

SOMMAIRE

SOMMAIRE

Introduction	1
1. Problématique	2
2. Objectifs de la thèse	3
3. Stratégie d'étude	3
Chapitre 1 : Synthèse bibliographique	5
1. Bactéries lactiques :	7
1.1. Description et utilisation	7
1.2. Taxonomie et classification phylogénétique	9
1.3. Le genre <i>Lactococcus</i>	12
2. Génomique des bactéries lactiques	16
2.1. Historique et caractéristiques générales des génomes des LAB	16
2.2. Apports de la génomique sur l'étude des bactéries lactiques	19
2.2.1. Informations issues d'un génome	19
2.2.2. Structure du génome	19
2.2.3. Capacités métaboliques	21
2.2.4. Comparaison de génomes	22
2.3. Génome de <i>Lactococcus lactis</i>	24
3. Les technologies NGS « Next Generation Sequencing »	28
3.1. Bioinformatique et biologie computationnelle	28
3.2. Le processus de séquençage	28
3.3. La méthode de Sanger	29
3.4. Automatisation avec le séquenceur 3730xl	31
3.5. La prochaine génération	32
3.5.1. Le pyroséquençage 454	32
3.5.2. Le séquençage Illumina	33
3.6. Autres technologies	34
3.6.1. Helicos	34
3.6.2. Pacific Biosciences	34
3.6.3. Oxford Nanopore	34

Sommaire

3.6.4. Ion Torrent Ion Torrent	34
3.6.5. Moleculo	35
3.7. La métagénomique	37
4. Séquençage des génomes de LAB	38
4.1. Analyse bio-informatique des génomes et bases de données	38
4.1.1. Outils d'analyse et d'annotation des séquences	38
4.1.2. Bases de données	39
4.2. Intérêts du séquençage des génomes de <i>Lactococcus lactis</i>	40
5. Réponses au stress chez les bactéries lactiques	41
5.1. La réponse générale au stress	41
5.2. Le stimulon de réponse au stress thermique	42
5.3. Les BL et la régulation des réponses aux stress	44
5.4. Le stress thermique Chez les BL	45
5.5. <i>Lactococcus lactis</i> et le stress hyperthermique	48
5.6. Les protéines de choc thermique chez <i>Lactococcus lactis</i> : synthèse et régulation; thermotolérance	51
5.6.1. Introduction	51
5.6.2. Les protéines du choc thermique	53
5.6.2.1. Les chaperonines : dissociation des agrégats et remodelage des protéines	54
5.6.2.2. La protéolyse ATP-dépendante chez les bactéries	54
5.6.2.3. Identification des substrats	55
5.6.2.4. Reconnaissance et fixation du substrat par l'ATPase	56
5.6.2.5. Transport du substrat jusqu'au site de protéolyse	56
5.6.2.6. Protéolyse et évacuation des produits de dégradation	56
5.6.2.7. Les principales protéases ATP-dépendantes bactériennes	57
5.6.2.7.1. La sérine-protéase Lon	57
5.6.2.7.2. La métalloprotéase FtsH	57
5.6.2.7.3. Les complexes protéasiques de ClpP	57
5.6.2.7.4. ClpP	57
5.6.2.7.5. Les chaperonines-ATPases partenaires de ClpP	58
5.7. Régulation de la synthèse des protéines de choc thermique	60
5.8. La thermotolérance	66

Sommaire

5.9. Conclusion	68
Chapitre 2 : Matériels et méthodes	70
1. Matériel biologique	70
2. Test de pureté	71
2.1. Culture sur milieu M17 liquide	71
2.2. Test de pureté sur milieu M17 solide et microscopie	71
2.2.1. Coloration de Gram	72
2.3. Conservation sur glycérol 30% stérile	72
3. Test de thermotolérance	72
4. Analyse physiologique	73
4.1. Test de Protéolyse en milieu solide Agar-lait	73
4.2. Conservation dans le lait	73
4.3. Screening des lactocoques thermotolérants	73
4.4. Etude de la cinétique de croissance en milieu M17 (à 30°C et 50°C)	73
4.5. Etude de la cinétique de croissance en milieu lait (à 30°C et 50°C)	73
4.6. Croissance sur différents milieux de cultures	74
4.7. Traitement par l'ultrasonification	74
5. Analyse génétique	74
5.1. Obtention des variants	74
5.2. Analyse des variants	75
5.2.1. Test de thermotolérance des variants	75
5.2.2. Test d'activité protéolytique et d'hydrolyse du lactose	75
5.3. Curage d'ADN plasmidique et transformation par électroporation	76
5.3.1. Curage	76
5.3.2. Transformation bactérienne par électroporation	76
5.3.2.1. Extraction de l'ADN plasmidique	76
5.3.2.2. Dosage et contrôle de pureté de l'ADN	77
5.3.2.3. Préparation des cellules compétentes	77
5.3.3. Conditions de transformation	77
5.4. Extraction de l'ADN génomique	78
5.4.1. Préparation de l'ADN génomique « à la protéinase K »	78
5.4.2. Extraction de l'ADN génomique par KIT	79

Sommaire

5.5. Amplification par PCR	79
5.6. Amplification des gènes de choc thermique par RT-PCR	80
5.6.1. Isolement de l'ARN	81
5.6.2. Synthèse du cDNA et RT-PCR	81
5.7. Séquençage	82
5.7.1. Séquençage partiel de l'ADNr 16S	82
5.7.2. Séquençage du gène <i>gdpP</i>	83
5.8. Construction phylogénétique	86
5.9. Séquençage du génome total d'un lactocoque thermotolerant	87
5.9.1. Choix de la souche	87
5.9.2. Stratégie de séquençage envisagée	87
5.9.3. Préparation de l'ADN génomique pour le séquençage haut-débit Illumina	88
5.9.3.1. Préparation de la librairie	88
5.9.4. Principe du séquençage par la technique Illumina (séquençage par terminaison cyclique réversible)	91
5.10. Analyse bio-informatique des données de séquençage	94
5.10.1. Types d'analyse	95
5.10.2. Alignement	95
5.10.3. Construction de gabarits « <i>Scaffolding</i> »	95
5.10.4. Estimation de la qualité globale des lectures et filtration des lectures	95
5.10.5. L'assemblage des données issues du séquençage nouvelle génération	96
5.10.5.1. Les difficultés de l'assemblage	96
5.10.5.2. Les méthodes d'assemblage des séquences	99
5.10.5.2.1. La méthode « <i>overlap-layout-consensus</i> »	99
5.10.5.2.2. L'approche du graphique « <i>De Bruijn</i> »	101
5.10.5.3. Vue d'ensemble des assembleurs et leur utilisation	103
5.11. Assemblage génomique et méta-génomique	105
5.11.1. Assembleurs utilisés	105
5.11.2. Annotation du génome et génomique fonctionnelle	106

Sommaire

5.11.3. Analyse méta-génomique	106
5.12. La microscopie électronique à balayage (MEB/SEM)	107
5.12.1. Principe de la microscopie électronique à balayage	107
5.12.1.1. Interactions du faisceau électronique avec l'échantillon	107
5.12.1.2. Émission d'électrons secondaires	108
5.12.1.3. Émission d'électrons rétrodiffusés	108
5.12.1.4. Émission de rayons X	109
5.12.1.5. Émission d'électrons Auger	109
5.12.1.6. Cathodoluminescence	109
5.12.1.7. Canalisation d'électrons	109
5.12.2. Équipement	111
5.12.3. Préparation des échantillons pour la microscopie électronique à balayage MEB	113
5.12.3.1. A partir du milieu liquide	113
5.12.3.2. A partir du milieu solide	114
5.12.3.3. Conditions opératoires	114
5.13. Ressources bioinformatiques en ligne ou logiciels utilisés	116

Chapitre 3 : Résultats et discussion

Partie I : Diversité naturelle des lactocoques dans le lait de chamelle d'Algérie	119
1. Introduction	119
2. Etude de la diversité des isolats	121
2.1. Caractéristiques morphologiques	121
2.1.2. Caractéristiques macroscopiques	121
2.1.3. Caractéristiques microscopiques	121
2.2. Test de la thermotolérance	122
3. Analyse physiologique	125
3.1. Test de Protéolyse en milieu solide Agar-lait	125
3.2. Screening des lactocoques thermotolérants	126
3.3. Etude de la cinétique de croissance en milieu M17 (à 30°C et 50°C)	128
3.4. Etude de la cinétique de croissance en milieu lait (à 30°C et 50°C)	130

Sommaire

4. Analyse génétique	135
4.1. Test de thermotolérance des variants	135
4.2. Test d'activité protéolytique et d'hydrolyse du lactose	135
4.3. Expression des gènes de choc thermique par RT-PCR	136
5. Analyse phylogénétique	140
5.1. Séquençage de l'ADNr 16S: le lait de chamelle présente une diversité naturelle parmi les sous espèces de <i>L. lactis</i>	141
5.2. Séquençage de <i>gdpP</i> : Variabilité génétique du gène <i>gdpP</i> parmi les sous espèces de <i>lactis</i> et signes de transfert horizontal de gènes	145
5.3. Lait de chamelle, une niche écologique propice au transfert entre les sous espèces <i>Lactococcus lactis</i>	148
5.4. <i>L. lactis</i> subsp <i>lactis</i> biovar <i>diacetylactis</i> HD20A est une chimère entre <i>ssp. cremoris</i> et <i>ssp. lactis</i>	150
Partie II : Analyse du génome de <i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> HD9B_GL2	153
1. Introduction	153
2. Définitions importantes	154
3. Analyse des données de séquençage	156
3.1. Vérification de la qualité des séquences	157
3.2. Suppression des adaptateurs/primers	157
3.3. Filtres des séquences	157
3.4. Assemblage des séquences	158
3.4.1. Assemblage de novo	159
3.4.2. <i>Mapping</i> des données sur une référence	159
3.5. Construction de gabarits « Scaffolding »	159
3.6. Annotation	159
3.7. Création de Workflows	159
4. Analyse des données génomique de <i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> HD9B_GL2	161
4.1. Phages et plasmides dans la HD9B_GL2	166
4.2. Diversité de la robustesse chez <i>Lactococcus lactis</i> HD9B_GL2 (<i>robustness related genes</i>)	171

Sommaire

4.3. Analyse de l'organisation génétique des gènes de choc thermique chez <i>L. lactis</i> ssp. <i>lactis</i> HD9B_GL2	174
Partie III : Analyse métagénomique d'un consortium bactérien « <i>Lactococcus</i>, <i>Leuconostoc</i> et <i>Enterococcus</i> » présent dans l'écosystème lait de chamelle	194
1. Introduction	194
2. Analyse de la micro-communauté 9B	196
2.1. Identification des trois espèces du « consortium »	196
2.2. Séparation des trois espèces	206
2.3. Caractérisation des trois espèces séparées du consortium microbien	211
2.3.1. Cinétique de croissance à différentes températures	211
2.3.2. Propriétés importantes	213
2.3.3. Observations morphologiques en microscopie électronique à balayage	220
3. Caractérisation moléculaire de la Biogenèse des pilis chez <i>Lactococcus lactis</i> HD9B_GL2 et leur rôle dans l'agrégation et la Formation de Biofilms	225
4. Analyse moléculaire d'un nouveau cluster de gènes <i>eps</i> chez <i>L. lactis</i> HD9B_GL2 et son investigation dans la micro-agrégation	230
Conclusion et perspectives	242
Références bibliographique	247
Annexes	291
Résumé	

Liste des tableaux

Tableau 1 :	Quelques utilisations des bactéries lactiques dans l'alimentation	08
Tableau 2 :	Bactéries lactiques dont les génomes sont complètement ou partiellement séquencés.	18
Tableau 3 :	Souches de ssp. <i>lactis</i> dont le génome total est complètement séquencé	25
Tableau 4 :	Caractéristiques des technologies de séquençage	36
Tableau 5 :	Souches utilisées dans ce travail	70
Tableau 6 :	Programmes d'amplification PCR/RT-PCR de différents gènes étudiés	84
Tableau 7 :	Amorces utilisées dans l'amplification et le séquençage des différents gènes étudiés	85
Tableau 8 :	Logiciels et programmes utilisés dans ce travail de thèse	116
Tableau 9 :	Précriblage de souches thermotolérantes à croissance rapide apparente	127
Tableau 10 :	Résultat de la cinétique de croissance en milieu M17 (30°C)	128
Tableau 11 :	Résultat de la cinétique de croissance en milieu M17 (50°C)	129
Tableau 12 :	Résultat de la cinétique de croissance en milieu lait (30°C)	132
Tableau 13 :	Résultat de la cinétique de croissance en milieu lait (50°C)	133
Tableau 14 :	Récapitulation des données de la cinétique	134
Tableau 15 :	Comparaison des différents assembleurs utilisés	162
Tableau 16 :	Caractéristiques générales du génome de <i>Lactococcus lactis</i> HD9B_GL2	163
Tableau 17 :	Gènes associés à la robustesse sous un stress thermique chez <i>L. lactis</i> ssp. <i>lactis</i> HD9B_GL2	173
Tableau 18 :	Gènes de choc thermique présents chez <i>Lactococcus lactis</i> . ssp. <i>lactis</i> HD9B_GL2	176
Tableau 19 :	Principales caractéristiques des gènes de choc thermique clonés et séquencés chez <i>Lactococcus lactis</i> ssp. <i>lactis</i> HD9B_GL2.	186

Tableau 20 :	Séquences nucléotidiques remarquables et conservées chez les gènes de choc thermique de <i>Lactococcus lactis</i> ssp. <i>lactis</i> biovar. <i>diacetylactis</i> HD9B_GL2.	187
Tableau 21:	Informations génomiques générales des trois espèces isolées du consortium: <i>Leuconostoc mesenteriodes</i> GL1, <i>Lactococcus lacti</i> ssp. <i>Lactis</i> GL2 et <i>Enterococcus faecium</i> GL3	202
Tableau 22 :	Analyse physiologique des trois composantes de l' isolat protéolytique HD9B	219
Tableau 23 :	Piline chez HD9B_GL2	226
Tableau 24 :	Propriétés des protéines codées par le cluster des gènes EPS chez <i>L. lactis</i> ssp. <i>lactis</i> GL2	233

Liste des figures

Figure 01 :	Organigramme de la stratégie d'étude	06
Figure 02 :	Arbre consensus, basé sur l'analyse comparative de l'ADN ribosomiques (ADNr) montrant les différents groupes phylogénétiques des bactéries lactiques à faible pourcentage en guanine et cytosine (GC%) et les genres Gram positif non reliés <i>Propionibacterium</i> et <i>Bifidobacterium</i> .	12
Figure 03:	Arbre phylogénétique des souches de <i>Lactococcus lactis</i> dont le génome est complètement séquencé.	15
Figure 04 :	Un séquenceur d'ADN.	30
Figure 05 :	Principe du séquençage selon la méthode de Sanger	31
Figure 06 :	Comparaison des régulateurs d'HrcA et CtsR chez différents firmicutes	44
Figure 07 :	Séquences CIRCE putatives détectées en amont des gènes codant les protéines de choc thermique.	47
Figure 08 :	Modèle de la régulation des opérons de réponse au choc thermique chez les lactobacilles.	48
Figure 09 :	Régulation de l'expression des gènes de choc thermique chez <i>Escherichia coli</i> et <i>Bacillus subtilis</i> .	64
Figure 10:	Protocole de préparation de la librairie « <i>paired-end</i> » selon Illumina	90
Figure 11 :	Amplification en pont de la technologie Illumina	92
Figure 12:	Aperçu de la technologie de séquençage Illumina	93
Figure 13 :	Schématisation du comportement d'un assembleur vis-à-vis d'une répétition.	98
Figure 14 :	Schématisation du principe de fonctionnement d'un assembleur utilisant la méthode <i>Overlap-Layout-Consensus</i> .	100
Figure 15 :	Représentation schématique de l'assemblage de deux séquences par la méthode graphique « <i>De Bruijn</i> ».	102
Figure 16:	Logiciels intégrés dans le pipeline d'assemblage et l'annotation génomique et méta-génomique de la composante HD9B	104

Figure 17 :	Représentation schématique de l'interaction entre un faisceau d'électron et la surface d'un échantillon	108
Figure 18 :	Poire de diffusion détaille un spectre de distribution d'énergie selon le type d'électrons émis.	110
Figure 19:	Représentation schématique de l'énergie de distribution des électrons émis par un échantillon	110
Figure 20:	Représentation schématique d'un microscope électronique à balayage équipé d'un système de microanalyse de rayons X	112
Figure 21:	JEOL 7500F	115
Figure 22 :	Aspect macroscopique des colonies de lactocoque après 18 heures de culture sur M17 à 30°C.	121
Figure 23 :	Caractéristiques microscopiques des isolats étudiés après coloration de Gram. [Grossissement x100]	121
Figure 24 :	Histogramme présentant le screening des lactocoques thermotolérants	123
Figure 25 :	Lactocoques ensemencés par la méthode des patchs	125
Figure 26 :	Résultats de la protéolyse	125
Figure 27 :	Comparaison des cinétiques de croissance à 30 et 50°C en milieu M17 des Lactocoques issus du lait cru de chamelle ou du lait de chèvre	130
Figure 28 :	Comparaison des cinétiques de croissance à 30 et 50°C en milieu lait des lactocoques issus du lait cru de chamelle ou du lait de chèvre	131
Figure 29:	Résultat du test de l'activité protéolytique de souches testées en milieu Lac-MAB.	136
Figure 30 :	Expression des principaux gènes de choc thermique appartenant à différents régulons.	137
Figure 31 :	Alignement des séquences partielles des gènes codant pour l'ARNr 16S des souches de référence utilisées pour dessiner les amorces.	143
Figure 32 :	Arbre phylogénétique basé sur 16S DNAr	144
Figure 33:	Arbre phylogénétique basé sur le gène <i>gdpP</i>	147
Figure 34:	Alignement des séquences partielles du gène <i>gdpP</i> (<i>recombination breakpoint region</i>).	151
Figure 35:	Séquençage en « paire end reads » selon Illumina	156
Figure 36 :	Interface utilisateur de FastQC	158
Figure 37 :	<i>Workflow</i> : Pipeline de l'assemblage et d'annotation.	160

Figure 38 :	Arbre phylogénétique basé sur les gènes marqueurs 16s RNA, <i>rpoA</i> et la protéine L2	165
Figure 39 :	Phages dans <i>L. lactis</i> ssp. <i>lactis</i> HD9B_GL2.	167
Figure 40 :	Organisations génomique des phages tempérés de <i>L. lactis</i> ssp. <i>lactis</i> HD9B_GL2	168
Figure 41:	A. Organisation transcriptionnelle des opérons <i>hrcA-grpE-dnaK</i> , <i>groES-groEL</i> et <i>dnaJ</i> chez <i>L. lactis</i> HD9B_GL2. B. Alignement de séquence des acides aminés de l'extrémité N-terminale de la protéine HrcA chez <i>L. lactis</i> ssp. <i>lactis</i> et ssp. <i>cremoris</i> , <i>Streptococcus mutans</i> UA159 et chez <i>B. subtilis</i> 168. C. Région 5'-en amont du gene <i>hrcA</i> chez <i>L. lactis</i> ssp. <i>lactis</i> et chez ssp. <i>cremoris</i> . D. Région 5'-en amont du gène <i>hrcA</i> chez <i>Streptococcus mutans</i> UA159. E. Structures secondaires prédites de la séquence UTR non traduite "leader" de <i>hrcA</i> dans <i>L. lactis</i> ssp. <i>lactis</i> .	181
Figure 42:	A. Organisation transcriptionnelle de l'opéron <i>ctsR-clpC</i> . B. Alignement de la séquence N-terminale de la protéine CtsR des quatre souches <i>L. lactis</i> et de <i>Streptococcus mutants</i> UA159 / <i>B. subtilis</i> 168. C. Région 5'-en amont du gène <i>ctsR</i> chez <i>L. lactis</i> ssp. <i>lactis</i> et chez ssp. <i>cremoris</i> . D. Alignement des régions de RecA1 et RecA2 de HD9B_GL2 contenant les motifs conservés Walker A et Walker B. E. Présence ou absence de RecA1 et RecA2 dans différentes souches de <i>L. lactis</i> .	185
Figure 43:	Comparaison des régulons d'HrcA et CtsR chez différents firmicutes	192
Figure 44:	Identification des trois espèces composantes du « consortium » lait. L'arbre phylogénétique est construit par le logiciel Megan5.	199
Figure 45:	Arbre phylogénétique basé sur 16S rRNA.	200
Figure 46:	Arbre phylogénétique visualisé par iTOL	201
Figure 47 :	(a) Comparaison de la qualité de l'assemblage chez <i>Leuconostoc mesenteroides</i> GL1	203
	(b) Comparaison de la qualité de l'assemblage chez <i>Lactococcus lactis</i> GL2	204
	(c) Comparaison de la qualité de l'assemblage chez <i>Enterococcus faecium</i> GL3	205

Figure 48:	Sites des amorces dans <i>rpoB</i> . Les séquences sont issues du projet de séquençage total de la composante HD9B	209
Figure 49 :	Amplification du gène <i>rpoB</i> chez les différentes souches de LAB	210
Figure 50 :	Croissance en milieu M17-glucose liquide á différentes températures.	212
Figure 51 :	Activité protéolytique du « consortium » HD_9B et ses trois espèces composantes GL1, GL2 et GL3 ensemencées sur milieu Lait glucosé.	214
Figure 52 :	Croissance sur milieu McDonald's HHD	216
Figure 53:	<i>Lactococcus lactis</i> subsp. <i>Lactis</i> DSM 20481 observée sous le microscope électronique à balayage (sur filtre de Nitrocellulose)	222
Figure 54:	<i>Lactococcus lactis</i> subsp. <i>Lactis</i> HD9B_GL2 observée sous le microscope électronique à balayage (sur filtre de Nitrocellulose)	223
Figure 55:	Observation au microscope électronique à balayage des trois constituants du consortium. A : <i>Leuconostoc mesenteroides</i> GL1. B : <i>Enterococcus faecium</i> GL3. C : <i>Lactococcus lactis</i> GL2. D : <i>Lactococcus lactis</i> DSM 20481. E et F: HD_9B mixture des trois espèces.	224
Figure 56:	Représentation schématique du cluster de pilis chez <i>L. lactis</i> HD9B_GL2.	230
Figure 57:	Représentation schématique du cluster de gènes <i>eps</i> chez <i>L. lactis</i> HD9B_GL2.	232
Figure 58 :	Comparaison de l'usage de codons dans la séquence génomique de HD9B_GL2.	237

Abréviations

ADN : Acide désoxyribonucléique

ARN : Acide ribonucléique

ARNt : Acide ribonucléique de transfert

ARNr : Acide ribonucléique ribosomal

ATCC : American Type Culture Collection (Collection Américaine de Cultures)

ATP : Adénosine tri-phosphate

BLAST : Basic local alignmentsearchtool (outil de recherche d'alignement local)

CDS : Coding DNA sequence (séquence codante d'ADN)

dNTP : Désoxyribonucléotidetri-phosphate

ddNTP : Didésoxyribonucléotidetri-phosphate

DSM : Deutsche Sammlung von Mikroorganismen (Collection Allemande de
d Microorganismes)

kb : kilobase

Mb : Mégabase

pb : paire de base

ORF : Open reading frame (phase ouverte de lecture)

PCR : Polymerasechainreaction (réaction de polymérisation en chaine)

RBS : Ribosome Binding Site

RT : Reverse Transcription

ssp. : sous-espèce

PM : Poids Moléculaire

ADNc : ADN complémentaire

AA : Acide aminé

ARNm : Acide ribonucléique messenger

BL : Bactérie Lactique

Da : Dalton

EDTA : éthylène diamine tétra acétique

L. lactis: *Lactococcuslactis*

ORF: phase ouverte de lecture (open reading frame)

RT : Reverse Transcription

SDS : sodium dodécil sulfate

ssp. : Sous-espèce

Tris : Tris-(hydrométhyl)aminométhane

UFC : Unité formant colonie

Nt : nucléotide

INTRODUCTION

Introduction

Depuis plus de quatre mille ans, l'homme utilise de façon empirique des bactéries lactiques pour la fermentation d'aliments. C'est en 1873, que Joseph Lister un élève de Pasteur a le premier isolé un ferment à partir d'un lait fermenté. Cette bactérie d'abord appelée *Bacterium lactis* est aujourd'hui nommée *Lactococcus lactis* (Ludwig et al. 1985).

Les bactéries lactiques doivent leur nom à la production massive d'acide lactique ; elles contribuent à la texture et à la saveur des aliments ainsi qu'à la production de composés aromatiques. Elles inhibent la prolifération de micro-organismes indésirables en abaissant le pH de leur environnement, mais aussi en produisant des composés inhibiteurs, comme le peroxyde d'hydrogène et les bactériocines (Gilliland et al. 1985; Ruiz-Barba et al. 1991; Marty-Teyssset et al. 2000; van de Guchte et al. 2001; van de Guchte et al. 2002). Certaines espèces de bactéries lactiques produisent aussi des exopolysaccharides qui modifient l'aspect et la texture des produits fermentés (Peant et al. 2005; Van der Meulen et al. 2007). Ainsi, ces microorganismes largement utilisés dans l'alimentation (Tableau 1) sont considérés comme non pathogènes et sont reconnus au niveau international comme « GRAS » pour “*Generally Regarded As Safe*” (Pochart et al. 1992; Aguirre and Collins 1993; Adams and Marteau 1995).

La technologie laitière représente le principal secteur d'application des bactéries lactiques et repose principalement sur la production du lait et la sélection de ferments lactiques performants. Selon les statistiques du FAO (« Food and Agriculture Organization » : Organisation des Nations Unies pour l'alimentation et l'agriculture) en 2005, la production mondiale du lait a avoisiné 532 millions de tonnes ; l'Union européenne est le premier producteur avec 142 millions de tonnes. Vu cette importante production laitière, les industriels ont souvent recouru à des procédés physiques, essentiellement thermiques, pour préserver les qualités biologiques de cette matière première fragile. D'autres méthodes reposent sur la transformation du lait afin de prolonger sa durée de conservation et d'éliminer le risque de prolifération d'agents pathogènes. L'exemple le plus concret est la fermentation lactique qui conduit à la fabrication de laits fermentés et de fromages.

Les bactéries lactiques sont ubiquitaires et ont été isolées à partir de différentes niches écologiques comme les végétaux, la viande, le poisson, la muqueuse vaginale et le système

Introduction

digestif. Néanmoins, les habitats les plus fréquents pour les lactocoques demeurent le lait, les laits fermentés et les fromages. L'adaptation à la croissance dans le lait repose sur un équipement enzymatique approprié pour cataboliser le lactose (sucre du lait) et hydrolyser les caséines du lait. À ce jour, les bactéries lactiques laitières caractérisées et utilisées comme ferments sont en majorité issues du lait de vache d'Europe (France, Angleterre et Pays Bas) ou d'Amérique (USA et Canada). C'est pour cette raison que nous nous sommes intéressés à un autre écosystème laitier, à savoir le lait cru de chamelle du sud de l'Algérie, cette région étant caractérisée par un climat chaud et considérée comme une zone aride.

La réponse au choc thermique est un phénomène général dans le monde vivant, caractérisé par la production rapide et transitoire de protéines spécifiques à une température supra-optimale. Cette réorientation des synthèses protéiques implique l'adoption par la cellule d'un programme d'expression génique original : répression des gènes « constitutifs » et activation des gènes de choc thermique. L'implication simultanée de processus de régulation transcriptionnelle, post-transcriptionnelle et traductionnelle souligne le dynamisme de cette réponse.

L. lactis est soumis à de fortes variations de températures dans le sol ou lors de la fabrication de fromages, pour lequel la température monte jusqu'à 40°C. Les mécanismes de perception du stress, essentiels pour que la cellule puisse se protéger avant que les effets de ce stress ne soient irréversibles, sont encore peu connus chez *Lactococcus lactis*. Évoquer la réponse au choc thermique, et plus particulièrement les hsp qui en sont la manifestation la plus flagrante, c'est décrire: (1) les circonstances de leur apparition, leurs nombre et qualités; (2) le mode de régulation de leur synthèse; (3) leur participation hypothétique au développement de la thermotolérance. Le séquençage des gènes isolés ou groupés en opérons et de leurs zones promotrices exclusivement végétatives remet en question l'universalité du seul modèle accepté ces dernières années.

1. Problématique

Au cours de la fabrication de certains fromages, les fortes températures rencontrées au cours de la « cuisson » peuvent être à l'origine d'altérations cellulaires majeures. Les protéines intracellulaires vont se dénaturer puis s'agréger, diverses macromolécules comme les ribosomes ou l'ARN vont être déstabilisées et la fluidité membranaire va augmenter posant des problèmes de perméabilité excessive et de transport. La réponse au choc thermique ou *heat shock* (HS) a été très étudiée, notamment chez les bactéries lactiques. Ce type de stress

Introduction

provoque, chez *L. lactis*, l'induction de protéines dont des protéines chaperonnes et des protéases. La régulation de ces protéines, dont différents gènes sont organisés en opérons, est assurée par divers régulateurs transcriptionnels. Étant l'un des organismes modèle dans le métabolisme microbien, *Lactococcus lactis* a été la principale cible de l'ingénierie métabolique chez les BL. Cette bactérie a de nouvelles applications biomédicales et biotechnologiques.

Les lactocoques sont réputés être des mésophiles. Néanmoins certaines souches de lactocoque issus récemment du lait cru de chamelle d'Algérie (Drici et al. 2010) sont capables de se développer dans une gamme de température de 10 à 50°C. De ce fait, ces lactocoques sont considérés comme atypiques et originaux de par de leur caractère de thermotolérance. Cette thermotérance peut s'expliquer par l'adaptation de ces lactocoques à leur écosystème saharien chaud et aride, et remet en question le caractère mésophile attribué aux lactocoques de façon générale.

2. Objectifs de la thèse

Dans ce travail de thèse, nous nous intéressons à l'étude du déterminisme génétique de la thermotolérance chez ces nouveaux lactocoques qui peuvent être utilisés comme ferments lactiques locaux adaptés au lait de chamelle et aux conditions climatiques du sud d'Algérie. Notre objectif principal est d'identifier et de caractériser les gènes de choc thermique, jusqu'à présent peu inventoriés, en particulier chez *Lactococcus lactis* et qui sont responsables de la production de protéines dites « *Hsp*: *heat shock protein* » et du caractère de la thermotolérance.

3. Stratégie d'étude :

Mon travail de thèse est constitué de cinq chapitres. Le premier chapitre correspond à une synthèse bibliographique ayant pour but de situer le travail dans son contexte scientifique. Elle comporte une étude sur la génomique des bactéries lactiques, la réponse au stress thermique chez les lactocoques, l'impact de la bioinformatique sur l'étude des bactéries lactiques, et les techniques du séquençage. L'analyse bibliographique est divisée en différentes parties. Elle commence par des définitions, taxonomie et classification puis une liste des utilisations des bactéries lactiques dans l'alimentation. Ensuite elle décrit la génomique des bactéries lactiques ainsi que leur réponse au stress. Elle détaille le mécanisme de résistance au stress thermique chez le genre *Lactococcus* ainsi que la synthèse et la régulation des protéines

Introduction

de choc thermique chez ces bactéries. Elle consacre une partie importante à la description des méthodes de séquençage des génomes bactériens et des outils de l'analyse bioinformatique des génomes séquencés offrant ainsi un large panorama de méthodes de technologies NGS « *next generation sequencing* ». Elle se termine par un bilan très utile des différentes méthodes utilisées actuellement.

