HAL
open science

Comparative genomic analysis of Carnobacterium maltaromaticum: Study of diversity and adaptation to different environments

Christelle Iskandar

To cite this version:

Christelle Iskandar. Comparative genomic analysis of Carnobacterium maltaromaticum: Study of diversity and adaptation to different environments. Food and Nutrition. Université de Lorraine, 2015. English. NNT : 2015LORR0245 . tel-01754646

HAL Id: tel-01754646

https://hal.univ-lorraine.fr/tel-01754646v1

Submitted on 30 Mar 2018

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 LORRAINE

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Université de Lorraine
École Nationale Supérieure d'Agronomie et des Industries Alimentaires
Ecole Doctorale Sciences et Ingénierie des Ressources, Procédés, Produits, Environnement (RP2E)
Laboratoire d'Ingénierie des Biomolécules (LIBio)

THESE
présentée devant L'Université de Lorraine
pour obtenir le grade de Docteur de l'Université de Lorraine
Spécialité : Procédés Biotechnologiques et Alimentaires

Par

Christelle ISKANDAR

Analyse par génomique comparée de *Carnobacterium maltaromaticum* : Etude de la diversité et de l'adaptation à différents environnements

Comparative genomic analysis of *Carnobacterium maltaromaticum*: Study of diversity and adaptation to different environments

Prof. Xavier DOUSSET	UMR INRA-SECALIM - ONIRIS Nantes	Rapporteur
Dr. Marie-Christine CHAMPOMIER-VERGES	UMR 1319 MICALIS – INRA Jouy en Josas	Rapporteur
Dr. Veronique DELCENSERIE	FARAH - Fundamental and Applied Research for Animals & Health - University of Liege	Examinateur
Dr. Fréderic BORGES	LIBIO Laboratoire d'Ingénierie des Biomolécules ENSAIA - Université de Lorraine	Examinateur
Prof. Anne-Marie REVOL-JUNELLES	LIBIO Laboratoire d'Ingénierie des Biomolécules ENSAIA - Université de Lorraine	Directrice de thèse
Dr. Catherine CAILLIEZ-GRIMAL	LIBIO Laboratoire d'Ingénierie des Biomolécules ENSAIA - Université de Lorraine	Co-directrice de thèse

Dédicace

Je dédie ce travail

A mes parents,

qui ont tout sacrifié pour me voir réussir.

Vous êtes la source de mes efforts et la flamme de mon cœur.

A mes frères

lumière de mes jours, mon soutien moral, source de joie et de bonheur.

Vous m'avez toujours aidée et encouragée durant tout le chemin.

Christelle

Remerciements

Ce travail de thèse a été réalisé au sein du Laboratoire d'Ingénierie des Biomolécules LIBio. La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Je remercie tout d'abord les membres du jury, Monsieur Xavier DOUSSET Professeur à ONIRIS Nantes et Madame Marie-Christine CHAMPOMIER-VERGES, Directrice de recherche à l'INRA, rapporteurs de thèse et Madame Veronique DELCENSERIE, assistant- professeur à l'université de Liège, examinatrice des travaux de thèse, pour avoir accepté de juger ce travail..

J'exprime ensuite toute ma reconnaissance à mes encadrants, Madame Anne-Marie REVOL-JUNELLES, directrice de thèse, Madame Catherine CAILLIEZ-GRIMAL co-directrice et Monsieur Fréderic BORGES pour avoir accepté de diriger ce travail. Merci pour votre aide, votre soutien et spécialement pour votre confiance en moi.

Je tiens à remercier aussi tout le personnel du LIBio, Monsieur Michel LINDER, directeur du LIBio, enseignant-chercheurs, doctorants, techniciens et stagiaires. C'était un grand plaisir de vous connaitre. Un grand merci à tous les membres de l'équipe de microbiologie, Myriam, Sylvie et Arnaud, pour leur soutien et leur bonne humeur. Je suis tout aussi redevable à Monsieur Emmanuel RONDAGS, Monsieur Fabrice BLANCHARD et Monsieur Cédric PARIS pour leur disponibilité et leurs judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je voudrais adresser un merci à ERASMUS MUNDUS qui m'a donné la chance de réaliser ce parcours en me procurant le financement nécessaire durant mes deux ans en France. Un grand merci à Madame Delphine LAURANT pour sa patience.

Je voudrais exprimer ma reconnaissance envers les amis, spécialement Georges et Elie, qui m'ont apporté leur support moral et intellectuel malgré la distance. Un très grand merci à Christelle SALAMEH pour sa confiance et son support inestimable. Tu es plus qu'une amie, tu es une sœur, un compagnon de route dans les moments agréables, et un soutien dans les moments pénibles. Une pensée particulière à mon cher ami Ibrahim. Ta présence m'a donnée la force et le courage pour pouvoir toujours avancer.

Enfin, un vif remerciement s'adresse à mes parents, mes frères et ma tante Roula, pour leur encouragement, leur confiance et l'entraide dans les moments difficiles.

Résumé

Carnobacterium est un genre bactérien ubiquiste appartenant au groupe des bactéries lactiques (LAB). Les souches de *Carnobacterium maltaromaticum* sont présentes dans une grande diversité d'environnements et de produits alimentaires. Elles présentent de nombreuses propriétés, dont un effet positif en fabrication fromagère et des potentialités probiotiques qui en font un microorganisme à forte potentialité industrielle. L'objectif de ce travail de thèse est d'approfondir les connaissances sur *C. maltaromaticum* par une étude génomique. Neuf génomes de bactéries du genre *Carnobacterium,* dont 5 de *C. maltaromaticum,* disponibles sur la plateforme MicroScope ont été analysés et comparés. *Carnobacterium maltaromaticum* DSM20342 MX5 possède un génome de 3,85 Mbp, le plus grand génome parmi les LAB connus, de 1 Mbp de plus que celui des autres *Carnobacterium* n'appartenant pas à l'espèce *C. maltaromaticum*. L'analyse génomique détaillée indique la présence de quelques voies métaboliques spécifiques et suggère que les souches de *C. maltaromaticum* et *C. divergens* présenteraient une surface cellulaire caractérisée par une grande diversité moléculaire.Les trois autres souches de *Carnobacterium* présenteraient une surface fortement dépourvue en molécules de surfaces dont les protéines. Les propriétés de surface qui en découlent seraient à relier à la capacité de ces deux espèces à coloniser différents habitats. Par ailleurs, les souches de *C. maltaromaticum* isolées de différents produits laitiers se caractérisent par une grande diversité génomique. Afin de caractériser le niveau de diversité, les gènes impliqués dans les voies métaboliques d'utilisation du lactose (voie du Tagatose-6Phosphate, gènes *lac*) et du galactose (voie de Leloir, gènes *gal*) ont été caractérisés par analyse génomique au sein du genre *Carnobacterium.* De plus, ces gènes ont été recherchés par PCR chez 42 souches de *C. maltaromaticum*. Les deux voies sont présentes et organisées différemment chez les différentes souches de *C. maltaromaticum* d'origine laitière et non laitière L'analyse de ces gènes au niveau de la population révèle une dissémination des deux voies au sein des 4 lignées de cette population. La voie de Leloir est présente dans les lignées I, II et III et la voie du Tagatose-6Phosphate dans les lignées I, II et IV. Ces résultats suggèrent que les gènes *lac* et *gal* ont évolué selon un schéma complexe, qui est le reflet du haut niveau de diversité génétique de la population. La recherche de ces gènes au sein des 237 génomes de LAB disponibles sur la plateforme MicroScope Mage indique que les gènes *lac* et *gal* présentent un niveau de diversité élevé parmi les différents genres bactériens constituant le groupe des LAB, avec une dominance de la présence de la voie de Leloir. Ce niveau de diversité génétique n'est retrouvé qu'à l'échelle du genre chez les LAB et non de l'espèce, comme c'est le cas de *C. maltaromaticum*.

Mots-clés : *Carnobacterium maltaromaticum*, diversité, adaptation, comparaison génomique, lactose.

Summary

Carnobacterium bacteria belong to the lactic acid bacteria (LAB) and are ubiquitous. *Carnobacterium maltaromaticum* strains are present in a wide variety of environments and foods. They have many positive properties in cheese manufacturing that make them microorganisms of industrial interest. The objective of this thesis is to deepen the knowledge on *C. maltaromaticum* by a comparative genomic approach. Nine *Carnobacterium* genomes, including 5 *C. maltaromaticum*, available on the platform MicroScope were analyzed and compared. *Carnobacterium maltaromaticum* DSM20342 MX5 is the largest genome among LAB, 1Mbp more than other *Carnobacterium* species. Detailed analysis indicates the presence of some specific metabolic pathways, and the presence of a high molecular diversity of cell surface proteins in *C. maltaromaticum* and *C. divergens*, while the three other *Carnobacterium* species present only few proteins on their surfaces. This can be related to the ability of these two species to colonize different habitats. Furthermore, *C. maltaromaticum* strains isolated from various dairy products are characterized by a large genomic diversity. In order to characterize this diversity, genes involved in lactose (Tagatose-6Phosphate pathway, *lac* genes) and galactose (Leloir pathway, *gal* genes) metabolic pathways were identified by genomic analysis within *Carnobacterium* genus and detected by PCR in 42 *C. maltaromaticum* strains. Both pathways are present and organized differently in the *C. maltaromaticum* strains. The analysis of these genes at the population level shows a spread of the two pathways within the 4 lineages of this population. The Leloir pathway is present in the lineages I, II and III and the Tagatose-6Phosphate in the lineages I, II and IV. These results suggest that the *lac* and *gal* genes have evolved in a complex pattern, which reflects the high level of genetic diversity of the population. These genes were characterized in 237 LAB genome and exhibit a high degree of diversity among the different bacterial genera of the LAB group, with a dominance of the presence of the Leloir pathway. This level of genetic diversity is found at the genus level in LAB and not at the species level, as for *C. maltaromaticum*.

Keywords: *Carnobacterium maltaromaticum*, diversity, adaptation, comparison genomics, lactose.

Table of content

Dédicace II

Remerciements III

Résumé IV

Summary IV

Table of content V

List of Tables VII

List of Figures VIII

List of abbreviations X

List of Annexes 1

INTRODUCTION 2

CHAPTER I: BIBLIOGRAPHY 6

1. Lactic Acid Bacteria and *Carnobacterium* 7

1.1. Taxonomy and genomics of Lactic Acid Bacteria 7

1.2. The genus *Carnobacterium* 11

2. Bacterial Cell Wall Proteins 24

2.1. Proteins anchorage to the Cell Surface 25

2.2. The function of Cell Wall Proteins 29

3. Bacterial adaptation to dairy products: lactose and galactose metabolism 36

3.1. Introduction 36

3.2. Sugar uptake systems 38

3.3. Lactose molecule cleavage: beta-galactosidase 39

3.4. Lactose/ galactose metabolism pathways 39

3.5. Regulation of the metabolic pathways 40

3.6. Genes mutation effects 41

3.7. Specificity and genes organization in LAB strains 42

4. Objectives 49

CHAPTER II: RESULTS AND DISCUSSION 51

1. Comparative genomic analysis reveals contrasting differences between *Carnobacterium* species 52

1.1. Introduction 52

1.2. Material and Methods 54

1.3. Results and discussion 55

1.4. Conclusion 72

2. Adaptation to dairy environment: lactose metabolism pathways in LAB and *Carnobacterium* 73

2.1. Genes associated to lactose metabolism illustrate the high diversity of *Carnobacterium maltaromaticum* 73

2.2. Diversity of lactose/galactose metabolic pathways within Lactic Acid Bacteria 87

CHAPTER III: CONCLUSIONS AND PERSPECTIVES 102

REFERENCES 107

ANNEXES 137

Résumé 156

Summary 156

List of Tables

Table 1: Ecology of eight *Carnobacterium* sp. 13

Table 2: Ecology of *Carnobacterium divergens* and *Carnobacterium maltaromaticum*. 14

Table 3: Genes implicated in lactose and galactose transportation systems, Tagatose-6P and Leloir pathways. 37

Table 4: *lac/gal* genes organization in some LAB strains. 43

Table 5: Origins of *Carnobacterium* sp. 54

Table 6 : General features of *Carnobacterium* genomes. 57

Table 7: Comparative analysis of the Pan/Core genome (50-80) between A: all the *Carnobacterium* genomes, B: *Carnobacterium* sp. and *C. inhibens* subsp. *gilichisnkyi* WN1359 genomes and C: *Carnobacterium divergens* and *C. maltaromaticum* genomes. 60

Table 8: Primer sequences used in this study 77

Table 9: Lactose/galactose metabolism putative genes in *Carnobacterium* sp. 79

Table 10: Presence of *lac* and *gal* genes within a collection of 42 *Carnobacterium maltaromaticum* strains. 83

Table 11 : Lactose/galactose metabolism genes present in LAB and non-LAB strains. 90

List of Figures

Figure 1: Schematic phylogenetic tree of some lactic acid bacteria, including some Gram-positive bacteria of the low GC subdivision, *Bacillus, Listeria* and *Staphylococcus.* 8

Figure 2: Genome size of some LAB genera. 9

Figure 3: eBURST analysis of 47 *C. maltaromaticum* strains. 19

Figure 4 : Cell envelope of lactobacilli with a schematic representation of cell wall and membrane-associated proteins. 25

Figure 5 : Isd-mediated heme-iron uptake across the cell wall of *Staphylococcus aureus.* 30

Figure 6 : Model of pilus biogenesis. 32

Figure 7 : Proteolysis pathway by LAB, with schematic representation of the action of peptidases found in lactic acid bacteria. 34

Figure 8: Schematic representation of the lactose and galactose catabolic pathways. ... 37

Figure 9: Relationship between *Carnobacterium* genus. A- phylogenetic tree of the sequenced genomes of *Carnobacterium* based on the conserved nucleic sequence of 10 essential genes (*dnaK gyrA polA lepA dnaB gyrB secA ftsZ recG ileS*). B- eBURST analysis of 49 *C. maltaromaticum* strains. 56

Figure 10: COG distribution in the different classes for 5 strains of *Carnobacterium*. *C. maltaromaticum* LMA28, *C. divergens* V41, *Carnobacterium* sp. 17.4, *C. inhibens* subsp. *gilichinskyi* WN1359. 59

Figure 11: Pan/Core genome analysis of *Carnobacterium maltaromaticum* strains with a cut-off of 70% amino-acid identity on 80% coverage. 61

Figure 12: Comparison of the distribution of the Core genome and the intraspecific Core genome in the different COG classes. 62

Figure 13: Secretome size in the different *Carnobacterium* genomes, LAB genomes and other Gram-positive bacteria. 64

Figure 14: Putative surface proteins containing predicted functional domains. 68

Figure 15: Gene cluster organization and synteny between *C. maltaromaticum* LMA28, *L. rhamnosus* GG and *E. faecalis* V583. 72

Figure 16: Organization and synteny between *lac* genes among *Carnobacterium maltaromaticum* LMA28, DSM20342-MX5, ATCC35586, *Streptococcus mutans,* and *Staphylococcus aureus*. 78

Figure 17: Organization and synteny between *gal* genes in *Carnobacterium maltaromaticum* 3.18, *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. 17.4 and *Carnobacterium* sp. AT7. 81

Figure 18: Diversity of the lactose and galactose metabolic pathways within A. genus and B. species of the different genus analyzed. 98

List of abbreviations

°C degree celcius

µl micro litre

µM micro Mole

16S rDNA 16S ribosomal Desoxyribo Nucleic Acid

16S rRNA 16S ribosomal RiboNucleic Acid

aa amino acid

ADN Acide DesoxyriboNucléique

AOP Appélation d'Origine Protégée

BBH Bidirectionnal Best Hit

CBL Club des Bactéries Lactiques

CC Clonal Complex

CDS Coding DNA Sequence

CEP Cell Envelope Protease

CFU.g-1 Colony Forming Unit per Gramme

CM Cell Membrane

COG Cluster of Orthologuous Genes

CRISPES-cas Cluster Regularly Interspaced Short Palindromic Repeats – associated genes

CWP Cell Wall Proteins

CWSS Cell Wall Signal Sequence

EC Enzyme Commission

EPS Exopolysaccharids

FBP Fructose Bi Phosphate

G6P Glucose 6 Phosphate

GC Guanine and Cytocine

GRAS Generally Recognized As Safe

Hb Hemoglobin

HGT Horizontal Gene Transfer

Isd Iron Surface Determinant

KEGG Kyoto Encyclopedia of Genes and Genomes

LAB Lactic Acid Bacteria

LTA LipoTeichoic Acid

LysM Lysine Motif

MAP Modified Atmosphere Packaging

mbar milli bar

Mbp Mega base pair

MLST Multi Locus Sequence Typing

NEAT NEAr-Transporter

ng nanogramme

P phosphate

PCR Polymerase Chain Reaction

PEP PhosphoEnolPyruvate

PG PeptidoGlycan

PRD PTS Regulatory Domains

PTS PhosphoTransferase System

SDP Sortase Dependant Protein

sec second

Sec Secretion system

SP Signal Peptide

List of Annexes

Annexe 1: CDS distribution of the intraspecific core genome in the different COG... 138

Annexe 2: Summary of the analysis of the cell wall proteins 142

Annexe 3: Genes encoding surface proteins of the 9 *Carnobacterium* strains, and the genes involved in the proteolysis. 143

Annexe 4: Percentage of identity between *lac* and *gal* genes within LAB. 148

Annexe 5: List of *C. maltaromaticum* LMA28 megaplasmid genes grouped in the different COG 149

Annexe 6: List of permeases found in *Carnobacterium* strains 151

INTRODUCTION

Les travaux de recherche effectués au sein du LIBio sont centrés sur la valorisation d'agro-ressources à des fins alimentaires et non alimentaires, en alliant des compétences en biochimie, physico-chimie et microbiologie. L'objectif de la recherche développée au laboratoire est de comprendre et de maitriser la structuration et la fonctionnalisation de la matière molle, les mécanismes de transferts et de relargage dans des systèmes complexes, ainsi que les interactions au sein des différents systèmes étudiés. L'ensemble des travaux prend en compte l'impact des paramètres biotique et abiotique afin de stabiliser les systèmes et/ou de concevoir des vecteurs et des matrices à fonctionnalité ciblée. Les composés actifs peuvent être des biomolécules - lipides polaires, acides gras polyinsaturés, antioxydants, composés antibactériens - ou des bactéries.

Des systèmes glucidiques et protéiques sont conçus pour la stabilisation et/ou la vectorisation des composés actifs particuliers que sont les bactéries. Le choix du système matrice/composé actif intègre les interactions avec la matrice ainsi que l'impact que pourrait avoir la bactérie incorporée sur la structuration des communautés microbiennes de l'écosystème considéré.

Dans ce contexte, la matrice alimentaire et la bactérie lactique *Carnobacterium maltaromaticum* servent de modèle dans le cadre de cette étude.

Carnobacterium appartient au groupe des bactéries lactiques (LAB). Sur les onze espèces que contient le genre *Carnobacterium*, les plus fréquemment isolées des aliments sont *C. mataromaticum* et *C. divergens* (Leisner et al., 2007). *Carnobacterium maltaromaticum* a été isolée d'une grande variété d'aliments dont des produits laitiers (Afzal et al., 2010). En effet, les travaux réalisés au LIBio ont démontré que *C. maltaromaticum* fait partie de la microflore complexe des fromages (Cailliez-Grimal et al., 2007, 2005; Millière et al., 1994; Milliere and Lefebvre, 1994; Rahman et al., 2014a). La présence de cette bactérie n'est pas négligeable car elle peut atteindre des concentrations élevées en fin d'affinage dans des fromages au lait de chèvre ou de vache (jusqu'à 10^9 UFC.g^{-1}) (Cailliez-Grimal et al., 2007) où elle jouerait un rôle positif (Afzal et al., 2010). C'est une bactérie lactique psychrotrophe, capable de se développer à des valeurs de pH alcalin et produisant un arome malté (Afzal et al., 2013b, 2012). De nombreuses souches de cette espèce sont connues pour la production d'une large gamme de bactériocines (Afzal et al., 2010). Les derniers travaux réalisés au laboratoire ont mis en évidence des potentialités probiotiques de *C. maltaromaticum* LMA28 (Rahman et al., 2014b) ainsi qu'une grande diversité

génétique au sein d'une collection de 42 souches de *C. maltaromaticum* d'origines diverses (Rahman et al., 2014a).

Le séquençage des génomes bactériens est de plus en plus rapide et accessible. Ceci ouvre la voie à de nouvelles possibilités d'analyse des propriétés des bactéries. Récemment, le génome de *C. maltaromaticum* LMA28 a été séquencé au laboratoire (Cailliez-Grimal et al., 2013). Les génomes *C. maltaromaticum* 3.18 et ML.1.97 ainsi que celui de *C. divergens* ont été mis à notre disposition grâce à des collaborations respectivement avec J. Leisner (University of Copenhagen, Denmark) et H. Prévost (ONIRIS, Nantes). Ceux de deux autres souches de *Carnobacterium* sp. AT7 et 17.4, de *C. maltaromaticum* DSM20342 MX5 et ATCC35586, ainsi que celui de *C. inhibens* subsp. *gilichinsky* WN1359 sont en accès libre.

C'est dans ce contexte que s'est déroulé ce travail de thèse, dont l'objectif est d'approfondir les connaissances sur *C. maltaromaticum* par une étude génomique.

Cette étude devrait permettre de répondre à différentes questions :

- *Carnobacterium maltaromaticum* est isolée d'une grande diversité d'environnement et la population présente une grande diversité. Une analyse a montré que les génomes des souches de *C. maltaromaticum* sont de plus grande taille que ceux des autres bactéries du même genre. Que renferme cet ADN ? Quelle est la fonction des gènes spécifiquement présents chez *C. maltaromaticum* ?

- *Carnobacterium maltaromaticum* est très représentée dans l'environnement laitier, avec une grande diversité de souches. Cette diversité peut-elle se retrouver au niveau des propriétés physiologiques, et notamment au niveau des voies métaboliques d'utilisation du lactose, composé clé en technologie fromagère ?

Afin de répondre à ces interrogations, une revue bibliographique, rédigée en anglais, a été réalisée sur la génomique des bactéries lactiques, l'écologie des bactéries du genre *Carnobacterium* et les voies de métabolisation du lactose et du galactose. Les analyses de génomique comparée ont montré la particularité des bactéries du genre *Carnobacterium* vis-à-vis de leur surface cellulaire. Une partie de la bibliographie est relative à la structure des surfaces bactériennes des bactéries à Gram-positive.

Les résultats sont rédigés en anglais et présentés en deux chapitres :

Le premier chapitre est dédié à l'analyse génomique comparée des différentes bactéries du genre *Carnobacterium*.

Le deuxième chapitre est dédié à la caractérisation des gènes impliqués dans les voies de métabolisation du lactose et du galactose chez les bactéries du genre *Carnobacterium*, chez *C. maltaromaticum* plus spécifiquement et chez l'ensemble des LAB.

Les conclusions générales et les perspectives que laissent entrevoir ce travail sont rédigées en français et présentées dans le dernier chapitre de cette thèse.

Ce travail a fait l'objet :

- d'une communication affichée au Club des Bactéries Lactiques (CBL) Lille, 17-19 juin 2015, France

Christelle F. ISKANDAR, Frédéric BORGES, Abdur RAHMAN, Emmanuel RONDAGS, Fabrice BLANCHARD, Benoît REMENANT, Monique ZAGOREC, Jorgen J. LEISNER, Catherine CAILLIEZ-GRIMAL, Anne-Marie REVOL-JUNELLES. Analyse génomique et diversité de *Carnobacterium maltaromaticum* : un exemple, le métabolisme du lactose

- d'une publication soumise à Food Microbiology

Christelle F. ISKANDAR, Catherine CAILLIEZ-GRIMAL, Abdur RAHMAN, Emmanuel RONDAGS, Benoît REMENANT, Monique ZAGOREC, Jorgen J. LEISNER, Frédéric BORGES and Anne-Marie REVOL-JUNELLES Genes associated to lactose metabolism illustrate the high diversity of *Carnobacterium maltaromaticum*

CHAPTER I: BIBLIOGRAPHY

1. Lactic Acid Bacteria and *Carnobacterium*

1.1.Taxonomy and genomics of Lactic Acid Bacteria

The interest for lactic acid bacteria (LAB) and foods begins with Pasteur's work on lactic acid fermentation in 1857 and the first isolation of a pure culture, *Bacterium lactis* (now called *Lactococcus lactis* subsp. *lactis*), by Leister in 1873. According to Orla-Jensen in 1919, the "true lactic acid bacteria" form a natural group of Gram-positive, non sporulating rods or cocci that ferment carbohydrates and higher alcohols to form lactic acid (Axelsson, 2004; Stackebrandt and Teuber, 1988; Stiles and Holzapfel, 1997).

The definition and classification of LAB change with the introduction of molecular biology and was the focus of intense taxonomic study with approaches involving both phenotypic and phylogenetic characterization of bacteria (Axelsson, 2004; Pot, 2008).

Actually, this term LAB relates to the metabolic capabilities of microorganisms to ferment various nutrients predominantly into lactic acid (Klaenhammer et al., 2005; Liu, 2003). The general characteristics of LAB are that they are Gram-positive, anaerobic, acid-tolerant and non-sporulating (Klaenhammer et al., 2005), catalase negative, although some LAB strains have a catalase activity mediated by a non-heme "pseudocatalase"(Engesser and Hammes, 1994), microaerophilic, rods and cocci (Axelsson, 2004).

LAB form a heterogeneous group of microorganisms, and the different genera included in the term LAB have been subject to several controversies. Historically, the four genera *Lactobacillus*, *Leuconostoc*, *Pediococcus* and *Streptococcus* form the core of the group, corresponding to the current genera *Enterococcus, Carnobacterium, Lactobacillus, Lactococcus, Leuconostoc, Oenococcus, Pediococcus, Streptococcus, Vagococcus* and *Weissella* (Axelsson, 2004; Pot, 2008; Stackebrandt and Teuber, 1988; Stiles and Holzapfel, 1997).

Several taxonomic revisions of these genera and the description of new genera lead to several change. Among domesticated bacteria studied and exploited, LAB are found in two distinct phyla, namely *Firmicutes* and *Actinobacteria*. LAB in *Actinobacterium* phylum include only *Atopobium* and *Bifidobacterium* genera (which produce lactic acid but always in combination with acetic acid), with a Guanine-Cytosine (GC) content of 36-46% and 58-61%, respectively (de Vos, 2011; Horvath et al., 2009; Makarova et al., 2006). Within the *Firmicutes* phylum, LAB belong to order *Lactobacillales* and include the genera *Aerococcus*, *Alliococcus*,

Carnobacterium, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus*, *Vagococcus* and *Weissella* which are all low GC content organisms (31-49%) (de Vos, 2011; Pot, 2008). In a phylogenetic tree of LAB published in 2004, *Carnobacterium* genus appeared to be more related to *Vagococcus* and *Enterococcus* than to any other LAB (Figure 1) (Axelsson, 2004). However, phylogenetic relationships among LAB are still the subject of many discussions (Zhang et al., 2011).

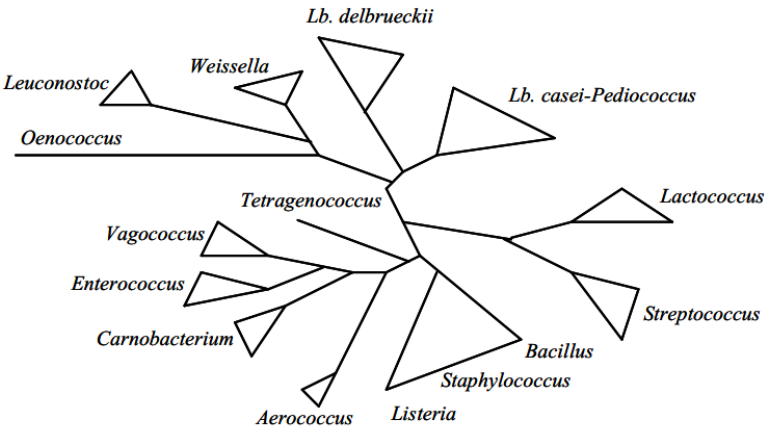


Figure 1: Schematic phylogenetic tree of some lactic acid bacteria, including some Gram-positive bacteria of the low GC subdivision, *Bacillus, Listeria* and *Staphylococcus* (Axelsson, 2004).

The first LAB that was completely sequenced, *L. lactis* subsp. *lactis* IL1403 came out in 2001 (Bolotin et al., 2001). Then in 2003, Kleerebezem et al. sequenced the first and complete *Lactobacillus* strain genome, *L. plantarum* WCFS1 (Kleerebezem et al., 2003). In 2014, approximately 6,800 bacterial genome sequences are available of which 182 *Lactobacillus* genomes, 14 *Lactococcus* genomes, 113 *Enterococcus* genomes, and much more *Streptococcus* genomes (Wassenaar and Lukjancenko, 2014).

Genome sequencing is a relatively new discipline allowing a rapid bacterial exploration. It offers a wealth of information to well-understand LAB when investigating their gene content, their properties and their role in human health and food fermentation (Johnson and Klaenhammer, 2014). The study of LAB sequenced bacteria showed that *Lactobacillales* present a diversity in genome size ranging from 1.8 Mbp for *Streptococcus* to 3.7 Mbp for *Carnobacterium* (Figure 2). It was hypothesized that *Lactobacillales* diverged from their common ancestor with *Bacilli* causing the loss of 600-1200 genes, some of them encoding biosynthesis enzymes (Makarova and Koonin, 2007). Genome size within a genus is homogenous in almost all LAB except in *Lactobacillus* (Wassenaar and Lukjancenko, 2014)

and *Carnobacterium* genera. Genome size of *Lactobacillus* range from 1.3 Mbp for *L. iners* AB-1 (Macklaim et al., 2011) to 3.3 Mbp for *L. plantarum* WCFS1 (Kleerebezem et al., 2003). The genome size of *Carnobacterium* species is estimated ranging from 1.9 (for *C. alterfunditum*) (Pikuta et al., 2005) to 3.7 Mbp (for *C. maltaromaticum*) (Cailliez-Grimal et al., 2013). *Carnobacterium maltaromaticum* strains possess bigger genomes compared to other LAB (Hols et al., 2005; Kelly et al., 2010; van de Guchte et al., 2006). It has been suggested that the large size of *Carnobacterium* genome may be the reason of its well adaptation to environmental challenges (Leisner et al., 2007). Indeed, in order to survive and grow in a variety of environments, a bacterium will need a large number of encoded genes, and thus a larger genome. Therefore, the number of predicted protein genes strongly correlates with genome size; that means that a larger genome holds a larger number of CDS (Coding DNA Sequence) (Wassenaar and Lukjancenko, 2014).

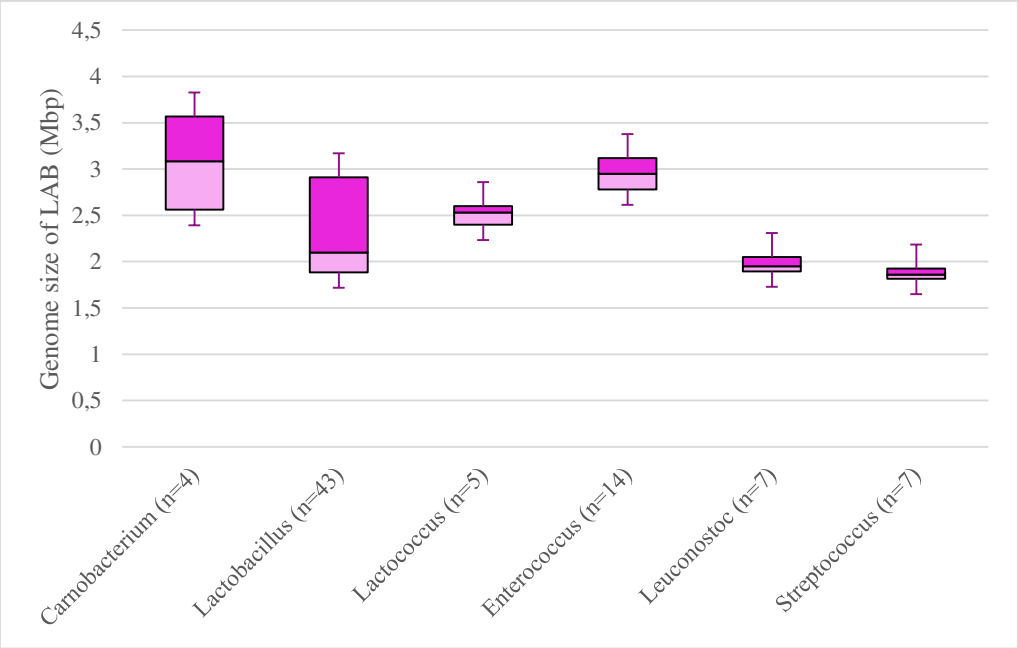


Figure 2: Genome size of some LAB genera.
according to (Cailliez-Grimal et al., 2013; Douillard and de Vos, 2014; Leisner et al., 2012; Leonard et al., 2013; Voget et al., 2011; Wassenaar and Lukjancenko, 2014). n = number of strains

LAB are found naturally in a variety of environmental habitats, including plant (fruits, vegeTable, cereal), meat and milk environment, and are involved in a large number of industrial and spontaneous food fermentations, notably those based on raw materials derived from these natural habitats. The traditional roles for many LAB have been as starter cultures (Klaenhammer et al., 2005), which lead to their widespread human consumption. Indeed, the LAB used in food are considered non-pathogenic and are awarded the qualification of Anglo-

Saxon agencies Generally Regarded As Safe (GRAS) (Aguirre and Collins, 1993). Their primary contribution is in rapid acid production and acidification of foods, but metabolic processes accompanying the growth of LAB impact also flavor, nutrition, and texture quality of a variety of fermented foods (Axelsson, 2004; Klaenhammer et al., 2005; Kleerebezem and Hugenholtz, 2003; Pot, 2008). The sequenced and annotated bacterial genomes permit the prediction of certain important characteristics researched by industrials. Thanks to comparative genomics, it became easy to search for metabolic pathways (Cogan et al., 2007). Indeed, *Lactobacillus helveticus* CNRZ32 had been discovered to possess homologs of genes involved in the proteolysis of casein found in milk (Broadbent et al., 2013). The research project consists in predicting the formation of metabolic compounds *via* the exploration of the genomic content of CDS in comparison to previously annotated genomes. This is the case also of the identification of flavor molecules (G. Smit et al., 2005), decarboxylation of branched-chain α-Keto acids in *L. lactis* (B. A. Smit et al., 2005). In addition, LAB have the ability to avoid the food product deterioration by inhibition of bacterial growth by the production of lactic acid and growth-inhibiting compounds such as acids, H_2O_2, CO_2 and bacteriocins (Cotter et al. 2005).

Some LAB are also closely associated with the mucosal surfaces of animals and human environment, including the gastrointestinal tract, the oral, the respiratory and the vaginal cavities. Moreover, many species of LAB are considered to be important components of the normal intestinal microbiota, which contribute to a variety of functions including intestinal integrity, immunomodulation, and pathogen resistance. Selected groups of LAB are used as probiotics for human consumption (Gilliland, 1989; Klaenhammer et al., 2005; Pot, 2008; Stiles and Holzapfel, 1997).

LAB colonize a great diversity of habitats. Adaptation to a particular niche is allowed by genomic remodeling. Comparative genomics of LAB revealed that loss and gain of genes occurred during the evolution of these bacteria, and it is accepted that these events allowed to improve their ecological performance (Makarova et al., 2006).

The more striking examples of evolution by gene loss are *Streptococcus thermophilus* (Bolotin et al., 2004), *Lactobacillus bulgaricus* (Kafsi et al., 2014; van de Guchte et al., 2006), *L. helveticus* (Callanan et al., 2008) and *L. lactis* subsp. *cremoris* UC509.9 (Ainsworth et al., 2013). The genome of these bacteria present many pseudogenes linked to amino acids biosynthesis and those of the fermentation of plant derived sugars (Cavanagh et al., 2015). It

was hypothesized that these lineages experienced a massive gene loss resulting in genome decay.

At another extent, the gene gain has been detected in several cases. For organisms mainly present in gastrointestinal tracts, such as *L. acidophilus* (Altermann et al., 2005), *L. plantarum* (Kleerebezem et al., 2003) and *L. gasseri* (Azcarate-Peril et al., 2008), many genes were suspected to be gained by Horizontal Gene Transfer (HGT) and contributed to the interaction of the bacterium with intestinal mucosa and its survival in gastric conditions. In this connection, many mucin-binding proteins and gene clusters for transport of a diverse group of carbohydrates are found in those genomes.

Gain of genes has been also well documented with the acquisition of plasmids holding metabolic pathways. This is the case of *Lactococcus garvieae* (Collins et al., 1983; Vendrell et al., 2006; Wyder et al., 2011), *L. lactis* (Kelly et al., 2010) and *Streptococcus macedonicus* (Malkin et al., 2008) isolated from different environments. They are known to possess plasmids containing genes responsible of the lactose and galactose utilization *via* the Tagatose-6Phosphate pathway (Cavanagh et al., 2015; Flórez and Mayo, 2015; Papadimitriou et al., 2015; Passerini et al., 2010).

1.2. The genus *Carnobacterium*

1.2.1. Taxonomy

The genus *Carnobacterium* was first proposed to clarify the taxonomic position of atypical heterofermentative *Lactobacillus* species (groupe III) isolated from chicken meat. They suggested classifying *Lactobacillus divergens* and *Lactobacillus piscicola* in a new genus, *Carnobacterium*, as *Carnobacterium divergens* and *Carnobacteriun piscicola*. Two new species, *Carnobacterium gallinarum* and *Carnobacterium mobile* were also added in this new genus (Collins et al., 1987).

The first *Carnobacterium* strain was isolated in 1974 and designated as *Lactobacillus maltaromicus* by Miller (Miller et al., 1974). Based on phenotypic and genotypic comparisons, this strain and *C. piscicola* described in 1987 were reclassified in 2003 as *C. maltaromaticum* (Mora et al., 2003).

The *Carnobacterium* genus forms a phylogenetically coherent group belonging to the LAB, which is quite distinct from other LAB, as shown by 16S rDNA sequencing (Wallbanks et al., 1990). It belongs to the phylum *Firmicutes*, Class *Bacilli*, order, *Lactobacillales*, family *Carnobacteriaceae*. This new family was created in 2009 (Ludwig, et al., 2009) with *Carnobacterium* as genus type and was formed of two paraphyletic groups: the first one consisted of eleven genera, *Alkalibacterium, Allofustis, Alloiococcus, Atopococcus, Atopostipes, Carnobacterium, Desemzia, Dolosigranulum, Isobaculum, Marinilactibacillus* and *Trichococcus*, and the second one included *Atopobacter* and *Granulicatella*. The phylogenetic position of this second group is still under discussion (Pikuta, 2014).

Strains of *Carnobacterium* are Gram-positive, non-spore-forming rods or coccobacilli, motile or not. These bacteria are catalase negative, but some species exhibit catalase activity in the presence of heme. Initially described as heterofermentative, carnobacteria could be considered to be homofermentative organisms that produce lactic acid from glucose (except for the species *C. pleistocenium*) or facultative heterofermentative. Some species could metabolize hexoses and pentoses to L(+)-lactic acid and, depending on the access of oxygen, may produce acetic acid, ethanol, CO_2 and formic acid in varying amounts. Respiration might occur in the presence of hematin. They do not reduce nitrate to nitrite. Production of $NH4^+$ from arginine is a result of its catabolism catalyzed by the arginine deaminase pathway. Some species possess the ability of converse tyrosine to tyramine. They are mesophilic and some species may be psychrotolerant and grow at 0 °C. NaCl is not required for growth. Some strains are halotolerant until 8% NaCl and alkaliphilic with growth until pH 9.5 (Cailliez-Grimal et al., 2014; Pikuta, 2014; Pikuta and Hoover, 2014).

The genus *Carnobacterium* included 11 species: *Carnobacterium alterfunditum, C. divergens, C. funditum, C. gallinarum, C. iners, C. inhibens, C. jeotgali, C. maltaromaticum, C. mobile, C. pleistocenium*, and *C. viridans*. Recently, *C. gilichinskyi* was reclassified as *C. inhibens* subsp. *gilichinskyi*, and the subspecies *C. inhibens* subsp. *inhibens* was created automatically for *C. inhibens* (Nicholson et al., 2015).

1.2.2. Ecology

Carnobacteria are ubiquitous LAB which have been frequently isolated from cold and temperate environment and from animals and foods of animal origin such as seafood, meat and dairy products. Table 1 lists the carnobacteria species and Table 2 focuses on ecology of the two species *C. divergens* and *C. maltaromaticum*.

The wide distribution in nature is due to their particular physiology. Moreover, these bacteria produce several bacteriocins, which allow them to compete with other bacteria for colonization of the environment.

A potential pathogenicity is associated with some *Carnobacterium* species.

Table 1: Ecology of eight *Carnobacterium* sp.

Species	Ecology	References
C. alterfunditum	Ace Lake	(Franzmann et al., 1991)
	MAP rough-head-grenadier	(Matamoros et al., 2009)
C. funditum	Ace Lake	(Franzmann et al., 1991)
	Hindgut region of Artic charr	(Ringo et al., 2002)
	Ice from a fjord in Norway	(Groudieva et al., 2004)
	Cooked tropical shrimps	(Jaffres et al., 2009; Matamoros et al., 2009)
	Marine sponge	(Li and Liu, 2006)
	Antartica	(Loperena et al., 2012)
C. gallinarum	Chicken meat.	(Collins et al., 1987)
C. iners	Littoral zone of an Antarctic lake	(Snauwaert et al., 2013)
C. inhibens subsp. *inhibens*	Gastrointestinal tract of Atlantic salmon	(Joborn et al., 1999)
	Siberian permafrost	(Nicholson et al., 2013)
C. inhibens subsp. *gilichinskyi*	Siberian permafrost	(Nicholson et al., 2013)
C. jeotgali	Traditional Korean fermented food (shrimp)	(Kim et al., 2009)
C. mobile	Chicken meat	(Collins et al., 1987)
	Altered raw salmon	(Mace et al., 2012)
C. pleistocenium	Pleistocene ice from the permafrost tunnel	(Pikuta et al., 2005)
C. viridans	Siberian permafrost	(Nicholson et al., 2013)
	Spent mushroom compost	(Ntougias et al., 2004)
	Antic toundra soil	(Kim et al., 2013)
	Refrigerated vacuum-packaged bologna	(Holley et al., 2002)

Table 2: Ecology of *Carnobacterium divergens* and *Carnobacterium maltaromaticum*.

Species	Ecology		References
C. divergens	Environment	Silage and mixed pasture of vegetal	(Tohno et al., 2012)
	Fish	Rainbow trout intestine	(Kim and Austin, 2008)
		Cold salmon	(Leroi et al., 1998)
		Fish products	(Pilet et al., 1994)
		Fresh cuts of suckling lamb	(Oses et al., 2013)
	Meat	Vacuum-packaged raw minced beef	(Holzapfel and Gerber, 1983)
		Ham	(Audenaert et al., 2010)
		Fresh pork sausage	(Benson et al., 2014; Curiel et al., 2011)
		Fresh beef	(Ercolini et al., 2009)
		Pork meat juice	(Rieder et al., 2012)
		Marinated pork product	(Schirmer et al., 2009)
	Poultry	Irradiated minced chicken	(Thornley, 1957)
		Chicken skin	(Vihavainen et al., 2007)
		Coocked MAP refrigerated poultry meat	(Barakat et al., 2000)
		MAP sliced chicken or turkey	(Audenaert et al., 2010)
	Dairy	French surface mould-ripened cheese	(Millière et al., 1994)
		Traditional Mozarella	(Morea et al., 1999)
	Alteration	Vacuum-packaged beef	(Jones, 2004; Youssef et al., 2014)
		MAP refrigerated broiler legs	(Bjorkroth et al., 2005)
		Refrigerated cooked tropical shrimps	(Jaffres et al., 2009)
		Altered raw salmon	(Mace et al., 2012)
C. maltaromaticum	Environment	Soil after 48 freeze-thaw cycles	(Walker et al., 2006)
		Farm field treated with whey	(Coombs and Brenchley, 1999)
		Japan lake water	(Yanagida et al., 2007)
	Fish	Rainbow trout intestine	(Kim and Austin, 2008)
		Cold salmon	(Leroi et al., 1998)
		Fish products	(Pilet et al., 1994)
		Vacuum packaged cold smoked salmon	(Duffes et al., 1999)
	Meat	Marinated pork product	(Schirmer et al., 2009)
		MAP meat	(Nieminen et al., 2011)
		Belgian artisan-type coocked ham	(Vasilopoulos et al., 2010)
	Poultry	Irradiated minced chicken	(Thornley, 1957)
		Chicken skin	(Vihavainen et al., 2007)
		Coocked MAP refrigerated poultry meat	(Barakat et al., 2000)
	Dairy	Milk	(Miller et al., 1974)
		French surface mould-ripened soft cheese,	(Cailliez-Grimal et al., 2007; Edima et al., 2007; Millière et al., 1994; Milliere and Lefebvre, 1994)
		Cheese of diverse origins	(Rahman et al., 2014a)
		Taleggio cheese	(Feligini et al., 2012)
		Traditional Mozarella	(Morea et al., 1999)
	Alteration	Spoilage of refrigerated beef	(Ercolini et al., 2009; Youssef et al., 2014)
		MAP refrigerated broiler legs	(Bjorkroth et al., 2005)
		Refrigerated MAP brined shrimp	(Mejlholm et al., 2008)
		Vacuum or MAP fresh salmon	(Emborg et al., 2002; Mace et al., 2012)
		Smoked salmon	(Paludan-Muller et al., 1998)
		Refrigerated cooked tropical shrimps	(Jaffres et al., 2009)

1.2.3. Environment

Cold environment

Carnobacteria are mesophilic psychotolerant species able to grow at 0-28°C and this psychrotolerance could explain their distribution in cold natural environments. Some *Carnobacterium* sp. are remarkably resistant to freezing. In fact, a *Carnobacterium* sp., closed to *C. maltaromaticum*, was isolated from soil after 48 freeze-thaw cycles (Walker et al., 2006).

Some strains are able to grow under low temperature (0°C), low pressure (7 mbar) and anoxic CO_2-dominated atmosphere (Nicholson et al., 2015).

Three species were initially isolated from cold environment with low nutrients contents in Alaska. *Carnobacterium funditum* and *C. alterfunditum* were isolated from the water of Ace Lake (Franzmann et al., 1991) and *C. pleistocenium* from Pleistocene ice from the permafrost tunnel (Pikuta et al., 2005). *Carnobacterium funditum* was also isolated from ice taken from a fjord in Norway (Groudieva et al., 2004), in Antartica (Loperena et al., 2012), and six *Carnobacterium* strains were isolated from the Siberian permafrost. Five of these strains were closely related to *C. inhibens* and one isolate to *C. viridans* (Nicholson et al., 2013).

Carnobacterium iners was isolated from a microbial mat actively growing in the littoral zone of an Antarctic lake (Forlidas Pond) in the Pensacola mountain (Snauwaert et al., 2013). *Carnobacterium* sp. are present in Lake Vanda, a permanently ice-covered Antartic lake (Bratina et al., 1998), in deep surface sediments (Newberry et al., 2004). Several new strains have been identified in cold-preserved tissue sample from a Siberian baby mammoth (Pikuta et al., 2011). However, the culture origin is not clear and their presence could be due to contamination during transport (Pikuta and Hoover, 2014).

Terrestrial environment

Some *Carnobacterium* strains closed to *C. maltaromaticum* are present in terrestrial environment such as a farm field treated with whey (Coombs and Brenchley, 1999). *Carnobacterium* strains closed to *C. viridans* are part of the Anctic toundra soil (Kim et al., 2013) and of Spent mushroom compost, composed of thermally treated cereal straw and animal manure mixture colonized by fungi (Ntougias et al., 2004). *Carnobacterium* sp. are one of the dominant strains found in biogas slurry compost and cow manure compost (Hong-yan et al., 2013). *Carnobacterium* is also present in freshwater habitats such as lake water in Japan (Yanagida et al., 2007).

Vegetal

Carnobacterium divergens is part of LAB strains isolated from mixed pasture of timothy (*Phleum pratense*) and orchadgrass (*Dactylis glomerata*) and its badly preserved silages (Table 2) (Tohno et al., 2012).

Carnobacterium funditum is present in the bacterial community of marine sponge, *Craniella australiensis* (Li and Liu, 2006) and a *Carnobacterium* sp. was isolated from a Sphagnum pond (Leisner et al., 2007).

Recently, a rhizobacterial isolate, *Carnobacterium* sp. SJ-5 (Jain and Choudhary, 2014) was characterized and it was demonstrated that it protected soybean against infection by *Fusarium oxysporum*.

Animals and products of animal origin

For years, the presence of *Carnobacterium* in foods was underreported due to the inability of these bacteria to grow in acetate containing media, such as Rogosa or MRS agar, usually used for LAB numeration and screening. Some modifications and adaptation of conventional microbial techniques (Edima et al., 2007; Holzapfel, 1992; Millière et al., 1994; Wasney et al., 2001) and development of molecular biology techniques adapted to detection of this genus (Brooks et al., 1992; Cailliez-Grimal et al., 2005; Connil et al., 1998; Nissen et al., 1994; Rachman et al., 2004), led to the detection and isolation of these bacteria in a great variety of products.

Due to their ability to grow at refrigerated temperatures and survive at elevated levels of CO_2 in the modified atmospheres, they can proliferate in multiple refrigerated food ecosystems and they frequently dominate in modified atmosphere-packed (MAP) products. Their presence and roles have been extensively studied these last years.

Fish and seafood

Carnobacterium are present in sea and fresh water and are usually found in the normal intestinal microbiota of fish (Pilet and Leroi, 2011). *Carnobacterium inhibens* subsp. *inhibens* was first isolated from the gastrointestinal tract of Atlantic salmon (Joborn et al., 1999). *Carnobacterium funditum*-like species were isolated from the hindgut region of Artic charr

Salvelinus alpinus L. (Ringo et al., 2002) and *C. maltaromaticum* and *C. divergens* from the rainbow trout intestine (Kim and Austin, 2008).

Carnobacterium funditum is present in the microbial ecosystem of cooked tropical shrimps (Jaffres et al., 2009; Matamoros et al., 2009) and one *C. alterfunditum* strain was isolated in MAP rough-head-grenadier (Matamoros et al., 2009).

Carnobacterium maltaromaticum and *C. divergens* are the predominant *Carnobacterium* species isolated from seafoods and eventually dominate the LAB in cold salmon (Leroi et al., 1998). *Carnobacterium maltaromaticum* was frequently isolated from different seafoods including fish products (Pilet et al., 1994), vacuum packaged cold smoked salmon (Duffes et al., 1999), and *C. divergens* was isolated from fish products (Pilet et al., 1994) and fresh cuts of suckling lamb (Oses et al., 2013).

Carnobacterium maltaromaticum is the only *Carnobacterium* species found in altered refrigerated MAP brined shrimp (Mejlholm et al., 2008), vacuum or MAP fresh (Emborg et al., 2002) or smoked salmon (Paludan-Muller et al., 1998), and is present with *C. divergens* in altered refrigerated products such as cooked tropical shrimps (Jaffres et al., 2009) or altered raw salmon (Mace et al., 2012).

Carnobacterium jeotgali was reported to be present in a traditional Korean fermented food, jeotgal, made with freshwater shrimp and salt (Kim et al., 2009).

Meat and meat products

No scientific report indicates the presence of *Carnobacterium* species in the gastro-intestinal or in the skin of animals. Leisner et al., (Leisner et al., 2007) suggested that the source of carnobacteria in meat products is most probably the processing plant. However, DNA from *C. maltaromaticum* was found in the faeces of polar bears, cheetahs and humans and DNA from other *Carnobacterium* species, including *C. divergens*, was found in human infant faeces and in cow rumen (Rahman, 2013).

Meats, meat products and poultry are substrates rich in proteins with water activity and neutral pH favorable for the growth of carnobacteria, reaching high levels (10^6-10^8 cfu.g^1). They are found in vacuum packaged meat and related products stored at low temperatures.

Carnobacterium viridans could be responsible of discoloration of refrigerated vacuum-packaged bologna upon opening the package (Holley et al., 2002).

Carnobacterium divergens was initially isolated from vacuum-packaged raw minced beef (Holzapfel and Gerber, 1983) and was further detected in other meats such as ham (Audenaert et al., 2010), fresh pork sausage (Benson et al., 2014; Curiel et al., 2011), beef (Ercolini et al., 2011, 2009), pork meat juice (Rieder et al., 2012), marinated pork product (Schirmer et al., 2009) and has been associated with alteration of vacuum-packaged beef (Jones, 2004; Youssef et al., 2014). *Carnobacterium maltaromaticum* is also associated with marinated pork product (Schirmer et al., 2009), MAP meat (Nieminen et al., 2011), spoilage microbiota of Belgian artisan-type cooked ham refrigerated at 7°C for four weeks (Vasilopoulos et al., 2010) or refrigerated beef where it can contribute to meat spoilage (Ercolini et al., 2009; Youssef et al., 2014). These two strains dominate among the spoilage LAB of MAP refrigerated broiler legs (Bjorkroth et al., 2005).

Poultry

Thornley in 1957 has isolated atypical *Lactobacillus* strains from irradiated minced chicken. These strains were identified as *C. maltaromaticum* and *C. divergens* in 1987 (Collins et al., 1987).

Carnobacterium maltaromaticum and *C. divergens* are the two most abundant *Carnobacterium* species identified in poultry and in poultry foods. These two species are present in the chicken skin in three manufacturers (Vihavainen et al., 2007) or in coocked MAP refrigerated poultry meat (Barakat et al., 2000). *Carnobacterium maltaromaticum* is also detected in the plant environment. This suggests that it was the origin of the chicken contamination. *Carnobacterium divergens* is detected in MAP sliced chicken or turkey (Audenaert et al., 2010).

Dairy product

Association of *Carnobacterium* with milk is very old. In 1974, the presence of *C. maltaromaticum* in milk is associated with a distinct malty or chocolate like flavor and aroma which correspond to the presence of aldehydes (Miller et al., 1974).

The presence of *C. maltaromaticum* in Brie cheeses, a variety of AOP (protected designation of origin) French surface mould-ripened soft cheese, was reported for the first time in 1994 (Millière et al., 1994; Milliere and Lefebvre, 1994). Out of 30 French soft-ripened or red-smear cheeses made from cow's, ewe's or goat's milk (raw or pasteurized), 10 contains this species (Edima et al., 2007). This species was also isolated from cheese of diverse origins (Rahman et al., 2014a) and associated with Taleggio cheese (Feligini et al., 2012). Two species, *C. divergens* and *C. maltaromaticum* is present in the dominant bacterial community of traditional Mozarella (Morea et al., 1999). *Carnobacterium divergens* is also isolated in French surface mould-ripened cheese (Millière et al., 1994).

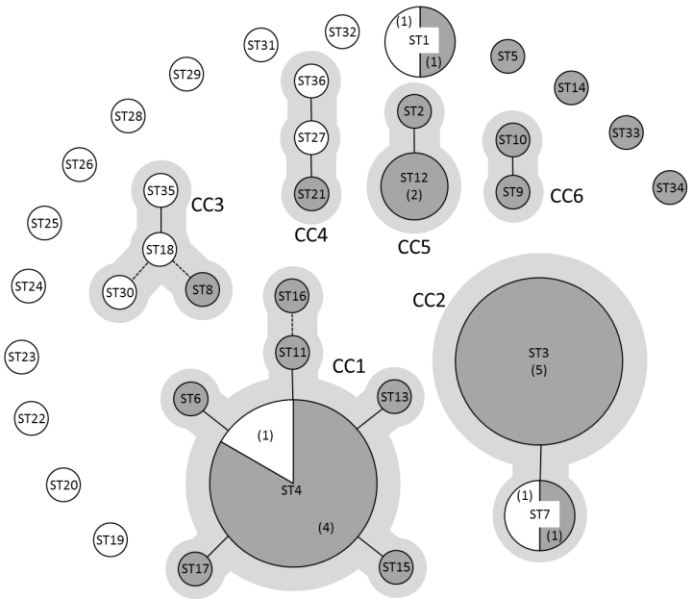


Figure 3: eBURST analysis of 47 *C. maltaromaticum* strains (Rahman et al., 2014).
Each sequence type ST is represented by a circle, the size of the circle is proportional to the number of strains belonging to a given ST. The number of strains is also indicated in brackets when the ST is represented by more than one strain. Dairy and non-dairy strains are represented by dark grey and white circles, respectively. Solid lines connect SLVs (Single Locus Variants) and broken lines connect DLVs (Double Locus Variants). Clonal Complexes (CC) are indicated in light grey.

Although *Carnobacterium* sp. can be isolated from diverse environments, little is known about their population structure. An MLST scheme based on the analysis of fragments of the genes *dapE*, *ddlA*, *glpQ*, *ilvE*, *pyc*, *pyrE*, and *leuS* was applied to a collection of 47 *C. maltaromaticum* strains of diverse origins (Figure 3) (Rahman et al., 2014a). The scheme allowed detecting 36 sequence types (STs) in the collection with eighty percent of the STs represented by a unique strain. Among the remaining STs, several are represented by strains from very diverse biotopes. Besides, eBURST analysis revealed 15 singletons and six clonal complexes. The two major clonal complexes CC1 and CC2 contained 11 and 7 strains, and connected 7 and 2 STs, respectively. The whole population is clustered in four deeply branched lineages, in which the dairy strains were spread.

Pathogenicity

There are some reports in the literature that consider *Carnobacterium* species as a fish pathogen, and *C. maltaromaticum* is usually associated with fish disease. In 1984, Hiu et al., (Hiu et al., 1984) proposed the name *Lactobacillus piscicola* for 17 strains isolated from diseased rainbow trout (*Salmo gairdneri*) and diseased cutthroat trout (*Salmon clarki*) (Michel et al., 1986). Some *C. maltaromaticum* strains were associated with mortalities of cultured striped bass and channel catfish (Baya et al., 1991) and caused important mortalities in market-size rainbow trout (Toranzo et al., 1993). More recently, *Carnobacterium* sp. was isolated from Chinook salmon from Michigan Lake affected with bacterial infections (Loch et al., 2012) and *C. maltaromaticum* seemed to be a significant cause of morbidity and mortality in juvenile salmon shark due to brain infection (Schaffer et al., 2013). The virulence often appears to be low and stressed fish more susceptible (Leisner et al., 2012). *Carnobacterium maltaromaticum* seems to be a common opportunistic pathogen (Michel et al., 1986).

Carnobacterium maltaromaticum was isolated from a case of multi-bacterial synergistic gangrene (Xu et al., 1997) and from the pus from traumatic amputation of the right hand (Chmelar et al., 2002). *Carnobacterium* sp. was also isolated from human blood culture (Hoenigl et al., 2010). However, the causal association with the clinical symptoms remains uncertain. The patient might have become infected through a micro-lesion in his hand while consuming raw or cooked fish (Hoenigl et al., 2010).

No *Carnobacterium* species were direct human pathogen. Their optimal growth temperature below 36.6 °C and growth maxima within 28-40 °C suggest that these bacteria are not pathogenic for warm-blood animals and humans (Pikuta and Hoover, 2014).

The presence of virulence factors in carnobacteria is not well documented. *Carnobacterium viridans* showed β-haemolytic activity on sheep blood agar (Holley et al., 2002). The presence of putative virulence genes encoding products involved in adhesion, capsule synthesis, haemolysis invasion and resistance to toxic compounds were identified in the genome of *C. maltaromaticum* ATCC 35586, isolated from a disease salmon (Leisner et al., 2012).

1.2.4. *Carnobacterium* : positive or negative flora ?

The presence of *Carnobacterium* in foods resulted in an increased number of scientific investigations to study their role and effect on these products. Among the 11 species of *Carnobacterium* genus, only *C. maltaromaticum* and *C. divergens* are frequently isolated from foods. They can be present as natural flora and/or as spoilers under specific storage conditions. The question was asked by Laursen et al., (Laursen et al., 2005), *C. divergens* and *C. maltaromaticum* as spoilers or protective cultures in meat and seafood? Leisner et al., in 2007 (Leisner et al., 2007) noted that *Carnobacterium* can have positive or negative effects in the environment and in foods. However, other authors underlined its positive technological role in cheese making (Afzal et al., 2010).

Spoilers in meat and seafood

Carnobacteria are frequently associated with food spoilage, but spoilage activity shows intraspecies and interspecies variation (Leisner et al., 2007), and depends on the product.

Carnobacterium divergens and *C. maltaromaticum* are involved in the spoilage of meat and seafoods products (Table 1 and 2). They can dominate the spoilage microbiota, particularly in refrigerated food packaged under modified or vacuum atmosphere (Remenant et al., 2015). These bacteria cause sensory spoilage due to the production of volatile compounds (Laursen et al., 2006; Leisner et al., 2007). In refrigerated meat, the alteration caused by *C. maltaromaticum* is due to production of volatile organic compounds, aldehydes, lactones and sulfur compounds (Ercolini et al., 2009). In cold-smoked salmon, *C. maltaromaticum* can also

produce 2,3-butanedione and 2,3-pentanedione, substances which give off a strong butter odor and are not unpleasant (Joffraud et al., 2001).

Some reports noticed the production of biogenic amine by *Carnobacterium*, but the level of tyramine produced varies depending on the food, the environmental conditions and the strains (Leisner et al., 2007; Masson et al., 1996). *Carnobacterium divergens* strains are able to produce tyramine in meat (Curiel et al., 2011; Masson et al., 1999) or seafood product (Laursen et al., 2006). *Carnobacterium maltaromaticum* produces tyramine, detected at the end of the shelf life, in fresh and MAP salmon (Emborg et al., 2002). However, histamine or tyramine are not produced by *C. alterfunditum* in seafood products (Matamoros et al., 2009) or in cheese artificially contaminated with *C. maltaromaticum* LMA28 (Edima et al., 2007).

In conclusion, the presence of *Carnobacterium* strains can be sometimes associated with spoilage of seafood products, but in many cases they are not directly responsible for off-flavors (Pilet and Leroi, 2011).

Positive role in dairy products

Carnobacterium maltaromaticum is the predominant *Carnobacterium* species isolated in dairy products, and has a positive role (Afzal et al., 2010). This strain did not cause off-flavours and could possibly even play a beneficial role in texture and taste during cheese ripening (Afzal et al., 2010; Millière et al., 1994; Milliere and Lefebvre, 1994). In cheeses tested, *C. maltaromaticum* constituted the main psychrotrophic LAB flora. Its growth was able to continue even at an alkaline pH value, reaching high levels (10^8–10^9 cfu.g^{-1}) at the end of a further cold (4 °C) storage period (Cailliez-Grimal et al., 2007). This species constitutes 70% of the curd of Mozzarella cheese and is a citrate-fermenting member of the microflora involved in fermentation (Morea et al., 1999). *Carnobacterium maltaromaticum* LMA28 is a slow acidifying species compared to commercial starter LAB, such as *L. lactis* or *S. thermophilus* (Edima et al., 2008) but can sustain low pH values in co-culture with these starters (Edima et al., 2008). Because of their slow acidifying activity, *Carnobacterium* strains cannot be used as a starter and can therefore be considered as NS (Non Starter) LAB. *Carnobacterium maltaromaticum* produces α-ketoisocaproic acid, 3-methylbutanal and 3-methylbutanol from leucine, 3-methylbutanoic acid and 2-methylbutanal from isoleucine, and 2-methylbutanol, 2-methylpropanal and 2-methylpropanol from valine (Afzal et al., 2013b, 2012). It has been

suggested that these aldehydes convey the malty aroma that is the characteristic of *C. maltaromaticum* in milk.

Protectrice culture

The genus *Carnobacterium* is well known for its ability to produces bacteriocins. Several *Carnobacterium* strains are known to produce anti-*Listeria* bacteriocins (Drider et al., 2006). Six bacteriocins belonging to class IIa, IIc and one cyclic, have been described for different *C. maltaromaticum* strains (Afzal et al., 2010).

Carnobacterium strains have been extensively studied for their role as protective flora against *Listeria monocytogenes* in dairy, meat or fish foods (reviewed by (Leisner et al., 2007)) and cold smoked salmon (review by (Pilet and Leroi, 2011)).

Potential probiotic bacteria?

Carnobacterium maltaromaticum and *C. divergens* are well known as a normal flora of fish intestine and induce the expression of genes encoding IL-1b and TNF-α of head kidney leucocytes of rainbow trout suggesting that these bacteria can stimulate the innate immunity of this fish species (Kim and Austin, 2008). In addition, feeding rainbow trout with these bacteria enhances the cellular and humoral immune response in rainbow trout (Kim and Austin, 2008). *Carnobacterium maltaromaticum* LMA28 was able to survive during the transit through the digestive tract in a murin model and to interact with the host in a cell line model (Rahman et al., 2014b) suggesting that the immunomodulatory properties of *Carnobacterium* are not restricted to interaction with fish and could be extended to mammals.

Carnobacteria, and notably *C. maltaromaticum*, show environmental diversity. Cell Wall Proteins (CWP) are well-known to be responsible of the bacterial interaction with its environment. They are involved in various important processes (Bierne and Cossart, 2007; Burgain et al., 2014a, 2014b). Their characterization would permit to understand why *Carnobacterium* are able to colonize such different environments.

2. Bacterial Cell Wall Proteins

Gram-positive bacteria possess a cell wall composed of several components, *i.e.* peptidoglycan polymer, teichoic acids, lipoteichoic acids and proteins.

Proteins are synthetized in the cytoplasm and possess a signal peptide (SP) on their N-terminal in order to be addressed to the membrane (Blobel, 1980). The secretion through the cytoplasmic membrane is possible *via* several pathways. They are conserved within bacteria: secretion system (Sec), twin-arginine translocation (Tat), flagella export apparatus (FEA), frimbrilin-protein exporter (FPE), pore forming (holin), peptide-efflux ABC and WXG100 secretion system (Wss) (for reviews Desvaux et al., 2009; Driessen and Nouwen, 2008; Lee et al., 2006; van Wely et al., 2001).

The Sec system assures i) the secretion of unfolded proteins across the membrane and ii) the insertion of proteins into the cytoplasmic membrane (Luirink et al., 2005). Proteins are synthetized as precursors containing the maturation protein with an N-terminal SP essential for the signature of protein secretion. The SP has a tripartite structure including a positively charged N-terminus, a hydrophobic core and a neutral or negatively charged C-terminus containing SP cleavage site (von Heijne, 2001).

What differ the secreted proteins and the anchored proteins is the presence of a conserved domain on the C-terminal of the protein. This binding is possible through either the covalent or the non-covalent interactions with the peptidoglycan or secondary wall polymers such as teichoic acids. CWP display a variety of functions. Some are involved in structural support and movement, others in enzymatic activity, and others in interaction with the environment.

The different protein anchorage mechanisms and of their main functions are detailed.

2.1. Proteins anchorage to the Cell Surface

Several types of anchorage systems are used by Gram-positive bacteria for protein attachment to the cell surface. Each type is characterized by the presence of a conserved sequence. Some proteins are non-covalently associated with the cell wall containing GW modules, WxL or LysM domains. Others are designated to be covalently attached to the peptidoglycan layer by sortases *via* a sortase dependent motif, usually called the LPxTG motif (Desvaux et al., 2006; Scott and Barnett, 2006). Several proteins are anchored to the membrane by a hydrophobic tail (Sutcliffe and Harrington, 2002) and some are attached through unknown domains and believed to have cytoplasmic functions (Doyle, 2007; Schaumburg et al., 2004; Trost et al., 2005) (Figure 4).

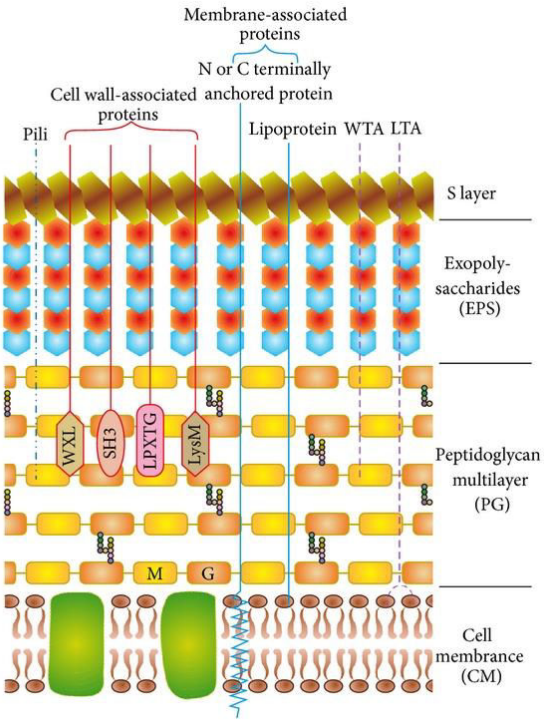


Figure 4 : Cell envelope of lactobacilli with a schematic representation of cell wall and membrane-associated proteins. The Figure was adapted from (Kleerebezem et al., 2010; Sengupta et al., 2013).
The cell membrane (CM) with embedded proteins is covered by a multilayered peptidoglycan (PG) decorated with lipoteichoic acids (LTA), wall teichoic acids (WTA), pili, proteins, and lipoproteins. Exopolysaccharides (EPS) form a thick covering closely associated with PG and are surrounded by an outer envelope of S-layer proteins. The proteins are attached to the cell wall either covalently (LPxTG proteins) or noncovalently (exhibiting LysM, SH3, or WXL domains), lipid are anchored to the CM (lipoproteins) or attached to the CM via N- or C-terminal transmembrane helix. M: N-acetyl-muramic acid; G: N-acetyl-glucosamine.

2.1.1. Proteins non-covalently associated to the cell wall

Proteins can be attached to the cell wall thanks to non-covalent interactions. They present in most cases repeated domains (Bierne and Cossart, 2007).

Lysine motif (LysM) domains are found in many extracellular enzymes involved in bacterial cell wall degradation, and in proteins involved in bacterial cell wall metabolism or binding activities (Bateman and Bycroft, 2000; Buist et al., 2008). Indeed, the C-terminal domain of the *E. faecalis* N-acetylglucosaminidase AtlA, which contains six LysM modules, bind to highly purified peptidoglycan (Eckert et al., 2006). In *L. monocytogenes*, 6 proteins, including P60 and MurA are known to be holding a LysM domain (Bierne and Cossart, 2007). It is also mainly found in *Lactobacillus* sp. (Kleerebezem et al., 2010), *L. lactis* (Steen et al., 2005), *S. thermophilus* (Borges et al., 2005), *Staphylococcus* strains (Hirschhausen et al., 2012) and *Escherichia coli* proteins (Bateman and Bycroft, 2000).

WxL domain (where x can be any amino-acid) allows the protein non-covalent anchorage to the peptidoglycan (Brinster et al., 2007). It is a type of cell wall association domain of 160 to 190 amino acids and it contains two conserved sequence motifs with the Trp-x-Leu signature. WxL proteins belong to the Csc family of surface proteins. *csc* gene clusters encode CscA, a protein with a domain of unknown function and a C-terminal transmembrane anchor, CscB and CscC, which display a C-terminal WxL domain and CscD, a small LPxTG protein (Bierne and Cossart, 2007; Siezen et al., 2006). WxL domains are found in many LAB such as *Lactobacillus* sp. (Boekhorst et al., 2006b; Chaillou et al., 2005; Siezen et al., 2006), *E. faecalis, L. lactis* (Brinster et al., 2007) and *L. monocytogenes* (Bierne and Cossart, 2007).

In addition of having a cell wall binding function, GW modules can exhibit an adhesion function to eukaryotic cells, as exemplified by the protein Ami of *L. monocytogenes* (Milohanic et al., 2001). It is found in many Gram-positive bacteria such as *L. monocytogenes* (Bierne and Cossart, 2007), staphylococcal species (Baba and Schneewind, 1996). The strength of the association of the GW domain with cell wall increases with the number of GW modules (Braun et al., 1997; Jonquières et al., 1999). Thus, the better binding of Ami compared to InlB is due to the presence of 8 GW modules in Ami while InlB possesses only 3 such modules (Braun et al., 1997; Doyle, 2007).

2.1.2. Proteins covalently attached to the peptidoglycan *via* sortases

Surface proteins with LPxTG motif are covalently anchored to the peptidoglycan layer of the cell wall by the sortase family of transpeptidases (Schneewind and Missiakas, 2014). An hydrophobic domain and a tail of positively charged amino-acid generally follow the LPxTG motif (Fischetti et al., 1990; Navarre and Schneewind, 1999).

Sortase catalyzes the formation of a covalent bond between these so called Sortase Dependant Proteins (SDPs) and the peptidoglycan, or between SDPs, depending on the type of sortase enzymes (Maresso and Schneewind, 2008; Rasinkangas et al., 2014; Suree et al., 2007). They are identified in most Gram-positive bacteria, but also in some Gram-negative bacteria (Comfort and Clubb, 2004; Pallen et al., 2001). They can be divided up based on their phylogeny into 6 classes, A to F (Bradshaw et al., 2015; Dramsi et al., 2005; Spirig et al., 2011), the sortase A from *S. aureus* (Sa-SrtA) being the most studied.

Since the discovery of Sa-SrtA (Mazmanian et al., 1999), more than 800 genes encoding related proteins have been identified in approximately 260 distinct bacterial species (Finn et al., 2010). When the full-length precursor protein containing SP is exported through the Sec pathway (Kline et al., 2009), the C-terminal Cell Wall Signal Sequence (CWSS) is then handled by a sortase enzyme. The positively charged C-terminal extremity likely delays export, positioning the protein for processing by the extracellular membrane associated sortase enzyme. The sortase catalyzes a transpeptidase reaction either between SDPs and peptidoglycan, or between SDPs depending on the type of sortase.

The different types of sortases differ by the type of CWSS motifs they recognize of their substrate that they involve in transpeptidation.

Sortase A: Sortase A exhibits a housekeeping function in the cell, because it can anchor a high number of functionally distinct proteins to the cell wall (Marraffini et al., 2006). Sortase A is thus predicted to display 43 distinct proteins in *L. monocytogenes* EGD-e, 22 in *S. aureus* subsp. *aureus* MSSA476, 12 in *L. lactis* IL1403 and 6 in *S. mutans* UA159 (Boekhorst et al., 2005). They are found in bacteria with low GC content such as *Bacillus subtilis, L. monocytogenes* and *S. pneumoniae* (Dramsi et al., 2005; Spirig et al., 2011). Sortase A recognizes the LPxTG motif within the CWSS (Maresso and Schneewind, 2008).

Sortase B: Sortases that belong to sortase B family have a reduced number of substrates compared to sortase A. They can be involved in the attachment of hemoproteins to the cell wall of certain pathogenic bacteria (Maresso et al., 2006; Mazmanian et al., 2002).

To a lesser extent they are also involved in pili assembly: only Spy0129 from *S. pyogenes* is described to be involved in pili assembly (Kang et al., 2011). Pili are extracellular appendages made up of multiple subunits called pilin that promote microbial adhesion and biofilm formation. Pili formation is described below in more details. Sortases B are found in low GC content Gram-positive bacteria, such as *B. cereus, B. anthracis, L. monocytogenes, S. pyogenes* and *Clostridium perfringens* (Bierne et al., 2004; Dramsi et al., 2005; Maresso et al., 2006). These enzymes recognize a different pentapeptide motif in the CWSS of target substrates, which varies slightly between bacterial species. Thus, sortase B recognizes the motif NPQTN instead of the LPxTG motif (Comfort and Clubb, 2004; Mazmanian et al., 2002). Unlike the gene *srtA*, the substrates of *srtB* are encoded by genes that localized in the vicinity of *srtB*.

Sortase C: Class C sortases are found in low GC content Gram-positive bacteria, such as bacilli, clostridia and streptococci, and in high GC content bacteria, such as *Corynebacterium* (Dramsi et al., 2005). There are 24 types of Sortase C enzymes. Sortases C are considered as pilus biogenesis dedicated sortases. These pilus-specific sortases mediate polymerization of pilin subunits via isopeptide bond formation between (i) the threonine and glycine of the cleaved LPxTG motif in one pilus subunit, and (ii) a conserved lysine present within the adjacent pilin subunit (Kang et al., 2009, 2007). Genes encoding class C sortases are often clustered with their substrate encoding genes (Hendrickx et al., 2011).

Sortase D: The Class D sortases are found in many bacilli, clostridia, and actinomycetales (Schneewind and Missiakas, 2014). They are poorly investigated and are believed to play a role in bacteria exhibiting a complex morphological differentiation cycle (Dramsi et al., 2005). These sortases seem to recognize LPNTA sorting signals.

Sortase E – F: The Class E and F sortases were recently defined. Sortase E enzymes are found in some high GC bacterial species instead of Sortase A (Ton-That and Schneewind, 2003). Genes encoding class A and E enzymes are never found in the same organism. In addition, *srtE* gene is not located in the same gene loci with its substrate. Sortase E recognizes and processes substrates containing a LAxTG motif (Comfort and Clubb, 2004). There is at least 57 Class F sortases identified in various *Actinobacteria*. The function of these Class F sortases remains to be determined, it is speculated that they play essential roles in sorting surface proteins (Spirig et al., 2011).

2.2. The function of Cell Wall Proteins

CWP display a variety of functions. Some functions which can be involved in environmental adaptation are described: heme uptake system, permitting the internalizing of iron molecules important for cell functions such as respiration, pili, important for adhesion, and the Cell Envelope associated Proteases (CEP) and thus the proteolytic system responsible of the utilization of milk casein.

2.2.1. Heme utilization

Iron is an essential nutrient for bacterial survival, normally found in the host cell in the form of a complex molecule formed of porphyrin ring that holds a central iron ion, called heme. It is a complex molecule holding the iron in the center. Gram-positive bacteria can sequester heme from host *via* cell wall embedded hemoprotein receptors. They hold an Iron Surface Determinant "Isd" specific for heme uptake. Isd proteins harbor one or more near-transporter (NEAT) domains. Despite low sequence identity, NEAT domains have similar immunoglobulin-like folds. Below, a description for heme uptake by Gram-positive bacteria is detailed.

Heme uptake system

The *S. aureus* heme uptake system has been extensively studied. Four heme-binding proteins (IsdA, IsdB, IsdH and IsdC) are anchored to the cell wall, and three other proteins (IsdD, IsdE and IsdF) are attached to the membrane (Figure 5).

IsdB and IsdH interact physically and capture heme from host Hemoglobin (Hb). The hemic compound is subsequently transferred to the cytoplasmic membrane *via* a heme transfer cascade to IsdC *via* IsdA. IsdC is considered to be involved in translocation of heme through the cell wall (Mazmanian et al., 2002; Zapotoczna et al., 2012). Heme is then internalized across the membrane by a membrane lipoprotein (IsdE) and an ABC transporter (IsdDF) (Hammer and Skaar, 2011). In some cases, heme uptake is facilitated by the action of hemophores. Hemophores, capture free heme or extract heme from hemoglobin in the external medium and present it to specific outer membrane receptors (Ghigo et al., 1997). Indeed, *B. anthracis*

secretes two hemophores IsdX1 and IsdX2 which acquire heme from MetHb (Honsa and Maresso, 2011).

Once in the cytoplasm, heme complex is captured by Heme-degrading monooxygenase proteins IsdG and IsdI that dissociate Heme in order to release iron (Haley et al., 2011) (Figure 5).

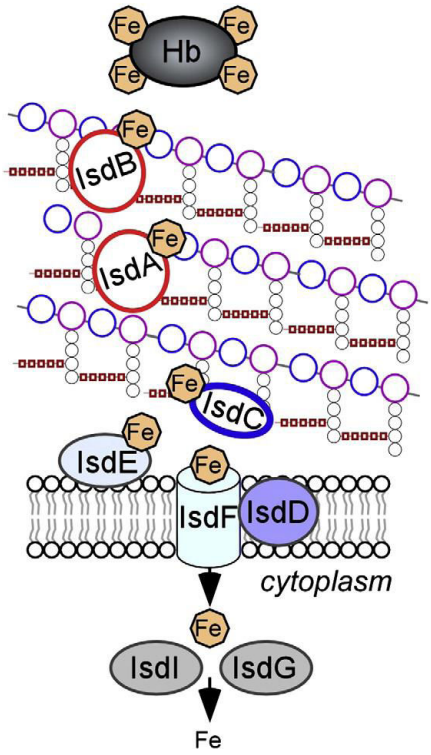


Figure 5 : Isd-mediated heme-iron uptake across the cell wall of *Staphylococcus aureus*.
IsdA, IsdB and receptors for hemoprotein ligands, such as hemoglobin (Hb). IsdC heme receptor, IsdDEF membrane transport system, IsdG and IsdI heme oxygenases (Schneewind and Missiakas, 2014).

Near-transport domains

The presence of NEAT domain characterizes Isd proteins. IsdB possesses two tandem NEAT domains (IsdBN1 and IsdBN2) (Bowden et al., 2014) while IsdH possesses three (IsdHN1, IsdHN2 and IsdHN3) (Pilpa et al., 2009). Each NEAT domain confers to Isd protein a specific linkage with the hemoprotein. Thus, IsdBN1 binds Hb whereas IsdBN2 binds heme (Gaudin et al., 2011; Torres et al., 2006). IsdHN1 and IsdHN2 are found to bind Hb and/or Hp, whereas IsdHN3 binds heme (Dryla et al., 2003; Pilpa et al., 2009).

As for *B. anthracis* hemophores, IsdX1 contains one NEAT domain whereas IsdX2 has five tandem NEAT domains. The *B. anthracis* Isd system also has a membrane-associated

heme-acquisition leucine rich repeat protein (Hal) (Balderas et al., 2012) that contains a NEAT domain capable of binding and acquiring heme from Hb.

Role of sortases in the heme uptake

SrtA recognizes and anchors the heme binding proteins IsdA, IsdB and IsdH, which have LPxTG motifs. SrtB recognizes and anchors only IsdC possessing an NPQTN motif (Balderas et al., 2012; Dryla et al., 2003; Malki et al., 2012; Tiedemann and Stillman, 2012). SrtA promotes association of hemoproteins receptors to the lipid II, and SrtB buries IsdC deeper into the cell wall envelope to preassembled peptidoglycan instead of lipid II (Schneewind and Missiakas, 2014).

2.2.2. Pili

Bacterial pili, also called fimbriae, are defined as non-flagellar, proteinaceous, multi-subunit surface appendages involved in adhesion to other bacteria, host cells, or environmental surfaces (Fronzes et al., 2008; Scott and Barnett, 2006; Telford et al., 2006). Pili have been implicated in critical host-pathogen interactions, colonization, biofilm formation, invasion, and signaling events (Kline et al., 2009; Proft and Baker, 2009). Pili were first studied in Gram-negative bacteria, and they were recently discovered in Gram-positive bacteria. *Corynebacterium* was the first Gram-positive bacterium to be described to have pili in 1968 (Yanagawa et al., 1968). Then, shortly after, it was discovered in streptococci that inhabit the oral cavity (Wu and Fives-Taylor, 2001). Gram-negative bacteria may hold 4 types of pili: Type I also called "*Chaperon-Usher*", type IV pili similar to those of Gram-negative organisms, curli pili and "*alternative Chaperon-Usher*" pili (Kline et al., 2010; Scott and Zähner, 2006). Gram-positive bacteria express pili assembled by sortase enzymes.

Pili structure and biogenesis

Pili in Gram-positive bacteria are composed of a single major subunit called pilin and usually 1–2 accessory subunits called minor pilin. The structure of pili varies between species. Pili are all composed of major pilin subunits, which are polymerized to form the pilus backbone, and of minor pilins, which can be linked to backbone subunits (*ie* major pilins) along the pili. They can also be attached at the tip and/or the base of the pilus. In *B. cereus*, the minor pilin BcpB is localized at the tip of pili (Budzik et al., 2009, 2007). In *S. pyogenes,* pili contain two

types of minor pilins : Cpa is on the tip and Spy0130 is at the base (Quigley et al., 2009). In group B streptococci, the cap protein PilA, and the base protein PilC are also bound along the pili where they are linked to the major pilin subunits PilB (Dramsi et al., 2006; Konto-Ghiorghi et al., 2009). In *C. diphtheriae*, SpaC is only localized on the tip, and SpaB is found on the base and along the backbone formed by SpaA (Mandlik et al., 2008b; Ton-That and Schneewind, 2003). These differences in pili composition result in diverse assembly processes of pili.

Pilus biogenesis is divided into two steps: polymerization of the subunits and anchorage to the cell wall by covalent linkage (Figure 6).

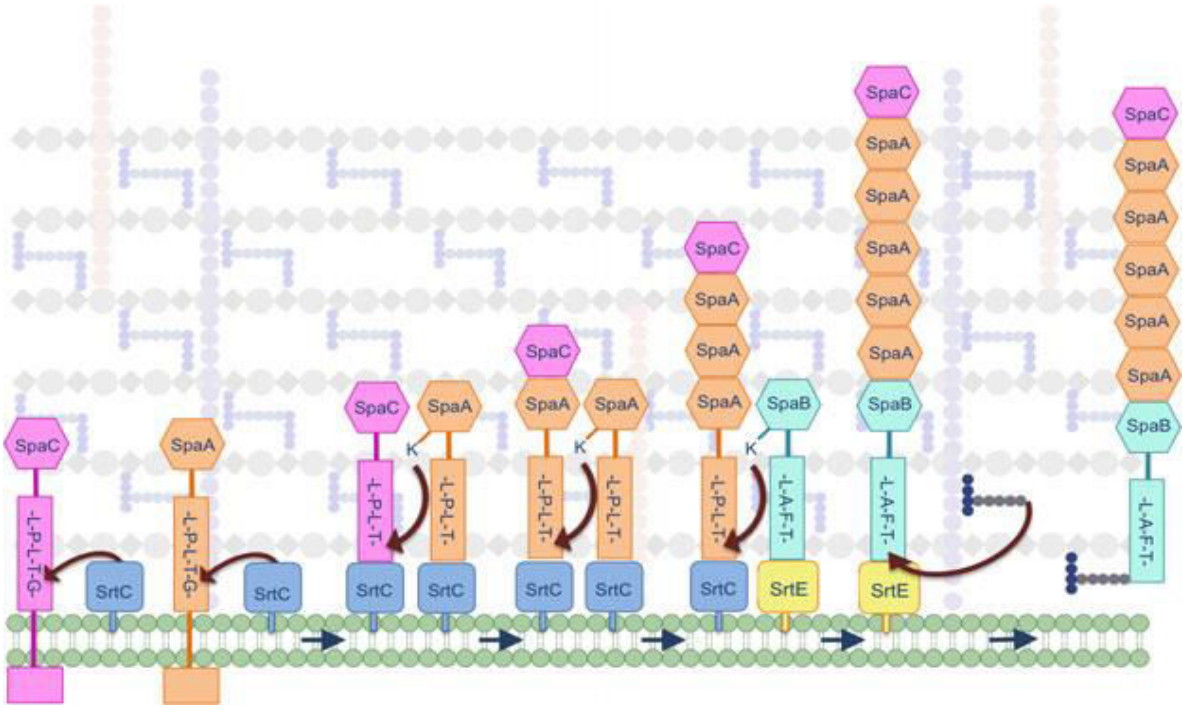


Figure 6 : Model of pilus biogenesis (Bradshaw et al., 2015).
SpaA backbone pillin, SpaC tip pillin, SpaB base pillin, srtA sortase A, srtC sortase C, srtE sortase E, LPLTG and LAFT anchorage domains of pillins.

After export through the Sec pathway, pilin subunits are linked to each other by Class C sortases and subsequently, the whole pilus is linked to the peptidoglycan thanks to Sortase A (Bradshaw et al., 2015). This anchorage model is typical for group B streptococci, where the mutation of the SrtA will lead to the secretion of the pilus instead of its anchorage (Konto-Ghiorghi et al., 2009; Swaminathan et al., 2007). In some cases, like in *S. pneumoniae*, SrtC is responsible for the anchorage of the pilus instead of SrtA (LeMieux et al., 2008).

Gene clusters involved in pili biogenesis

The pilus genes are all clustered in genetic loci and always contain one or more sortase genes. Several Gram-positive organisms such as *C. diphtheriae* and *E. faecium* encode more than one pilin locus (Hendrickx et al., 2008; Ton-That and Schneewind, 2003), although genes of each locus encode components specific for the assembly of a unique pilus type (Gaspar and Ton-That, 2006).

Pilin-specific sortase can be present in one or more copies. In organisms holding only one pilin-specific sortase, the latter is essential for fiber formation and incorporation of all minor subunits (Mora et al., 2005; Nallapareddy et al., 2006). As for organisms with pilus loci encoding more than one sortase, it was shown both functional specificity and redundancy among the enzymes. For instance, in a pilus locus of *S. agalactiae* (Group B *Streptococcus*) which contains two pilin sortases, each sortase was required for the incorporation of a single specific minor subunit (Rosini et al., 2006). *Streptococcus pneumoniae* encodes three SrtC. While deletion of any single SrtC does not abolish pilus formation (LeMieux et al., 2008; Swaminathan et al., 2007), the absence of both SrtC-1 and SrtC-2 eliminates piliation and indicates that SrtC-3 is not sufficient for pilus polymerization (Fälker et al., 2008).

Functions of pili

Pili contribute to the interaction of bacteria with their environment or between them. They possess also a virulent characteristic.

Pili are involved in adhesion. When the bacterium produces a capsule, most CWP are hidden, except pili that can be long enough to cross the capsule (Konto-Ghiorghi et al., 2009; Quigley et al., 2009). Adhesins present in pili interact with several surfaces such as epithelial cells (Dramsi et al., 2006), keratinocyte (Abbot et al., 2007), mucin (Kankainen et al., 2009), collagen, milk proteins (Burgain et al., 2014b) and solid surfaces (Konto-Ghiorghi et al., 2009).

Pili also interact with each other and result in bacterial aggregation (Edwards et al., 2008; Konto-Ghiorghi et al., 2009; Tripathi et al., 2013). *Lactococcus lactis* had been shown to aggregate when it is present in saliva. It was demonstrated that saliva contain a protein rich in cysteine residues that promotes bacterial aggregation (Edwards et al., 2008).

2.2.3. Cell Envelope associated Proteases (CEP) and proteolytic system

CEP are responsible of the utilization of casein in order to assure the amino acid pool necessary for the bacterial growth. It permits to LAB to be independent in milk and thus it is a sign of adaptation to the dairy environment.

LAB are unable to synthetize many of the amino acids required for protein synthesis. Therefore, they necessitate an exogenous source of amino acid and can have complex proteolytic system to liberate the amino acids from the proteins in their environment. The proteolytic system plays a key role in milk and in LAB auxotrophy, and is especially found after adaptation of the bacteria in milk. The same LAB strain found in other environment might be prototrophic for all the amino acids (Georges and François-Marie, 2008). Proteolysis by LAB present several biotechnological benefits in textural and sensorial cheese properties, since amino acids are precursors of several flavor compounds (alcohols, aldehydes, acids, esters and sulfur compounds) and bacterial proteolysis is involved in bitterness reduction (McSweeney, 2004).

Three steps are involved in the proteolysis by LAB. First, in milk, LAB hydrolyze extracellular casein using CEP to form oligopeptides. Secondly, these oligopeptides are transferred to the cytoplasm *via* specific peptides transport systems. Thirdly, intracellular degradation into shorter peptides and amino acids occurs through numerous aminopeptidases (Figure 7) (Christensen et al., 1999; Doeven et al., 2005; Kunji et al., 1996; Savijoki et al., 2006).

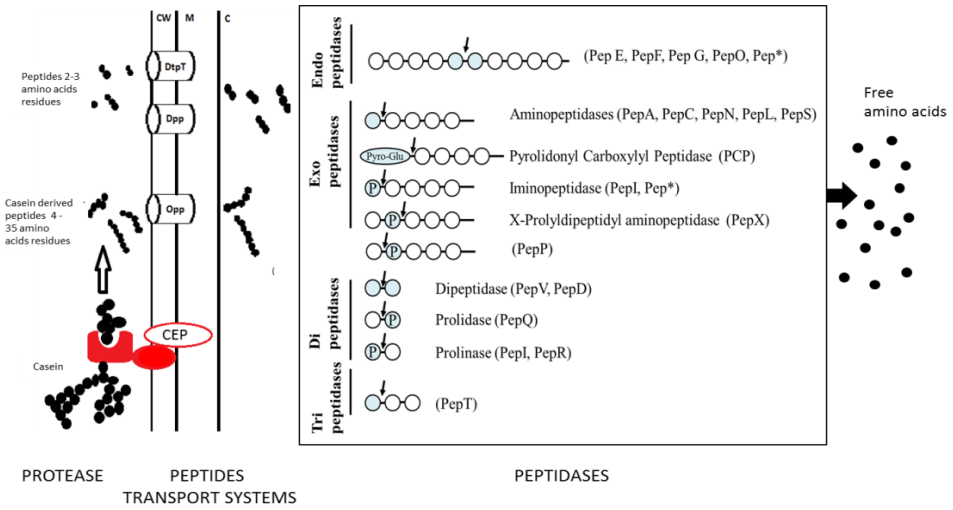


Figure 7 : Proteolysis pathway by LAB, adapted from Doeven (Doeven et al., 2005), Christensen (Christensen et al., 1999) and Savijoki (Savijoki et al., 2006) with schematic representation of the action of peptidases found in lactic acid bacteria.
Amino acids are symbolized by a circle and the arrow indicates the cleavage site. P: proline; *, unclassified Endopeptidases.

Five different major types of CEP were well characterized at the genetic and biochemical levels in LAB (Savijoki et al., 2006): PrtP in *L. lactis* (Kok et al., 1988), PrtH in *L. helveticus* (Smeianov et al., 2007), PrtR in *L. rhamnosus* (Pastar et al., 2003), PrtS in *S. thermophilus*, PrtB in *L. bulgaricus*. Many LAB are known to hold only one Prt system except for *L. helveticus* that exhibit a great diversity of *prtH* genes with 4 paralogs (*prt*H, *prt*H2, *prt*H3 and *prt*H4) among strains (Broadbent et al., 2011; Griffiths and Tellez, 2013; Sadat-Mekmene et al., 2011).

After protein hydrolysis, oligopeptides enter the cells *via* specific transport systems. There are three functional peptide transport systems (Doeven et al., 2005). DtpT, a symport peptides/H^+, is a secondary transporter belonging to the Peptide TRansport (PTR) family of peptide transporters (Doeven et al., 2005). It possesses specificity for di- (and tri) peptides (Hagting et al., 1994). The Opp system is an ABC transporter, which catalyses the uptake of oligopeptide containing 4 up to 35 amino acid residues (Doeven et al., 2005). An *opp* knockout strain was completely blocked in the uptake of oligopeptides, indicating that the *opp* system is essential for the organism to grow on milk (Picon et al., 2000; Tynkkynen et al., 1993). Dpp peptide transport system prefers hydrophobic di- and tripeptides (Foucaud et al., 1995) and depend on ATP. *Streptococcus thermophilus* possesses an ABC transporter named Ami. It is characterized by the transport of oligo and tri peptides (Garault et al., 2002).

After internalization, LAB hydrolyze oligopeptides into free amino acids. Many different peptidases have been characterized biochemically and genetically from *S. thermophillus*, different *Lactococcus* and *Lactobacillus* strains (Christensen et al., 1999). Based on their substrate specificity and site of enzymatic action, peptidases are classified into different groups (Figure 7). Not all peptidases are present in all LAB strains, and a great substrate specificity and variability exists among these peptidases and the same type of peptidase (Christensen et al., 1999; Savijoki et al., 2006; Upadhyay et al., 2004).

In *L. lactis*, proteolysis is controlled at the transcriptional level and includes a pleiotropic repressor CodY. In presence of a high content of intracellular branched amino acids (isoleucine, leucine, valine), codY binds to the codY-box (Guédon et al., 2001; Hengst et al., 2005; Petranovic et al., 2004) located upstream of its target gene leading to a repression of several components of the proteolytic system. However, different mechanism(s) of regulation controlling proteolysis in certain lactobacilli may be present (Azcarate-Peril et al., 2005).

Degradation of proteins other than caseins has been characterized, and several studies have been made in order to characterize LAB species with proteolytic activities and their

incidence on proteolysis of the derived product. Some *Lactobacillus sp.* strains, grown in specific medium, showed the ability to hydrolyze soy proteins. The degree of hydrolysis and the pattern of peptides produced varied among strains (Aguirre et al., 2014; Pescuma et al., 2013). During dough fermentation, wheat proteins (glutenin and gliadins) are degraded by cereal and bacterial proteases, and several sourdough LAB strains with proteolytic activities were characterized and contribute to the degradation of proteins (Gänzle et al., 2007). The proteolysis that occurs during the ripening of meat products is due to the action of the two participants, meat enzymes such as cathepsins and bacterial proteases (Díaz et al., 1997; Fadda et al., 2001).

3. Bacterial adaptation to dairy products: lactose and galactose metabolism

3.1. Introduction

Lactose is a disaccharide consisting of the combination of two simple sugars, α/β-D-glucose and β-D-galactose, linked by a β (1-4) glycosidic bond. It is the main carbon source of milk. Lactose is metabolized by LAB and fermented into lactic acid and/or other final products, mainly ethanol, CO_2 and acetic acid using homofermentative or heterofermentative pathways (Thompson and Gentry-Weeks, 1994). Only *C. divergens* was found to produce formic acid from glucose (De Bruyn et al., 1988).

Because of its industrial relevance, lactose/galactose utilization by LAB has been studied intensively and is well documented (Cocaign-Bousquet et al., 1996; Kandler, 1983; Neves et al., 2005; Wu et al., 2015). The end products of the lactose fermentation confer to food products many properties such as preservation, development of flavors, texture and nutritional values (Leroy and De Vuyst, 2004).

Lactose/galactose metabolism occurs *via* two main pathways: Leloir and Tagatose-6-phosphate (Tagatose-6P) (Kandler, 1983) (Figure 8; Table 3). Lactose and/or galactose enter the cell *via* two systems, permease or phosphoenolpyruvate system (Thompson, 1987; van Rooijen et al., 1991). Lactose is then hydrolyzed *via* a β-galactosidase or a P-β-galactosidase respectively. Glucose enters the glycolysis, galactose the Leloir pathway, and galactose-6P the Tagatose-6P pathway.

This bibliographic synthesis reviews lactose and galactose metabolic pathways and focuses on the diversity within the LAB.

Table 3: Genes implicated in lactose and galactose transportation systems, Tagatose-6P and Leloir pathways. (E. C.: Enzyme Commission; P: phosphate)

Genes (synonyms)	Associated enzyme	E.C. number	References
lacAB	Galactose-6P isomerase	5.3.1.26	(Zeng et al., 2010)
lacC	Tagatose-6P kinase	2.7.1.144	(Zeng et al., 2010)
lacD	Tagatose-1,6 diP aldolase	4.1.2.40	(Zeng et al., 2010)
lacFE	Enzyme II lactose-PTS	2.7.1.69	(Grossiord et al., 1998; Zeng et al., 2010)
lacZ lacLM	ß-galactosidase	3.2.1.23	(Leong-Morgenthaler et al., 1991)
lacA	Galactoside Acetyltransferase	2.3.1.18	(Griffin and Gasson, 1994)
lacG	Intracellular phospho-ß-galactosidase	3.2.1.85	(Abranches et al., 2004; Zeng et al., 2010)
galP (galA)	Galactose permease		(Grossiord et al., 1998; Neves et al., 2010)
lacS	Lactose permease		(Poolman et al., 1989; Schmidt et al., 1989; Vaughan et al., 1996)
galM	Galactose mutarotase	5.1.3.3	(Grossiord et al., 1998; Neves et al., 2010)
galK	Galactokinase	2.7.1.6	(Neves et al., 2010)
galT	Galactose-1P uridyltransferase	2.7.7.12	(Neves et al., 2010)
galE	UDP-glucose4-epimerase	5.1.3.2	(Neves et al., 2010)
galR lacR	Promoteur regulator		(Grossiord et al., 1998)
Pgm	Phospho-gluco-mutase		(Neves et al., 2010)

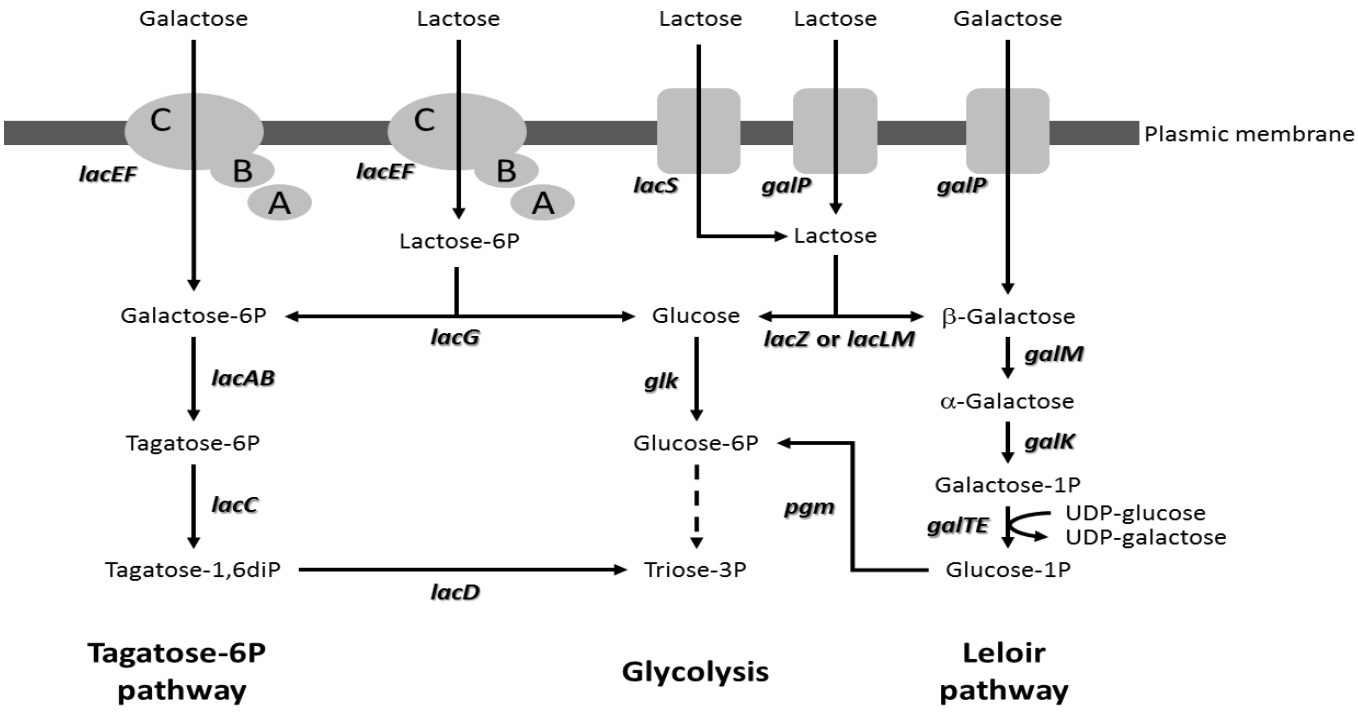


Figure 8: Schematic representation of the lactose and galactose catabolic pathways. (according to (Neves et al., 2010; Price et al., 2012; Zeng et al., 2010)) (for abbreviations see Table 3) The dotted line symbolizes several reactions.

3.2. Sugar uptake systems

There are several type of sugar transport systems, specific or unspecific, which have been extensively studied and documented (Deutscher et al., 2006).

The Tagatose-6P pathway involves a phoshoenolpyruvate (PEP): carbohydrate phosphotransferase system (PTS) which assure the transport and the phosphorylation of the molecule at the expense of PEP.

The PTS-PEP system consists of several membranous sub-units collaborating to transport and phosphorylate the incoming sugar. Indeed, there are three sub-units EFIIA (subunit A), EFIIB C (subunit B) and EFIIC (subunit C), each encoded by a specific gene. Sugars are phosphorylated during transfer across the cell membrane *via* a chain of enzymatic reactions. Each sugar possesses a specific PTS-PEP system, but in some cases, one PTS system has the ability to transport another type of sugar but with lower affinity (Solopova et al., 2012). In the Tagatose-6P pathway, lactose/galactose enter the cell *via* $EIIABC^{Lac}/EIIABC^{Gal}$, a transporter depending on the PTS system encoded by the genes *lacFE* (Abranches et al., 2004; Zeng et al., 2012, 2010).

The Leloir pathway uses an ion-linked sugar transport system. An ion gradient is necessary in order to internalize sugar into the cytoplasm (Poolman, 1993). The galactose can enter the cell *via* a permease encoded by *galP,* and the lactose *via* a permease encoded by *galP* or *lacS*. An ATP hydrolysis through membranous ATPase is coupled to the proton output. A proton force is created through the membrane which leads to the return of protons accompanied by lactose/galactose molecules through a symport specific H^+/sugar (Fortina et al., 2003; Grossiord et al., 1998; Leong-Morgenthaler et al., 1991; Maxwell et al., 1962; Thomas et al., 1980; Thompson, 1980).

The choice of the sugar uptake system is not random. The PTS system may be more favorable than the permease system. It costs only one ATP (in the form of PEP) when transporting and phosphorylating the sugar molecule (two reactions in one single step). In comparison, the permease system consumes more than one ATP molecule (Neves et al., 2005).

3.3. Lactose molecule cleavage: beta-galactosidase

After entering the cell, lactose is cleaved by a hydrolase. The hydrolase type depends on the type of transport used. Phosphorylated lactose which entered *via* a PTS system is hydrolyzed by a phospho-β-galactosidase into glucose and galactose-6P encoded by the gene *lacG*. The symport permeases are associated to a ß-galactosidase responsible for the hydrolysis of lactose in LAB into glucose and ß-galactose (Premi et al., 1972). This enzyme is encoded by *lacZ* or *lacLM* (Figure 8, Table 3).

Those two hydrolases involved in lactose hydrolysis are part of the glycosyl hydrolases families and are not the only enzymes responsible for sugar cleavage in nature. Fourty-five glycosyl hydrolases families with 52 E.C. (Enzyme Commission) entries, classified based on amino acid sequence comparison of 482 sequences, were identified (Henrissat, 1991; Henrissat and Bairoch, 1993). Four families (1, 2, 35 and 42) contain enzymes with β-galactosidase activities. The P-β-galactosidase enzyme encoded by *lacG* is a family 1 member, and the one encoded by *lacZ* is a family 2 member.

Lactose is mainly found in dairy products, but galactose can be present in several forms. It can be found in the anomer form or associated to other compounds such as glucose to give the lactose molecules or to proteins and lipids to give glycoconjugates that are important to the bacterial cells (Frey, 1996). Therefore, bacteria may contain several types of glycosyl hydrolases in order to utilize carbohydrates. Two surface β-galactosidase proteins, coded by *bgaA* (β1,4-galactosidase) and *bgaC* (β1,3-galactosidase), are present in *S. pneumoniae* (Jeong et al., 2009; Zähner and Hakenbeck, 2000). They are not included in lactose metabolism but may be involved in interaction with other bacterial cells, by hydrolyzing the glycan. This may be a sign of pathogenicity. In *C. maltaromaticum* strain BA (Coombs and Brenchley, 2001, 1999) three hydrolyses, α-galactosidase coded by *aga*A (family 36) and two β-galactosidases, coded by *bga*B (family 42) and *bga*C (family35), were present and can degrade polysaccharides.

3.4. Lactose/ galactose metabolism pathways

In tagatose-6P pathway, lactose is absorbed by a specific PTS (*lacFE*) and the resulted lactose-6P is hydrolyzed by a phospho-β-galactosidase (*lacG*) into glucose and galactose-6P (Morse et al., 1968). The latter is converted into tagatose-6P, tagatose-1,6 diP and then into glyceraldehyde-3P and dihydroxyacetone-P by galactose-6P isomerase (*lacAB*), tagatose-6P

kinase (*lacC*) and tagatose-1,6 diP aldolase (*lacD*) respectively (Abranches et al., 2004; Jagusztyn-Krynicka et al., 1992) (Figure 8, Table 3).

Leloir pathway is encoded by a *gal* operon consisting of four genes *galMKTE*. After entering through the permease, lactose is hydrolyzed by a ß-galactosidase (*lacZ* or *lacLM*) into glucose and galactose. Glucose enters the glycolysis while the galactose undergoes several reactions. First, ß-galactose is isomerized by a mutarotase (*galM*) to give α-galactose (Bouffard et al., 1994) which is phosphorylated through galactose-1P kinase (*galK*) to give galactose-1P. The latter is transformed into glucose-1P through an uridyltransferase (*galT*) and UDP-glucose 4-epimerase (*galE*) and then into glucose-6P through phosphoglucomutase (*pgm*) to enter under this form the glycolytic pathway (Figure 8, Table 3) (Maxwell et al., 1962; Poolman et al., 1989; van Rooijen et al., 1991).

3.5. Regulation of the metabolic pathways

There are several genes regulating the metabolic pathways of lactose/galactose, which is subject to complex regulation systems.

In the Tagatose-6P pathway, *lacR* is one of the regulation systems. It is an auto regulation gene that binds to *lacA* promoter and represses its activity when tagatose-6P is absent (de Vos and Vaughan, 1994; van Rooijen et al., 1991). The molecules interfering in the *lacR* type regulation are mostly phosphorylated carbohydrates that are intermediates in the metabolic pathways controlled by regulator protein. In this case, galactose-6P and tagatose-6P have been suggested to bind each one of them to a homologous *lacR* to facilitate the dissociation of its binding site (Zeng et al. 2010). The gene *lacT* is a gene for a transcriptional anti-terminator protein present within the *lac* operon of many LAB, such as *S. gordonii* (Zeng et al., 2012). The gene *lacT* is a member of the BglG/SacY family of proteins commonly associated with the regulation of carbohydrate metabolism (Declerck et al., 2002). It prevents the formation of a terminator to allow RNA polymerase to continue the transcription of downstream genes (Declerck et al., 2002). It also contains two PTS regulatory domains (PRDs) phosphorylated by components of the PTS (van Tilbeurgh and Declerck, 2001). When the PTS for lactose or galactose is engaged in sugar transport, the PRD in *LacT* would be dephosphorylated, allowing for the anti-termination of the *lac* operon genes (Fujita, 2009).

In the Leloir pathway, the *galR* regulation gene is considered as an activator of the *galKTE lacSZLM* operons, but its product is a negative regulator of its own expression. (Vaughan et al., 2001).

It is known that the intracellular levels of fructose-1,6 diP (FBP) and glucose-6P (G6P) are closely associated with the regulation of sugar transport and carbon-catabolite repression (Postma et al., 1993; Vadeboncoeur and Pelletier, 1997). Catabolism of galactose through the Tagatose-6P pathway avoids the upper portion of the glycolytic pathway, and thus the levels of G6P and FBP should be lower than the lactose/galactose metabolism *via* the Leloir pathway exclusively. Thus, lactose/galactose catabolism may have a substantially different impact on physiology and gene expression of strains.

3.6. Genes mutation effects

In order to understand the importance of each gene involved in lactose/galactose metabolism, *lac* and *gal* genes were mutated and their effect on bacterium growth and sugar utilization were studied. Results show that some of the genes, when absent, cause damages to the bacterium.

Every sugar type possesses a specific transport system, but when absent, bacterium cell tries to modulate its genes in order to adapt. *Lactococcus lactis* MG1363 PTS^{lac} deficient strain would utilize lactose through a $PTS^{cel\text{-}lac}$ (PTS system specific for cellulose) (Solopova et al., 2012), when a single point mutation is performed in the promoter region of the *cel* cluster (Barabote and Saier, 2005). *Streptococcus gordonii* PTS^{Lac} mutant failed to grow on lactose, but in medium rich in galactose this strain grows normally with an optical density similar to the wild type (Abranches et al. 2004).

Permeases are known to be selective, and in *L. lactis* MG1363 *galP* deficient strain was not able to ferment galactose (Grossiord et al., 2003).

Mutation of hydrolases inducts the inability of strain to use lactose. *Streptococcus gordonii lacG* deficient strain is impaired to grow on lactose but grew on galactose (Zeng et al., 2012). *Streptococcus salivarius lacZ* deficient strain could not grow on lactose, indicating that intracellular β-galactosidase is essential for lactose metabolism (Chen et al., 2002).

In Tagatose-6P pathway, accumulation of galactose-6P in mutant lacking galactose-6P isomerase (*lacAB*), accumulation of tagatose-1,6-diP in absence of tagatose-1,6-diP aldolase

(*lacD*), cause inhibition of growth on lactose/galactose medium. Contrariwise, *lacC* (Tagatose-6P kinase) mutant showed normal cells growing on medium containing lactose and/or galactose as a source of carbohydrate (Abranches et al., 2004; Zeng et al., 2012, 2010).

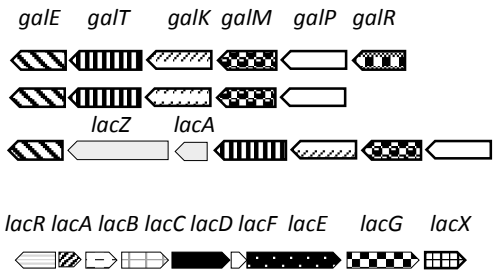
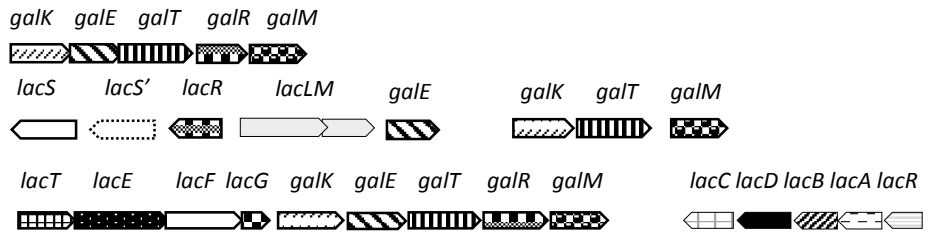
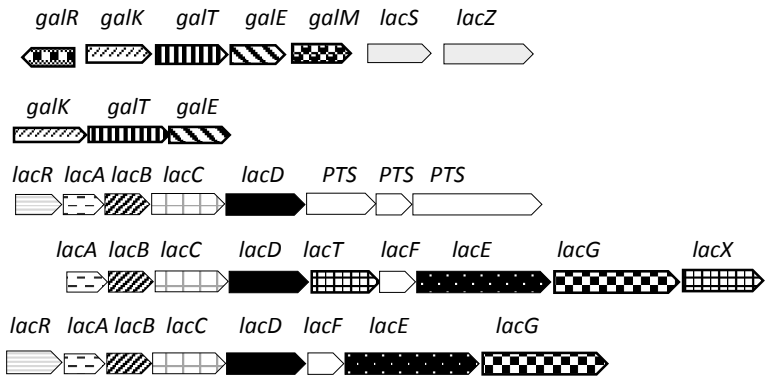
In Leloir pathway, the deletion of *galK* impairs severely the ability of *S. mutans* (Abranches et al., 2004), *S. gordonii* (Zeng et al., 2012) and *S. salivarius* (Chen et al., 2002) to grow on a medium containing galactose although the first two strains possess the two pathways and can metabolize galactose *via* the Tagatose-6P pathway. *Lactococcus lactis galT* deficient strain accumulate galactose-1P which leads to the inhibition of growth on lactose and galactose (Vaughan et al., 1998), but when *galM* is mutate, no difference in growth rate was observed, because the isomerization may be done spontaneously (Neves et al., 2010).