Le deuxième chapitre Matériel et méthode décrit les expérimentations faites dans le cadre de l'identification des déterminants génétiques de la thermotolérance chez les souches étudiées. Ce chapitre commence par une analyse physiologique des souches, leur caractérisation phénotypique et culturale. Une analyse génétique est ainsi utilisée pour dévoiler le support de la thermotolérance des souches. Une analyse phylogénétique est effectuée pour établir les relations phylogénétiques entre les souches utilisées dans ce travail de thèse. L'approche moléculaire débute par un séquençage du génome total d'un lactocoque thermotolerant. Ensuite, il est montré comment, à partir des données brutes, il est possible de les traiter pour effectuer l'assemblage des séquences génomiques et la prédiction des gènes de choc thermique ainsi que leurs protéines fonctionnelles. Les traitements effectués paraissent judicieusement choisis en regard de l'application traitée : il s'agit tout d'abord de définir des lectures brutes issues du séquençage du génome total à partir de génomes préexistantes ou d'un assemblage de novo, de traiter les séquences, de les corriger et de vérifier leur qualité, d'établir les contigs ainsi que la construction des gabarits « *scaffolding* ». Cette analyse moléculaire se base sur une liste des logiciels bioinformatiques utilisés en cours des traitements. Enfin une analyse métagénomique et génomique comparative est réalisée afin de discuter les différentes interactions et le phénomène d'agrégation entre les différentes populations et espèces d'un même écosystème laitier « le lait de chamelle d'Algérie » et qui constituent un consortium microbien particulier. Cette partie est intéressante également et répond au problème de la prise en compte de la propriété d'agrégation due à la production de protéines particulières « *Piline* » en ce qui concerne la présence des résultats pertinents en tête de liste.

Le troisième chapitre traite ensuite les résultats obtenus tout au long de ce travail colossal de thèse. C'est en fait de nouveaux concepts moléculaire génomique et bioinformatique. Cette heuristique semble intéressante et une étude systématique des gènes de choc thermique est analysée pour la première fois en détail chez ces lactocoques, ce qui représente à nouveau un travail expérimental très important. Ce chapitre se termine en illustrant le fonctionnement des

Introduction

logiciels réalisés. On obtient alors un inventaire de gènes impliqués dans le problème du stress thermique. La conclusion dégage les perspectives importantes qui font suite à ce travail.

Le protocole d'étude est schématisé dans la figure (1).

Introduction

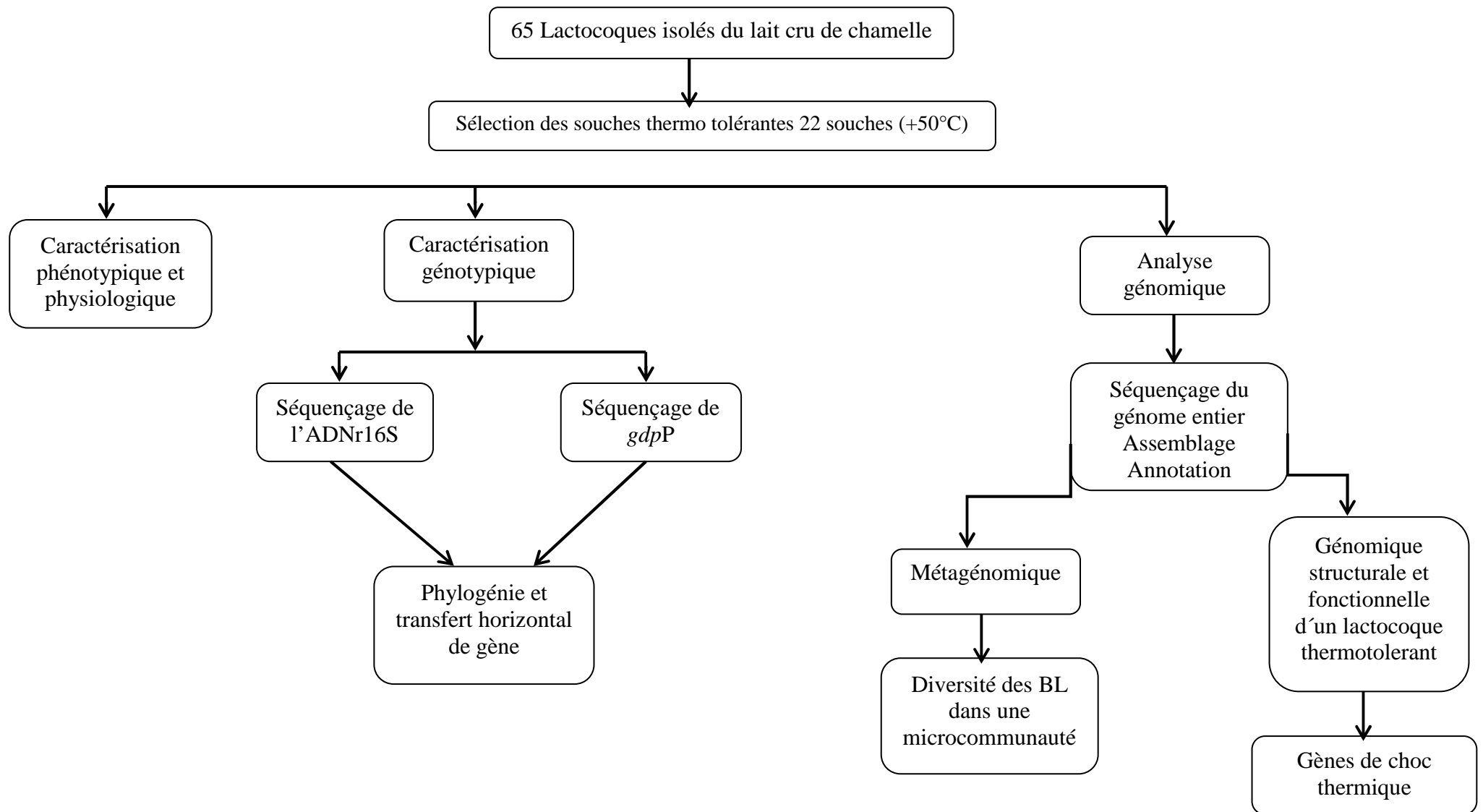


Figure 1: Organigramme de la stratégie d'étude

CHAPITRE 1

SYNTHESE BIBLIOGRAPHIQUE

Chapitre 1 : Synthèse bibliographique

1. Bactéries lactiques

Les bactéries lactiques sont des micro-organismes à Gram positif, non sporulants, nonmobiles, anaérobis mais aérotoleurants et ne possédant aucune catalase. Toutes les bactéries lactiques possèdent un métabolisme fermentaire leur permettant, en utilisant des sucres fermentescibles, de produire principalement de l'acide lactique (Halász, 2009) mais aussi d'autres acides organiques (acide acétique, acide formique...).

1.1. Description et utilisation

Les caractéristiques métaboliques des bactéries lactiques en font des acteurs indispensables au cours des fermentations alimentaires. L'utilisation de la fermentation par l'Homme remonte à des temps très anciens. Les premiers produits fermentés ont certainement été obtenus par acidification spontanée des jus végétaux (vins, bières...) ou par contamination naturelle du lait (yaourts, fromages...). Les premières preuves de l'existence des produits laitiers fermentés remontent à 8000 ans avant JC dans le croissant fertile au Moyen Orient (plaines du Nil, du Jourdain, de l'Euphrate et du Tigre), époque où les végétaux et les animaux sont domestiqués (Fox, 1993). La fermentation des végétaux (vins, bières) et la production de levain apparaissent entre 4000 et 2000 avant JC chez les Egyptiens.

Les bactéries lactiques sont classiquement impliquées dans un grand nombre de fermentations alimentaires, seules ou avec d'autres micro-organismes (transformation du lait, boissons fermentées, salaison, fermentation des végétaux), et sont également étroitement associées à l'environnement humain (Schroeter and Klaenhammer 2009) (Tableau 1). Le principal atout de ces bactéries réside donc dans leur capacité à acidifier les produits alimentaires. L'acide lactique, qui est le produit principal du métabolisme fermentaire, joue un rôle majeur dans la conservation des aliments puisqu'il inhibe fortement la croissance des bactéries pathogènes à bas pH (Stiles 1996). Il a également un rôle direct dans l'industrie laitière puisqu'il permet la formation du caillé à bas pH. Les bactéries lactiques participent également à la texture (production d'exopolysaccharides) et à la saveur des produits laitiers. Les arômes sont multiples, parfois indésirables (amines biogènes) et peuvent provenir d'origines diverses, soit du catabolisme des hydrates de carbone présents dans le lait (lactose, citrate...), soit du métabolisme des acides aminés ou encore des matières grasses (Deutsch et al. 2000).

Analyse bibliographique

**Tableau 1 : Quelques utilisations des bactéries lactiques dans l'alimentation
(d'après Atlan *et al.*, 2000)**

Utilisation	Bactéries lactiques* (et autres micro-organismes associés)
Végétaux concombres, choux olives	<i>P. acidilactici</i> et <i>P. pentosaceus</i> <i>Leu. mesenteroides</i> , <i>Lb. plantarum</i> , <i>Lb. brevis</i> , <i>Lb. sake</i> <i>Leu. mesenteroides</i>
Sauce de soja	<i>T. halophilus</i>
Viandes fermentées (saucisse) semi-conservées (sous vide)	<i>P. acidilactici</i> et <i>P. pentosaceus</i> , <i>C. divergens</i> , <i>C. piscicola</i> <i>Lb. sake</i> , <i>Lb. curvatus</i> <i>Leu. carnosum</i> et <i>Le. gelidum</i> <i>C. divergens</i> , <i>C. piscicola</i>
Poissons marinés	<i>T. halophilus</i> , <i>Lb. brevis</i> , <i>Lb. buchneri</i> , <i>Leu. mesenteroides</i>
Produits laitiers Yaourt Laits fermentés	<i>Lb. delbrueckii ssp. bulgaricus</i> et <i>S. thermophilus</i> <i>Lb. acidophilus</i> , <i>Lb. casei</i>
Fromages affinés Camembert et Brie Emmental et Gruyère	<i>L. lactis ssp. lactis</i> et <i>L. lactis ssp. cremoris</i> <i>S. thermophilus</i> et <i>Lb. helveticus</i> ou <i>Lb. delbrueckii ssp. bulgaricus</i> ou <i>Lb. delbrueckii ssp. lactis</i>
Mozzarella Cheddar	<i>S. thermophilus</i> et <i>Lb. helveticus</i> ou <i>Lb. delbrueckii ssp. bulgaricus</i> <i>L. lactis ssp. lactis</i> et <i>L. lactis ssp. cremoris</i> et <i>L. lactis ssp. lactis</i> <i>biovar. diacetylactis</i>
Roquefort	<i>S. thermophilus</i> et <i>L. lactis ssp. lactis</i> et <i>L. lactis ssp. cremoris</i> et <i>L. lactis ssp. lactis biovar. diacetylactis</i>
Boulangerie (panification)	<i>Lb. sanfrancisco</i> , <i>Lb. fermentum</i> , <i>Lb. brevis</i> <i>Lb. plantarum</i> (<i>Saccharomyces cerevisiae</i>)
Boissons vins kéfirs	<i>Oenococcus oeni</i> (<i>Saccharomyces cerevisiae</i>) <i>Lb. kefir</i> , <i>Lb. kefiranofaciens</i> , <i>Lb. kefirgranum</i> , <i>Lb. acidophilus</i> , <i>Lb. brevis</i>

* C : *Carnobacterium* ; L. : *Lactococcus* ; Leu : *Leuconostoc* ; Lb : *Lactobacillus* ; P : *Pediococcus* ; S : *Streptococcus* ; T : *Tetragenococcus*

De nos jours, les bactéries lactiques font l'objet de recherches intensives qui sont améliorées par la disponibilité de la séquence complète du génome de nombreuses bactéries lactiques : *Lactococcus lactis* IL1403 (Bolotin *et al.* 2001), *Lactobacillus plantarum* WCFS1 (Kleerebezem *et al.* 2003), *Lactobacillus johnsonii* NCC 533 (Pridmore *et al.* 2004), *Lactobacillus acidophilus* NCFM (Altermann *et al.* 2005). La production sans cesse en augmentation de produits laitiers fermentés et surtout de fromage, 18 millions de tonnes en

Analyse bibliographique

2004 alors que seulement 15 millions de tonnes étaient produites en 1999 (source statistique FAO), conduit aujourd'hui à une rationalisation indispensable de l'utilisation des ferments lactiques dans l'industrie. Les technologies laitières constituent le secteur principal d'application des bactéries lactiques. La compréhension de la physiologie de ces microorganismes contribue ainsi à un meilleur contrôle des procédés de même qu'à l'amélioration et à la diversification des qualités organoleptiques et texturales des produits laitiers fermentés. Pour cela diverses stratégies ont été mises en place, il s'agit tout d'abord d'une meilleure sélection des souches et leur utilisation en mélanges complexes dans des levains de culture. Une optimisation métabolique des souches a également été envisagée afin d'augmenter la production d'arômes (Ferain et al. 1996a; Ferain et al. 1996b; Delcour et al. 1999; Delcour et al. 2000; Kleerebezemab et al. 2000), de polysaccharides (Nyysola et al. 2005), de vitamines (Wegkamp et al. 2007), d'agents antimicrobiens (Arques et al. 2015) ou encore de la protéolyse des protéines du lait (Pontes et al. 2011)..

Les bactéries lactiques sont également impliquées dans de nouveaux types de produits en tant que « probiotiques » (Marco et al. 2010). Il s'agit de micro-organismes vivants qui une fois ingérés, vont conférer un effet physiologique bénéfique à leur hôte animal grâce à leurs propriétés microbiennes (Fuller 1991). Les genres *Lactobacillus*, *Bifidobacterium* et *Enterococcus* abritent des espèces considérées comme probiotiques (Reid et al. 2005; Goldin and Gorbach 2008). D'autres bactéries, qui ne colonisent pas naturellement le tractus digestif des mammifères, mais sont utilisées comme starters dans l'industrie laitière sont également considérées comme des probiotiques, *Lb. bulgaricus* et *St. thermophilus*.

Leur classification dans les probiotiques ainsi que parmi les microorganismes GRAS (*Generally Regarded As Safe*) fait de ces bactéries des acteurs potentiellement important dans les domaines de la médecine et de la santé : amélioration de la digestion du lactose, stimulation du système immunitaire, vecteurs de molécules à effets thérapeutiques (Amar et al. 2011) .

1.2. Taxonomie et classification phylogénétique

Orla-Jensen en 1919 (Orla-Jensen 1921) définit les bactéries lactiques comme un groupe de bactéries à Gram positif, non mobiles, asporogènes et fermentant les hydrates de carbone et certains alcools en lactate. Cependant, plus récemment, l'apport de nouvelles approches de taxonomie bactérienne (physiologie, génétique), a permis d'affiner cette classification (Stackebrandt et al. 1987; Vandamme et al. 1996). La première classification des bactéries

Analyse bibliographique

lactiques a été établie en 1919 par Orla-Jensen sur divers critères morphologiques et physiologiques (activités catalase et nitrite réductase, type de fermentation) (Stiles and Holzapfel 1997). Les critères phénotypiques permettant la classification des bactéries se sont ensuite étendus à la composition de la paroi, le type d'acides gras cellulaires, le type de quinones (accepteur d'électrons). Cependant ces méthodes phénotypiques ne rendent pas compte des relations phylogénétiques entre les groupes. En 1977, Woese et Fox (Fox et al. 1977) introduisent la phylogénie moléculaire basée sur la séquence des ARN ribosomiques. Cette méthode va révolutionner la taxonomie des bactéries et la classification des BL va être profondément modifiée. D'autres méthodes génotypiques (basés sur les acides nucléiques) sont aussi utilisées en classification, comme le pourcentage en GC ou l'hybridation ADN:ADN. L'approche consistant à prendre en compte les méthodes phénotypiques et génotypiques s'appelle la taxonomie polyphasique (Vandamme et al. 1996). La technique de MLST (pour *Multi Locus Sequence Typing*), basée sur la divergence nucléique de gènes de ménage, est utilisée pour la classification des BL pathogènes (notamment streptocoques ; (Enright and Spratt 1999). Cependant cette technique n'a pas été utilisée pour d'autres BL.

Le groupe des BL ne peut pas être considéré comme un groupe phylogénétique. Elles appartiennent toutes au groupe des bactéries à Gram positif, mais si la plupart d'entre elles appartiennent au groupe des bactéries à Gram positif à bas G+C (phylum des Firmicutes), le genre *Bifidobacterium* appartient au groupe des bactéries à Gram positif à haut G+C (phylum des Actinomycètes) (Figure 2). Si les bifidobactéries sont phylogénétiquement éloignées des BL *sensu stricto* (BL à bas G+C), elles sont tout de même traditionnellement incluses dans les BL car elles partagent certaines caractéristiques avec elles (elles produisent de l'acide lactique et sont utilisées dans les laits fermentés ; (Holzapfel et al. 2001). Parmi les bactéries lactiques *sensu stricto*, on peut distinguer quatre groupes phylogénétiques principaux. Les genres *Enterococcus*, *Melissococcus*, *Tetragenococcus*, *Vagococcus*, *Carnobacterium*, *Lactosphaera* et *Aerococcus* forment un premier groupe (Stiles and Holzapfel 1997). Dans ce groupe, l'effort de séquençage des génomes a porté sur les espèces pathogènes opportunistes de l'homme, *Enterococcus faecalis* et *E. faecium*. Les genres *Streptococcus* et *Lactococcus* n'ont été séparés qu'en 1985 avec l'apparition du genre *Lactococcus* (Schleifer et al., 1985) et sont phylogénétiquement proches de *Lactococcus lactis* est une bactérie très utilisée en fermentation fromagère. Le premier génome de BL disponible a été celui d'une souche de *Lactococcus lactis* ssp. *lactis* en 2001 (Bolotin et al. 2001). Dans le genre *Streptococcus*, la seule espèce utilisée comme ferment dans l'industrie alimentaire est *S. thermophilus*, notamment pour la production de yaourt avec *L. bulgaricus*. Ce genre contient un grand

Analyse bibliographique

nombre de bactéries plus ou moins pathogènes (*S. pneumoniae*, *S. pyogenes*, *S. mutans*, *S. agalactiae*...) pour lesquels 13 séquences génomiques sont d'ores et déjà disponibles dans les bases de données. Le troisième groupe est formé des genres *Oenococcus*, *Leuconostoc* et *Weissella*. Dans ce groupe, on retrouve des bactéries très utilisées dans l'industrie agroalimentaire comme *Oenococcus oeni*, bactérie importante dans la production de vin pour la fermentation malolactique ou *Leuconostoc citreum* utilisé pour la fermentation de certains légumes et *Lc. mesenteroides* ssp. *mesenteroides* utilisé dans la production de choucroute et de dextran.

Enfin, le dernier groupe phylogénétique est constitué des genres *Lactobacillus* et *Pediococcus*.

Les bactéries lactiques impliquées dans l'industrie alimentaire appartiennent à différents genres : *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Tetragenococcus*, *Vagococcus* et *Weissella* (Vandamme et al. 1996).

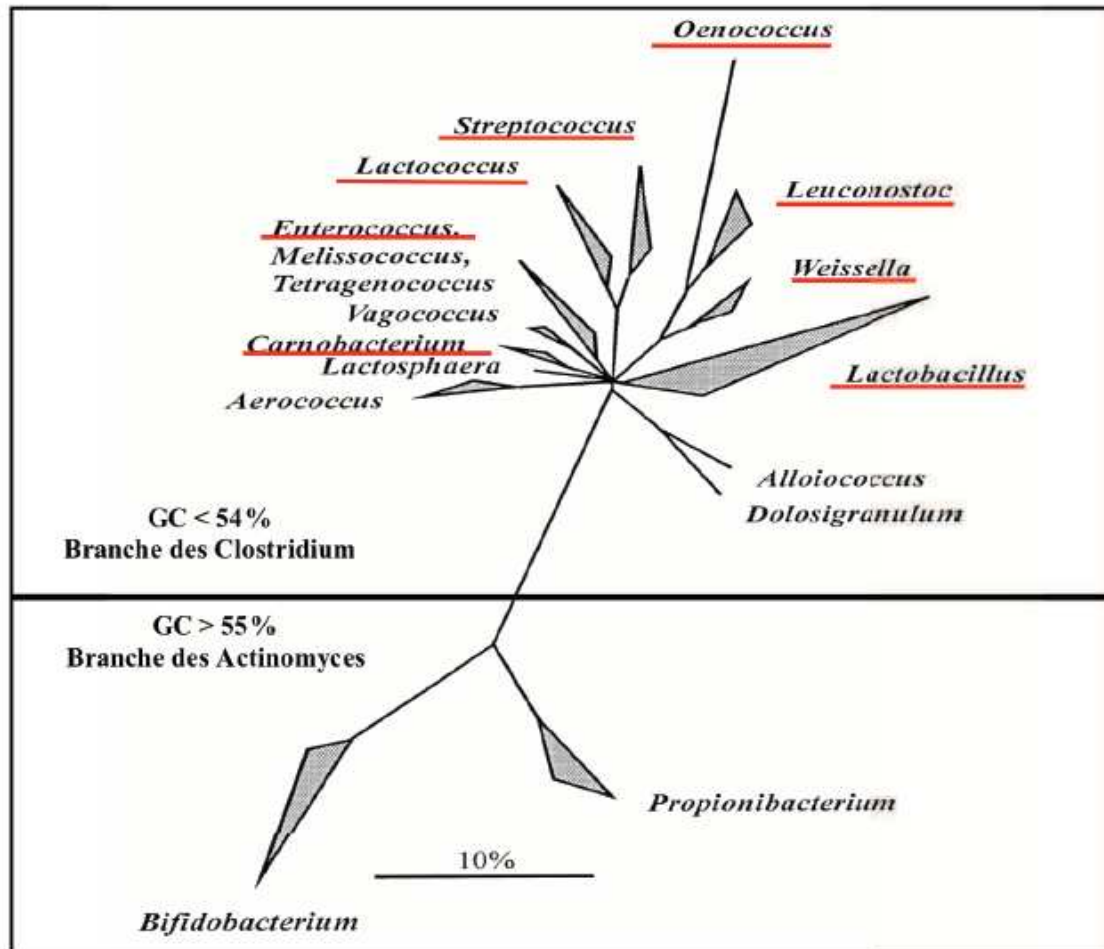


Figure 2 : Arbre consensus, basé sur l'analyse comparative de l'ADN ribosomiques (ADNr) montrant les différents groupes phylogénétiques des bactéries lactiques à faible pourcentage en guanine et cytosine (GC%) et les genres Gram positif non reliés *Propionibacterium* et *Bifidobacterium*. (d'après (Holzapfel et al. 2001). Les genres soulignés en rouge correspondent aux genres les plus utilisés en industrie alimentaire. La barre indique une divergence de séquence estimée de 10%.

1.3. Le genre *Lactococcus* :

Le genre *Lactococcus*, qui nous intéresse plus particulièrement, était autrefois apparenté aux streptocoques. Il s'agit donc de bactéries en forme de coques, regroupées ou non en chaînes de longueurs variables. Elles présentent un métabolisme homolactique, et sont mésophiles puisque leur température optimale de croissance est proche de 30°C. Ce sont les études de Schleifer et coll. (1985) qui ont justifié la création de ce nouveau genre bactérien regroupant la quasi-totalité des streptocoques du groupe N de la classification de Lancefield (Lancefield

Analyse bibliographique

and Perlmann 1952). Le genre *Lactococcus* inclut différentes espèces, *garviae*, *lactis*, *piscium*, *plantarum*, *fujimensis* et *raffinolactis*. L'espèce *lactis* regroupe elle-même trois sous-espèces, *cremoris*, *tructae*, *hordniae* et *lactis* qui elle-même comprend le biovar *diacetylactis*, capable de surproduire le diacétyl.

Les différentes sous-espèces de *Lactococcus lactis* peuvent être différenciées entre autre par leur résistance aux stress. Ainsi, la sous-espèce *lactis* est plus robuste que la *cremoris* puisqu'elle est capable de croître à 40°C, ou en présence de 4 % de NaCl. De plus, la sous-espèce *lactis* est capable de produire de l'ammoniac à partir d'arginine grâce à la voie de l'arginine déiminase (ADI) qui est très rare chez *cremoris*.

Les bactéries du genre *lactis* sont les bactéries lactiques les plus importantes d'un point de vue commercial. Les habitats les plus importants des lactocoques demeurent le lait, les laits fermentés ainsi que les fromages où ils constituent la flore dominante. Cependant, on peut également les isoler des plantes (Premi et al. 1972) ainsi que de la peau de certains animaux, et il semble que la contamination qui a lieu au cours de la traite ait pour origine principale le fourrage.

Les lactocoques sont des bactéries lactiques coccoïdes. Habituellement, ces bactéries croissent dans 4% de NaCl et à une température optimale comprise entre 20 et 30°C, elles sont capables de se développer à 10°C mais pas à 45°C (Stiles and Holzapfel 1997; Mofredj et al. 2007; Cavanagh et al. 2015a; Cavanagh et al. 2015b); Axelsson, 2004) ; elles produisent l'acide lactique de forme L(+) par métabolisme homofermentaire et elles sont couramment présentes dans les produits fermentés mais certaines espèces notamment *Lactococcus lactis* sont souvent isolées des végétaux (Stiles and Holzapfel 1997).

Le genre *Lactococcus* (Lc.) comprend plusieurs espèces : *Lc. garviae* associée à des infections animales (Stiles and Holzapfel 1997; Teuber and Perreten 2000) et humaines telles que les endocardites (James et al. 2000), *Lc. piscium* isolée à partir de poisson salmonidés, *Lc. plantarum* isolée à partir de petits pois congelés, *Lc. raffinolactis* isolée à partir de lait cru (Stiles and Holzapfel 1997) et *Lc. lactis* qui est la plus étudiée et la plus utilisée en industrie alimentaire (Stiles and Holzapfel 1997; Teuber and Perreten 2000)

L'espèce *Lc. lactis* est divisée en quatre sous- espèces : *Lc. lactis* ssp. *lactis* et *Lc. lactis* ssp. *cremoris* isolées de produits laitiers, *Lc. lactis* ssp. *hordniae* isolée de la cicadelle *Hordnia circellata* (Latorreguzman et al. 1977) et *Lc. Lactis* ssp. *tructae* isolée du tractus digestif poisson de l'eau douce (Itoi et al. 2008; Itoi et al. 2009). Les deux sous espèces *hordniae* et

Analyse bibliographique

tractae sont rarement rencontrées dans les fermentations lactières. Les souches ayant la capacité de produire le diacétyle à partir du citrate sont classées comme *Lc. lactis* ssp. *diacetylactis* (Stiles and Holzapfel 1997). *Lc. lactis* ssp. *cremoris* se distingue de *Lc. lactis* ssp. *lactis* par son incapacité à se développer à 40°C, à pH 9.2, en présence de 4% de NaCl, à hydrolyser l'arginine et à fermenter le ribose (Kelly and Ward 2002); (Rademaker et al. 2007), Schleifer et al., 1985 ; Stiles et Holzapfel, 1997 ; Pot, 2008). Ces deux sous-espèces se sont également révélées génétiquement distinctes par des études d'homologies ADN-ADN et par comparaison de leurs séquences d'ARNr 16S (Salama et al. 1991; Ward et al. 1998).

Les souches de *Lc. lactis* sont largement utilisées comme starters dans la fabrication de divers produits fermentés (Florez et al. 2008), notamment les fromages où elles interviennent dans la formation du caillé et le développement des caractéristiques organoleptiques du produit fini tout en inhibant la croissance des germes indésirables grâce aux agents antimicrobiens qu'elles produisent (Florez et al. 2008). Cependant, la principale préoccupation concernant ces souches est la stabilité de certaines de leurs caractéristiques technologiques d'origine plasmidique (Stiles and Holzapfel 1997).

Analyse bibliographique

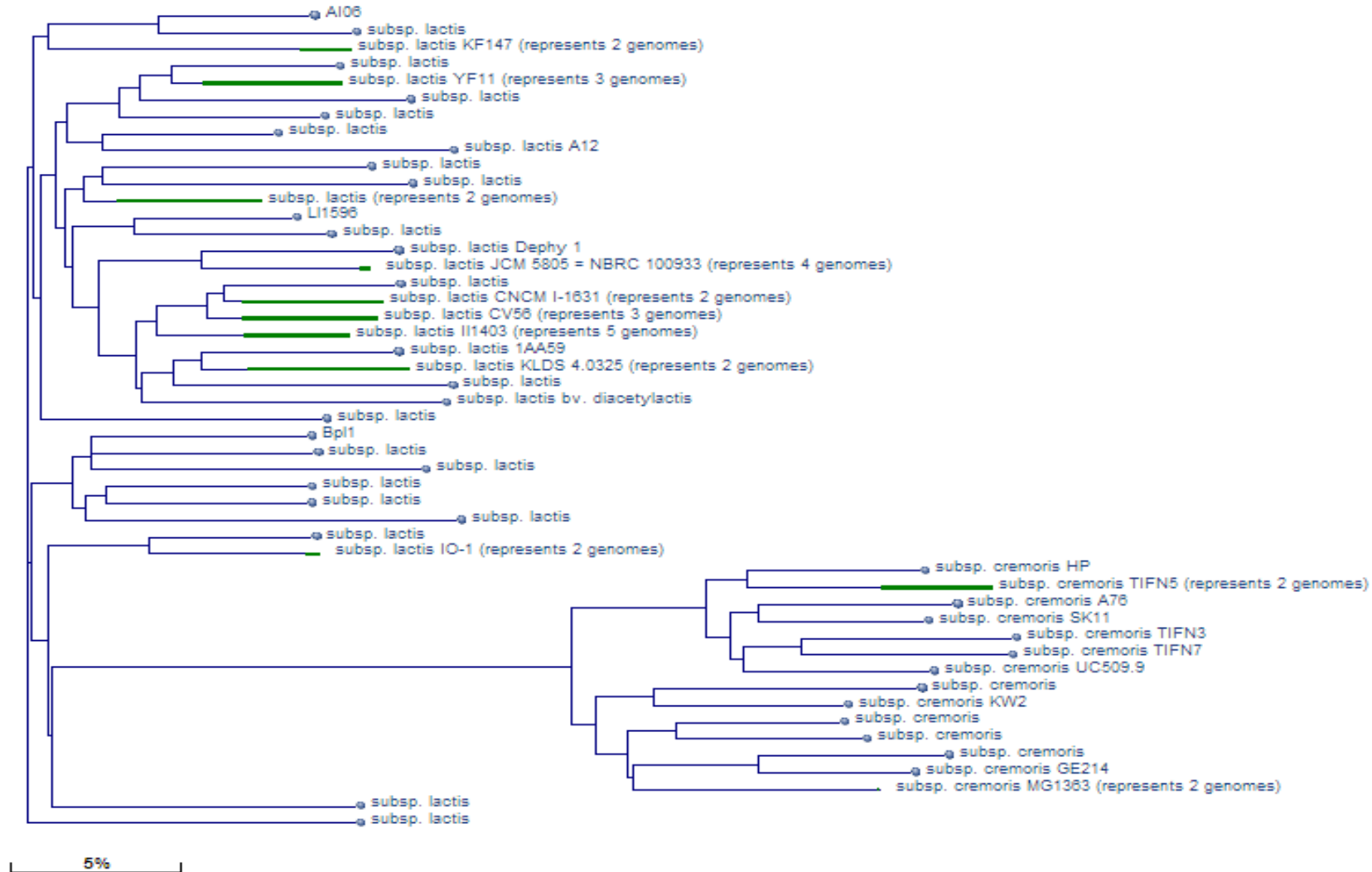


Figure 3 : Arbre phylogénétique des souches de *Lactococcus lactis* dont le génome est complètement séquencé. (29 Février 2016 NCBI)

2. Génomique des bactéries lactiques

Nous sommes exposés à une grande variété de microorganismes parmi eux, les bactéries d'acide lactique (LAB) ; ce groupe est composé de bactéries microaérophiles, batonnets et/ou des cocci non sporulants et qui ont la capacité de produire principalement de l'acide lactique à partir de sucres hexoses (Makarova and Koonin 2007). La grande diversité des LAB leur permet de coloniser une variété de niches écologiques allant des aliments comme les produits laitiers, les viandes, les légumes, le pain le vin et les muqueuses telles que la cavité buccale, et le tractus gastro-intestinal (Pfeiler and Klaenhammer 2007). Les caractéristiques métaboliques des LAB ont été exploitées pour la conservation des aliments. Les aliments fermentés à l'aide des LAB sont encore largement consommés. Récemment, les LAB commensaux ont reçu une attention accrue en raison de leur rôle important dans le maintien de la santé et la prévention des infections (Reid et al. 2003; Klaenhammer et al. 2005). Actuellement, 31 génomes complet de LAB ont été séquencés et sont accessibles au public. En outre, des analyses génomiques comparatives et fonctionnelles considérables ont accompagné l'apparition de l'information de la séquence génomique.