Not all mutations cause damages to bacterium cells. In fact, some dairy strains are Gal negative phenotype due to the defect in the *galK* induction, which seems to be the rate-limiting enzyme of the Leloir pathway (deVos, 1996; Vaillancourt et al., 2002). However, under appropriate selective conditions, such as limited lactose and excess of galactose concentrations, Gal positive derivatives can be obtained (Hutkins et al., 1985; Thomas and Crow, 1984). The differences between Gal^+ and Gal^- strains are in the *galR-galK* promoter, and are due to deletion and substitution of some nucleotides (Anbukkarasi et al., 2014).

3.7. Specificity and genes organization in LAB strains

Lactose and galactose utilization by LAB revealed various combinations of the *gal* and *lac* genes. While the *gal* genes are usually located in chromosome, *lac* genes of the Tagatose-6P pathway may be plasmid or chromosomal. Although genes present high homology between them, the co-localizations are not always the same (Table 4).

Table 4: *lac/gal* genes organization in some LAB strains.
For abbreviations, see Table 3

Strains (references)	Gene organization
Lactococcus garvieae DSM 20684^{T} (Fortina et al, 2009) *Lactococcus lactis* MG1363 (Grossiord et al, 2003) *L. lactis* ATCC 7962 (NCOD 2054) (Vaughan et al, 1998) *L. lactis* plasmid pSK11 (Siezen et al, 2005) pLP712 (Wegmann et al, 2012) *L. lactis* (de Vos and Hugenholtz, 2004)	
Lactobacillus casei 64H (Bettenbrock and Alpert, 1998) *Lactobacillus helveticus* ATCC 15009^{T} (Mollet et Pilloud, 1991; Fortina et al, 2003) *Lactobacillus rhamnosus* GG (Tsai and Lin, 2006)	
Streptococcus thermophilus CNRZ 302 (Vaughan et al, 2001) *Streptococcus gordonii* DL1 (Zeng et al, 2012) *Streptococcus mutans* UA159 (Zeng et al, 2012)	

3.7.1. *Lactococcus* genus

Like mentioned above, in the *Lactococcus* genus, the *galPMKTE* operon of the Leloir pathway is fully distributed on chromosome (Thompson, 1980) with mostly high synteny (Table 4) (example *L. garviae* ATCC49156, *L. lactis* MG1363) (Fortina et al., 2009). This genetic organization reflects the order of the metabolic conversions during galactose utilization *via* the Leloir pathway (Grossiord et al, 2003). Some differences can occurs like in *L. lactis* ATCC7962 (also known as NCDO2054) or *L. lactis* KF147 which possess a *gal* operon where *galT* and *galE* are separated by *lacAZ* (Table 4) (Vaughan et al., 1998). The authors pointed out the greater degree of similarity of the *lac*A and *lac*Z genes to homologs of Gram-negative organisms than to those of Gram-positive bacteria. These results suggest that the genes have been recently acquired within the *gal* operon of the strain *L. lactis* NCDO2054 rather than being lost *via* the generation of a *lac* gene-free *gal* operon as in the case of the strain *L. lactis* LM0230. The procurement of these genes could be through Horizontal Gene Transfer (HTG) or insertion sequence-mediated events (Vaughan et al., 1998).

Among the *Lactococcus* genus*,* some strains of *L. lactis* subsp. *cremoris*, species found in various types of cheeses, may possess both pathways and metabolize lactose *via* a plasmidic *lac* operon specific for the Tagatose-6P pathway (de Vos and Hugenholtz, 2004; de Vos and Vaughan, 1994; Gasson, 1983; Siezen et al., 2005; Wegmann et al., 2012). Contrary to Leloir pathway, the latter is not fully distributed in this genus, probably due to the absence of plasmid in some strains. The large plasmids (47 kb for pSK11L (Gasson, 1983; Siezen et al., 2005), 55 kb for pPL712 (van Rooijen et al., 1991)), encoding the lactose metabolism in *Lactococcus* strains have been shown to possess the same organization (Table 4). A great number of IS elements is also present in these plasmids that may add to the mobility of the plasmid-located genes *via* homologous or site-specific recombination (Bachmann et al., 2012; Siezen et al., 2005; Wegmann et al., 2012).

In Tagatose-6P pathway, lactose enters the cell *via* a specific PTS^{Lac} (Crow et al., 1983). As for galactose, it may enter *via* a plasmidic PTS^{Gal} or a chromosomic *galP* which showed higher affinity for galactose than the PTS^{Gal} (Thompson, 1980). In the absence of the lactose plasmid and when mutating *galP*, lactose and galactose may be internalized *via* other systems (Neves et al., 2010). The plasmid free strains *L. lactis* IL1403 (Aleksandrzak-Piekarczyk et al., 2005) and *L. lactis* MG1363 (Solopova et al., 2012), derived from the industrial dairy starter *L. lactis* NCDO712, are unable to grow in a medium rich in lactose. It is also noted that *L. lactis* strain devoid of its lactose plasmid may utilize lactose *via* alternative pathway.

The plasmid free strain MG1363 may utilize lactose *via* the PTS^{cel} specific for cellobiose uptake in some conditions (Solopova et al., 2012). As for galactose, it has been shown that it may enter the cell *via* a PTS, phosphorylated into galactose-6P, dephosphorylated in order to regenerate galactose that will be metabolized *via* the Leloir pathway (Neves et al., 2010).

3.7.2. *Lactobacillus* genus

Among the *Lactobacillus* genus, the geography of the genes of the Leloir pathway shows high synteny for the *galRTEK* genes (example *L. casei* 64H – Table 4) (Bettenbrock and Alpert, 1998; Grossiord et al., 1998; Wu et al., 2015).

For the thermophilic starter, *L. helveticus,* species of industrially importance for the fermentation of food products, lactose metabolism cluster showed an unusual organization with genes widespread in the genome in comparison to other LAB. *galK* and *galT* genes of the Leloir pathway are transcribed together (Mollet and Pilloud, 1991) while *galE*, *lacS* and *lacR* are organized in one operon (*L. helveticus* ATCC 15009^T Table 4) (Fortina et al., 2003). The presence of putative inserted sequences may be the cause of the presence of *galE* outside the *gal* operon. β-galactosidase, that is 71% similar to that of *L. acidophilus* but only 30% similar to *L. bulgaricus* and *S. thermophilus,* is present (Fortina et al., 2003). *galE* shows high similarity to other LAB such as *L. casei* (Bettenbrock and Alpert, 1998), *S. thermophilus* (Vaillancourt et al., 2002) and *L. lactis* (Vaughan et al., 1998). Downstream *lacS*, an inactivated *lacS'* is present due to a serie of mutations (Fortina et al., 2003). Additionally to the Leloir pathway, in *Lactobacillus* genus, the presence of the Tagatose-6P pathway seems to be species-specific in *L. casei*, *L. paracasei* and *L. rhamnosus* and strains specific in *L. johnsonii* FI9785 and *L. gasseri* ATCC 33323 (Wu et al, 2015).

For those species or strains, the organization of the two pathways are quite similar with genes *lacT-lacE-lacG-lacF* together, sometimes carry on plasmid (Chassy et al., 1978; Chassy and Alpert, 1989) or on the chromosome (Gosalbes et al., 1997), followed or not by the *gal* gene *(galKETRM)* (Tsai and Lin, 2006) upstream the *lacC-lacD-lacB-lacA-lacR* (Example *L. rhamnosus* GG Table 4).

3.7.3. *Streptococcus* genus

In *Streptococcus* genus, *S. thermophilus* is known for its ability to ferment lactose in order to produce lactic acid in yogurt, cheeses and other food products. Other strains, *S. salivarius and S. gordonii*, are considered as human oral bacteria that can colonize teeth and soft tissues of the mouth (Bowden, 2000). *Streptococcus agalactiae* is known to be a human and bovine pathogen (Ben Hamida Nouaili et al., 2011; Zadoks et al., 2011) and *S. mutans* is considered the primary agent of dental caries.

Streptococcus thermophilus is generally found to have an incomplete fermentation of lactose, only glucose is metabolized and galactose is discharged in the medium through *lacS* which is responsible for the lactose/galactose exchange (Crow and Thomas, 1984; Hutkins et al., 1985; Poolman, 1993). Several studies showed that *S. thermophilus* hold mostly the Leloir pathway even though P-β-galactosidase activity was found in some strains (Anbukkarasi et al., 2014). The gene organization and sequences of *gal* operon (*galR-galK-galT-galE-galM-lacS-lacZ*) are the same in 4 strains of *S. thermophilus* possessing for some Gal^+ phenotype and for others Gal^- phenotype; (Table 4) (Anbukkarasi et al., 2014; Vaughan et al., 2001). *gal* and *lac* genes are transcribed separately, and *lacSZ* form an independent operon, different from the *gal* operon (Poolman et al., 1990).

The genome of *S. gordonii* contains 16 genes dedicated to the metabolism of lactose/galactose. Two operons, encoding the Tagatose-6P pathway, were identified, one for the lactose and the other for the galactose. These genes show a great similarity (more than 70% identity) between them (Zeng et al., 2010). It also holds *galKTE* that is not a significant contributor to the use of galactose. For regulation, *S. gordonii* possesses two genes: *lacT* within the lactose operon which possesses positive role in the Tagatose-6P pathway and a negative regulator *lacR* within the galactose operon (Zeng et al., 2012).

Streptococcus mutans can metabolize galactose *via* Leloir pathway (*galKTE*) (Ajdic and Ferretti, 1998) and lactose *via* the Tagatose-6P pathway (*lacABCDEFG*) (Jagusztyn-Krynicka et al., 1992). *galKTE* operon is present upstream of the *msm* operon special for multiple sugar metabolism; the latter carries and metabolize melibiose and raffinose containing galactose residues (Russell et al., 1992). A deletion of the *msm* operon and part of the *gal* operon inhibits the strain's growth on galactose (Robinson et al., 2003). A study showed that *msn* operon is not implicated in the transport of galactose alone in *S. mutans*, and thus a true galactose permease remains to be characterized (Abranches et al., 2004). Although *S. mutans* lacks a *galP* responsible of the galactose uptake (Zeng et al., 2010) it possesses 14 putative PTS with one

lactose specific PTS (Ajdic and Ferretti, 1998). Galactose is transported by other PTS specific for other sugar molecules but with lower affinity (Abranches et al., 2004; Zeng et al., 2010). *Streptococcus mutans* can effectively metabolize galactose in the form of monosaccharide only when both pathways, Leloir and Tagatose-6P, are functional. Tagatose-6P pathway is mainly responsible of the use of the galactose-6P moiety that comes from the hydrolysis of lactose-6P (Abranches et al., 2004).

Streptococcus salivarius metabolizes lactose and galatose *via* the Leloir pathway. It is not yet clear whether there are specific permeases for each type of sugar, or if both elements are transported by the same system (Chen et al., 2002; Postma et al., 1993; Vadeboncoeur and Pelletier, 1997). Organization of the *gal* operon was found to be essentially identical to that of *S. thermophilus* (Vaughan et al., 2001). *galK* and *lacZ* showed 95% similarities with *S. thermophilus*. The *galK*-deficient strain was able to use only glucose and lactose as the sole carbohydrate source, and not galactose. The *lacZ* deficient strain was unable to utilize lactose which gives a justification of the role of β-galactosidase in hydrolyzing lactose into glucose and galactose (Chen et al., 2002).

In order to grow in bovine milk, *S. agalactiae* has been shown to possess 2 *lac* operons in the genome; lac.1 *lacR-IIA-IIB-IIC*-neuraminidase-*ABCDX* and lac.2 *lacRABCDFEGX* (in bovine isolate strains) *lacRABCDTFEGX* (in NEM316 strain) (Richards et al., 2013, 2011). Lac.1 has been suggested to be involved in virulent reactions due to the presence of neuraminidase in the operon (Soong, 2006) and lac.2 is involved in lactose metabolism (Richards et al., 2013).

3.7.4. Lactose metabolism by none LAB bacteria

Lactose and galactose are found in animals and plants, and the metabolic pathways are required for their utilization and the production of metabolites essential for the glycoproteins, glycolipids and cell wall biosynthesis (Frey, 1996).

Lactose and galactose metabolism are not only found in LAB but also in other organisms such as *E. coli*. *Escherichia coli* is a well-known bacteria and many studies had been organized in order to describe its metabolic pathway. *Escherichia coli* showed the capacity of utilizing lactose and galactose *via* the Leloir pathway (Bouffard et al., 1994; Buttin, 1963; Shapiro and Adhya, 1969). It consist of the *galKTEM* operon regulated by *galR*, with *galP* elsewhere.

Lactose and galactose metabolism is also well studied in some food pathogenic bacteria for several reasons: i) to understand the bacterium behavior when present in food products, ii) to inhibit its growth by mutating *gal* or *lac* genes and then to produce vaccines.

Salmonella sp. has been demonstrated to be a pathogenic strains and especially a foodborne pathogen. It has been showed that *Salmonella* possesses a *gal* operon responsible for the galactose metabolism in the order *galE*, *galT* and *galK* (Houng et al., 1990). *Salmonella typhimirium gal* genes had been intensively studied for medical reasons. A simple mutation of *galE* is efficient to obtain antiviral strains used as vaccines against typhoid (Hone et al., 1988).

The presence of the *lacABCDEFG* operon that is capable of metabolizing lactose and galactose in *S. aureus* has been demonstrated (Crow et al., 1983; Rosey and Stewart, 1992). *Staphylococcus aureus* is known to be a pathogenic bacteria-contaminating milk. Lactose metabolism is necessary for the survival of *S. aureus* in medium such as milk, because the non-metabolized D-galactose inhibits bacterium growth (Krispin and Allmansberger, 1998).

4. Objectives

Carnobacterium est un genre bactérien ubiquiste appartenant au groupe LAB. Les souches de *C. maltaromaticum* sont présentes dans une grande diversité d'environnements et sont, avec *C. divergens*, les plus fréquemment isolées de produits alimentaires (Tableaux 1, 2). Ceci laisse supposer une plus grande capacité d'adaptation pour cette espèce comparativement aux autres espèces du genre *Carnobacterium*.

Il a été noté que *C. maltaromaticum* possède un génome plus grand que celui des autres LAB (Hols et al., 2005; Kelly et al., 2010; van de Guchte et al., 2006). La taille des génomes est estimée comme variant de 1,9 Mbp pour *C. alterfunditum* (Pikuta et al., 2005) à 3,3 Mbp pour *C. maltaromaticum* (Cailliez-Grimal et al., 2013). Celle des souches appartenant à d'autres espèces, telles que *Carnobacterium* sp. 17-4 (Voget et al., 2011) et AT7, serait respectivement de 2,6 Mbp et 2,4 Mbp. Ainsi, il y aurait environ 1,4 Mbp de différence entre le plus petit et le plus grand génome connu de différentes espèces de *Carnobacterium*. Il a été émis l'hypothèse que ceci expliquerait l'adaptation des bactéries à une grande diversité d'environnements (Leisner et al., 2007; Wassenaar and Lukjancenko, 2014). Le génome de *C. maltaromaticum* LMA28 possède des gènes évoquant une aptitude à interagir avec les surfaces du tractus gastro-intestinal (Ainsworth et al., 2013; Bolotin et al., 2004; Callanan et al., 2008; Rahman et al., 2014b; van de Guchte et al., 2006). Les bactéries à Gram-positif sont connues pour posséder une grande variété de propriétés de surfaces, qui peuvent jouer un rôle clé dans les interactions avec l'environnement (Burgain et al., 2014a, 2014b).

Actuellement les génomes de trois souches de *C. maltaromaticum* isolés de divers environnements sont disponibles sur la plateforme MicroScope : *C. maltaromaticum* LMA28, isolée de fromage, *C. maltaromaticum* DSM20342 MX5, isolée de lait, *C. maltaromaticum* ATCC35586, isolée de truite malade. Les génomes de *C. maltaromaticum* 3.18, isolée de produit à base de porc, *C. maltaromaticum* ML.1.97, isolée de poisson frais, et de *C. divergens* V41 ont été mis à notre disposition. Les génomes de *Carnobacterium* sp. AT7 et 17.4, souches isolées respectivement des fosses Aléoutiennes et d'eau de mer et celui de *C. inhibens* subsp. *gilichinskyi* WN1359, souche isolée du permafrost, sont en accès libre.

Des analyses et comparaisons de génomes ont été réalisées sur les neufs génomes de *Carnobacterium* afin (i) d'évaluer la spécificité ainsi que la diversité des génomes des différentes espèces de *Carnobacterium* et (ii) d'examiner la variabilité existante au sein des

génomes afin de déterminer les implications éventuelles dans le succès évolutif de l'espèce *C. maltaromaticum*.

Les résultats de ces travaux sont présentés dans le premier chapitre intitulé « Comparative genomic analysis reveals contrasting differences between *Carnobacterium* species».

Les bactéries lactiques jouent un rôle majeur dans les fermentations alimentaires et notamment dans les fermentations des produits laitiers. L'acidification liée à l'utilisation du lactose y est particulièrement importante et la capacité à utiliser le lactose et le galactose est un critère majeur de sélection des bactéries lactiques. Ainsi, les gènes constitutifs des voies du Tagatose-6P et de la voie de Leloir ont été caractérisés chez différentes espèces appartenant au groupe des bactéries lactiques. Les études réalisées indiquent un grand niveau de diversité d'organisation génétique. Cependant, peu d'études ont exploré la diversité de la répartition des gènes des voies du Tagatose-6P et de la voie de Leloir au sein des différents genres de bactéries lactiques.

La population de souches laitières de C. *maltaromaticum* est caractérisée par un très haut niveau de diversité génétique, comme l'a montré une étude MLST réalisée au laboratoire (Rahman et al., 2014a). Les souches laitières sont dispersées au sein des quatre lignées distinctes (I, II, III et IV) mises en évidence par MLST. Cette diversité observée dans l'environnement laitier pourrait se traduire par une diversité des propriétés technologiques des souches dans le contexte fromager.

L'objectif des travaux présentés dans le chapitre 2 est d'analyser si cette diversité se retrouve au niveau d'autres caractéristiques et notamment au niveau des gènes impliqués dans le métabolisme du lactose, les gènes *lac* (voie du Tagatose-6P) et *gal* (voie de Leloir).

Les résultats de ces travaux sont présentés dans deux sous-chapitres, intitulés « Genes associated to lactose metabolism illustrate the high diversity of *Carnobacterium maltaromaticum* » et «Diversity of lactose/galactose metabolic pathways within Lactic Acid Bacteria».

Le dernier chapitre indique les grandes conclusions de ce travail et souligne quelques perspectives.

CHAPTER II: RESULTS AND DISCUSSION

1. Comparative genomic analysis reveals contrasting differences between *Carnobacterium* species

Abstract

Carnobacterium are LAB and two species, *C. maltaromaticum* and *C. divergens*, are the species most frequently isolated from food. Moreover, *C. maltaromaticum* is the dominant species in cheese, where its population is characterized by a high diversity and is able to survive in the digestive tracts of mammals. The genes involved in the adaptation of these bacteria in food and the digestive tract are not well known. Comparative genomic analysis of *Carnobacterium* was performed on the basis of 9 genome sequences: 5 strains of *C. maltaromaticum*, *C. divergens* V41, *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. AT7, and *Carnobacterium* sp. 17.4. Genome analyses revealed that the strains *C. maltaromaticum* and *C. divergens* V41 exhibit a large secretome compared to other *Carnobacterium* strains. The secretome is characterized by a large amount of variable cell-surface proteome, predicted to confer enzyme activity, nutrient-uptake, and adhesive properties that might contribute to the fitness of *C. maltaromaticum* and *C. divergens* in food and in the digestive tract. By contrast, the cell-surface of *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. AT7, and *Carnobacterium* sp. 17.4 seem to be decorated with a very small diversity of surface proteins. These results revealed the very high heterogeneity of the genus *Carnobacterium* regarding the cell-surface and might explain the apparent higher fitness of the two species *C. maltaromaticum* and *C. divergens* in food and the digestive tract.

1.1. Introduction

The genus *Carnobacterium* belongs to the LAB, which are Gram-positive organisms with a low GC content that belong to the phylum *Firmicutes*. It is a member of the *Clostridium-Bacillus* subdivision of Gram-positive eubacteria (Pot et al., 1994). When using sequence analysis of rRNA 16S genes, Wallbanks et al. (Wallbanks et al., 1990) has shown that the species of this genus possess a high degree of similarity (96-98%) and forme a group phylogenetically distinct from other LAB.

Carnobacterium genus consists of 11 species: *C. alterfunditum, C. divergens, C. funditum, C. gallinarum, C. iners, C. inhibens* (with two sub-species, *C inhibens* subsp. *inhibens and C. inhibens* subsp. *gilichinskyi), C. jeotgali, C. mobile, C. maltaromaticum, C. pleistocenium* and *C. viridans* (Snauwaert et al., 2013). They are mesophilic and some species

may be psychrotolerant and grow at 0 °C. Some strains are halotolerant until 8% NaCl and alkaliphilic with growth until pH 9.5 (Cailliez-Grimal et al., 2014; Pikuta, 2014; Pikuta and Hoover, 2014). Their particular physiology could explain their wide distribution in nature. They have frequently been isolated from cold and temperate environment and from animals and foods of animal origin such as seafood, meat and dairy products. They can be present as natural flora and/or as spoilers. Spoilage activity shows intraspecies and interspecies variations, depends on the product, occurs under specific storage conditions (Leisner et al., 2007) and in seafood products they are not directly responsible for off-flavors (Pilet and Leroi, 2011). However, *C. maltaromaticum* positive technological role in cheese making has been underlined (Afzal et al., 2010; Millière et al., 1994; Milliere and Lefebvre, 1994).

Two species, *C. maltaromaticum* and *C. divergens* seem to be particularly well adapted to food and digestive tract (Rahman et al., 2014b). The bacterial cell surface is decorated with a diverse array of proteins that contribute in several functions such as sensing and protecting from environmental stresses, adhesion, and interaction with the immune system (Bierne and Cossart, 2007). These proteins are synthetized in the cytoplasm (Blobel, 1980) and their secretion through the cytoplasmic membrane is possible *via* several pathways. Secreted and anchored proteins differ by the presence or not of a conserved domain on the C-terminal of the protein (Doyle, 2007) and binding is possible through either covalent or non-covalent interactions with the cell wall. Some cell wall proteins enable each bacteria to interact with its environment and their characterization would permit to better understand the success of *Carnobacterium* in colonizing different environments. Moreover, it was already noted that *C. maltaromaticum* possesses larger genome compared to other LAB (Hols et al., 2005; Kelly et al., 2010; van de Guchte et al., 2006). The genome size of *Carnobacterium* species is estimated ranging from 1.9 (for *C. alterfunditum*) (Pikuta et al., 2005) to 3.7 Mb (for *C. maltaromaticum*) (Cailliez-Grimal et al., 2013). It has been suggested that the big *C. maltaromaticum* genome may be the reason of its well adaptation to environmental challenges (Leisner et al., 2007).

At present, the genomes of some *Carnobacterium* strains from different environment such as permanent cold seawater, Aleutian trench, and siberian permafrost, foods such as milk, soft cheese, pork, fish and diseased salmon have been sequenced. A genome-based comparison of different *Carnobacterium* species was made in order to analyze genomic specificity and diversity in relation with their adaptation to different environmental conditions. Cell-surface

proteome, well known to be responsible of the bacterial interaction with its environment, was detailed.

1.2. Material and Methods

1.2.1. *Carnobacterium* strains

The current public database contains nine *Carnobacterium* genomes, with three complete genomes (*C. maltaromaticum* LMA28, *Carnobacterium sp.* 17.4 and *C. inhibens* subps. *gilishinsky* WN1359) including five strains of *C. maltaromaticum.* The genome sequenced were determined in strains from different origins, some from food such as *C. maltaromaticum* strains and *C. divergens* V41, and others from environment (Table 5).

Table 5: Origins of *Carnobacterium* sp.

Strains	Code	Origins	References
C. divergens	V41	Fish	(Pilet et al., 1994)
C. inhibens subsp. *gilichinskyi*	WN1359	Siberian permafrost	(Leonard et al., 2013)
C. maltaromaticum	ATCC35586	Diseased trout	(Leisner et al., 2012)
	LMA28	Soft ripened cheese	(Millière et al., 1994)
	DSM20342 MX5	Milk with malty flavour	(Miller et al., 1974)
	3.18	Porc products	(Leisner et al., 2007)
	ML.1.97	Fresh salmon	(Leisner et al., 2007)
Carnobacterium sp.	AT7	Aleutian trench	(Lauro et al., 2007)
	17.4	Cold seawater	(Voget et al., 2011)

1.2.2. Genome analysis

The sequenced genomes were integrated in the MicroScope platform (Vallenet et al., 2013) to perform automatic and expert annotation of the genes, as well as comparative analysis. MicroScope thus hosts the genome of two *Carnobacterium* sp., one *C. inhibens* subsp. *gilichinskyi* WN1359, one *C. divergens* V41 and five *C. maltaromaticum* strains (Table 5). Gene phyloprofile Tool interface allows searching for common or specific genes/regions between a query genome and other genomes or replicons chosen from the ones available in Prokaryotic Genome DataBase (PkGDB) (*i.e*, (re)annotation of bacterial genomes) or complete proteome downloaded from the RefSeq/WGS sections.

Pan/core genome was calculated using two methods: Pan/core genone tool accessible in the comparative genomics section and using as MICFAM parameter 50 or 80% identity of

the amino acid (aa) on 80% coverage and phyloprofile tool from MicroScope platform with a cut-off of 70% aa identity on 80% coverage with the best Bidirectional Best Hit (BBH).

Synteny, defined as orthologous gene set having the same local organization in species A and in species B, was researched by sequence similarity by BlastP BBH with at least 30% identity on 80% of the shortest sequence (minLrap 0.8) analyses and co-localization.

Metabolic pathways were predicted using Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway and MicroCyc databases. A percentage of 70% minimum identity is used to recuperate the specific and common genes for *C. maltaromaticum*, excluding genes with 30% minimum identity with the four other *Carnobacterium* strains.

Neighbor-joining-based phylogenetic reconstruction was based on the aa sequence of ten essential genes (*dnaK gyrA polA lepA dnaB gyrB secA ftsZ recG ileS*) and was performed using MEGA6 by using the Kimura two-parameter model, including transitions and transversions. The candidate tree was tested with 1,000 bootstrap replications (Tamura et al., 2013). The Type (MLST) was updated using e-BURST analysis from Rahman et al., (Rahman et al., 2014a). the resulting tree was rooted using *Enterococcus faecalis* V583 as outgroup.

Prophages were searched using PHAST Search Tool (Zhou et al., 2011).

1.2.3. Data availability

The annotations were deposited at DDBJ/EMBL/GenBank under the following references: PRJEB9002 for *C. maltaromaticum* ML.1.97, and PRJEB8756 for *C. maltaromaticum* 3.18. The annotations are publicly available for consultation in MicroScope (http://www.genoscope.cns.fr/agc/microscope/home/)

1.3. Results and discussion

1.3.1. Comparative genome analysis of *Carnobacterium* strains

1.3.1.1. Phylogeny

A phylogenetic tree shows that *C. maltaromaticum* and *C. divergens* on one hand, and *Carnobacterium* sp. 17.4, *Carnobacterium* ps. AT7 and *C. inhibens* subps. *gilichinskyi* WN1359 on the other hand, are closely related. They form two monophyletic taxons sharing a common ancestor, as shown by the outgroup *E. faecalis* V583 (Figure 9A). The *C. maltaromaticum* strains LMA28 and ATCC35586 were previously described to exhibit MLST genotypes ST4 and ST36, respectively (Rahman et al., 2014a). Enquiry of the

C. maltaromaticum MLST database revealed the strain *C. maltaromaticum* DSM20342 MX5 exhibits the genoype ST4 as the strain *C. maltaromaticum* LMA28 (Figure 9B).

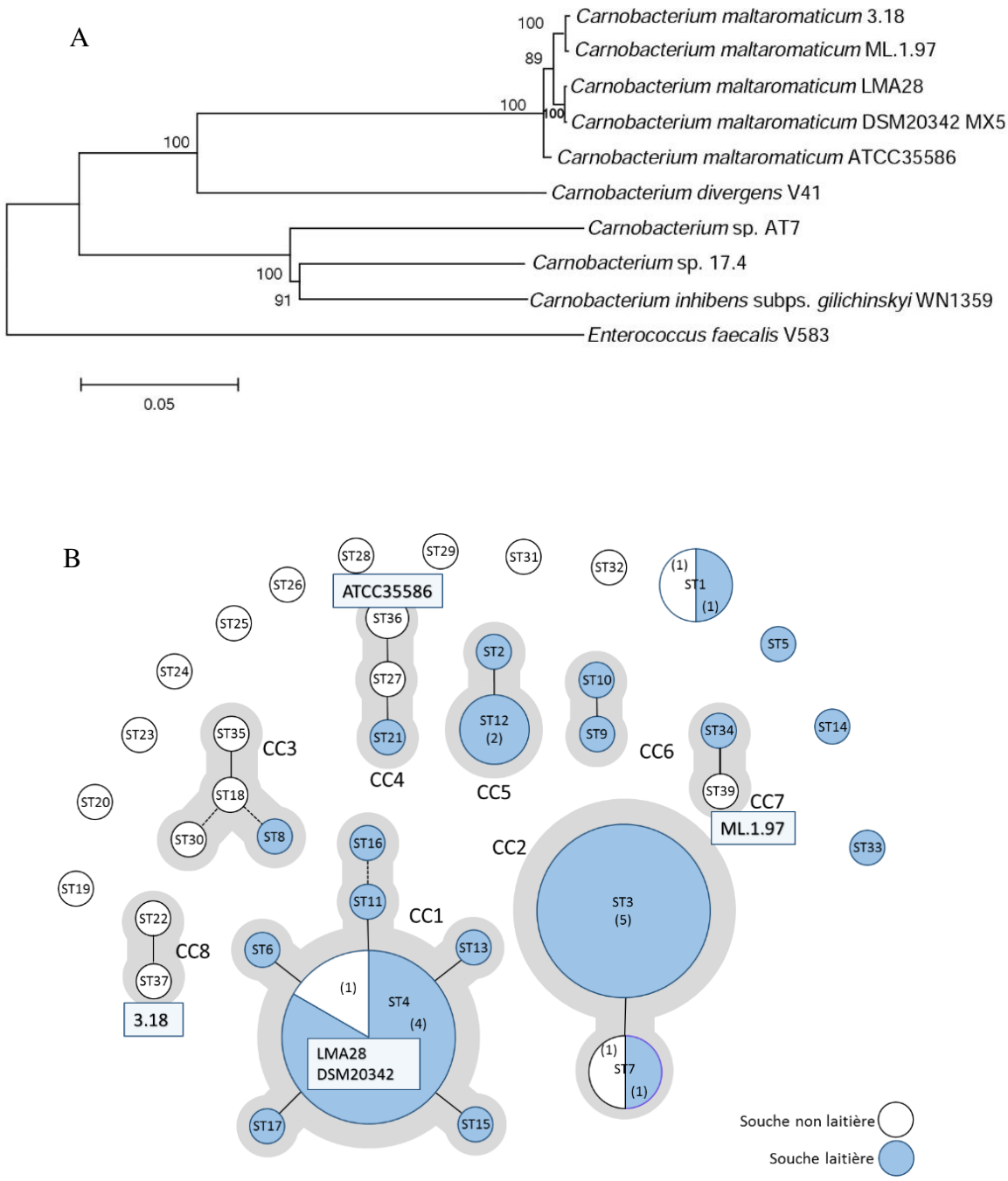


Figure 9: Relationship between *Carnobacterium* genus.
A- phylogenetic tree of the sequenced genomes of *Carnobacterium* based on the conserved nucleic sequence of 10 essential genes (*dnaK gyrA polA lepA dnaB gyrB secA ftsZ recG ileS*).
B- eBURST analysis of 49 *C. maltaromaticum* strains.
Each sequence type ST is represented by a circle, the size of the circle is proportional to the number of strains belonging to a given ST. The number of strains is also indicated in brackets when the ST is represented by more than one strain. Dairy and nondairy strains are represented by bleu and white circles, respectively. Solid lines connect single locus variants, and broken lines connect double locus variants. Clonal complexes (CC) are indicated in light gray.

By contrast, the two strains *C. maltaromaticum* 3.18, and ML.1.97 exhibit previously undescribed genotypes: ST37 and ST39. eBURST analysis revealed that the genotypes ST37 and ST39 are closely related to the previously described genotypes ST22, and ST34, respectively, and thus form two new clonal complexes: CC7 and CC8 (Figure 9B). These results show that strains *C. maltaromaticum* LMA28 and *C. maltaromaticum* DSM20342 MX5 are very closely related and that overall, the strains from this species, for which the genome is available, represent four different clonal complexes: CC1, CC4, CC7, and CC8 (Figure 9B).

1.3.1.2.General features

The general features of the nine genome sequences are compiled (Table 6). The genomes of the three strains *C. maltaromaticum* LMA28, *C. inhibens* subsp. *gilichinskyi* WN1359 and *Carnobacterium* sp. 17.4 are complete. The draft genome sequence of the other strains *C. maltaromaticum* DSM20342 MX5, *C. maltaromaticum* ATCC35586, *C. maltaromaticum* ML.1.97, *C. maltaromaticum* 3.18 and *Carnobacterium* sp. AT7 have a number of contigs comprised between 5 to 229 (Table 6).

Table 6 : General features of *Carnobacterium* genomes.
*** intact and **region incomplete**

Organisms	*C. maltaromaticum*					*C. divergens*	*C. inhibens* subsp. *gilichinskyi*	*Carnobacterium* sp.	
	LMA28	**DSM20342**	**ATCC35586**	**ML.1.97**	**3.18**	**V41**	**WN1359**	**AT7**	**17.4**
Sequence length (Mbp)	3.65	3.877	3.54	3.33	3.57	2.74	2.35	2.45	2.63
GC content (%)	34.5	34.4	34.5	34.4	34.4	35.3	35.2	35.2	35.2
Number of plasmids	3	ND	ND	ND	ND	ND	5	ND	1
Number of CDS	3,687	3,812	3,448	3,368	3,465	2,633	2,268	2,431	2,584
Number of COG	2,671	2,800	2,636	2,639	2,672	2,089	2,257	1,986	2,155
fCDS	48	12	12	49	8	15	69	12	25
Number of tRNA	59	64	61	37	59	8	75	71	67
Number of 16S-RNA	6	6	1	1	3	1	8	7	8
Prophage clusters	2* (3**)	2 (2)	1 (4)	1	1 (5)	0 (3)	0	0 (3)	0
CRISPR	0	0	0	0	0	8	0	0	0
Scaffolds	1	5	74	229	160	32	1	69	1
Contigs	1	5	74	229	160	32	1	69	1

The genome of *C. maltaromaticum* strains ranges from 3.33 to 3.87 Mbp. By contrast, the genome size of the other *Carnobacterium* ranges from 2.35-2.74 Mbp, with the smallest being the genome of *C. inhibens* subsp. *gilichinskyi* WN1359. The genome of *C. maltaromaticum* strains is 0.8-1.52 Mbp larger than the genome of other *Carnobacterium* strains (Table 6).

Accordingly, the number of predicted CDS in each genome ranged from 3,368 to 3,812 for *C. maltaromaticum*, and from 2,268 to 2,633 for the other *Carnobacterium* (Table 6).

In LAB, high genome size variations are also observed for *Enterococcus* and *Lactobacillus* for which the genomes range from 2.5 to 3.5 Mbp and from 1.3 to 3.4 Mbp, respectively (Douillard and de Vos, 2014; Wassenaar and Lukjancenko, 2014). Compared to other LAB genomes accessible in public databases, the genome of *C. maltaromaticum* and more specifically of *C. maltaromaticum* DSM20342 MX5 is the largest. The genome size differences between *C. maltaromaticum* and the other *Carnobacterium* species suggest massive gene gains in the ancestor of *C. maltaromaticum* or massive gene losses, both in the ancestor of *C. divergens* V41 and in the ancestor of *Carnobacterium* sp. 17.4, *Carnobacterium* sp. AT7, and *C. inhibens* subsp. *gilichinskyi* WN1359. The most parsimonious hypothesis would be a massive gain of genes.

Plasmids

The complete genome sequences of *C. inhibens* subsp. *gilichinskyi* WN1359 and *Carnobacterium* sp.17.4 are described to contain five and one plasmid(s), respectively. Overall, the amount of extrachromosomal DNA is of 148 kbp and 50 kbp for *C. inhibens* subsp. *gilichinskyi* WN1359 and *Carnobacterium* sp.17.4, respectively (Leonard et al., 2013; Voget et al., 2011). The genome of *C. maltaromaticum* LMA28 contains one megaplasmid of 97.3 kbp, and two plasmids of 58.9 kbp, and 12 kbp. Globally, this strain contains 168.2 kbp of extrachromomal DNA. The megaplasmid contains the *lac* genes, which encode the Tagatose-6P pathway allowing to metabolize lactose and galactose. Similarly to *L. lactis,* by carrying the *lac* genes, this megaplasmid could play a role in adaptation to the dairy environment.

Prophages

The *Carnobacterium* genomes of all strains contain one to four prophages and/or prophage remnants, ranging in size from 11.5 to 74.4 kbp (data not shown) expect for two strains *C. inhibens* subsp. *gilichinskyi* and *Carnobacterium* sp. 17.4. In dairy niches, phage predation is ubiquitous as shown in many studies of LAB. Indeed, prophages and phage remnants can make up a substantial fraction of novel sequences in the genome (Pfeiler and Klaenhammer, 2007).

CRISPR-cas

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) (Barrangou et al., 2007), and the CRISPR associated (cas) genes confer resistance to phage infection. The role and mechanistic of the CRISPR-cas system in bacterial species have been extensively studied and indicate that the spacer sequences can be considered as a signature of past exposure to exogenous DNA. Among all *Carnobacterium* genomes, *C. divergens* V41 is the only strain having CRISPR-cas system. More precisely, two loci were predicted in the genome of this strain. The presence of these genes suggests that strain *C. divergens* V41 would have systems acting as barriers against HGT, thereby that would limit genome expansion in this taxon.

COG

More than 70% of the CDS encode proteins that can be classified in at least one Cluster of Orthologous Group (COG). All CDS are distributed into 21 different classes of COG. The distribution of the genes of each species of *Carnobacterium* in the different COG groups was performed and shows no significance differences between the strains (Figure 10).

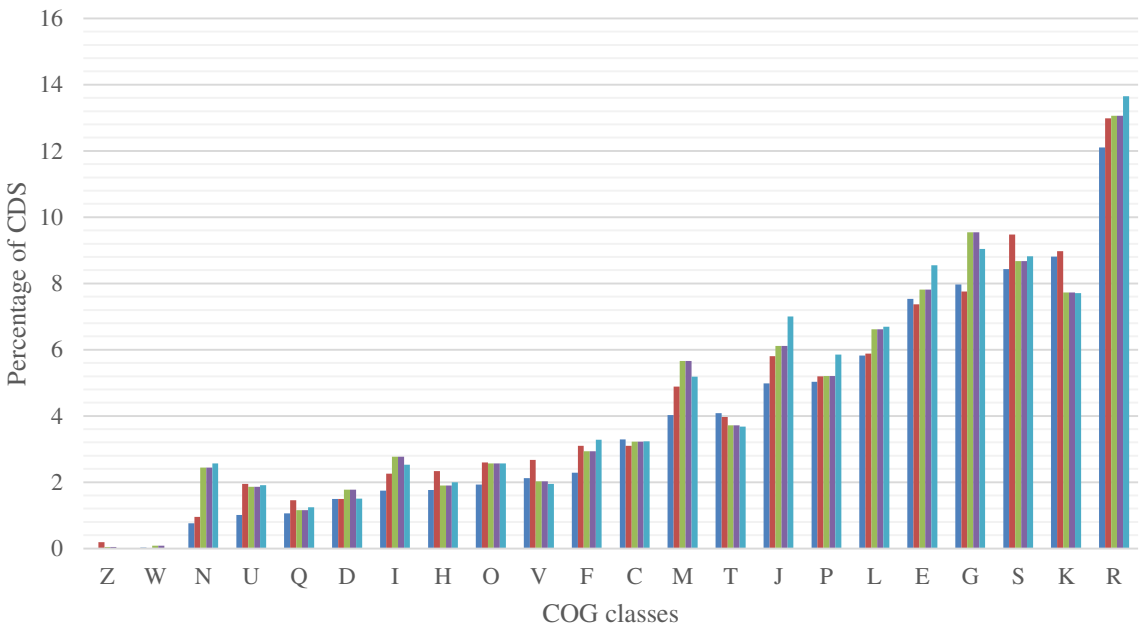


Figure 10: COG distribution in the different classes for 5 strains of *Carnobacterium*. *C. maltaromaticum* LMA28 (dark bleu), *C. divergens* V41 (red), *Carnobacterium* sp. 17.4 (green), *C. inhibens* subsp. *gilichinskyi* WN1359 (light bleu).
C Energy production and conversion; D Cell cycle control, cell division, chromosome partitioning; M Cell wall/membrane/envelope biogenesis; N Cell motility; T Signal transduction mechanisms; U Intracellular trafficking, secretion, and vesicular transport; V Defense mechanisms; W Extracellular structures; J Translation, ribosomal structure and biogenesis; K Transcription; L Replication, recombination and repair; C Energy production and conversion. E Amino acid transport and metabolism; F Nucleotide transport and metabolism; G Carbohydrate transport and metabolism; H Coenzyme transport and metabolism; I Lipid transport and metabolism; P Inorganic ion transport and metabolism; Q Secondary metabolites biosynthesis, transport and catabolism; R General function prediction only; S Function unknown.

1.3.2. Pan/core genome

The pan/core genome analysis of the nine *Carnobacterium* (50% aa identity on 80 % coverage) shows that the core genome represents 1,130-1,215 (31-47%) of the predicted CDS (Table 7A). When comparison was made with *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. 17-4, and *Carnobacterium* sp. AT7, the core genome increased up to 1,745-1,785 (67-73 %) of the predicted CDS (Table 7B). Similarly, when considering the five *C. maltaromaticum* strains and *C. divergens* V41, the core genome increased up to 1,723-1,825 (68-76%) of the predicted CDS (Table 7C). This result is in agreement with the phylogenetic tree showing close phylogenetic proximity of *C. maltaromaticum* and *C. divergens* on the one hand, and of *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. 17.4, and *Carnobacterium* sp. AT7 on the other hand (Figure 9).

Table 7: Comparative analysis of the Pan/Core genome (50-80) between A: all the *Carnobacterium* genomes, B: *Carnobacterium* sp. and *C. inhibens* subsp. *gilichisnkyi* WN1359 genomes and C: *Carnobacterium divergens* and *C. maltaromaticum* genomes.

	Organism	Pan	Core	Variable	Strain specific	Core (%)	Variable (%)	Strain specific (%)
A	*C. maltaromaticum* DSM20342 MX5	3822	1215	2607	447	31.79	68.21	11.695
	C. maltaromaticum LMA28	3883	1205	2678	508	31.033	68.967	13.083
	C. maltaromaticum ATCC35586	3447	1195	2252	450	34.668	65.332	13.055
	C. maltaromaticum ML.1.97	3367	1200	2167	506	35.64	64.36	15.028
	C. maltaromaticum 3.18	3462	1209	2253	223	34.922	65.078	6.441
	C. divergens V41	2618	1139	1479	727	43.506	56.494	27.769
	C. inhibens subsp. *gilichisnkyi* WN1359	2427	1164	1263	346	47.96	52.04	14.256
	Carnobacterium sp. AT7	2418	1130	1288	390	46.733	53.267	16.129
	Carnobacterium sp. 17.4	2626	1162	1464	455	44.25	55.75	17.327
B	*C. inhibens* subsp. *gilichisnkyi* WN1359	2427	1782	645	372	73.424	26.576	15.328
	Carnobacterium sp. AT7	2418	1745	673	410	72.167	27.833	16.956
	Carnobacterium sp. 17.4	2626	1785	841	532	67.974	32.026	20.259
C	*C. maltaromaticum* DSM20342 MX5	3822	1825	1997	460	47.75	52.25	12.036
	C. maltaromaticum LMA28	3883	1811	2072	537	46.639	53.361	13.83
	C. maltaromaticum ATCC35586	3447	1792	1655	455	51.987	48.013	13.2
	C. maltaromaticum ML.1.97	3367	1801	1566	514	53.49	46.51	15.266
	C. maltaromaticum 3.18	3462	1825	1637	231	52.715	47.285	6.672
	C. divergens V41	2618	1723	895	762	65.814	34.186	29.106

To be more restrictive, the pan/core genome analysis was performed with a cut-off of 70% aa identity on 80% coverage. It revealed that the strains of *C. maltaromaticum* share 2,665 CDS, and each of the five strains possesses between 279 and 644 specific genes (Figure 11) indicating substantial strain-to-strain variations.

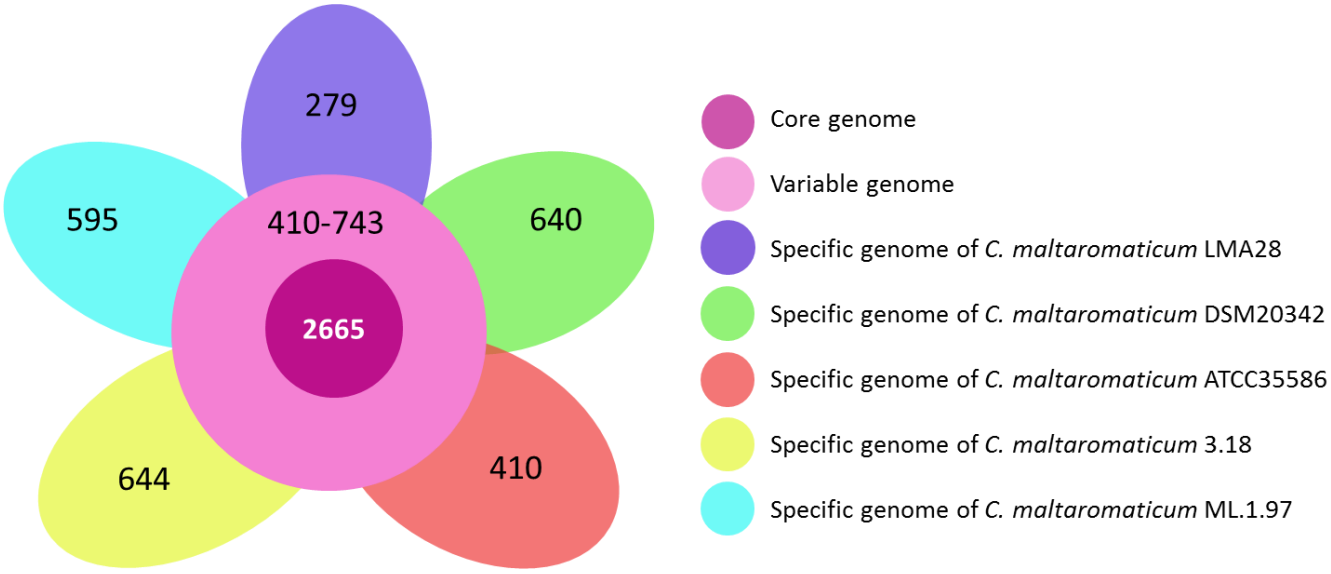


Figure 11: Pan/Core genome analysis of *Carnobacterium maltaromaticum* strains with a cut-off of 70% amino-acid identity on 80% coverage.

1.3.2.1.Intraspecific core genome

In order to identify the genes characterizing the difference between the species *C. maltaromaticum* and the other species, the set of genes commonly found in the strains of *C. maltaromaticum* without homologues in other *Carnobacterium* species was delineated. To do so, homologues were identified by using a cut-off of 70% identity at the aa level between the *C. maltaromaticum* strains and among these homologues, the ones that are specifically found in *C. maltaromaticum* were predicted by excluding all *Carnobacterium* CDS showing more than 30% identity with those of *C. maltaromaticum*. This analysis allowed to delimit a set of 433 CDS defined as the *C. maltaromaticum* intraspecific core genome.

COG analysis of the intraspecific core genome

The COG comparison analysis between the *C. maltaromaticum* intraspecific core genome and the whole core genome of *C. maltaromaticum* was performed. The comparison revealed that several classes were over-represented in the intraspecific core genome: the class E, K, R, S and T (Figure 12). The more contrasting classes are class S and R, which represent

together more than 50% of the CDS of the intraspecific core genome. These classes contain poorly characterized genes or genes with unknown functions, indicating that efforts have to be made in order to understand adaptation factors of *C. maltaromaticum*. The class E contains genes involved in the metabolism of amino acids. Classes K and class T are involved in transcription and signal transduction, which are functions allowing to adapt to variations of the environment. This suggests that *C. maltaromaticum* would have a higher ability to adapt to environment changes than other *Carnobacterium.*

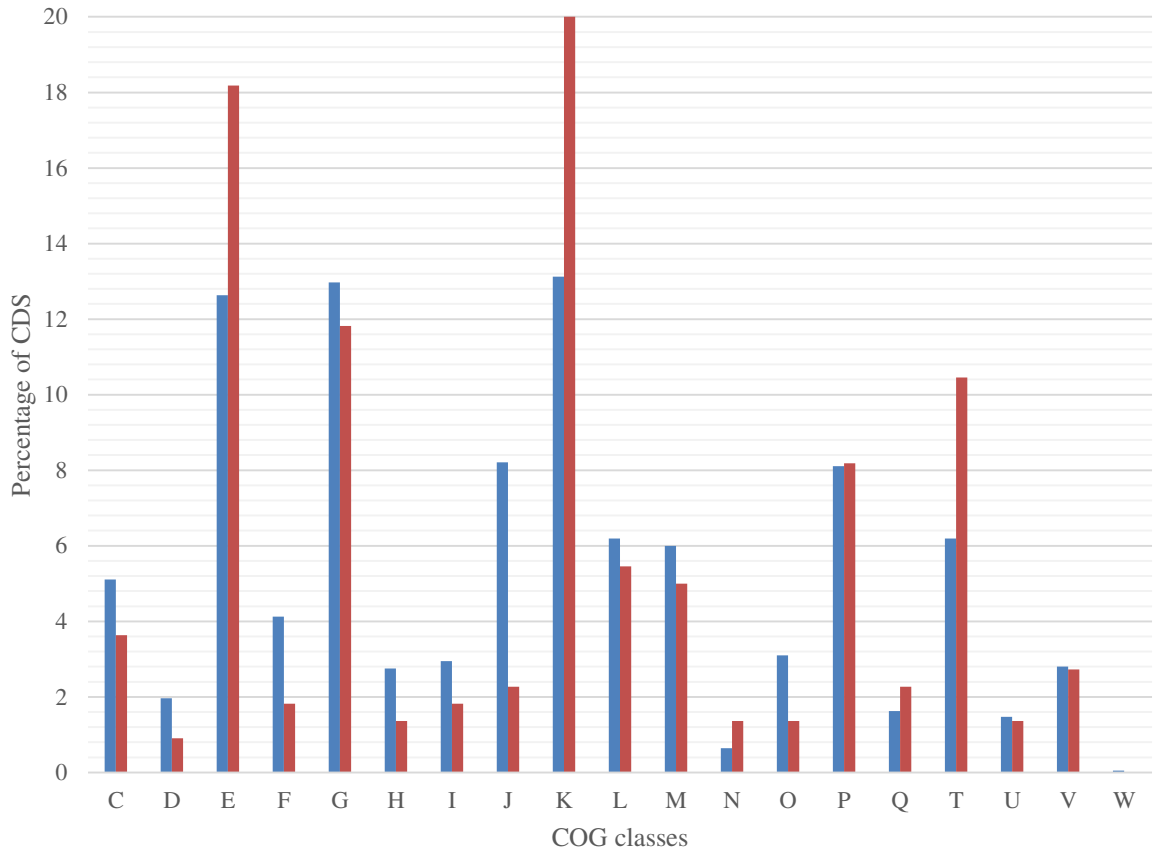


Figure 12: Comparison of the distribution of the Core genome (bleu) and the intraspecific Core genome (red) in the different COG classes.

1.3.2.2.Metabolic pathways

The *C. maltaromaticum* intraspecific core genome contains CDS encoding pathways for tryptophan biosynthesis and ethanolamine degradation.

Tryptophan is one of the most expensive aa produced in the cell. To synthesize tryptophan, *C. maltaromaticum* involves five enzymes encodes by seven genes. Homologues of these genes are arranged as a single cluster *trpEGDCFBA*. The presence of these genes might

confer to *C. maltaromaticum* a selective advantage over other *Carnobacterium* in environments depleted in amino acid and amino acid-containing nutrients.

Ethanolamine is a compound derived from the degradation of phosphatidylethanolamine, an abundant phospholipids in membranes, by the action of phosphodiesterases (Randle et al., 1969). Ethanolamine is highly found in the intestine due to its presence in the cell membranes and in human diet (Randle et al., 1969; Roof and Roth, 1988). Indeed, it is the result of the turnover and exfoliation of cells (Cotton, 1972). Bacteria that carry *eut* genes allowing degradation of ethanolamine are found mainly in the intestine. The *C. maltaromaticum* intraspecific core genome contains the two central genes for ethanolamine utilization *eutBC* as well as 14 additional genes predicted to be involved in this metabolic pathway. These genes could be involved in adaptation of *C. maltaromaticum* to the intestinal medium.

1.3.3. Secretome

All genomes of *Carnobacterium* contain homologs of genes encoding a functional Sec pathway system, while no homologue of the Tat pathway was found (data not shown). The secretome was therefore predicted by identifying genes suspected to encode signal-peptide containing proteins. The comparison of the resulting secretome encoded by the *Carnobacterium* genomes revealed that strains could be divided into two groups on the size criterion. On one hand, the strains of the species *C. maltaromaticum* and *C. divergens* exhibit a large secretome by containing between 272 and 375 predicted proteins. On the other hand, the three strains *Carnobacterium* sp. AT7, *Carnobacterium* sp. 17.4, and *C. inhibens* ssp. *gilichinsky* WN1359, would have a smaller secretome with a number of proteins comprised between 131 and 141. Compared to other LAB, the size of the secretome of *C. maltaromaticum* and *C. divergens* are among the biggest (Figure 13).

A COG analysis revealed that the number of secretome genes is in general higher in *C. maltaromaticum* and *C. divergens* compared to other *Carnobacterium*. More specifically, COG families E, G, M, P, R, S, and T are highly represented in *C. maltaromaticum* and *C. divergens* compared to the other *Carnobacterium*. This suggests that extracellular functions are more diversified in *C. maltaromaticum* and *C. divergens* than the other *Carnobacterium.* For instance, *C. maltaromaticum* and *C. divergens* exhibit between 22 and 35 genes in COG R, *i.e.* general function prediction, while the three other strains *Carnobacterium* sp. AT7,

Carnobacterium sp. 17.4, and *C. inhibens* ssp. *gilichinsky* WN1359, would have less than 10 genes encoding a SP within this COG subclass (Annexe 1).

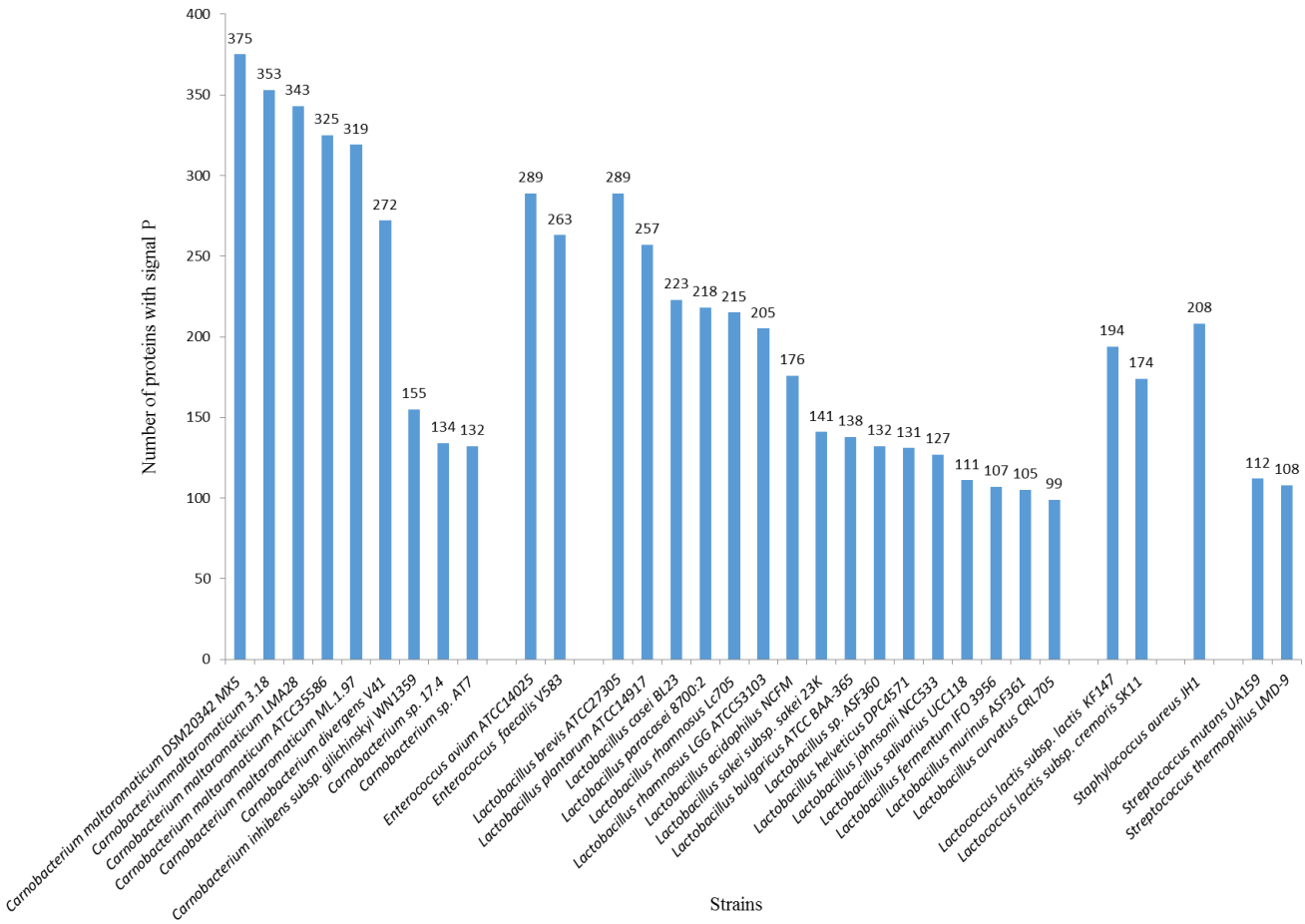


Figure 13: Secretome size in the different *Carnobacterium* genomes, LAB genomes and other Gram-positive bacteria.

More strikingly, the strains of *C. maltaromaticum* and *C. divergens* V41 contain between 116 and 155 unclassified secreted proteins, while the other three strains would contain between 0 and 23 unclassified secreted proteins. The detailed analysis of the gene content of those unclassified genes, revealed that approximately half of them are hypothetical proteins or conserved exported protein of unknown function, and 30% encode secreted proteins associated to the cell wall. These results strongly suggest that *C. maltaromaticum* and *C. divergens* compared to *Carnobacterium* sp. AT7, *Carnobacterium* sp. 17.4, and *C. inhibens* ssp. *gilichinsky* WN1359 are highly contrasting regarding their surface, the surface of *C. maltaromaticum* and *C. divergens* V41 being much more decorated with diversified proteins. Therefore, the genes linked to the cell wall were analyzed in more details.

1.3.3.1.Cell wall proteins

Proteins designated to be attached to the cell wall should be equipped with specific domains on their C-terminal. A large set of non-covalently bound proteins were predicted for *C. maltaromaticum* and *C. divergens* V41 whereas a small number of proteins were predicted for the other *Carnobacterium*. These proteins contain at least one LysM domain or one WxL domain.

All *Carnobacterium* strains contain between four and six LysM proteins. Two of them are conserved in all strains, three are found only in *C. maltaromaticum* and *C. divergens* V41, and three are variable within *Carnobacterium* sp. strains (Annexe 2). LysM domain proteins are found in several copies in many Gram-positive bacteria such as *E. faecalis* (Eckert et al., 2006), *L. monocytogenes* (Bierne and Cossart, 2007), *Lactobacillus* sp. (Kleerebezem et al., 2010), *L. lactis* (Steen et al., 2005), *Staphylococcus* strains (Hirschhausen et al., 2012) and *E. coli* (Bateman and Bycroft, 2000).

WxL domain anchors protein non-covalently to the cell wall (Brinster et al., 2007). It was discovered rather lately; it is a domain of 160 to 190 amino acids and it contains two conserved sequence motifs with the Trp-x-Leu signature. Strikingly, the number of WxL containing proteins is comprised between 42 and 53 proteins distributed only in *C. maltaromaticum* and *C. divergens* V41 strains. No WxL motif proteins were found in the other genome of *Carnobacterium*. Few are specific for each genome of *C. maltaromaticum* strain (between one and seven genes) and 15 are specific for *C. divergens* V41. When 56 genes are variable within the strains, only three are conserved within the *C. maltaromaticum* (Annexe 3). Little is known about WxL containing proteins, however they are commonly present in several Gram-positive bacteria such as *Lactobacillus* sp. (Boekhorst et al., 2006b; Chaillou et al., 2005; Siezen et al., 2006), *E. faecalis, L. lactis* (Brinster et al., 2007) and *L. monocytogenes* (Bierne and Cossart, 2007). It seems that *C. maltaromaticum* and *C. divergens* V41 genomes contain a high number of WxL proteins (Annexe 3).

Some proteins are covalently anchored to the CW, and possess an LPxTG like motif at their C-terminal. They are called Sortase Dependent Proteins (SDP) due to the fact that they are attached to the peptidoglycan layer of the cell walls by the sortase family of transpeptidases (Schneewind and Missiakas, 2014). Those later are found in *C. maltaromaticum* and *C. divergens* V41; only one LPxTG domain protein is found in *Carnobacterium* sp. 17.4 and *Carnobacterium* sp. AT7. However, *C. inhibens* subsp. *gilichinskyi* WN1359 possesses no gene encoding a SDP (Annexes 2 and 3).

Sortase enzymes decorate the surfaces of Gram-positive bacteria with a diverse array of proteins that enable each microbe to interact with its environment. *C. maltaromaticum* strains and *C. divergens* V41 possess many putative sortase A and B genes. *C. inhibens* subsp. *gilichinskyi* WN1359 holds only one putative sortase encoding gene. *Carnobacterium* sp. AT7 and *Carnobacterium* sp 17.4 possess a putative sortase gene encoding a protein devoided of the SP (Annexe 2 and 3). When proteins are exported through the Sec pathway, sortase enzymes handle the C-terminal LPxTG-like domain (Kline et al., 2009). Sortase-protein are then bond by trans-peptidase reactions. The difference between the several sortases is the type of CWSS motifs and amino groups that will recognize and define the transpeptidase reactions (Comfort and Clubb, 2004; Maresso and Schneewind, 2008).

The genomes of *C. maltaromaticum* are predicted to encode between 22 and 35 SDP, and the genome of *C. divergens* V41 would encode 18 SDP. In *C. maltaromaticum*, 13 SDP proteins are conserved in all strains, of which seven are also conserved with *C. divergens* (Annexe 2). In addition, in *C. maltaromaticum,* 24 SDP encoding genes are not conserved and can be found in at least one strain. The variability observed either resulted from the absence of homologues or the presence of genes fragments, likely pseudogenes. As for the three bacteria *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. AT7 and *Carnobacterium* sp 17.4, those latters possess no SDP on their CW (Annexe 3).

1.3.3.2.Surface protein functions

An analysis of the CWP amino acid revealed the presence for some proteins of functional domains (Figure 14).

Enzymes

Several surface proteins are predicted to contain domains involved in peptidoglycan hydrolysis (Annexe 3 and Figure 14). Peptidoglycan hydrolases are involved in fundamental aspect of bacterial biology: peptidoglycan assembly and disassembly, thereby in growth, turnover and recycling of the peptidoglycan, cell division, cell shape determination, biofilm formation, bacterial cell lysis as autolysis or as weapon in interbacterial species competition (Frirdich and Gaynor, 2013). These proteins, likely anchored thanks to LysM domains, are predicted to contain domain with peptidoglycan hydrolase function such as glucosaminidase and autolysin domains. Two other LysM proteins, found in *Carnobacterium* sp. AT7 and *Carnobacterium* sp. 17.4 would contain a domain either with peptidase activity or with polysaccharide deacetylase activity.

Among the conserved proteins in *C. maltaromaticum* and *C. divergens,* there are homologues of nucleotidases/metallophosphatases and PrtB (Annexe 3 and Figure 14). The nucleotidase/metallophosphatase proteins (BN424_190 and BN424_3529 in *C. maltaromaticum* LMA28 are predicted to be a multidomain protein. Extracellular 5'-nucleotidases domain catalyzes dephosphorylation of exogenous adenine 5'-nucleotides to adenosine and phosphate (Bengis-Garber and Kushner, 1982). It is described in aquatic ecosystems as a key enzyme for phosphorous regeneration (Ammerman and Azam, 1985). The metallophosphatase domain superfamily is functionally diverse. The model protein of this family is YhcR from *Bacillus subtilis*, which is a cell-wall endonuclease (Oussenko et al., 2004).

PrtB is a Cell Envelope associated Protease (CEP) of *Lactobacillus delbrueckii* subsp. *bulgaricus* allowing to degrade casein into peptides (Siezen, 1999). These peptides are subsequently internalized in order to be hydrolyzed by intracellular peptidases into small peptides and free amino acids (Savijoki et al., 2006). The presence of a PrtB homologue in the cell wall of *C. maltaromaticum* and *C. divergens* (Annexe 3 and Figure 14), raise the question of the presence of an equipment allowing to exploit extracellular oligopeptides. The analysis of the genomes of *C. maltaromaticum* and *C. divergens* revealed the conserved presence of homologues of the oligopeptides transporter systems OppABCDF, DtpT, and permitted to predict the presence of a whole arsenal of intracellular peptidases (Annexe 3). All these data strongly suggest that *C. maltaromaticum* and *C. divergens* are able to exploit amino acids from proteins of their environments. Interestingly, homologues of the Opp system and of intracellular peptidases were also found in *Carnobacterium* sp. 17-4, *Carnobacterium* sp. AT7 and *C. inhibens* subsp. *gilichinskyi* WN1359. However, no CEP were predicted from the genome of these latter three *Carnobacterium.* Therefore, *C. maltaromaticum* and *C. divergens*, compared to the other three *Carnobacterium,* are characterized by their predicted ability to produce an extracellular CEP. In the dairy environment, CEP are believed to confer a selective advantage by allowing bacteria to exploit the amino acids from milk casein. Similarly, the presence of PrtB homologue in *C. maltaromaticum* and *C. divergens* would be a selective advantage in the dairy environment.

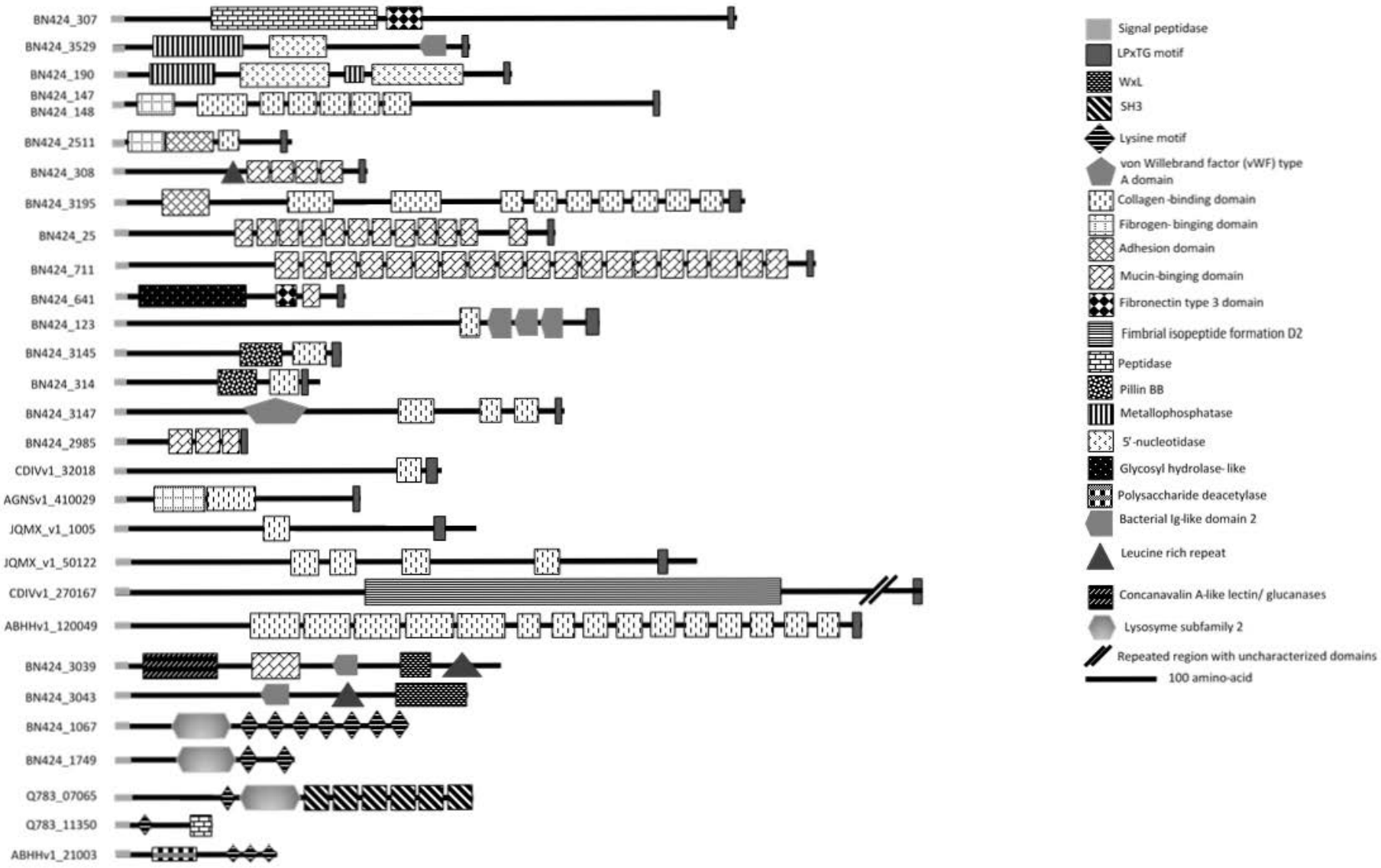


Figure 14: Putative surface proteins containing predicted functional domains.

Several strains of *C. maltaromaticum* produce a SDP protein with glycoside hydrolase activity (BN424_641 in *C. maltaromaticum* LMA28) (Annexe 3 and Figure 14). This multi-domain protein would also contain a Fibronectin type III-like module of unknown function, which is usually associated with glycoside hydrolase domains (Alahuhta et al., 2010). Such enzymatic activity could allow *C. maltaromaticum* to exploit extracellular carbohydrate polymers. Interestingly, this protein is also predicted to contain a mucin-binding domain. Mucin is the major component of the intestinal mucus. As a glycoprotein it is tempting to speculate that the putative mucin-binding glycoside hydrolase of *C. maltaromaticum* would hydrolyze mucin glycan moieties as was described for gut bacteria (Tailford et al., 2015).

Nutrient uptake: heme compounds

Two SDP, predicted as heme-binding proteins in *C. maltaromaticum* and *C. divergens*, are homologues of IsdA and IsdC (Annexe 3). In *S. aureus*, these two proteins are part of a heme uptake system composed of IsdABCDEFG where IsdA, IsdB, and IsdC are cell wall anchored heme-binding proteins, IsdD, IsdE, and IsdF form a membrane associated heme transporter, and IsdG is a heme degrading monooxygenase allowing to release iron in the cytoplasm (Haley et al., 2011; Schneewind and Missiakas, 2014). In addition to IsdA and IsdC homologues, homologues of IsdE, IsdF, and IsdG were found in the genome of *C. maltaromaticum* and *C. divergens*. Similarly to *S. aureus,* a gene putatively encoding a sortase B is predicted in the vicinity of the *isd* homologues in *C. maltaromaticum*. In *S. aureus*, the sortase B anchors IsdC to the cell wall, while IsdA and IsdB are anchored by sortase A (Balderas et al., 2012; Tiedemann and Stillman, 2012). Although we have found an IsdA homologue in *C. maltaromaticum* it exhibits several differences. In *C. maltaromaticum*, IsdA is longer than IsdA from *S. aureus* (713aa / 350 aa). The IsdA homologue in *C. maltaromaticum* is indeed predicted to contain 4 NEAT domains while IsdA from *S. aureus* contains 2. Neat domains bind heme compounds or proteins containing heme compound (Balderas et al., 2012; Gaudin et al., 2011). In addition, IsdA from *C. maltaromaticum* is not predicted to be anchored by sortase A since no LPxTG domain were found. This suggests that the IsdA homologue of *C. maltaromaticum* either could be secreted in the medium or could be anchored to the cell wall thanks to another mechanism. Although several Isd homologues where found in *C. maltaromaticum,* no homologue was found for IsdD and IsdB. Several differences

were found between the heme-uptake system from *S. aureus* and *C. maltaromaticum*, suggesting that if *C. maltaromaticum* have a functional heme-uptake system, it is working differently from that of *S. aureus*.

In addition, it can be noticed that *C. maltaromaticum* is described to exhibit a better growth when the culture medium is supplemented with hematin. This suggests that heme is used by *C. maltaromaticum* to respire oxygen. Consistently, *C. maltaromaticum* (formerly *Lactobacillus maltaromicus*) is reported to produce cytochrome b and d types when grown aerobically in the presence of hematin (Meisel et al., 1994). The genome of *C. maltaromaticum* contain homologues suggesting the ability to produce a functional respiratory chain. Indeed, we found homologues of the electron-donors encoding genes *noxB* and *ndh* ; homologues of the genes *yhdB, menE, menB, menD, menF, ispA, ispB* and *menA*, which are required for the synthesis of the electron shuttle menaquinone. Homologues of *cydABCD*, which are required for the synthesis of cytochrome quinol oxidase, a heme-requiring protein that perform the final electron transfer to the acceptor oxygen, are also found (Lechardeur et al., 2011). Comparative genomic analysis revealed that homologues of these genes in the genomes of *C. maltaromaticum* and *C. divergens*, except strains *C. maltaromaticum* ML.1.97 for which the homologue of *menF* appears as a pseudogenes. All these genomic data suggest that components of the cell wall proteome might be involved in extracellular heme utilization for oxygen respiration in *C. maltaromaticum* and *C. divergens*.

Microbial adhesion

The previous analysis of the genome of *C. maltaromaticum* LMA28 revealed the presence of putative cell surface adhesins (Rahman et al., 2014b). Analysis of the *Carnobacterium* genomes revealed that the strains of *C. maltaromaticum* and *C. divergens* produce on their surface several proteins predicted to contain domains reminiscent of adhesion function: collagen-binding surface protein Cna-like, mucBP, Leucin-Rich Repeat (LRR).

Ten LPxTG proteins (Annexe 3) are predicted to have a collagen binding domain (Figure 14). Functional domains are either annotated Collagen binding protein, Cna or bacterial adhesin. Collagen binding proteins are found in *S. aureus* (Elasri et al., 2002) and *L. monocytogenes* (Bierne and Cossart, 2007), and are suggested to participate in their infectious process. However, the probiotic character of *L. plantarum*WCFS1 was examined by the miming of at least 12 proteins

that are putatively involved in adherence to host components such as collagen and mucin (Boekhorst et al., 2006b).

Three LPxTG proteins (Annexe 3 and Figure 14) contain a MucBP (mucin binding protein) domain. MucBP domains allow adhesion to mucus material (Lukić et al., 2012). *Lactobacillales* and *Listeria* bacterial species can have between 1 and 14 MucBP-containing proteins (Bierne and Cossart, 2007; Boekhorst et al., 2006a).

Two LRR domain were found in some strains of *C. maltaromaticum* holding a WxL anchorage domain (Annexe 3 and Figure 14). In *L. monocytogenes*¸ LRR containing proteins are differently anchored to the cell wall, they are indeed SDPs. These domains are involved in protein-protein interactions. They are described to be associated to domains exhibiting Ig-like (Immunoglobulin) fold, of which function is thought to facilitate the presentation of the adjacent LRR domain (Bierne and Cossart, 2007). Accordingly, next to the LRR domain of the putative *C. maltaromaticum* LMA28 surface proteins BN424_3043, an Ig-like domain was found at the C-terminal end of the LRR region (Annexe 3 and Figure 14). In *L. monocytogenes* strains, internalins are LRR containing proteins and are associated to virulence. LRR containing proteins are rather uncommon in *Lactobacillus*.

The presence of such putative adhesins suggests that strains of these two species exhibit adhesive ability to intestinal mucosa and extracellular matrices of animals. Overall, 19 putative adhesins were predicted from the genomes of *C. maltaromaticum* and *C. divergens*. All these proteins are SDP except two LRR-containing proteins exhibiting a WxL binding domain. Of these 19 proteins, two are conserved in all *C. maltaromaticum* strains: one putative collagen-binding SDP, and one mucin-binding protein.

Genes encoding the pili proteins previously described for *C. maltaromaticum* LMA28 (Rahman et al., 2014b), were found only in *C. maltaromaticum* DSM20342 MX5. Both strains belong to clonal complex CC1, which is suspected to be a lineage well-adapted to dairy environment. Pili were described as surface component promoting adhesion to dairy matrix in *L. rhamnosus* GG and thereby might contribute to bacterial fitness in dairy (Burgain et al., 2014b). In general, two or three subunit genes encoding a Gram-positive pilus are organized into an operon, along with at least one sortase gene (Mandlik et al., 2008a; Proft and Baker, 2009). The closest homologues were found in *E. faecalis*, for which the genetic organization of the pili encoding

genes is highly similar to *C. maltaromaticum*. In *E. faecalis* V583, the major pilin PilB is homologous to BN424_3145 (Figure 15).

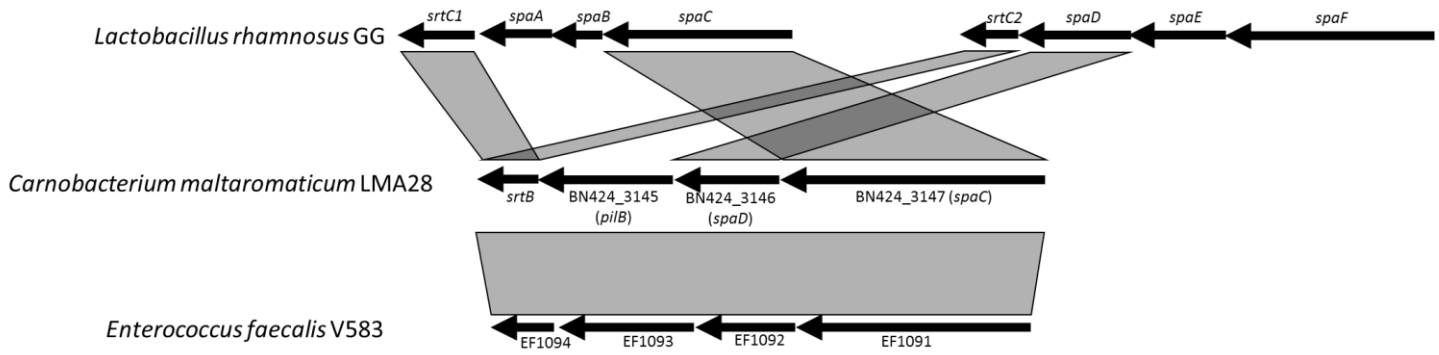


Figure 15: Gene cluster organization and synteny between *C. maltaromaticum* LMA28, *L. rhamnosus* GG and *E. faecalis* V583.

Homologues for BN424_3146 and BN424_3147 can be also found in *L. rhamnosus* GG, however they were found in different pili loci; *L. rhamnosus* GG indeed exhibits two pili encoding loci (Kankainen et al., 2009). More precisely, BN424_3146 is homologous to *spaD*, the major pilin encoded by one pili locus in *L. rhamnosus* GG, and BN424_3147 is homologous to adhesin encoding gene *spaC*, which is found in the other pili locus in *L. rhamnosus* GG and encodes a minor pilin. This suggests that the genetic structures of pili loci in *C. maltaromaticum, E. faecalis* V583, and *L. rhamnosus* GG are the result of gene rearrangements that occurred during evolution. In addition, these results suggest that the pili locus of *C. maltaromaticum* and *E. faecalis* V583 encodes two major pilins.

Overall, these comparative genomic analyses suggest that all strains from *C. maltaromaticum* and *C. divergens* would have adhesive properties and that strains might exhibit differences in this regard.

1.4. Conclusion

Comparative genomic analyses revealed striking differences between strains of *C. maltaromaticum* and *C. divergens*, and strains that belong to other species: *Carnobacterium* sp. AT7, *Carnobacterium* sp.17-4, and *C. inhibens* ssp. *gilichinsky* WN1359. The cell-surface of *C. maltaromaticum* and *C. divergens* V41 would be decorated with a large array of surface proteins that would contribute to their ability to adapt to food products and the intestinal tract. These surface components might also contribute to the virulence of these bacteria; they are indeed described as

fish pathogens (Leisner et al., 2007). Thus, regarding their surface, *C. maltaromaticum* and *C. divergens* are more similar to other LAB than to other *Carnobacterium* species, which might reflect common ecological properties linked to food and the digestive tract. As a LAB, *C. maltaromaticum* exhibit the largest cell-surface proteome that could explain the ability of this species to colonize multiple environments.

2. Adaptation to dairy environment: lactose metabolism pathways in LAB and *Carnobacterium*

2.1. Genes associated to lactose metabolism illustrate the high diversity of *Carnobacterium maltaromaticum*

Abstract

Carnobacterium maltaromaticum is a LAB of technological interest as bioprotective and ripening agent in cheese. The dairy population of *C. maltaromaticum* is characterized by a high diversity suggesting a high diversity of the genetic traits linked to the dairy process. As lactose is the main carbon source in milk, the genetics of lactose metabolism was investigated in this LAB. Comparative genomic analysis revealed that the species *C. maltaromaticum* exhibits genes related to the Leloir and the Tagatose-6P pathways. More precisely, strains can bear genes related to one or both pathways and several strains apparently do not contain homologs related to these pathways. Analysis at the population scale revealed that the Tagatose-6P and the Leloir encoding genes are disseminated in multiple phylogenetic lineages of *C. maltaromaticum*: genes of the Tagatose-6P pathway are present in the lineages I, II and III, and genes of the Leloir pathway are present in the lineages I, III and IV. These data suggest that these genes evolved thanks to horizontal transfer, genetic duplication and translocation. We hypothesize that the *lac* and *gal* genes evolved in *C. maltaromaticum* according to a complex scenario that mirrors the high population diversity.

Keywords: *Carnobacterium*, Lactose, Galactose, Comparative Genomic Analysis, Diversity

2.1.1. Introduction

Lactic Acid Bacteria (LAB) are naturally found in a variety of environmental habitats, including plant (fruits, vegeTables, and cereals), meat and milk environments. They are involved in a large number of industrial food fermentations where they are used as processing aids (Klaenhammer et al., 2005). Whereas the selection of LAB starters and adjunct cultures allows to standardize industrial processes, it also induces a striking reduction of microbial biodiversity in the food product. This leads to the need for the design of new generations of processing aids allowing to reconcile industrial requirements and microbial diversity.

Carnobacterium maltaromaticum is a promising candidate for culture innovation. This LAB could be used as ripening flora since it is able to produce malty flavors in cheeses (Afzal et al., 2013a), and as a bioprotective culture by inhibiting *Listeria monocytogenes* (Leroi et al., 1996; Mathieu et al., 1994; Wan et al., 1997). It can reach high concentrations in different cheeses manufactured with ewe, cow or goat raw milk (Cailliez-Grimal et al., 2007), and does not interfere with the acidification process involving the starters *Lactococcus lactis* or *Streptococcus thermophilus* (Edima et al., 2008). It is an atypical LAB displaying valuable properties such as growth at low temperatures (psychrotrophy) and at alkaline pH values (up to 9.6), which allow this bacterium to grow mainly during ripening. Unlike well-established starter cultures such as *L. lactis* or *S. thermophilus* (Delorme et al., 2010; Passerini et al., 2010), *C. maltaromaticum* dairy strains exhibit a high diversity. Indeed, the population structure of the species, determined by MultiLocus Sequence Typing (MLST), revealed that dairy strains are spread into four deeply branched lineages (I, II, III and IV). Consistently, the genetic diversity of the dairy population is as high as the diversity of the non-dairy population (Rahman et al., 2014a). The high genetic diversity of *C. maltaromaticum* suggests that this species is characterized by a high diversity of properties of technological interest, such as the ability to metabolize lactose.

Lactose is the main carbon source found in milk. The ability of starter cultures to metabolize lactose and the resulting galactose, which is massively released in the curd when LAB such as *S. thermophilus* are used for acidification, is therefore a major aspect of starter selection. Lactose is a disaccharide consisting of the combination of two simple sugars, α/β-D-glucose and β-D-galactose, linked by a β (1-4) glycosidic bond. Lactose and its galactose moiety can be metabolized *via* two main pathways: the Tagatose-6P (*lac* genes) and the Leloir (*gal* genes)

pathways (Kandler, 1983) (Figure 8). In the Tagatose-6P pathway, lactose/galactose enter the cell via $EIIABC^{Lac}/EIIABC^{Gal}$, a transporter depending on the phosphoenol pyruvate phosphotransferase system (PTS) encoded by the genes *lacEF* (Abranches et al., 2004; Zeng et al., 2012, 2010). Galactose is phosphorylated into galactose-6P by the EII PTS component. As well, lactose is phosphorylated through the PTS into lactose-6P, which is then hydrolyzed by a phospho-β-galactosidase (encoded by *lacG*) into glucose and galactose-6P. This latter is then transformed into 2 trioses *via* a series of 3 reactions encoded by *lacAB* (galactose-6-phosphate isomerase), *lacC* (tagatose-6-phosphate kinase) and *lacD* (tagatose-1,6-diphosphate aldolase). For the Leloir pathway, the galactose can enter the cell *via* a permease encoded by *galP,* and the lactose can enter the cell *via* a permease encoded by *galP* or *lacS* (Fortina et al., 2003; Leong-Morgenthaler et al., 1991; Neves et al., 2010). Lactose is then hydrolyzed *via* a β-galactosidase (*lacZ* or *lacLM*) into glucose and galactose. Galactose is subsequently converted into glucose-6-phosphate within 4 reactions. First, it is isomerized from β-galactose to α-galactose *via* a galactose mutarotase (*galM*). Then, a galactose kinase encoded by *galK* phosphorylates α-galactose. The resulting galactose-1-phosphate is converted into glucose-1-phosphate *via* a galactose-1-phosphate uridyltransferase (*galT*) / UDP-glucose-4-epimerase (*galE*) system. Glucose-1-phosphate is finally converted into glucose-6-phosphate *via* a phosphoglucomutase (Maxwell et al., 1962; Poolman et al., 1989; van Rooijen et al., 1991). While lactose and galactose metabolism in typical cheese starters is well documented (Wu et al., 2015), very little is known in LAB such as *C. maltaromaticum.*

The aim of this study was to investigate at the genome level the metabolism of lactose and galactose in *C. maltaromaticum.* To do so, the genome sequence of two strains of *C. maltaromaticum,* and the sequence of three *C. maltaromaticum* LMA28 plasmids were determined. Genome data mining, including genome comparisons, were performed at the intraspecific and interspecific levels. In addition, the *lac* and *gal* genes were searched in a collection of 42 strains of *C. maltaromaticum* of diverse origins in order to characterize the genetic diversity of the genes associated to lactose and galactose metabolism in this LAB species.

2.1.2. MATERIALS AND METHODS

Cultivation of strains

Carnobacterium maltaromaticum strains used in this study are listed in Table 1. They were stored in 20% glycerol at -80°C and were isolated by surface plating on Trypticase Soja (TS) Agar (Biomérieux, France), inoculated in TS Broth and incubated at 30 °C for 24 h before use.

Sequencing and annotation

The genome sequence of *C. maltaromaticum* 3.18 and ML.1.97 were determined by using the 454 pyrosequencing GS-FLX system (Roche 454 Life Sciences, Mannheim, Germany) as described by (Leisner et al., 2012). The three plasmids of *C. maltaromaticum* LMA28 were sequenced as previously described (Cailliez-Grimal et al., 2013).

Genome analysis

The newly sequenced genomes were integrated in the MicroScope platform (Vallenet et al., 2013) to perform automatic and expert annotation of the genes, as well as comparative analysis. MicroScope thus hosts the genome of three *Carnobacterium* sp. and five *C. maltaromaticum* strains (Table 2). Synteny analyses were performed by using 30% of identity and 80% of coverage as thresholds at the amino acid level.

PCR analysis

DNA was extracted from *C. maltaromaticum* as previously described (Rahman et al., 2014a). PCR reactions were performed in a volume of 50 µL containing: 23.5 µl of Emerald Amp GT PCR master mix (Takara), 0.4 µM of each primer (Table 8), and 3 to 5 ng of the genomic DNA. The PCR conditions were 5 min at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec at the adapted hybridization temperature and 45 sec at 72°C and the final extension for 5 min at 72°C. The PCR products were subsequently analyzed by agarose gel electrophoresis.

Table 8: Primer sequences used in this study

Targeted gene	Primer sequence (5' to 3') forward	Primer sequence (5' to 3') reverse	Hybridization temperature (°C)	Expected product size (bp)
lacA2	GCAGATAATGCCGGATTTGA	AGCTGCGTAATCAGCATGG	55	366
lacB2	TGACCACATTGTCACGGATG	CGTGATAGAATCCTTCAGACCA	55	498
lacA1	TGCAGATTTAGCGGGTAAGG	ACGCGAATTTGATGTCTTCC	49	387
lacB1	TTGAAATCGCTGGGTCATGAAG	TGATATTCACCGTTATCCCATTTCA	49	446
lacC1	TGTATTGCGATCCTCCATGA	GTTTCTCTTGGGCATTCAGC	49	598
lacD	GCTGCTGATGCTCGTGATAA	GCAGCTTGCTCACCTTCTTT	49	659
lacE1	TATCTCAGCGATGCCTGTTG	GCATGCTGTCCAGCTTGTAA	49	723
lacF1	GCAGGTGAAGCGCGATCATA	TCATCAAGTGATCCTGTCCATGC	49	205
lacG	CCGACTGGTTTTGGTGAAGT	ATCCAGTCCCAATCAGTTCG	56	800
lacE2	TTCCATTTCCATGGGAGCTA	TGGCTCTTCCTCCTGCTAAA	55	497
lacF2	GGTTTTGAAATTGTCGCGTA	AACTTCTTCTCCGCCTGCTT	55	186
lacC2	AACTAGGAGCGCCCATTACA	CAATGCCATCGAATTTTTCA	55	504
galK	ACAATGAACGTCGCAGTGAA	TTGATCTTCAGCAACGATGG	49	389
galE	TCGTTTTAGGTGGAGCTGGT	GGTTGGTTGCAGCATTTTCT	49	399
galT	TTTTGCAGGTTCAAATGCTG	AAAATCCCATCGGGGTAGTC	49	402

Data availability

The annotations were deposited at DDBJ/EMBL/GenBank under the following references: PRJEB9002 for *C. maltaromaticum* ML.1.97, and PRJEB8756 for *C. maltaromaticum* 3.18.

The annotations are publicly available for consultation in MicroScope (http://www.genoscope.cns.fr/agc/microscope/home/)

2.1.3. Results

2.1.3.1. Analysis of *lac/gal* genes by comparative genome analysis

Tagatose-6P pathway

The strains *C. maltaromaticum* LMA28, DSM20342 MX5 and ATCC35586 exhibit orthologs of the Tagatose-6P pathway genes (Table 9), named $lacA_1B_1C_1D_1F_1E_1G_1$. In the public databases, the closest homologues are the genes *lacABCDFEG* encoding the Tagatose-6P pathway in *Staphylococcus aureus* and *Streptococcus mutans*, with which they share an overall 73% of aa identity. The similarity with the *lac* genes of *S. aureus* and *S. mutans* is unexpectedly high since housekeeping genes share on average 64% and 57% aa identity with *Streptococcus* and *Staphylococcus* species, respectively.

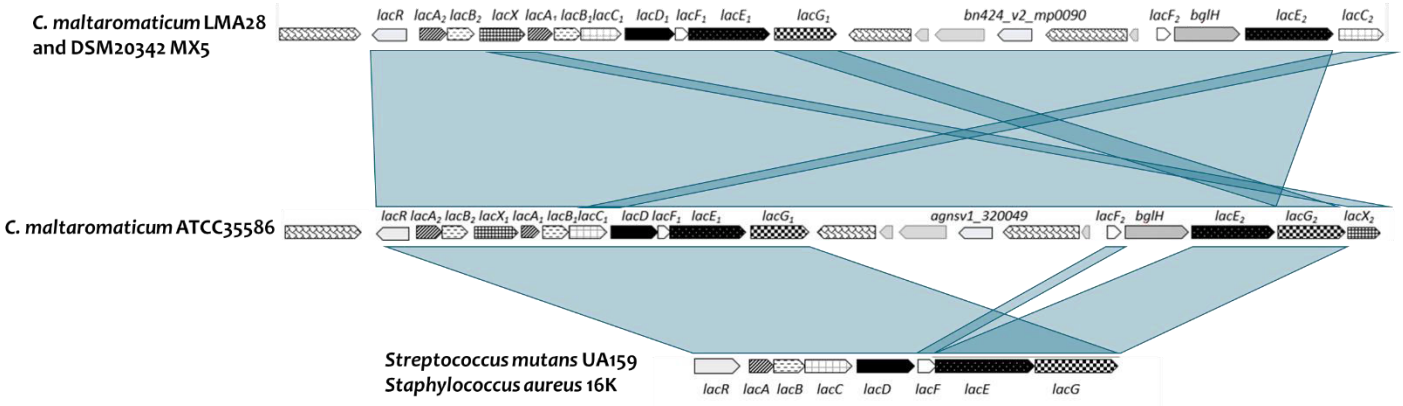


Figure 16: Organization and synteny between *lac* genes among *Carnobacterium maltaromaticum* LMA28, DSM20342-MX5, ATCC35586, *Streptococcus mutans,* and *Staphylococcus aureus*.
The arrows symbolize the coding sequences; homologous genes between strains are connected with grey bars. *lacAB*, galactose isomerase; *lacC*, tagatose-6-phosphate kinase; *lacD*, tagatose-1,6-diphosphate aldolase; *lacFE*, EIIABClac/Gal; *lacG*, phospho-β-galactosidase; *lacR*, transcriptional regulator; *lacX*, putative mutarotase; *bglH*, glucosidase, bn424_v2_mp0090, transcriptional regulator; agnsv1_320049, transcriptional regulator.
gene of unknown function **transposase.**

Table 9: Lactose/galactose metabolism putative genes in *Carnobacterium* sp.

Light grey indicates that one copy of the gene is present and the dark grey means that more than one copy are present. White boxes indicates that no homologue was found. *lacAB*, galactose isomerase; *lacC*, tagatose6-phosphate kinase; *lacD*, tagatose-1,6-diphosphate aldolase; *lacEF*, PTS system; *lacG*, phospho-β-galactosidase; *lacR*, regulation genes; *lacS*, lactose permease; *lacLM*, β-galactosidase; *lacZ*, β-galactosidase; *galP*, galactose permease; *galM*, galactose epimerase; *galK*, galactose kinase; *galT*, galactose-1-phosphate-uridyltransferase; *galE*, UDP-glucose-4-epimerase; Mpl, megaplasmid.

Species		*C. maltaromaticum*					*Carnobacterium* sp.		*C. inhibens* subsp. *gilichinskyi*
Strain code		LMA28 Mpl	DSM20342 MX5	ATCC 35586	ML.1.97	3.18	17.4	AT7	WN1359
Ecology		Soft ripened cheese	Milk with malty flavor	Diseased trout	Fresh salmon	Pork products	Cold seawater	Aleutian trench	Siberian permafrost
References		(Millière et al., 1994)	(Miller et al., 1974)	(Leisner et al., 2012)	(Leisner et al., 2007)	(Leisner et al., 2007)	(Voget et al., 2011)	(Lauro et al., 2007)	(Leonard et al., 2013)
Tagatose-6P pathway genes	*lacA*								
	lacB								
	lacC								
	lacD								
	lacE								
	lacF								
	lacG								
	lacR								
	lacT								
Leloir pathway genes	*galR*								
	lacZ/ lacLM								
	galP/ lacS/lacY								
	galM								
	galK								
	galT								
	galE								

These lac_1 genes are localized in a locus of 20 kbp highly conserved between the strains *C. maltaromaticum* LMA28, DSM20342 MX5 and ATCC35586 (Figure 16). This locus comprises other genes predicted to be involved in carbohydrate metabolism. Among them paralogs of the lac_1 genes with approximately 50% identity were found upstream and downstream of the lac_1 cluster (Figure 16, Annexe 4). More precisely, upstream and downstream of the lac_1 cluster of *C. maltaromaticum* LMA28 and DSM20342 MX5, the paralogs $lacA_2B_2$, and $lacF_2E_2C_2$ were identified, respectively. In the strain *C. maltaromaticum* ATCC35586, instead of an additional

copy of *lacC,* a *lacG* homologue was identified (called $lacG_2$, 86% aa identity with $lacG_1$) downstream of $lacE_2$. In addition, downstream of $lacG_2$, a potential mutarotase encoding gene ($lacX_2$) was predicted. In the vicinity of $lacF_2E_2$ paralogs, the three strains *C. maltaromaticum* LMA28, DSM20342 MX5 and ATCC35586, possess a putative *bglH* gene encoding a cytoplasmic hydrolase targeting carbohydrate polymer substrates other than lactose and galactose such as aryl-phospho-β-D-glucosides. BglH contributes to the intracellular release of glucose and glucose-6P (Setlow et al., 2004). The presence of multiple copies of the *lac* genes, and of genes involved in the metabolism of other carbohydrates (i.e. *bglH* and *lacX*), may suggest that the genetic evolution of this locus ended to carbohydrate metabolism diversification in *C. maltaromaticum* with the second *lac* cluster involved in PEP-dependent transport and catabolism of sugars different than lactose.

Genomic comparison revealed that those *lac* genes are localized in different position in the genome of the strain *C. maltaromaticum* ATCC35586 compared to the strains *C. maltaromaticum* LMA28 and DSM20342 MX5. In *C. maltaromaticum* LMA28, for which the whole genome sequence is available, this carbohydrate gene cluster is localized on a megaplasmid of 97.2 kbp. This megaplasmid contains other genes putatively involved in carbohydrate metabolism, mainly transporters (Annexe 5). In *C. maltaromaticum* ATCC35586, which does not contain the aforementioned megaplasmid, the *lac* gene cluster is surrounded by genes likely involved in other cellular functions than carbohydrate metabolism.

Even though *C. maltaromaticum* 3.18 contains genes homologous to *lacA, lacB, lacC, lacD, lacE, lacF,* and *lacG* (Table 9), these homologs are not co-localized in the genome. Thus, the *lacB* homologous gene is linked to genes putatively involved in ribose metabolism and the *lacCD* homologous genes are surrounded by genes predicted to be involved in N-acetylglucosamine degradation. It is indeed known that *lacCD* belong to a gene family involved in multiple carbohydrate pathways (Bissett and Anderson, 1980; Bork et al., 1993). Similarly, the genomes of *C. maltaromaticum* ML.1.97, *Carnobacterium* sp. 17.4, and *Carnobacterium* sp. AT7, contain one or several *lacCD* homologs but do not have the other genes known to be required for an active Tagatose-6P pathway (Table 9).

The Leloir pathway

Interestingly, in the genome of *C. maltaromaticum* 3.18, genes similar to *lacEF* are in the vicinity of genes expected to encode the Leloir pathway, i.e. *galMKTER* (Figure 18). In addition, three putative genes encoding hydrolases -*agaR*, encoding an α-galactosidase (family 36), *bgaB* and *bgaC* both encoding β-galactosidases (family 42, family 35, respectively)- are localized downstream of the *lacEF* paralogs (Figure 18). They are described to be responsible of the hydrolysis of carbohydrate polymers other than lactose, and containing galactose monomers (Coombs and Brenchley, 2001, 1999). However, lactose could be used by these strains thanks to a β-galactosidase of the *lacZ*/*lacLM* family (Table 9).

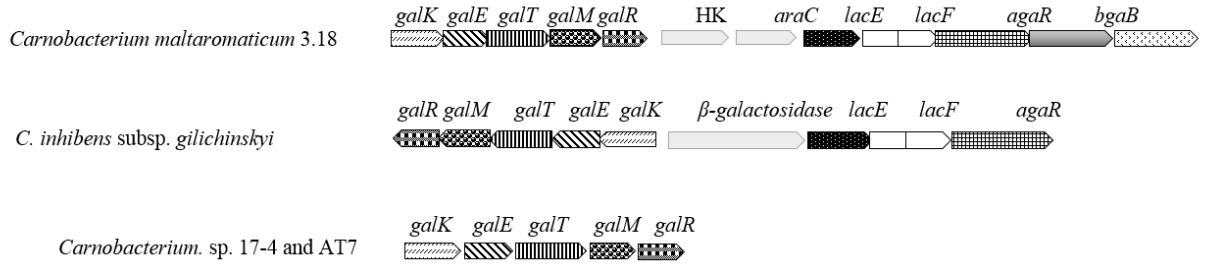


Figure 17: Organization and synteny between *gal* genes in *Carnobacterium maltaromaticum* 3.18, *C. inhibens* subsp. *gilichinskyi* WN1359, *Carnobacterium* sp. 17.4 and *Carnobacterium* sp. AT7.
Arrows represent the coding sequences. *galK*, galactose kinase; *galM*, galactose mutarotase; *galT*, galactose-1-phospohate uridyltransferase; *galE*, UDP-glucose-4-epimerase; *lacZ*, β-galactosidase; *lacLM*, β-galactosidase; *bgaB*, β-galactosidase; *bgaC*, β-galactosidase; *lacS*, lactose permease; *agaR*, alpha-galactosidase; *araC*, transcriptional regulator, HK putative histidine kinase.

Genes similar to those encoding the Leloir pathway, with an organization similar to that of the *galMKTER* cluster, were also found in the genome of *Carnobacterium* sp. 17.4, *Carnobacterium* sp. AT7 and *Carnobacterium inhibens* subsp. *gilichinskyi* WN1359 suggesting that this pathway is not specifically found in the species *C. maltaromaticum* (Table 9, Figure 17).

In *Carnobacterium* sp., only the strain *Carnobacterium* sp. 17.4 possesses a *galP* paralog (Table 9), with 20 to 25% aa id. with those of other LAB species (data not shown). This suggests that galactose or galactose containing carbohydrates polymers would be internalized by other transporters. Indeed, the genomes of *Carnobacterium* are predicted to encode between 35 and 80 sugar transport systems such as PTS cellobiose, mannose, glucose and other sugar types, sugar ABC transporters (Annexe 6).

From the Leloir pathway's encoding genes, the strains *C. maltaromaticum* ATCC35586, and *C. maltaromaticum* ML.1.97 seem to contain only one homologue of *galE* suggesting that

these strains does not have this pathway. This *galE* homologue could be involved in other general function reactions such as the cell wall biosynthesis (Grossiord et al., 2003).

2.1.3.2. Presence of *lac* and *gal* genes in *C. maltaromaticum* of diverse origins

The genes lac_1 *(*$lacA_1B_1C_1D_1F_1E_1G_1$*)* and lac_2 *(*$lacA_2B_2F_2E_2C_2$*)* were searched by PCR in 42 *C. maltaromaticum* strains isolated from different ecological origins (Table 10). The collection contains 29 dairy strains and 13 non-dairy strains.

Genes of the *lac1* cluster were detected in 10 *C. maltaromaticum* dairy strains (F43, CP23, F86, CP5, L1, F33, F7, F44, LMA28, and DSM20342; Table 10) and one non-dairy strain (ATCC35586, Table 10). Among those strains, six strains also contain $lacA_2B_2F_2E_2C_2$ (CP23, CP5, F33, F44, LMA28, DSM20342^T), and five strains lack $lacA_2$, $lacF_2$, $lacE_2$ and/or $lacC_2$ genes (F43, F86, L1, F7, ATCC355896). In addition, two dairy strains (DSM20344 and CIP102035, Table 10), were found to contain only the *lac2* genes but no *lac1* genes (Table 7). For 17 strains, only $lacE_2$ and/or $lacC_2$ were detected, suggesting they do not possess the entire Tagatose-6P pathway. The *gal* cluster was detected in 13 *C. maltaromaticum* strains: seven from dairy products (F43, F42, F4, F84, F87, CP7, and L11; Table 10), six from non-dairy environments (9.4, 5.27, N15, CIP100481, CIP101254 and 3BO4; Table 10). One out of the seven dairy strains (F43) also contains the *lac1* genes. Overall, these results suggest that 25 strains possess genes encoding the Tagatose-6P pathway and/or the Leloir pathway, and that 17 strains (10 dairy strains and 7 nondairy strains) do not possess these pathways (Table 10).

The population structure of *C. maltaromaticum* is divided in four main lineages (Rahman et al., 2014a). The *lac* genes are present in strains that belong to lineages I, II, and III. More precisely, *lac* genes were detected in strains belonging to the clonal complex CC1, except for the strain CP32 (ST17/CC1) and the non-dairy strain F29-1 (ST4/CC1). The *lac* genes were also found in the dairy strains CIP102035 (ST21/CC4) and DSM20344 (ST12/CC5), and the only non-dairy strain ATCC 35586 (ST36/CC4) (Table 10).

The *gal* genes were detected in strains that belong to lineages I, III, and IV. More precisely, they were detected in dairy and non-dairy strains belonging to CC2 and CC3, except for the dairy strain CP14 (ST3/CC2) and the non-dairy strain 710 (ST35/CC3).

Table 10: Presence of *lac* and *gal* genes within a collection of 42 *Carnobacterium maltaromaticum* strains.

					Tagatose-6P pathway genes												Leloir pathway genes		
Strain code	**Lineage**	CC	ST	Origin	*lacA2*	*lacB2*	*lacA1*	*lacB1*	*lacC1*	*lacD1*	*lacF1*	*lacE1*	*lacG*	*lacF2*	*lacE2*	*lacC2*	*galK*	*galT*	*galE*
710		3	35	Pork	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9.4		3	30	Beef	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
CIP100481		3	18	Human blood	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
F43		3	8	St Marcellin	-	+	+	+	+	+	+	+	+	-	-	-	+	+	+
CIP101342	**I**		19	Human blood	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
DSM20344		5	12	Raw milk	+	+	-	-	-	-	-	-	-	+	+	+	-	-	-
F88		5	12	Pérail	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
F2		5	2	Camembert	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
F14			5	Morbier	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
CP32		1	17	Brie	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CP23		1	16	Brie	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
F86		1	11	Cabrales	+	+	+	+	+	+	+	+	+	+	-	+	-	-	-
CP5		1	15	Brie	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
L1		1	13	Raw milk	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-
F33	**II**	1	6	St Maure de Touraine	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
F7		1	4	Halloum	+	+	+	+	+	+	+	+	+	-	+	+	-	-	-
F44		1	4	Mont D'or	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
LMA28		1	4	Brie	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
DSM20342^{T}		1	4	Raw milk	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-
F29-1		1	4	Lumpfish roe	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CIP101354			20	Deer	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
F42		2	7	Rocamadour	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
N15		2	7	Halibut-sugar salted	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
F4		2	3	Le Montagnard	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
F84		2	3	Kernhemmer	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+
F87		2	3	Tynjethaler	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+
CP7	**III**	2	3	Brie	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+
CP14		2	3	Brie	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DSM20624			23	Vaccum packaged minced meat	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
CIP102035		4	21	Raw milk	+	+	-	-	-	-	-	-	-	+	+	+	-	-	-
ATCC35586		4	36	Diseased rainbow trout	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
F1			1	Camembert	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
C4			1	Soil	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G117			34	Raw milk	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
G97			33	Raw milk	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
6.2			29	Beef	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
L11			14	Raw milk	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
5.27	**IV**		28	Beef	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
DSM20590			22	Vaccum packaged beef	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
3BO4			31	Sphagnum	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
F73		6	10	Selles-sur-Cher	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
F48		6	9	Mascarpone gorgonzola	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

2.1.4. Discussion

The detailed genomic investigation of the lactose and galactose utilization in *C. maltaromaticum* revealed that strains could exhibit genes related to the Leloir pathway, while others contain genes related to the Tagatose-6P pathway, and some strains apparently do not harbor any. Two strains (*C. maltaromaticum* 3.18 and F43) possess both *lac* and *gal* related genes. A similar genetic variability was also reported for other dairy LAB species (deVos, 1996; de Vos and Vaughan, 1994; Grossiord et al., 1998). *Lactococcus lactis* is able to metabolize lactose via the *lac* operon (de Vos and Vaughan, 1994) and galactose thanks to *galPMKTER* (Thompson, 1980). *Lactobacillus helveticus* metabolizes lactose *via* the Leloir pathway (Fortina et al., 2003) whereas *L. casei* exhibits genes encoding both the Tagatose-6P and the Leloir pathways (Chassy et al., 1978; Chassy and Alpert, 1989; Gosalbes et al., 1997). Among streptococci, the dairy starter species *S. thermophilus* (Anbukkarasi et al., 2014) as well as *S. salivarius* (Chen et al., 2002) hold the Leloir pathway genes, whereas *S. gordonii* (Zeng et al., 2012, 2010) and *S. mutans* (Abranches et al., 2004) possess the genes encoding the two pathways. Thus, regarding the presence of the Tagatose-6P and Leloir pathways, *C. maltaromaticum* seems to be a typical LAB.

In *S. mutans*, the *lac* cluster allows to metabolize lactose and galactose (Zeng et al., 2010). Similarly, the lac_1 genes of the three *C. maltaromaticum* strains could be involved in lactose and galactose metabolism.

The strain *C. maltaromaticum* 3.18 contains *lacA, lacB, lacC, lacD, lacE, lacF,* and *lacG* genes. However, in this strain, they are not organized in a cluster and synteny differs with that of other strains. This suggests that these genes are not involved in the same cellular process and that this strain does not have a functional Tagatose-6P pathway. The strains *C. maltaromaticum* ML.1.97, *Carnobacterium* sp. 17.4, and *Carnobacterium* sp. AT7, contain one or several *lacCD* homologous genes but would miss the other genes of the Tagatose-6P pathway, suggesting that a functional Tagatose-6P pathway is missing.

The *lac* genes are localized in a megaplasmid in *C. maltaromaticum* LMA28. The presence of lactose metabolism genes on plasmids is also described in the other LAB *L. lactis* (de Vos and Vaughan, 1994), *Lactococcus garvieae* IPLA 31405 (Flórez and Mayo, 2015), or *Streptococcus macedonicus* ACA-DC 198 (Papadimitriou et al., 2015).

The highest identity with the *Carnobacterium* lac_1 genes was found in *Streptococcus* and *Staphylococcus* which suggest that these genes evolved by horizontal transfer as it was previously suggested for other LAB (Cavanagh et al., 2015; Flórez and Mayo, 2015; Papadimitriou et al., 2015). Our results suggest that, in addition to horizontal transfer, the *lac* genes were subjected to translocation genetic event in the ancestor of *C. maltaromaticum* LMA28, DSM20342 MX5, and ATCC35586.

Multiple *lac* genes are present in two copies in *C. maltaromaticum*. Thus, in addition to the lac_1 locus whose organization is similar to that of other LAB and *Staphylococcus,* $lacA_2B_2F_2E_2C_2$ paralogs were found in their vicinity in strains *C. maltaromaticum* LMA28 and *C. maltaromaticum* DSM20342 MX5. The presence of such paralogs for the *lac* genes suggests that a duplication genetic event took place in the ancestor to *C. maltaromaticum*. Following duplication, the copies could have had different fates. One copy could have kept the same function, related to lactose transport and catabolism, and the other could either have retained the same function, could have been inactivated by mutation, or could have gained another function (Zhang, 2003). The analysis of the vicinity of $lacF_2E_2$ revealed the presence of other genes predicted to be involved in carbohydrate metabolism. Thus, in the three strains *C. maltaromaticum* LMA28, DSM20342 MX5 and ATCC35586, several putative carbohydrate polymer hydrolases were predicted. This suggests that the lac_2 genes are not involved in the lactose and galactose metabolism and could be involved in the metabolism of complex carbohydrate polymers. In this scenario, the duplication of *lac* genes would have been conducted to carbohydrate metabolism diversification. This metabolic diversification could contribute to the particularly high adaptation ability of *C. maltaromaticum*, which is one of the two most frequently isolated species of the genus *Carnobacterium* (Leisner et al., 2007). In the strain *C. maltaromaticum* ATCC35586, a second copy of *lacG* was found instead of a second copy of *lacC*. This suggests that this diversification process would have undergone in the evolutionary history of the species *C. maltaromaticum* and could explain the observed carbohydrate intraspecific variability described in the literature (Laursen et al., 2005).