2.1. Historique et caractéristiques générales des génomes des LAB

La disponibilité des génomes séquencés a permis une meilleure compréhension de la divergence évolutive des LAB et révèle une tendance de réduction relativement récente de la taille du génome (van de Guchte et al. 2006). Le dernier ancêtre commun des Lactobacillales semble avoir perdu c. 600 – 1200 gènes et avoir acquis < 100 gènes durant sa divergence de l'ancêtre des bacilli (Makarova and Koonin 2007).

L'ampleur de la réduction du génome varie considérablement entre les LAB ; *Oenococcus oeni* a seulement c. 1700 ORFs prédites par rapport au c. 3000 chez *Lactobacillus plantarum* (Pfeiler and Klaenhammer 2007). L'analyse des génomes de LAB disponibles suggère que la majeure partie des gènes qui ont été perdus en raison de l'adaptation aux environnements alimentaires riches en nutriments, en particulier ces microorganismes qui se sont adaptés au lait et d'autres environnements riches en protéines et en glucides. La bactérie du yaourt *Lactobacillus delbrueckii* ssp. *bulgaricus* montre une grande différence dans la teneur en G-C % (49,7 %) de l'espèce *Lactobacillus acidophilus* (34,7 %), un organisme commensal gastro-intestinale. Fait intéressant, la différence était principalement due à la position du troisième codon moins conservé, qui avait une teneur en G-C 65 % ce qui implique une évolution rapide à un contenu plus élevé de GC. En outre, le nombre de gènes ARNr et des ARNt chez

Analyse bibliographique

L. bulgaricus est de 50 % supérieur à la moyenne de la taille de son génome. Ces nombres correspondraient à un génome de 3 – 4 Mb, significativement plus grande que sa taille réelle de 1,8 Mbp (van de Guchte et al. 2006). L'adaptation spécialisée au lait est particulièrement intéressante car ce milieu de fermentation n'existerait pas sans l'intervention humaine. La pression de sélection est venue non seulement de l'environnement naturel, mais aussi des milieux anthropiques créés par l'homme, qui a essentiellement domestiqué ces microorganismes au cours des 5000 dernières années par un transfert répété des cultures de LAB pour la production de produits laitiers fermentés.

Analyse bibliographique

Tableau 2 : Bactéries lactiques dont les génomes sont complètement ou partiellement séquencés. (NCBI,29-02-2016)

Genre	Nombre de génomes séquencés	Référence
<i>Lactobacillus</i>	747	http://www.ncbi.nlm.nih.gov/assembly/organism/1578/latest/
<i>Lactobacillus casei</i>	89 (8 génomes complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/652 http://www.ncbi.nlm.nih.gov/assembly/organism/655183/latest/
<i>Lactobacillus paracasei</i>	53 (7 complets)	http://www.ncbi.nlm.nih.gov/genome/2032 http://www.ncbi.nlm.nih.gov/assembly/organism/1597/latest/
<i>Lactobacillus delbrueckii (bulgaricus)</i>	29 (5 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/514? http://www.ncbi.nlm.nih.gov/assembly/organism/1584/latest/
<i>Lactobacillus plantarum</i>	53 (12 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/1108 http://www.ncbi.nlm.nih.gov/assembly/organism/1590/latest/
<i>Pediococcus</i>	27	http://www.ncbi.nlm.nih.gov/assembly/organism/1253/latest/
<i>Pediococcus pentosaceus</i>	7 (2 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/1109
<i>Lactococcus</i>	88	http://www.ncbi.nlm.nih.gov/assembly/organism/1357/latest/
<i>Lactococcus lactis</i>	68 (14 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/156?
<i>L. lactis</i> ssp. <i>lactis</i>	52 (8 complets)	http://www.ncbi.nlm.nih.gov/assembly/organism/1357/latest/
<i>L. lactis</i> ssp. <i>lactis</i> bv diacetylactis	5 (incomplets)	http://www.ncbi.nlm.nih.gov/assembly/organism/1357/latest/
<i>L. lactis</i> ssp. <i>cremoris</i>	11 (6 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/156?
<i>Streptococcus</i>	7177	http://www.ncbi.nlm.nih.gov/assembly/organism/1301/latest/
<i>Streptococcus thermophilus</i>	25 (11 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/420? http://www.ncbi.nlm.nih.gov/assembly/organism/1308/latest/
<i>Enterococcus</i>	961	http://www.ncbi.nlm.nih.gov/assembly/organism/1350/latest/
<i>Enterococcus faecalis</i>	442 (7 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/808? http://www.ncbi.nlm.nih.gov/assembly/organism/1351/latest/
<i>Enterococcus faecium</i>	414 (8 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/871? http://www.ncbi.nlm.nih.gov/assembly/organism/1352/latest/
<i>Tetragenococcus</i>	5	http://www.ncbi.nlm.nih.gov/assembly/organism/51668/latest/
<i>Leuconostoc</i>	42	http://www.ncbi.nlm.nih.gov/assembly/organism/1243/latest/
<i>Leuconostoc mesenteroides</i>	14 (4 complete)	http://www.ncbi.nlm.nih.gov/genome/genomes/1078?
<i>Fructobacillus fructosus</i>	3	http://www.ncbi.nlm.nih.gov/genome/genomes/11762?
<i>Oenococcus</i>	65	http://www.ncbi.nlm.nih.gov/assembly/organism/46254/latest/
<i>Bifidobacterium</i>	271	http://www.ncbi.nlm.nih.gov/assembly/organism/1678/latest/
<i>Bifidobacterium longum</i>	57 (16 complets)	http://www.ncbi.nlm.nih.gov/genome/genomes/183? http://www.ncbi.nlm.nih.gov/assembly/organism/216816/latest/
<i>Bifidobacterium bifidum</i>	24 (5 complete)	http://www.ncbi.nlm.nih.gov/genome/genomes/881? http://www.ncbi.nlm.nih.gov/assembly/organism/1681/latest/
<i>Weissella</i>	22	http://www.ncbi.nlm.nih.gov/assembly/organism/46255/latest/
<i>Carnobacterium</i>	7 (1 complet)	http://www.ncbi.nlm.nih.gov/genome/genomes/11762?
<i>Olsenella</i>	4 (1 complet)	http://www.ncbi.nlm.nih.gov/genome/genomes/13725?

2.2. Apports de la génomique sur l'étude des bactéries lactiques

La génomique est l'étude de l'ensemble de l'information génétique d'un organisme. La génomique est une discipline récente et en pleine expansion. Le premier génome bactérien séquencé et publié est celui de *Haemophilus influenzae*, agent pathogène responsable de méningites et dont le génome fait 1,8 Mb (Fleischmann et al. 1995).

La même année, le génome de *Mycoplasma genitalium*, agent d'infections uro-génitales qui possède le plus petit génome bactérien connu (580 kb) est publié (Fraser et al. 1995). Cette publication donnera déjà lieu à la comparaison des deux seuls génomes disponibles, et ouvrira la voie de la génomique comparative. De 1995 à 2000, 32 génomes bactériens sont publiés dont ceux d'*Escherichia coli* (Blattner et al. 1997), bactérie modèle des Gram (-) et de *Bacillus subtilis* (Kunst et al. 1997), bactérie modèle des Gram (+). Les progrès techniques et la création de grands centres dédiés au séquençage ont permis à la génomique de se développer. De 2001 à 2005, 237 génomes dont 63 pour l'année 2005 ont été séquencés et publiés (source <http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>).

2.2.1. Informations issues d'un génome

2.2.2. Structure du génome

La connaissance de la séquence d'un génome nous donne accès à un grand nombre d'informations parmi lesquelles la structure du génome étudié. La connaissance de la structure des génomes bactériens a largement profité des progrès de la génomique. Les paramètres les plus utilisés pour la description de la structure des génomes sont la taille, la géométrie, le nombre de réplicons et le pourcentage en G+C. La taille des génomes bactériens séquencés varie de 580 kb (*Mycoplasma genitalium*) à 9,1 Mb (*Bradyrhizobium japonicum*) (Bentley and Parkhill 2004). La taille des génomes est la résultante de différents évènements comme la duplication de gènes, l'acquisition de matériel génétique par transfert horizontal et la perte de gènes par délétion (Mira et al. 2001). Les plus petits génomes appartiennent à des espèces vivant dans des niches écologiques stables, le plus souvent en association avec un hôte. Les plus grands génomes, à l'inverse, se retrouvent dans des espèces vivant dans des milieux plus complexes et variables, comme le sol par exemple. La densité des gènes varie peu (environ 1 gène par kb), la taille du génome est donc en général proportionnelle au nombre de gènes présents (Bentley and Parkhill 2004).

Analyse bibliographique

La majorité des espèces bactériennes possède un chromosome circulaire unique. Cependant, certaines espèces possèdent un chromosome linéaire comme celles des genres *Streptomyces* (Kieser et al. 1992) ou *Borrelia* (Ferdows and Barbour 1989). Certaines espèces possèdent plusieurs chromosomes, codant chacun des fonctions essentielles à la survie cellulaire, par exemple les opérons ribosomiques (Ochman 2002). *Agrobacterium tumefaciens* contient des chromosomes linéaires et circulaires (Allardet-Servent et al. 1993).

La génomique a révélé que la composition nucléotidique des génomes n'est pas homogène, il est proposé que cela résulte de contraintes qui sont à ce jour encore mal comprises. Le pourcentage en G+C des génomes séquencés se situe entre 26,5% pour *Wigglesworthia glossinidia* (Akman et al. 2002) et 72,1% pour *Streptomyces coelicolor* (Bentley et al. 2002).

Une corrélation semble exister entre la taille des génomes et le pourcentage en G+C mais son explication n'est pas claire aujourd'hui (Bentley and Parkhill 2004). Au sein d'un même chromosome, le pourcentage en G+C varie. Les séquences intergéniques sont plus riches en A+T, et notamment les régions promotrices des gènes, facilitant ainsi l'ouverture des brins d'ADN lors de la transcription (Pedersen et al. 2000). Certainement pour des raisons similaires, l'origine de réplication est, elle aussi, riche en A+T (Bentley and Parkhill 2004). Il existe aussi des régions plus étendues pour lesquelles le pourcentage en A+T est plus élevé que la moyenne du génome. Ces régions de composition atypique sont souvent assimilées à de l'ADN issu de transfert horizontal (Daubin et al. 2003). Un autre paramètre est souvent regardé dans les génomes, le «GC skew». Le brin précoce est plus riche en G qu'en C, et inversement pour le brin retardé. Cela crée une inversion du «GC skew» (G-C/G+C) à l'origine de réplication et au terminus, ce qui permet de les situer sur le chromosome (McLean et al. 1998).

L'obtention de la séquence génomique d'un organisme permet aussi de mettre en évidence des éléments pouvant intervenir dans la plasticité du génome. Les séquences répétées (rDNA, séquences d'insertion) sont susceptibles de recombiner entre elles et d'induire des réarrangements dans le génome (Hughes 2000). Les éléments mobiles sont une autre source de plasticité du génome. Parmi ces éléments, on trouve les prophages, les transposons, les séquences d'insertion et les éléments conjugatifs qui lors de leur excision du génome peuvent emmener de l'ADN de l'hôte qui sera intégré ensuite dans le chromosome receveur (Dobrindt et al. 2004). La compétence naturelle est aussi un élément qui intervient dans la plasticité des génomes puisqu'elle permet à la bactérie compétente d'introduire dans son génome de l'ADN issu du milieu extérieur.

2.2.3. Capacités métaboliques

La séquence d'un génome donne accès à l'ensemble des gènes présents dans un organisme. Cependant, l'identification d'un cadre ouvert de lecture ne permet pas toujours d'avoir accès à sa fonction. Encore aujourd'hui, 20 à 40% des gènes détectés dans les génomes sont de fonction inconnue (Zheng et al. 2002). La génomique est donc tributaire des avancées dans les autres disciplines comme la physiologie ou la génétique classique. Lorsqu'un gène est détecté dans un génome, la première étape est d'attribuer une fonction (putative) au produit du gène. La plupart du temps, cette fonction est attribuée par homologie avec des gènes déjà étudiés, par identification de domaines conservés avec d'autres protéines ou plus occasionnellement par rapport à un contexte génétique (Gerlt et Babbitt, 2000). Ce processus, appelé annotation, est une partie essentielle à l'utilisation des données génomiques. L'annotation présente plusieurs limites : seules les fonctions déjà connues peuvent être identifiées, la fonctionnalité de la protéine n'est que putative (possibilité de mutations qui inactivent la protéine ce qui génère des faux positifs), et dans le cas de remplacements non orthologues (fonction assurée par une protéine non homologue), une fonction peut ne pas être identifiée (génère des faux négatifs).

Sur la base des fonctions putatives assignées aux gènes, il est possible de reconstruire *in silico* les capacités métaboliques de l'organisme. Certaines bases de données se sont spécialisées dans cette démarche, comme KEGG (Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.ad.jp/kegg/>), ERGO (<http://ergo.integratedgenomics.com/ERGO/>) ou MetaCyc (<http://metacyc.org/>). La reconstruction des voies métaboliques présentes dans un génome a été réalisée pour plusieurs organismes (Borodina and Nielsen 2005). Pour *L. plantarum*, la reconstruction des voies métaboliques a été effectuée dans l'environnement Pathway Tools (Karp et al. 2002) à partir des EC numbers (Enzyme Classification number : numéro permettant de caractériser la réaction réalisée par une enzyme) issus de l'annotation et des noms des produits des gènes (Teusink et al. 2005). Les reconstructions métaboliques représentent donc un moyen puissant de prédiction de voies métaboliques fonctionnelles chez un organisme à partir des données génomiques. Cependant, ces reconstructions, pour être fiables, nécessitent un travail de nettoyage manuel important et ne peuvent être effectuées que sur des voies de biosynthèse bien caractérisées. L'étape supplémentaire dans l'analyse des données génomiques est la comparaison entre plusieurs génomes.

2.2.4. Comparaison de génomes

La comparaison de génomes de souches différentes de la même espèce permet de mettre en évidence les régions variables entre les souches. Cette comparaison a surtout été réalisée pour des génomes de bactéries pathogènes. Chez ces bactéries, il était connu de longue date que les différences de phénotype entre deux souches étaient souvent dues à des gènes groupés physiquement dans le génome (notion de « cluster »). Ces régions génomiques, appelées « îlots de pathogénicité » (Hacker et al. 1990) ou « îlots génomiques » (Boyd et al. 2000) suivant qu'ils contiennent ou non des gènes associés à la virulence, ont été retrouvées dans les comparaisons entre génomes de souches d'une même espèce, comme dans le cas de *E. coli* (pour revue : (Whittam and Bumbaugh 2002).

Chez les bactéries lactiques, la seule comparaison de génomes de deux souches d'une même espèce à être publiée, a été réalisée pour *Streptococcus thermophilus* (Bolotin et al. 2004). Les deux souches, isolées de yaourt, ont plus de 90% des séquences codantes en commun et diffèrent par 55 insertions ou délétions de plus de 50 bases qui représentent 4% de la longueur du génome. 23 différences sont liées à des séquences d'insertion (IS) qui sont présentes à des endroits différents entre les deux génomes. Ces insertions sont soit restreintes à l'IS (pour 18 d'entre elles), soit accompagnées de groupes de gènes supplémentaires (pour 5 d'entre elles). Les gènes qu'on retrouve associés aux IS sont les groupes de gènes de biosynthèse des exopolysaccharides (EPS) dans deux cas, des groupes de gènes de fonction inconnue dans deux cas et un groupe de gènes impliqués dans la biosynthèse de bactériocine. Les autres différences les plus importantes non associées à des éléments génétiques mobiles concernent des gènes impliqués dans la biosynthèse des polysaccharides à rhamnose (4 opérons), un prophage, l'immunité et la synthèse de bactériocines (4 loci).

Une étude portant sur la diversité génomique entre 20 souches de *L. plantarum* a été réalisée par puces à ADN (Molenaar et al. 2005). Cette comparaison est partielle puisque l'ADN présent sur les puces représente 80% du génome de la souche *L. plantarum* WCFS1 séquencée (Kleerebezem et al. 2003). Les séquences manquantes dans les 19 autres souches peuvent donc être détectées mais pas les séquences supplémentaires. Parmi les régions qui sont absentes d'au moins 30% des souches, on retrouve en particulier des clusters de gènes dont les produits sont impliqués dans l'utilisation de sucres (9 clusters), la synthèse de protéines extracellulaires (3 clusters), la synthèse non ribosomale de protéines (1 cluster) et la synthèse d'un cofacteur, la molybdoptérine (1 cluster). On retrouve aussi des groupes de gènes classiquement variables entre souches comme les deux opérons de synthèse d'EPS, trois

Analyse bibliographique

prophages, un cluster de production d'une bactériocine, la plantaricine et un système de restriction-modification. Aucune corrélation n'a pu être mise en évidence entre la variabilité génomique et la provenance des souches (10 provenant d'isolats humains et 9 de produits végétaux fermentés) par cette méthode.

La comparaison de génomes peut aussi être réalisée entre les génomes d'espèces différentes mais proches phylogénétiquement. Chez les lactobacilles, jusqu'à aujourd'hui, les séquences de seulement quatre génomes sont disponibles, tous d'espèces différentes. Les espèces les plus proches pour lesquelles les génomes ont été séquencés sont *L. johnsonii* (Pridmore et al. 2004) et *L. acidophilus* (Altermann et al. 2005). Aucune étude n'a été publiée sur la comparaison de ces deux génomes. Cependant certaines caractéristiques sont communes aux deux génomes séquencés. La taille des deux génomes est d'environ 2 Mb et le pourcentage en G+C est de 34,71% pour *L. acidophilus* NCFM et 34,6% pour *L. johnsonii* NCC 533. Les deux bactéries semblent incapables de synthétiser la plupart des acides aminés et des vitamines mais possèdent des protéases et un grand nombre de transporteurs de peptides et d'acides aminés et de peptidases. Les deux génomes codent pour de nombreux systèmes de transports et de métabolisme de sucres, notamment d'hexoses. *L. johnsonii* NCC 533 a été isolée de l'intestin humain (Haller et al. 2002) et *L. acidophilus* NCFM a été isolée de fèces d'enfant (Altermann et al. 2005) et sont toutes les deux étudiées pour leur propriétés probiotiques. Leurs caractéristiques génomiques communes semblent liées à l'environnement de ces deux souches, le tractus digestif, qui est un milieu riche où les bactéries peuvent se procurer un grand nombre de nutriments. La comparaison entre les génomes de *L. plantarum* et *L. johnsonii* a été publiée en 2004 (Boekhorst et al. 2004).

La comparaison de génomes, qu'elle soit réalisée entre des souches proches phylogénétiquement ou non, a permis de mettre en évidence une certaine plasticité des génomes et notamment les transferts horizontaux. Ces transferts de matériel génétique peuvent représenter une grande partie d'un génome bactérien, comme le montre la comparaison de trois souches différentes d'*E. coli* qui n'ont qu'environ 40% de leurs gènes en commun (Welch et al. 2002).

Dans ce cas particulier, un grand nombre de différences entre les génomes est dû à des îlots génomiques ou de pathogénicité et à des prophages.

Deux méthodes sont utilisées pour détecter les gènes acquis par transfert horizontal : l'approche phylogénétique et l'approche paramétrique (Lawrence and Ochman 2002).

Analyse bibliographique

L'approche phylogénétique se base sur la distribution ou la forte similarité d'un gène avec des homologues dans des bactéries non apparentées (Lawrence and Hartl 1992; Ragan 2001). Un gène acquis par transfert horizontal aurait donc une histoire phylogénique différente de celle de son hôte. Cette méthode a besoin, pour être utilisée à l'échelle d'un génome, de plusieurs séquences génomiques de bactéries proches, voire de celles de plusieurs souches d'une même espèce.

Les méthodes paramétriques sont basées sur l'hypothèse que les gènes dont la composition nucléique et l'utilisation des codons synonymes sont différentes de celles de la majorité des gènes de l'organisme, sont probablement issus de transferts horizontaux (Medigue et al. 1991; Moszer et al. 1999; Mrazek and Karlin 1999). Ces méthodes ne nécessitent pas d'autres données génomiques que celles de l'organisme. Cependant, elles sont incapables de détecter des gènes issus de transfert horizontaux anciens (les gènes sont soumis au même biais mutationnel une fois intégrés dans le génome receveur et deviennent de moins en moins atypiques) ainsi que les gènes provenant de bactéries ayant le même biais mutationnel et la même utilisation de codons synonymes. Cette méthode détecte donc de manière plus fiable les gènes acquis récemment et sous-estime probablement leur nombre.

Ces deux méthodes, utilisées pour le génome d'*E. coli* ne donnent pas les mêmes résultats (Koski et al. 2001). Néanmoins, les résultats issus des deux analyses diffèrent plus sur la taille des régions transférées et donc sur le nombre de gènes que sur le nombre de transferts à proprement parler (Lawrence and Ochman 2002).

2.3. Génome de *Lactococcus lactis*



































Lactococcus lactis est une bactérie très utilisée en fermentation fromagère. Le premier génome de BL disponible a été celui d'une souche de *Lactococcus lactis* ssp. *lactis* IL1403 en 2001 (Bolotin et al. 2001). Plusieurs génomes de *lactococcus* ont été séquencés et de nombreux projets de séquençage sont en cours.

Analyse bibliographique





Tableau 3: Souches de *ssp. lactis* dont le génome total est complètement séquencé




Organisme (souche)	Niveau	Taille (Mb)	GC%	Gènes	Protéines	Références
<i>Lactococcus lactis</i> subsp. <i>lactis</i> I11403	■	2.36559	35.30	2406	2277	(Bolotin et al. 2001)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363	●	2.52948	35.70	2583	2400	(Wegmann et al. 2007)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11	●	2.59835	35.82	2682	2412	(Makarova et al. 2006)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> KF147	■	2.63565	34.86	2595	2445	(Siezen et al. 2010)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> NZ9000	■	2.53029	35.70	2583	2404	(Linares et al. 2010)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> CV56	■	2.51874	35.09	2533	2378	(Gao et al. 2011)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> A76	●	2.5771	35.88	2679	2382	Bolotin, et al., 2011
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> UC509.9	●	2.45735	35.76	2489	2188	(Ainsworth et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> IO-1	■	2.42147	35.10	2342	2230	(Kato et al. 2012)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> KW2	●	2.42705	35.70	2345	2223	(Kelly et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> NCDO 2118	■	2.59226	34.86	2545	2382	(Oliveira et al. 2014)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> KLDS 4.0325	●	2.59549	35.39	2648	2448	(Yang et al. 2013)
<i>Lactococcus lactis</i> , "AI06"	■	2.39809	35.00	2320	2178	(McCulloch et al. 2014)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> , "S0"	●	2.4887	35.20	2482	2311	Zhong,et al., 2014
<i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> gl2	●	2.2454	35.20	2179	2022	(Gabed et al. 2015)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> Dephy 1	■	2.60355	35.10	2634	2459	Velly,et al.2013
<i>Lactococcus lactis</i> subsp. <i>lactis</i> CNCM I-1631	■	2.51133	34.90	2546	2403	McNulty,N.P et al 2011
<i>Lactococcus lactis</i> subsp. <i>lactis</i> YF11	■	2.52731	34.90	2452	2328	(Du et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> A12	■	2.7	35.30	2663	2425	(Passerini et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> TIFN1	■	2.67978	35.50	2828	2285	(Erkus et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> str. TIFN2	■	2.50507	35.10	2603	2296	(Erkus et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> TIFN3	■	2.72521	35.50	2869	2291	(Erkus et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> str. TIFN4	■	2.55039	35.00	2635	2349	(Erkus et al. 2013)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> bv. <i>diacetylactis</i> str. LD61	■	2.59924	35.00	2687	2490	(Falentin et al. 2014)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> HP	■	2.26951	35.60	2325	2042	(Lambie et al. 2014)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> JCM 5805=NBRC 100933	■	2.527	35.20	2537	2210	(Kato et al. 2012)
<i>Lactococcus lactis</i> subsp. <i>lactis</i> , "511 »	■	2.48081	35.10	2492	2304	(Lambie et al. 2014)
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> GE214	■	2.80103	35.40	2835	2603	(Ladero et al. 2014)
<i>Lactococcus lactis</i> , "Bp11"	■	2.3057	34.90	2200	2092	Chaston,et al., 2014
<i>Lactococcus lactis</i> subsp. <i>lactis</i> , "CECT 4433"	■	2.57915	34.90	2629	2290	(Tschoeke et al. 2014)

Analyse bibliographique

Lactococcus lactis subsp. lactis 1AA59		2.57654	35.00	2570	2406	(Ladero et al. 2015)
Lactococcus lactis subsp. cremoris,"A17"		2.67994	35.60	2551	2372	(Leong et al. 2014)
Lactococcus lactis subsp. lactis,"JCM 5805"		2.54579	35.20	2553	2359	(Fujii et al. 2015)
Lactococcus lactis subsp. cremoris,"Mast36"		2.60534	35.40	2616	2414	(Plumed-Ferrer et al. 2015)
Lactococcus lactis subsp. cremoris,"DPC6860"		2.60744	35.60	2677	2261	(Cavanagh et al. 2015a)
Lactococcus lactis subsp. lactis,"DPC6853"		2.50715	35.00	2535	2116	(Cavanagh et al. 2015a)
Lactococcus lactis subsp. lactis bv. diacetylactis,"CRL264"		2.57372	35.00	2650	2446	Zuljan,et al., 2015
Lactococcus lactis subsp. lactis,"ATCC 19435"		2.54729	35.30	2558	2373	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"E34"		2.37566	34.90	2306	2217	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"K231"		2.33604	34.80	2283	2178	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF7"		2.36676	34.90	2335	2209	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF24"		2.61922	34.90	2645	2483	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"K337"		2.44552	34.70	2399	2263	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF67"		2.6843	34.80	2665	2514	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF146"		2.57452	34.70	2529	2408	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF196"		2.44589	34.80	2387	2282	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF282"		2.65125	34.80	2589	2471	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"LMG8520"		2.43558	34.80	2452	2060	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"LMG8526"		2.47749	34.80	2422	2304	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"LMG9446"		2.4884	34.80	2481	2324	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"LMG9449"		2.70754	34.80	2745	2555	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"LMG14418"		2.41093	34.80	2423	2275	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF134"		2.4634	34.80	2383	2282	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"KF201"		2.37639	34.90	2327	2222	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"Li-1"		2.47593	34.90	2422	2303	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"M20"		2.67432	34.80	2675	2535	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"ML8"		2.52187	35.10	2560	2373	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"N42"		2.74392	35.10	2738	2540	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"UC317"		2.49842	35.00	2522	2357	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis,"NCDO895"		2.47306	34.90	2450	2319	(Ercan et al. 2015)
Lactococcus lactis subsp. lactis JCM 5805 = NBRC 100933		2.54762	35.20	2557	2406	(Ercan et al. 2015)
Lactococcus lactis,"L11596"		2.39296	35.10	2377	2237	(Ptitsyn et al. 2015)
Lactococcus lactis subsp. cremoris,"DPC6856"		2.2454	35.20	2179	2022	(Cavanagh et al. 2015a)
Lactococcus lactis subsp. lactis,"LMG 7760"		2.86238	35.30	2980	2606	(Sun et al. 2015)

Analyse bibliographique

Lactococcus lactis subsp. cremoris TIFN5,"TIFN5"		2.24545	35.10	2202	2072	(Ainsworth et al. 2013)
Lactococcus lactis subsp. cremoris TIFN6,"TIFN6"		2.54151	35.50	2555	2232	(Ainsworth et al. 2013)
Lactococcus lactis subsp. cremoris TIFN7		2.59151	35.70	2669	2334	(Ainsworth et al. 2013)
Lactococcus lactis subsp. lactis JCM 7638		2.63409	35.60	2853	2505	(Shimizu-Kadota et al. 2013)

 Complet [14]  Scaffold [4]  Contig [50]

3. Les technologies NGS « *Next Generation Sequencing* »

3.1. Bioinformatique et biologie computationnelle

La bio-informatique est souvent vue comme l'application de l'informatique à la biologie, tandis que la biologie computationnelle est davantage un champ de recherche qui vise à développer de nouvelles méthodes computationnelles. Le professeur Pavel Pevzner de l'Université de Californie à San Diego a écrit plusieurs articles sur ce sujet (Pevzner 2004; Pevzner et al. 2008; Pevzner and Shamir 2009). Essentiellement, l'informatique a révolutionné la biologie et les cursus universitaires de biologie doivent inclure plus d'informatique. Je considère que mes travaux de doctorat sont classifiés dans la discipline de la biologie computationnelle.

La structure de l'acide désoxyribonucléique a été découverte par James D. Watson et Francis H. C. Crick (Watson and Crick 1953). Cette découverte était très importante et ils ont reçu un prix Nobel. La structure de l'ADN a aussi été étudiée dans plusieurs autres articles dans les années 1950 (voir <http://www.nature.com/nature/dna50/archive.html>). L'un de ceux-ci est l'article de James D. Watson et Francis H. C. Crick et dans celui-ci, les auteurs ont suggéré que les deux brins de l'ADN pourraient être utilisés par un mécanisme de copie du matériel génétique. Le séquençage de l'ADN a pour but de donner la séquence de nucléotides de molécules d'ADN. Cette information peut être utilisée, par exemple, pour identifier des mutations qui causent des phénotypes particuliers, ou encore pour détecter des agents infectieux dans des échantillons. Plusieurs publications ont comparé les technologies de séquençage (Loman et al. 2012; Quail et al. 2012; Junemann et al. 2013).

3.2. Le processus de séquençage

Plusieurs méthodes pour séquencer l'ADN se sont succédé au cours de l'évolution technologique des séquenceurs d'ADN. Un séquenceur d'ADN produit des fichiers contenant des séquences numériques d'ADN (Shendure and Ji 2008) (figure 2.1). Ces séquences sont des chaînes de caractères et l'alphabet utilisé est Le symbole N est aussi utilisé pour représenter une ambiguïté. Les types d'erreurs sont les substitutions, les insertions, et les délétions. Une substitution est présente lorsqu'un nucléotide est remplacé par un autre nucléotide différent. Le polymorphisme de simple nucléotide ou SNP (Single Nucleotide Polymorphism) est une substitution de nucléotide par un autre. Il y a une substitution entre les séquences ATGCATCGATGCTACGCAT et ATGGATCGATGCTACGCAT (la substitution est soulignée). Une insertion est produite lorsqu'une séquence est ajoutée à une position dans

Analyse bibliographique

une séquence. Par exemple, il faut faire une insertion pour transformer la séquence ATGATCGATGCTACGCAT en la séquence ATGATCGAGCGCGCGCGCTGCTACGCAT (l'insertion est soulignée). Finalement, une délétion est le contraire d'une insertion. La séquence ATGATCGAGCGCGCGCGCTGCTACGCAT doit subir une délétion (la séquence à enlever est soulignée) pour obtenir la séquence ATGATCGATGCTACGCAT.

3.3. La méthode de Sanger

Les deux premières méthodes de séquençage de l'ADN étaient celle de Maxam et Gilbert (Maxam and Gilbert 1977) et celle de Sanger (Sanger et al. 1977). Ces deux méthodes ont été publiées en 1977. L'une des premières méthodes qui a été décrite pour séquencer l'ADN était la méthode Sanger- une méthode décrite par Frederick Sanger en 1977 (Sanger et al. 1977). Frederick Sanger a reçu un prix Nobel en 1980 pour cette méthode, et il avait reçu un autre prix Nobel pour le séquençage de l'insuline en 1958. La méthode de séquençage de l'ADN de Sanger (Sanger et al. 1977) utilise des terminateurs aléatoires qui bloquent la polymérisation de l'ADN par l'ADN polymérase.

Dans la méthode Sanger, 4 réactions de polymérisation sont réalisées en parallèle. Chacune des réactions contient les 4 nucléotides en concentration égale. Dans chacune d'entre elles, un des quatre nucléotides est aussi présent avec une plus faible concentration sous une forme contenant un terminateur. Au final, chaque réaction contiendra des molécules de plusieurs longueurs. Chaque réaction est déposée sur un gel, et une électrophorèse est réalisée. La méthode Sanger utilise la radioactivité pour détecter des bandes sur un gel. Ce gel contient 4 lignes de migration – une ligne par nucléotide. Puisque les molécules migrent selon leur longueur, une lecture du gel produit directement la séquence de l'ADN initial. En partant avec la bande la plus éloignée du point de départ, et en lisant les 4 lignes en direction du point de départ, le biologiste moléculaire obtient la séquence d'ADN directement. Cette méthode était longue et fastidieuse.

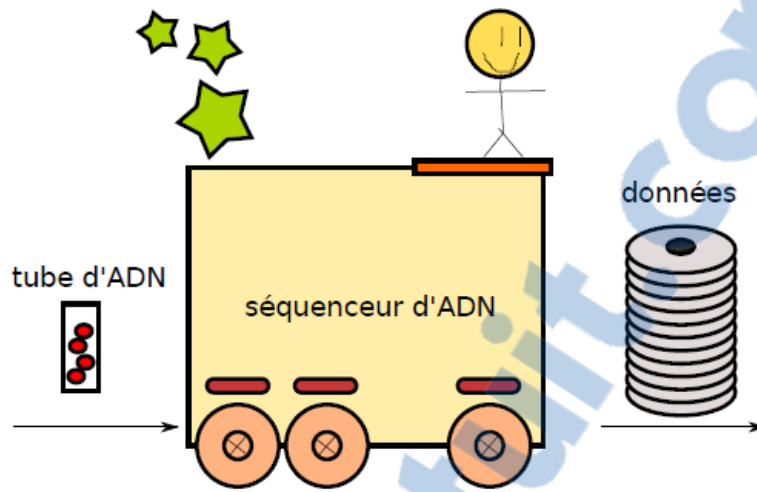


Figure 4: Un séquenceur d'ADN.

Un séquenceur est un système opéré par un humain qui a comme entrées des molécules d'ADN (ou bien des bibliothèques contenant des molécules d'ADN) et qui a comme sorties des séquences d'ADN numérisées dans des fichiers informatiques.

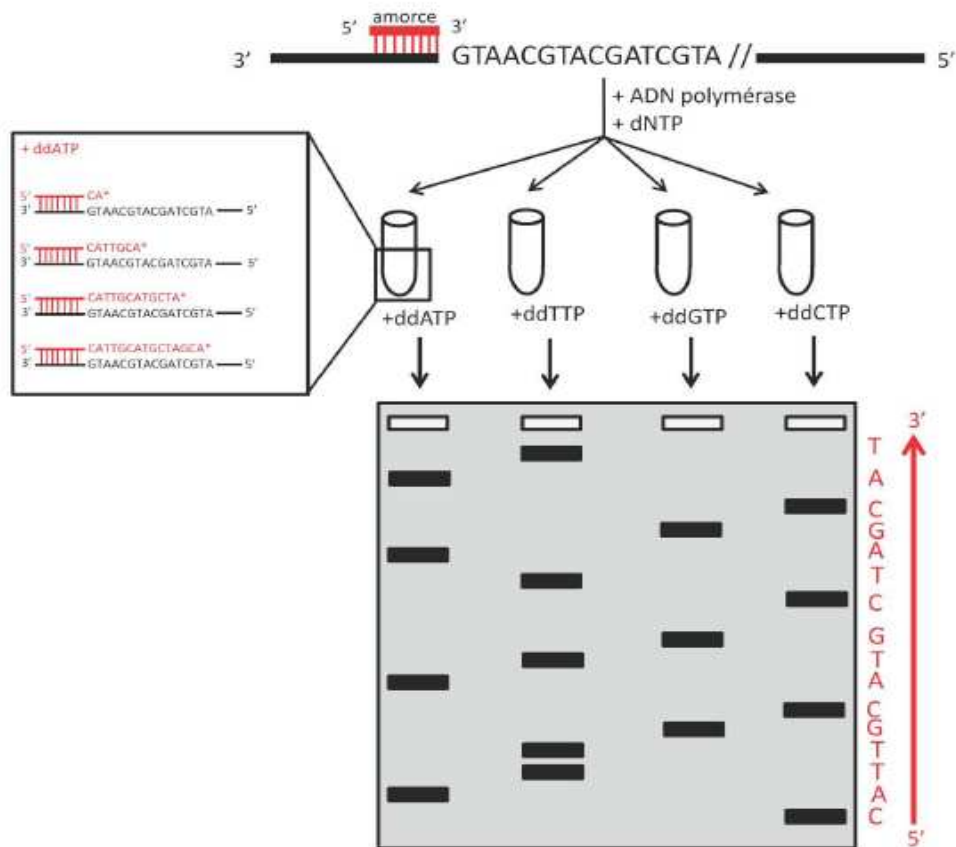


Figure 5 : Principe du séquençage selon la méthode de Sanger. Chacun des quatre tubes contient la matrice ADN, des amorces, de l'ADN polymérase, des dNTP et des ddNTP. A la fin de la réaction de polymérisation, chaque tube contient des fragments de tailles variables, correspondant à l'arrêt de la synthèse pour chacun des sites d'incorporation du ddNTP. A l'origine, la séquence était déduite de la séparation sur gel de polyacrylamide des fragments d'ADN de chaque réaction de polymérisation.

3.4. Automatisation avec le séquenceur 3730xl

Par la suite, la méthode de Sanger a été automatisée avec un ordinateur, l'électrophorèse capillaire, et la fluorescence (Smith et al. 1986). La migration sur un gel d'agarose a été remplacée par une séparation à l'aide d'une électrophorèse capillaire. L'avantage d'utiliser des terminateurs aléatoires est que le signal reçu de la fluorescence provient d'une seule position et non de plusieurs. Cette méthode a été commercialisée dans le produit 3730xl de la compagnie Applied Biosystems. Cette technologie a été utilisée par le Projet Génome Humain (Lander et al. 2001; Lander 2011). Avec cette méthode, une seule réaction (souvent appelée réaction de BigDye Terminator) est requise et aucune radio-activité n'est nécessaire. Les

molécules d'ADN sont séparées par taille avec l'électrophorèse capillaire (habituellement 96 capillaires indépendants).

3.5. La prochaine génération

Les concepts fondamentaux pour le séquençage de l'ADN à très haut débit ont été décrits par Church en 1988 (Marx 2013). Dans ce papier, le multiplexage des cibles est un concept important. De plus, une revue de la littérature des développements les plus significatifs dans le domaine des essais génomiques parallèles a été écrite en 2006 (Fan et al. 2006). Les principes des essais génomiques parallèles de Fan et al. 2006 s'appliquent aux situations suivantes : la réaction en chaîne de la polymérase (PCR) multiplexe, les puces de génotypage (essai GoldenGate ou essai avec sondes moléculaires à inversion (MIP : molecular inversion probe)), les puces d'hybridation comparative en génomique, l'analyse des mutations contribuant à l'expression spécifique à des allèles, et évidemment le séquençage parallèle de l'ADN. Ces principes sont aussi applicables pour le ChIP-on-chip, et le ChIP-seq (Fan et al. 2006).

3.5.1. Le pyroséquençage 454

La technologie 454 (Ronaghi et al. 1996; Ronaghi et al. 1998; Ronaghi 2001; Margulies et al. 2005) n'utilise pas de terminateurs aléatoires de polymérisation, contrairement à la méthode de Sanger (Sanger et al. 1977; Smith et al. 1986). La technologie 454 utilise plutôt une enzyme appelée la luciférase pour interpréter la séquence avec de la lumière. Dans le milieu réactionnel, il y a les enzymes suivantes : la polymérase d'ADN, ATP sulfurylase, la luciférase et l'apyrase. Les substrats suivants sont aussi présents : adénosine 5' phosphosulfate (APS), les didéoxynucléotides (dNTPs) et la luciférine.

Premièrement, une amorce s'hybride à l'amplicon simple brin d'ADN. Lorsque l'ADN polymérase incorpore un dNTP à l'ADN, un pyrophosphate est libéré. L'enzyme ATP sulfurylase convertit le pyrophosphate en adénosine triphosphate en présence d'APS. Par la suite, la luciférase convertit la luciférine en oxyluciférine, ce qui génère de la lumière. La hauteur du pic dans le pyrogramme est proportionnelle au nombre de didéoxynucléotides incorporés (Rothberg and Leamon 2008). La technologie 454 (Margulies et al. 2005) produit des erreurs de séquençage au travers des événements de type carry-forward et de la polymérisation dans les régions d'homopolymères. Étant donné qu'il n'y a pas de terminateurs aléatoires dans la technologie 454, le signal devient difficile à interpréter en présence d'homopolymères. Par exemple, la technologie 454 produit un signal très similaire

pour la séquence TTTTTT (6 T) et la séquence TTTTTTT (7 T). Les produits utilisant la technologie 454 sont : Roche 454 GS, Roche 454 GS FLX, Roche 454 GS FLX+ et Roche 454 GS Junior.

Le séquençage 454 est apprécié pour les séquences longues qu'il génère. L'alignement d'une séquence de 750 nucléotides est plus spécifique que l'alignement d'une séquence ayant entre 100 et 200 nucléotides (Wommack et al. 2008).

3.5.2. Le séquençage Illumina

Le séquençage Illumina est une approche de séquençage par synthèse. La technologie de séquençage Illumina domine le marché. Cette technologie est utilisée dans les séquenceurs de type Illumina Genome Analyzer, Illumina HiSeq et Illumina MiSeq. Différents kits de préparation des bibliothèques sont disponibles. La bibliothèque d'ADN est disposée sur une cellule à flux et une amplification en pont est réalisée pour augmenter la quantité de molécules afin de pouvoir les détecter. Lors du processus de séquençage, les monomères sont incorporés d'une manière similaire à la méthode Sanger c'est-à-dire qu'il y a des terminateurs (Fiegler et al. 2006). Les produits utilisant la technologie Illumina sont : Illumina Genome Analyzer, Illumina Genome Analyzer II, Illumina Genome Analyzer Iix, Illumina HiSeq 1000, Illumina HiSeq 1500, Illumina HiSeq 2000, Illumina HiSeq 2500 et le Illumina MiSeq.

Biais

Plusieurs articles ont souligné que la technologie Illumina produit des erreurs reproductibles (Rutherford et al. 2000). Avec la technologie de séquençage de l'ADN Illumina, des substitutions de simple nucléotide peuvent parfois être causées par l'oxydation de l'ADN pendant la préparation de l'échantillon (Costello et al. 2013). De plus, un biais concernant un des deux brins de l'ADN peut être observé lors du séquençage de l'ADN (Guo et al. 2012). Par ailleurs, ces biais spécifiques aux brins ne proviennent pas de source biologique et peuvent créer beaucoup de sommets interconnectés dans le graphe d'assemblage (Abnizova et al. 2012).

Un autre biais présent dans la technologie de séquençage Illumina (ainsi que dans d'autres technologies de séquençage) est que les régions riches en guanines et cytosines ont en règle générale une profondeur de séquençage plus basse. Un appariement entre une guanine et une cytosine contient trois liaisons hydrogène. Un appariement entre une adénine et une thymine

contient deux liens hydrogènes. Donc, un appariement entre une guanine et une cytosine est plus difficile à briser qu'un appariement entre une adénine et une thymine.

3.6. Autres technologies

3.6.1. Helicos

La technologie de séquençage Helicos analyse des molécules individuellement (Harris et al. 2008; Lipson et al. 2009; Ozsolak et al. 2009; Ozsolak and Milos 2011). En 2009, des chercheurs ont séquencé un génome humain pour \$50 000 avec cette technologie (Pushkarev et al. 2009). Le séquençage direct de molécules individuelles permet d'éviter de devoir amplifier les gabarits qui seront séquencés. La compagnie Helicos Biosciences a fait faillite au mois de novembre 2012. Il s'ensuit donc que cette technologie n'est plus disponible commercialement.

3.6.2. Pacific Biosciences

La technologie Pacific Biosciences analyse des molécules individuelles et en temps réel (Taddese et al. 2014). Les séquences Pacific Biosciences sont les moins biaisées, car la méthode analyse des molécules individuelles – il n'y a pas d'amplification (Ross et al. 2013). Aussi, la technologie SMRT permet de mesurer l'activité enzymatique de l'ADN polymérase directement et est très sensible.

3.6.3. Oxford Nanopore

La technologie Oxford Nanopore est en développement et utilise un pore pour y faire passer une molécule d'ADN (Smith et al. 1986). La nature de ce qui passe dans le pore change les propriétés du pore. En particulier, le courant résiduel du pore change selon le tri-nucléotide présent dans le pore (Venkatesan and Bashir 2011). Ce type de technologie serait très avantageux car le coût de fabrication serait plus bas et pourrait être fait dans les mêmes usines qui fabriquent les microprocesseurs. Par ailleurs, une technologie utilisant des circuits de semi-conducteurs, est décrite ci-après.

3.6.4. Ion Torrent Ion Torrent, Inc. a été acheté par Life Technologies, Inc. en 2010. Les deux produits sont le Ion Proton Sequencer et le Ion PGM Sequencer. Cette technologie utilise une technologie de séquençage avec des semi-conducteurs (Rothberg et al. 2011). Les gabarits d'ADN sont amplifiés sur des billes et ces billes sont déposées dans les puits d'une puce, avec exactement une bille par puits. La puce est ensuite déposée sur une grille de

Analyse bibliographique

senseurs qui mesurent le signal (le voltage). Lorsqu'un nucléotide triphosphate est incorporé au brin séquencé, un proton est libéré et celui-ci change l'acidité du milieu (le pH). Étant donné qu'il n'y a pas de terminateurs lors de la polymérisation, la technologie produit des erreurs de séquençage dans les homopolymères.

3.6.5. Moleculo La technologie de séquençage Moleculo (commercialisé par Illumina, Inc.) permet d'obtenir des séquences dont la longueur est plusieurs milliers de nucléotides.

Analyse bibliographique

	1ère génération	2e génération										3e génération						
Plateforme	3730	454			Illumina					SOLID		Ion Torrent			HelixScope	PacBio	Starlight	
Entreprise actuelle	Life Technologies	Roche			Illumina					Life Technologies		Life Technologies			Helicos	Pacific Biosciences	Life Technologies	
Entreprise d'origine	Applied Biosystems	454			Solexa					Applied Biosystems		Ion Torrent						
Commercialisé depuis	2002	2005			2006					2007		2010			Ne l'est plus	2011	Pas encore	
Renommée	Représentatif de la 1ère génération	1er de la nouvelle génération, longs reads			1er séquenceur à courts reads, leader actuel					2e séquenceur à courts reads, faible taux d'erreurs		1er séquenceur sans système de détection optique, 1er à moins de 100.000\$			1er séquenceur d'une molécule unique	1er séquenceur en temps réel d'une molécule unique	Séquençage d'une molécule unique avec quantum dots	
Méthode de séquençage	Sanger	Synthèse (pyroséquençage)			Synthèse					Ligation		Synthèse (détection de H ⁺)			Synthèse	Synthèse	Synthèse	
Méthode d'amplification	PCR	PCR en émulsion			PCR en ponts					PCR en émulsion		PCR en émulsion			Aucune	Aucune	Aucune	
Instrument	3730xl	GS Junior Titanium	FLX Titanium	FLX+	MiSeq	GA IIx	HiSeq 1000	HiSeq 2000	HiSeq 2000 v3	SOLID 4	SOLID 5500	SOLID 5500xl	Ion PGM, puce 314	Ion PGM, puce 316	Ion PGM, puce 318	Helicos ³	PacBio RS	Starlight
Coût à l'achat (\$ US)	376 000	108 000	500 000	500 000	125 000	250 000	560 000	690 000		475 000	349 000	595 000		49 500		nd	695 000	nd
Temps de run ¹	2h	10h	10h	18-20h	26h	14 jours	8 jours	8 jours	10 jours	12 jours	8 jours	8 jours		2h		nd	0,5-2h	nd
Millions de reads/run	0,000096	0,10	1	1	3,4	320	500	1000	≤ 3000	> 840	> 700	> 1410	0,10	1	4-8	800	0,01	0,01
Nb bases/read	650	400	400	700	150+150	150+150	100+100	100+100	100+100	50+35	75+35	75+35	100	>100	>100	35	860-1100	> 1000
Rendement Mb/run	0,06	50	500	900	1 020	96 000	100 000	200 000	≤ 600 000	71 400	77 000	155 100	>10	>100	>1000	28 000	5-10	nd
Coût des réactifs/run ²	96	1 100	6 200	6 200	750	11 524	10 220	20 120	23 470	8 128	6 101	10 503	500	750	925	nd	110-900	nd
Coût des réactifs/Mb	1500	22	12,4	7	0,74	0,12	0,1	0,1	<0,04	<0,11	<0,08	<0,07	<50	7,5	0,93	nd	11-180	nd
Erreurs principales	Substitutions	Insertions, délétions			Substitutions					Biais A-T		Insertions, délétions			nd	Délétions CG	nd	
Taux d'erreur (%) ⁴	0,1-1	1			>0,1					>0,06	>0,01	>0,01	~1			nd	16	nd
Applications principales		a*, b, c*, d, g, h*			a*, b, c*, d, e, f, g, h					c*, e, f, h		a, b, c, d, h			e, h	a, b, c, g, h	a, b, g, h	

1 : pour la longueur de lecture maximale

2 : de la préparation de la librairie au séquençage. Pour le séquençage capillaire, uniquement le séquençage

3 : instruments et réactifs ne sont désormais plus commercialisés

4 : pourcentage d'erreur par base sur les reads de longueur maximale (voir ligne nombre de bases/read)

nd: non disponible

Tous les coûts sont donnés en \$ US

a : séquençage de novo de génomes microbiens

b : caractérisation du transcriptome

c : reséquençage ciblé

d : reséquençage de novo de génomes animaux et végétaux

e : reséquençage

f : détection de mutation

g : métagénomique

h : autres applications (ChIP-Seq, MeDIP-Seq, RNA-Seq,...)

* : indexage nécessaire pour une utilisation optimale de cette application

Point fort en vert

Point faible en rouge

Tableau 4 : Caractéristiques des technologies de séquençage. Chiffres d'après Glenn 2011 (Glenn 2011). Ces technologies de séquençage ont été, sont ou seront commercialisées sur le marché.

3.7. La méta-génomique

Un système possède des composantes dans sa structure, et son comportement est défini par l'agencement de ces composantes. À l'intérieur de ces composantes figurent des processus qui transforment des entrées en sorties. Ces processus définissent un comportement. Un tel système interagit avec son environnement.

La biologie des systèmes s'intéresse aux systèmes biologiques, comme les communautés d'organismes vivants. La méta-génomique consiste à analyser les génomes de communautés de microorganismes (Lorenz and Eck 2005; Meyer et al. 2008). En particulier, l'intérêt de la méta-génomique est de pouvoir étudier des microorganismes sans les cultiver (Handelsman et al. 1998; Handelsman 2004). L'outil principalement utilisé pour obtenir des données en méta-génomique est le séquençage de l'ADN. Par exemple, il est possible de faire le diagnostic d'infections bactériennes avec la méta-génomique (Nakamura et al. 2008). Plusieurs tâches sont associées à la méta-génomique. Une des tâches est la classification taxonomique des séquences obtenues par séquençage de l'ADN (Krause et al. 2008; Wu et al. 2011; Segata et al. 2012). Une autre tâche en méta-génomique est l'assemblage de novo de méta-génomes c'est-à-dire de construire des séquences longues qui correspondent au matériel génétique des organismes vivant dans un environnement donné à partir des données de séquençage (Sanger et al. 1977; Laserson et al. 2011; Peng et al. 2011; Namiki et al. 2012). Avec l'assemblage de novo, il est possible d'analyser le contenu d'un échantillon plus efficacement. En effet, des séquences plus longues sont plus spécifiques pour la méta-génomique (Wommack et al. 2008). Des séquences longues peuvent être obtenues avec l'assemblage de novo (Chaisson et al. 2009) ou avec une technologie de séquençage offrant des séquences plus longues comme les technologies Pacific Biosciences ou Moleculo. Les chercheurs étudient des environnements très diversifiés en méta-génomique, comme les océans en général (Rusch et al. 2007) et l'Arctique ou l'Antarctique (Varin et al. 2012), en particulier. La méta-génomique comparative consiste à comparer des méta-génomes en utilisant des mesures comme le contenu en gènes (Tringe et al. 2005). Un des objectifs principaux de la méta-génomique est de prédire les réseaux métaboliques et donc les capacités enzymatiques qui sont présentes dans un système pour lequel des données de séquençage de l'ADN sont disponible.

4. Séquençage des génomes

La première bactérie séquencée a été *L. lactis* IL1403. Une première séquence diagnostique a été publiée par Bolotin et al. En 1999, et une séquence de haute qualité a été publiée en 2001. Depuis, le nombre de génomes séquencés augmente régulièrement (Liu et al. 2005) et le tableau 2 récapitule les génomes de toutes les souches appartenant à l'espèce *lactis* répertoriées par GenBank en 2016.

Actuellement, les stratégies de séquençage génomique débutent par une préparation de librairie (une banque de fragments de taille connue), ensuite est séquencée selon une approche de NGS choisie. Après le séquençage, un premier assemblage est constitué, en général avec une suite de logiciels bioinformatique, pour obtenir une série de « contigs » (séquence consensus rassemblant des lectures chevauchantes) qui couvriront 95 à 98 % du génome (Kim et al. 2001a). Une autre étape, souvent nommée « finishing », consiste à ordonner les « contigs », à combler les séquences qui les séparent, et à améliorer la qualité des séquences dans les régions où subsistent des indéterminations (dus, par exemple, à un nombre insuffisant de lecture). Avec l'augmentation de la disponibilité de génomes séquences, une autre stratégie est utilisée pour faciliter l'assemblage. Son principe est basé sur le fait que l'organisation génétique, entre des bactéries phylogénétiquement proches, est parfois suffisamment conservée pour permettre d'aligner les « contigs » par rapport aux génomes connus d'autres bactéries (en général, cela se limite aux bactéries phylogénétiquement proches). Cette méthode a été utilisée pour assembler la séquence de *L. lactis* ssp. *cremoris* à l'aide du génome de *L. lactis* ssp. *lactis* IL1403 (van Hijum et al. 2005).

4.1. Analyse bio-informatique des génomes et bases de données

La séquence brute d'un génome n'est pas une information utilisable sans une panoplie d'outils pour l'analyser (Madhumalar and Bansal 2005; Siezen et al. 2005). Nous distinguons deux types d'outils nécessaires à l'exploitation des données génomiques : les outils d'analyse et les bases de données. Ces outils ne sont pas spécifiques aux bactéries lactiques.

4.1.1. Outils d'analyse et d'annotation des séquences

L'identification correcte des gènes codés par le génome est une étape déterminante en génomique. Ce travail est largement basé sur des processus informatiques automatiques (Karp et al. 2002; Iliopoulos et al. 2003). Les processus d'annotation sont de plus en plus sophistiqués et combinent généralement de multiples informations (Reed et al. 2006a; Reed et al. 2006b). Il existe des logiciels de détection des phases ouvertes de lecture basés sur différents principes. Ils combinent, en général, la recherche de codon d'initiation et d'arrêt de

Analyse bibliographique

la traduction, des analyses probabilistes basées sur les séquences nucléotidiques ou sur la similarité de leurs produits avec ceux de gènes déjà caractérisés (Liou et al. 2004).

La similitude entre le produit possible d'une phase de lecture ainsi déterminé et d'une protéine déjà identifiée est un facteur supplémentaire permettant de valider l'identification des gènes. Les algorithmes les plus utilisés à cet effet sont FASTA et BLAST. BLAST, la référence pour initier une recherche d'homologie, a donné lieu à de nombreux programmes dérivés, plus fiables et plus rapides. Ces programmes utilisent également une matrice de score pour identifier les régions les plus longues présentant une homologie « non-due au hasard ». L'utilisation de FASTA (ou de ses dérivés) permet d'optimiser les alignements avec les homologues identifiés. En effet, cet algorithme est plus exigeant en terme de calcul puisqu'il cherche à maximiser les scores d'alignement en calculant des scores de manière récurrente. Enfin, des logiciels comme CLUSTALW permettent de réaliser des alignements multiples de l'ensemble des protéines présentant une forte similitude. Il est alors possible de vérifier que les nouvelles protéines présentent les caractéristiques essentielles des protéines des groupes auxquels elles appartiennent. La présence de motifs dans une séquence protéique est une information supplémentaire qui permet de proposer une fonction. Ainsi, même si une protéine ne présente pas une similarité suffisante avec une famille de protéines pour qu'une fonction précise lui soit attribuée, certains motifs sont caractéristiques de la fixation de cofacteurs, comme l'ATP ou le NAD, d'autres permettent d'identifier des séquences de reconnaissance de l'ADN, des sites de phosphorylation, etc. Enfin, les protéines comportent un certain nombre de domaines qui peuvent être identifiés. La combinaison de ces domaines permet de classer les protéines en familles, même en absence d'une homologie très forte sur l'ensemble de la protéine. Ces recherches permettent ainsi de guider les biologistes pour émettre des hypothèses sur le rôle d'une protéine. Au-delà des similarités de séquences, certains logiciels recherchent des similitudes dans le repliement des protéines. En effet, les structures secondaires et tertiaires des protéines jouent parfois un rôle déterminant dans leur fonction, sans qu'il y ait conservation de la séquence primaire (Aloy et al. 2003; Holscher et al. 2012). Mis à part les premières étapes de l'identification des séquences codantes, l'assignation d'une fonction requiert des connaissances préalables et des éléments de comparaison. Ces données sont stockées dans des bases de données qui constituent des références incontournables dans l'exploitation des données génomiques.

4.1.2. Bases de données

Les bases de données constituent un enjeu essentiel en bio-informatique. Elles sont nécessaires pour la structuration et l'exploitation de la masse de données produites par les

Analyse bibliographique

programmes de génomique (Siezen et al. 2004). Comme nous l'avons vu précédemment, de nombreux logiciels d'analyses s'appuient sur les informations stockées dans des bases de données. Parmi les bases de données, on peut distinguer :

- Les grandes banques généralistes internationales de séquences nucléotidiques (EMBL et Genbank) et protéiques (Swissprot et NBRF-PIR). Ces banques, comme l'EMBL, intègrent en général de nombreuses données (Kasprzyk et al. 2004);
- Les banques spécialisées dans les motifs (PROSITE, PRODOM, PFAM, RFAM...), les données structurales (par exemple, PDB qui donne les coordonnées 3D des protéines dont la structure a été déterminée), les données métaboliques (KEGG, *Kyoto Encyclopedia of Gene and Genomes*, permet de visualiser voies et composés métaboliques, enzymes, cartes génomiques, gènes) ;
- Les bases de données génomiques, issues directement du séquençage de génomes particuliers (par exemple SGD, *Saccharomyces Genome Database*).

Parmi, les bases de données spécialisées sur des bactéries particulières, mentionnons EcoCyc (*Encyclopedia of Escherichia coli K12 Genes and Metabolism* : (Keseler et al. 2005). Cette base de données est devenue une référence pour la biologie d'*E. coli* intégrant séquences génomiques, annotations, données transcriptionnelles (opérons, promoteurs, facteurs de transcription...), données structurales, voies métaboliques (voies, cofacteurs, données enzymatiques...) et données bibliographiques spécialisées. L'un des points essentiels de cette base est la fiabilité des données qu'elle contient grâce à un travail de vérification constant mené par différents scientifiques de la communauté « *E. coli* ». Il n'existe pas de base comparable chez les bactéries lactiques. De nombreuses données sont cependant consultables dans les bases de données généralistes, et dans quelques bases de données construites pour permettre l'exploitation des données de séquençage. Parmi les bases généralistes, PAREO (*Pathway RELational Organisation*) a été mise à jour pour les données concernant *L. lactis* IL1403 (<http://genome.jouy.inra.fr/pareo>). Cette base intègre en particulier les données de la base KEEG, et permet de comparer les voies métaboliques de différentes bactéries. De nombreuses données sont également disponibles sur le site du JGI (*Joint Genome Institut*, http://genome.jgi-psf.org/mic_home.html) où les données génomiques de différentes bactéries lactiques ont été analysées pour produire des schémas de reconstruction métabolique. Cependant, ces données ont été générées automatiquement, sans vérification de leur qualité.

4.2. Intérêts du séquençage des génomes de *Lactococcus lactis*

Étant l'un des organismes modèle dans le métabolisme microbien, *Lactococcus lactis* a été la principale cible de l'ingénierie métabolique chez les BL Holscher et al. (2012). Cette bactérie

a de nouvelles applications biomédicales : comme adjuvants, immunostimulants ou comme un système de livraison des composés thérapeutiques tel que la livraison des enzymes digestives pour compléter l'insuffisance pancréatique chez l'homme ; livraison de l'ADN plasmidique aux cellules eucaryotes ou elle est considérée comme « usine » pour produire des molécules thérapeutiques et plus récemment l'utilisation potentielle de *L. lactis* comme « *Gram-positive enhancer matrix* » (GEM), comme un système de diffusion pour la vaccination des muqueuses (Wyszynska et al. 2015). *Lactococcus lactis*, est une usine cellulaire pour la production de protéines recombinantes utilisées dans diverses applications biotechnologiques (Morello et al. 2008).