In the genome of the strain *C. maltaromaticum* 3.18, the genes *galMKTER* are in the vicinity of the three genes predicted to encode galactosidases described to be involved in the hydrolysis of carbohydrate polymers other than lactose, and containing galactose monomers (Coombs and Brenchley, 2001, 1999). This suggests that *C. maltaromaticum* 3.18 is able to

hydrolyze diverse carbohydrate polymers containing galactose and that the resulting galactose monomers could be subsequently metabolized through the Leloir pathway.

At the population scale of the species *C. maltaromaticum*, the results revealed that homologues related to the Tagatose-6P and the Leloir pathways are disseminated in multiple lineages: Tagatose-6P is present in the lineages I, II and III, and the Leloir pathway in the lineages I, III and IV. It can therefore be speculated that the evolution of the *lac* and *gal* genes involved multiple genetic events.

The population of *C. maltaromaticum* is characterized by the major clonal complex CC1 (Rahman et al., 2014a). Genotypes that belong to a clonal complex suggest that they evolve by clonal expansion (Feil et al., 2004). Almost all strains of this clonal complex were isolated from dairy products, suggesting that clonal expansion of lineage CC1 took place in the dairy environment. The presence of the genes related to the Tagatose-6P pathway in almost all strains of CC1 suggests that these genes significantly contribute to the fitness of this lineage in the dairy environment as suggested for other LAB (Cavanagh et al., 2015; Flórez and Mayo, 2015; Papadimitriou et al., 2015).

2.1.5. Conclusion

In conclusion, we hypothesize that the *lac* and *gal* genes evolved in *C. maltaromaticum* according to a complex scenario of multiple genetic events including horizontal transfer, duplication and translocation. It can be speculated that these events shaped the genome of this LAB and contributed to the ability of *C. maltaromaticum* to adapt to a wide range of environments and to the variability of properties of technological interest.

Acknowledgement

The authors would like to thank Charlène Thiebaut, Arnaud Khemisti, Myriam Michel and Sylvie Wolff for excellent technical assistance. They are also grateful to the Genoscope, the MicroScope team, and the National Infrastructure “France Génomique”.

2.2. Diversity of lactose/galactose metabolic pathways within Lactic Acid Bacteria

Abstract

LAB are naturally found in a variety of habitats including milk environments, and the ability of starter cultures to metabolize lactose is a major aspect of LAB selection. Comparative genomic analysis of genes related to the Leloir and Tagatose-6P pathways was performed in 237 LAB and other related Gram-positive genomes with low GC content available in the MicroScope platform. All the genes involved in the Leloir (*gal* genes) and Tagatose-6P (*lac* genes) pathways have been investigated and have highlighted the great diversity of *lac* and *gal* presence and profiles among the 237 different LAB genomes investigated. Leloir pathway, present in 63% of the strains analyzed, is more commonly found within the LAB than the Tagatose-6P pathway, present in 45% of the strains analyzed. The genetic diversity of the lactose metabolic pathways remains at the genus level, except for *Carnobacterium maltaromaticum,* for which diversity is also present at the species level. This diversity may be then explained by the presence of many genetic events that conducted to the transfer of genes from one organism to another, the presence of galactose in other products than milk and the metabolic impact of galactose.

Keywords: Comparative genomics, Leloir, Tagatose-6Phosphate, Lactose, LAB, *Carnobacterium maltaromaticum*

2.2.1. Introduction

The term LAB (Lactic Acid Bacteria) relates to the metabolic capabilities of microorganisms to ferment various carbohydrates predominantly into lactic acid (Klaenhammer et al., 2005; Liu, 2003). LAB form a heterogeneous group of microorganisms and historically, the four genera *Lactobacillus*, *Leuconostoc*, *Pediococcus* and *Streptococcus* form the core of the group, corresponding to the current genera *Carnobacterium, Enterococcus, Lactobacillus, Lactococcus, Leuconostoc, Oenococcus, Pediococcus, Streptococcus, Vagococcus* and *Weissella* (Axelsson, 2004; Pot, 2008; Stackebrandt and Teuber, 1988; Stiles and Holzapfel, 1997). LAB are naturally found in a variety of natural habitats, including plants, meat and milk environment, and are involved in a large number of industrial and spontaneous food fermentations, notably those based on milk. The function of LAB is particularly important in dairy products, where their primary contribution is mainly in the rapid production of acid, which led to the acidification of

milk products (Klaenhammer et al., 2005). Lactose is the main carbon source in milk and the ability of starter cultures to metabolize lactose and the resulting galactose is therefore a major aspect of LAB selection. Because of this industrial relevance, lactose/galactose utilization by LAB has been well studied and documented (Cocaign-Bousquet et al., 1996; Kandler, 1983; Leroy and De Vuyst, 2004; Neves et al., 2005; Wu et al., 2015).

Lactose/galactose metabolism occurs *via* two main pathways, Leloir (*gal* genes) and Tagatose-6P (*lac* genes) (Kandler, 1983) (Table 3).

In the Leloir pathway, the galactose can enter the cell *via* a permease encoded by *galP,* and the lactose *via* a permease encoded by *galP* or *lacS* (Fortina et al., 2003; Leong-Morgenthaler et al., 1991; Neves et al., 2010). Lactose is then hydrolyzed *via* a β-galactosidase (*lacZ* or *lacLM*) into glucose and β-galactose (Maxwell et al., 1962; Poolman et al., 1989; van Rooijen et al., 1991). In the Tagatose-6P pathway, a phoshoenolpyruvate (PEP): carbohydrate phosphotransferase system (PTS) encoded by the genes *lacFE* assures the transport and the phosphorylation of lactose and galactose at the expense of PEP (Abranches et al., 2004; Zeng et al., 2012, 2010). Lactose-6P is then hydrolyzed by a phospho-β-galactosidase (encoded by *lacG*) into glucose and galactose-6P. This latter is then transformed into 2 trioses *via* a series of 3 reactions encoded by *lacAB* (galactose-6-phosphate isomerase), *lacC* (tagatose-6-phosphate kinase) and *lacD* (tagatose-1,6-diphosphate aldolase).

Various combinations of the *gal* and *lac* genes had been revealed in lactose and galactose utilization by LAB (Breidt et al., 1987; Chassy and Thompson, 1983; Loughman and Caparon, 2007; Rosey and Stewart, 1992) and, although individually genes present a high synteny level, each bacterium assembles them differently (Grossiord et al., 1998). However, studies were performed during the two last decades and lactose and galactose metabolic pathways diversity and variability between the different LAB genus and species is largely unexplored today.

Genome sequencing and the use of microbial genomic tools creates a number of new possibilities for strain characterization with a high level of resolution as well as for understanding the physiological properties of bacteria (Kankainen et al., 2009). Actually, approximately 240 LAB genomes are available in the MicroScope platform (Vallenet et al., 2013) to perform automatic and expert annotation of the genes. The aim of this paper is to use genome comparaisons to actualize the analysis of *gal* and *lac* genes of the Leloir and Tagatose-6P pathways.

Genome comparisons were performed in 237 accessible genomes (complete or draft) of LAB and some Gram-positive bacteria of the low GC subdivision, like *Bacillus, Listeria* and *Staphylococcus* (Axelsson, 2004) and results were discussed in terms of genetic diversity at the genus and species level.

2.2.2. Material and methods

Genome analysis

The sequenced genomes were integrated in the MicroScope platform (www.genoscope.cns.fr) (Vallenet et al., 2013) to perform automatic and expert annotation of the genes, as well as comparative analysis. From the current database, the genome of 2 *Aerococcus*, 1 *Bacillus subtilis*, 8 *Carnobacterium*, 44 *Enterococcus*, 36 *Lactobacillus*, 7 *Lactococcus*, 7 *Leuconostoc*, 51 *Listeria*, 3 *Oenococcus*, 2 *Pediococcus*, 18 *Staphylococcus aureus*, 52 *Streptococcus* and 6 *Weissella* were explored (Table 11). Sequence similarity analyses were performed by using 30% of identity and 80% of coverage as thresholds at the amino acid level.

2.2.3. Results and discussion

2.2.3.1. In *silico* analysis of *lac/gal* genes

The analysis of presence and number of putative genes from Leloir (*gal* genes) or Tagatose-6P (*lac* genes) pathways indicated that, among the 237 strains, 4 different groups can be defined (Table 11): strains holding only the Leloir pathway genes (31%), strains holding only Tagatose-6P pathway genes (14%), strains holding both pathways genes (31%) and strains holding no pathway (24%).

A great genetic profile diversity exists inside the same group and 83 different profiles were identified among the 237 strains analyzed. Presence and/or duplication of some essential genes can led to additional functions, whereas other ones are not essentials.

Table 11 : Lactose/galactose metabolism genes present in LAB and non-LAB strains (MicroScope platform).

The colored cases represent the presence of the genes, light gray indicates that only one copy of the gene is present and the dark gray means that more than one copy are present. The numbers indicate the number of gene replication. *lacAB* (galactose isomerase), *lacC* (tagatose-6P kinase), *lacD* (tagatose-1,6P aldolase), *lacFE* (PTS system), *lacG* (P-β galactosidase), *lacR* (regulation genes), *lacS* (lactose permease), *lacLMZ* (β-galactosidase), *galP* (galactose permease), *galM* (galactose epimerase), *galK* (galactose kinase), *galT* (galactose-1P-uridyltransferase), *galE* (UDP-glucose-4-epimerase).

	Tagatose-6Phosphate pathway									Leloir pathway						
	lacR	***lacA***	***lacB***	*lacC*	*lacD*	***lacE***	***lacF***	*lacG*	*lacT*	*galP lacS lacY*	*lacZ lacLM*	galR	*galM*	***galK***	*galT*	*galE*
Bacillus subtilis 168	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Carnobacterium gilichinskyi WN1359	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Carnobacterium maltaromaticum 3.18	1	2	0	1	2	1	1	1	0	0	0	1	1	1	1	2
Carnobacterium sp. 17-4	0	0	0	2	1	0	0	0	0	1	0	1	1	1	1	2
Carnobacterium sp. AT7	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1
Enterococcus durans ATCC6056	0	0	0	1	1	0	0	0	0	1	0	1	1	1	1	1
Enterococcus faecium D344SRF, DO NZ_AAAK, TX0133a04, TX0133B, TX0082, E1071, E1636, U0317, E1679, DO NZ_ACIY, TX0133a01, TX0133A, TX0133C, E1162, E1039,	0	0	0	0	0	1	1	1	0	1	0	1	1	1	1	1
Enterococcus hirae ATCC9790	0	0	0	1	1	0	0	0	0	1	0	1	1	1	1	1
Enterococcus mundtii QU25	0	0	0	0	2	1	1	1	0	1	0	1	1	1	1	1
Lactobacillus acidophilus NCFM	0	0	0	0	0	0	0	0	0	1	2	1	1	1	1	1
Lactobacillus brevis ATCC27305	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	1
Lactobacillus brevis ATCC14869, KB290	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Lactobacillus brevis ATCC367	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Lactobacillus fermentum IFO3956	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Lactobacillus helveticus DPC4571	0	0	0	0	0	0	0	0	0	1	2	1	1	1	1	1
Lactobacillus johnsonii NCC533	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	2
Lactobacillus plantarum ATCC14917	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	1
Lactobacillus plantarum JDM1, ST-III	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	2
Lactobacillus plantarum WCFS	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2
Lactobacillus rhamnosus ATCC5462	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Lactobacillus sakei 23K	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	2
Lactobacillus salivarius ATCC11741, GJ-24, NIAS840	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	2
Lactobacillus salivarius SMXD51, UCC118	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	2
Lactobacillus sp. ASF360	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	3
Lactococcus garvieae ATCC49156	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Lactococcus lactis KF147	0	0	0	0	0	0	0	0	0	1	1	1	1	2	2	1
Lactococcus lactis MG1363, IL403	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Lactococcus piscium MKFS47	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	1
Leuconostoc citreum KM20	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Leuconostoc gasicomitatum LMG18811	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

Leuconostoc gelidum JB7	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Leuconostoc kimchii IMSNU11154	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Leuconostoc mesenteroides J18	0	0	0	0	0	0	0	0	0	3	1	1	1	2	1	1
Leuconostoc mesenteroides ATCC8293	0	0	0	0	0	0	0	0	0	3	1	1	1	3	2	2
Listeria grayi DSM20601	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Oenococcus oeni AWRIB429, ATCC BAA-1163, PSU-1	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Pediococcus pentosaceus ATCC25745	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Streptococcus salivarius CCHSS3	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Streptococcus salivarius SK126	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1
Streptococcus sanguinis SK353	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Streptococcus thermophilus CNRZ1066, LMG18311, LMD-9	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2
Weissella cibaria KACC11862	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Weissella confusa LBAE C39-2	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Weissella koreensis KCTC3621, KACC15510	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Weissella paramesenteroides ATCC33313	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1
Enterococcus avium ATCC14025	1	1	1	2	2	2	2	2	0	1	0	1	1	1	1	1
Enterococcus casseliflavus ATCC12755	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1
Enterococcus faecalis V583, TX0635, TX0411, HH22, DAPTO 512, DAPTO 516, R712, S613, TX0104, TX4248, TX1322, TX2134, TX0855, TX0860, TUSoD Ef11, TX0470, PC1.1, ATCC29200, TX0109, TX0102	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1
Enterococcus faecium E980, PC4.1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1
Enterococcus faecium TX1330	2	2	2	2	2	2	2	2	0	1	0	1	1	1	1	1
Enterococcus italicus DSM15952	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1
Lactobacillus curvatus CRL705	1	2	2	1	2	1	1	1	0	0	0	1	1	1	1	2
Lactobacillus johnsonii FI 9785	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1
lactobacillus murinus ASF361	1	1	1	0	2	1	1	1	1	1	1	1	1	1	1	1
Lactobacillus paracasei 8700:2	1	1	1	2	2	2	1	1	0	1	0	1	1	1	1	1
Lactobacillus rhamnosus CASL, HN001, R0011, LMS2_1, ATCC21052, LC705	1	1	1	1	1	2	2	2	2	0	0	1	1	1	1	1
Lactobacillus rhamnosus GG ATCC53103	1	1	1	1	2	1	1	1	1	0	0	1	1	1	1	1
Lactocbacillus casei ATCC334	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1
Lactocbacillus casei BL23	1	1	1	1	2	1	1	1	1	0	0	1	1	1	1	1
Lactococcus lactis SK11	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1
Pediococcus acidilactici DSM20284	1	1	1	0	1	0	0	0	0	1	1	1	1	1	1	1
Streptococcus agalactiae A909	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	2
Streptococcus gordonii CH1	1	2	2	2	3	2	2	1	1	0	0	0	0	1	1	1
Streptococcus infantarius ATCC-BA-102	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	2
Streptococcus mutans UA159	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	2
Streptococcus pneumoniae D39	1	1	1	1	1	2	2	2	1	0	0	1	1	1	2	3
Streptococcus pneumoniae R6, TIGR4	1	1	1	1	1	1	1	1	1	0	0	1	1	1	2	3
Streptococcus sanguinis ATCC29667, ATCC49296, 2908, SK36, SK335, SK678, SK1087, SK1057	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1
Streptococcus sanguinis SK1, SK1056, SK1058, SK1059, SK115, SK160, SK340, SK405, SK408, SK49, SK72, SK150, SK330	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	2
Streptococcus sanguinis VMC 66	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1

Streptococcus suis 05ZYH33, ST3, ST1, SS12, JS14, 98HAH33	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1
Aerococcus viridans ATCC11563	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Carnobacterium maltaromaticum ATCC35586	1	2	2	2	3	2	2	1	0	0	0	0	0	0	0	1
Carnobacterium maltaromaticum LMA28, DSM20342 MX5	1	2	2	2	1	2	2	2	0	0	0	0	0	0	0	0
Lactococcus raffinolactis 4877	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Staphylococcus aureus 16K, JKD6008, MRSA252, TW20, Z172, T0131, S0385, COL, N315, USA300-TSACH1516, CA-347, USA300-FPR3757, Mu50, MW2, 08BA02176, JH1, 11819-97, Newman NCTC8325	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Streptococcus agalactiae 2603 V/R	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1
Streptococcus agalactiae NEM 316	2	2	2	2	2	1	1	1	1	0	0	0	0	0	0	1
Streptococcus pyogenes Manfredo, SSI-1, SF370, MGAS9429, MGAS8232, MGAS6180, MGAS10750, MGAS10394	2	2	2	2	2	1	1	1	0	0	0	0	0	0	0	0
Aerococcus urinae ACS-120-V-Col10a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carnobacterium maltaromaticum ML.1.97	1	0	0	2	2	0	0	0	0	0	0	0	0	0	0	1
Lactobacillus delbruckii ATCC11842, ATCC BAA-365	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
Leuconostoc carnosum JB16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Listeria innocua FSL S4-378, ATCC33091, FSL S1-023, Clip11262	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Listeria ivanovii FSL F6-596	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Listeria marthii FSL S4-120	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Listeria monocytogenes FSL F2-515, FSL F2-208, FSL N1-017, FSL R2-503, FSL J1-194, HCC23, L99, M7, serotype 4b, 08-5923, 10403S, 08-5578, FSL R2-561, SLCC2372, HPB2262, Finland 1988, EGD-e, FSL J1-175, SLCC2479, SLCC5850, FSL J2-003, SLCC179, 07PF0776, FSL N3-165, F6900, J0161, J2818, J1816, serotype 1-2a, FSL J2-071, L312, J1-220, SLCC2540, SLCC540, SLCC2376, SLCC2378, SLCC2755, serotype 7, ScottA, ATCC19117, L028, FSL J2-064.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Listeria seeligeri 1/2b SLCC5334, FSL N1-067	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Listeria welshimeri 6b SLCC5334	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Streptococcus agalactiae FSL S3-026	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Weissella ceti NC36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Leloir pathway

Putative *gal* genes are presents in 72 strains holding the Leloir pathway alone and in 74 strains possessing Leloir and Tagatose-6P pathways (Table 11). Some strains holding only Leloir pathway possess also some *lac* genes conferring or not supplementary functions. In the genus *Carnobacterium*, *C. maltaromaticum* 3.18 strain possesses *gal* and *lac* genes (except *lacB*). *lac* genes are dispersed in the genome and are implicated in other clusters than lactose metabolism (data not shown).

Putative lactose/galactose permease *galP* gene is not always present in all genome possessing the Leloir pathway and never found in genome holding only Tagatose-6P pathway. Among genome holding no pathway, only *Lactobacillus delbruckii* strains possess a putative *galP* gene (Table 11). Few studies describe the permeases (Forrest et al., 2011; Yan, 2013), but the absence of *galP* gene does not necessarily imply a deficiency in lactose/galactose use. Indeed, in *Lactococcus lactis* NZ9000, *galP* mutating strain may internalize the galactose molecules *via* PTS system. Galactose-6P will then be dephosphorylated to join the Leloir pathway (Neves et al., 2010). *Streptococcus mutans* (holding Leloir and Tagatose-6P pathways) lacks a *galP* responsible for the galactose uptake (Zeng et al., 2010). It possesses 14 putative PTS with one lactose specific PTS (Ajdic and Ferretti, 1998). Galactose is then transported by other PTS specific for other sugar molecules but with lower affinity (Abranches et al., 2004; Zeng et al., 2010). In general, permeases are non-specific transport systems (Forrest et al., 2011; Yan, 2013) except for *lacS* (lactose permease) and *galP* (galactose permease) found in *Streptococcus thermophilus.* Indeed, *S. thermophilus* strains have only Leloir pathway. This species is generally found to have a partial fermentation of lactose, with the metabolization of glucose moiety, and excretion of galactose in the medium through *lacS*, responsible for the lactose/galactose exchange (Hutkins et al., 1985; Poolman, 1993).

The role of β-galactosidase encoded by *lacZ/lacLM* consists of the cleavage of lactose molecules into glucose and β-galactose. Putative *lacZ/lacLM* genes are not always found in all genomes possessing Leloir pathway suggesting the use of this pathway for metabolizing galactose (Table 11). In *Streptococcus salivarius* 57.I *lacZ/lacLM* mutation inducts the inability of the strain to use lactose and in those strains, only galactose may be metabolized. *lacZ* deficient strain could not grow on lactose, indicating that intracellular β-galactosidase is essential for lactose metabolism

(Chen et al., 2002). Among the 146 strains holding putative Leloir pathway genes alone or with the tagatose pathway, less than 40 strains exhibit synteny with homologue of *lacZ/LM* of *S. thermophilus*. In some strains, like *C. maltaromaticum* 3.18, *Carnobacterium* sp.17.4, *Carnobacterium* sp. AT7 and *E. casseliflavus* ATCC12755 putative genes encoding β-galactosidase (*bgaB* and *bgaC*) are in the vicinity of the Leloir pathway genes. The presence of these genes encoding enzymes responsible to the hydrolysis of carbohydrate polymers containing galactose monomers (combeys and Brenchley, 2001), suggests a possible degradation of diverse polysaccharides with subsequently metabolization of galactose.

Lactobacillus delbruckii ATCC 11482 and *L. delbruckii* BAA-365 strains holding a putative *galP* gene possess also *lacLM* putative genes (Table 11). This may lead to the hypothesis that these strains are capable of internalizing and hydrolyzing the lactose molecules, but only glucose can be metabolized and, due to the absence of complete Leloir and Tagose-6P pathway, the galactose is exported to the extracellular medium, like in *S. thermophilus* strains (Crow and Thomas, 1984; Kafsi et al., 2014; Leong-Morgenthaler et al., 1991).

Almost all the genomes studied, possessing the Leloir pathway alone or with Tagatose- 6P pathway, hold one putative *galM* gene (Table 11). *galM* incoding a galactose mutarotase is not an essential enzyme because its mutation inducts no differences in growth rate. It promotes the isomerization of β-galactose into α-galactose and the isomerization may be done spontaneously (Neves et al., 2010). It can be noted that *galM* is present in strains without *gal* and *lac* genes but does not confer new function.

In the next step of the Leloir pathway, the galactokinase (*galK*) phosphorylates α-D-galactose and then galactose-1-phosphate uridylyltransferase (*galT*) transforms galactose-1P into glucose-1P. Putative *galK* and *galT* genes are essentials and are found in all genomes holding the Leloir pathway, and not present in all genomes without this pathway (table 11). In fact, the deletion of *galK* impairs severely the ability of the strain *S. mutans* UA159 to metabolize lactose and galactose (Abranches et al., 2004). In the same way, *L. lactis* NCD02054 *galT* deficient strain accumulates galactose-1P, which leads to the inhibition of growth on lactose and galactose (Vaughan et al., 1998).

UDP-galactose 4-epimerase (*galE*) plays a role in the last step of the Leloir pathway. It is an essential enzyme and putative genes are present in all genomes holding or not the Leloir

pathway (Table 11). It should be noted that this enzyme activity also acts in other general metabolic functions such as the cell wall formation (Grossiord et al., 2003) and this fact can explain its presence in some strains without the Leloir pathway.

The *galR* regulation gene is considered as an activator of the *galKTE lacSZLM* operons, and its product is a negative regulator of its own expression. When a strain is deficient in *galR* gene, no expression of genes is observed in media containing lactose and/or galactose (Vaughan et al., 2001). Thus, it is present in genomes holding this pathway (Table 11), except for *B. subtilis* 168 and S*treptococcus. gordonii* CH1. Indeed, the former is known to be unable to use lactose. For the latter, the contribution of the Leloir pathway to galactose (Krispin and Allmansberger, 1998) or lactose catabolism appears to be not sufficiently active (Zeng et al., 2012).

Tagatose-6P pathway

Tagatose-6P pathway is present alone in 33 strains and in 74 strains holding also the Leloir pathway. Putative *galE* gene and putative *galR* are also found in few strains conferring no supplementary functions as described above (Table 11).

Putative genes encoding PTS (*lacFE*) are found, duplicated in some cases, in all genomes possessing this pathway, except in *Pediococcus acidilactici* DSM20284, *S. agalactiae* A909 and *S. agalactiae* 2603 V/R (Table 11). In general, PTS seems to be specific for each type of sugar. *Streptococcus gordonii* DL1 dedicates one PTS for lactose and another one for galactose. Strain deficient in PTS^{lac} would not internalize lactose and the one deficient in PTS^{Gal} grows slowly in medium containing galactose as a carbon source (Zeng et al., 2012). In some cases, one PTS system has the ability to transport another type of sugar but with lower affinity (Solopova et al., 2012). In *S. gordonii* DL1, galactose may enter the cell *via* PTS^{Lac}, PTS^{Man} or *via* PTS^{Gal} (Zeng et al., 2012). For the three strains lacking PTS system, lactose or galactose molecules may enter the cell *via* other PTS systems.

When the phosphorylated lactose entered *via* a PTS system, it is hydrolyzed by a 6P-β-galactosidase into glucose and galactose-6P. This essential enzyme is encoded by the gene *lacG* and the mutation of *lacG* gene inducts the inability of the strain to use lactose: *S. gordonii* DL1 *lacG* deficient strain impaired to grow on lactose but grew on galactose (Zeng et al., 2012). In fact, putative *lacG* genes were found in all genomes possessing Tagatose-6P pathway and in some cases this gene is duplicated (Table 11).

The putative *lacFE* and/or *lacG* genes were also detected in 18 genomes holding a cluster responsible of the Leloir pathway. *lacFE* putative genes are present in *C. maltaromaticum* 3.18 and *Lactococcus piscium* MKFS47, and *lacFEG* putative genes in 15 *Enterococcus faecium* strains and *Enterococcus mundtii* QU25 (Table 11). Those strains may then be capable of internalizing and phosphorylating lactose or galactose molecules. Lactose-6P is then cleaved *via* 6P-β-galactosidase encoded by *lacG* if present. A preliminary dephosphorylation of the resulting galactose-6P is necessary in order to be metabolized by the Leloir pathway as described previously (Neves et al., 2010).

Galactose-6P isomerase encoded by *lacAB* transforms galactose-6P into tagatose-6P, essential for galactose utilization. Consequently, putative *lacAB* genes presence is essential and found in all genomes holding this pathway. No *lacAB* were found in the other genomes (Table 11). In fact, accumulation of galactose-6P in mutant lacking galactose-6P isomerase (*lacAB*) causes inhibition of growth on lactose/galactose medium (Abranches et al., 2004; Zeng et al., 2012, 2010).

Enzyme encoded by *lacC* phosphorylates tagatose-6P in order to form tagatose-1,6diP. It is present in all genomes possessing the Tagatose-6P pathway except in *Lactobacillus johnsonii* FI9785, *Lactobacillus murinus* ASF361 and *P. acidilactici* DSM20284 (Table 11). Contrary to *lacAB* mutant, *lacC* mutant showed normal cells growing on medium containing lactose and/or galactose as a source of carbohydrate (Abranches et al., 2004; Zeng et al., 2012, 2010). *LacC* belongs to the pfkB family of carbohydrate kinases, which include ribokinase (RbsK) and 1-phosphofructokinase (FruK) (Bork et al., 1993). It may present an enzymatic activity on fructose-6P (Bissett and Anderson, 1980). This might explain the presence of *lacC* in several copies and in five other genomes holding no Tagatose-6P pathway.

Enzyme encoded by *lacD* is indispensable and promotes the cleavage of Tagatose-1,6 diP into 2 trioses. The *lacD* mutation causes the accumulation of Tagatose-1,6diP in the medium which leads to the inhibition of growth of the cells (Abranches et al., 2004; Zeng et al., 2012, 2010). In other ways, Tagatose 1,6 diP aldolase may play a role in the cleavage of fructose-1,6diP (Bissett and Anderson, 1980) and *lacD* can be found in genomes even in the absence of the Tagatose-6P pathway (Table 11).

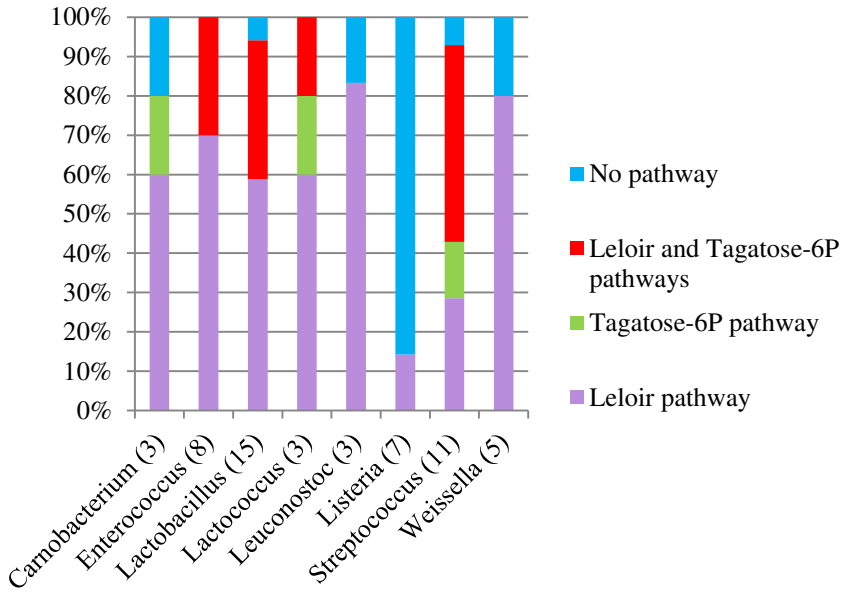
lacT is a gene encoding a transcriptional antiterminator protein present within the *lac* operon of many LAB, that may play a role in the regulation down-stream *lacFEG* expression

(Richards et al., 2013; Zeng et al., 2012). *LacT* is a member of the BglG/SacY family of proteins commonly associated with the regulation of carbohydrate metabolism (Declerck et al., 2002). *LacT* is present in some strains holding the Tagatose-6P pathway, but *lacR* is present in all genomes possessing the Tagatose-6P pathway (Table 11). *lacR* is one component of the regulation system. In the Tagatose-6P pathway, it is an auto regulation gene that binds to *lacA* promoter and represses its activity when tagatose-6P is absent (de Vos and Vaughan, 1994; van Rooijen et al., 1991).

2.2.3.2.Diversity of the lactose and galactose metabolic pathways

Leloir and Tagatose-6P pathways contribute differently to lactose and galactose metabolism depending on genus and species. Repartition of these two pathways between the different genus and species is analyzed (Table 11, Figure 18 A-B). Only genus containing more than three species and species containing more than eight strains are taking into consideration. In the different genus studied, the Leloir pathway is always present and can be found in 60% or more of species in each genus, except for *Listeria,* where the absence of *lac* and *gal* genes is observed for more than 85% of *Listeria* species. The concomitant presence of Leloir and Tagatose-6P pathways is present in four different genus, *Enterococcus*, *Lactobacillus*, *Lactococcus* and *Streptococcus.* The Tagatose-6P pathway is rarely present alone in the different genus and species analyzed and was ideed found in 20% of *Carnobacterium*, *Lactococcus*, and *Streptococcus* (Figure 18.A). On the species level, the genetic profies are rather homologeous (Figure 18.B). Fifty percent (5/9) of the species studied present only one profile, *i.e.* two holding the two pathways together (*Enterococcus faecalis*, *Streptococcus suis*), two possessing only the Tagatose-6P pathway (*Staphylococcus aureus*, *Streptococcus pyogenes*) and one lacking of pathways (*Listeria. monocytogenes*). Thirteen percent (3/9) of species show two different profiles (*Enterococcus faecium*, *Lactobacillus rhamnosus*, *Streptococcus sanguinis*), some strains of the same species possessing only Leloir pathway and the others the two pathways together. *Carnobacterium maltaromaticum* is the only species showing intraspecific variability in profiles. Some possess the Leloir pathway, some others the Tagatose-6P pathway, one strain hold the *lac* and *gal* genes and some none of the pathways (Figure 18-B).

A.



B.

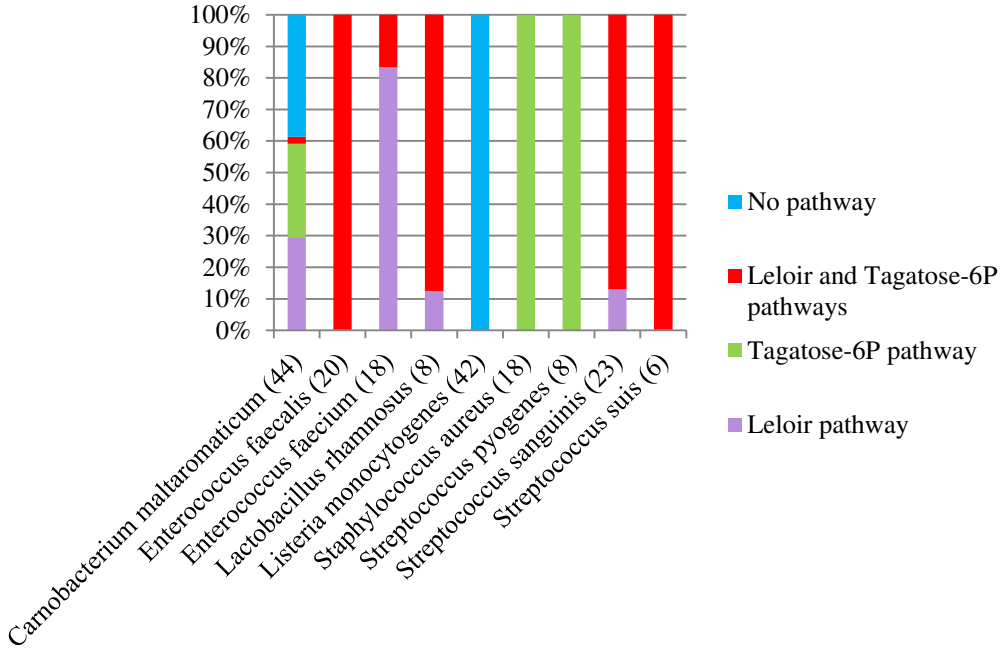


Figure 18: Diversity of the lactose and galactose metabolic pathways within A. genus and B. species of the different genus analyzed.
***Carnobacterium maltaromaticum* strains hold the 5 strains of the MicroScope platform and 42 strains tested by PCR in a previous study.**

This analysis permits to highlight the variation level of the presence of the pathways. It is clear that the lactose and galactose metabolic pathway diversity is interspecific, except for *C. maltaromaticum* for which an intraspecific variability was observed. Two hypothesis can explain the dominant presence of the Leloir pathway over the Tagatose-6P pathway.

Diversity within genomes can be relied to the functionality of each pathway in the cell. These two pathways play different roles in metabolism of lactose and galactose.

The Tagatose-6P pathway is mainly dedicated to galactose-6P metabolism. Tagatose-6P pathway transforms galactose-6P into triose-diP that will be implicated in the glycolysis. Strains like *L. casei* (Bettenbrock et al., 1999), *L. rhamnosus* (Tsai and Lin, 2006), *L. lactis* (Siezen et al., 2005), *S. mutans* (Jagusztyn-Krynicka et al., 1992) (Table 11) utilize the Tagatose-6P pathway for lactose and galactose metabolism. *Streptococcus gordonii* CH1 presents duplicated *lac* cluster. Paralogs have become more specific; one cluster is implicated in the lactose utilization and the other in the galactose metabolism (Zeng et al., 2010). The Tagatose-6P pathway is also found to contribute to the pathogenicity of some strains. It either contributes to the adaptation of the pathogenic strain in dairy environment, like in *S. aureus* (Rosey et al., 1991) or is directly involved in pathogenic reactions like in *S. agalactiae* NEM316 and *S. pyogenes* strains possessing a duplicated *lac* cluster. *Streptococcus agalactiae* NEM316 has been demonstrated to possess one *lac* cluster involved in virulent reactions due to the presence of putative genes encoding a neuraminidase in the operon and the other cluster conserves its function in the lactose metabolism (Richards et al., 2013). *Streptococcus pyogenes* strains possess one cluster with truncated *lacC* preceding a *lacD* involved in the regulation of the expression of virulent genes (Loughman and Caparon, 2007). Leloir pathway is implicated in galactose metabolism and in UDP-glucose and UDP-galactose synthesis, which are precursors of glucose- and galactose-containing exopolysaccharides biosynthesis (Boels et al., 2003; Welman and Maddox, 2003). Thus, galactose will not be imperatively redirected to the central metabolism of sugars. EPS play important role in adhesion to solid surfaces, biofilm formation, cellular recognition and pathogenicity. In addition to that, they play vital role in protection of the microbes from adverse conditions as desiccation, nutrient shortage, toxic compounds, bacteriophages, osmotic stress and antagonists (Looijesteijn et al., 2001). In fact, the strain *S. gordonii* CH1 possesses Leloir and Tagatose-6P pathways (Table 11). In addition to the diversification in functionality of the two *lac* clusters as described before, the *gal* genes forming the Leloir pathway have been demonstrated not to be involved in the

galactose metabolism, *i.e.* mutation of all *lac* genes does not imply the activation of the Leloir pathway for lactose or galactose metabolism (Zeng et al., 2010). On the other hand, Leloir pathway of *L. lactis* SK11 (Thompson, 1980) and *S. mutans* UA159 (Ajdic and Ferretti, 1998) has been demonstrated to be involved in the galactose metabolism.

The origin of lactose and galactose molecules may be involved in the diversity of the presence of one or another metabolic pathway. While lactose is only found in the dairy environment, plants associated media can be rich in galactose and galactose-containing molecules. Notably, the intestinal tract of mammals can contain, next to a variety of diet-derived carbohydrates, substantial quantities of galactose which can be hydrolyze by different galactosidases (Flint et al., 2012).

Lactose molecules, in order to be metabolized *via* the Leloir or the Tagatose-6P pathway, should be hydrolyzed by the action of two enzyme encoded by *lacZ/lacLM* and *lacG* respectively (Premi et al., 1972). Those two hydrolases are part of the glycosyl hydrolases families and are not the only enzymes responsible for sugar cleavage in nature. Fourty-five glycosyl hydrolases families with 52 E.C. entries, classification based on amino acid sequence comparison of 482 sequences, were identified (Henrissat, 1991; Henrissat and Bairoch, 1993). Four families (1, 2, 35 and 42) contains enzymes with β-galactosidase activities. The P-β-galactosidase *lacG* is a family 1 member, and *lacZ* is a family 2 member. Bacteria may contain several type of glycosyl hydrolases in order to utilize carbohydrates from nature. Two surface proteins β-galactosidase - BgaA (β1,4-galactosidase) and BgaC (β1,3-galactosidase) - are present in *S. pneumoniae* (Jeong et al., 2009; Zähner and Hakenbeck, 2000). They are not included in lactose metabolism but may be involved in interaction with other bacterial cells, by hydrolyzing the glycan. In *C. maltaromaticum* strain BA *(piscicola)* (Coombs and Brenchley, 2001, 1999) three hydrolyses, α-galactosidase AgaA (family 36) and two β-galactosidase BgaB (family 42) and bgaC (family 35), not involved in lactose utilization, were present. Galactose-containing substrates will then be hydrolyzed in the extracellular medium and galactose molecules enter the cell in order to be metabolized *via* the Leloir pathway. Indeed, Leloir pathway supports the highest galactose consumption rate (Neves et al., 2010).

Therefore, it is not surprising to find that the Leloir pathway is more dispersed in all genera and is present in the majority of strains, whereas the Tagatose-6P pathway is less disseminated within the order *Lactobacillales*.

2.2.3.3.Conclusion

Comparative genome analysis of *gal* and *lac* genes of the Leloir and Tagatose-6P pathways permits to highlight the variation level in gene presence among the LAB genus studied. Leloir pathway is present in 63% of the 237 strains analyzed whereas Tagatose-6P pathway in only 13%. The lactose and galactose metabolic pathway diversity remains genus dependent, except for *Carnobacterium* genus for which it is also species dependent. Between all the species studied, *C. maltaromaticum* is the only one to possess such a diversity in the metabolic profiles, which can be related to its high genetic diversity, visualized in an MLST scheme (Rahman et al., 2014a).

CHAPTER III: CONCLUSIONS AND PERSPECTIVES

L'objectif des travaux réalisés était d'approfondir les connaissances sur *C. maltaromaticum* par des études d'analyse génomique comparée et de répondre à différentes questions :

- Quels sont les gènes impliqués dans l'adaptation au contexte alimentaire et digestif ?

- La diversité des souches de *C. maltaromaticum* présentes dans l'environnement laitier peut-elle se retrouver au niveau des voies métaboliques d'utilisation du lactose ?

Caractéristique du génome

La taille des génomes des différentes espèces de *C. maltaromaticum* est de 1,45 Mbp supérieure à celles des autres espèces de *Carnobacterium*. Deux hypothèses peuvent expliquer ces différences de tailles. Soit un gain massif de gènes aurait eu lieu chez l'ancêtre de *C. maltaromaticum,* soit deux pertes massives de gènes auraient eu lieu : une chez l'ancêtre de *C. divergens* et une autre chez l'ancêtre commun de *Carnobacterium* sp. AT7, *Carnobacterium* sp. 17.4, et *C. inhibens* subsp. *gilichinskyi* WN1359. L'hypothèse la plus parcimonieuse est en faveur d'un gain massif chez l'ancêtre de *C. maltaromaticum*. Par ailleurs, il est attendu qu'une évolution par perte de gènes se traduise par la présence d'une proportion importante de pseudogènes. Or, l'analyse des génomes de ces carnobactéries a révélé une faible proportion de pseudogènes, renforçant ainsi l'hypothèse d'un gain massif de gènes.

La distribution des CDS dans les différents groupes de COG n'est pas différente entre les souches de *C. maltaromaticum*. Ces génomes renferment des voies métaboliques spécifiques, absentes chez les autres espèces de *Carnobacterium.* Par ailleurs, une majorité des gènes spécifiques mis en évidence est de fonctions inconnues. Il serait intéressant de déterminer la fonction de ces gènes, afin de mettre en évidence la particularité de *C. maltaromaticum* et leur implication dans l'adaptation à différents environnements. Une première étape pour caractériser ces gènes de fonction inconnue pourrait consister à déterminer si ces gènes sont exprimés et dans quelles conditions. Cela pourrait permettre de déterminer dans quel habitat ces gènes sont utiles pour la bactérie. Ainsi, des analyses de transcriptomes pourraient être réalisées à partir de carnobactéries cultivées dans différents aliments ainsi que dans des conditions mimant celles du tube digestif.

L'étude des protéines de surface de *C. maltaromaticum* et de *C. divergens* V41 a mis en évidence un sécrétome de taille importante. Le sécrétome de *C. maltaromaticum* serait le plus

grand jamais identifié chez les bactéries lactiques. En revanche, les analyses de génome de *Carnobacterium* sp. AT7, *Carnobacterium* sp. 17.4, et *C. inhibens* subsp. *gilichinskyi* WN1359 ont révélé un sécrétome de très petite taille. De plus, les génomes de *C. maltaromaticum* et de *C. divergens* présentent de multiples homologues de sortases et une grande diversité de protéines ancrées de façon covalente à la paroi. La surface cellulaire de ces carnobactéries présenterait également une grande diversité de protéines ancrées de façon non covalente à la paroi et notamment de nombreuses protéines présentant des domaines d'attachement WxL. En revanche, des trois autres carnobactéries, seule *Carnobacterium* sp. 17.4 présenterait un homologue de sortase et une protéine ancrée de façon covalente à la paroi. Ainsi, *Carnobacterium* sp. 17.4, *Carnobacterium* sp. AT7 et *C. inhibens* subsp. *gilichinskyi* WN1359 présenteraient des surfaces cellulaires très faiblement diversifiées en protéines de surface. Une situation similaire a été décrite dans le genre *Streptococcus* dans lequel une grande diversité de protéines de surface est décrite pour les espèces du genre, excepté pour *S. thermophilus* dont la surface présenterait très peu de protéines de surface différentes. *S. thermophilus* est une espèce dont le seul habitat connu est l'environnement laitier. Le génome de cette bactérie a évolué par réduction, il est en effet riche en pseudogènes. Il est supposé que ces caractéristiques sont la conséquence d'un confinement écologique à l'environnement laitier. A l'opposé, *C. maltaromaticum* est une espèce ubiquiste qui présenterait une surface cellulaire très diversifiée. Cette surface diversifiée pourrait lui permettre d'établir une grande diversité d'interactions avec son environnement et ainsi de coloniser une grande diversité d'habitats.

Les études de comparaison génomique ont été faites sur une collection de 5 souches de *C. maltaromaticum* et une souche de *C. divergens* V41. Ces deux bactéries présentent des propriétés de surfaces proches. Il serait intéressant d'élargir l'étude de comparaison génomique à un plus grand nombre de souches de *C. divergens* et d'autres espèces de *Carnobacterium*.

Métabolisme du lactose

L'analyse génomique comparée réalisée sur l'ensemble des *Carnobacterium* a montré que les gènes de la voie du Tagatose-6P et de la voie de Leloir sont présents avec un haut niveau de diversité. En effet, les génomes des différentes souches peuvent contenir les gènes impliqués uniquement dans une des deux voies, dans les deux voies ou bien ne pas posséder ces gènes. Par ailleurs, la voie du Tagatose-6P renferme des gènes homologues au sein d'un même génome, avec cependant une organisation spécifique des gènes. L'organisation des gènes de la voie de Leloir est similaire dans les différents génomes.

L'analyse de la présence des gènes de la voie du Tagatose-6P et de la voie de Leloir au sein de la population de 42 souches de *C. maltaromaticum* indique que la voie du Tagatose-6P est présente dans les lignées I, II et III et la voie de Leloir dans les lignées I, II et IV. Par ailleurs la majorité des souches du complexe clonal 1 (CC1) issues de l'environnement laitier possède les gènes de la voie du Tagatose-6P, qui peuvent être transmis d'une bactérie à une autre *via* des HTG. Ceci suggère que l'expansion de la lignée II et spécialement du CC1, peut être due à la présence de ces gènes. Il est à noter que cette voie est présente sur le mégaplasmide de *C. maltaromaticum* LMA28. Il serait intéressant de rechercher le caractère plasmidique de ces gènes *lac* chez les autres souches du CC1.

Les études de génomique comparée réalisées sur les 237 génomes disponibles sur la plateforme MicroScope ont mis en évidence un niveau de variation élevé dans les gènes impliqués dans la voie du Tagatose-6P et de la voie de Leloir. La voie de Leloir est présente chez 63% des souches analysées alors que la voie du Tagatose-6P n'est présente que dans 45% des souches. Cette voie n'est présente seule que chez 14% des souches, comparativement à 31% pour la voie de Leloir. Par ailleurs une grande variabilité dans la présence des gènes est constatée au sein des différentes voies mises en évidence.

Cette prépondérance de la présence de la voie de Leloir peut s'expliquer de différentes façons : le devenir des métabolites issus de chacune des voies et la disponibilité en substrat lactose ou galactose.

La diversité de la répartition des voies du Tagatose-6P et de la voie de Leloir est constatée au niveau des genres bactériens alors qu'elle est observée au niveau de l'espèce chez *C. maltaromaticum*.

À l'échelle de la population de *C. maltaromaticum*, les résultats ont révélé que les homologues liés aux gènes des voies du Tagatose-6P et de Leloir sont diffusés dans de multiples lignées. En conclusion, ces résultats indiquent que la capacité à métaboliser le lactose est génétiquement diversifiée au sein de l'espèce *C. maltaromaticum* et suggèrent que l'histoire évolutive ayant conduit à la capacité à métaboliser le lactose chez *C. maltaromaticum* est le résultat d'événements génétiques multiples. Ces évènements auraient contribué à moduler le génome de cette bactérie et son aptitude à s'adapter à une large gamme d'environnements.

Il serait intéressant de déterminer le rôle effectif des gènes dans le métabolisme du lactose et du galactose. Pour cela deux approches pourraient être envisagées : soit une approche par inactivation de gènes, soit par expression hétérologue des gènes d'intérêt chez un hôte ne métabolisant ni le lactose ni le galactose.

***Carnobacterium maltaromaticum*, espèce pathogène** ?

Carnobacterium maltaromaticum a été classée en 1986 en tant que pathogène opportuniste (Michel et al., 1986). Des gènes de virulences putatifs ont été identifiés chez *C. maltaromaticum* ATCC35586 isolée d'un poisson malade (Leisner et al., 2012). Il serait alors intéressant de séquencer le génome d'autres *C. maltaromaticum* isolées de poissons malades. Une analyse génomique comparée pourrait permettre de mettre en évidence des gènes potentiels de virulence et d'établir les relations phylogénétiques entre ces souches et les autres. Ceci permettra de déterminer si ces souches isolées de poissons malades forment une lignée distincte.

***Carnobacterium maltaromaticum*, espèce du microbiote intestinal** ?

Des travaux antérieurs ont démontré la capacité de *C. maltaromaticum* LMA28 à survivre durant le transit gastro-intestinal de souris (Rahman et al., 2014b). Les résultats de cette thèse semblent indiquer que d'autres souches de *C. maltaromaticum* présenteraient des composants de surface cellulaire d'adaptation au tube digestif. Notamment, de nombreuses adhésines ont été mises en évidence, suggérant que la bactérie pourrait non seulement survivre dans le tube digestif mais pourrait également persister en adhérant à la muqueuse intestinale. Il serait intéressant de déterminer la présence de *C. maltaromaticum* et des carnobactéries en général, en analysant le métagénome du microbiote intestinal.

REFERENCES

Abbot, E.L., Smith, W.D., Siou, G.P.S., Chiriboga, C., Smith, R.J., Wilson, J.A., Hirst, B.H., Kehoe, M.A., 2007. Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. Cell. Microbiol. 9, 1822–1833. doi:10.1111/j.1462-5822.2007.00918.x

Abranches, J., Chen, Y.-Y.M., Burne, R.A., 2004. Galactose Metabolism by *Streptococcus mutans*. Appl. Environ. Microbiol. 70, 6047–6052. doi:10.1128/AEM.70.10.6047-6052.2004

Afzal, M.I., Ariceaga, C.C.G., Lhomme, E., Ali, N.K., Payot, S., Burgain, J., Gaiani, C., Borges, F., Revol-Junelles, A.-M., Delaunay, S., Cailliez-Grimal, C., 2013a. Characterization of *Carnobacterium maltaromaticum* LMA 28 for its positive technological role in soft cheese making. Food. Microbiol. 36, 223–230. doi:10.1016/j.fm.2013.05.008

Afzal, M.I., Boulahya, K.-A., Paris, C., Delaunay, S., Cailliez-Grimarl, C., 2013b. Effect of oxygen on the biosynthesis of flavor compound 3-methylbutanal from leucine catabolism during batch culture in *Carnobacterium maltaromaticum* LMA 28. J. Dairy Sci. 96, 352–359. doi:10.3168/jds.2012-6088

Afzal, M.I., Delaunay, S., Paris, C., Borges, F., Revol-Junelles, A.-M., Cailliez-Grimal, C., 2012. Identification of metabolic pathways involved in the biosynthesis of flavor compound 3-methylbutanal from leucine catabolism by *Carnobacterium maltaromaticum* LMA 28. Int. J. Food Microbiol. 157, 332–339. doi:10.1016/j.ijfoodmicro.2012.05.010

Afzal, M.I., Jacquet, T., Delaunay, S., Borges, F., Milliere, J.-B., Revol-Junelles, A.-M., Cailliez-Grimal, C., 2010. *Carnobacterium maltaromaticum*: Identification, isolation tools, ecology and technological aspects in dairy products. Food. Microbiol. 27, 573–579. doi:10.1016/j.fm.2010.03.019

Aguirre, L., Hebert, E.M., Garro, M.S., Savoy de Giori, G., 2014. Proteolytic activity of *Lactobacillus* strains on soybean proteins. LWT - Food Science and Technology 59, 780–785. doi:10.1016/j.lwt.2014.06.061

Aguirre, M., Collins, M. d., 1993. Lactic acid bacteria and human clinical infection. J. Appl. Bacteriol. 75, 95–107. doi:10.1111/j.1365-2672.1993.tb02753.x

Ainsworth, S., Zomer, A., de Jager, V., Bottacini, F., van Hijum, S.A.F.T., Mahony, J., van Sinderen, D., 2013. Complete Genome of *Lactococcus lactis* subsp. *cremoris* UC509.9, Host for a Model Lactococcal P335 Bacteriophage. Genome Announc. 1. doi:10.1128/genomeA.00119-12

Ajdic, D., Ferretti, J.J., 1998. Transcriptional regulation of the *Streptococcus mutans gal* operon by the *GalR* repressor. J. Bacteriol. 180, 5727–5732.

Alahuhta, M., Xu, Q., Brunecky, R., Adney, W.S., Ding, S.-Y., Himmel, M.E., Lunin, V.V., 2010. Structure of a fibronectin type III-like module from Clostridium thermocellum. Acta Cryst. F 66, 878–880. doi:10.1107/S1744309110022529

Aleksandrzak-Piekarczyk, T., Kok, J., Renault, P., Bardowski, J., 2005. Alternative Lactose Catabolic Pathway in *Lactococcus lactis* IL1403. Appl. Environ. Microbiol. 71, 6060–6069. doi:10.1128/AEM.71.10.6060-6069.2005

Altermann, E., Russell, W.M., Azcarate-Peril, M.A., Barrangou, R., Buck, B.L., McAuliffe, O., Souther, N., Dobson, A., Duong, T., Callanan, M., Lick, S., Hamrick, A., Cano, R., Klaenhammer, T.R., 2005. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. Proc. Natl. Acad. Sci. U.S.A. 102, 3906–3912. doi:10.1073/pnas.0409188102

Ammerman, J.W., Azam, F., 1985. Bacterial 5-nucleotidase in aquatic ecosystems: a novel mechanism of phosphorus regeneration. Science 227, 1338–1340. doi:10.1126/science.227.4692.1338

Anbukkarasi, K., Nanda, D.K., UmaMaheswari, T., Hemalatha, T., Singh, P., Singh, R., 2014. Assessment of expression of Leloir pathway genes in wild-type galactose-fermenting *Streptococcus thermophilus* by real-time PCR. Eur. Food Res. Technol. 239, 895–903. doi:10.1007/s00217-014-2286-9

Audenaert, K., D'Haene, K., Messens, K., Ruyssen, T., Vandamme, P., Huys, G., 2010. Diversity of lactic acid bacteria from modified atmosphere packaged sliced cooked meat products at sell-by date assessed by PCR-denaturing gradient gel electrophoresis. Food Microbiol. 27, 12–18. doi:10.1016/j.fm.2009.04.006

Axelsson, L., 2004. Lactic acid bacteria: classification and physiology, in: Lactic Acid Bacteria: Microbiological and Functional Aspects. Salminen, S., von Wright, A., Ouwehand, A., pp. 1–66.

Azcarate-Peril, M.A., Altermann, E., Goh, Y.J., Tallon, R., Sanozky-Dawes, R.B., Pfeiler, E.A., O'Flaherty, S., Buck, B.L., Dobson, A., Duong, T., Miller, M.J., Barrangou, R., Klaenhammer, T.R., 2008. Analysis of the genome sequence of *Lactobacillus gasseri* ATCC 33323 reveals the molecular basis of an autochthonous intestinal organism. Appl. Environ. Microbiol. 74, 4610–4625. doi:10.1128/AEM.00054-08

Azcarate-Peril, M.A., McAuliffe, O., Altermann, E., Lick, S., Russell, W.M., Klaenhammer, T.R., 2005. Microarray Analysis of a Two-Component Regulatory System Involved in Acid Resistance and Proteolytic Activity in *Lactobacillus acidophilus*. Appl. Environ. Microbiol. 71, 5794–5804. doi:10.1128/AEM.71.10.5794-5804.2005

Baba, T., Schneewind, O., 1996. Target cell specificity of a bacteriocin molecule: a C-terminal signal directs lysostaphin to the cell wall of *Staphylococcus aureus*. EMBO J. 15, 4789–4797.

Bachmann, H., Starrenburg, M.J.C., Molenaar, D., Kleerebezem, M., Vlieg, J.E.T. van H., 2012. Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. Genome Res. 22, 115–124. doi:10.1101/gr.121285.111

Balderas, M.A., Nobles, C.L., Honsa, E.S., Alicki, E.R., Maresso, A.W., 2012. Hal Is a *Bacillus anthracis* heme acquisition protein. J. Bacteriol. 194, 5513–5521. doi:10.1128/JB.00685-12

Barabote, R.D., Saier, M.H., 2005. Comparative Genomic Analyses of the Bacterial Phosphotransferase System. Microbiol. Mol. Biol. Rev. 69, 608–634. doi:10.1128/MMBR.69.4.608-634.2005

Barakat, R.K., Griffiths, M.W., Harris, L.J., 2000. Isolation and characterization of *Carnobacterium, Lactococcus*, and *Enterococcus* spp. from cooked, modified atmosphere packaged, refrigerated, poultry meat. Int. J. Food Microbiol. 62, 83–94.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Sci. 315, 1709–1712. doi:10.1126/science.1138140

Bateman, A., Bycroft, M., 2000. The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). J. Mol. Biol. 299, 1113–1119. doi:10.1006/jmbi.2000.3778

Baya, A.M., Toranzo, A.E., Lupiani, B., Li, T., Roberson, B.S., Hetrick, F.M., 1991. Biochemical and serological characterization of *Carnobacterium* spp. isolated from farmed and natural populations of striped bass and catfish. Appl. Environ. Microbiol. 57, 3114–20.

Bengis-Garber, C., Kushner, D.J., 1982. Role of membrane-bound 5'-nucleotidase in nucleotide uptake by the moderate halophile *Vibrio costicola*. J. Bacteriol. 149, 808–815.

Ben Hamida Nouaili, E., Abidi, K., Chaouachi, S., Marrakchi, Z., 2011. Epidemiology of maternal-fetal group b streptococcal infections. Médecine et Maladies Infectieuses 41, 123–125. doi:10.1016/j.medmal.2010.09.004

Benson, A.K., David, J.R.D., Gilbreth, S.E., Smith, G., Nietfeldt, J., Legge, R., Kim, J., Sinha, R., Duncan, C.E., Ma, J., Singh, I., 2014. Microbial Successions Are Associated with Changes in Chemical Profiles of a Model Refrigerated Fresh Pork Sausage during an 80-Day Shelf Life Study. Appl. Environ. Microbiol. 80, 5178–5194. doi:10.1128/AEM.00774-14

Bettenbrock, K., Alpert, C.-A., 1998. The *gal* Genes for the Leloir Pathway of *Lactobacillus casei* 64H. Appl. Environ. Microbiol. 64, 2013–2019.

Bettenbrock, K., Siebers, U., Ehrenreich, P., Alpert, C.-A., 1999. *Lactobacillus casei* 64H Contains a Phosphoenolpyruvate-Dependent Phosphotransferase System for Uptake of Galactose, as Confirmed by Analysis of ptsH and Differentgal Mutants. J. Bacteriol. 181, 225–230.

Bierne, H., Cossart, P., 2007. *Listeria monocytogenes* Surface Proteins: from Genome Predictions to Function. Microbiol. Mol. Biol. Rev. 71, 377–397. doi:10.1128/MMBR.00039-06

Bierne, H., Garandeau, C., Pucciarelli, M.G., Sabet, C., Newton, S., Garcia-del Portillo, F., Cossart, P., Charbit, A., 2004. Sortase B, a New Class of Sortase in *Listeria monocytogenes*. J. Bacteriol. 186, 1972–1982. doi:10.1128/JB.186.7.1972-1982.2004

Bissett, D.L., Anderson, R.L., 1980. Lactose and D-galactose metabolism in *Staphylococcus aureus* . III. Purification and properties of D-tagatose-6-phosphate kinase. J. Biol. Chem. 255, 8745–8749.

Bjorkroth, J., Ristiniemi, M., Vandamme, P., Korkeala, H., 2005. *Enterococcus* species dominating in fresh modified-atmosphere-packaged, marinated broiler legs are overgrown by *Carnobacterium* and *Lactobacillus* species during storage at 6 degrees C. Int. J. Food Microbiol. 97, 267–276. doi:10.1016/j.ijfoodmicro.2004.04.011

Blobel, G., 1980. Intracellular protein topogenesis. Proc. Natl. Acad. Sci.. U. S. A. 77, 1496–1500.

Boekhorst, J., Helmer, Q., Kleerebezem, M., Siezen, R.J., 2006a. Comparative analysis of proteins with a mucus-binding domain found exclusively in lactic acid bacteria. Microbiology (Reading, Engl.) 152, 273–280. doi:10.1099/mic.0.28415-0

Boekhorst, J., Wels, M., Kleerebezem, M., Siezen, R.J., 2006b. The predicted secretome of *Lactobacillus plantarum* WCFS1 sheds light on interactions with its environment. Microbiol. 152, 3175–3183. doi:10.1099/mic.0.29217-0

Boels, I.C., van Kranenburg, R., Kanning, M.W., Chong, B.F., de Vos, W.M., Kleerebezem, M., 2003. Increased Exopolysaccharide Production in *Lactococcus lactis* due to Increased Levels of Expression of the NIZO B40 *eps* Gene Cluster. Appl. Environ. Microbiol. 69, 5029–5031. doi:10.1128/AEM.69.8.5029-5031.2003

Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E., Mazur, M., Pusch, G.D., Fonstein, M., Overbeek, R., Kyprides, N., Purnelle, B., Prozzi, D., Ngui, K., Masuy, D., Hancy, F., Burteau, S., Boutry, M., Delcour, J., Goffeau, A., Hols, P., 2004. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. Nat. Biotechnol. 22, 1554–1558. doi:10.1038/nbt1034

Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D., Sorokin, A., 2001. The Complete Genome Sequence of the Lactic Acid Bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome Res. 11, 731–753. doi:10.1101/gr.169701

Borges, F., Layec, S., Thibessard, A., Fernandez, A., Gintz, B., Hols, P., Decaris, B., Leblond-Bourget, N., 2005. cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein Involved in Cellular Segregation in *Streptococcus thermophilus*. J. Bacteriol. 187, 2737–2746. doi:10.1128/JB.187.8.2737-2746.2005

Bork, P., Sander, C., Valencia, A., 1993. Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. Protein Sci 2, 31–40.

Bouffard, G.G., Rudd, K.E., Adhya, S.L., 1994. Dependence of Lactose Metabolism upon Mutarotase Encoded in the *gal* Operon in *Escherichia coli*. J. Mol. Biol. 244, 269–278. doi:10.1006/jmbi.1994.1728

Bowden, C.F.M., Verstraete, M.M., Eltis, L.D., Murphy, M.E.P., 2014. Hemoglobin binding and catalytic heme extraction by IsdB near iron transporter domains. Biochem. 53, 2286–2294. doi:10.1021/bi500230f

Bowden, G.H.W., 2000. The microbial ecology of dental caries. Microbial ecology in health and disease 12, 138–148.

Bradshaw, W.J., Davies, A.H., Chambers, C.J., Roberts, A.K., Shone, C.C., Acharya, K.R., 2015. Molecular features of the sortase enzyme family. FEBS J. n/a–n/a. doi:10.1111/febs.13288

Bratina, B.J., Stevenson, B.S., Green, W.J., Schmidt, T.M., 1998. Manganese reduction by microbes from oxic regions of the Lake Vanda (Antarctica) water column. Appl. Environ. Microbiol. 64, 3791–3797.

Braun, L., Dramsi, S., Dehoux, P., Bierne, H., Lindahl, G., Cossart, P., 1997. InlB: an invasion protein of *Listeria monocytogenes* with a novel type of surface association. Mol. Microbiol. 25, 285–294. doi:10.1046/j.1365-2958.1997.4621825.x

Breidt, F., Hengstenberg, W., Finkeldei, U., Stewart, G.C., 1987. Identification of the genes for the lactose-specific components of the phosphotransferase system in the *lac* operon of *Staphylococcus aureus*. J. Biol. Chem. 262, 16444–16449.

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, 1244–1253. doi:10.1128/JB.00773-06

Broadbent, J.R., Cai, H., Larsen, R.L., Hughes, J.E., Welker, D.L., De Carvalho, V.G., Tompkins, T.A., Ardö, Y., Vogensen, F., De Lorentiis, A., Gatti, M., Neviani, E., Steele, J.L., 2011. Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains1. J. Dairy Sci. 94, 4313–4328. doi:10.3168/jds.2010-4068

Broadbent, J.R., Hughes, J.E., Welker, D.L., Tompkins, T.A., Steele, J.L., 2013. Complete Genome Sequence for *Lactobacillus helveticus* CNRZ 32, an Industrial Cheese Starter and Cheese Flavor Adjunct. Genome Announc. 1. doi:10.1128/genomeA.00590-13

Brooks, J.L., Moore, A.S., Patchett, R.A., Collins, M.D., Kroll, R.G., 1992. Use of the polymerase chain reaction and oligonucleotide probes for the rapid detection and identification of *Carnobacterium* species from meat. J Appl Bacteriol 72, 294–301.

Budzik, J.M., Marraffini, L.A., Schneewind, O., 2007. Assembly of pili on the surface of *Bacillus cereus* vegetative cells. Mol. Microbiol. 66, 495–510. doi:10.1111/j.1365-2958.2007.05939.x

Budzik, J.M., Oh, S.-Y., Schneewind, O., 2009. Sortase D Forms the Covalent Bond That Links BcpB to the Tip of *Bacillus cereus* Pili. J. Biol. Chem. 284, 12989–12997. doi:10.1074/jbc.M900927200

Buist, G., Steen, A., Kok, J., Kuipers, O.P., 2008. LysM, a widely distributed protein motif for binding to (peptido)glycans. Mol. Microbiol. 68, 838–847. doi:10.1111/j.1365-2958.2008.06211.x

Burgain, J., Scher, J., Francius, G., Borges, F., Corgneau, M., Revol-Junelles, A.M., Cailliez-Grimal, C., Gaiani, C., 2014a. Lactic acid bacteria in dairy food: Surface characterization and interactions with food matrix components. Adv. Colloid Interface Sci. 213, 21–35. doi:10.1016/j.cis.2014.09.005

Burgain, J., Scher, J., Lebeer, S., Vanderleyden, J., Cailliez-Grimal, C., Corgneau, M., Francius, G., Gaiani, C., 2014b. Significance of bacterial surface molecules interactions with milk proteins to enhance microencapsulation of *Lactobacillus rhamnosus* GG. Food Hydrocolloids 41, 60–70. doi:10.1016/j.foodhyd.2014.03.029

Buttin, G., 1963. Mécanismes régulateurs dans la biosynthèse des enzymes du métabolisme du galactose chez *Escherichia coli* K12: II. Le Déterminisme génétique de la régulation. J. Mol. Biol. 7, 183–. doi:10.1016/S0022-2836(63)80045-5

Cailliez-Grimal, C., Afzal, M.I., Revol-Junelles, A.-M., 2014. *Carnobacterium*, in: Encyclopedia of Food Microbiology, Elsevier, Academic Press. Batt, C. A;, Tortorello, M. L., pp. 379–383.

Cailliez-Grimal, C., Chaillou, S., Anba-Mondoloni, J., Loux, V., Afzal, M.I., Rahman, A., Kergourlay, G., Champomier-Vergès, M.-C., Zagorec, M., Dalgaard, P., Leisner, J.J., Prévost, H., Revol-Junelles, A.-M., Borges, F., 2013. Complete Chromosome Sequence of *Carnobacterium maltaromaticum* LMA 28. Genome Announc. 1, e00115–12. doi:10.1128/genomeA.00115-12

Cailliez-Grimal, C., Edima, H.C., Revol-Junelles, A.-M., Millière, J.-B., 2007. Short communication: *Carnobacterium maltaromaticum*: the only *Carnobacterium* species in French ripened soft cheeses as revealed by polymerase chain reaction detection. J. Dairy Sci. 90, 1133–1138. doi:10.3168/jds.S0022-0302(07)71599-0

Cailliez-Grimal, C., Miguindou-Mabiala, R., Leseine, M., Revol-Junelles, A.M., Milliere, J.B., 2005. Quantitative polymerase chain reaction used for the rapid detection of *Carnobacterium* species from French soft cheeses. FEMS Microbiol. Lett. 250, 163–169. doi:10.1016/j.femsle.2005.05.037

Callanan, M., Kaleta, P., O'Callaghan, J., O'Sullivan, O., Jordan, K., McAuliffe, O., Sangrador-Vegas, A., Slattery, L., Fitzgerald, G.F., Beresford, T., Ross, R.P., 2008. Genome Sequence of *Lactobacillus helveticus*, an Organism Distinguished by Selective Gene Loss and Insertion Sequence Element Expansion. J. Bacteriol. 190, 727–735. doi:10.1128/JB.01295-07

Cavanagh, D., Fitzgerald, G.F., McAuliffe, O., 2015. From field to fermentation: The origins of *Lactococcus lactis* and its domestication to the dairy environment. Food Microbiol. 47, 45–61. doi:10.1016/j.fm.2014.11.001

Chaillou, S., Champomier-Vergès, M.-C., Cornet, M., Crutz-Le Coq, A.-M., Dudez, A.-M., Martin, V., Beaufils, S., Darbon-Rongère, E., Bossy, R., Loux, V., Zagorec, M., 2005. The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. Nat. Biotech. 23, 1527–1533. doi:10.1038/nbt1160

Chassy, B.M., Alpert, C.-A., 1989. Molecular characterization of the plasmid-encoded lactose-PTS of *Lactobacillus casei*. FEMS Microbiol. Lett. 63, 157–165. doi:10.1111/j.1574-6968.1989.tb14112.x

Chassy, B.M., Gibson, E.M., Guiffrida, A., 1978. Evidence for plasmid-associated lactose metabolism in *Lactobacillus casei* subsp. *casei*. Current Microbiol. 1, 141–144. doi:10.1007/BF02601666

Chassy, B.M., Thompson, J., 1983. Regulation and characterization of the galactose-phosphoenolpyruvate-dependent phosphotransferase system in *Lactobacillus casei*. J. Bacteriol. 154, 1204–1214.

Chen, Y.-Y.M., Betzenhauser, M.J., Snyder, J.A., Burne, R.A., 2002. Pathways for lactose/galactose catabolism by *Streptococcus salivarius*. FEMS Microbiol. Lett. 209, 75–79.

Chmelar, D., Matusek, A., Korger, J., Durnova, E., Steffen, M., Chmelarova, E., 2002. Isolation of *Carnobacterium piscicola* from human pus - Case report. Folia Microbiol. 47, 455–457. doi:10.1007/BF02818708

Christensen, J.E., Dudley, E.G., Pederson, J.A., Steele, J.L., 1999. Peptidases and amino acid catabolism in lactic acid bacteria. Antonie van Leeuwenhoek 76, 217–246.

Cocaign-Bousquet, M., Garrigues, C., Loubiere, P., Lindley, N.D., 1996. Physiology of pyruvate metabolism in *Lactococcus lactis*. Antonie Van Leeuwenhoek 70, 253–267.

Cogan, T.M., Beresford, T.P., Steele, J., Broadbent, J., Shah, N.P., Ustunol, Z., 2007. Invited review: Advances in starter cultures and cultured foods. J. Dairy Sci. 90, 4005–4021. doi:10.3168/jds.2006-765

Collins, M.D., Farrow, J. a. E., Phillips, B.A., Ferusu, S., Jones, D., 1987. Classification of *Lactobacillus divergens, Lactobacillus piscicola* , and Some Catalase-Negative, Asporogenous, Rod-Shaped Bacteria from Poultry in a New Genus, *Carnobacterium* . Int. J. Syst. Bacteriol. 37, 310–316. doi:10.1099/00207713-37-4-310

Collins, M.D., Farrow, J.A., Phillips, B.A., Kandler, O., 1983. *Streptococcus garvieae* sp. nov. and *Streptococcus plantarum* sp. nov. J. Gen. Microbiol. 129, 3427–3431.

Comfort, D., Clubb, R.T., 2004. A Comparative Genome Analysis Identifies Distinct Sorting Pathways in Gram-Positive Bacteria. Infect. Immun. 72, 2710–2722. doi:10.1128/IAI.72.5.2710-2722.2004

Connil, N., Dousset, X., Onno, B., Pilet, M.F., Breuil, M.F., Montel, M.C., 1998. Enumeration of *Carnobacterium divergens* V41, *Carnobacterium piscicola* V1 and *Lactobacillus brevis* LB62 by in situ hybridization-flow cytometry. Lett Appl Microbiol 27, 302–6.

Coombs, J., Brenchley, J.E., 2001. Characterization of Two New Glycosyl Hydrolases from the Lactic Acid Bacterium *Carnobacterium piscicola* Strain BA. Appl. Environ. Microbiol. 67, 5094–5099. doi:10.1128/AEM.67.11.5094-5099.2001

Coombs, J.M., Brenchley, J.E., 1999. Biochemical and phylogenetic analyses of a cold-active β-galactosidase from the lactic acid bacterium *Carnobacterium piscicola* BA. Appl. Environ. Microbiol. 65, 5443–5450.

Cotton, P.B., 1972. Non-dietary lipid in the intestinal lumen. Gut 13, 675–681.

Crow, V.L., Davey, G.P., Pearce, L.E., Thomas, T.D., 1983. Plasmid linkage of the D-tagatose 6-phosphate pathway in *Streptococcus lactis*: effect on lactose and galactose metabolism. J. Bacteriol. 153, 76–83.

Crow, V.L., Thomas, T.D., 1984. Properties of a *Streptococcus lactis* strain that ferments lactose slowly. J. Bacteriol. 157, 28–34.

Curiel, J.A., Ruiz-Capillas, C., de las Rivas, B., Carrascosa, A.V., Jimenez-Colmenero, F., Munoz, R., 2011. Production of biogenic amines by lactic acid bacteria and enterobacteria isolated from fresh pork sausages packaged in different atmospheres and kept under refrigeration. Meat Sci. 88, 368–373. doi:10.1016/j.meatsci.2011.01.011

De Bruyn, I.N., Holzapfel, W.H., Visser, L., Louw, A.I., 1988. Glucose metabolism by *Lactobacillus divergens*. J. Gen. Microbiol. 134, 2103–2109. doi:10.1099/00221287-134-8-2103

Declerck, N., Minh, N.L., Yang, Y., Bloch, V., Kochoyan, M., Aymerich, S., 2002. RNA recognition by transcriptional antiterminators of the *BglG/SacY* family: mapping of *SacY* RNA binding site. J. Mol. Biol. 319, 1035–1048. doi:10.1016/S0022-2836(02)00373-X

Delorme, C., Bartholini, C., Bolotine, A., Ehrlich, S.D., Renault, P., 2010. Emergence of a Cell Wall Protease in the *Streptococcus thermophilus* Population. Appl. Environ. Microbiol. 76, 451–460. doi:10.1128/AEM.01018-09

Desvaux, M., Dumas, E., Chafsey, I., Hébraud, M., 2006. Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. FEMS Microbiol. Lett. 256, 1–15. doi:10.1111/j.1574-6968.2006.00122.x

Desvaux, M., Hébraud, M., Talon, R., Henderson, I.R., 2009. Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. Trends Microbiol. 17, 139–145. doi:10.1016/j.tim.2009.01.004

Deutscher, J., Francke, C., Postma, P.W., 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol. Mol. Biol. Rev. 70, 939–+. doi:10.1128/MMBR.00024-06

de Vos, W.M., 2011. Systems solutions by lactic acid bacteria: from paradigms to practice. Microb Cell Fact 10, S2.

deVos, W.M., 1996. Metabolic engineering of sugar catabolism in lactic acid bacteria. Int. J. Gen. Mol. Microbiol. 70, 223–242.

de Vos, W.M., Hugenholtz, J., 2004. Engineering metabolic highways in *Lactococci* and other lactic acid bacteria. Trends Biotechnol. 22, 72–79. doi:10.1016/j.tibtech.2003.11.011

de Vos, W.M., Vaughan, E.E., 1994. Genetics of lactose utilization in lactic acid bacteria. FEMS Microbiol. Rev. 15, 217–237.

Díaz, O., Fernandez, M., De Fernando, G.D., de la Hoz, L., Ordoñez, J.A., 1997. Proteolysis in dry fermented sausages: The effect of selected exogenous proteases. Meat Sci. 46, 115–128.

Doeven, M.K., Kok, J., Poolman, B., 2005. Specificity and selectivity determinants of peptide transport in *Lactococcus lactis* and other microorganisms: Peptide transport in *Lactococcus lactis*. Mol. Microbiol. 57, 640–649. doi:10.1111/j.1365-2958.2005.04698.x

Douillard, F.P., de Vos, W.M., 2014. Functional genomics of lactic acid bacteria: from food to health. Microb. Cell Fact. 13, S8. doi:10.1186/1475-2859-13-S1-S8

Doyle, R.J., 2007. Glycomicrobiology. Springer Science & Business Media.

Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prévost, M.-C., Kojadinovic, M., Lalioui, L., Poyart, C., Trieu-Cuot, P., 2006. Assembly and role of pili in group B streptococci. Mol. Microbiol. 60, 1401–1413. doi:10.1111/j.1365-2958.2006.05190.x

Dramsi, S., Trieu-Cuot, P., Bierne, H., 2005. Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. Res. Microbiol. 156, 289–297. doi:10.1016/j.resmic.2004.10.011

Drider, D., Fimland, G., Hechard, Y., McMullen, L.M., Prevost, H., 2006. The continuing story of class IIa bacteriocins. Microbiol Mol Biol Rev 70, 564–82.

Driessen, A.J.M., Nouwen, N., 2008. Protein Translocation Across the Bacterial Cytoplasmic Membrane. Annual Review of Biochemistry 77, 643–667. doi:10.1146/annurev.biochem.77.061606.160747

Dryla, A., Gelbmann, D., Von Gabain, A., Nagy, E., 2003. Identification of a novel iron regulated staphylococcal surface protein with haptoglobin-haemoglobin binding activity. Mol. Microbiol. 49, 37–53. doi:10.1046/j.1365-2958.2003.03542.x

Duffes, F., Leroi, F., Boyaval, P., Dousset, X., 1999. Inhibition of *Listeria monocytogenes* by *Carnobacterium* spp. strains in a simulated cold smoked fish system stored at 4 degrees C. Int. J. Food Microbiol. 47, 33–42.

Eckert, C., Lecerf, M., Dubost, L., Arthur, M., Mesnage, S., 2006. Functional analysis of AtlA, the major N-acetylglucosaminidase of *Enterococcus faecalis*. J. Bacteriol. 188, 8513–8519. doi:10.1128/JB.01145-06

Edima, H.C., Cailliez-Grimal, C., Revol-Junelles, A.-M., Rondags, E., Milliere, J.-B., 2008. Short communication: Impact of pH and temperature on the acidifying activity of *Carnobacterium maltaromaticum*. J. Dairy Sci. 91, 3806–3813. doi:10.3168/jds.2007-0878

Edima, H.C., Cailliez-Grimal, C., Revol-Junelles, A.-M., Tonti, L., Linder, M., Milliere, J.-B., 2007. A selective enumeration medium for *Carnobacterium maltaromaticum*. J. Microbiol. Methods 68, 516–521. doi:10.1016/j.mimet.2006.10.006

Edwards, A.M., Manetti, A.G.O., Falugi, F., Zingaretti, C., Capo, S., Buccato, S., Bensi, G., Telford, J.L., Margarit, I., Grandi, G., 2008. Scavenger receptor gp340 aggregates group A streptococci by binding pili. Mol. Microbiol. 68, 1378–1394. doi:10.1111/j.1365-2958.2008.06220.x

Elasri, M.O., Thomas, J.R., Skinner, R.A., Blevins, J.S., Beenken, K.E., Nelson, C.L., Smeltzer, M.S., 2002. *Staphylococcus aureus* collagen adhesin contributes to the pathogenesis of osteomyelitis. Bone 30, 275–280.

Emborg, J., Laursen, B.G., Rathjen, T., Dalgaard, P., 2002. Microbial spoilage and formation of biogenic amines in fresh and thawed modified atmosphere-packed salmon (Salmo salar) at 2 degrees C. J Appl Microbiol 92, 790–9.

Engesser, D., Hammes, W., 1994. Nonheme Catalase Activity of Lactic-Acid Bacteria. Syst. Appl. Microbiol. 17, 11–19.

Ercolini, D., Ferrocino, I., Nasi, A., Ndagijimana, M., Vernocchi, P., La Storia, A., Laghi, L., Mauriello, G., Guerzoni, M.E., Villani, F., 2011. Monitoring of Microbial Metabolites and Bacterial Diversity in Beef Stored under Different Packaging Conditions. Appl. Environ. Microbiol. 77, 7372–7381. doi:10.1128/AEM.05521-11

Ercolini, D., Russo, F., Nasi, A., Ferranti, P., Villani, F., 2009. Mesophilic and psychrotrophic bacteria from meat and their spoilage potential in vitro and in beef. Appl Environ Microbiol.

Fadda, S., Vignolo, G., Aristoy, M.C., Oliver, G., Toldrá, F., 2001. Effect of curing conditions and *Lactobacillus casei* CRL705 on the hydrolysis of meat proteins. J. Appl. Microbiol. 91, 478–487.

Fälker, S., Nelson, A.L., Morfeldt, E., Jonas, K., Hultenby, K., Ries, J., Melefors, Ö., Normark, S., Henriques-Normark, B., 2008. Sortase-mediated assembly and surface topology of adhesive pneumococcal pili. Mol. Microbiol. 70, 595–607. doi:10.1111/j.1365-2958.2008.06396.x

Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. 186, 1518–1530.

Feligini, M., Panelli, S., Buffoni, J.N., Bonacina, C., Andrighetto, C., Lombardi, A., 2012. Identification of Microbiota Present on the Surface of Taleggio Cheese Using PCR-DGGE and RAPD-PCR. J. Food Sci. 77, M609–M615. doi:10.1111/j.1750-3841.2012.02932.x

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., Bateman, A., 2010. The Pfam protein families database. Nucleic Acids Res 38, D211–D222. doi:10.1093/nar/gkp985

Fischetti, V.., Pancholi, V., Schneewind, O., 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from Gram-positive cocci. Mol. Microbiol. 4, 1603–1605. doi:10.1111/j.1365-2958.1990.tb02072.x

Flint, H.J., Scott, K.P., Duncan, S.H., Louis, P., Forano, E., 2012. Microbial degradation of complex carbohydrates in the gut. Gut Microb. 3, 289–306. doi:10.4161/gmic.19897

Flórez, A.B., Mayo, B., 2015. The Plasmid Complement of the Cheese Isolate *Lactococcus garvieae* IPLA 31405 Revealed Adaptation to the Dairy Environment. PLoS ONE 10, e0126101. doi:10.1371/journal.pone.0126101

Forrest, L.R., Krämer, R., Ziegler, C., 2011. The structural basis of secondary active transport mechanisms. Biochim. Biophys. Acta 1807, 167–188. doi:10.1016/j.bbabio.2010.10.014

Fortina, M.G., Ricci, G., Borgo, F., 2009. A Study of Lactose Metabolism in *Lactococcus garvieae* Reveals a Genetic Marker for Distinguishing between Dairy and Fish Biotypes. J. Food Prot. 72, 1248–1254.

Fortina, M.G., Ricci, G., Mora, D., Guglielmetti, S., Manachini, P.L., 2003. Unusual Organization for Lactose and Galactose Gene Clusters in *Lactobacillus helveticus*. Appl. Environ. Microbiol. 69, 3238–3243. doi:10.1128/AEM.69.6.3238-3243.2003

Foucaud, C., Kunji, E.R., Hagting, A., Richard, J., Konings, W.N., Desmazeaud, M., Poolman, B., 1995. Specificity of peptide transport systems in *Lactococcus lactis*: evidence for a third system which transports hydrophobic di- and tripeptides. J. Bacteriol. 177, 4652–4657.

Franzmann, P.D., Hopfl, P., Weiss, N., Tindall, B.J., 1991. Psychrotrophic, lactic acid-producing bacteria from anoxic waters in Ace Lake, Antarctica; *Carnobacterium funditum* sp. nov. and </i>Carnobacterium alterfunditum</i> sp. nov. Arch. Microbiol. 156, 255–62.

Frey, P.A., 1996. The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. FASEB J 10, 461–470.

Frirdich, E., Gaynor, E.C., 2013. Peptidoglycan hydrolases, bacterial shape, and pathogenesis. Curr. Opin. Microbiol. 16, 767–778. doi:10.1016/j.mib.2013.09.005

Fronzes, R., Remaut, H., Waksman, G., 2008. Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. EMBO J. 27, 2271–2280. doi:10.1038/emboj.2008.155

Fujita, Y., 2009. Carbon catabolite control of the metabolic network in *Bacillus subtilis*. Biosci. Biotechnol. Biochem. 73, 245–259. doi:10.1271/bbb.80479

Gänzle, M.G., Vermeulen, N., Vogel, R.F., 2007. Carbohydrate, peptide and lipid metabolism of lactic acid bacteria in sourdough. Food Microbiology 24, 128–138. doi:10.1016/j.fm.2006.07.006

Garault, P., Bars, D.L., Besset, C., Monnet, V., 2002. Three Oligopeptide-binding Proteins Are Involved in the Oligopeptide Transport of *Streptococcus thermophilus*. J. Biol. Chem. 277, 32–39. doi:10.1074/jbc.M107002200

Gaspar, A.H., Ton-That, H., 2006. Assembly of Distinct Pilus Structures on the Surface of *Corynebacterium diphtheriae*. J. Bacteriol. 188, 1526–1533. doi:10.1128/JB.188.4.1526-1533.2006

Gasson, M.J., 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. 154, 1–9.

Gaudin, C.F.M., Grigg, J.C., Arrieta, A.L., Murphy, M.E.P., 2011. Unique Heme-Iron Coordination by the Hemoglobin Receptor IsdB of *Staphylococcus aureus*. Biochem. 50, 5443–5452. doi:10.1021/bi200369p

Georges, C., François-Marie, L., 2008. Bactéries lactiques. De la génétique aux ferments. Lavoisier.

Ghigo, J.M., Létoffé, S., Wandersman, C., 1997. A new type of hemophore-dependent heme acquisition system of Serratia marcescens reconstituted in *Escherichia coli*. J. Bacteriol. 179, 3572–3579.

Gilliland, S.E., 1989. Acidophilus Milk Products: A Review of Potential Benefits to Consumers. J. Dairy Sci. 72, 2483–2494. doi:10.3168/jds.S0022-0302(89)79389-9

Gosalbes, M.J., Monedero, V., Alpert, C.-A., Pérez-Martínez, G., 1997. Establishing a model to study the regulation of the lactose operon in *Lactobacillus casei*. FEMS Microbiol. Lett. 148, 83–89. doi:10.1016/S0378-1097(97)00017-7

Griffin, H., Gasson, M., 1994. The Gene (*lacA*) Encoding Galactoside Acetyltransferase from *Lactococcus lactis*. Biotechnol. Lett. 16, 1125–1130. doi:10.1007/BF01020837

Griffiths, M.W., Tellez, A.M., 2013. *Lactobacillus helveticus*: the proteolytic system. Front. Microbiol. 4. doi:10.3389/fmicb.2013.00030

Grossiord, B.P., Luesink, E.J., Vaughan, E.E., Arnaud, A., de Vos, W.M., 2003. Characterization, expression, and mutation of the *Lactococcus lactis galPMKTE* genes, involved in galactose utilization via the Leloir pathway. J. Bacteriol. 185, 870–878. doi:10.1128/JB.185.3.870-878.2003

Grossiord, B., Vaughan, E.E., Luesink, E., de Vos, W.M., 1998. Genetics of galactose utilisation via the Leloir pathway in lactic acid bacteria. Lait 78, 77–84.

Groudieva, T., Kambourova, M., Yusef, H., Royter, M., Grote, R., Trinks, H., Antranikian, G., 2004. Diversity and cold-active hydrolytic enzymes of culturable bacteria associated with Arctic sea ice, Spitzbergen. Extremophiles 8, 475–488. doi:10.1007/s00792-004-0409-0

Guédon, E., Renault, P., Ehrlich, S.D., Delorme, C., 2001. Transcriptional Pattern of Genes Coding for the Proteolytic System of *Lactococcus lactis* and Evidence for Coordinated Regulation of Key Enzymes by Peptide Supply. J. Bacteriol. 183, 3614–3622. doi:10.1128/JB.183.12.3614-3622.2001

Hagting, A., Kunji, E.R., Leenhouts, K.J., Poolman, B., Konings, W.N., 1994. The di- and tripeptide transport protein of *Lactococcus lactis*. A new type of bacterial peptide transporter. J. Biol. Chem. 269, 11391–11399.

Haley, K.P., Janson, E.M., Heilbronner, S., Foster, T.J., Skaar, E.P., 2011. *Staphylococcus lugdunensis* IsdG Liberates Iron from Host Heme. J. Bacteriol. 193, 4749–4757. doi:10.1128/JB.00436-11

Hammer, N.D., Skaar, E.P., 2011. Molecular mechanisms of Staphylococcus aureus iron acquisition. Annu. Rev. Microbiol. 65. doi:10.1146/annurev-micro-090110-102851

Hendrickx, A.P.A., Bonten, M.J.M., van Luit-Asbroek, M., Schapendonk, C.M.E., Kragten, A.H.M., Willems, R.J.L., 2008. Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate. Microbiol. 154, 3212–3223. doi:10.1099/mic.0.2008/020891-0

Hendrickx, A.P.A., Budzik, J.M., Oh, S.-Y., Schneewind, O., 2011. Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. Nat. Rev. Microbiol. 9, 166–176. doi:10.1038/nrmicro2520

Hengst, C.D. den, Hijum, S.A.F.T. van, Geurts, J.M.W., Nauta, A., Kok, J., Kuipers, O.P., 2005. The *Lactococcus lactis CodY* Regulon IDENTIFICATION OF A CONSERVED cis-REGULATORY ELEMENT. J. Biol. Chem. 280, 34332–34342. doi:10.1074/jbc.M502349200

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309–316.

Henrissat, B., Bairoch, A., 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. j 293, 781–788.

Hirschhausen, N., Schlesier, T., Peters, G., Heilmann, C., 2012. Characterization of the Modular Design of the Autolysin/Adhesin Aaa from *Staphylococcus aureus*. PLoS ONE 7, e40353. doi:10.1371/journal.pone.0040353

Hiu, S.F., Holt, R.A., Sriranganathan, N., Seidler, R.J., Fryer, J.L., 1984. *Lactobacillus piscicola*, a New Species from Salmonid Fish. Int. J. Syst. Bacteriol. 34, 393–400. doi:10.1099/00207713-34-4-393

Hoenigl, M., Grisold, A.J., Valentin, T., Leitner, E., Zarfel, G., Renner, H., Krause, R., 2010. Isolation of *Carnobacterium* sp from a human blood culture. J. Med. Microbiol. 59, 493–495. doi:10.1099/jmm.0.016808-0

Holley, R.A., Guan, T.Y., Peirson, M., Yost, C.K., 2002. *Carnobacterium viridans* sp. nov., an alkaliphilic, facultative anaerobe isolated from refrigerated, vacuum-packed bologna sausage. Int. J. Syst. Evol. Microbiol. 52, 1881–5.

Hols, P., Hancy, F., Fontaine, L., Grossiord, B., Prozzi, D., Leblond-Bourget, N., Decaris, B., Bolotin, A., Delorme, C., Dusko Ehrlich, S., Guédon, E., Monnet, V., Renault, P., Kleerebezem, M., 2005. New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. FEMS Microbiol. Rev. 29, 435–463. doi:10.1016/j.femsre.2005.04.008

Holzapfel, W.H., 1992. Culture media for non-sporulating gram-positive food spoilage bacteria. Int J Food Microbiol 17, 113–33.

Holzapfel, W.H., Gerber, E.S., 1983. *Lactobacillus divergens* sp. nov., a New Heterofermentative *Lactobacillus* Species Producing L(+)-Lactate. Syst. Appl. Microbiol. 4, 522–534. doi:10.1016/S0723-2020(83)80010-1

Hone, D.M., Attridge, S.R., Forrest, B., Morona, R., Daniels, D., LaBrooy, J.T., Bartholomeusz, R.C., Shearman, D.J., Hackett, J., 1988. A *galE* via (Vi antigen-negative) mutant of *Salmonella typhi* Ty2 retains virulence in humans. Infect. Immun. 56, 1326–1333.

Hong-yan, Z., Jie, L., Jing-jing, L., Yu-cai, L., Xiao-fen, W., Zong-jun, C., 2013. Microbial Community Dynamics During Biogas Slurry and Cow Manure Compost. J. Integr. Agric. 12, 1087–1097. doi:10.1016/S2095-3119(13)60328-7

Honsa, E.S., Maresso, A.W., 2011. Mechanisms of iron import in anthrax. Biometals 24, 533–545. doi:10.1007/s10534-011-9413-x

Horvath, P., Coûté-Monvoisin, A.-C., Romero, D.A., Boyaval, P., Fremaux, C., Barrangou, R., 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. International Journal of Food Microbiology, 15th Meeting of the Club des Bactéries Lactiques 131, 62–70. doi:10.1016/j.ijfoodmicro.2008.05.030

Houng, H.S., Kopecko, D.J., Baron, L.S., 1990. Molecular cloning and physical and functional characterization of the *Salmonella typhimurium* and *Salmonella typhi* galactose utilization operons. J. Bacteriol. 172, 4392–4398.

Hutkins, R., Morris, H.A., McKay, L.L., 1985. Galactokinase activity in *Streptococcus thermophilus*. Appl. Environ. Microbiol. 50, 777–780.

Jaffres, E., Sohier, D., Leroi, F., Pilet, M.F., Prevost, H., Joffraud, J.J., Dousset, X., 2009. Study of the bacterial ecosystem in tropical cooked and peeled shrimps using a polyphasic approach. Int. J. Food Microbiol. 131, 20–29. doi:10.1016/j.ijfoodmicro.2008.05.017

Jagusztyn-Krynicka, E.K., Hansen, J.B., Crow, V.L., Thomas, T.D., Honeyman, A.L., Curtiss, R., 1992. *Streptococcus mutans* serotype c tagatose 6-phosphate pathway gene cluster. J. Bacteriol. 174, 6152–6158.

Jain, S., Choudhary, D.K., 2014. Induced defense-related proteins in soybean (Glycine max *L. Merrill*) plants by *Carnobacterium* sp. SJ-5 upon challenge inoculation of *Fusarium oxysporum*. Planta 239, 1027–1040. doi:10.1007/s00425-014-2032-3

Jeong, J.K., Kwon, O., Lee, Y.M., Oh, D.-B., Lee, J.M., Kim, S., Kim, E.-H., Le, T.N., Rhee, D.-K., Kang, H.A., 2009. Characterization of the *Streptococcus pneumoniae* BgaC Protein as a Novel Surface -Galactosidase with Specific Hydrolysis Activity for the Gal 1-3GlcNAc Moiety of Oligosaccharides. J. Bacteriol. 191, 3011–3023. doi:10.1128/JB.01601-08

Joborn, A., Dorsch, M., Olsson, J.C., Westerdahl, A., Kjelleberg, S., 1999. *Carnobacterium inhibens* sp. nov., isolated from the intestine of Atlantic salmon (Salmo salar). Int. J. Syst. Bacteriol. 49 Pt 4, 1891–8.

Joffraud, J.J., Leroi, F., Roy, C., Berdague, J.L., 2001. Characterisation of volatile compounds produced by bacteria isolated from the spoilage flora of cold-smoked salmon. Int. J. Food Microbiol. 66, 175–84.

Johnson, B.R., Klaenhammer, T.R., 2014. Impact of genomics on the field of probiotic research: historical perspectives to modern paradigms. Antonie Van Leeuwenhoek 106, 141–156. doi:10.1007/s10482-014-0171-y

Jones, R.J., 2004. Observations on the succession dynamics of lactic acid bacteria populations in chill-stored vacuum-packaged beef. Int. J. Food Microbiol. 90, 273–282. doi:10.1016/S0168-1605(03)00310-6

Jonquières, R., Bierne, H., Fiedler, F., Gounon, P., Cossart, P., 1999. Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of gram-positive bacteria. Mol. Microbiol. 34, 902–914.

Kafsi, H.E., Binesse, J., Loux, V., Buratti, J., Boudebbouze, S., Dervyn, R., Hammani, A., Maguin, E., Guchte, M. van de, 2014. Genome Sequence of *Lactobacillus delbrueckii* subsp. *lactis* CNRZ327, a Dairy Bacterium with Anti-Inflammatory Properties. Genome Announc. 2, e00328–14. doi:10.1128/genomeA.00328-14

Kandler, O., 1983. Carbohydrate metabolism in lactic acid bacteria. Antonie van Leeuwenhoek 49, 209–224. doi:10.1007/BF00399499

Kang, H.J., Coulibaly, F., Clow, F., Proft, T., Baker, E.N., 2007. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science 318, 1625–1628. doi:10.1126/science.1145806

Kang, H.J., Coulibaly, F., Proft, T., Baker, E.N., 2011. Crystal Structure of Spy0129, a *Streptococcus pyogenes* Class B Sortase Involved in Pilus Assembly. PLoS One 6. doi:10.1371/journal.pone.0015969

Kang, H.J., Paterson, N.G., Gaspar, A.H., Ton-That, H., Baker, E.N., 2009. The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. Proc. Natl. Acad. Sci. U. S. A. 106, 16967–16971. doi:10.1073/pnas.0906826106

Kankainen, M., Paulin, L., Tynkkynen, S., von Ossowski, I., Reunanen, J., Partanen, P., Satokari, R., Vesterlund, S., Hendrickx, A.P.A., Lebeer, S., De Keersmaecker, S.C.J., Vanderleyden, J., Hämäläinen, T., Laukkanen, S., Salovuori, N., Ritari, J., Alatalo, E., Korpela, R., Mattila-Sandholm, T., Lassig, A., Hatakka, K., Kinnunen, K.T., Karjalainen, H., Saxelin, M., Laakso, K., Surakka, A., Palva, A., Salusjärvi, T., Auvinen, P., de Vos, W.M., 2009. Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. Proc. Natl. Acad. Sci. U.S.A. 106, 17193–17198. doi:10.1073/pnas.0908876106

Kelly, W.J., Ward, L.J.H., Leahy, S.C., 2010. Chromosomal Diversity in *Lactococcus lactis* and the Origin of Dairy Starter Cultures. Genome Biol Evol 2, 729–744. doi:10.1093/gbe/evq056

Kim, D.-H., Austin, B., 2008. Characterization of probiotic carnobacteria isolated from rainbow trout (*Oncorhynchus mykiss*) intestine. Lett. Appl. Microbiol. 47, 141–147. doi:10.1111/j.1472-765X.2008.02401.x

Kim, H.M., Chae, N., Jung, J.Y., Lee, Y.K., 2013. Isolation of facultatively anaerobic soil bacteria from Ny-lesund, Svalbard. Polar Biol. 36, 787–796. doi:10.1007/s00300-013-1302-z

Kim, M.-S., Roh, S.W., Nam, Y.-D., Yoon, J.-H., Bae, J.-W., 2009. *Carnobacterium jeotgali* sp nov., isolated from a Korean traditional fermented food. Int. J. Syst. Evol. Microbiol. 59, 3168–3171. doi:10.1099/ijs.0.010116-0

Klaenhammer, T.R., Barrangou, R., Buck, B.L., Azcarate-Peril, M.A., Altermann, E., 2005. Genomic features of lactic acid bacteria effecting bioprocessing and health. FEMS Microbiol. Rev. 29, 393–409. doi:10.1016/j.fmrre.2005.04.007

Kleerebezem, M., Boekhorst, J., Kranenburg, R. van, Molenaar, D., Kuipers, O.P., Leer, R., Tarchini, R., Peters, S.A., Sandbrink, H.M., Fiers, M.W.E.J., Stiekema, W., Lankhorst, R.M.K., Bron, P.A., Hoffer, S.M., Groot, M.N.N., Kerkhoven, R., Vries, M. de, Ursing, B., Vos, W.M. de, Siezen, R.J., 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc. Natl. Acad. Sci. U. S. A. 100, 1990–1995. doi:10.1073/pnas.0337704100

Kleerebezem, M., Hols, P., Bernard, E., Rolain, T., Zhou, M., Siezen, R.J., Bron, P.A., 2010. The extracellular biology of the Lactobacilli. FEMS Microbiol. Rev. 34, 199–230. doi:10.1111/j.1574-6976.2009.00208.x

Kleerebezem, M., Hugenholtz, J., 2003. Metabolic pathway engineering in lactic acid bacteria. Curr. Opin. Biotechnol. 14, 232–237. doi:10.1016/S0958-1669(03)00033-B

Kline, K.A., Dodson, K.W., Caparon, M.G., Hultgren, S.J., 2010. A tale of two pili: assembly and function of pili in bacteria. Trends Microbiol. 18, 224–232. doi:10.1016/j.tim.2010.03.002

Kline, K.A., Kau, A.L., Chen, S.L., Lim, A., Pinkner, J.S., Rosch, J., Nallapareddy, S.R., Murray, B.E., Henriques-Normark, B., Beatty, W., Caparon, M.G., Hultgren, S.J., 2009. Mechanism for Sortase Localization and the Role of Sortase Localization in Efficient Pilus Assembly in *Enterococcus faecalis*. J. Bacteriol. 191, 3237–3247. doi:10.1128/JB.01837-08

Kok, J., Leenhouts, K.J., Haandrikman, A.J., Ledeboer, A.M., Venema, G., 1988. Nucleotide sequence of the cell wall proteinase gene of *Streptococcus cremoris* Wg2. Appl. Environ. Microbiol. 54, 231–238.

Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Duménil, G., Caliot, E., Trieu-Cuot, P., Dramsi, S., 2009. Dual Role for Pilus in Adherence to Epithelial Cells and Biofilm Formation in *Streptococcus agalactiae*. PLoS Pathog. 5. doi:10.1371/journal.ppat.1000422

Krispin, O., Allmansberger, R., 1998. The *Bacillus subtilis galE* gene is essential in the presence of glucose and galactose. J. Bacteriol. 180, 2265–2270.

Kunji, E.R., Mierau, I., Hagting, A., Poolman, B., Konings, W.N., 1996. The proteolytic systems of lactic acid bacteria. Antonie Van Leeuwenhoek 70, 187–221.

Lauro, F.M., Chastain, R.A., Blankenship, L.E., Yayanos, A.A., Bartlett, D.H., 2007. The unique 16S rRNA genes of piezophiles reflect both phylogeny and adaptation. Appl. Environ. Microbiol. 73, 838–845. doi:10.1128/AEM.01726-06

Laursen, B.G., Bay, L., Cleenwerck, I., Vancanneyt, M., Swings, J., Dalgaard, P., Leisner, J.J., 2005. *Carnobacterium divergens* and *Carnobacterium maltaromaticum* as spoilers or protective cultures in meat and seafood: phenotypic and genotypic characterization. Syst. Appl. Microbiol. 28, 151–164. doi:10.1016/j.syapm.2004.12.001

Laursen, B.G., Leisner, J.J., Dalgaard, P., 2006. *Carnobacterium* species: Effect of metabolic activity and interaction with *Brochothrix thermosphacta* on sensory characteristics of modified atmosphere packed shrimp. J. Agric. Food Chem. 54, 3604–3611. doi:10.1021/jf053017f

Lechardeur, D., Cesselin, B., Fernandez, A., Lamberet, G., Garrigues, C., Pedersen, M., Gaudu, P., Gruss, A., 2011. Using heme as an energy boost for lactic acid bacteria. Curr. Opin. Biotechnol. 22, 143–149. doi:10.1016/j.copbio.2010.12.001

Lee, P.A., Tullman-Ercek, D., Georgiou, G., 2006. The Bacterial Twin-Arginine Translocation Pathway. Annu. Rev. Microbiol. 60, 373–395. doi:10.1146/annurev.micro.60.080805.142212

Leisner, J.J., Hansen, M.A., Larsen, M.H., Hansen, L., Ingmer, H., Sørensen, S.J., 2012. The genome sequence of the lactic acid bacterium, *Carnobacterium maltaromaticum* ATCC 35586 encodes potential virulence factors. Int. J. Food Microbiol. 152, 107–115. doi:10.1016/j.ijfoodmicro.2011.05.012

Leisner, J.J., Laursen, B.G., Prévost, H., Drider, D., Dalgaard, P., 2007. *Carnobacterium*: positive and negative effects in the environment and in foods. FEMS Microbiol. Rev. 31, 592–613. doi:10.1111/j.1574-6976.2007.00080.x

LeMieux, J., Woody, S., Camilli, A., 2008. Roles of the Sortases of *Streptococcus pneumoniae* in Assembly of the RlrA Pilus. J. Bacteriol. 190, 6002–6013. doi:10.1128/JB.00379-08

Leonard, M.T., Panayotova, N., Farmerie, W.G., Triplett, E.W., Nicholson, W.L., 2013. Complete Genome Sequence of *Carnobacterium gilichinskyi* Strain WN1359T (DSM 27470T). Genome Announc. 1, e00985–13–e00985–13. doi:10.1128/genomeA.00985-13

Leong-Morgenthaler, P., Zwahlen, M., Hottinger, H., 1991. Lactose Metabolism in *Lactobacillus bulgaricus* - Analysis of the Primary Structure and Expression of the Genes Involved. J. Bacteriol. 173, 1951–1957.

Leroi, F., Arbey, N., Joffraud, J.J., Chevalier, F., 1996. Effect of inoculation with lactic acid bacteria on extending the shelf-life of vacuum-packed cold smoked salmon. International journal of food microbiology 31, 497–504.

Leroi, F., Joffraud, J.-J., Chevalier, F., Cardinal, M., 1998. Study of the microbial ecology of cold-smoked salmon during storage at 8°C. Int. J. Food Microbiol. 39, 111–121. doi:10.1016/S0168-1605(97)00126-8

Leroy, F., De Vuyst, L., 2004. Lactic acid bacteria as functional starter cultures for the food fermentation industry. Trends Food Sci. Tech. 15, 67–78. doi:10.1016/j.tifs.2003.09.004

Liu, S.-Q., 2003. Practical implications of lactate and pyruvate metabolism by lactic acid bacteria in food and beverage fermentations. Int. J. Food Microbiol. 83, 115–131. doi:10.1016/S0168-1605(02)00366-5

Li, Z.-Y., Liu, Y., 2006. Marine sponge *Craniella austrialiensis*-associated bacterial diversity revelation based on 16S rDNA library and biologically active *Actinomycetes* screening, phylogenetic analysis. Lett. Appl. Microbiol. 43, 410–416. doi:10.1111/j.1472-765X.2006.01976.x

Loch, T.P., Scribner, K., Tempelman, R., Whelan, G., Faisal, M., 2012. Bacterial infections of Chinook salmon, *Oncorhynchus tshawytscha* (Walbaum), returning to gamete collecting weirs in Michigan. J. Fish Dis. 35, 39–50. doi:10.1111/j.1365-2761.2011.01322.x

Looijesteijn, P.J., Trapet, L., de Vries, E., Abee, T., Hugenholtz, J., 2001. Physiological function of exopolysaccharides produced by *Lactococcus lactis*. Int. J. Food Microbiol. 64, 71–80.

Loperena, L., Soria, V., Varela, H., Lupo, S., Bergalli, A., Guigou, M., Pellegrino, A., Bernardo, A., Calvino, A., Rivas, F., Batista, S., 2012. Extracellular enzymes produced by microorganisms isolated from maritime Antarctica. World J. Microbiol. Biotechnol. 28, 2249–2256. doi:10.1007/s11274-012-1032-3

Loughman, J.A., Caparon, M.G., 2007. Comparative functional analysis of the *lac* operons in *Streptococcus pyogenes*. Mol. Microbiol. 64, 269–280. doi:10.1111/j.1365-2958.2007.05663.x

Ludwig, W., Schleifer, K.-H., Whitman, W.B., 2009. Family III. *Carnobacteriaceae* fam. nov, in: Bergey's Manual of Systematic Bacteriology. P. De Vos, G. M. Garrity, D. Jones, N. R. Krieg, W. Ludwig, F. A. Rainey, K.-H. Schleifer & W. B. Whitman., New York: Springer, p. 549.

Luirink, J., Heijne, G. von, Houben, E., Gier, J.-W. de, 2005. Biogenesis of Inner Membrane Proteins in *Escherichia Coli*. Annu. Rev. Microbiol. 59, 329–355. doi:10.1146/annurev.micro.59.030804.121246

Lukić, J., Strahinić, I., Jovčić, B., Filipić, B., Topisirović, L., Kojić, M., Begović, J., 2012. Different Roles for Lactococcal Aggregation Factor and Mucin Binding Protein in Adhesion to Gastrointestinal Mucosa. Appl. Environ. Microbiol. 78, 7993–8000. doi:10.1128/AEM.02141-12

Mace, S., Cornet, J., Chevalier, F., Cardinal, M., Pilet, M.-F., Dousset, X., Joffraud, J.-J., 2012. Characterisation of the spoilage microbiota in raw salmon (Salmo salar) steaks stored under vacuum or modified atmosphere packaging combining conventional methods and PCR-TTGE. Food Microbiol. 30, 164–172. doi:10.1016/j.fm.2011.10.013

Macklaim, J.M., Gloor, G.B., Anukam, K.C., Cribby, S., Reid, G., 2011. At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners* AB-1. Proc. Natl. Acad. Sci. U. S. A. 108, 4688–4695. doi:10.1073/pnas.1000086107

Makarova, K.S., Koonin, E.V., 2007. Evolutionary Genomics of Lactic Acid Bacteria. J. Bacteriol. 189, 1199–1208. doi:10.1128/JB.01351-06

Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., others, 2006. Comparative genomics of the lactic acid bacteria. PNAS 103, 15611–15616.

Malki, I., Amorim, G.C. de, Simenel, C., Prochnicka-Chalufour, A., Delepierre, M., Izadi-Pruneyre, N., 2012. 1H, 13C and 15N resonance assignments of the periplasmic signalling domain of HasR, a TonB-dependent outer membrane heme transporter. Biomol. NMR Assign. 7, 43–46. doi:10.1007/s12104-012-9377-y

Malkin, J., Kimmitt, P.T., Ou, H.-Y., Bhasker, P.S.S., Khare, M., Deng, Z., Stephenson, I., Sosnowski, A.W., Perera, N., Rajakumar, K., 2008. Identification of *Streptococcus gallolyticus* subsp. *macedonicus* as the etiological agent in a case of culture-negative multivalve infective endocarditis by 16S rDNA PCR analysis of resected valvular tissue. J. Heart Valve Dis. 17, 589–592.

Mandlik, A., Das, A., Ton-That, H., 2008a. The molecular switch that activates the cell wall anchoring step of pilus assembly in gram-positive bacteria. PNAS 105, 14147–14152.

Mandlik, A., Swierczynski, A., Das, A., Ton-That, H., 2008b. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. Trends Microbiol. 16, 33–40. doi:10.1016/j.tim.2007.10.010

Maresso, A.W., Chapa, T.J., Schneewind, O., 2006. Surface Protein IsdC and Sortase B Are Required for Heme-Iron Scavenging of *Bacillus anthracis*. J. Bacteriol. 188, 8145–8152. doi:10.1128/JB.01011-06

Maresso, A.W., Schneewind, O., 2008. Sortase as a Target of Anti-Infective Therapy. Pharmacol Rev 60, 128–141. doi:10.1124/pr.107.07110

Marraffini, L.A., DeDent, A.C., Schneewind, O., 2006. Sortases and the Art of Anchoring Proteins to the Envelopes of Gram-Positive Bacteria. Microbiol. Mol. Biol. Rev. 70, 192–221. doi:10.1128/MMBR.70.1.192-221.2006

Masson, F., Johansson, G., Montel, M.C., 1999. Tyramine production by a strain of *Carnobacterium divergens* inoculated in meat-fat mixture. Meat Sci. 52, 65–69. doi:10.1016/S0309-1740(98)00149-1

Masson, F., Talon, R., Montel, M.C., 1996. Histamine and tyramine production by bacteria from meat products. Int. J. Food Microbiol. 32, 199–207.

Matamoros, S., Pilet, M.F., Gigout, F., Prevost, H., Leroi, F., 2009. Selection and evaluation of seafood-borne psychrotrophic lactic acid bacteria as inhibitors of pathogenic and spoilage bacteria. Food Microbiol. 26, 638–44.

Mathieu, F., Michel, M., Lebrihi, A., Lefebvre, G., 1994. Effect of the bacteriocin carnocin CP5 and of the producing strain Carnobacterium piscicola CP5 on the viability of Listeria monocytogenes ATCC 15313 in salt solution, broth and skimmed milk, at various incubation temperatures. Int. J. Food Microbiol. 22, 155–172.

Maxwell, E., Kurahashi, K., Kalckar, H., 1962. Enzymes of the Leloir Pathway. Methods Enzymol. 5, 174–189. doi:10.1016/S0076-6879(62)05204-0

Mazmanian, S.K., Liu, G., Ton-That, H., Schneewind, O., 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. Science 285, 760–763.

Mazmanian, S.K., Ton-That, H., Su, K., Schneewind, O., 2002. An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. Proc. Natl. Acad. Sci. U.S.A. 99, 2293–2298. doi:10.1073/pnas.032523999

McSweeney, P.L.H., 2004. Biochemistry of cheese ripening. Int. J. Dairy Technol. 57, 127–144. doi:10.1111/j.1471-0307.2004.00147.x

Meisel, J., Wolf, G., Hammes, W.P., 1994. Heme-dependent cytochrome formation in *Lactobacillus maltaromicus*. System. Appl. Microb. 17, 20–23.

Mejlholm, O., Kjeldgaard, J., Modberg, A., Vest, M.B., Boknaes, N., Koort, J., Bjorkroth, J., Dalgaard, P., 2008. Microbial changes and growth of *Listeria monocytogenes* during chilled storage of brined shrimp (Pandalus borealis). Int. J. Food Microbiol. 124, 250–259. doi:10.1016/j.ijfoodmicro.2008.03.022

Michel, C., Faivre, B., Kerouault, B., 1986. Biochemical-Identification of *Lactobacillus-Piscicola* Strains from France and Belgium. Dis. Aquat. Org. 2, 27–30. doi:10.3354/dao002027

Miller, A., Morgan, M.E., Libbey, L.M., 1974. *Lactobacillus maltaromicus*, a New Species Producing a Malty Aroma. Int. J. Syst. Bacteriol. 24, 346–354. doi:10.1099/00207713-24-3-346

Milliere, J.B., Lefebvre, G., 1994. *Carnobacterium piscicola*, a common species of French soft cheeses from cow's raw milk. Neth. Milk Dairy J-NE 48, 19–30.