5. Réponses au stress chez les bactéries lactiques

5.1. La réponse générale au stress

Les facteurs σ sont utilisés par de nombreuses bactéries pour moduler l'expression de leurs gènes en fonction de leur environnement. Parmi les facteurs σ contrôlant des régulons de réponse aux stress identifiés chez les bactéries à Gram (+), le plus étudié est probablement σ^B , qui contrôle la réponse générale au stress de *B. subtilis*. σ^B est induit en condition de stress osmotique et oxydant, en présence d'éthanol et en phase stationnaire (Boylan et al. 1993).

D'autres études ont montré que σ^B répond à d'autres stimuli environnementaux comme le stress hypothermique (Brigulla et al. 2003) et hyperthermique (Benson and Haldenwang 1993), le stress acide (Kovacs et al. 1998) et des conditions de carence en glucose, oxygène (Volker et al. 1994) ou phosphate. L'activation du régulon σ^B permet à la bactérie d'améliorer sa résistance à différents stress (Volker et al. 1994). Des analyses transcriptomiques ont montré que σ^B contrôle l'expression de plus d'une centaine de gènes (Petersohn et al. 2001) en condition de stress, dont la plupart sont de fonction inconnue. Parmi les gènes régulés par σ^B se trouvent des gènes pour lesquels la relation avec la résistance aux stress est évidente comme par exemple ceux codant pour la catalase (*katX*), la thioredoxine (*trxA*) et une protéine de liaison au fer (*dps*), impliquées dans la résistance au stress oxydant, des transporteurs de proline (*opuE*) et de glycine-bétaine (*opuD*) impliqués dans la résistance au stress osmotique ou *clpC*, codant une sous-unité des protéases Clp impliquées dans la résistance au stress hyperthermique (Hoper et al. 2005). La régulation de σ^B est complexe et a été largement étudiée chez *B. subtilis*.

Analyse bibliographique

Les protéines qui régulent l'expression de σ^B sont au nombre de sept : RsbR, RsbS, RsbT, RsbU, RsbV, RsbW et RsbU (Regulator of sigmaB). Le gène sigB est co-transcrit avec ces protéines régulatrices dans un opéron dont le promoteur est dépendant du sigma primaire. Lors de l'induction de sB, un promoteur interne dépendant de σ^B est induit, permettant l'induction des gènes rsbV, rsbW, sigB et rsbX (les quatre derniers gènes de l'opéron) (Wise and Price 1995).

Bien que σ^B ne soit pas conservé chez tous les firmicutes (notamment chez les bactéries lactiques), des homologues de σ^B existent, par exemple chez *Listeria monocytogenes* et *Staphylococcus aureus*. Chez ces deux bactéries, σ^B est non seulement impliqué dans la réponse générale au stress mais aussi dans la virulence (Kazmierczak et al. 2003; Bischoff et al. 2004).

L'organisation génétique de l'opéron sigB et la régulation de ce facteur s diffèrent entre *B. subtilis* et *S. aureus*, car seuls sont présents, dans le génome de *S. aureus*, les gènes rsbU, rsbV, rsbW et sigB.

D'autres facteurs s ont été impliqués dans la résistance aux stress mais ils contrôlent des régulons moins étendus que celui contrôlé par σ^B . Citons par exemple σ^X de *B. subtilis*, impliqué dans la résistance au stress hyperthermique (Huang et al. 1997) ou σ^V de *Enterococcus faecalis* impliqué dans la survie à des stress hyperthermique, acide et à la présence d'éthanol (Benachour et al. 2005).

5.2. Le stimulon de réponse au stress thermique : exemple de régulons coordonnés

Chez *B. subtilis*, les réponses aux stress ne dépendent pas uniquement de σ^B . Les relations entre le régulon de réponse générale au stress σ^B et la réponse à un stress spécifique peuvent être illustrée par le stimulon (ensemble des gènes induits en réponse à un stimulus) du stress hyperthermique. Les gènes de réponse au choc hyperthermique sont classés en six grandes classes chez *B. subtilis*, mais les mécanismes de régulation sont seulement connus pour les classes I, II, III et V (Zellmeier et al. 2003). Les gènes de la classe II sont les gènes régulés par σ^B et ne sont donc pas spécifiques du stress hyperthermique.

Présentons maintenant succinctement la régulation des classes I, III et IV.

Les gènes de la classe I sont régulés par un régulateur de transcription, HrcA. Chez *B. subtilis*, ils sont organisés en deux opérons, les opérons groES/EL et

Analyse bibliographique

hrcA/*grpE*/*dnaK*/*dnaJ*/*yqeT*/*yqeU*/*yqeV* (Homuth et al. 1997). HrcA est un répresseur transcriptionnel qui se fixe sur des sites appelés CIRCE (Controlling Inverted Repeat of Chaperone Expression) de consensus : TTAGCACTC-N9-GAGTGCTAA (Zuber and Schumann 1994). L'activité de HrcA est modulée par la protéine GroEL (Mogk et al. 1997).

L'hypothèse est que, pour être active, HrcA a besoin de GroEL. Lors d'un choc hyperthermique, GroEL est titrée par les protéines dénaturées, HrcA est alors inactivé et la répression des deux opérons de chaperons est levée.

Les gènes de la classe III sont au nombre de six et sont régulés par le régulateur transcriptionnel CtsR. Ils sont organisés en trois unités transcriptionnelles, *ctsR*/*mcsA*/*mcsB*/*clpC*, et les unités monocistroniques *clpE* et *clpP* (Kruger and Hecker 1998; Derre et al. 1999a; Derre et al. 1999b). CtsR réprime la transcription en se fixant sur la CtsR-box dont le consensus est : A/GGTCAAANANA/GGTCAAA. L'activité de CtsR est régulée post traductionnellement de deux manières. CtsR est dégradée à 37°C par les protéases ATP dépendantes ClpP et ClpX, ce qui permet de garder un niveau basal faible de CtsR (Derre et al. 2000). De plus, CtsR est phosphorylée par la tyrosine kinase McsA/McsB (McsA est l'activateur de la kinase McsB), ce qui inhibe sa fixation sur la CtsR-Box (Kirstein et al. 2005). L'activité de McsB est inhibée par ClpC (codée par le même opéron que CtsR, McsA et McsB). En condition de stress hyperthermique, ClpC serait titrée par les protéines dénaturées, McsB phosphorylerait alors CtsR, provoquant son inactivation ce qui lèverait la répression des gènes de la classe III.

La classe V compte deux membres, *htrA* et *htrB* codant des protéases ancrées à la membrane (Darmon et al. 2002). Ces gènes sont contrôlés par le système à deux composants C_{ss}RS, qui répond non seulement au stress hyperthermique mais aussi au stress de sécrétion (induit en laboratoire par surproduction d' α -amylase). C_{ss}S détecterait l'accumulation de protéines dénaturées dans le milieu extérieur et activerait, via la phosphorylation de C_{ss}R et sa fixation sur les opérateurs, l'expression des protéases de surface (Hyyrylainen et al. 2001).

La régulation du stress hyperthermique chez *B. subtilis* fait donc intervenir différents types de régulateurs. De plus, certains gènes sont sous double contrôle, comme les gènes de classe III, *clpC* et *clpP* sous contrôle de *sB* et de CtsR.

Chez les autres firmicutes, l'organisation observée chez *B. subtilis* n'est pas toujours conservée. Par une étude de génomique comparative, Chastan et coll. ont montré que les sites de liaison de CtsR se retrouvent en amont des opérons *groESL* des streptocoques

Analyse bibliographique

(Chastanet et al. 2003) et cette régulation a été démontrée chez *S. pneumoniae* (Chastanet et al. 2001). Chez *Streptococcus salivarius*, *clpP* est sous contrôle de CtsR et d'HrcA (Chastanet and Msadek 2003). Chez *Staphylococcus aureus*, CtsR contrôle non seulement les gènes *clpC*, *clpB* et *clpP* mais aussi les opérons *dnaK* et *groESL*, qui se retrouvent sous le double contrôle de CtsR et HrcA (Chastanet et al. 2003). Chez la bactérie lactique *Oenococcus oeni*, seul CtsR est présent, HrcA est absent de la séquence du génome. Chez cette bactérie, les gènes *clp* et les opérons *groESL* et *dnaK* sont sous le contrôle unique de CtsR (Grandvalet et al. 2005).

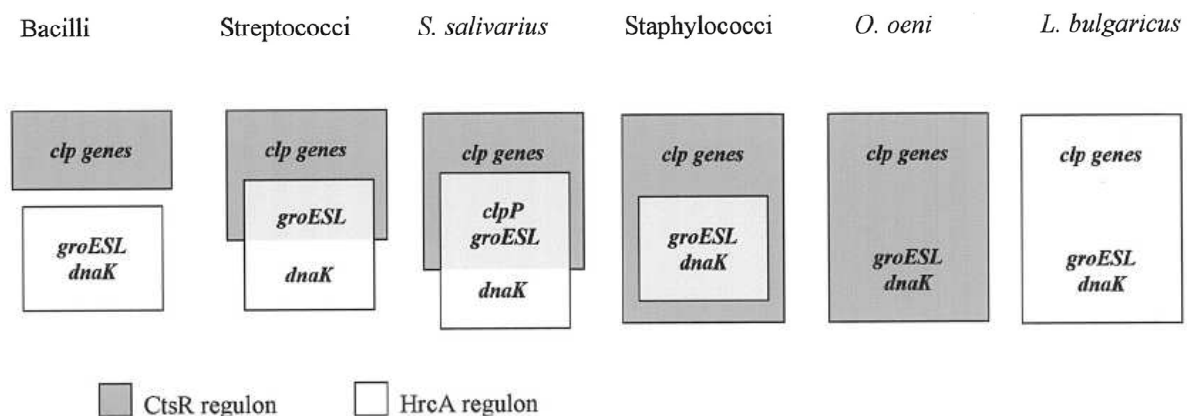


Figure 6 : Comparaison des régulons d'HrcA et CtsR chez différents firmicutes (Grandvalet et al. 2005)

5.3. Les BL et la régulation des réponses aux stress

Chez les bactéries lactiques, peu de régulons de stress sont caractérisés. La plupart des études menées jusqu'à aujourd'hui sont des études globales permettant d'identifier des gènes induits ou des protéines surproduites en condition de stress. Cependant, ces études ne se sont généralement pas focalisées sur la régulation des mécanismes de résistance mis en jeu. Les génomes de bactéries lactiques dont la séquence a été publiée montrent que ces bactéries possèdent en général peu ou pas de facteurs σ alternatif et notamment pas d'homologue de σ B.

Parmi les systèmes de régulation, les systèmes à deux composants semblent de bons candidats pour la régulation coordonnée de la réponse aux changements environnementaux.

Certaines études ont été réalisées par l'inactivation spécifique de ces systèmes chez *L. lactis* (O'Connell-Motherway et al. 2000) et chez *L. sakei* (Morel-Deville et al. 1998). Chez *L. lactis*, l'inactivation d'un des systèmes à deux composants induit la perte de l'activité de la

voie arginine déiminase et une augmentation de la sensibilité au stress acide. Chez *L. sakei*, certains mutants sont plus sensibles à certains stress comme le stress hyperthermique, acide ou oxydant.

Cependant, dans les deux cas, les régulateurs des systèmes à deux composants étudiés ne sont pas déterminés. Récemment, une étude portant sur un système à deux composants de *L. acidophilus* a été réalisée en transcriptomique (Azcarate-Peril et al. 2005). Le mutant est plus sensible au stress acide et la comparaison entre la souche sauvage et le mutant montre que la plupart des gènes différentiellement exprimés appartiennent aux systèmes de protéolyse et de transport d'acides aminés ou de peptides. Cependant, encore une fois, le mécanisme d'activation de ce système à deux composants est inconnu et rien ne permet de conclure à une action directe sur les promoteurs de ces gènes ou à une action indirecte.

Les bactéries lactiques rencontrent, à priori, au cours des procédés industriels, dans leur environnement ou lors de la traversée du transit digestif in vivo, de nombreux stress, comme le stress thermique (hyper ou hypothermique), le stress osmotique, le stress oxydant ou le stress biliaire.

De part leur métabolisme fermentaire, les bactéries lactiques sont soumises au stress thermique, stress auquel nous nous sommes plus particulièrement intéressés.

5.4. Le stress thermique chez les BL

Chez les bactéries à Gram +, la réponse au stress hyperthermique fait intervenir deux familles d'acteurs principaux : les protéines chaperons et les complexes multiprotéiques Clp.

Chez *L. bulgaricus*, l'analyse in silico du génome permet de retrouver ces deux grandes familles de protéines mais aussi quelques autres protéines référencées comme «*Heat Shock Protein*» (HSP) chez d'autres micro-organismes. En termes de fonctions potentiellement associées au stress thermique chez *L. bulgaricus*, le bilan est le suivant.

Les protéines chaperons ont pour rôle de permettre aux protéines dénaturées par la température de retrouver leur conformation native. Chez *L. bulgaricus*, les deux principaux opérons de protéines chaperons sont codés par les opérons groES/groEL et hrcA/grpE/dnaK/dnaJ. Chez de nombreux micro-organismes, hrcA code le régulateur négatif de ces deux opérons. Nous reviendrons sur son rôle chez *L. bulgaricus* à la fin de ce paragraphe.

Analyse bibliographique

Les autres protéines pouvant être impliquées dans la réponse au stress thermique détectées dans le génome de ATCC 11842 sont :

La protéine codée par ldb0289, homologue à Hsp18, protéine chaperon associée à la membrane, étudiée chez *O. oeni* (Delmas et al. 2001; Coucheney et al. 2005). Chez *L. bulgaricus*, ldb0289 semble être en opéron avec deux gènes codant des peptidases (pepW et pepG).

La protéine Ldb0364, homologue à Hsp15 d'*E. coli* chez qui cette protéine est induite en condition de choc thermique (Richmond et al. 1999) et qui s'associe aux ribosome (Korber et al. 2000).

La protéine Ldb0369 (codée par hslO), homologue à Hsp33, protéine chaperon associée au stress oxydant chez *E. coli* (Winter et al. 2005).

Les complexes multiprotéiques Clp sont la seconde famille de protéines importantes pour la résistance au stress thermique. Les complexes Clp sont composés de deux sous-unités : une sous unité protéolytique, ClpP et une sous-unité liant l'ATP qui permet la reconnaissance du substrat.

Sous forme de complexes, les Clp ont pour fonction de dégrader les protéines qui sont définitivement dénaturées lors du stress. Mais la sous-unité ATPasique peut, sous forme libre, remplir un rôle de protéine chaperon. Citons notamment le cas de ClpB qui ne présente pas le motif (LIV)G(FL) requis pour l'interaction avec ClpP et qui, induite en conditions de stress, aurait un rôle de chaperon protéique (Chastanet et al. 2004). Les gènes clpP codant la sous-unité protéolytique et clpL, clpX, clpE, clpC codant différentes sous-unités liant l'ATP sont présents dans le génome de *L. bulgaricus*. Les gènes clpY et clpQ, présents eux aussi, codent pour un système de dégradation indépendant du système ClpP et dont il a récemment été montré qu'il a un rôle mineur dans la résistance aux stress de *S. aureus* (Frees et al. 2005).

Chez *B. subtilis*, bactérie modèle des firmicutes, les régulateurs des protéines chaperons et des complexes Clp sont respectivement HrcA et CtsR. Notre analyse de *L. bulgaricus* révèle la présence d'un homologue de HrcA mais aucun gène susceptible de coder CtsR. Soulignons qu'en amont des clpP, clpC et clpE d'ATCC 11842, aucune séquence comparable au site de fixation de CtsR (A/GGTCAAANANA/GGTCAAA, deux dégénérescences admises pour la recherche du motif dans le génome d'ATCC 11842) n'est identifiée. Au contraire, une séquence proche du consensus TTAGCACTC-N9-GAGTGCTAA correspondant au site de

Analyse bibliographique

fixation d'HrcA est présente en amont des opérons *groESL* et *hrcA/grpE/dnaK/dnaJ*. Cette séquence est également retrouvée en amont de *clpP*, *clpE* et de *ldb0289* codant l'homologue de Hsp18 (figure 14) mais pas en amont des gènes codant les autres sous-unités Clp (*clpL*, *clpX* et *clpC*), ni en amont de *ldb0364* (Hsp15), de *ldb0369* (*hslO*, Hsp33), *clpY* et *clpQ*. Ces observations nous conduisent à proposer que chez *L. bulgaricus*, le rôle de HrcA soit plus important que chez les autres firmicutes et qu'il pallie en partie l'absence de CtsR. Dans l'état actuel, nous ne pouvons cependant pas affirmer qu'HrcA est le seul régulateur de la réponse au stress thermique chez ATCC 11842. En effet, bien qu'aucune séquence conservée ne soit identifiée en amont de *clpX*, *clpL*, *clpC*, *ldb0364*, *ldb0369*, *clpY* et *clpQ*, il faudrait tester si ces gènes répondent au choc thermique avant de conclure.

<i>clpP</i>	TTAGCACTGTATCTAAATGAGTGCTAA
<i>groES (1)</i>	TTAGCACTTGATCAAGATGAGTGCTAA
<i>groES (2)</i>	TTAGCACTCTCGACTAAAAAGTGCTAT
<i>ldb0289</i>	TTAGCAGTCAAAAAAGTTGAGTGCTAA
<i>clpE</i>	TTAGCACTTGACTAACTAGAGTGCTAA
<i>hrcA</i>	TTAGCATTCTCCTTGACTAACTGCTAA
consensus	TTAGCAcTcnnnnnnnnngAgTGCTAa

Figure 7 : séquences CIRCE putatives détectées en amont des gènes codant les protéines de choc thermique. Les bases conservées sont en bleu.

Chez *L. bulgaricus*, la régulation des gènes de réponse au stress hyperthermique semble donc à l'opposé de celle d'*O. oeni* qui ne possède pas HrcA mais dont les opérons codant les chaperons DnaK et GroESL sont sous contrôle de CtsR (Grandvalet et al. 2005).

Nous avons regardé quelle était la situation chez les autres lactobacilles dont les génomes sont séquencés et disponibles. Sur les quatre autres génomes de lactobacilles disponibles, seuls *L. sakei* et *L. plantarum* possèdent CtsR. Les lactobacilles du groupe *L. acidophilus* ne possèdent que HrcA. Chez *L. johnsonii* et *L. acidophilus*, on retrouve des boîtes CIRCE en amont des deux opérons codant les chaperons (dont l'organisation est bien conservée) mais aussi en amont de *clpP* et des deux copies de *clpE*. Les homologues de *ldb0289* (*lba0205* et *lj_0181*) semblent eux aussi régulés par HrcA puisqu'ils possèdent une boîte CIRCE. Aucune autre séquence CIRCE n'est détectée dans ces génomes. Le régulon HrcA semble donc conservé au

Analyse bibliographique

sein du groupe *L. acidophilus*. Chez *L. plantarum* et *L. sakei* qui n'appartiennent pas au groupe *L. acidophilus*, où *ctsR* est présent, on retrouve des boîtes CIRCE en amont des opérons *groESL* et *hrcA-grpEdnaK- dnaJ* ainsi qu'en amont des homologues de *ldb0289* (*lsa0050*, *lp_3352* et *lp_0129*). Cependant, aucune boîte CIRCE n'est détectée en amont des gènes codant les sous-unités Clp. D'autre part, la boîte de reconnaissance de CtsR est détectée en amont de *ctsR/clpC* et de *clpE* dans ces deux génomes (figure 15).

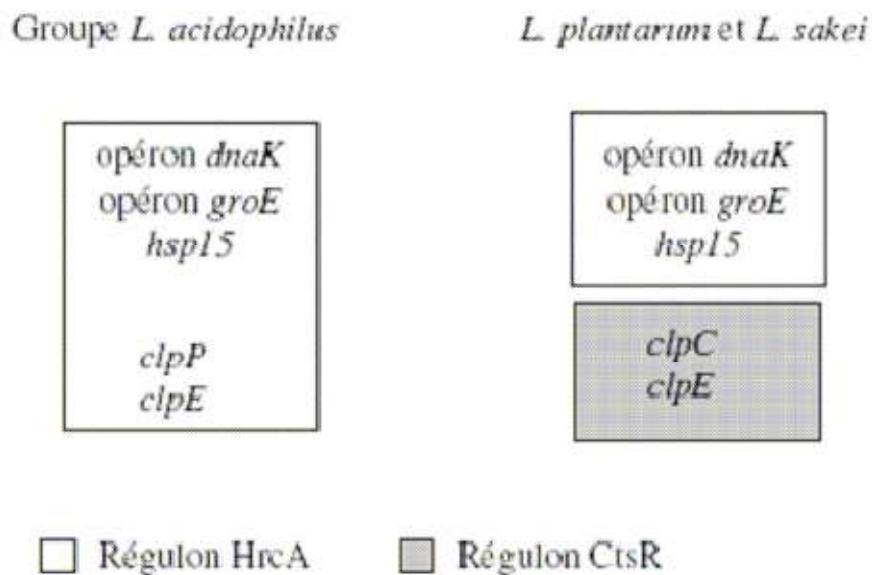


Figure 8 : Modèle de la régulation des opérons de réponse au choc thermique chez les lactobacilles. Modèle basé sur l'analyse in silico.

5.5. *Lactococcus lactis* et le stress hyperthermique

Les bactéries lactiques, et en particulier *Lactococcus lactis*, jouent un rôle primordial dans les premières étapes de la transformation du lait, mais elles interviennent aussi, directement et indirectement, dans la phase d'affinage et dans la qualité sanitaire des produits.

La croissance des bactéries en général, et de *Lactococcus lactis* en particulier, dépend des conditions nutritionnelles (sources de carbone et d'azote...) et physico-chimiques (pH, température, salinité...). Les conditions de culture en laboratoire sont souvent optimisées et contrôlées pour assurer une croissance rapide. A l'inverse, *Lactococcus lactis*, que ce soit dans son habitat naturel (plantes, sol) ou lors de sa mise en œuvre dans les processus industriels, doit faire face à de multiples stress nutritionnels ou physico-chimique (thermique,

Analyse bibliographique

oxydatif, acide, osmotique) parfois concomitants (Bunthof et al. 1999; Sanders et al. 1999; Stuart et al. 1999).

En effet, *L. lactis* est soumis à de fortes variations de températures dans le sol ou lors de la fabrication de fromages, pour lequel la température monte jusqu'à 40°C. A l'inverse, la température est beaucoup plus faible pendant l'affinage (8-16°C) ou le stockage du fromage. De même, la production de lactate dans les levains provoque, quant à elle, une acidification croissante du milieu (lait). Dans ce cas, les bactéries sont elles-mêmes à l'origine du stress acide. La pression osmotique est également susceptible de varier significativement, notamment lors du pressage. Cette étape peut également conduire à une limitation carbonée par l'élimination du lactose dans le lactosérum (Stuart et al. 1999).

Les mécanismes de perception du stress, essentiels pour que la cellule puisse se protéger avant que les effets de ce stress ne soient irréversibles, sont encore peu connus chez *Lactococcus lactis*, mais semblent se faire chez les autres bactéries, soit au niveau membranaire, soit au niveau du cytoplasme (Rallu et al. 2000).

Les effets du stress peuvent être visualisés à plusieurs niveaux : modification du catabolisme, de l'anabolisme, de l'état énergétique, des constituants cellulaires...

Face aux différents effets du stress, les bactéries disposent de plusieurs types de réaction qui se répartissent en trois catégories (Rallu et al. 2000). La fuite consiste, pour les bactéries capables de motilité, à rechercher des conditions plus favorables à la croissance, mais *Lactococcus lactis* qui est dépourvu de flagelle ne peut pas utiliser cette tactique. Le refus a pour but de diminuer l'intensité du stress, par exemple en utilisant une autre source de nutriment lorsqu'elle est épuisée. Enfin, la résistance devient nécessaire lorsque le stress est persistant et/ou dommageable pour les constituants cellulaires.

Elle peut être acquise soit par une modification irréversible du matériel génétique (mutation, remaniement) conduisant à une meilleure tolérance aux conditions environnementales, soit par l'induction de systèmes de protection et de lutte contre le stress ou de réparation contre les dommages causés.

Au cours de la fabrication de certains fromages, les fortes températures rencontrées au cours de la « cuisson » peuvent être à l'origine d'altérations cellulaires majeures. Les protéines intracellulaires vont se dénaturer puis s'agréger, diverses macromolécules comme les ribosomes ou l'ARN vont être déstabilisées et la fluidité membranaire va augmenter posant

Analyse bibliographique

des problèmes de perméabilité excessive et de transport (Earnshaw et al. 1995; Teixeira et al. 1997; Hansen et al. 2001).

La réponse au choc thermique ou *heat shock* (HS) a été très étudiée, notamment chez les bactéries lactiques. Ce type de stress provoque, chez *L. lactis*, l'induction de 17 protéines dont des protéines chaperonnes (DnaK, DnaJ, GrpE, GroES, GroEL) et des protéases (Clp, HtrA, FtsH) (Kilstrup et al. 1997). Les chaperonnes sont impliquées dans le repliement des protéines néo-synthétisées et la re-conformation des protéines dénaturées.

La régulation de ces protéines, dont les gènes sont organisés en deux opérons, est assurée par la protéine HrcA qui reconnaît une séquence opératrice appelée CIRCE identifiée en amont de différents gènes chez *L. lactis* : dnaJ (van Asseldonk et al. 1993), hrcA-grpE-dnaK (Eaton et al. 1993).

Différents types de protéases sont induits par un stress hyper-thermique. HtrA, présent chez les lactococci (Poquet et al. 2000), est une sérine-protéase qui présente des fonctions de chaperonne ou de protéase selon la température, chaperonne à basse température et protéase quand celle-ci est normale ou élevée (Spiess et al. 1999). La protéine FtsH, qui est une métallo-protéase ATP-dépendante liée à la membrane, est essentielle chez *E. coli*. Le gène, également identifié chez *L. lactis* (Nilsson et al. 1994), est induit par une augmentation de la température, et RecA pourrait être impliqué dans sa régulation (Duwat et al. 1995).

Le dernier type de protéase regroupe les protéases caséinolytiques, Clp. ClpP est une sérine-protéase multimérique qui dégrade les peptides inférieurs à 7 acides aminés. Lui sont associées des Clp-ATPases, 4 chez *L. lactis* (clpB, clpC, clpE, clpX), qui peuvent avoir une activité de chaperonne. ClpP peut s'associer à une Clp-ATPase pour former un complexe protéolytique dont la spécificité de substrat sera déterminée par l'ATPase. Après un choc thermique, ces protéines vont être responsables de la dégradation des protéines qui n'ont pût être repliées correctement. La transcription des gènes clpB, C, E et P est réprimée par la protéine régulatrice CtsR (Varmanen et al. 2000), mais TrmA semble également impliquée dans le contrôle de ces gènes.

5.6. Les protéines de choc thermique chez *Lactococcus lactis* : synthèse et régulation; thermotolérance

5.6.1. INTRODUCTION

Il y a une cinquantaine d'années, Ritossa (1962) (Ritossa 1996) rapportait, dans une courte note, un étrange phénomène qu'il avait observé dans un élevage de drosophiles, soustrait à son environnement thermique habituel (25 °C): «Si les larves écrivait-il, sont soumises à un choc thermique à 30 °C pendant 30 minutes environ, il s'ensuit un brutal changement du profil des «putts» ou boursouflures des chromosomes ». Ces «gonflements» qui intéressent toujours les mêmes régions des chromosomes polytènes des glandes salivaires des diptères traduisent une intense activité transcriptionnelle : «Les puffs sont des gènes actifs» ajoutait Ritossa.

L'apparition de ces hernies, ou nodules, est suivie, quelques minutes après par la synthèse de protéines inhabituelles (Tissieres et al. 1974), qui furent appelées hsps (*heat shock proteins*). Cette première observation ouvrit des perspectives fertiles; le même phénomène, marqué par un bouleversement du spectre protéique, consécutif à une augmentation de température, fut reconnu chez de très nombreuses espèces animales ou végétales avant d'être identifié chez *Escherichia coli* (Lemaux et al. 1978; Yamamori et al. 1978). La réponse à ce «choc» thermique, comme toutes celles déclenchées par d'autres stress physico-chimiques, permet aux organismes de mettre en place des dispositifs physiologiques temporaires qui assurent un ajustement des voies métaboliques aux nouvelles contraintes ou menaces de l'environnement. Ces réponses sont bénéfiques en ceci qu'elles confèrent aux organismes la capacité de survivre, du moins transitoirement, à des conditions ou épreuves létales, homologues ou proches de celles qui en ont commandé le déclenchement.

On nommera tolérance cette amélioration phénotypique de la survie et, plus précisément thermotolérance, l'exhaussement de la résistance «naturelle» de l'organisme à la chaleur. Les deux phénomènes synthèse d'hsps d'une part, thermotolérance d'autre part, coexistent. Sont-ils pour autant liés par un principe de causalité (Boutibonnes et al. 1992): la question est toujours irrésolue ... Quelle que soit l'assise biochimique de la tolérance, la réponse bipartite à médiation biochimique et physiologique est universelle: présente chez tous les êtres vivants, chez toutes les cellules en culture, qu'elles soient embryonnaires, adultes ou néoplasiques, elle se déroule selon les mêmes modalités qui font intervenir successivement: un «stresseur» ou stimulateur, la chaleur bien sûr, mais aussi tout autre agent dit thermomimétique qui, comme l'alcool, produit un déclenchement analogue; un capteur ou «senseur» qui est une cible

Analyse bibliographique

cellulaire sensible et qui dans le cas présent est une (ou des) protéine (s) ; un signal de nature inconnue, sinon à considérer que la protéine altérée assure elle-même cette fonction; un effecteur responsable de la bascule transcriptionnelle qui conduit à la «différenciation métabolique» ; des modulateurs ou régulateurs négatifs qui tempèrent l'amplitude de la réponse.

Universelle, la réponse est transitoire: elle est manifeste quelques minutes après le choc mais elle est de courte durée. En général 1 ou 2 h après le déclenchement du dispositif, les organismes ou les cellules retournent à un état de repos qui est une nécessité, pour obéir à une nouvelle sollicitation du milieu extérieur. Universelle et transitoire, la réponse qui s'est sans doute mise en place précocement au cours de l'évolution, est conservée dans tout le monde vivant: on retrouve, même si elles portent des noms différents les mêmes protéines présentant, au moins dans leurs zones fonctionnelles, les mêmes structures, et assurant les mêmes fonctions; aussi en est-il de la hsp 60, de la cpn 60 des cellules eucaryotes et de GroEL d'E coli.

Elles possèdent toutes trois la propriété de veiller au repliement orthodoxe des protéines immatures, ou de restaurer la conformation native de leurs structures anormales irrégulières ou agrégées, provoquées par des agents chaotropiques.

Parce qu'elle est universelle et conservée, parce qu'elle est la plus anciennement connue, parce qu'elle permet aux organismes de s'affranchir de l'impérieux pouvoir du réel, la réponse au choc thermique peut être considérée comme l'exemple des stratégies d'adaptation ou d'ajustement (coping) à l'environnement. Elle représente aussi un modèle simple et efficace de différenciation programmatique : son recours puis sa mise en œuvre révèle, au sein du génome, l'existence de projets enfouis, occasionnellement réactivés. En cela, elle ne cesse de nous fasciner ...

Il semble inévitable, lorsqu'on aborde chez les bactéries la physiologie ou la régulation de la réponse au choc thermique, de se référer, à ce qui, aujourd'hui, fait figure de système codifié, voire de paradigme: le modèle E coli; non seulement parce qu'il est accompli un deuxième régulon intriqué au système multigénique classique et le complétant vient d'être individualisé (Raina et al. 1995; Rouviere et al. 1995) mais aussi à cause de sa «beauté» et, de son évidente simplicité: ainsi la bascule métabolique se noue autour d'un relais de facteurs sigma. Nous y ferons bien sûr référence.