Millière, J. b., Michel, M., Mathieu, F., Lefebvre, G., 1994. Presence of *Carnobacterium* spp. in French surface mould-ripened soft-cheese. J. Appl. Bacteriol. 76, 264–269. doi:10.1111/j.1365-2672.1994.tb01626.x

Milohanic, E., Jonquières, R., Cossart, P., Berche, P., Gaillard, J.-L., 2001. The autolysin Ami contributes to the adhesion of *Listeria monocytogenes* to eukaryotic cells via its cell wall anchor. Mol. Microbiol. 39, 1212–1224. doi:10.1111/j.1365-2958.2001.02208.x

Mollet, B., Pilloud, N., 1991. Galactose utilization in *Lactobacillus helveticus*: isolation and characterization of the galactokinase (*galK*) and galactose-1-phosphate uridyl transferase (*galT*) genes. J. Bacteriol. 173, 4464–4473.

Mora, D., Scarpellini, M., Franzetti, L., Colombo, S., Galli, A., 2003. Reclassification of *Lactobacillus maltaromicus* (Miller et al 1974) DSM 20342 (T) and DSM 20344 and

Carnobacterium piscicola (Collins et al 1987) DSM 20730(T) and DSM 20722 as *Carnobacterium maltaromaticum* comb. nov. Int. J. Syst. Evol. Microbiol. 53, 675–678. doi:10.1099/ijs.0.02405-0

Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G.O., Maggi, T., Taddei, A.R., Grandi, G., Telford, J.L., 2005. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc. Natl. Acad. Sci. U.S.A. 102, 15641–15646. doi:10.1073/pnas.0507808102

Morea, M., Baruzzi, F., Cocconcelli, P.S., 1999. Molecular and physiological characterization of dominant bacterial populations in traditional mozzarella cheese processing. J. Appl. Microbiol. 87, 574–82.

Morse, M.L., Hill, K.L., Egan, J.B., Hengstenberg, W., 1968. Metabolism of Lactose by *Staphylococcus aureus* and Its Genetic Basis. J. Bacteriol. 95, 2270–2274.

Nallapareddy, S.R., Singh, K.V., Sillanpaa, J., Garsin, D.A., Hook, M., Erlandsen, S.L., Murray, B.E., 2006. Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. J. Clin. Invest. 116, 2799–2807. doi:10.1172/JCI29021

Navarre, W.W., Schneewind, O., 1999. Surface Proteins of Gram-Positive Bacteria and Mechanisms of Their Targeting to the Cell Wall Envelope. Microbiol. Mol. Biol. Rev. 63, 174–229.

Neves, A.R., Pool, W.A., Kok, J., Kuipers, O.P., Santos, H., 2005. Overview on sugar metabolism and its control in *Lactococcus lactis* - The input from in vivo NMR. Fems Microbiol. Rev. 29, 531–554. doi:10.1016/j.femsre.2005.04.005

Neves, A.R., Pool, W.A., Solopova, A., Kok, J., Santos, H., Kuipers, O.P., 2010. Towards Enhanced Galactose Utilization by *Lactococcus lactis*. Appl. Environ. Microbiol. 76, 7048–7060. doi:10.1128/AEM.01195-10

Newberry, C.J., Webster, G., Cragg, B.A., Parkes, R.J., Weightman, A.J., Fry, J.C., 2004. Diversity of prokaryotes and methanogenesis in deep subsurface sediments from the Nankai Trough, Ocean Drilling Program Leg 190. Environ. Microbiol. 6, 274–287. doi:10.1111/j.1462-2920.2004.00568.x

Nicholson, W.L., Krivushin, K., Gilichinsky, D., Schuerger, A.C., 2013. Growth of *Carnobacterium* spp. from permafrost under low pressure, temperature, and anoxic atmosphere has implications for Earth microbes on Mars. Proc. Natl. Acad. Sci. U. S. A. 110, 666–671. doi:10.1073/pnas.1209793110

Nicholson, W.L., Zhalnina, K., de Oliveira, R.R., Triplett, E.W., 2015. Proposal to rename *Carnobacterium inhibens* as *Carnobacterium inhibens* subsp *inhibens* subsp nov and description of *Carnobacterium inhibens* subsp *gilichinskyi* subsp nov., a psychrotolerant bacterium isolated from Siberian permafrost. Int. J. Syst. Evol. Microbiol. 65, 556–561. doi:10.1099/ijs.0.067983-0

Nieminen, T.T., Vihavainen, E., Paloranta, A., Lehto, J., Paulin, L., Auvinen, P., Solismaa, M., Bjorkroth, K.J., 2011. Characterization of psychrotrophic bacterial communities in modified atmosphere-packed meat with terminal restriction fragment length polymorphism. Int. J. Food Microbiol. 144, 360–366. doi:10.1016/j.ijfoodmicro.2010.10.018

Nissen, H., Holck, A., Dainty, R.H., 1994. Identification of *Carnobacterium* spp. and *Leuconostoc* spp. in meat by genus-specific 16S rRNA probes. Lett. Appl. Microbiol. 19, 165–8.

Ntougias, S., Zervakis, G.I., Kavroulakis, N., Ehaliotis, C., Papadopoulou, K.K., 2004. Bacterial diversity in spent mushroom compost assessed by amplified rDNA restriction analysis and sequencing of cultivated isolates. Syst. Appl. Microbiol. 27, 746–754. doi:10.1078/0723202042369857

Oses, S.M., Diez, A.M., Melero, B., Luning, P.A., Jaime, I., Rovira, J., 2013. Characterization by culture-dependent and culture-independent methods of the bacterial population of suckling-lamb packaged in different atmospheres. Food Microbiol. 36, 216–222. doi:10.1016/j.fm.2013.05.005

Oussenko, I.A., Sanchez, R., Bechhofer, D.H., 2004. *Bacillus subtilis* YhcR, a High-Molecular-Weight, Nonspecific Endonuclease with a Unique Domain Structure. J. Bacteriol. 186, 5376–5383. doi:10.1128/JB.186.16.5376-5383.2004

Pallen, M.J., Lam, A.C., Antonio, M., Dunbar, K., 2001. An embarrassment of sortases – a richness of substrates? Trends Microbiol. 9, 97–101. doi:10.1016/S0966-842X(01)01956-4

Paludan-Muller, C., Dalgaard, P., Huss, H.H., Gram, L., 1998. Evaluation of the role of *Carnobacterium piscicola* in spoilage of vacuum- and modified-atmosphere-packed cold-smoked salmon stored at 5 degrees C. Int. J. Food Microbiol. 39, 155–66.

Papadimitriou, K., Anastasiou, R., Maistrou, E., Plakas, T., Papandreou, N.C., Hamodrakas, S.J., Ferreira, S., Supply, P., Renault, P., Pot, B., Tsakalidou, E., 2015. Acquisition through Horizontal Gene Transfer of Plasmid pSMA198 by *Streptococcus macedonicus* ACA-DC 198 Points towards the Dairy Origin of the Species. PLoS One 10. doi:10.1371/journal.pone.0116337

Passerini, D., Beltramo, C., Coddeville, M., Quentin, Y., Ritzenthaler, P., Daveran-Mingot, M.L., Le Bourgeois, P., 2010. Genes but not genomes reveal bacterial domestication of *Lactococcus lactis*. PloS one 5, e15306.

Pastar, I., Tonic, I., Golic, N., Kojic, M., van Kranenburg, R., Kleerebezem, M., Topisirovic, L., Jovanovic, G., 2003. Identification and genetic characterization of a novel proteinase, PrtR, from the human isolate *Lactobacillus rhamnosus* BGT10. Appl. Environ. Microbiol. 69, 5802–5811.

Pescuma, M., Turbay, M.B.E., Mozzi, F., Valdez, G.F. de, Giori, G.S. de, Hebert, E.M., 2013. Diversity in proteinase specificity of thermophilic *Lactobacilli* as revealed by hydrolysis of dairy and vegeTable proteins. Appl. Microbiol. Biotechnol. 97, 7831–7844. doi:10.1007/s00253-013-5037-0

Petranovic, D., Guédon, E., Sperandio, B., Delorme, C., Ehrlich, D., Renault, P., 2004. Intracellular effectors regulating the activity of the *Lactococcus lactis* CodY pleiotropic transcription regulator. Mol. Microbiol. 53, 613–621. doi:10.1111/j.1365-2958.2004.04136.x

Pfeiler, E.A., Klaenhammer, T.R., 2007. The genomics of lactic acid bacteria. Trends Microbiol. 15, 546–553. doi:10.1016/j.tim.2007.09.010

Picon, A., Kunji, E.R.S., Lanfermeijer, F.C., Konings, W.N., Poolman, B., 2000. Specificity Mutants of the Binding Protein of the Oligopeptide Transport System of *Lactococcus lactis*. J. Bacteriol. 182, 1600–1608.

Pikuta, E.V., 2014. The family *Carnobacteriaceae*. Blackwell Science Publ, Oxford.

Pikuta, E.V., Fisher, D., Hoover, R.B., 2011. Anaerobic cultures from preserved tissues of baby mammoth. P. SOC. PHOTO-OPT. INS. 8152, 81520U. doi:10.1117/12.897080

Pikuta, E.V., Hoover, R.B., 2014. The genus *Carnobacterium*. Blackwell Science Publ, Oxford.

Pikuta, E.V., Marsic, D., Bej, A., Tang, J., Krader, P., Hoover, R.B., 2005. *Carnobacterium pleistocenium* sp nov., a novel psychrotolerant, facultative anaerobe isolated from permafrost of the Fox Tunnel in Alaska. Int. J. Syst. Evol. Microbiol. 55, 473–478. doi:10.1099/ijs.0.63384-0

Pilet, M.F., Dousset, X., Barré, R., Novel, G., Desmazeaud, M., Piard, J.C., 1994. Evidence for two bacteriocins produced by *Carnobacterium piscicola* and *Carnobacterium divergens* isolated from fish and active against *Listeria monocytogenes*. J. Food Prot. 58, 256–262.

Pilet, M.-F., Leroi, F., 2011. Applications of protective cultures, bacteriocins and bacteriophages in fresh seafood and seafood products. Woodhead Publ. Food Sci. Technol. Nutr. 324–347.

Pilpa, R.M., Robson, S.A., Villareal, V.A., Wong, M.L., Phillips, M., Clubb, R.T., 2009. Functionally Distinct NEAT (NEAr Transporter) Domains within the *Staphylococcus aureus* IsdH/HarA Protein Extract Heme from Methemoglobin. J. Biol. Chem. 284, 1166–1176. doi:10.1074/jbc.M806007200

Poolman, B., 1993. Energy transduction in lactic acid bacteria. FEMS Microbiol. Rev. 12, 125–147. doi:10.1111/j.1574-6976.1993.tb00015.x

Poolman, B., Royer, T.J., Mainzer, S.E., Schmidt, B.F., 1990. Carbohydrate utilization in *Streptococcus thermophilus*: characterization of the genes for aldose 1-epimerase (mutarotase) and UDPglucose 4-epimerase. J. Bacteriol. 172, 4037–4047.

Poolman, B., Royer, T.J., Mainzer, S.E., Schmidt, B.F., 1989. Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. J. Bacteriol. 171, 244–253.

Postma, P.W., Lengeler, J.W., Jacobson, G.R., 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. 57, 543–594.

Pot, B., 2008. The taxonomy of lactic acid bacteria., in: Bactéries Lactiques, de La Génétique Aux Ferments. G. Corrieu, F.M. Luquet, pp. 1–152.

Pot, B., Ludwig, W., Kersters, K., Schleifer, K.-H., 1994. Taxonomy of Lactic Acid Bacteria, in: Vuyst, L.D., Vandamme, E.J. (Eds.), Bacteriocins of Lactic Acid Bacteria. Springer US, pp. 13–90.

Premi, L., Sandine, W.E., Elliker, P.R., 1972. Lactose-hydrolyzing enzymes of *Lactobacillus* species. Appl. Microbiol. 24, 51–57.

Price, C.E., Zeyniyev, A., Kuipers, O.P., Kok, J., 2012. From meadows to milk to mucosa – adaptation of *Streptococcus* and *Lactococcus* species to their nutritional environments. FEMS Microbiol. Rev. 36, 949–971. doi:10.1111/j.1574-6976.2011.00323.x

Proft, T., Baker, E.N., 2009. Pili in Gram-negative and Gram-positive bacteria — structure, assembly and their role in disease. Cell. Mol. Life Sci. 66, 613–635. doi:10.1007/s00018-008-8477-4

Quigley, B.R., Zähner, D., Hatkoff, M., Thanassi, D.G., Scott, J.R., 2009. Linkage of T3 and Cpa pilins in the *Streptococcus pyogenes* M3 pilus. Molecular Microbiology 72, 1379–1394. doi:10.1111/j.1365-2958.2009.06727.x

Rachman, C., Kabadjova, P., Valcheva, R., Prévost, H., Dousset, X., 2004. Identification of *Carnobacterium* Species by Restriction Fragment Length Polymorphism of the 16S-23S rRNA Gene Intergenic Spacer Region and Species-Specific PCR. Appl. Environ. Microbiol. 70, 4468–4477. doi:10.1128/AEM.70.8.4468-4477.2004

Rahman, A., 2013. Étude de la diversité génétique chez la bactérie lactique *Carnobacterium maltaromaticum* et de son adaptation à l'environnement gastro-intestinal de mammifères. Université de Lorraine.

Rahman, A., Cailliez-Grimal, C., Bontemps, C., Payot, S., Chaillou, S., Revol-Junelles, A.-M., Borges, F., 2014a. High Genetic Diversity among Strains of the Unindustrialized Lactic Acid Bacterium *Carnobacterium maltaromaticum* in Dairy Products as Revealed by Multilocus Sequence Typing. Appl. Environ. Microbiol. 80, 3920–3929. doi:10.1128/AEM.00681-14

Rahman, A., Gleinser, M., Lanhers, M.-C., Riedel, C.U., Foligne, B., Hanse, M., Yen, F.T., Klouj, A., Afzal, M.I., Back, A., Mangavel, C., Cailliez-Grimal, C., Revol-Junelles, A.-M., Borges, F., 2014b. Adaptation of the lactic acid bacterium *Carnobacterium maltaromaticum* LMA 28 to the mammalian gastrointestinal tract: From survival in mice to interaction with human cells. Int. Dairy J. 34, 93–99. doi:10.1016/j.idairyj.2013.07.003

Randle, C.L., Albro, P.W., Dittmer, J.C., 1969. The phosphoglyceride composition of gram-negative bacteria and the changes in composition during growth. Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism 187, 214–220. doi:10.1016/0005-2760(69)90030-7

Rasinkangas, P., Reunanen, J., Douillard, F.P., Ritari, J., Uotinen, V., Palva, A., Vos, W.M. de, 2014. Genomic Characterization of Non-Mucus Adherent Derivatives of *Lactobacillus rhamnosus* GG Reveals Genes Affecting Pilus Biogenesis. Appl. Environ. Microbiol. AEM.02006–14. doi:10.1128/AEM.02006-14

Remenant, B., Jaffrès, E., Dousset, X., Pilet, M.-F., Zagorec, M., 2015. Bacterial spoilers of food: Behavior, fitness and functional properties. Food Microbiology, Spoilers, wonder spores and diehard microorganisms: New insights to integrate these super foes in food spoilage risk management 45, Part A, 45–53. doi:10.1016/j.fm.2014.03.009

Richards, V.P., Choi, S.C., Bitar, P.D.P., Gurjar, A.A., Stanhope, M.J., 2013. Transcriptomic and genomic evidence for *Streptococcus agalactiae* adaptation to the bovine environment. BMC genomics 14, 920. doi:10.1186/1471-2164-14-920

Richards, V.P., Lang, P., Pavinski Bitar, P.D., Lefébure, T., Schukken, Y.H., Zadoks, R.N., Stanhope, M.J., 2011. Comparative genomics and the role of lateral gene transfer in the evolution of bovine adapted *Streptococcus agalactiae*. Infect. Genet. Evol. 11, 1263–1275. doi:10.1016/j.meegid.2011.04.019

Rieder, G., Krisch, L., Fischer, H., Kaufmann, M., Maringer, A., Wessler, S., 2012. *Carnobacterium divergens* - a dominating bacterium of pork meat juice. FEMS Microbiol. Lett. 332, 122–130. doi:10.1111/j.1574-6968.2012.02584.x

Ringo, E., Lodemel, J.B., Myklebust, R., Jensen, L., Lund, V., Mayhew, T.M., Olsen, R.E., 2002. The effects of soybean, linseed and marine oils on aerobic gut microbiota of *Arctic charr Salvelinus alpinus L.* before and after challenge with *Aeromonas salmonicida* ssp *salmonicida*. Aquac. Res. 33, 591–606. doi:10.1046/j.1365-2109.2002.00697.x

Robinson, W.G., Old, L.A., Shah, D.S.H., Russell, R.R.B., 2003. Chromosomal insertions and deletions in *Streptococcus mutans*. Caries Res. 37, 148–156. doi:10.1159/000069023

Roof, D.M., Roth, J.R., 1988. Ethanolamine utilization in *Salmonella typhimurium*. J. Bacteriol. 170, 3855–3863.

Rosey, E.L., Oskouian, B., Stewart, G.C., 1991. Lactose metabolism by *Staphylococcus aureus*: characterization of *lacABCD*, the structural genes of the tagatose 6-phosphate pathway. J. Bacteriol. 173, 5992–5998.

Rosey, E.L., Stewart, G.C., 1992. Nucleotide and deduced amino acid sequences of the *lacR, lacABCD*, and *lacFE* genes encoding the repressor, tagatose 6-phosphate gene cluster, and sugar-specific phosphotransferase system components of the lactose operon of *Streptococcus mutans*. J. Bacteriol. 174, 6159–6170.

Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzo, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., Telford, J.L., 2006. Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. Mol. Microbiol. 61, 126–141. doi:10.1111/j.1365-2958.2006.05225.x

Russell, R.R., Aduse-Opoku, J., Sutcliffe, I.C., Tao, L., Ferretti, J.J., 1992. A binding protein-dependent transport system in *Streptococcus mutans responsible* for multiple sugar metabolism. J. Biol. Chem. 267, 4631–4637.

Sadat-Mekmene, L., Genay, M., Atlan, D., Lortal, S., Gagnaire, V., 2011. Original features of cell-envelope proteinases of *Lactobacillus helveticus*. A review. Int. J. Food Microbiol. 146, 1–13. doi:10.1016/j.ijfoodmicro.2011.01.039

Savijoki, K., Ingmer, H., Varmanen, P., 2006. Proteolytic systems of lactic acid bacteria. Appl. Microbiol. Biotechnol. 71, 394–406. doi:10.1007/s00253-006-0427-1

Schaffer, P.A., Lifland, B., Van Sommeran, S., Casper, D.R., Davis, C.R., 2013. Meningoencephalitis Associated With *Carnobacterium maltaromaticum* -Like Bacteria in Stranded Juvenile Salmon Sharks (*Lamna ditropis*). Vet. Pathol. 50, 412–417. doi:10.1177/0300985812441033

Schaumburg, J., Diekmann, O., Hagendorff, P., Bergmann, S., Rohde, M., Hammerschmidt, S., Jänsch, L., Wehland, J., Kärst, U., 2004. The cell wall subproteome of *Listeria monocytogenes*. Proteomics 4, 2991–3006. doi:10.1002/pmic.200400928

Schirmer, B.C., Heir, E., Langsrud, S., 2009. Characterization of the bacterial spoilage flora in marinated pork products. J. Appl. Microbiol. 106, 2106–16.

Schmidt, B.F., Adams, R.M., Requadt, C., Power, S., Mainzer, S.E., 1989. Expression and nucleotide sequence of the *Lactobacillus bulgaricus* beta-galactosidase gene cloned in *Escherichia coli*. J. Bacteriol. 171, 625–635.

Schneewind, O., Missiakas, D., 2014. Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. Biochim. Biophys. Acta, Protein trafficking and secretion in bacteria 1843, 1687–1697. doi:10.1016/j.bbamcr.2013.11.009

Scott, J.R., Barnett, T.C., 2006. Surface Proteins of Gram-Positive Bacteria and How They Get There. Annu. Rev. Microbiol. 60, 397–423. doi:10.1146/annurev.micro.60.080805.142256

Scott, J.R., Zähner, D., 2006. Pili with strong attachments: Gram-positive bacteria do it differently. Mol. Microbiol. 62, 320–330. doi:10.1111/j.1365-2958.2006.05279.x

Sengupta, R., Altermann, E., Anderson, R.C., McNabb, W.C., Moughan, P.J., Roy, N.C., Sengupta, R., Altermann, E., Anderson, R.C., McNabb, W.C., Moughan, P.J., Roy, N.C., 2013. The Role of Cell Surface Architecture of *Lactobacilli* in Host-Microbe Interactions in the Gastrointestinal Tract, The Role of Cell Surface Architecture of *Lactobacilli* in Host-Microbe Interactions in the Gastrointestinal Tract. Mediat inflamm 2013, 2013, e237921. doi:10.1155/2013/237921, 10.1155/2013/237921

Setlow, B., Cabrera-Hernandez, A., Cabrera-Martinez, R.M., Setlow, P., 2004. Identification of aryl-phospho-beta-D-glucosidases in *Bacillus subtilis*. Arch. Microbiol. 181, 60–67. doi:10.1007/s00203-003-0628-2

Shapiro, J.A., Adhya, S.L., 1969. The Galactose Operon of *E. coli* K-12. II. a Deletion Analysis of Operon Structure and Polarity. Genetics 62, 249–264.

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. doi:10.1186/1471-2164-7-126

Siezen, R.J., 1999. Multi-domain, cell-envelope proteinases of lactic acid bacteria. Antonie Van Leeuwenhoek 76, 139–155.

Siezen, R.J., Renckens, B., Swam, I. van, Peters, S., Kranenburg, R. van, Kleerebezem, M., Vos, W.M. de, 2005. Complete Sequences of Four Plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 Reveal Extensive Adaptation to the Dairy Environment. Appl. Environ. Microbiol. 71, 8371–8382. doi:10.1128/AEM.71.12.8371-8382.2005

Smeianov, V.V., Wechter, P., Broadbent, J.R., Hughes, J.E., Rodríguez, B.T., Christensen, T.K., Ardö, Y., Steele, J.L., 2007. Comparative high-density microarray analysis of gene expression during growth of *Lactobacillus helveticus* in milk versus rich culture medium. Appl. Environ. Microbiol. 73, 2661–2672. doi:10.1128/AEM.00005-07

Smit, B.A., van Hylckama Vlieg, J.E.T., Engels, W.J.M., Meijer, L., Wouters, J.T.M., Smit, G., 2005. Identification, cloning, and characterization of a *Lactococcus lactis* branched-chain alpha-keto acid decarboxylase involved in flavor formation. Appl. Environ. Microbiol. 71, 303–311. doi:10.1128/AEM.71.1.303-311.2005

Smit, G., Smit, B.A., Engels, W.J.M., 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS Microbiol. Rev. 29, 591–610. doi:10.1016/j.fmrre.2005.04.002

Snauwaert, I., Hoste, B., De Bruyne, K., Peeters, K., De Vuyst, L., Willems, A., Vandamme, P., 2013. *Carnobacterium iners* sp nov., a psychrophilic, lactic acid-producing bacterium from the littoral zone of an Antarctic pond. Int. J. Syst. Evol. Microbiol. 63, 1370–1375. doi:10.1099/ijs.0.042861-0

Solopova, A., Bachmann, H., Teusink, B., Kok, J., Neves, A.R., Kuipers, O.P., 2012. A Specific Mutation in the Promoter Region of the Silent *cel* Cluster Accounts for the Appearance of Lactose-Utilizing *Lactococcus lactis* MG1363. Appl. Environ. Microbiol. 78, 5612–5621. doi:10.1128/AEM.00455-12

Soong, G., 2006. Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. J. Clin. Invest. 116, 2297–2305. doi:10.1172/JCI27920

Spirig, T., Weiner, E.M., Clubb, R.T., 2011. Sortase enzymes in Gram-positive bacteria. Mol. Microbiol. 82, 1044–1059. doi:10.1111/j.1365-2958.2011.07887.x

Stackebrandt, E., Teuber, M., 1988. Molecular taxonomy and phylogenetic position of lactic acid bacteria. Biochimie 70, 317–324. doi:10.1016/0300-9084(88)90204-0

Steen, A., Buist, G., Horsburgh, G.J., Venema, G., Kuipers, O.P., Foster, S.J., Kok, J., 2005. AcmA of *Lactococcus lactis* is an N-acetylglucosaminidase with an optimal number of LysM domains for proper functioning. FEBS J. 272, 2854–2868. doi:10.1111/j.1742-4658.2005.04706.x

Stiles, M.E., Holzapfel, W.H., 1997. Lactic acid bacteria of foods and their current taxonomy. Int. J. Food Microbiol. 36, 1–29. doi:10.1016/S0168-1605(96)01233-0

Suree, N., Jung, M.E., Clubb, R.T., 2007. Recent advances towards new anti-infective agents that inhibit cell surface protein anchoring in *Staphylococcus aureus* and other gram-positive pathogens. Mini Rev. Med. Chem. 7, 991–1000.

Sutcliffe, I.C., Harrington, D.J., 2002. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. Microbiol. 148, 2065–2077.

Swaminathan, A., Mandlik, A., Swierczynski, A., Gaspar, A., Das, A., Ton-That, H., 2007. Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. Mol. Microbiol. 66, 961–974. doi:10.1111/j.1365-2958.2007.05968.x

Tailford, L.E., Crost, E.H., Kavanaugh, D., Juge, N., 2015. Mucin glycan foraging in the human gut microbiome. Front. Genet 6, 81. doi:10.3389/fgene.2015.00081

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30, 2725–2729. doi:10.1093/molbev/mst197

Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., Grandi, G., 2006. Pili in Gram-positive pathogens. Nature Rev. Microbiol. 4, 509–519. doi:10.1038/nrmicro1443

Thomas, T.D., Crow, V.L., 1984. Selection of Galactose-Fermenting *Streptococcus thermophilus* in Lactose-Limited Chemostat Cultures. Appl. Environ. Microbiol. 48, 186–191.

Thomas, T.D., Turner, K.W., Crow, V.L., 1980. Galactose fermentation by *Streptococcus lactis* and *Streptococcus cremoris*: pathways, products, and regulation. J. Bacteriol. 144, 672–682.

Thompson, J., 1987. Regulation of sugar transport and metabolism in lactic acid bacteria. FEMS Microbiol. Lett. 46. doi:10.1111/j.1574-6968.1987.tb02462.x

Thompson, J., 1980. Galactose transport systems in *Streptococcus lactis*. J. Bacteriol. 144, 683–691.

Thompson, J., Gentry-Weeks, C., 1994. Métabolisme des sucres par les bactéries lactiques., in: Bactéries Lactiques. De Roissart H, Luquet F, Lorica M, Uriage France, pp. 239–290.

Thornley, 1957. Observations on the microflora of minced chicken meat irradiated with 4MeV cathode rays. Journal of Applied Bacteriology 368–376.

Tiedemann, M.T., Stillman, M.J., 2012. Heme binding to the IsdE(M78A; H229A) double mutant: challenging unidirectional heme transfer in the iron-regulated surface determinant protein heme transfer pathway of *Staphylococcus aureus*. J. Biol. Inorg. Chem. 17, 995–1007. doi:10.1007/s00775-012-0914-z

Tohno, M., Kobayashi, H., Nomura, M., Uegaki, R., Cai, Y., 2012. Identification and characterization of lactic acid bacteria isolated from mixed pasture of timothy and orchardgrass, and its badly preserved silage. Anim. Sci. J. 83, 318–330. doi:10.1111/j.1740-0929.2011.00955.x

Ton-That, H., Schneewind, O., 2003. Assembly of pili on the surface of *Corynebacterium diphtheriae*. Mol. Microbiol. 50, 1429–1438.

Toranzo, A., Romalde, J., Nunez, S., Figueras, A., Barja, J., 1993. An Epizootic in Farmed, Market-Size Rainbow-Trout in Spain Caused by a Strain of Carnobacterium-Piscicola of Unusual Virulence. Dis. Aquat. Org. 17, 87–99. doi:10.3354/dao017087

Torres, V.J., Pishchany, G., Humayun, M., Schneewind, O., Skaar, E.P., 2006. *Staphylococcus aureus* IsdB Is a Hemoglobin Receptor Required for Heme Iron Utilization. J. Bacteriol. 188, 8421–8429. doi:10.1128/JB.01335-06

Tripathi, P., Beaussart, A., Alsteens, D., Dupres, V., Claes, I., von Ossowski, I., de Vos, W.M., Palva, A., Lebeer, S., Vanderleyden, J., Dufrêne, Y.F., 2013. Adhesion and nanomechanics of pili from the probiotic *Lactobacillus rhamnosus* GG. ACS Nano 7, 3685–3697. doi:10.1021/nn400705u

Trost, M., Wehmhöner, D., Kärst, U., Dieterich, G., Wehland, J., Jänsch, L., 2005. Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. Proteomics 5, 1544–1557. doi:10.1002/pmic.200401024

Tsai, Y.-K., Lin, T.-H., 2006. Sequence, organization, transcription and regulation of lactose and galactose operons in *Lactobacillus rhamnosus* TCELL-1. J. Appl. Microbiol. 100, 446–459. doi:10.1111/j.1365-2672.2005.02790.x

Tynkkynen, S., Buist, G., Kunji, E., Kok, J., Poolman, B., Venema, G., Haandrikman, A., 1993. Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. J. Bacteriol. 175, 7523–7532.

Upadhyay, V.K.., McSweeney, P.L.H., Magboul, A.A.A., Fox, P.F., 2004. Proteolysis in cheese during ripening., in: Cheese, Chemistry, Physics and Microbiology. P. F. Fox, T. M. Cogan and T. P. Guinee Elsevier.

Vadeboncoeur, C., Pelletier, M., 1997. The phosphoenolpyruvate:sugar phosphotransferase system of oral *streptococci* and its role in the control of sugar metabolism. FEMS Microbiol. Rev. 19, 187–207.

Vaillancourt, K., Moineau, S., Frenette, M., Lessard, C., Vadeboncoeur, C., 2002. Galactose and lactose genes from the galactose-positive bacterium *Streptococcus salivarius* and the phylogenetically related galactose-negative bacterium *Streptococcus thermophilus*: organization, sequence, transcription, and activity of the *gal* gene products. J. Bacteriol. 184, 785–793.

Vallenet, D., Belda, E., Calteau, A., Cruveiller, S., Engelen, S., Lajus, A., Le Fevre, F., Longin, C., Mornico, D., Roche, D., Rouy, Z., Salvignol, G., Scarpelli, C., Thil Smith, A.A., Weiman, M., Medigue, C., 2013. MicroScope--an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. Nucleic Acids Res. 41, D636–D647. doi:10.1093/nar/gks1194

van de Guchte, M., Penaud, S., Grimaldi, C., Barbe, V., Bryson, K., Nicolas, P., Robert, C., Oztas, S., Mangenot, S., Couloux, A., Loux, V., Dervyn, R., Bossy, R., Bolotin, A., Batto, J.-M., Walunas, T., Gibrat, J.-F., Bessières, P., Weissenbach, J., Ehrlich, S.D., Maguin, E., 2006. The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. Proc. Natl. Acad. Sci. U.S.A. 103, 9274–9279. doi:10.1073/pnas.0603024103

van Rooijen, R.J., van Schalkwijk, S., de Vos, W.M., 1991. Molecular cloning, characterization, and nucleotide sequence of the tagatose 6-phosphate pathway gene cluster of the lactose operon of *Lactococcus lactis*. J. Biol. Chem. 266, 7176–7181.

van Tilbeurgh, H., Declerck, N., 2001. Structural insights into the regulation of bacterial signalling proteins containing PRDs. Curr. Opin. Struct. Biol. 11, 685–693.

Vasilopoulos, C., De Maere, H., De Mey, E., Paelinck, H., De Vuyst, L., Leroy, F., 2010. Technology-induced selection towards the spoilage microbiota of artisan-type cooked ham packed under modified atmosphere. Food Microbiol. 27, 77–84. doi:10.1016/j.fm.2009.08.008

Vaughan, E.E., David, S., Vos, W.M. de, 1996. The lactose transporter in *Leuconostoc lactis* is a new member of the LacS subfamily of galactoside-pentose-hexuronide translocators. Appl. Environ. Microbiol. 62, 1574–1582.

Vaughan, E.E., Pridmore, R.D., Mollet, B., 1998. Transcriptional regulation and evolution of lactose genes in the galactose-lactose operon of *Lactococcus lactis* NCD02054. J. Bacteriol. 180, 4893–4902.

Vaughan, E.E., van den Bogaard, P.T.C., Catzeddu, P., Kuipers, O.P., de Vos, W.M., 2001. Activation of silent gal genes in the *lac-gal* regulon of *Streptococcus thermophilus*. J. Bacteriol. 183, 1184–1194. doi:10.1128/JB.183.4.1184-1194.2001

Vendrell, D., Balcázar, J.L., Ruiz-Zarzuela, I., de Blas, I., Gironés, O., Múzquiz, J.L., 2006. *Lactococcus garvieae* in fish: a review. Comp. Immunol. Microbiol. Infect. Dis. 29, 177–198. doi:10.1016/j.cimid.2006.06.003

Vihavainen, E., Lundstrom, H.-S., Susiluoto, T., Koort, J., Paulin, L., Auvinen, P., Bjorkroth, K.J., 2007. Role of broiler carcasses and processing plant air in contamination of modified-atmosphere-packaged broiler products with psychrotrophic lactic acid bacteria. Appl. Environ. Microbiol. 73, 1136–1145. doi:10.1128/AEM.01644-06

Voget, S., Klippel, B., Daniel, R., Antranikian, G., 2011. Complete Genome Sequence of *Carnobacterium* sp 17-4. J. Bacteriol. 193, 3403–3404. doi:10.1128/JB.05113-11

von Heijne, G., 2001. Signal Peptides, in: eLS. John Wiley & Sons, Ltd.

Walker, V.K., Palmer, G.R., Voordouw, G., 2006. Freeze-Thaw Tolerance and Clues to the Winter Survival of a Soil Community. Appl. Environ. Microbiol. 72, 1784–1792. doi:10.1128/AEM.72.3.1784-1792.2006

Wallbanks, S., Martinez-Murcia, A.J., Fryer, J.L., Phillips, B.A., Collins, M.D., 1990. 16S rRNA Sequence Determination for Members of the Genus *Carnobacterium* and Related Lactic Acid Bacteria and Description of *Vagococcus salmoninarum sp. nov.* . Int. J. Syst. Bacteriol. 40, 224–230. doi:10.1099/00207713-40-3-224

Wan, J., Harmark, K., Davidson, B.E., Hillier, A.J., Gordon, J.B., Wilcock, A., Hickey, M.W., Coventry, M.J., 1997. Inhibition of *Listeria monocytogenes* by piscicolin 126 in milk and Camembert cheese manufactured with a thermophilic starter. Journal of applied microbiology 82, 273–80.

Wasney, M.A., Holley, R.A., Jayas, D.S., 2001. Cresol red thallium acetate sucrose inulin (CTSI) agar for the selective recovery of *Carnobacterium* spp. Int. J. Food Microbiol. 64, 167–74.

Wassenaar, T.M., Lukjancenko, O., 2014. Comparative genomics of *Lactobacillus* and other LAB, in: Holzapfel, W.H., Wood, B.J.B. (Eds.), Lactic Acid Bacteria. John Wiley & Sons, Ltd, pp. 55–69.

Wegmann, U., Overweg, K., Jeanson, S., Gasson, M., Shearman, C., 2012. Molecular characterization and structural instability of the industrially important composite metabolic plasmid pLP712. Microbiol. 158, 2936–2945. doi:10.1099/mic.0.062554-0

Welman, A.D., Maddox, I.S., 2003. Exopolysaccharides from lactic acid bacteria: perspectives and challenges. Trends in Biotechnology 21, 269–274. doi:10.1016/S0167-7799(03)00107-0

Wely, K.H.M. van, Swaving, J., Freudl, R., Driessen, A.J.M., 2001. Translocation of proteins across the cell envelope of Gram-positive bacteria. FEMS Microbiol. Rev. 25, 437–454. doi:10.1111/j.1574-6976.2001.tb00586.x

Wu, H., Fives-Taylor, P.M., 2001. Molecular Strategies for Fimbrial Expression and Assembly. C.R.O.B.M. 12, 101–115. doi:10.1177/10454411010120020101

Wu, Q., Cheung, C.K.W., Shah, N.P., 2015. Towards galactose accumulation in dairy foods fermented by conventional starter cultures: Challenges and strategies. Trends Food Sci. Tech. 41, 24–36. doi:10.1016/j.tifs.2014.08.010

Wyder, A.B., Boss, R., Naskova, J., Kaufmann, T., Steiner, A., Graber, H.U., 2011. *Streptococcus* spp. and related bacteria: Their identification and their pathogenic potential for chronic mastitis – A molecular approach. Res. Vet. Sci. 91, 349–357. doi:10.1016/j.rvsc.2010.09.006

Xu, J.G., Yang, H.M., Lai, X.H., Fu, X.L., Wu, J.M., Huang, L.B., Yu, X.J., Wu, Y.P., Wu, Y.J., Liu, B.Y., 1997. Etiological study for a case of multi-bacterial synergistic gangrene. Chin. Sci. Bull. 42, 511–&. doi:10.1007/BF02882606

Yanagawa, R., Otsuki, K., Tokui, T., 1968. Electron microscopy of fine structure of *Corynebacterium renale* with special reference to pili. JPN J. VET. RES. 16, 31–37.

Yanagida, F., Chen, Y.-S., Yasaki, M., 2007. Isolation and characterization of lactic acid bacteria from lakes. J. Basic Microbiol. 47, 184–190. doi:10.1002/jobm.200610237

Yan, N., 2013. Structural advances for the major facilitator superfamily (MFS) transporters. Trends Biochem. Sci. 38, 151–159. doi:10.1016/j.tibs.2013.01.003

Youssef, M.K., Gill, C.O., Tran, F., Yang, X., 2014. Unusual Compositions of Microflora of Vacuum-Packaged Beef Primal Cuts of Very Long Storage Life. J. Food Prot. 77, 2161–2167. doi:10.4315/0362-028X.JFP-14-190

Zadoks, R.N., Middleton, J.R., McDougall, S., Katholm, J., Schukken, Y.H., 2011. Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. J. Mammary Gland Biol. Neoplasia 16, 357–372. doi:10.1007/s10911-011-9236-y

Zähner, D., Hakenbeck, R., 2000. The *Streptococcus pneumoniae* Beta-Galactosidase Is a Surface Protein. J. Bacteriol. 182, 5919–5921.

Zapotoczna, M., Heilbronner, S., Speziale, P., Foster, T.J., 2012. Iron-Regulated Surface Determinant (Isd) Proteins of </i>Staphylococcus lugdunensis</i>. J. Bacteriol. 194, 6453–6467. doi:10.1128/JB.01195-12

Zeng, L., Das, S., Burne, R.A., 2010. Utilization of Lactose and Galactose by *Streptococcus mutans*: Transport, Toxicity, and Carbon Catabolite Repression. J. Bacteriol. 192, 2434–2444. doi:10.1128/JB.01624-09

Zeng, L., Martino, N.C., Burne, R.A., 2012. Two Gene Clusters Coordinate Galactose and Lactose Metabolism in *Streptococcus gordonii*. Appl. Environ. Microbiol. 78, 5597–5605. doi:10.1128/AEM.01393-12

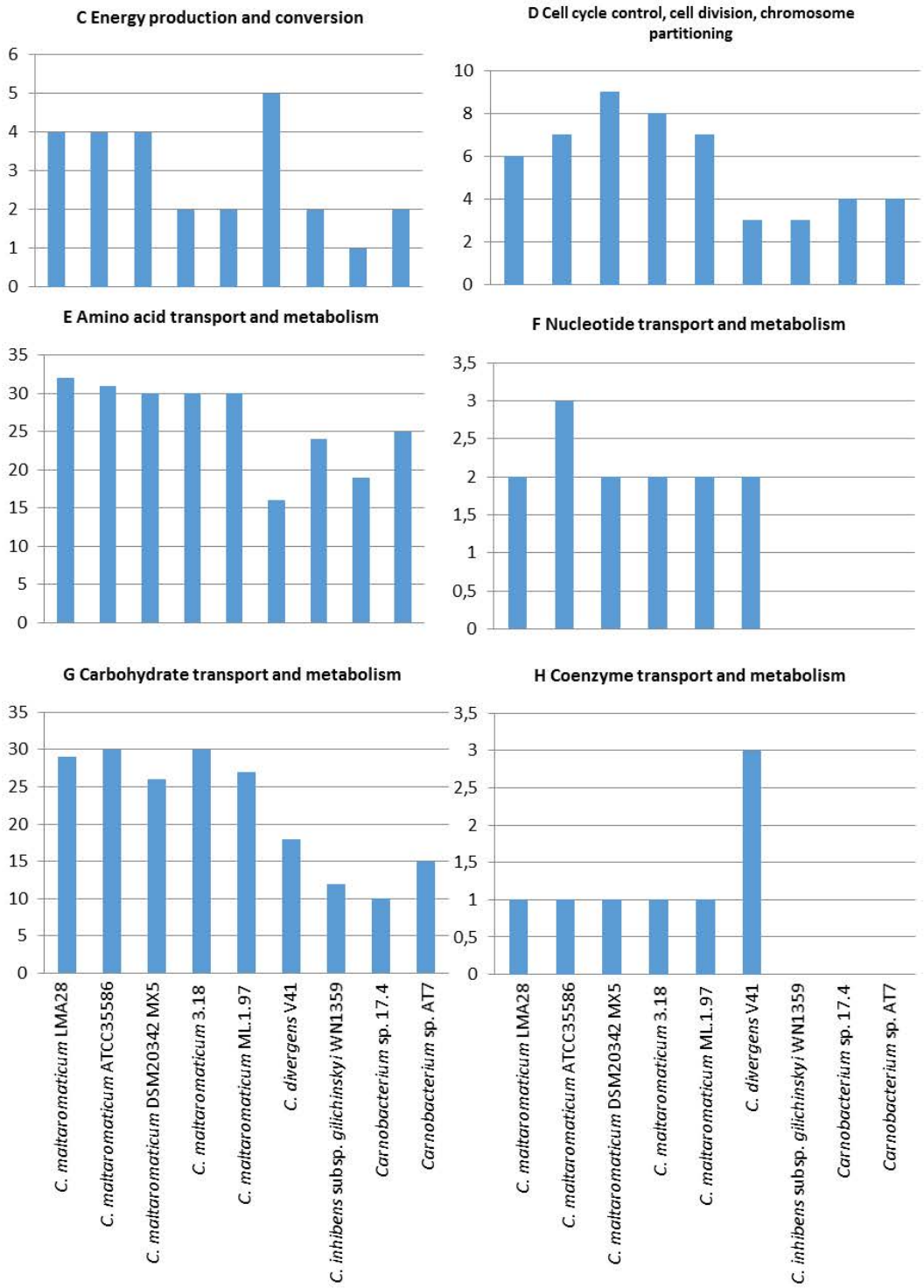
Zhang, J., 2003. Evolution by gene duplication: an update. trend. Ecol. Evol. 18, 292–298. doi:10.1016/S0169-5347(03)00033-8

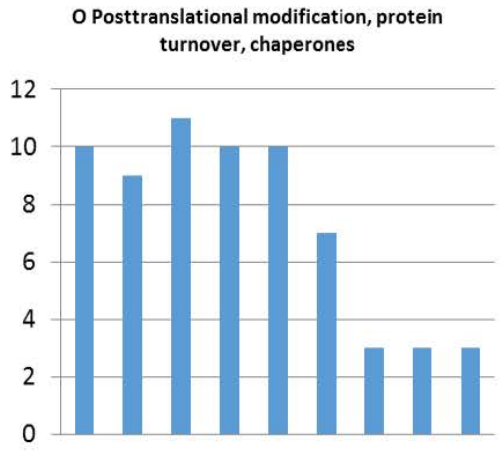
Zhang, Z.-G., Ye, Z.-Q., Yu, L., Shi, P., 2011. Phylogenomic reconstruction of lactic acid bacteria: an update. BMC Evol. Biol. 11, 1. doi:10.1186/1471-2148-11-1

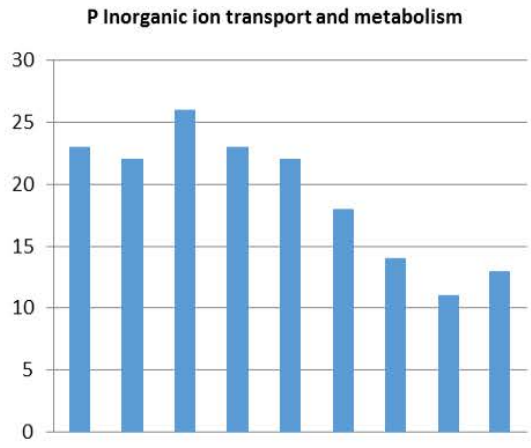
Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S., 2011. PHAST: A Fast Phage Search Tool. Nucl. Acids Res. gkr485. doi:10.1093/nar/gkr485

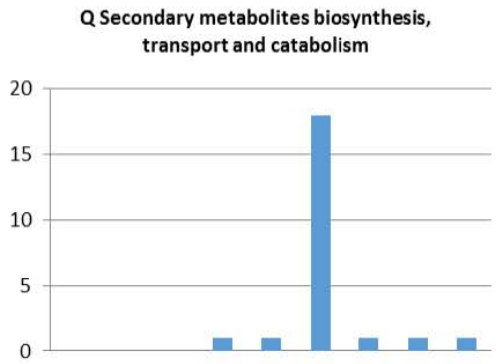
ANNEXES

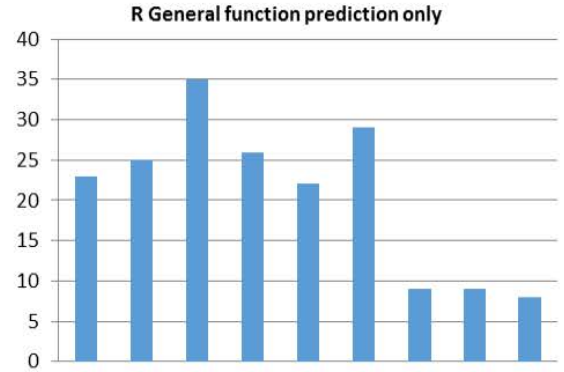
Annexe 1: CDS distribution of the intraspecific core genome in the different COG

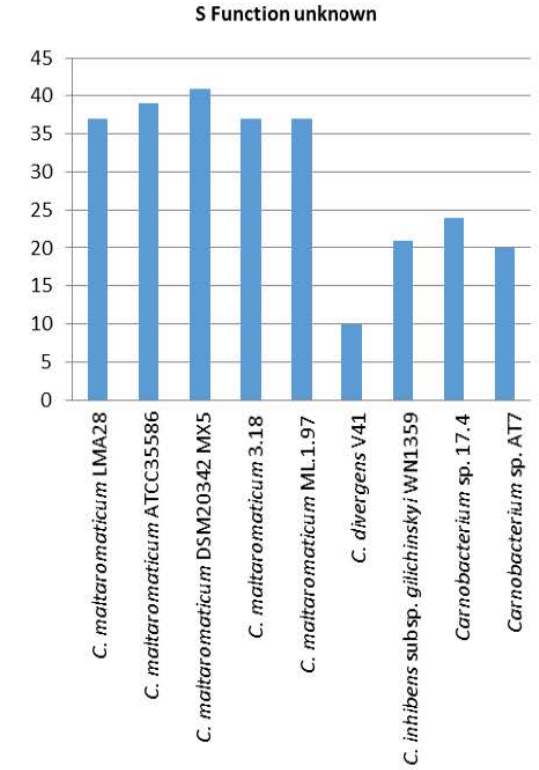


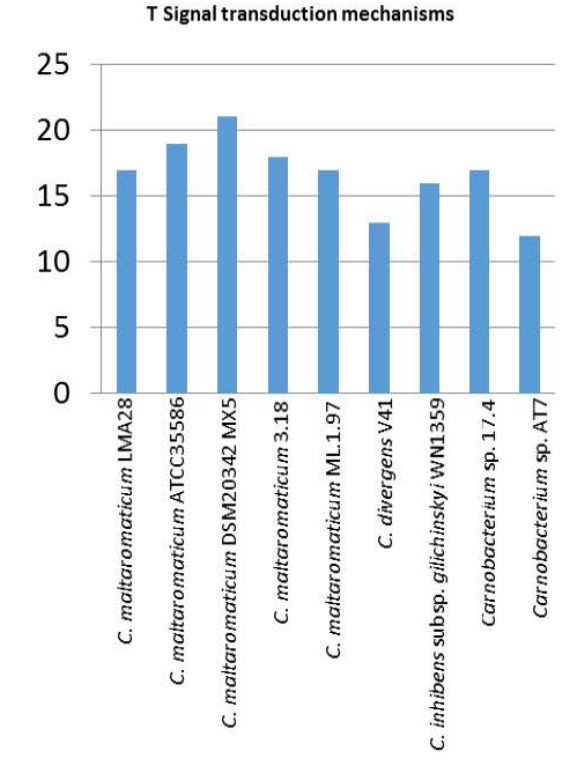

O Posttranslational modification, protein turnover, chaperones
12
10
8
6
4
2
0

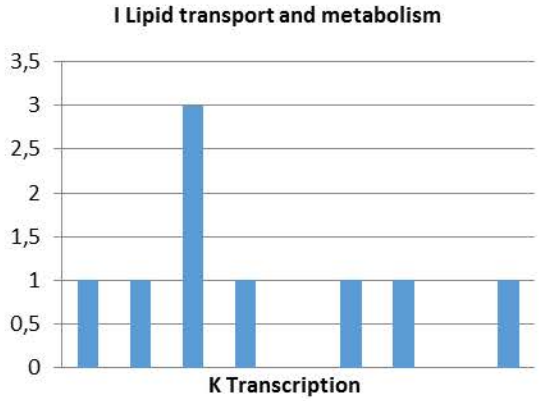

P Inorganic ion transport and metabolism
30
25
20
15
10
5
0

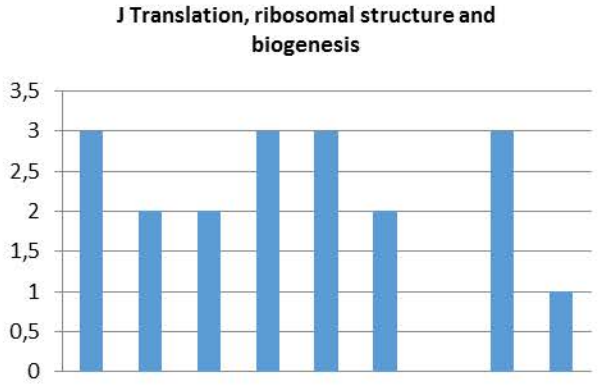

Q Secondary metabolites biosynthesis, transport and catabolism
20
15
10
5
0


R General function prediction only
40
35
30
25
20
15
10
5
0

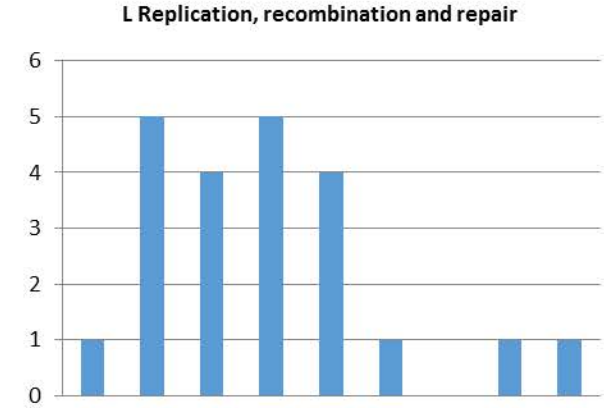

S Function unknown
45
40
35
30
25
20
15
10
5
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7

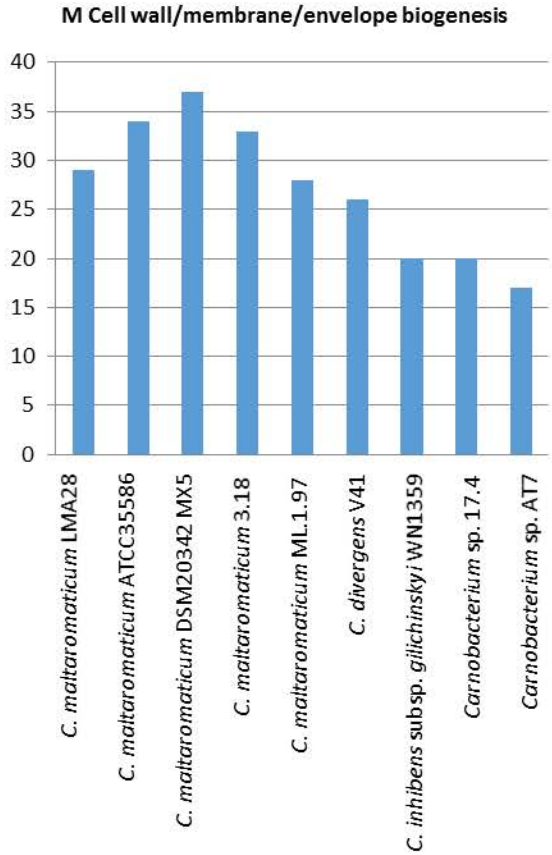

T Signal transduction mechanisms
25
20
15
10
5
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7

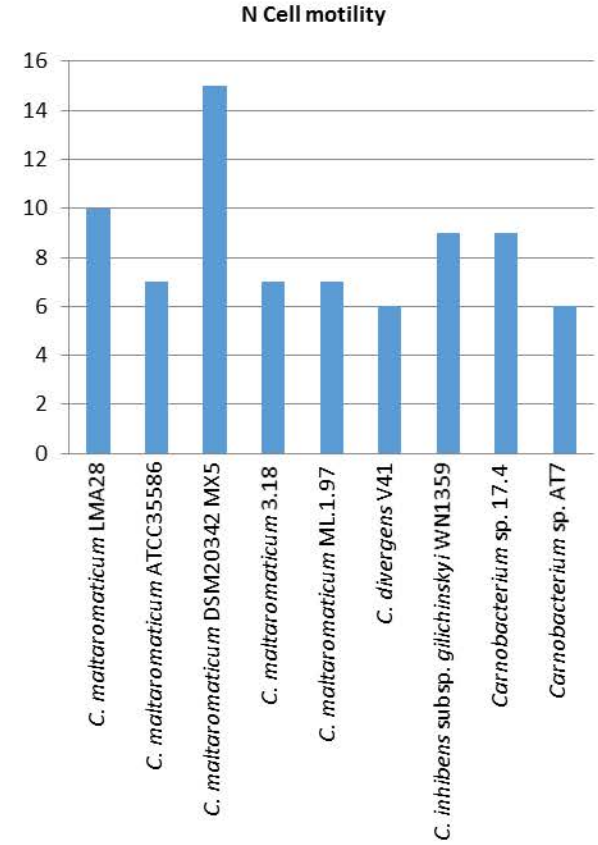

I Lipid transport and metabolism
3,5
3
2,5
2
1,5
1
0,5
0

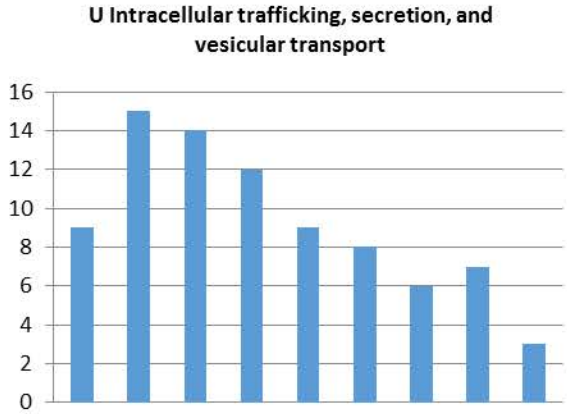

J Translation, ribosomal structure and biogenesis
3,5
3
2,5
2
1,5
1
0,5
0

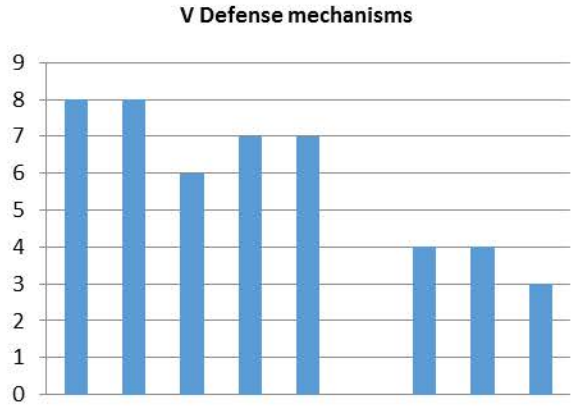
K Transcription
6
5
4
3
2
1
0

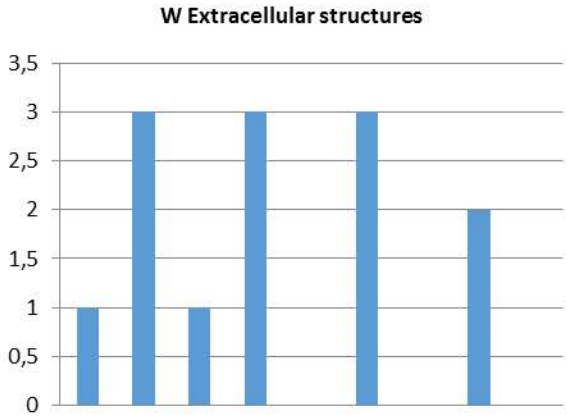

L Replication, recombination and repair
6
5
4
3
2
1
0

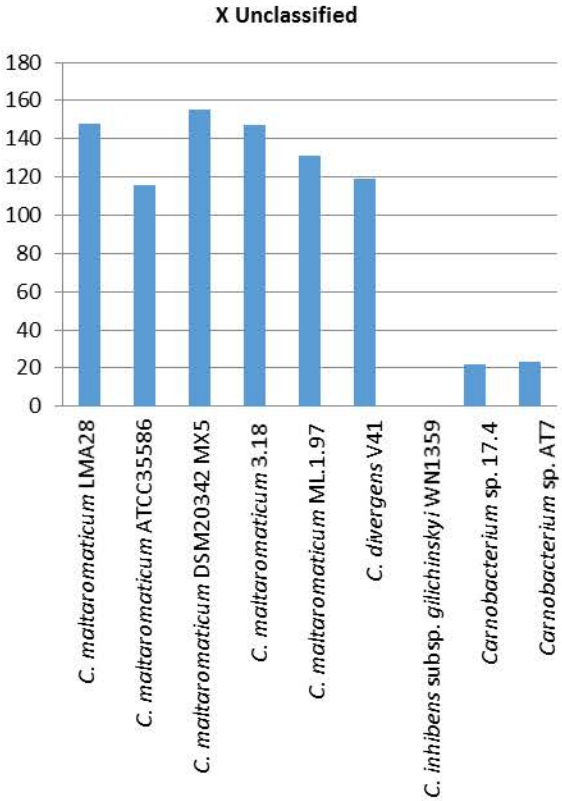

M Cell wall/membrane/envelope biogenesis
40
35
30
25
20
15
10
5
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML.1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7

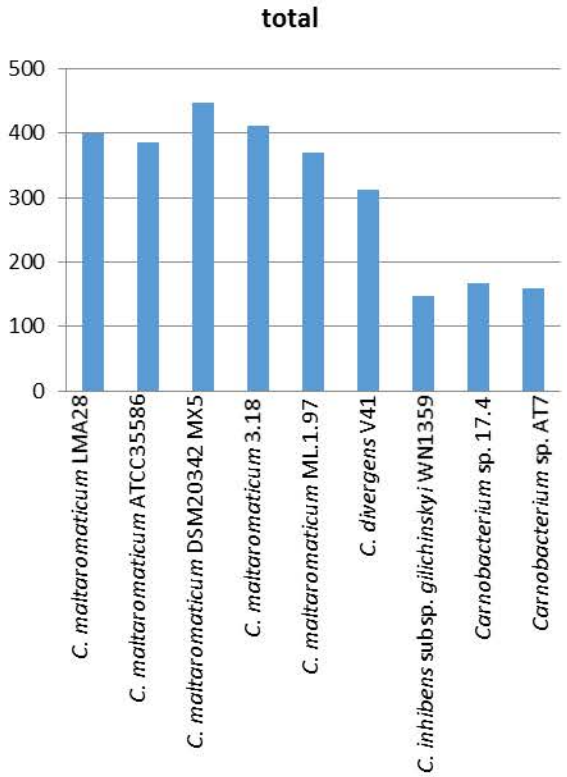

N Cell motility
16
14
12
10
8
6
4
2
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML.1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7


U Intracellular trafficking, secretion, and vesicular transport
16
14
12
10
8
6
4
2
0


V Defense mechanisms
9
8
7
6
5
4
3
2
1
0


W Extracellular structures
3,5
3
2,5
2
1,5
1
0,5
0


X Unclassified
180
160
140
120
100
80
60
40
20
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML.1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7


total
500
400
300
200
100
0
C. maltaromaticum LMA28
C. maltaromaticum ATCC35586
C. maltaromaticum DSM20342 MX5
C. maltaromaticum 3.18
C. maltaromaticum ML.1.97
C. divergens V41
C. inhibens subsp. gilichinskyi WN1359
Carnobacterium sp. 17.4
Carnobacterium sp. AT7

Annexe 2: Summary of the analysis of the cell wall proteins

	C. maltaromaticum					*C. divergens*	*C. inhibens* ssp. *gilichinskyi*	*Carnobacterium* sp.	
	LMA28	**DSM20342 MX5**	**ATCC35586**	**3.18**	**ML.1.97**	**V41**	**WN1359**	**AT7**	**17.4**
LysM	6	5	5	5	5	6	4	5	5
WxL	**48**	**49**	**37**	**53**	**42**	**46**	0	0	0
SDP	**35**	**35**	**28**	**29**	**26**	**26**	0	1	1
Sortase	**6**	**8**	**6**	**6**	**6**	**5**	1	0	0
Total	**95**	**97**	**76**	**93**	**79**	**83**	5	6	6

		C. maltaromaticum					*C. divergens*	*C. inhibens* subdp. *gilichinskyi*	*Carnobacterium* sp.	
		LMA28	DSM20342	ATCC35586	3.18	ML.1.97	V41	WN1359	AT7	17.4
Sortase A	Total	4	6	5	5	5	4			1
	Conserved between *C. maltaromaticum* and *C. divergens* V41	4	5	5	5	5	4			
	Specific		1							1
	Duplicated		1	1	1	1				
Sortase B	Total	2	2	1	1	1	1			
	Conserved between *C. maltaromaticum* and *C. divergens* V41	1	1	1	1	1	1			
	conserved LMA28+DSM	1	1							
	Fragments				1	1				
SDP	Total	35	35	28	29	26	26			1
	Specific	2	2	0	0	1	10			1
	duplicated				2		2			
	Fragments	1	2	4	2	3	1			
WxL	Total	48	49	37	53	42	46			
	Specific	1	1	5	7	4	15			
	duplicated		2	2	1	2				
LysM	Total	6	5	5	5	5	6	4	5	5
	Specific	1					1			
	Fragments				1	1	1			

Annexe 3: Genes encoding surface proteins of the 9 *Carnobacterium* strains, and the genes involved in the proteolysis. Yellow cells indicate the presence of one homologue in a genome, the green cells indicate the presence of duplicated genes in the same genome and red cells indicate the presence of pseudogenes.