Cependant *Lactococcus lactis* se distingue de l'archétype bactérien par quatre traits singuliers: c'est une bactérie Gram positive, mésophile, microaérophile qui a emprunté des voies évolutives distinctes de celle des protéobactéries ; son GC % est bas, 34 à 38 % (Premi et al. 1972) contre 48 à 52 % pour *E coli* (Hickman-Brenner et al. 1984); son génome est de petite taille: de 2 à 2,7 MB (Tanskanen et al. 1990; Le Bourgeois et al. 1995), alors qu'il est de 4,2 à 4,5 MB chez *E coli*. Enfin sa niche écologique, le lait, les produits laitiers, est éloignée de celle des bactéries entériques (Premi et al. 1972); *L lactis* y est soumis, au cours des processus de biotransformation auxquels il participe à de nombreuses sautes thermiques. Ces constatations laissent supposer que la réponse, même si elle est phylogéniquement conservée est, chez cette espèce, assez divergente du modèle conventionnel accepté. Évoquer la réponse au choc thermique, et plus particulièrement les hsp qui en sont la manifestation la plus flagrante, c'est décrire: (1) les circonstances de leur apparition, leurs nombre et qualités; (2) le mode de régulation de leur synthèse; (3) leur participation hypothétique au développement de la thermotolérance.

5.6.2. Les protéines de choc thermique

La plupart des bactéries ont développé au cours de l'évolution des réponses moléculaires rapides qui leur permettent de survivre aux conditions hostiles auxquelles elles sont exposées au cours de leur vie saprophytique. Elles supportent ainsi des écarts de température, des carences nutritionnelles, des variations de pH et d'osmolarité, des stress chimiques ou encore la compétition d'autres micro-organismes. Un réseau complexe et finement régulé de protéines effectrices s'est constitué pour maintenir la viabilité cellulaire lors de l'exposition au stress, en prévenant notamment l'accumulation délétère de protéines endommagées. Chez *Escherichia coli* et *Bacillus subtilis*, ainsi sans doute que chez la plupart des espèces bactériennes, c'est l'induction d'homologues des *heat shock proteins* (Hsp) eucaryotes qui permet le recyclage des protéines dénaturées par des conditions hostiles. Ces enzymes agissent selon l'état d'altération des protéines, soit en restaurant une conformation active (action des « chaperonines »), soit en assurant leur protéolyse (action des protéases). Ce contrôle de qualité des protéines est nécessaire au maintien du fonctionnement des voies physiologiques essentielles des bactéries agressées. Chez les bactéries pathogènes, ces mécanismes peuvent être mis à profit pour résister à la réponse de l'hôte infecté, et sont même indispensables à l'expression de la virulence de certaines espèces.

5.6.2.1. Les chaperonines : dissociation des agrégats et remodelage des protéines

Les principales protéines impliquées dans le remodelage de protéines anormales (Ben-Zvi and Goloubinoff 2001) sont les chaperonines GroEL et DnaK, équivalents respectifs chez *E. coli* des Hsp60 et Hsp70 eucaryotes, et les chaperonines-ATPases homologues des Hsp100 eucaryotes HslU, ClpA, ClpX, ClpC, ClpE, ClpB, ClpR... Ces dernières, ou « Clp-ATPases », sont réunies avec la protéase ClpP sous le vocable de caseinolytic proteins bien que seule ClpP soit dotée d'activité protéasique sensu stricto.

Chez *E. coli*, GroEL prend en charge des protéines anormalement repliées ou dénaturées et leurs agrégats insolubles, ainsi que des chaînes polypeptidiques non repliées. Associée à la protéine GroES, GroEL assure la dissociation des agrégats et la restauration d'une conformation active de ces substrats. Chez la bactérie pathogène *Listeria monocytogenes*, cette activité « chaperonne » est indispensable à l'expression de la virulence, et la transcription du gène groEL est fortement accrue lors de l'infection.

DnaK et ses cofacteurs DnaJ et GrpE ont une activité similaire à celle de GroESL (y compris la dissociation d'agrégats protéiques toxiques), mais qui s'exerce sur des substrats différents. De plus, chez *E. coli*, DnaK intervient lors de la dégradation de certains substrats par les protéases Lon ou FtsH. Il existe un homologue de dnaK chez plusieurs espèces pathogènes, dont le rôle n'est qu'incomplètement élucidé.

Les homologues bactériens des Hsp100 (ClpA, ClpX, ClpC, ClpE, ClpL, ClpB, HslU...) forment un groupe d'ATPases très conservées à activité chaperonne. Assemblées en hexamères, elles catalysent le dépliement de substrats protéiques spécifiques. Les polypeptides libérés peuvent alors connaître deux sorts distincts, soit un remodelage en protéine native, soit une dégradation si leur structure est irrémédiablement altérée. Cette dégradation est alors assurée par des complexes qui associent une Clp-ATPase et ClpP ou HslU et la protéase HslV.

5.6.2.2. La protéolyse ATP-dépendante chez les bactéries

La protéolyse est un élément fondamental de la régulation des processus métaboliques et de la réponse cellulaire aux stimuli environnementaux. Elle contrôle les voies métaboliques, notamment en modulant les quantités des protéines régulatrices. Chez les procaryotes et les organites cellulaires d'origine procaryote comme les mitochondries ou les chloroplastes, l'essentiel de l'activité protéolytique est assurée par des protéases ATP-dépendantes. Ces

Analyse bibliographique

enzymes jouent un rôle central dans de nombreuses cascades régulatrices (Gottesman 1999; Msadek 1999) et sont indispensables à l'élimination des protéines anormales continuellement produites par dénaturation spontanée, erreurs de biosynthèse ou accumulation de mutations. Leur action est particulièrement vitale lors de l'exposition bactérienne aux stress qui détériorent les protéines. La plupart des protéases sont d'ailleurs inductibles par ces stress pour prévenir l'accumulation de polypeptides anormaux potentiellement toxiques et maintenir ainsi l'homéostasie cellulaire et les activités métaboliques de la bactérie.

Les cinq principales protéases ATP-dépendantes caractérisées chez *E. coli* sont ClpAP, ClpXP, HslUV, Lon et la protéase membranaire FtsH. Des homologues de ces protéases ont été identifiés chez la plupart des espèces bactériennes. Leurs substrats restent largement méconnus. Quelques protéines et polypeptides naturels ou semi-synthétiques spécifiquement dégradés par l'une ou l'autre ont néanmoins été identifiés, ce qui a permis d'appréhender les mécanismes par lesquels les substrats sont reconnus, captés puis transformés. Il est ainsi possible d'établir un schéma de la protéolyse, commun à toutes les protéases ATP-dépendantes, dans lequel la reconnaissance du substrat, son dépliement et son introduction dans la chambre protéolytique de la partie protéase sont assurés par la partie ATPase. Les activités protéase et chaperonine/ATPase peuvent être présentes sur la même molécule (Lon et FtsH), ou le fait d'éléments distincts au sein d'un complexe protéolytique (ClpAP, ClpXP et HslUV).

5.6.2.3. Identification des substrats

Afin d'éviter la dégradation inappropriée de composants cellulaires essentiels, les protéases assurent une reconnaissance précise de leurs substrats. Dans les cellules eucaryotes, les substrats destinés à la protéolyse sont reconnus par des systèmes enzymatiques variés, mais qui réalisent tous un « étiquetage » par une même protéine, l'ubiquitine ; c'est ce marquage qui les désigne au seul protéasome 26S dans lequel s'effectue la protéolyse (Ulrich 2002). À l'inverse, chez les procaryotes, il n'existe ni étiquetage ubiquitaire ni protéasome unique. L'orientation vers l'une des protéases dépend de la nature des substrats eux-mêmes et d'un étiquetage spécifique. Il peut s'agir d'une séquence d'acides aminés du substrat natif, ou d'un ajout co- ou post-translationnel, dont plusieurs modèles ont été récemment décrits. Parmi ceux-ci, l'étiquetage SsrA consiste en l'ajout de 11 acides aminés à l'extrémité C-terminale de tous les polypeptides restés bloqués sur leur ribosome en raison d'une troncature de leur ARNm (Keiler et al. 1996). La protéine anormale qui en résulte est adressée par cette étiquette à la protéolyse ATP-dépendante. Dans d'autres situations, c'est une protéine accessoire qui

tient lieu d'étiquette et désigne le substrat à sa protéase ATP-dépendante (Zhou et al. 2001). Les étiquettes de reconnaissance identifiées à ce jour sont toutes situées aux extrémités C- ou N-terminale des substrats. Lorsqu'un substrat protéique porte les motifs de reconnaissance de plusieurs protéases différentes, il peut être pris en charge par chacune d'entre elles.

5.6.2.4. Reconnaissance et fixation du substrat par l'ATPase

L'amarrage des substrats destinés à la protéolyse s'effectue exclusivement au niveau de la partie chaperonine. Les mécanismes de reconnaissance des substrats étiquetés semblent impliquer des domaines appelés SSD (sensor-substrate discrimination), présents à la fois chez Lon, ClpA, ClpX, ClpB, ClpE et ClpY et capables de fixer les étiquettes de protéolyse. La spécificité de ces régions d'amarrage n'est cependant pas absolue et il faut sans doute plusieurs séquences « réceptrices » différentes à une chaperonine pour assurer la reconnaissance de ses substrats. Enfin, l'amarrage nécessite la fixation d'ATP, mais pas son hydrolyse.

5.6.2.5. Transport du substrat jusqu'au site de protéolyse

L'amarrage seul ne modifie pas la conformation du substrat. La déstabilisation des structures tertiaires et secondaires requiert l'hydrolyse d'ATP. Il est vraisemblable que la linéarisation indispensable à l'accès au site de protéolyse s'effectue dans une cavité formée par la structure tertiaire ou par l'oligomérisation des sous-unités de chaperonine (Bukau and Horwich 1998). Le maintien de la chaîne dans une conformation linéaire puis sa translation vers le site protéolytique nécessitent à leur tour l'hydrolyse de nouvelles molécules d'ATP. Au fur et à mesure du dépliement, de nouveaux déterminants du substrat sont exposés et peuvent réagir avec de nouvelles régions de la chaperonine, favorisant la translation (Ishikawa et al. 2001). Celle-ci est directionnelle, l'extrémité C-terminale entrant par exemple la première dans la chambre de protéolyse lorsqu'un substrat porteur d'une étiquette SsrA est pris en charge par la protéase ClpAP (Reid et al. 2001).

5.6.2.6. Protéolyse et évacuation des produits de dégradation

Le site actif des protéases est isolé du milieu extérieur dans une cavité peu accessible formée par la structure tertiaire ou l'assemblage de plusieurs sous-unités. La chaîne polypeptidique doit être dévidée en un brin unique pour franchir l'étroit passage d'environ 1 nm qui commande l'accès à cette chambre de protéolyse. À l'intérieur de la chambre, le substrat est tronçonné dans une nouvelle réaction ATP-dépendante. Les oligopeptides ainsi produits semblent modifier l'équilibre ionique dans la chambre de protéolyse, créant des ouvertures

dans les parois, par lesquelles ils s'échappent. Enfin, la dégradation en acides aminés est assurée dans le cytoplasme par des peptidases non spécifiques ATP-indépendantes.

5.6.2.7. Les principales protéases ATP-dépendantes bactériennes

5.6.2.7.1. La sérine-protéase Lon

La sérine-protéase Lon est une protéase cytosolique tétramérique dont chaque sous-unité porte un site sérine-protéase et un domaine ATPase distincts. Chez *E. coli*, Lon est la protéase la plus impliquée dans la dégradation des protéines de conformation anormale (Gottesman 1996). Étonnamment, il n'existe pas d'homologue du gène lon chez des bactéries pathogènes comme *L. monocytogenes* ni chez *Streptococcus pneumoniae*, bien que Lon soit assez ubiquitaire et présente par exemple chez les mycoplasmes dont le génome réduit n'encode aucune autre protéase ATP-dépendante.

5.6.2.7.2. La métalloprotéase FtsH

La métalloprotéase FtsH est ancrée dans la membrane cytoplasmique. Son domaine cytoplasmique contient la région ATPase et le site métalloprotéase Zn-dépendant. Plusieurs substrats de FtsH ont été identifiés chez *E. coli*, notamment des polypeptides étiquetés par le système SsrA. FtsH est capable de linéariser puis de dégrader des protéines enchâssées dans la membrane cytoplasmique, après reconnaissance de leur extrémité C- et/ou N-terminale.

5.6.2.7.3. Les complexes protéasiques de ClpP

5.6.2.7.4. ClpP

ClpP est une sérine-protéase présente chez les bactéries à l'exception des mycoplasmes, ainsi que dans les mitochondries et les chloroplastes des cellules eucaryotes. Chez ces derniers, ClpP joue sans doute un rôle essentiel, puisqu'elle est l'une des quatre protéines encodées par l'ADN du chloroplaste vestigial de *Epifagus virginiana*, une plante parasite ayant perdu toute capacité de photosynthèse. Chez les bactéries, ClpP est présente dans les conditions normales de croissance, mais fortement hyperproduite dans de nombreux types de stress, notamment les chocs thermiques. L'inactivation du gène clpP affecte la viabilité bactérienne en conditions de stress, surtout chez des bactéries à Gram positif comme *B. subtilis*, *L. monocytogenes*, *S. pneumoniae* et *Staphylococcus aureus*. L'oligomérisation de ClpP en deux cycles heptamériques accolés (ClpP₁₄) est indispensable à son activité biologique, laquelle se limite

à la dégradation de peptides de moins de sept acides aminés. C'est l'association de ClpP14 à une chaperonine-ATPase qui permet la protéolyse de substrat de plus grande taille.

5.6.2.7.5. Les chaperonines-ATPases partenaires de ClpP

ClpA, ClpX, ClpC et ClpE sont les principaux partenaires possibles de ClpP. Dans chaque génome bactérien séquencé à ce jour et contenant un gène *clpP*, on a pu identifier au moins un gène de l'une de ces Clp-ATPases. Toutes les Clp-ATPases possèdent une structure commune, incluant un domaine N-terminal SSD de reconnaissance des substrats précédé d'un ou deux domaines de fixation de l'ATP. ClpA de *E. coli* a été la première identifiée et constitue un bon modèle du fonctionnement des autres Clp-ATPases qui interagissent avec ClpP. Il n'existe cependant pas d'orthologue de *clpA* chez les bactéries à Gram positif, chez lesquelles elle semble remplacée par ClpC. En présence d'ATP, six sous-unités de cette structure sont nécessaires à l'activité chaperonne (linéarisation et remodelage) et à l'arrimage d'un tétradécamère ClpP14. Il existe probablement une certaine souplesse des hexamères de Clp-ATPase qui autorise la liaison dissymétrique avec les heptamères annulaires de ClpP (Kim et al. 2001b). Parmi les substrats reconnus par ClpA et dégradés par le complexe ClpAP, on trouve des protéines à étiquette SsrA et la protéine RepA. Ishikawa et al. ont récemment montré par cryomicroscopie électronique le processus de translation de RepA dans la protéase ClpAP. L'attachement à l'ATPase a été confirmé ainsi que le transfert ATP-dépendant du substrat jusqu'à la chambre de protéolyse, via l'intérieur de l'anneau ClpA6 (Ishikawa et al. 2001).

ClpX est une ATPase ubiquitaire retrouvée chez les bactéries à Gram positif comme à Gram négatif. Elle se distingue de ClpA par une spécificité de substrats différente. On connaît cependant des substrats communs à ClpXP et ClpAP. Pour ceux-ci, la dégradation par ClpXP est moins efficace que par ClpAP, sans doute en raison de la présence de deux domaines ATPases chez ClpA alors qu'il n'en existe qu'un seul chez ClpX.

ClpC et ClpE ont une structure qui suggère un mode d'appariement et un fonctionnement voisins de ceux de ClpA. ClpC semble être avec ClpX le principal partenaire de ClpP chez les bactéries à Gram positif, dans lesquelles elle est essentielle au catabolisme des protéines anormales et à la survie en conditions de stress (Msadek et al. 1994; Rouquette et al. 1996). Le rôle de ClpE est encore méconnu ; le gène *clpE* caractérisé chez *B. subtilis* puis *L. monocytogenes* n'est pas indispensable à la survie de ces bactéries, mais jouerait un rôle dans la virulence listérienne (Rouquette et al. 1996).

Analyse bibliographique

Chez *Bacillus subtilis*, ClpP peut être associée à ClpC, ClpX et ClpE, chacune de ces protéines étant inductible par différents stress. ClpCP et ClpXP sont directement impliquées dans la dégradation des protéines dénaturées et des substrats étiquetés par SsrA. ClpP, ClpC et ClpX sont nécessaires à la croissance à température élevée ou en milieu hypersalé et à la mobilité bactérienne (Msadek et al. 1998). À l'inverse, aucun phénotype ne semble associé à l'inactivation de clpE. En plus de la résistance au stress, ClpP, ClpC et ClpX sont indispensables au fonctionnement de deux voies essentielles du développement de *B. subtilis*, la sporulation et la compétence. L'inactivation de ClpP confère le déficit de sporulation le plus net, alors que la sporulation des mutants clpC n'est perturbée qu'à haute température. ClpP est également indispensable au développement de la compétence, c'est-à-dire de la propriété qu'ont les bactéries en fin de phase exponentielle de capter et d'internaliser de l'ADN exogène (Msadek 1999).

Le rôle des Clp des mycobactéries n'a pas été élucidé, mais il existe des anticorps spécifiques dirigés contre ClpC de *Mycobacterium leprae* chez des patients atteints de lèpre ou de tuberculose. Chez les *Streptomyces*, plusieurs isomères de ClpP ont été caractérisés et sont impliqués dans la formation du mycélium de ces bactéries telluriques. La surexpression de ClpX permet la pigmentation tandis que la surexpression de ClpP1 et ClpP2 active la formation du mycélium de *Streptomyces coelicolor* (de Crecy-Lagard et al. 1999).

Chez *Lactococcus lactis*, ClpP est inductible par la chaleur et les variations de pH. L'inactivation de ClpP chez *L. lactis* entraîne une réponse médiocre à ces stress et une perte importante de viabilité des bactéries (Frees and Ingmer 1999).

Chez *Staphylococcus aureus*, clpC fait partie des gènes plus exprimés dans les bactéries constituées en biofilm que dans les cellules planctoniques.

Le génome de *Streptococcus pneumoniae* contient des orthologues de clpP, clpC, clpX et clpE dont l'organisation est similaire à celle des gènes de *B. subtilis*. ClpC a été identifiée comme un possible « facteur de virulence » lors d'un criblage génomique par transposon-tagged mutagenesis [24], alors que d'autres travaux mettent plutôt en avant le rôle de ClpE, ClpX et surtout ClpP dans la résistance au stress, le développement de la compétence et la virulence du pneumocoque [25]. Chez *Streptococcus agalactiae*, ClpP module l'activité de la chaperonine DnaK, mais ne semble jouer qu'un rôle mineur dans la virulence [26].

Lorsque l'on considère les agressions que subit *Listeria monocytogenes* lors du processus infectieux (variations de pH et de température, stress oxydatif dans les phagolysosomes des

cellules infectées, carences nutritionnelles, etc.), on conçoit que les protéases ATP-dépendantes soient essentielles à la survie de cette bactérie chez l'hôte. L'organisation génétique et la régulation de l'expression des Clp-protéases de *L. monocytogenes* sont identiques à celles décrites chez *B. subtilis*. ClpC est la première ATPase à avoir été impliquée dans la résistance au stress de *L. monocytogenes* et dans la physiopathologie de la listériose expérimentale [15]. ClpC est nécessaire à l'échappement des bactéries des phagosomes et à leur survie dans les macrophages [27] et semble moduler la production d'internaline, principale adhésine impliquée dans l'entrée de *Listeria* dans les cellules eucaryotes [28]. ClpE est requise pour la survie prolongée à haute température et la division cellulaire, mais seule l'inactivation simultanée de *clpE* et *clpC* produit un mutant avirulent [16]. En revanche, l'inactivation du gène *clpP* entraîne à elle seule la disparition complète de la virulence, ClpP étant indispensable à la croissance dans la plupart des conditions de stress [29]. Le plus intrigant chez *L. monocytogenes* est que ClpP paraît moduler l'activité biologique du principal facteur de virulence de la bactérie, la listériolysine O (LLO). Chez un mutant *clpP* – exposé un stress thermique, l'activité cytotolytique in vitro de la LLO est fortement réduite, alors que la quantité de LLO produite est identique à celle d'une souche isogénique sauvage [30]. Il est possible qu'en situation de stress, ClpP dégrade un inhibiteur encore inconnu de la LLO, favorisant ainsi l'expression de ce facteur de virulence lorsque la bactérie en a expressément besoin.

5.7. Régulation de la synthèse des protéines de choc thermique

Par quelle voie, et selon quel mode de régulation s'effectue la synthèse des hsp ?

Construire un modèle pertinent, c'est en premier lieu identifier les structures géniques qui commandent la formation pour, ensuite, se référer aux schémas bactériens établis: en particulier, à celui d'*E. coli* élevé au rang de modèle canonique mais aussi à celui moins élaboré, peut-être aussi plus archaïque de *Bacillus subtilis* (Arnosti et al, 1986). Six gènes des hsp majeures de *L. lactis* ont été jusqu'à ce jour reconnus, clonés puis séquencés. Un septième gène thermorégulé, homologue de *ftsH* (*hf/8*) d' *E. coli* vient d'être identifié (Nilsson et al, 1994) ; son produit, la protéine transmembranaire HfIB, douée d'une faible activité ATPasique participerait - comme DnaK, DnaJ, GrpE - à la régulation négative de la réponse subordonnée à σ^{32} dont il contrôlerait la dégradation et par là, le niveau endocellulaire (Herman et al, 1995 ; Duwat et al, 1995). A cet ensemble non clos, il faut ajouter le gène *recA*, pièce centrale du régulon SOS: son produit, la protéine Rec A, exercerait un

Analyse bibliographique

effet régulateur sur le système multigénique de choc thermique en supervisant le taux de HflB (Duwat et al, 1995). Les caractéristiques essentielles de ces gènes figurent dans les tableaux III et IV ; nous résumerons, ci-dessous, leurs particularités les plus saillantes.

Les gènes sont isolés ou groupés en petits opérons, qui rappellent ceux d'E coli (groEL-ES), (Kim et Batt, 1993), de 8subtilis (grpE-dnaK) (Barril et al, 1994), ou des deux espèces (rpoO-dnaE) (8 subtilis) ou (rpoO - dnaG) (E coli). Cependant, chez L lactis l'opéron dnaK comprend deux autres cadres ouverts de lecture (ort) (Eaton et al, 1993), et dnaJ, habituellement couplé à dnaK fait défaut (Van Asseldonk et al, 1993).

De la structure opéronique rpoO-dnaE, constituant l'opéron MMS (macro molecular synthesis) est absent, en amont, le gène rpsU dont le produit est impliqué dans la traduction (Araya et al, 1993). L'alignement des séquences en acides aminés des produits des gènes clonés, déduites des séquences nucléotidiques, avec celles des protéines homologues de diverses espèces bactériennes (E coli, 8 subtilis, 8 megaterium, Staphylococcus aureus), démontre une homologie importante (tableau IV). Cependant, chez L lactis comme chez les autres bactéries Grampositives on note la délétion de certaines parties: ainsi la région V2 du facteur sigma majeur d'E coli (rpoO) (Eaton et al, 1993) qui compte chez les bactéries Gram-négatives 250 acides aminés est réduite ici à quelques résidus (Araya et al, 1993 ; Gansel et al, 1993b) ; de même, la protéine DnaK, est amputée dans sa partie N terminale d'une séquence de 24 acides aminés ce qui justifie une masse moléculaire calculée plus faible que celle de la molécule homologue d'E coli (Eaton et al, 1993).

Les cadres ouverts de lecture ou les opérons de L lactis sont précédés:

- par un promoteur végétatif présentant des séquences -35 et -10 caractéristiques (TTGACA et TAAAAT) (Eaton et al, 1993) ;
- par une IR, (répétition inversée) de 9 à 12 pb formant une structure en tige-etboucle (stem-loop), identique à celle repérée chez 8 subtilis et chez d'autres bactéries Gram-positives (Eaton et al, 1993) ;
- par un site de fixation des ribosomes, (séquence de Shine Dalgarno) complémentaire de l'ARN 16S (AGGAG pour dnaK) ; chez certains d'entre eux (rpoO) sont inclus une séquence TGN, présente dans environ 50 % des gènes des bactéries Gram-positives et une autre répétition inversée faisant office de terminateur potentiel (Araya et al, 1993) ; aucun ne

Analyse bibliographique

montre le motif consensuel qui signe l'existence d'un promoteur reconnu par un facteur sigma alternatif, comme σ^{32} ou σ^{24} (σ^E) d'E coli (Lindquist, 1986) ;

- parmi une soixantaine de marqueurs, quatre de ces gènes sont localisés, (Le Bourgeois et al, 1992, 1993) (*rpoO*, *dnaK*, *dnaJ* et *grpE*) sur la carte génétique de *L. lactis* 1403 (Le Bourgeois et al, 1993), et de *L. lactis* subsp *cremoris* MG 1363 (Le Bourgeois et al, 1995). Sur la première, *dnaJ* flanque *rpoO* ; sur la seconde *dnaJ*, *rpoO* et *dnaK* divisent le chromosome en trois parties égales.

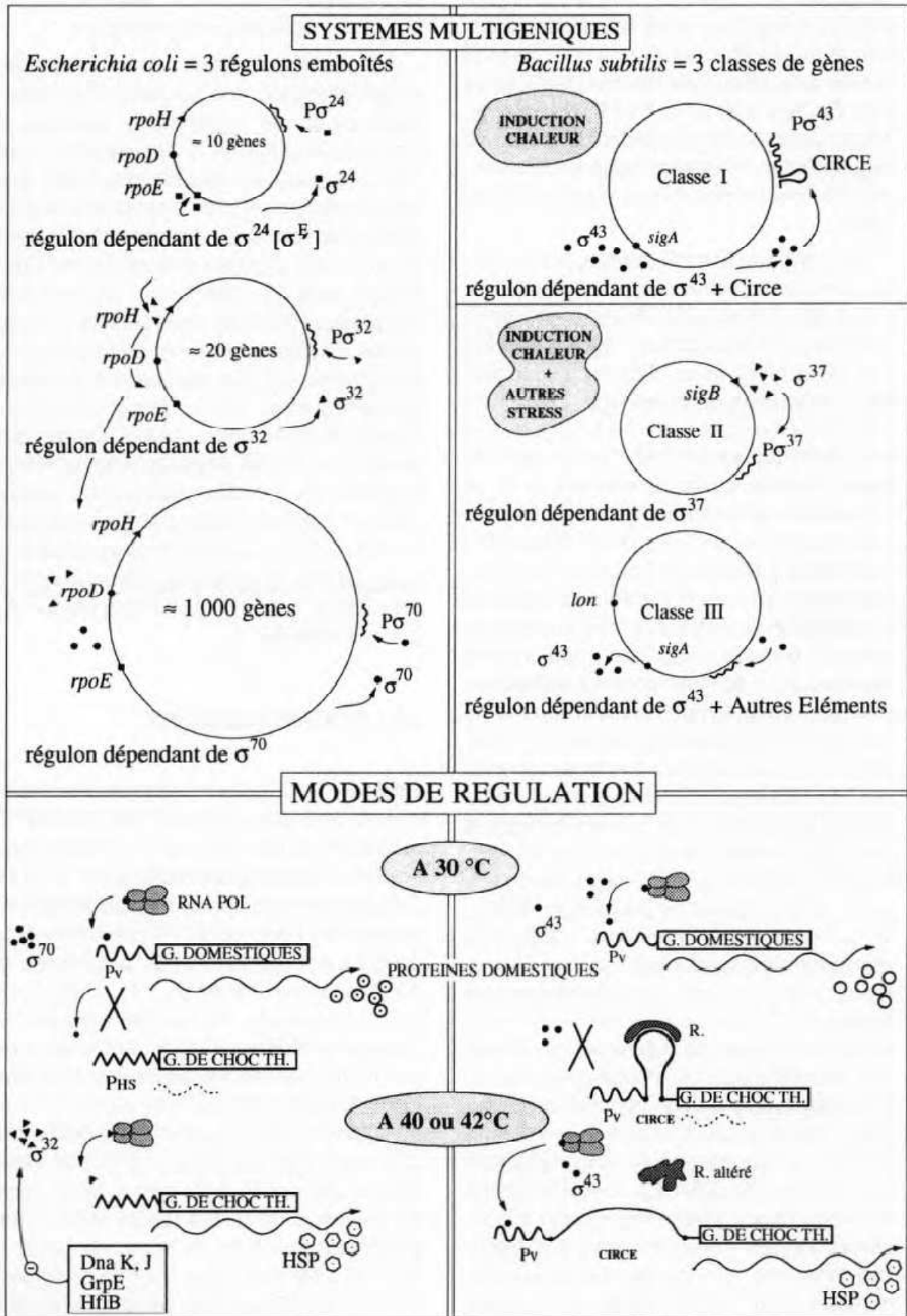
Le séquençage des gènes isolés ou groupés en opérons et de leurs zones promotrices exclusivement végétatives remet en question l'universalité du seul modèle accepté ces dernières années (Neidhardt et al, 1984 ; Neidhardt et VanBogelen, 1987 ; Lindquist et Craig, 1988), celui d' E coli, qui se résume à une commutation transcriptionnelle. Classiquement, la réponse à une stimulation thermique s'accompagne dans les minutes qui suivent par le remplacement du facteur sigma majeur (σ^{70}), dégradé à température non permissive, par le facteur sigma alternatif (σ^{32}) (Calendar et al, 1988) stable et même activé à haute température et dont la synthèse est par ailleurs amplifiée sous l'effet d'un autre facteur (σ^{24} ou σ^E) qui reconnaît le promoteur P3 de *rpoH* et son propre promoteur P2, assurant ainsi son autorégulation (Bukau, 1993 ; Yura et al, 1993). L'inflexion de la réponse et le retour à état de repos est sous la dépendance de 4 hsp, *DnaK*, *GroEL*, *GrpE*, *HflB* ou *FtsH* (Herman et al, 1995). Le facteur σ^{32} conduit la transcription de la vingtaine de gènes qu'il reconnaît grâce à leur motif hs (heat-shock) et qui composent ce régulon conventionnel (fig 2).

La majeure partie des protéines dont les gènes sont inclus dans le régulon dépendant de σ^{32} présentent une affinité pour d'autres protéines cellulaires, cibles de la chaleur, qu'elles assistent, désagrègent, restaurent ou dégradent. Elles participent vraisemblablement à la protection des cellules ou à leur sauvetage lorsqu'elles sont menacées par leur environnement thermique.

Le facteur σ^{24} ou σ^E classé dans le groupe des facteurs ECF (extracytoplasmic function) contrôle l'expression d'une dizaine de gènes (dont *rpoH* et *rpoE*) dont les produits sont impliqués dans la croissance aux températures élevées (42-44 °C) ; dans la survie aux épreuves thermiques létale (50 °C) (Rouvière et al, 1995 ; Raina et al, 1995). Les produits des gènes de ce second régulon interviendraient dans le compartiment extracytoplasmique (Meccas et al, 1993).

Figure 9 : Régulation de l'expression des gènes de choc thermique chez *Escherichia coli* et *Bacillus subtilis*. A. Systèmes multigéniques impliqués dans la réponse au choc thermique chez ces 2 espèces. B. Modes de régulation de l'expression des gènes. (JIO (rpoO) ; (J32 (rpoH) ; (J24 ou (JE (rpoE) ; P : ARNpolymérase ; Pv : promoteur végétatif dépendant de (JIO (*E. coli*) ou de (J43 (*B. subtilis*) ; Phs: promoteur de choc thermique reconnu par (J32 ; hse : séquence IR (inversion répétée) constituant l'élément de choc thermique (CIRCE) de *B. subtilis* ; ORF : cadre ouvert de lecture; T : terminateur potentiel (IR) ; R : protéine «répresseur» hypothétique.

Analyse bibliographique



Analyse bibliographique

Au contraire chez *B subtilis*, les promoteurs hs sont absents; ceux d'E colifusionnés à des gènes rapporteurs sont inactifs (Wetzstein et Schumann, 1990). L'expression de la réponse au choc thermique procède de promoteurs canoniques, reconnus par O⁷⁰ (homologue de O⁷⁰ d'E coli), qui ont été identifiés en amont d'un groupe de gènes thermorégulés de classe 1, et d'une séquence inversée répétée de 9 pb appelée CIRCE (controlling IR of chaperone expression) (Zuber et Schumann, 1994), habituellement intercalée entre le signal de transcription et le début du cadre ouvert de lecture.

La fixation supposée d'une protéine répresseur sur ce motif récurrent limiterait ou interdirait l'expression du gène. Ce même répresseur constituerait le capteur ou «senseur», l'une des cibles sensibles, qui obéirait à la sollicitation thermique par un changement conformationnel, lequel provoquerait son détachement de l'IR suivi par la transcription du gène correspondant (Wetzstein et al, 1992 ; Zuber et Schumann 1994 ; Volker et al, 1994 ; Yura et al, 1993). À la régulation transcriptionnelle positive s'exprimant, chez E coli, par une permutation des facteurs O⁷⁰ aux facteurs O³² et O²⁴ eux-mêmes emboîtés, se substitue chez *B subtilis*, au moins pour un groupe de gènes, un mécanisme à régulation négative en cis faisant intervenir, outre le couple promoteurfacteur sigma végétatif, la reconnaissance d'un élément hs (CIRCE), couplé à l'altération d'un hypothétique répresseur.

Deux modèles nous sont ainsi proposés: l'un complexe, étranger semble-t-il aux éléments aujourd'hui reconnus, par l'analyse des gènes séquencés, chez *L lactis* ; l'autre moins élaboré, qui traduit une riposte plus fruste aux sollicitations thermiques. Ne possédant que des données éparses et des informations lacunaires, nous retiendrons ce dernier modèle pour expliciter le mécanisme global de la réponse chez *L lactis*. La présence des IR ayant été, par ailleurs, reconnue chez une dizaine de bactéries Gram positives et apparentées, appartenant à des groupes très variés, de *Clostridium* sp à *Mycobacterium tuberculosis*, chez lesquelles elles représentent le motif le plus fortement conservé, il apparaît qu'un tel système est plus ancien que le mécanisme régulateur mis en place, chez E coli (Zuber and Schumann 1994). Cependant Duwat et al (1995) (Duwat et al. 1995) proposent une autre voie originale de contrôle impliquant à la fois, par une réaction en cascade, Rec A et HflB. L'intervention de cette dernière protéine sous-entend la médiation par un facteur sigma alternatif du type cr32, qui, jusqu'à ce jour, n'a pas été identifié.

5.8. La thermotolérance

De 30 à 43 °C, le taux de croissance d'une culture de *L lactis* varie peu. De 40 à 48 °C, la viabilité d'une population cellulaire prélevée en phase exponentielle (à 30 °C) n'est pas affectée; elle décroît très rapidement en revanche à partir de 50 °C pour n'être que de 8 % et 0,06 % après 15 minutes et 30 minutes de chauffage à 51°C. À 55 °C et au-delà, les cellules meurent en quelques minutes. Cependant le taux de survie des cellules exposées aux températures élevées dépend des conditions de croissance. Ainsi, il est considérablement plus élevé si les bactéries, avant d'être exposées à une épreuve thermique létale, sont soumises, durant des temps qui correspondent à peu près au temps de génération (30 à 40 min) à des températures de 5 à 10°C supérieures à la température habituelle de croissance. L'appréciation quantitative de ce phénomène d'exhaussement de la survie cellulaire basale, appelé résistance induite ou thermotolérance, observé chez tous les représentants du monde vivant (bactéries, champignons, plantes et animaux) (Lindquist and Craig 1988) est fournie par le facteur de thermotolérance, c'est à- dire le rapport du pourcentage de survie d'une population adaptée ou conditionnée à telle épreuve sur le pourcentage de survie d'une population témoin, «naïve» cultivée à 30°C (Boutibonnes et al. 1991). Ce facteur *covarie avec* la température de préculture : égal à 1 à 30 °C, il est maximal à 42 °C où sa valeur dépasse 100 , puis il décroît très vite quand la température d'adaptation se rapproche de la température d'épreuve (52 °C). Si les cellules «conditionnées» par la chaleur, qui anticipent ou se préparent à une épreuve thermique létale, meurent lentement, le traitement d'adaptation entraîne de surcroît une tolérance hétérologue *vis-à-vis* d'autres modalités qui menacent la survie cellulaire; ces résistances croisées s'expriment *vis-à-vis* d'agents létaux, dont la nature et les fonctions cytotoxiques sont éloignées de celles de la chaleur (tableau V). Ainsi un prétraitement à 42 °C élève considérablement la résistance «constitutive » des bactéries à l'alcool et au pH acide (on parle d'alcoolotolérance ou de pH - tolérance). Il n'a qu'un effet discret sur la résistance à l'eau oxygénée, au sel et aux UV254 nm- Le traitement préalable par les températures modérées a également pour conséquence d'inhiber la lyse cellulaire des bactéries lyse-gènes (ici *L lactis* A 15, abritant le phage A15) provoquée par l'induction expérimentale du prophage (Boutibonnes et al, 1991) : le choc thermique (de 30 à 40 °C) est d'autant plus efficace qu'il est plus précoce, c'est-à-dire plus proche de l'administration de l'agent SOS-inducteur qui déclenche la lyse (mitomycine C ou UV254 nm).

La synthèse d'hsps, mais aussi le développement de la thermotolérance sont conditionnées aux traitements des cellules par diverses formes de stress, lesquels dans la majorité des cas entraînent des dommages protéiques (Hightower 1991). Toutefois, une question reste toujours

Analyse bibliographique

posée: elle concerne la relation hypothétique qui lierait, sans réserve, la néosynthèse d'une ou de quelques hsps à la capacité nouvellement acquise par les cellules d'outre passer l'épreuve létale (Nover and Hightower 1991; Parsell et al. 1994a; Parsell et al. 1994b). De nombreux travaux, portant sur les bactéries mais aussi sur les cellules eucaryotes en culture in vitro, étayent ou infirment l'existence de cette interdépendance; par exemple chez *E coli*, la surexpression des gènes du régulon *sigma 32*, obtenue en plaçant le gène *rpoH* sous la commande d'un promoteur *tac* inductible par l'IPTG (isopropyl thio-galactoside), suivi de la synthèse d'hsps à hauts niveaux endocellulaires, ne conduit pas au développement de thermotolérance (Jones et al. 1987; VanBogelen et al. 1987a; VanBogelen et al. 1987b). Chez la levure, la thermotolérance peut se développer en l'absence de l'hsp majeure ou hsp 104 identifiée par Sanchez et Lindquist (1990), et de toute néosynthèse protéique (Sanchez and Lindquist 1990) et de toute néosynthèse protéique (De Virgilio et al. 1991; Kirk and Piper 1991).

Ainsi chez *L lactis* des thermomimétiques qui induisent la synthèse des hsps, impliquées dans le surcroît de résistance à la chaleur (DnaK, GroEL, GrpE, hsp 104 ou ClpB) n'affectent pas la résistance thermique naturelle des cellules: c'est le cas de la plupart des antibiotiques. Inversement, l'acquisition de l'aptitude nouvelle à tolérer l'épreuve thermique létale, acquise après traitement par le chlorure mercurique et l'azide de sodium ou consécutive à l'entrée des cellules en phase stationnaire est découplée de la capacité de surexprimer les hsps. Soulignons enfin, que chez *L lactis*, à l'inverse de ce qui se passe chez *E coli*, des agents réputés pour déclencher chez des cellules eucaryotes en culture la synthèse des hsps et/ou la thermotolérance, sont sans effet.

Bien que leur présence soit nécessaire, la responsabilité des hsps dans la thermotolérance est sans cesse remise en question par des résultats expérimentaux contradictoires. Il y a vraisemblablement trois explications à cette incertitude: 1) les hsps naturellement présentes, quelquefois même abondantes, ne nécessiteraient qu'une «activation» par les traitements de conditionnement pour devenir fonctionnelles et intervenir dans les processus de sauvetage thermique. 2) Les hsps «classiques» ne seraient peut-être pas les seuls agents prévenant ou restaurant les dégâts protéiques, impliqués dans les tolérances (Parsell et Lindquist, 1994). D'autres facteurs de protection pourraient être associés à l'induction de systèmes multigéniques discrets (tel le régulon 024 d'*E coli* dont nous méconnaissons la fonction exacte). 3) L'hyperproduction ou l'activation des hsps, par mise en conformité des chaînes peptidiques, augmente la capacité de protéolyse endocellulaire.

Analyse bibliographique

Certaines hsps qui sont aussi des protéases, exercent leur activité sur d'autres hsps les faisant «disparaître» après qu'elles ont assuré leur office de restauration: leur trace même ne serait pas repérable par analyse électrophorétique.

5.9. Conclusion

Si depuis quelque temps l'on s'achemine vers la connaissance plus précise des termes (inducteurs ou stressseurs, cibles ou senseurs, protéines, gènes) qui, chez *L. lactis*, constituent la phénoménologie de la réponse au choc thermique, il reste encore beaucoup à faire pour comprendre le déclenchement et la cascade des événements qui aboutissent à ce phénomène «chaotique», obéissant à d'imperceptibles fluctuations des conditions initiales (ici une augmentation de quelques °C). Demeurent notamment ces trois questions: Quel est le mode de régulation de l'expression des gènes de choc thermique? Thermotolérance et hsps sont-elles en étroite connexion? Si c'est le cas, comment se réalisent, surveillance, prévention, mais aussi désagrégation et restauration de l'arsenal protéique immature ou endommagé par la machinerie de chaperon nage ?

La problématique des stress, c'est-à dire la prise en compte de l'environnement et de ses paramètres, qui forment avec l'hôte des niches un couple insécable, s'est doublement élargie. À l'instar du choc thermique, les fluctuations de pH (acide ou alcalin), d'osmolarité, d'oxygénation, l'abaissement de la température (4 °C) (Panoff et al, 1994), la pénurie en nutriment (par exemple l'épuisement de la source de carbone : Hartke et al, 1994) déclenchent chez les bactéries lactiques mésophiles –et comme on s'y attendait-, une réponse duelle, couplant la synthèse de nouvelles protéines de stress à des états de tolérance, ou de multitolérance, quelquefois élevée.

La réponse au choc thermique, qui semble constituer une réserve de mémoire déposée dans le génome sous forme de traces mnésiques, de projets prêts à être réactualisés en certaines occasions, est maintenant reconnue chez d'autres bactéries lactiques *senso lato* atypiques: *Enterococcus faecalis*, dont l'un des traits singuliers est une résistance peu ordinaire à de nombreuses conditions «extrêmes» (Boutibonnes et al, 1993) ; *Streptococcus thermophilus* dont la synthèse d'hsps, stimulée à haute température (52 OC), est accompagnée d'une viabilité importante à 58 °C (Auffray et al, 1995).

Le phénomène s'impose comme le modèle d'étude de la régulation de l'expression de systèmes multigéniques conduisant à une authentique différenciation métabolique : la cellule stressée n'est plus tout à fait la même que la cellule «naïve». Les forces sélectives, mais aussi les

Analyse bibliographique

contraintes endogènes (génétiques ou «mécaniques») ont agi pour maintenir au sein de la cellule (ou des organismes) ce que les processus évolutifs avaient érigé afin que cette cellule (ou ces organismes) échappent aux conditions adverses qui menaçaient à la fois l'intégrité de l'une (ou des autres) et la descendance de l'espèce. Il n'y a pas, dans ce type de réponse, de commune mesure entre l'ampleur des conséquences due à une véritable commotion métabolique (le plus souvent bénéfique) et l'insignifiance des causes (quelques degrés thermiques). Ce qui paraît être une riposte désordonnée et disproportionnée, n'est pas, semble-t-il, un défaut ou une crise de l'ordre biologique habituel; elle n'est qu'une des conditions qui a autorisé l'existence et la permanence de cet ordre.

C'est parce que, par essence, le désordre est versatile, qu'il se renouvelle sans cesse pour se convertir en un autre désordre jusqu'à paraître incernable, qu'une relation intelligible entre une cause et sa conséquence peut s'installer et légiférer de manière immuable la fonction des systèmes biologiques.

Rapport-Gratuit.com

CHAPITRE 2

MATERIELS ET METHODES

Chapitre 2 : Matériels et Méthodes

1. Matériel biologique :

Lactocoques lyophilisée, conditionnée dans des ampoules en verre.

- **60 souches à étudier:** lactocoques isolées à partir du lait cru de chamelle d'Algérie (Drici et al. 2010): (*Lactococcus lactis*), classés dans des séries comme suit:

Tableau 5 : Souches utilisées dans ce travail

série	1	2	3	4	5	6	7
souches	10 A ₂	2 B	14 B	5 C	18 C	9 D ₁	18 D
	10 A ₃	3 B	15 B	6 C	19 C	10 D	19 D
	12 A	4 B	16 B	7 C	20 C	11 D	20 D
	14 A	6 B	17 B	8 C	1 D	12 D	/
	15 A	7 B	18 B	10 C	2 D	13 D	/
	16 A	8 B	19 B	11 C	3 D	14 D	/
	17 A	9 B	1 C	12 C	5 D	15 D	/
	18 A	10 B	2 C	13 C	6 D	16 D	/
	20 A	11 B	3 C	17 C	8 D	17 D	/
	1 B	12 B	4 C	/	/	/	/

- **Souches de références:**

Lactococcus lactis DSM 20481 et *Leuconostoc mesenteroides* DSM 20343 (achetées de DSMZ Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Allemagne), *Enterococcus faecium* ATCC6057 proviennent de Institut für Hygiene und Mikrobiologie der Universität Würzburg). La souche témoin (3 MD) Lactocoque isolée à partir du lait de chèvre de la région d'Oran; qui fait partie de la collection du Laboratoire de Microbiologie Appliqué de l'université d'Oran « LMA » témoin positif qui se développe à des températures élevées jusqu'à 50°C.

2. Test de pureté :

2.1. Culture sur milieu M17 liquide :

En utilisant une anse de platine comme moyen d'ensemencement, on a prélevé un peu du lyophilisat bactérien qu'on a dissout dans le milieu réparti dans des tubes à vis, ainsi les souches lyophilisées préalablement peuvent être repiquées par cet ensemencement. L'incubation est faite à 30°C pendant 16 à 18 h.

On remarque une croissance bactérienne qui se traduit par un trouble plus ou moins intense avec un dépôt dans certains cas. On remarque ainsi l'absence totale du trouble auprès de la surface en formant un halo clair indiquant la microaérophilie des bactéries cultivées.

2. 2. Test de pureté sur milieu M17 solide et microscopie :

À partir des tubes positifs qui ont présenté un trouble, on a ensemencé sur milieu solide par méthode de stries en suivant la méthode de quatre quadrants. On a incubé à 30°C pendant 16 à 18 h. La pureté des souches est vérifiée par épuisement sur milieu M17 jusqu'à l'obtention de colonies isolées. Il est conseillé d'observer les colonies à l'aide d'une loupe binoculaire afin de vérifier leur pureté après épuisement.

La pureté des cultures est vérifiée par observation au microscope optique des préparations de cultures entre lame et lamelle à l'état frais ce qui permet de vérifier l'homogénéité des cellules, les différentes formes présentes, et leur mobilité.

2.2.1. Coloration de Gram :

Un contrôle supplémentaire par la coloration de Gram permet de différencier sur la base de la composition de la paroi cellulaire entre les bactéries qui sont à Gram négatif et les autres bactéries à Gram positif. La coloration de Gram doit son nom au médecin danois Hans Christian Gram qui mirent au point le protocole en 1884 (Microsoft ® Encarta ® 2006. © 1993-2005 Microsoft Corporation. Tous droits réservés).

Les étapes du protocole sont les suivantes : Les cellules sont fixées à la flamme. Ensuite le violet de Gentiane est ajouté (1-2 min). Les bactéries se colorent alors en violet. Le colorant est par la suite fixé à l'aide du lugol (iodure de potassium KI) (2 x 30 sec). La préparation est abondamment rincée à l'eau distillée afin d'évacuer l'excès de colorant. La préparation est décolorée (éthanol 95 %, 5 sec) et rincée de nouveau abondamment à l'eau distillée. Puis, une deuxième coloration avec de la fuschine ou safranine est réalisée (1-2min).

De nouveau, la préparation est rincée à l'eau et séchée à température ambiante. L'observation microscopique au grossissement X1000 avec huile à immersion permet de voir si les cellules sont colorées en violet (Gram-positif) ou bien en rose (Gram-négatif).



2. 3. Conservation sur glycérol 30% stérile :

À partir des tubes positifs, on a prélevé grâce à une micropipette 500 µl de culture bactérienne qu'on a mis dans un tube Eppendorf stérile contenant 500 µl du glycérol (30%) stérile étant donné que le glycérol est un cryoprotecteur. Les tubes sont scellés avec du parafilm et on a conservé à (-20°C). Les souches congelées peuvent être conservées une longue durée (-20°C de 1 année à 3 ans).

3. Test de thermotolérance:

À partir des tubes conservés à (-20°C), on a réalisé une culture sur milieu solide ou liquide de la même manière décrite précédemment. Et on a incubé à 5 températures différentes : 37°C, 45°C, 50°C, 60°C, 65°C.

Remarque :

On a lancé les 60 souches, dans les 5 températures, et sur milieu solide ou liquide à la fois, en utilisant le même volume d'inoculum (75µl) dans chaque tube. Le test a été effectué 3 fois successivement pour confirmer le résultat obtenu.

Ce test a été effectué à différentes températures : 45°, 50°, 60°, 65°C afin de déterminer la capacité de 60 lactocoques issus de lait cru de chamelle d'Algérie (Drici et al. 2010) en présence de *Lc. Lactis* subsp. *Lactis biovar diacetylactis* 3MD, pris comme témoin thermotolérant ce dernier a été isolé à partir de lait cru de chèvre de la région d'Oran et faisant partie de la collection du Laboratoire de microbiologie appliquée - Université d'Oran.

La souche témoin ainsi que les souches témoin ainsi que les souches destinées au test de thermotolérance sont également cultivées à 37°C qui représente la température contrôle de croissance.

4. Analyse physiologique:

On a préparé un milieu lait - agar « lait gélosé » à une concentration de 5 % en utilisant un lait stérile. On a appliqué la méthode des patches montrée par la figure pour faire une culture de 48h d'incubation à 30°C.

Remarque :

Le milieu gélose au lait à 5 % était très dense ; il a rendu difficile la visibilité des halos clairs formés par dégradation des caséines. On a opté pour refaire les tests avec un milieu lait –agar « lait gélosé » à une concentration de 3%.

4.1. Test de Protéolyse en milieu solide Agar-lait :

Les 22 lactocoques thermotolérants criblés, ont déjà montré leur capacité à utiliser les nutriments du lait essentiellement les caséines et le lactose. En effet, les phénotypes Prt+, Lac+ des lactocoques thermotolérants se sont manifestés à 30 °C, nous voulions savoir par le test de protéolyse si les phénotypes Prt+ est exprimé à 50°C.

4.2. Conservation dans le lait:

On a réparti le lait stérile dans des tubes Eppendorf à 1ml dans chacun. On a inoculé nos souches dans ces tubes, on a incubé à 30°C pendant 24h et on a conservé à (-20°C) après avoir un résultat positif exprimé par une coagulation.

4.3. Screening des lactocoques thermotolérants :

Après avoir fait une préculture (culture de nuit) des 22 souches thermotolérantes sélectionnées préalablement sur milieu M17 liquide, on a mesuré les DO des cultures mères le lendemain. On a prélevé un volume défini calculé selon la concentration de la culture mère de chaque préculture, on aensemencé dans un tube de 8 ml de M17 liquide et on a complété avec le même milieu pour un volume final de 10 ml afin d'avoir une même densité optique initiale dont $DO_i=0.05$ dans les cultures de T_0 et en appliquant la formule $DO_1.V_1 = DO_2.V_2$

On a incubé à 50°C et on a mesuré les DO à $T_0=0$ heures, $T_1=4$ heures, $T_2= 6$ heures. On a effectué trois tests successifs pour confirmer le résultat obtenu. (Reproductibilité).

4.4. Etude de la cinétique de croissance en milieu M17 (à 30°C et 50°C):

Pour l'étude cinétique le choix a été orienté vers les souches suivantes : 9B, 13C, « voir résultats ». La souche « 3MD » est prise comme souche témoin.

De la même manière décrite auparavant, on a établi la cinétique de croissance des 3 souches pendant 8 heures en mesurant la DO à 600 nm chaque 30 minutes.

4.5. Etude de la cinétique de croissance en milieu lait (à 30°C et 50°C) :

La cinétique de croissance est étudiée de la même manière qu'en milieu M17 , sauf qu'en milieu lait l'étape de transparisation est impérative pour pouvoir relever et lire les DO , cette

Matériels et Méthodes

transparisation est faite par l'addition de 3,5ml d'EDTA (20g/l, pH= 12) à 1ml de culture, et la lecture est faite par rapport à un témoin composé de 3,5ml d'EDTA (20g/l, pH=12) ajouté à 1ml du milieu lait stérile (échantillon à blanc).

4.6. Croissance sur différents milieux de cultures :

Pour déterminer le type fermentaire, l'activité protéolytique, la capacité de fermenter le lactose et l'utilisation du citrate, les souches à testé sont cultivées sur différents milieux de cultures.

Les souches (gl1, gl2, gl3 et les trois souches de référence *Lactococcus lactis* DSM20431, *Leuconostoc mesenteroides* DSM20343, *Enterococcus faecium* ATCC6057,) sont cultivées sur différents milieux HHD « homofermentative-heterofermentative differential medium » (McDonald et al. 1987) pour voir le type fermentaire), sur milieu KMK pour utilisation de citrate (Kempler and McKay 1980), et sur M17 X-Gal (pour fermenter lactose) : M17 + 0,25% lactose+2 ml/l de X-Gal (Sigma Chem) (5-Bromo-4chloro-3-indolyl- β -D-galactopyranoside). Le X-Gal est préparé dans une solution stock (20mg/ml) dans (diméthylsulfoxyde DMSO.). Une solution de l'antibiotique vancomycin est préparée dans de l'eau distillée (100 μ g/ml) directement avant son utilisation.

4.7. Traitement par l'ultrasonification :

Les cultures, tenues dans la glace sont soumises rapidement au traitement par l'ultrasonification dans l'appareil Branson Sonifier 250 (Branson Ultrasonics Corporation, Danbury, CT, USA). Deux 2 ml de chaque pré-culture est utilisé. Il faut régler d'abord l'amplitude sur 40% avec une microsonde (200 watt maximum) ensuite immerger la microsonde à moitié dans l'échantillon 3 fois pendant 15 secondes. Des dilutions sont ensuite préparées jusqu'à 10^{-6} et mises en culture sur différents milieux de culture cité précédemment.

5. Analyse génétique :

5.1. Obtention des variants:

Dans le but de chercher le support génétique de la thermotolérance chez les lactococques 9B et 13 C ainsi que 3MD, nous avons adopté comme stratégie le curage de ces lactococques des plasmides qui peuvent être hébergées par ces lactococques selon la technique de Fantuzzi *et al.*, 1991 (Gonzalez et al. 1990).

Matériels et Méthodes

On a effectué des repiquages successifs, et on a suivi l'évolution de la DO pour 11 générations. Après une série de dilutions décimales à partir de la G11 jusqu'à 10^{-6} , on a effectué un étalement sur M17 glucose solide, on a incubé à 30°C pendant 18h, on a divisé chaque boîte en 4 quadrants et on a choisi au hasard 3 colonies de chaque quadrant. Selon la méthode des Patches on a fait une culture sur boîtes contenant M17 glucose et on a incubé à 30 °C pendant 18h.

On a inoculé à partir des patches en milieu M17 glucose liquide, et on a incubé à 30°C pendant 18h. A partir de cette culture on a conservé les variants en milieu lait selon la méthode décrite auparavant « sous-titre3 ».

Après plusieurs repiquages successifs « sur une période de 11 jours », nous avons choisis aléatoirement des colonies dont les cellules peuvent correspondre aux variants recherchés et qui ont perdu la capacité de se développer à 50°C.

5.2. Analyse des variants :

5.2.1. Test de thermotolérance des variants:

On a préparé deux boîtes de M17-glucose solide qu'on aensemencé par patches puis incubé à 30 et 50°C. On remarque que les variants ont perdu leur capacité à se développer à 50°C, alors que leur développement était normal à 30°C. Ce qui suppose que le caractère de la thermotolérance n'est pas stable car il est perdu après 55 générations. Ce résultat nous permet d'émettre l'hypothèse que le caractère de thermotolérance est peut être porté soit par un ou plusieurs plasmides ou encore localisé sur un transposon conjuguatif, sachant que le déterminant génétique des protéines du choc thermique chez les lactocoques est chromosomique (Bolotin et al. 2001).

Ces résultats nous orientent vers l'exploration des protéines du choc thermique HSP chez ces lactocoques pour les comparer à celle déjà décrites et caractérisées. Ce résultat nous incite aussi à déterminer la structure et le mode de régulation des protéines qui s'expriment à 50°C. Ces nouvelles molécules semblent jouer un rôle crucial dans le maintien de l'intégrité des lactocoques thermotolérants.

5.2.2. Test d'activité protéolytique et d'hydrolyse du lactose:

On a préparé deux boîtes du milieu GMAB « glucose milk agar bromocrésol » selon la méthode des patches on les aensemencées et on a incubé à 30 et 50°C. On a préparé des

Matériels et Méthodes

boîtes identiques que celles qui précèdent, mais en utilisant M17-lactose et Lac-MAB. Le virage de l'indicateur du pH « pourpre de bromocrésol » contenu dans le milieu Lac-MAB du pourpre au jaune autour des patchs (figure 8), indique la présence de l'acide lactique obtenu par fermentation du lactose présent dans le milieu de culture (Lac-MAB). Ce résultat est en faveur d'un phénotype Lac⁺ pour les variants testés. L'utilisation du lactose par les variants est confirmée par comparaison à un témoin GMAB.

5.3. Curage d'ADN plasmidique et transformation par électroporation :

5.3.1. Curage :

Dans le but d'éliminer certains plasmides de la souche 9B, cette dernière a été repiquée successivement pendant huit jours en milieu M17-Glu liquide (Gonzalez et al. 1990). 10 mL de ce milieu M17 sont inoculés à chaque repiquage avec 100 µL de la culture bactérienne précédente. Au 13^{ème} repiquage ce qui correspond à la 65^{ème} génération, 50 ou 100 µL de la suspension bactérienne sont étalées sur milieu solide M17-Glu. Après incubation pendant 18 h à 30°C, les colonies bactériennes obtenues sont prélevées et patchées sur un milieu M17-Glu gélosé.

5.3.2. Transformation bactérienne par électroporation (Gerber and Solioz 2007)

5.3.2.1. Extraction de l'ADN plasmidique :

Des cultures de nuit de 50 mL en milieu M17 incubées à 30°C ont été réalisées pour les souches de lactocoques à tester. Les cellules sont récoltées par centrifugation à 10000 rpm pendant 10 min. Le culot bactérien est ensuite repris par 1mL d'une solution I contenant du sucrose (20%), du lysozyme (10 mg/mL), 200 µL de mutanolysine à (1000 U/mL) et 20 µL d'une solution d'ARNase (5 mg/mL). Le mélange est incubé à 37°C pendant 15 min. Cette étape correspond à la lyse de la paroi bactérienne. On ajoute à ces préparations 2 mL de la solution II qui permet la lyse cellulaire et la dénaturation de tous les ADN. Cette étape est suivie d'une incubation dans la glace pendant 10 min, puis on ajoute 1,5 mL d'une solution III qui a pour effet la neutralisation du milieu réactionnel. Cette neutralisation rapide conduit à la renaturation de l'ADN plasmidique alors que l'ADN chromosomique reste dénaturé, donc non soluble. Une centrifugation à 10000 rpm pendant 10 min permet de culoter l'ADN chromosomique dénaturé et le surnageant est récolté après filtration sur de la gaze stérile. Ensuite une extraction avec 4 mL de phénol/chloroforme/ isomalyalcool (BIOSOLVE, Valkenswaard, Netherlands) (24/24/1 ; V/V/V) est réalisée afin d'éliminer les protéines. La

Matériels et Méthodes

phase supérieure aqueuse est prélevée et l'ADN plasmidique présent est précipité par l'addition de 2,5 volumes d'éthanol absolu à -20°C pendant 40 min (ou à -80°C pendant 15 min). L'ADN plasmidique est culoté par centrifugation à 10000 rpm durant 15 min. Ce culot est à nouveau précipité par l'addition d'un mélange de 320 µL d'eau Aguetant, 80 µL de NaCl 4M et 400 µL de PEG 8000 13% et une incubation de 30 min dans la glace. L'ADN plasmidique est récolté par centrifugation à 10000 rpm durant 15 min. Un lavage du culot d'ADN est réalisé avec un petit volume (200-300 µL) d'éthanol à 70°. Après une centrifugation de 15 min à 10000 rpm, le culot d'ADN plasmidique est finalement remis en suspension dans 150 µl de Tampon TE (Tris-HCl 10 mM pH 8 ; EDTA 1mM).

5.3.2.2. Dosage et contrôle de pureté de l'ADN :

La pureté et la quantité de l'ADN plasmidique est mesurée à l'aide du Nanodrop™ (Peqlab, VWR international GmbH, Erlangen, Allemagne).

5.3.2.3. Préparation des cellules compétentes :

Une pré-culture de nuit de 5mL en milieu SGM17 : M17 media (Terzaghi and Sandine 1975) additionné à 1 mM MgSO₄, 0.5 M sucrose, 2% glycine, 1.5% glucose de la souche réceptrice (un lactococque mésophile) a servi pour inoculer 50 mL de milieu SGM17 avec une concentration cellulaire équivalente à une DO de 0,1 unité. Le milieu est incubé à 30°C pendant 18h. Inoculer 450 ml du milieu SGM17 avec 50 ml (culture de nuit). Le milieu est incubé à 30°C jusqu'à DO 0,4 (fin de la phase exponentielle). La culture est alors centrifugée 10 min à 8000 × g et à une température de 4°C. Resuspendre le culot bactérien dans 400 ml du tampon d'électroporation sans EDTA (0.5 M sucrose, 10% glycerol, pH 7.0), et centrifuger à 8000 × g à 4°C. Resuspendre le culot bactérien dans 200 ml du tampon d'électroporation avec EDTA (0.5 M sucrose, 10% glycerol, 50 mM Na-EDTA, pH 7.0), incubé à 4°C pendant 15 min puis centrifuger à 8000 × g à 4°C. Laver les cellules dans 400 ml avec du tampon d'électroporation sans EDTA (0.5 M sucrose, 10% glycerol, pH 7.0). Resuspendre le culot final dans 4 ml du tampon d'électroporation à 4°C. Conserver des aliquotes de 120 à 500 µl dans de l'azote liquide et stocker à -80 °C.