		LMA28	DSM20342	ATCC35586	3.18	ML-1-97	*C. divergens* V41	*C. gilichiskyi* WN1359	C. sp AT7	C. sp 17-4
Proteinase	prtB with LPXTG motif anchor domain	BN424_3077	JQMX_v1-12801	AGNSv1_370130	CM318v1_460013	BN1423_v1_1290020	CDIVv1_140219			
Oligopeptide transport system	OppA	BN424_931 BN424_1168 BN424_1169 BN424_1263 BN424_1936	JQMX_v1_10423 JQMX_v1_11013 JQMX_v1_11112 JQMX_v1_11113 JQMX_v1_11403 JQMX_v1_11856 JQMX_v1_12043	AGNSv1_230077 AGNSv1_310016 AGNSv1_310017 AGNSv1_390072 AGNSv1_480018 AGNSv1_520062 AGNSv1_200038	CM318v1_50005 CM318v1_50006 CM318v1_210036 CM318v1_220040 CM318v1_440127 CM318v1_560063 CM318v1_1500018	BN1423_v1_730035 BN1423_v1_740006 BN1423_v1_740007 BN1423_v1_950019 BN1423_v1_1320013	CDIVv1_320648 CDIVv1_320649 CDIVv1_120089 CDIVv1_50025 CDIVv1_40028 CDIVv1_20057	Q783_03285 Q783_03290 Q783_07975	ABHHv1_110071 ABHHv1_110072 ABHHv1_010131	CAR_c07390 CAR_c07400 CAR_c17340
	OppB	BN424_2393	JQMX_v1_13500	AGNSv1_220058	CM318v1_430043	BN1423_v1_210001	CDIVv1_320639	Q783_03325	ABHHv1_110064	CAR_c07470
	OppC	BN424_2392	JQMX_v1_13501	AGNSv1_220059	CM318v1_430044	BN1423_v1_210002	CDIVv1_320638	Q783_03330	ABHHv1_110063	CAR_c07480
	OppD	BN424_2391	JQMX_v1_13502	AGNSv1_220060	CM318v1_430045	BN1423_v1_210003	CDIVv1_320637	Q783_03335	ABHHv1_110062	CAR_c07490
	OppF	BN424_2390	JQMX_v1_13503	AGNSv1_220061	CM318v1_430046	BN1423_v1_210004	CDIVv1_320636	Q783_03340	ABHHv1_110061	CAR_c07500
	DtpT	BN424_1736	JQMX_v1_10552	AGNSv1_500044	CM318v1_210217	BN1423_v1_180008	CDIVv1_320386			
Tripeptidase	PepT	BN424_1746 BN424_1935	JQMX_v1_10424 JQMX_v1_10543	AGNSv1_500035 AGNSv1_390073	CM318v1_210208 CM318v1_210037	BN1423_v1_1540007 BN1423_v1_1250021	CDIVv1_320392	Q783_05365 Q783_09335	ABHHv1_90071 ABHHv1_100008	CAR_c11730 CAR_c20370
Dipeptidase	PepV	BN424_1096		AGNSv1_550032	CM318v1_590029	BN1423_v1_2190004	CDIVv1_40254	Q783_02400	ABHHv1_250032	CAR_c05690
Endopeptidase	PepF	BN424_2445 BN424_2215	JQMX_v1_10206	AGNSv1_220008 AGNSv1_610001	CM318v1_630028 CM318v1_170009	BN1423_v1_1970002 BN1423_v1_360007	CDIVv1_320632 CDIVv1_40179	Q783_03360 Q783_02810 Q783_09330	ABHHv1_110057 ABHHv1_70101 ABHHv1_100007	CAR_c07540 CAR_c06450 CAR_c20360
Prolidase	PepQ	BN424_2226	JQMX_v1_10196	AGNSv1_590002	CM318v1_160014	BN1423_v1_1840002 BN1423_v1_1840001	CDIVv1_40192	Q783_02750	ABHHv1_470003	CAR_c06350
Proline peptidase	PepP/PepQ	BN424_2304	JQMX_v1_10117	AGNSv1_2200092	CM318v1_430077	BN1423_v1_1530028		Q783_10240	ABHHv1_60006	
Pyrrolidone carboxylate peptidase	PCP	BN424_3450	JQMX_v1_12456	AGNSv1_320128	CM318v1_1470042	BN1423_v1_890019	CDIVv1_270082			CAR_c14840
Aminopeptidase	PepA	BN424_1132		AGNSv1_460012	CM318v1_60012	BN1423_v1_1170012	CDIVv1_320680	Q783_03205	ABHHv1_140014	CAR_c07230
	PepS	BN424_2586	JQMX_v1_13289	AGNSv1_230273	CM318v1_580017	BN1423_v1_450006	CDIVv1_230162	Q783_07180	ABHHv1_300004	CAR_c15790
	PepC							Q783_08765	ABHHv1_20055	CAR_c19130
Sortase	srtA	BN424_577	JQMX_v1_11737	AGNSv1_640058	CM318v1_500010	BN1423_v1_50010	CDIVv1_270320			
		BN424_1760	JQMX_v1_10530	AGNSv1_500022	CM318v1_210195	BN1423_v1_1060017	CDIVv1_320409			
		BN424_2317	JQMX_v1_10008 JQMX_v1_10101	AGNSv1_640058 AGNSv1_260136	CM318v1_500010 CM318v1_520044	BN1423_v1_50010 BN1423_v1_490025	CDIVv1_270320			
		BN424_2789	JQMX_v1_13088	AGNSv1_260136	CM318v1_520044	BN1423_v1_490025	CDIVv1_140233			
			JQMX_v1_40069							
	srtB à coté de IsdE	BN424_3051	JQMX_v1_12830	AGNSv1_690008	CM318v1_200004	BN1423_v1_710010	CDIVv1_320484			
		BN424_748	JQMX_v1_11575 JQMX_v1_11576 (fragment)	AGNSv1_400067	CM318_1560036	BN1423_v1_370006				

	srtB avec IsdC	BN424_3054	JQMX_v1_12827	AGNSv1_690005	CM318v1_460038 (fragment) CM318v1_200001 (fragment)	BN1423_v1_710006 (fragment) BN1423_v1_710007 (fragment)				
	srtB avec PilB	BN424_3144	JQMX_v1_12728							
	putative sortase							Q783_9540	ABHHv1_120050 (pas de signalP)	CAR_c15170 (pas de signalP)
LPXTG	PrtB	BN424_3077	JQMX_v1_12801	AGNSv1_370130	CM318V1_460013	BN1423_v1_1290020	CDIVv1-140219			
		BN424_3226	JQMX_v1_12648	AGNSv1_450052	CM318V1_1570046	BN1423_v1_290004	CDIVv1_140229			
	LPxTG+WxL	BN424_2474	JQMX_v1_13426	AGNSv1_300004	CM318V1_630004	BN1423_v1_700003	CDIVv1_140241 (fragment)			
		BN424_310	JQMX_v1_11988	AGNSv1_740031	CM318V1_570033	BN1423_v1_240001 (fragment)	CDIVv1_270272 (fragment) CDIVv1_270271 (fragment) CDIVv1_270270			
	nucleotidase	BN424_3529	JQMX_v1_12376	AGNSv1_320001 (fragment) AGNSv1_410033 (fragment)	CM318V1_670001	BN1423_v1_1450001	CDIVv1_10018			
	nucleotidase	BN424_190	JQMX_v1_12101	AGNSv1_90007	CM318V1_750003	BN1423-v1_1270006	CDIVv1_270033			
		BN424_3553	JQMX_v1_12353	AGNSv1_410009	CM318V1_620001	BN1423_v1_420009	CDIVv1_140031			
		BN424_3557	JQMX_v1_12350	AGNSv1_410006	CM318V1_140002	BN1423_v1_420006	CDIVv1_280039 CDIVv1_280058			
		BN424_3223	JQMX_v1_12651	AGNSv1_450056	CM318V1_1570049	BN1423_v1_290007	CDIVv1_280061			
	collagen binding protein (2 parts: one with signal P and one with LPXTG motif)	BN424_147 (part1) BN424_148 (part2)	JQMX_v1_12145 (part1) JQMX_v1_12144 (part2)	AGNSv1_360180 (part1) AGNSv1_360181 (part2) AGNSv1_360182 (part3)	CM318V1_1510055	BN1423_v1_1070059	CDIVv1_270205			
		BN424_1109	JQMX_v1_11170	AGNSv1_290005	CM318V1_600005	BN1423_v1_1000005				
		BN424_324	JQMX_v1_11977	AGNSv1_740043	CM318V1_570044	BN1423_v1_240012				
		BN424_1966	JQMX_v1_10393	AGNSv1_390045	CM318V1_210008	BN1423_v1_1340011				
	Collagen-binding surface protein Cna-like	BN424_2511	JQMX_v1_13389	AGNSv1_710004	CM318V1_340003	BN1423_v1_130003				
	Copper resistance protein CopC/internalin, Leucine-rich repeat-containing adjacent domain	BN424_3087	JQMX_v1_12787	AGNSv1_370121	CM318V1_1460060 (fragment)	BN1423_v1_1290008 (fragment)				
		BN424_3173	JQMX_v1_12702	AGNSv1_370018	CM318V1_1490021	BN1423_v1_280014				
	Collagen-binding surface protein Cna, B-type domain	BN424_3195	JQMX_v1_12678	AGNSv1_450081	CM318V1_1570085	BN1423_v1_300001 (fragment) BN1423_v1_300002 (fragment)				
		BN424_220	JQMX_v1_12070	AGNSv1_90019	CM318V1_90018	BN1423_v1_1280003				
		BN424_201	JQMX_v1_12089	AGNSv1_90019	CM318V1_230009 CM318V1_370020	BN1423-v1_750010				
	mucBP	BN424_25	JQMX_v1_12240	AGNSv1_360081	CM318V1_1470001	BN1423_v1_920005				
	mucBP	BN424_711	JQMX_v1_11615 (fragment) JQMX_v1_11616 (fragment)	AGNSv1_400036	CM318V1_1560001	BN1423_v1_880005				
		BN424_183	JQMX_v1_12108	AGNSv1_90001 (fragment) AGNSv1_360218 (fragment)	CM318V1_660001 (fragment) CM318V1_1510097 (fragment)					
	Glycoside hydrolase, superfamily	BN424_641	JQMX_v1_11684	AGNSv1_70077		BN1423_v1_1020014				
		BN424_115	JQMX_v1_12180		CM318V1_1510018		CDIVv1_280025			

	Bacterial Ig-like, Collagen-binding surface protein Cna-like, B-type domain	BN424_123	JQMX_v1_12172		CM318V1_1510026		CDIVv1_280015		
		BN424_159	JQMX_v1_12132	AGNSv1_360194					
	pilB	BN424_3145	JQMX_v1_12727						
	spaD	BN424_3146	JQMX_v1_12726						
	spaC	BN424_3147	JQMX_v1_12725						
		BN424_1627	JQMX_v1_10652						
	mucBP	BN424_2985	JQMX_v1_12898						
	Collagen-binding surface protein Cna-like, B-type domain			AGNSv1_230132 (fragment) AGNSv1_230133 (fragment)			CDIVv1_320181		
	Bacterial adhesins			AGNSv1_410029	CM318V1_620020				
					CM318V1_1570066	BN1423_v1_290027			
		BN424_2364							
		BN424_2984							
	Collagen-binding surface protein Cna-like, B-type domain		JQMX_v1_10050						
	Collagen-binding Cna B-type domain		JQMX_v1_50122						
						BN1423_v1_790008			
	LPxTG+WxL						CDIVv1_10071		
							CDIVv1_140202		
							CDIVv1_120082		
							CDIVv1_120102		
							CDIVv1_140029		
							CDIVv1_140064		
							CDIVv1_140240		
	Adhesin						CDIVv1_270167		
							CDIVv1_270186		
							CDIVv1_50151		
	Collagen-binding Cna B-type, Adhesin							ABHHv1_120049	CAR_c15171
WxL	LPxTG+WxL	BN424_116	JQMX_v1_12179	AGNSv1_90019	CM318V1_1510019	BN1423_v1_290028	CDIVv1_280026		
		BN424_117	JQMX_v1_12178	AGNSv1_90020	CM318V1_1510020	BN1423_v1_1280004	CDIVv1_280027		
		BN424_118	JQMX_v1_12177	AGNSv1_90021	CM318V1_1510021	BN1423_v1_290030	CDIVv1_280028		
		BN424_201	JQMX_v1_12089	AGNSv1_340018	CM318V1_230009	BN1423_v1_870010	CDIVv1_280026		
		BN424_202	JQMX_v1_12088	AGNSv1_340019	CM318V1_230010	BN1423_v1_750011	CDIVv1_280027		
		BN424_203	JQMX_v1_12087	AGNSv1_340020	CM318V1_230011	BN1423_v1_870012	CDIVv1_280028		
		BN424_308	JQMX_v1_11990	AGNSv1_740028	CM318V1_570031	BN1423_v1_220030	CDIVv1_140281		
		BN424_325	JQMX_v1_11975 JQMX_v1_11974	AGNSv1_740045 AGNSv1_740046	CM318V1_570046 CM318V1_570047	BN1423_v1_240014 BN1423_v1_240015	CDIVv1_10072 CDIVv1_10073		
		BN424_326	JQMX_v1_11973	AGNSv1_740047	CM318V1_570048	BN1423_v1_240016	CDIVv1_10074		
		BN424_327	JQMX_v1_11972	AGNSv1_740048	CM318V1_570049	BN1423_v1_240017	CDIVv1_10075		
		BN424_328	JQMX_v1_11971	AGNSv1_740049	CM318V1_570050	BN1423_v1_240018	CDIVv1_10076		
		BN424_329	JQMX_v1_11970	AGNSv1_740050	CM318V1_570051	BN1423_v1_240019	CDIVv1_10077		
		BN424_2469	JQMX_v1_13431	AGNSv1_300009	CM318V1_630009	BN1423_v1_700008	CDIVv1_140245		

	BN424_2471	JQMX_v1_13429	AGNSv1_300007	CM318V1_630007	BN1423_v1_700006	CDIVv1_140244
	BN424_2472	JQMX_v1_13428	AGNSv1_300006	CM318V1_630006	BN1423_v1_700005	CDIVv1_140243
	BN424_2473	JQMX_v1_13427	AGNSv1_300005	CM318V1_630005	BN1423_v1_700004	CDIVv1_140242
	BN424_2474	JQMX_v1_13426	AGNSv1_300004	CM318V1_630004	BN1423_v1_700003	CDIVv1_140241
	BN424_2791	JQMX_v1_13086		CM318V1_520042	BN1423_v1_490023	CDIVv1_140231
	BN424_2792	JQMX_v1_13085	AGNSv1_260134	CM318V1_520041	BN1423_v1_490022	CDIVv1_140230
	BN424_2938	JQMX_v1_12942	AGNSv1_260015	CM318V1_1540011	BN1423_v1_350003	CDIVv1_280060
	BN424_2939	JQMX_v1_12941	AGNSv1_260014	CM318V1_1540010	BN1423_v1_350002	CDIVv1_280059
	BN424_3224	JQMX_v1_12650	AGNSv1_450054	CM318V1_1570048	BN1423_v1_290006	CDIVv1_270188
	BN424_3225	JQMX_v1_12649	AGNSv1_450053	CM318V1_1570047	BN1423_v1_290005	CDIVv1_270187
	BN424_3554	JQMX_v1_12352	AGNSv1_410008	CM318V1_140004	BN1423_v1_420008	CDIVv1_280041 CDIVv1_140030
	BN424_3555	JQMX_v1_12351	AGNSv1_410007	CM318V1_140003	BN1423_v1_420007	CDIVv1_280040 CDIVv1_140029
	BN424_3580	JQMX_v1_12324	AGNSv1_340020	CM318V1_240086	BN1423_v1_870012	CDIVv1_280028
	BN424_3581	JQMX_v1_12323	AGNSv1_340019	CM318V1_240085	BN1423_v1_870011	CDIVv1_280027
	BN424_3582	JQMX_v1_12322	AGNSv1_340018	CM318V1_240084	BN1423_v1_870010	CDIVv1_280026
similar to BN424_354 of LMA28 holding WxL and LPxTG		JQMX_v1_11976	AGNSv1_740044	CM318V1_570045	BN1423_v1_240013	CDIVv1_10072
	BN424_3084	JQMX_v1_12790	AGNSv1_370123	CM318V1_460003 CM318V1_460002	BN1423_v1_1290011	
	BN424_3085	JQMX_v1_12789	AGNSv1_370123	CM318V1_460004 CM318V1_460002	BN1423_v1_1290011	
	BN424_264	JQMX_v1_12033	AGNSv1_480029	CM318V1_1500028	BN1423_v1_950030	
	BN424_120	JQMX_v1_12175		CM318V1_1510023		CDIVv1_280030
	BN424_122	JQMX_v1_12173		CM318V1_1510025		CDIVv1_280032
LPxTG+WxL	BN424_220	JQMX_v1_12070		CM318V1_90018		
	BN424_221	JQMX_v1_12069		CM318V1_90019		
	BN424_222	JQMX_v1_12068		CM318V1_90020		
	BN424_223	JQMX_v1_12067		CM318V1_90021		
	BN424_1624	JQMX_v1_10655				
	BN424_1625	JQMX_v1_10654				
	BN424_1626	JQMX_v1_10653				
	BN424_1800	JQMX_v1_10489	AGNSv1_370123		BN1423_v1_1290011	
	BN424_2961	JQMX_v1_12919	AGNSv1_360094			
	BN424_2963	JQMX_v1_12917	AGNSv1_360092			
LRR_5, carbohydrate binding proteins (Lectin-L-type superfamily), Bacterial Ig-like (Bid-2)	BN424_3039	JQMX_v1_12842		CM318V1_200014	BN1423_v1_760003	
LRR_5, Bacterial Ig-like (Bid-2)	BN424_3043	JQMX_v1_12838				
	BN424_3045	JQMX_v1_12836		CM318V1_200009	BN1423_v1_1400001	
	BN424_3049	JQMX_v1_12832		CM318V1_200006	BN1423_v1_710012	
				CM318V1_1570068	BN1423_v1_290029	
				CM318V1_1570071	BN1423_v1_290032	
				CM318V1_1570072	BN1423_v1_290033	
	BN424_1621					

			JQMX_v1_12190							
				AGNSv1_370063						
				AGNSv1_370059						
				AGNSv1_370057						
				AGNSv1_190016						
				AGNSv1_190017						
					CM318V1_360001					
					CM318V1_370020					
					CM318V1_370021					
					CM318V1_370022					
					CM318V1_370024					
					CM318V1_1570067					
					CM318V1_1570069					
						BN1423_v1_790009				
						BN1423_v1_790011				
	LPxTG+WxL					BN1423_v1_1280003				
						BN1423_v1_2160001				
							CDIVv1_10079			
							CDIVv1_50148			
							CDIVv1_50150			
							CDIVv1_140065			
							CDIVv1_140066			
							CDIVv1_140226			
							CDIVv1_140265			
							CDIVv1_140280			
							CDIVv1_270257			
							CDIVv1_270258			
							CDIVv1_270259			
							CDIVv1_270260			
							CDIVv1_270261			
							CDIVv1_270262			
							CDIVv1_270263			
LysM		BN424_800	JQMX_v1_11522	AGNSv1_620033	CM318V1_440002	BN1423_v1_1440002	CDIVv1_140075	Q783_07720	ABHHv1_10084	CAR_c16890
	Glucosaminidase, Lysozyme subfamily LytF	BN424_1067	JQMX_v1_11209	AGNSv1_180007	CM318V1_40001 (fragment)	BN1423_v1_2240001 (fragment)	CDIVv1_290015 (fragment) CDIVv1_290016 (fragment)	Q783_02360	ABHHv1_370007	CAR_c05630
		BN424_292	JQMX_v1_12003	AGNSv1_740014	CM318V1_570017	BN1423_v1_220016	CDIVv1_320266			
		BN424_333	JQMX_v1_11966	AGNSv1_740055	CM318V1_570055	BN1423_v1_240024	CDIVv1_50135			
	Glucosaminidase	BN424_1749	JQMX_v1_10540	AGNSv1_500032	CM318V1_210205	BN1423_v1_1540004	CDIVv1_320396			
		BN424_3392								
							CDIVv1_50081			
	autolysin + SH3							Q783_07065	ABHHv1_130044	CAR_c15550
	peptidase							Q783_11350	ABHHv1_380002	CAR_c25080
	Polysaccharide deacetylase								ABHHv1_210037	CAR_c10900

Annexe 4: Percentage of identity between *lac* and *gal* genes within LAB. * % of identity between the duplicates of the same strain. ** % of identity between the genome and LMA28

		lacR	*lacA2*	*lacB2*	*lacA1*	*lacB1*	*lacC1*	*lacD1*	*lacE1*	*lacF1*	*lacG1*	*lacF2*	*lacE2*	*lacC2*	*galK*	*galE*	*galT*	*galM*	*galR*
Carnobacterium maltaromaticum	LMA 28	100	100	100	100	100	100	100	100	100	100	100	100	100					
	*		54,61	67,68	54,61	67,68	44,91		58,35	45,1		45,1	58,35	44,91					
	DSM 20342	98,77	100	97,98	100	100	100	100	100	100	100	100	100	100					
	MX5																		
	*		54,61	62,43	54,61	62,43	44,91		57,2	45,1		45,1	57,2	44,91					
	ATCC 35586	98,77	99,3	97,98	100	100	100	100	100	100	100	100	96,45	44,91					
	*		55,32	62,43	55,32	62,43			45,1	60,18	86,05	45,1	60,18						
	ML-1-97																		
	3-18														100	100	100	100	100
Carnobacterium sp.	17-4														67,96	75,84	61,51	51,11	68,39
	AT 7														71,32	74,77	61,02	48,27	68,09
C. gilichinskyi	WN 1359														67,18	75,23	61,1	50,87	69,6
Streptococcus gordonii	str. Challis substr. CH1	40,95	50,35	67,68	72,34	84,8	59,68	72,98	73,73	79,41	84,12	47,06	53,16	44,53	58,91	59,34	52,24	35,28	47,89
	**		51,06	67,68	73,76	84,8	60,97	74,15						44,91					
Streptococcus mutans	UA159	43,62	51,77	68,69	72,34	75,44	70,65	70,15	76,53	76,92	86,05	46,08	54,66	43,02	60,47	59,58	53,36		48,64
Staphylococcus aureus	16K	39,76	53,9	71,72	74,47	84,8	63,55	76	77,1	74,51	86,97	49,02	55,89	47,17					

Annexe 5: List of *C. maltaromaticum* LMA28 megaplasmid genes grouped in the different COG

Class ID	CDS	Label	Gene	Product
G	28	LMAv2_mp0020	_	putative transport system permease
		LMAv2_mp0021	_	Sugar ABC superfamily ATP binding cassette transporter, membrane protein
		LMAv2_mp0022	_	Sugar ABC superfamily ATP binding cassette transporter, sugar-binding protein
		LMAv2_mp0023	_	YhcH/YjgK/YiaL family protein
		LMAv2_mp0027	nanE	putative N-acetylmannosamine-6-phosphate 2-epimerase
		LMAv2_mp0030	_	conserved protein of unknown function
		LMAv2_mp0046	_	membrane protein of unknown function
		LMAv2_mp0049	_	Permease of the major facilitator superfamily (fragment)
		LMAv2_mp0056	_	membrane protein of unknown function
		LMAv2_mp0067	_	conserved protein of unknown function
		LMAv2_mp0075	lacR	Lactose phosphotransferase system repressor
		LMAv2_mp0076	lacA	Galactose-6-phosphate isomerase subunit LacA
		LMAv2_mp0077	_	ribose 5-phosphate isomerase B/allose 6-phosphate isomerase (fragment)
		LMAv2_mp0078	_	Galactose-6-phosphate isomerase subunit LacB (fragment)
		LMAv2_mp0079	_	conserved protein of unknown function
		LMAv2_mp0080	lacA	Galactose-6-phosphate isomerase subunit LacA 2
		LMAv2_mp0081	lacB	Galactose-6-phosphate isomerase subunit LacB
		LMAv2_mp0082	lacC	Tagatose-6-phosphate kinase
		LMAv2_mp0083	lacD	Tagatose 1,6-diphosphate aldolase
		LMAv2_mp0084	lacF	Lactose-specific phosphotransferase enzyme IIA component
		LMAv2_mp0085	lacE	PTS system lactose-specific EIICB component [Includes: Lactose permease IIC component ; Lactose-specific phosphotransferase enzyme IIB component]
		LMAv2_mp0086	lacG	6-phospho-beta-galactosidase
		LMAv2_mp0090	_	Regulatory protein

Class ID	CDS	Label	Gene	Product
L	17	LMAv2_mp0011	_	Cytosine-specific methyltransferase
		LMAv2_mp0012	_	Phage DNA methylase
		LMAv2_mp0018	_	protein of unknown function
		LMAv2_mp0031	_	transposase (fragment)
		LMAv2_mp0042	_	transposase
		LMAv2_mp0044	_	transposase (fragment)
		LMAv2_mp0047	_	transposase (fragment)
		LMAv2_mp0050	_	conserved protein of unknown function
		LMAv2_mp0051	tnpA	transposase
		LMAv2_mp0063	yoeC	putative integrase/recombinase YoeC
		LMAv2_mp0069	_	DNA repair protein RadC
		LMAv2_mp0074	_	conserved protein of unknown function
		LMAv2_mp0091	_	transposase
		LMAv2_mp0105	_	Site-specific recombinase
		LMAv2_mp0110	yoeC	putative integrase/recombinase YoeC
		LMAv2_mp0117	_	Addiction module antitoxin, RelB/DinJ family (fragment)
		LMAv2_mp0119	xerS	Tyrosine recombinase XerS
R	6	LMAv2_mp0003	_	protein of unknown function
		LMAv2_mp0032	_	conserved protein of unknown function
		LMAv2_mp0046	_	membrane protein of unknown function
		LMAv2_mp0049	_	Permease of the major facilitator superfamily (fragment)
		LMAv2_mp0056	_	membrane protein of unknown function
		LMAv2_mp0065	ybbM	putative permease of an ABC transporter

		LMAv2_mp0093	lacF	Lactose-specific phosphotransferase enzyme IIA component
		LMAv2_mp0094	bglH	Aryl-phospho-beta-D-glucosidase BglH
		LMAv2_mp0095	lacE	PTS system lactose-specific EIICB component [Includes: Lactose permease IIC component ; Lactose-specific phosphotransferase enzyme IIB component]
		LMAv2_mp0096	_	protein of unknown function
		LMAv2_mp0097	lacC	Tagatose-6-phosphate kinase
K	7	LMAv2_mp0024	_	conserved protein of unknown function
		LMAv2_mp0030	_	conserved protein of unknown function
		LMAv2_mp0059	_	Rrf2 family transcriptional regulator
		LMAv2_mp0062	_	conserved protein of unknown function
		LMAv2_mp0075	lacR	Lactose phosphotransferase system repressor
		LMAv2_mp0090	_	Regulatory protein
		LMAv2_mp0100	_	protein of unknown function
D	3	LMAv2_mp0113	_	Adenosine monophosphate-protein transferase NmFic (fragment)
		LMAv2_mp0114	_	Adenosine monophosphate-protein transferase NmFic (fragment)
		LMAv2_mp0122	_	putative partition protein/ATPase
V	2	LMAv2_mp0103	_	Daunorubicin resistance ABC superfamily ATP binding cassette transporter
		LMAv2_mp0104	_	ABC superfamily ATP binding cassette transporter
S	1	LMAv2_mp0116	_	toxin of the YafQ-DinJ toxin-antitoxin system (fragment)
U	1	LMAv2_mp0041	_	protein of unknown function

E	5	LMAv2_mp0029	nanA	N-acetylneuraminate lyase
		LMAv2_mp0046	_	membrane protein of unknown function
		LMAv2_mp0049	_	Permease of the major facilitator superfamily (fragment)
		LMAv2_mp0056	_	membrane protein of unknown function
		LMAv2_mp0066	ybbL	putative transporter subunit: ATP-binding component of ABC superfamily
T	5	LMAv2_mp0054	_	putative cell growth regulatory protein
		LMAv2_mp0055	_	PemK family protein
		LMAv2_mp0108	_	Toxin-antitoxin system, toxin component, MazF family
		LMAv2_mp0111	_	putative transcriptional regulator/antitoxin, related to MazE
		LMAv2_mp0112	_	Transcriptional modulator of MazE/toxin, MazF
P	4	LMAv2_mp0046	_	membrane protein of unknown function
		LMAv2_mp0049	_	Permease of the major facilitator superfamily (fragment)
		LMAv2_mp0056	_	membrane protein of unknown function
		LMAv2_mp0064	zosA	Zinc-transporting ATPase
M	2	LMAv2_mp0029	nanA	N-acetylneuraminate lyase
		LMAv2_mp0034	_	conserved protein of unknown function
C	1	LMAv2_mp0058	ykgC	putative pyridine nucleotide-disulfide oxidoreductase YkgC
O	1	LMAv2_mp0106	clpL	ATP-dependent Clp protease ATP-binding subunit ClpL
Q	1	LMAv2_mp0048	_	putative multicopper oxidase

Annexe 6: List of permeases found in *Carnobacterium* strains

3.18 (76 sugar transport system)	Gene	ML.1.97 (71 sugar transport system)	Gene	DSM20342 MX5 (76 sugar transport system)	Gene	ATCC35586 (73 sugar transport system)	Gene	LMA28 (77 sugar transport system)	Gene	17.4 (40 sugar transport system)	Gene	AT7 (48 sugar transport system)	Gene	WN1359 (40 sugar transport system)	Gene
CM318V1_1460025	msmK	CM197v1_180011	o	JQMX_v1_10555	o	AGNSv1_260122	_	BN424_816	glcU	CAR_c22100	CAR_c22100	ABHHv1_50065	–	Q783_00120	Q783_00120
CM318V1_1500007	sgaT	CM197v1_220019	ycjV	JQMX_v1_12000	ycjV	AGNSv1_340032	yteP	BN424_1733	o418	CAR_c24840	msmX	ABHHv1_70008	–	Q783_01945	Q783_01945
CM318V1_70032	fruA	CM197v1_820024	msmK	JQMX_v1_12334	yteP	AGNSv1_370086	msmK	BN424_3078	_	CAR_c05940	ywbF	ABHHv1_230011	ycjV	Q783_01975	Q783_01975
CM318V1_200020	celB	CM197v1_850008	yteP	JQMX_v1_12750	msmK	AGNSv1_370129	_	BN424_241	_	CAR_c01090	dhaM	ABHHv1_290028	–	Q783_02980	Q783_02980
CM318V1_200039	celB	CM197v1_970012	_	JQMX_v1_12799	_	AGNSv1_500047	o	BN424_1402	ulaA	CAR_c01100	dhaL	ABHHv1_680001	–	Q783_06455	Q783_06455
CM318V1_200082	bglF	CM197v1_1290019	_	JQMX_v1_12054	sgaT	AGNSv1_740017	ycjV	BN424_64	_	CAR_c01110	dhaK	ABHHv1_680002	–	Q783_06460	Q783_06460
CM318V1_210144	mtlA	CM197v1_950007	sgaT	JQMX_v1_12334	yteP	AGNSv1_480007	sgaT	BN424_129	levE	CAR_c01290	scrA	ABHHv1_10117	_	Q783_06465	Q783_06465
CM318V1_210146	mtlA	CM197v1_40032	_	JQMX_v1_10374	fruA	AGNSv1_70079	celA	BN424_130	levE	CAR_c02070	bglP1	ABHHv1_10124	_	Q783_06470	Q783_06470
CM318V1_220098	_	CM197v1_60006	pts7C	JQMX_v1_10477	mtlA	AGNSv1_70080	celB	BN424_131	levF	CAR_c03430	CAR_c03430	ABHHv1_10161	_	Q783_08735	Q783_08735
CM318V1_220165	ulaA	CM197v1_130016	_	JQMX_v1_10479	mtlA	AGNSv1_90010	ptsC	BN424_132	levG	CAR_c04060	manX	ABHHv1_20075	_	Q783_08950	Q783_08950
CM318V1_220167	ulaC	CM197v1_130018	_	JQMX_v1_10880	ulaC	AGNSv1_120006	scrA	BN424_135	pts30BCA	CAR_c05070	levF1	ABHHv1_20083	chbA	Q783_10175	Q783_10175
CM318V1_240038	dhaK	CM197v1_230007	celB	JQMX_v1_10882	ulaA	AGNSv1_180008	nagP	BN424_143	_	CAR_c05080	levG1	ABHHv1_20091	ulaA	Q783_10185	Q783_10185

CM318V1_240039	dhaL	CM197v1_230008	_	JQMX_v1_10951	_	AGNSv1_200144	_	BN424_193	ptsC	CAR_c07730	ptsH	ABHHv1_20092	_	Q783_11220	Q783_11220
CM318V1_240040	dhaM	CM197v1_260022	bglP	JQMX_v1_11208	nagP	AGNSv1_220031	ptsH	BN424_195	celA1	CAR_c07740	ptsI	ABHHv1_20095	scrA	Q783_00120	Q783_00120
CM318V1_240043	dhaK	CM197v1_260026	_	JQMX_v1_11329	celB	AGNSv1_220032	ptsI	BN424_240	_	CAR_c09140	bglP2	ABHHv1_30022	manZ	Q783_00655	Q783_00655
CM318V1_240049	treB	CM197v1_260028	ulaB	JQMX_v1_11348	ptnD	AGNSv1_230072	bglP	BN424_241	_	CAR_c09920	ptsG1	ABHHv1_30023	manY	Q783_01245	Q783_01245
CM318V1_340016	_	CM197v1_260029	ptnC	JQMX_v1_11349	ptnC	AGNSv1_230162	bglP	BN424_255	mtlA	CAR_c12680	fruA1	ABHHv1_30024	manX	Q783_01565	Q783_01565
CM318V1_340018	agaV	CM197v1_260030	ptnD	JQMX_v1_11350	ulaB	AGNSv1_230165	_	BN424_364	malX	CAR_c14220	mtlA1	ABHHv1_30108	_	Q783_02000	Q783_02000
CM318V1_340019	agaC	CM197v1_290044	chbB	JQMX_v1_11352	_	AGNSv1_230167	ulaB	BN424_501	scrA	CAR_c14900	bglP3	ABHHv1_50025	sgcB	Q783_02005	Q783_02005
CM318V1_340020	agaD	CM197v1_290045	pts7C	JQMX_v1_11356	bglP	AGNSv1_230168	ptnC	BN424_643	celA	CAR_c14990	bglP4	ABHHv1_50034	_	Q783_02010	Q783_02010
CM318V1_350016	scrA	CM197v1_390004	celB	JQMX_v1_11410	bglP	AGNSv1_230169	ptnD	BN424_644	celB	CAR_c15080	CAR_c15080	ABHHv1_50035	_	Q783_04470	Q783_04470
CM318V1_380031	yfeU	CM197v1_490016	celC	JQMX_v1_11561	ptsA	AGNSv1_230186	celB	BN424_649	licA	CAR_c15090	CAR_c15090	ABHHv1_50036	_	Q783_05390	Q783_05390
CM318V1_380032	_	CM197v1_760009	celB	JQMX_v1_11602	ptbA	AGNSv1_260033	bglF	BN424_716	_	CAR_c17270	exp5	ABHHv1_50037	_	Q783_05750	Q783_05750
CM318V1_410003	celA	CM197v1_780013	ptsH	JQMX_v1_11606	malP	AGNSv1_260079	celB	BN424_721	glvC	CAR_c17800	ptsG2	ABHHv1_50071	_	Q783_06330	Q783_06330
CM318V1_410004	celB	CM197v1_780014	ptsI	JQMX_v1_11610	_	AGNSv1_260264	celB	BN424_725	ptbA	CAR_c18320	CAR_c18320	ABHHv1_90066	_	Q783_06335	Q783_06335
CM318V1_430016	ptsH	CM197v1_790013	celB	JQMX_v1_11636	_	AGNSv1_260268	chbA	BN424_761	ptsA	CAR_c19350	malP	ABHHv1_100032	glvC	Q783_06595	Q783_06595
CM318V1_430017	ptsI	CM197v1_800009	celB	JQMX_v1_11637	yfeU	AGNSv1_320044	lacE	BN424_923	bglP	CAR_c19430	licA	ABHHv1_100045	mtlA	Q783_07895	Q783_07895

CM318V1_440163	bglP	CM197v1_820030	_	JQMX_v1_11681	celB	AGNSv1_320054	lacE	BN424_976	bglP	CAR_c19440	licB	ABHHv1_100047	mtlA	Q783_07940	Q783_07940
CM318V1_440167	_	CM197v1_820032	chbA	JQMX_v1_11682	celA	AGNSv1_320201	pts7C	BN424_979	_	CAR_c19550	sacP	ABHHv1_110037	ptsI	Q783_08130	Q783_08130
CM318V1_440169	ulaB	CM197v1_820033	chbB	JQMX_v1_11812	scrA	AGNSv1_360012	treB	BN424_981	ulaB	CAR_c20740	mtlA2	ABHHv1_110038	ptsH	Q783_08870	Q783_08870
CM318V1_440170	ptnC	CM197v1_820039	fruA	JQMX_v1_11940	malX	AGNSv1_360015	dhaK	BN424_982	ptnC	CAR_c21710	CAR_c21710	ABHHv1_150003	glcB	Q783_08905	Q783_08905
CM318V1_440171	ptnD	CM197v1_820040	fruA	JQMX_v1_12040	mtlA	AGNSv1_360018	dhaM	BN424_983	ptnD	CAR_c21720	levG2	ABHHv1_170020	_	Q783_08910	Q783_08910
CM318V1_440189	celB	CM197v1_820041	fruA	JQMX_v1_12055	_	AGNSv1_360019	dhaL	BN424_1001	celB	CAR_c21730	levF2	ABHHv1_220005	dhaM	Q783_08950	Q783_08950
CM318V1_470123	pts7C	CM197v1_880015	glvC	JQMX_v1_12098	ptsC	AGNSv1_360020	dhaK	BN424_1068	nagE	CAR_c22050	CAR_c22050	ABHHv1_220006	dhaL	Q783_09475	Q783_09475
CM318V1_520036	celC	CM197v1_880016	glvC	JQMX_v1_12149	_	AGNSv1_360126	crr	BN424_1329	_	CAR_c23140	fruA2	ABHHv1_220007	dhaK	Q783_09570	Q783_09570
CM318V1_530055	_	CM197v1_880020	ptbA	JQMX_v1_12160	bglF	AGNSv1_360161	manX	BN424_1812	mtlA	CAR_c23410	manP	ABHHv1_270008	fruA	Q783_09580	Q783_09580
CM318V1_540003	ptsA	CM197v1_920043	crr	JQMX_v1_12163	manZ	AGNSv1_360162	manX	BN424_1981	fruA	CAR_c25130	treB	ABHHv1_280002	mtlA	Q783_10280	Q783_10280
CM318V1_570079	malX	CM197v1_930025	ulaA	JQMX_v1_12164	manY	AGNSv1_360163	manY	BN424_2420	ptsI	CAR_50p050	CAR_50p050	ABHHv1_280004	mtlA	Q783_11370	Q783_11370
CM318V1_580008	celB	CM197v1_930027	ulaC	JQMX_v1_12165	manX	AGNSv1_360164	manZ	BN424_2421	ptsH	CAR_50p070	fruA	ABHHv1_280025	_	Q783_12000	Q783_12000
CM318V1_580012	chbA	CM197v1_950006	_	JQMX_v1_12166	manX	AGNSv1_360167	bglF	BN424_2493	agaD	CAR_50p170	CAR_50p170	ABHHv1_330010	_	Q783_12020	Q783_12020
CM318V1_590002	nagE	CM197v1_950022	mtlA	JQMX_v1_12203	crr	AGNSv1_360176	_	BN424_2494	agaC			ABHHv1_330012	_		
CM318V1_750006	_	CM197v1_1020016	celA	JQMX_v1_12283	dhaK	AGNSv1_370001	_	BN424_2496	agaB			ABHHv1_330013	_		

CM318V1_1090001	_	CM197v1_1020017	celB	JQMX_v1_12284	dhaL	AGNSv1_370035	_	BN424_2498	_	ABHHv1_330014	_
CM318V1_1160002	_	CM197v1_1070031	manX	JQMX_v1_12285	dhaM	AGNSv1_370081	fruA	BN424_2592	celC	ABHHv1_380006	treB
CM318V1_1460031	_	CM197v1_1070032	manX	JQMX_v1_12288	dhaK	AGNSv1_370082	fruA	BN424_2593	celA3	ABHHv1_410001	_
CM318V1_1460032	chbA	CM197v1_1070033	manY	JQMX_v1_12291	treB	AGNSv1_370083	fruA	BN424_2594	celA3	ABHHv1_410002	yfeU
CM318V1_1460033	chbB	CM197v1_1070034	manZ	JQMX_v1_12532	pts7C	AGNSv1_370092	_	BN424_2595	celA3	ABHHv1_500001	_
CM318V1_1460039	fruA	CM197v1_1070041	_	JQMX_v1_12627	ptcA	AGNSv1_370093	chbA	BN424_2596	celB	ABHHv1_570001	bglF
CM318V1_1460040	fruA	CM197v1_1070042	_	JQMX_v1_12674	chbB	AGNSv1_370094	chbB	BN424_2796	celC2		
CM318V1_1460041	fruA	CM197v1_1070046	bglF	JQMX_v1_12675	pts7C	AGNSv1_370100	fruA	BN424_2864	celB		
CM318V1_1470036	crr	CM197v1_1070055	_	JQMX_v1_12682	celB	AGNSv1_370101	fruA	BN424_2865	licB4		
CM318V1_1490004	celB	CM197v1_1110007	dhaK	JQMX_v1_12745	fruA	AGNSv1_370102	fruA	BN424_2972	bglF		
CM318V1_1500006	_	CM197v1_1110008	dhaL	JQMX_v1_12746	fruA	AGNSv1_390026	fruA	BN424_3002	pts3C		
CM318V1_1500021	mtlA	CM197v1_1110009	dhaM	JQMX_v1_12747	fruA	AGNSv1_400015	yfeU	BN424_3013	celB2		
CM318V1_1510032	manX	CM197v1_1110012	dhaK	JQMX_v1_12758	chbA	AGNSv1_400016	_	BN424_3033	celB		
CM318V1_1510033	manX	CM197v1_1110018	treB	JQMX_v1_12759	chbB	AGNSv1_400047	glvC	BN424_3105	fptC		
CM318V1_1510034	manY	CM197v1_1160014	scrA	JQMX_v1_12765	fruA	AGNSv1_400051	ptbA	BN424_3106	frwB		

CM318V1_1510035	manZ	CM197v1_1200001	mtlA	JQMX_v1_12766	fruA	AGNSv1_450078	chbB	BN424_3107	fruA
CM318V1_1510038	bglF	CM197v1_1200004	mtlA	JQMX_v1_12767	fruA	AGNSv1_450079	pts7C	BN424_3109	celB
CM318V1_1510042	_	CM197v1_1270009	ptsC	JQMX_v1_12849	celB	AGNSv1_480006	_	BN424_3112	celB
CM318V1_1510051	_	CM197v1_1370012	_	JQMX_v1_12871	celB	AGNSv1_480021	mtlA	BN424_3113	licB
CM318V1_1540030	bglF	CM197v1_1390001	ptsA	JQMX_v1_12909	bglF	AGNSv1_510001	mtlA	BN424_3114	licA
CM318V1_1540040	_	CM197v1_1460026	fruA	JQMX_v1_13016	celB	AGNSv1_510003	mtlA	BN424_3125	_
CM318V1_1540042	chbB	CM197v1_1490010	malX	JQMX_v1_13081	celC	AGNSv1_600012	ulaC	BN424_3127	fruA
CM318V1_1540043	_	CM197v1_1590001	celB	JQMX_v1_13279	celB	AGNSv1_670034	bglF	BN424_3191	celB
CM318V1_1540044	_	CM197v1_1770024	yfeU	JQMX_v1_13283	chbA	AGNSv1_680007	ptsA	BN424_3199	celA
CM318V1_1560011	glvC	CM197v1_1770025	_	JQMX_v1_13403	_	AGNSv1_690014	celB	BN424_3247	ptcA
CM318V1_1560015	ptbA	CM197v1_2220002	nagE	JQMX_v1_13405	agaV	AGNSv1_690039	celB	BN424_3249	licB2
CM318V1_1570082	chbB	CM197v1_2230039	bglF	JQMX_v1_13406	agaC	AGNSv1_710018	_	BN424_3360	pts7C
CM318V1_1570083	pts7C	CM197v1_2260001	agaC	JQMX_v1_13407	agaD	AGNSv1_710020	agaV	BN424_3612	treP
CM318V1_1580024	celB	CM197v1_2260002	agaD	JQMX_v1_13473	ptsH	AGNSv1_710021	agaC	BN424_3629	dhaM
				JQMX_v1_13474	ptsI	AGNSv1_710022	agaD	BN424_3630	dhaL
				JQMX_v1_50021	lacE	AGNSv1_740079	malX	BN424_3631	dhaK
				JQMX_v1_50031	lacE			LMAv2_mp0021	–
				JQMX_v1_50032	lacF			LMAv2_mp0022	–
				JQMX_v1_50024	lacF			LMAv2_mp0085	lacE
								LMAv2_mp0095	lacE

Résumé

Analyse par génomique comparée de *Carnobacterium maltaromaticum* :
Etude de la diversité et de l'adaptation à différents environnements

Carnobacterium est un genre bactérien ubiquiste appartenant au groupe des bactéries lactiques (LAB). Les souches de *Carnobacterium maltaromaticum* sont présentes dans une grande diversité d'environnements et de produits alimentaires. Elles présentent de nombreuses propriétés, dont un effet positif en fabrication fromagère et des potentialités probiotiques qui en font un microorganisme à forte potentialité industrielle. L'objectif de ce travail de thèse est d'approfondir les connaissances sur *C. maltaromaticum* par une étude génomique. Neuf génomes de bactéries du genre *Carnobacterium,* dont 5 de *C. maltaromaticum,* disponibles sur la plateforme MicroScope ont été analysés et comparés. *Carnobacterium maltaromaticum* DSM20342 MX5 possède un génome de 3,85 Mbp, le plus grand génome parmi les LAB connus, de 1 Mbp de plus que celui des autres *Carnobacterium* n'appartenant pas à l'espèce *C. maltaromaticum*. L'analyse génomique détaillée indique la présence de quelques voies métaboliques spécifiques et suggère que les souches de *C. maltaromaticum* et *C. divergens* présenteraient une surface cellulaire caractérisée par une grande diversité moléculaire.Les trois autres souches de *Carnobacterium* présenteraient une surface fortement dépourvue en molécules de surfaces dont les protéines. Les propriétés de surface qui en découlent seraient à relier à la capacité de ces deux espèces à coloniser différents habitats. Par ailleurs, les souches de *C. maltaromaticum* isolées de différents produits laitiers se caractérisent par une grande diversité génomique. Afin de caractériser le niveau de diversité, les gènes impliqués dans les voies métaboliques d'utilisation du lactose (voie du Tagatose-6Phosphate, gènes *lac*) et du galactose (voie de Leloir, gènes *gal*) ont été caractérisés par analyse génomique au sein du genre *Carnobacterium.* De plus, ces gènes ont été recherchés par PCR chez 42 souches de *C. maltaromaticum*. Les deux voies sont présentes et organisées différemment chez les différentes souches de *C. maltaromaticum* d'origine laitière et non laitière L'analyse de ces gènes au niveau de la population révèle une dissémination des deux voies au sein des 4 lignées de cette population. La voie de Leloir est présente dans les lignées I, II et III et la voie du Tagatose-6Phosphate dans les lignées I, II et IV. Ces résultats suggèrent que les gènes *lac* et *gal* ont évolué selon un schéma complexe, qui est le reflet du haut niveau de diversité génétique de la population. La recherche de ces gènes au sein des 237 génomes de LAB disponibles sur la plateforme MicroScope Mage indique que les gènes *lac* et *gal* présentent un niveau de diversité élevé parmi les différents genres bactériens constituant le groupe des LAB, avec une dominance de la présence de la voie de Leloir. Ce niveau de diversité génétique n'est retrouvé qu'à l'échelle du genre chez les LAB et non de l'espèce, comme c'est le cas de *C. maltaromaticum*.

Mots-clés : *Carnobacterium maltaromaticum*, diversité, adaptation, comparaison génomique, lactose.

Summary

Comparative genomic analysis of *Carnobacterium maltaromaticum*:
Study of diversity and adaptation to different environments

Carnobacterium bacteria belong to the lactic acid bacteria (LAB) and are ubiquitous. *Carnobacterium maltaromaticum* strains are present in a wide variety of environments and foods. They have many positive properties in cheese manufacturing that make them microorganisms of industrial interest. The objective of this thesis is to deepen the knowledge on *C. maltaromaticum* by a comparative genomic approach. Nine *Carnobacterium* genomes, including 5 *C. maltaromaticum*, available on the platform MicroScope were analyzed and compared. *Carnobacterium maltaromaticum* DSM20342 MX5 is the largest genome among LAB, 1Mbp more than other *Carnobacterium* species. Detailed analysis indicates the presence of some specific metabolic pathways, and the presence of a high molecular diversity of cell surface proteins in *C. maltaromaticum* and *C. divergens*, while the three other *Carnobacterium* species present only few proteins on their surfaces. This can be related to the ability of these two species to colonize different habitats. Furthermore, *C. maltaromaticum* strains isolated from various dairy products are characterized by a large genomic diversity. In order to characterize this diversity, genes involved in lactose (Tagatose-6Phosphate pathway, *lac* genes) and galactose (Leloir pathway, *gal* genes) metabolic pathways were identified by genomic analysis within *Carnobacterium* genus and detected by PCR in 42 *C. maltaromaticum* strains. Both pathways are present and organized differently in the *C. maltaromaticum* strains. The analysis of these genes at the population level shows a spread of the two pathways within the 4 lineages of this population. The Leloir pathway is present in the lineages I, II and III and the Tagatose-6Phosphate in the lineages I, II and IV. These results suggest that the *lac* and *gal* genes have evolved in a complex pattern, which reflects the high level of genetic diversity of the population. These genes were characterized in 237 LAB genome and exhibit a high degree of diversity among the different bacterial genera of the LAB group, with a dominance of the presence of the Leloir pathway. This level of genetic diversity is found at the genus level in LAB and not at the species level, as for *C. maltaromaticum*.

Keywords: *Carnobacterium maltaromaticum*, diversity, adaptation, comparison genomics, lactose.