5.3.3. Conditions de transformation :

La transformation des bactéries compétentes est réalisée par électroporation. L'appareil utilisé est le « Bio-rad Gene pulser ». 120 µL de bactéries compétentes maintenues dans la glace sont mis en contact avec 2 µg d'ADN plasmidique. Le mélange bactéries / ADN plasmidique est

Matériels et Méthodes

placé au fond de la cuve d'électroporation (0,1 cm). Le courant électrique généré par l'appareil perméabilise la membrane cellulaire (formation de pores) et l'ADN peut alors pénétrer dans la cellule. Les valeurs des paramètres de l'électroporation utilisées sont : 200 Ω (50 W), 25 μ F, 1.3 kV. Un témoin négatif sans ADN plasmidique est traité de la même manière, afin de Contrôler la pureté des souches compétentes utilisées.

La suspension de bactéries électroporées est immédiatement transférée dans un microtube stérile qui contient 0,90 mL de milieu SM17MC : M17 additionné de 0.5 M sucrose, 10 mM $MgCl_2$ et 2 mM $CaCl_2$. L'ensemble est incubé 5 min sur la glace, puis 90 min à 30 °C. 50 à 100 μ L de cette suspension sont étalés sur un milieu gélosé de sélection (M17 additionné à 5 μ g/ml d'Erythromycine) et incubé à 30°C pendant 16h.

5.4. Extraction de l'ADN génomique :

Les cellules bactériennes sont les plus couramment étudiées et utilisées parmi les microorganismes. Tout d'abord, une courbe de croissance est obtenue pour un type cellulaire donné en mesurant la densité optique (OD) à 600 nm, à différents temps durant la culture. Pour l'extraction d'ADN, les cellules bactériennes sont cultivées en milieu liquide jusqu'à ce qu'elles atteignent une densité maximum de $2-3 \times 10^9$ cellules/ml, elles sont ensuite récoltées. La croissance des cellules peut être surveillée par l'analyse de la turbidité de la culture, et quantifiée en mesurant la densité optique à 600 nm et en la corrélant à la courbe de croissance établie précédemment. Les cellules collectées peuvent être ensuite traitées pour l'extraction de l'ADN en utilisant différents kits. La lyse cellulaire pour l'extraction d'ADN bactérien est souvent effectuée par des méthodes chimiques utilisant le lysozyme, l'EDTA, le lysozyme et l'EDTA, des détergents etc. Comme décrit dans l'introduction, la lyse cellulaire est suivie par l'élimination des composants cellulaires de l'ADN à l'aide d'une méthode d'extraction organique, ou d'une technologie utilisant la silice etc. La dernière étape implique la précipitation de l'ADN pour obtenir une solution d'ADN pur de concentration élevée.

L'ADN utilisé est obtenu selon plusieurs méthodes d'extraction:

5.4.1. Préparation de l'ADN génomique « à la protéinase K » :

Une colonie pure, fraîchement isolée sur milieu gélosé M17, est resuspendue dans 50 μ l du tampon 1X PCR, contenant 1 μ l de lysosyme (120 mg/ml) (Lysozyme from chicken egg, Sigma-Aldrich CHEMIE GmbH., Steinheim, Germany) et incubée à 37°C pendant 15 minutes. La protéinase K (Roth, Karlsruhe, Germany) est ajoutée (12 mg/ml), la préparation est ensuite incubée à 56°C pendant 15 minutes. Le mélange est mis à incuber 5 minutes à 110°C, pour

dénaturer la protéinase K. Les tubes sont stockés à -20°C . 2 μl de cellules traitées à la protéinase k sont utilisées pour la réaction de la PCR.

5.4. 2 Extraction de l'ADN génomique par KIT

Beaucoup de kits sont disponibles pour l'extraction d'ADN à partir de microbes comme les bactéries, levures, et champignons. Bien que le principe derrière les différentes étapes soit le même, les kits présentent beaucoup de modifications dans le tampon de lyse, la méthode de séparation de l'ADN, la membrane utilisée, le tampon de lavage et d'éluion de ADN. Le but de ces modifications est d'améliorer l'efficacité de l'extraction d'ADN du kit.

L'extraction et la purification d'ADN génomique a été réalisée par le Kit DNeasy Blood & Tissue Kit (QIAGEN GmbH, Hilden, Allemagne) après culture de la souche à identifier dans 5 ml du milieu M17 liquide additionné à 1% glucose pendant 16 heures à 30°C . La détermination de la pureté et la concentration d'ADN a été obtenue par le dosage au spectrophotomètre NanodropTM (Peqlab, VWR international GmbH, Erlangen, Allemagne). Le Nanodrop est un petit spectrophotomètre UV/Visible qui permet la mesure, sur 1 μl d'échantillon, de la concentration en acides nucléiques (ARN/ADN, linéaire entre 5 ng/ μl et 3 $\mu\text{g}/\mu\text{l}$), en sonde Cy3 ou Cy5, en protéines et la mesure de densités cellulaires. La pureté de l'ADN a été évaluée par le rapport des densités optiques DO 260/280 nm. Un extrait d'ADN de bonne qualité donne un rapport compris entre 1, 8 et 2.

5.5. Amplification par PCR:

La PCR (Mullis and Faloona 1987), de l'anglais *Polymerase Chain Reaction* ou en Français : Réaction de Polymérisation en Chaîne, permet l'amplification sélective in vitro de séquences d'acides nucléiques.

- Nous avons amplifié plusieurs gènes au cours de cette thèse : l'ADNr16S, *gdpP* et *rpoB* pour l'analyse phylogénétique, et le génome total pour étudier l'organisation génétique des gènes *heat shock*.
- La composition du tampon pour la PCR diffère selon les amorces utilisées et le gène à amplifier, le programme d'amplification est spécifique à chaque gène.
- Pour tous les gènes l'ADN matrice est purifié selon la méthode à la protéinase K ou à l'aide de kit.
- Les programmes d'amplification spécifiques à chaque gène sont regroupés dans le tableau (6).
- Les amorces utilisées dans l'amplification et le séquençage des différents gènes étudiés sont regroupés dans le tableau (7).

Matériels et Méthodes

L'amplification est réalisée sur 2 µL d'ADN obtenu par traitement à la protéinase K. Le volume du mélange réactionnel est de 50 µL et contient : 2 µL d'ADN, 5 µL de tampon de réaction 10X (10XPCR buffer (1,5 mM Mg²⁺ final concentration), MolZym, Bremen, Germany), 1µL de dNTPs chacun à 10 mM (GENAXXON, Bioscience), 2µL de chacune des deux amorces à 10 mM (Thermoscientific), 1 unité de MolTaq DNA polymérase (Taq DNA polymerase (5U/ µL), MolZym, Bremen, Germany) et 37µL d'eau milliQ stérile. Un contrôle négatif contenant 2 µL d'eau milliQ stérile au lieu de l'ADN est inclus dans chaque réaction PCR. L'amplification est réalisée dans un thermocycleur (BiometraGmbH, Göttinger, Germany) avec le programme qui correspond (Tableau 8).

Les produits d'amplification sont visualisés par électrophorèse (A. Hartenstein) en prenant 1µL d'ADN amplifié de chaque souche avec 5µl de tampon de charge (Thermoscientific) sur un gel d'agarose 1 % (Agarose NEEO Ultra-quality, Roth, Karlsruhe, Allemagne) préparé dans du tampon 1X TBE (Dark Hood DH-40/50, Biostep, Burkhardtsdorf, Allemagne) et mélangé avec 10 µl de HD green (INTAS Science Imaging instruments GmbH, Götting, Allemagne). La migration se fait dans du tampon TBE 1X (40 mM Tris-acétate, 1 mM EDTA, pH 8,3). Le gel est photographié (Dark Hood DH-40/50, Biostep). La taille du produit d'amplification attendu varie suivant le gène amplifié.

La totalité de l'amplifié est visualisé et découpé après migration sur gel d'agarose à 1 %. La bande correspondant à la taille du gène amplifié est découpée et purifiée à l'aide d'un kit (DNeasy Blood & Tissue Kit, QIAGEN GmbH, Hilden, Germany) à température ambiante, visuellement en le comparant aux bandes du Smart Ladder (Frgeneruler, Thermoscientific). Le Smart Ladder est un marqueur de poids moléculaire qui sert d'étalon puisqu'il est composé de fragments d'ADN double-brin linéaires de tailles connues, régulièrement réparties respectivement entre 200 et 10000 pb. De plus il permet après migration, de quantifier approximativement la quantité d'ADN déposé en fonction de l'intensité des bandes puisque le mélange est composé d'une quantité connue de chaque fragment. L'ADN élué est ensuite évalué pour sa pureté et sa concentration à l'aide du NanodropTM (Peqlab, VWR international GmbH, Erlangen, Allemagne).

5.6. Amplification des gènes de choc thermique par RT-PCR :

Pour effectuer la PCR utilisant l'ARN comme matrice, l'ARN doit d'abord être inversement transcrit en cDNA dans une réaction de transcription inverse (RT). RT et PCR peuvent être réalisées soit séquentiellement dans le même tube (en une seule étape RT-PCR) ou séparément (en deux étapes RT-PCR).

Matériels et Méthodes

Cinq gènes sont sélectionnés pour être analysés par une amplification par RT-PCR : trois genes *groEL* et *dnaK groES* codent pour les protéines GroEL, Dnak et GroES, sont des composantes de la machinerie de chaperonnage : chargées d'assurer le repliement régulier des oligopeptides, elles participent par un mécanisme ATP-dépendant à la désagrégation et à la restauration des protéines fonctionnelles endommagées par la chaleur. Les deux autres gènes *clpX* et *clpP* codant les protéines Clp X et ClpP, Chaperon moléculaire et protéase endocellulaire.

5.6.1. Isolement de l'ARN :

L'ARN est isolé par RNeasy Mini kit (Quiagen) selon le protocole suivant :

Les cellules doivent être récoltées en phase logarithmique de croissance et sont centrifugées 1000xg pendant 5 minutes à 4°C. Le culot bactérien est resuspendu dans 350 µl du tampon RLT contenant du lysozyme (120 mg/ml) et incubé pendant 30 minutes à 37°C. Ajouter 250 µl d'éthanol (96-100 %) au lysat et mélanger bien en pipetant également. Transférer l'échantillon à une colonne RNeasy placée dans un tube de prélèvement de 2 ml et centrifuger pendant 1 minute à $\geq 8000 \times g$ ($\geq 10,000$ tr/min). Ajouter 350 µl de tampon RW1 la colonne RNeasy et centrifuger pendant 1 minute à $\geq 8000 \times g$ ($\geq 10,000$ rpm). Pour éviter toute contamination par l'ADN génomique, 10 µl du DNase est ajoutée en présence de 70 µl du tampon RDD et incubé pendant 30 minutes à température ambiante. DNase est particulièrement sensible à la dénaturation physique. L'agitation devrait être effectuée en retournant doucement le tube. La DNase est éliminée par un second lavage avec 350 µl du tampon RW1 puis centrifuger pendant 1 minute à $\geq 8000 \times g$ ($\geq 10,000$ rpm). Ajouter 500 µl du tampon RPE, centrifuger ensuite 1 minute à $\geq 8000 \times g$. Ajouter encore une fois 500 µl du tampon RPE, centrifuger 2 minutes à $\geq 8000 \times g$. Transférer la colonne dans un nouveau tube Eppendorf et la centrifuger vide pendant 1 minute. Placer la colonne RNeasy dans un nouveau tube Eppendorf de 1,5 ml. Ajouter 20-30 µl d'eau exempte de RNase (RNase free water) directement sur la membrane de la colonne. Laisser 2 minutes à température ambiante. Centrifuger 1 minute à $\geq 8000 \times g$ ($\geq 10,000$ rpm) pour éluer l'ARN. La quantité et la pureté de l'ARN extrait est mesurée par l'appareil NanodropTM (Peqlab, VWR international GmbH, Erlangen, Allemagne).

5.6.2. Synthèse du cDNA et RT- PCR :

Plusieurs souches de lactocoques adaptés et tolérants des conditions extrêmes de température ont été évaluées au cours de ce travail. Pour survivre au stress thermique, les bactéries doivent répondre par l'expression de gènes liés à la résistance à haute températures. Plusieurs gènes

Matériels et Méthodes

impliqués dans la réponse au stress thermique ont été identifiés. Pour la recherche de gènes impliqués dans la tolérance à la température, la souche 9B a été sélectionnée sur la base de sa bonne et rapide croissance à des températures élevées (50°C) en milieu lait et en milieu M17. Le lactocoque 9Btesté a été cultivé sous stress thermique afin de permettre l'expression de gènes impliqués dans sa résistance à la haute température.

Les réactions de synthèse du cDNA (y compris le contrôle sans reverse transcriptase pour vérifier l'absence de toute contamination par l'ADN) sont réalisées à partir de cultures incubées à 30°C et à 50°C en milieu M17 et selon le protocole suivant : 1 µg de ARN purifié est additionné à 1 µl Random primer (Random Hexamer primer, Thermoscientific, 0,2 µg/µl) et 12 µl eau ultra pure puis incubé à 65°C pendant 5 minutes ensuite à 4°C pendant 5 minutes. 4µl du tampon (X5) de l'amplification additionné à 1µl de Ribolock RNase Inhibitor (Thermoscientific) et 2 µl d'un mix de(10 mM) dNTP (GENAXXON, Bioscience), 1 µl de la reverse transcriptase (Thermoscientific) est ajouté, le mélange est incubé à 25°C pendant 5 minutes puis à 42°C pendant 60 minutes enfin à 70°C pendant 5 minutes. La quantité et la pureté de l'ARN extrait est mesurée par l'appareil Nanodrop™.

Les programmes et les amorces utilisées en RT-PCR sont mentionnés dans le tableau 9.

5.7. Séquençage :

Les amorces utilisées pour le séquençage des différents gènes sont regroupées dans le tableau 9.

5.7.1. Séquençage partiel de l'ADNr 16S :

L'étude de l'opéron ribosomique (codant pour les ARNr 5S, 16S et 23S) constitue un outil de choix pour étudier les relations phylogénétiques grâce à sa présence chez toutes les bactéries, sa fonction constante et la présence de zones très conservées ainsi que de parties à séquence variable (Woese 1987; Hopfl et al. 1989; Goebel and Stackebrandt 1994; Rainey et al. 1994). Actuellement la plus importante base de données pour le développement de la classification des bactéries lactiques et de l'ensemble des bactéries est constituée par les séquences du gène codant pour l'ARNr16S. Notre analyse porte sur la partie terminale de 251pb de l'ADNr 16S qui permet une discrimination entre les différentes sous espèces de *Lactococcus lactis* : ssp *lactis*, ssp *cremoris* et ssp *hordniae*. Par comparaison des séquences complètes des ADNr des trois sous-espèces *L.l.lactis* IL1403, *L.l.cremoris* MG1363 et *L.l.hordniae* NCDO2181; on a pu déterminer une séquence dans la région 5-prime des gènes contenant des différences spécifiques qu'on a utilisé pour dessiner des amorces pour l'amplification et pour séquençage. *L. lactis* subsp. *lactis* se laisse différencier des autres sous-espèces par un "G" spécifique (pos.

91, souligné), pour *L. lactis* subsp. *Hordniae* le “T” (pos. 143souligné) est typique et chez *L. lactis* subsp. *cremoris* on trouve plusieurs nucléotides caractéristiques dans cette région.

Après l’amplification par PCR du gène codant pour l’ARNr16S. L’ADN est utilisé (40 à 50 ng) pour une réaction de séquence, et ensuite analysé dans un séquenceur Sanger 3730 XL de ABI (SeqLab, Sequence Laboratories GmbH, Göttinger, Allemagne) en utilisant les mêmes amorces pour l’amplification. Les séquences d’ADN sont obtenues sous forme de pics colorés avec 4 couleurs différentes correspondant chacune à une base de l’ADN. Elles sont ensuite corrigées à l’aide de différents logiciels. Les séquences partielles du gène 16S rDNA sont identifiées d’abord par Blast et sont analysées ensuite phylogénétiquement.

5.7.2. Séquençage du gène *gdpP*:

16S rRNA a plusieurs caractéristiques importantes: sa fonction essentielle, son ubiquité et ses propriétés évolutives l’ont permis d’être le marqueur moléculaire communément le plus utilisé dans l’identification des bactéries. Cependant, ce gène est présent en plusieurs copies dans la bactérie allant de 1 à 15 copies (en moyenne 4 copie par génome). Ces copies intra génomiques peuvent être différentes en leurs séquences, conduisant à l’identification de multiples ribotypes pour un seul organisme. Cette hétérogénéité intragénomique a influencé le pouvoir discriminant et la performance du gène 16S rRNA dans la phylogénie bactérienne. On a choisi un autre gène additionnel le *gdpP* se trouve en une seule copie dans le génome bactérien et capable d’identifier les relations phylogénétiques à l’échelle des sous espèces qui sont indétectable en utilisant seulement le gène 16S rRNA. En plus, le *gdpP* confère à la bactérie une résistance vis-à-vis de l’acide, de la chaleur et de divers agents de stress de la paroi cellulaire (Smith et al. 2012) et il est présent uniquement chez les deux sous espèces ssp. *lactis* et ssp. *cremoris* et absent pas chez les deux autres ssp. *hordniae* et ssp. *tructae*. Pour toutes ces raisons, nous avons procédé à la détermination de la séquence du gène *gdpP*. Les séquences partielles obtenues sont identifiées et analysées phylogénétiquement.

Matériels et Méthodes

Tableau 6 : Programmes d'amplification PCR/RT-PCR de différents gènes étudiés

Gene	16SrDNA	<i>gdpP</i>	<i>dnaK</i>	<i>groEL</i>	<i>clpX</i>
Programme	04 min 94°C	05 min 94°C	04 min 94°C	04 min 94°C	04 min 94°C
	30 sec 94°C 30sec 59°C 1 min 72°C	30 sec 94°C 30 sec 50°C 30 sec 72°C 30 sec 94°C	30 sec 94°C 30 sec 59°C 1 min 72°C	30 sec 94°C 30 sec 47°C 1 min 72°C	30 sec 94°C 30 sec 49°C 1 min 72°C
	10 min 72°C Pause 8°C	10 min 72°C Pause 8°C	10 min 72°C Pause 8°C	10 min 72°C Pause 8°C	10 min 72°C Pause 8°C

Suite du tableau 8

Gene	<i>dnaJ</i>	<i>grpE</i>	<i>groES</i>	<i>hflB</i>	<i>rpoD</i>
Programme	05 min 94°C	05 min 95°C	04 min 94°C	05 min 95°C	05 min 95°C
	02 min 94°C 01 min 59°C 01 min 30 sec 72°C	45 sec 94°C 01 min 60°C 1min 74°C	30 sec 94°C 30 sec 55°C 07 min 72°C	30sec 94°C 01 min 50°C 02 min 72°C	30 sec 94°C 01 min 50°C 02 min 72°C
	05 min 72°C Pause 8°C	5 min 72°C Pause 8°C	10 min 72°C Pause 8°C	30 sec 94°C 01 min 50°C 01 min 72°C	30 sec 94°C 01 min 50°C 01 min 72°C
				07 min 72°C Pause 8°C	10 min 72°C Pause 8°C

<i>rpoB</i>
03 min 95°C
30 sec 95°C 30 sec 56°C 30 sec 52°C 45 sec 72°C
8 min 72°C Pause 8°C

Matériels et Méthodes

Tableau 7 : Amorces utilisées dans l'amplification et le séquençage des différents gènes étudiés

Gène	Amorce	Réaction	Direction	Longueur de l'amorce (bases)	Séquence de l'amorce (5'→3')	Fragment amplifié pb
<i>groEL</i>	groELF1	PCR, séquence	forward	20	TGCGCAAGTAGCTACCGTTT	582
<i>groEL</i>	groELR1	PCR, séquence	reverse	20	GCAGAACCAGCACCTTCAAC	
<i>dnaK</i>	dnaKF1	PCR, séquence	forward	20	TGGTGGTGGTACATTGACG	542
<i>dnaK</i>	dnaKR1	PCR, séquence	reverse	20	CACCTTGAATTGCTGCACCC	
<i>clpX</i>	clpXF1	PCR, séquence	forward	20	TGATTGCCGGTTCAGACGTT	351
<i>clpX</i>	clpXR1	PCR, séquence	reverse	20	CTGCAATCGCAAACGGAACA	
<i>dnaJ</i>	dnaJF1	PCR	forward	20	CCAGAGACTTCAACACCGCT	821
<i>dnaJ</i>	dnaJR1	PCR	reverse	20	TCAGGTGGCTTTGGTGGTTT	
<i>dnaJ</i>	dnaJF2	PCR	forward	19	TGGAGGTTCTGGAGCGAAA	245
<i>dnaJ</i>	dnaJR2	PCR	reverse	20	AGCTTTCGCTCCAGAACCTC	
<i>dnaJ</i>	dnaJF3	PCR	forward	21	GCGGAGCTTCTGGTTTCTCTA	226
<i>dnaJ</i>	dnaJR3	PCR	reverse	20	CGGAAATTAGCCCCGTTTA	
<i>groES</i>	groESF1	PCR	forward	20	GGACAAGGTGATGCTGGTGA	157
<i>groES</i>	groESR1	PCR	reverse	20	CAAGCGCTTCCCAAGAAAAA	
<i>grpE</i>	grpEF1	PCR	forward	21	CACCATCCATTTTTACGGTCG	251
<i>grpE</i>	grpER1	PCR	reverse	20	AACGTGCACTTGCTGTTGAA	
<i>hflB</i>	hflBF1	PCR	Forward	20	GAGAAGCCCCACTGAGTCAC	319
<i>hflB</i>	hflBR1	PCR	Reverse1	21	CGGAAGCGCAGAAAGAATCGT	
<i>hflB</i>	hflBF2	PCR	Forward	19	CTTTCTGCGCTTCCGTCAT	844
<i>hflB</i>	hflBR2	PCR	Reverse	20	GCAAAGGTGCCATGTCGTTT	
<i>hflB</i>	hflBF3	PCR	Forward	22	GACGTGCCATCAGGTCATCAAT	581
<i>hflB</i>	hflBR3	PCR	Reverse3	20	GACGGAAGCGCAGAAAGAGC	
<i>rpoD</i>	rpoDF1	PCR	Forward	18	ACATGGCGCCAGATAAAG	354
<i>rpoD</i>	rpoDR1	PCR	Reverse	21	TTAGAACGACGTGGGTGACGT	
<i>rpoD</i>	rpoDF2	PCR	Forward	19	GGCGCCAGATAAAGTTCGT	242
<i>rpoD</i>	rpoDR2	PCR	Reverse	21	GCGTCCATCATCAAGACCGAG	
<i>rpoD</i>	rpoDF3	PCR	Forward	22	ACGCGATCCTTCACCAGAACGC	400
<i>rpoD</i>	rpoDR3	PCR	Reverse	21	AACGGTTTAGAACGACGCGCG	
<i>rpoD</i>	rpoDF4	PCR	Forward	22	ATCCTTCACCAGAAGAACGGTT	394
<i>rpoD</i>	RpoDR4	PCR	Reverse	21	TAGAACGACGTGGGAGCCAAG	
<i>gdpP</i>	gdpPF1	PCR, séquence	forward	20	TCCATCAGAAGATTCATTCA	667
<i>gdpP</i>	gdpPR1	PCR, séquence	reverse	19	TGACGTGCTGACAGTAATT	
<i>gdpP</i>	gdpPF2	PCR, séquence	forward	20	CAAGAATACGATAATCACAG	923
<i>gdpP</i>	gdpPR2	PCR, séquence	reverse	20	AATGCTTATTTTAGAACTGC	
16SrDNA	NG_Lactis_16s1	PCR, séquence	forward	21	GTGCCTAATACATGCAAGTTG	251
16SrDNA	NG_Lactis_16s1	PCR, séquence	reverse	20	CTATGTATCATCGCCTTGGT	
16SrDNA	NG_Lactis_16s2	séquence	reverse	21	CGGGATCATCTTTGAGTGATG	
<i>rpoB</i>	Lac-RpoF2	PCR	forward	17	GTTGAAGCTTCAACAGC	F2 x R2 200
<i>rpoB</i>	Lac-RpoR2	PCR	reverse	17	TCACGAGCAGTTTGATG	
<i>rpoB</i>	Lac-RpoR4	PCR	reverse	17	TTTCTCCAAGTCGAGCC	F2 x R4 370
<i>rpoB</i>	Leu-RpoF3	PCR	forward	17	TGACAAGCGTGTGGTG	250
<i>rpoB</i>	Leu-RpoR2	PCR	reverse	17	ACTTTAACACGGCCTGC	
<i>rpoB</i>	Ent-RpoF	PCR	forward	17	GCTGTAAAAACGGTTCG	330
<i>rpoB</i>	Ent-RpoR2	PCR	reverse	17	TGTGAATTCCCAAGAG	

5.8. Construction phylogénétique :

L'analyse phylogénétique est composée principalement par alignement de séquence et construction des arbres phylogénétiques.

Pour l'exploitation des données de séquences, ces dernières sont tout d'abord corrigées grâce à certains logiciels disponibles sur Internet tels que CLUSTALX ou CHROMAS-Pro et comparées avec d'autres séquences stockées dans des bases de données GENE BANK. Cette recherche est dite par « BLAST », du nom du programme informatique (Altschul et al. 1997) et se fait via l'internet (<http://www.ncbi.nlm.nih.gov/blast>). Le résultat est donné sous forme d'une liste des 100 séquences les plus proches de celle que l'on a soumis.

Les séquences corrigées doivent être ensuite alignées avec celles de souches de référence pour le même gène. Cet alignement peut être réalisé automatiquement par Dialign-TX software (Subramanian et al. 2008), DIALIGN (version 2.2, (Morgenstern 1999; Al Ait et al. 2013) et/ou CLUSTALW (Thompson et al. 1994).

Le fichier d'alignement doit être transféré sous un format reconnaissable par le logiciel de construction d'arbre phylogénétique. Les formats possibles sont FASTA, NEXUS ou PIR.

Les logiciels de construction d'arbres phylogénétiques utilisés dans cette étude sont PhyML3.0 (Criscuolo 2011), Seaview (Galtier et al. 1996), Splitstree 4 (Huson and Bryant 2006) et Figtree (Vlad I Morariu et al. 2008). Avant d'utiliser ces logiciels, il convient de choisir la méthode de construction phylogénétique adéquate. Trois méthodes sont les plus utilisées en phylogénie : celle de distance (Algorithmique), de parcimonie (Minimum Evolution) et de vraisemblance (Maximum-Likelihood).

- Pour les méthodes de distance, il s'agit tout d'abord de choisir le critère de distance entre les futures feuilles de l'arbre. Dans le cas où les feuilles sont des séquences d'ADN, on peut choisir comme distance le nombre de nucléotides qui diffèrent. Pour déterminer cette valeur, on est amené à en effectuer un alignement. Puis on peut utiliser la méthode UPGMA ou celle du Neighbour-Joining (Saitou and Nei 1987) pour en déduire l'arbre.
- Les méthodes de parcimonie consistent à trouver l'arbre qui minimise le nombre de mutations, délétions, ou insertions ponctuelles pour passer d'une séquence à l'autre. C'est une méthode très lente si l'on génère tous les arbres possibles pour en calculer la parcimonie.

Matériels et Méthodes

Les méthodes de vraisemblance sont plus probabilistes. En se basant sur le taux de substitution pour chaque élément de base (nucléotide pour les séquences d'ADN) au cours du temps, on estime la vraisemblance de la position et de la longueur des branches.

En taxonomie actuelle, les méthodes de maximum de vraisemblance et de Bayesian permettent la meilleure optimisation des données de séquences et l'obtention de résultats de phylogénie les plus robustes.

5.9. Séquençage du génome total d'un lactocoque thermotolerant :

Certains termes techniques ne trouvent pas de traduction ni d'équivalent corrects dans la langue française. Ils seront donc employés en langue anglaise tout au long de ce manuscrit, et ce en italique.

5.9.1. Choix de la souche :

La souche de *Lactococcus lactis* ssp. *Lactis* bv. *diacetylactis* HD9B_GL2 a été sélectionnée pour le séquençage. Cette souche, isolée en 2009 du lait cru de chamelle à Tindouf, une région chaude du sud algérien et rassemblant plusieurs propriétés d'intérêt biotechnologique laitier (Drici et al. 2010). L'originalité de l'isolat 9B réside aussi dans sa capacité à se développer à 50°C et à une spécificité d'hydrolyse des caséines différente de celles des PrtPs de type PI et PIII des lactocoques laitiers actuellement connus. La souche HD9B_GL2 héberge plusieurs plasmides dont trois dénommés pPrt9B (10 kb), pLac9B (12 kb) et pCit9B (25 kb) sont impliqués dans trois différentes fonctions métaboliques : l'hydrolyse des caséines du lait (pPrt9B), la fermentation du lactose (pLac9B) et du citrate (pCit9B). La souche HD9B_GL2 a été retenue aussi car elle montre un phénomène de transfert horizontal de gène car le système cit responsable du transport et du métabolisme du citrate porté par le plasmide pCit 9B (25 kb), est identique à celui de la souche *Leuconostoc paramesenteroides* J1. Il est composé du gène citI et de l'opéron cit-MCDEFGRP au lieu de l'opéron citQRP habituel des biovariants diacetylactis, le cluster citI-citMCDEFGRP a été acquis d'autres bactéries lactiques (*Leuconostoc*).

5.9.2. Stratégie de séquençage envisagée :

La majeure partie des protéines codées par les gènes *heat shock* présentent une affinité pour d'autres protéines cellulaires, cibles de la chaleur, qu'elles assistent, désagrègent, restaurent ou dégradent. Elles participent vraisemblablement à la protection des cellules ou à leur sauvetage

Matériels et Méthodes

lorsqu'elles sont menacées par leur environnement thermique. Ces gènes sont isolés ou groupés en petits opérons. Pour pouvoir caractériser ces gènes *heat choc* et identifier leur organisation génétique au niveau du chromosome chez les lactocoques, on a adopté une stratégie de séquençage du génome total d'un lactocoque atypique thermotolerant à croissance rapide désigné HD9B_GL2.

Pour obtenir la séquence génomique complète de la souche *Lactococcus lactis* ssp. *Lactis* bv. *diacetylactis* HD9B_GL2, un séquençage par la technique MiSeq d'Illumina a été commandé auprès d'un prestataire, StarSEQ (StarSEQ GmbH, Mainz, Allemagne) en utilisant le kit MiSeq reagent v2 contenant tous les réactifs nécessaires aux différentes étapes de la préparation de la librairie *Paired-End* selon le fournisseur Illumina. Le séquenceur utilisé est le MiSeq qui produit des lectures d'une taille moyenne de 250 nt. Une librairie de type « *paired end* » a été réalisée, ce type de librairie dite « *paired end* » ou lecture double, permet de séquencer les deux extrémités d'un fragment matrice. Ainsi, en plus de l'information apportée par la séquence, la distance séparant les deux lectures est connue (cette distance ne peut dépasser ~800 pb). Le protocole permettant la création d'une telle banque de séquence est donné en figure 13. Ce type de librairie permet un assemblage plus efficace des lectures en passant au travers des répétitions. La distance séparant les lectures de la banque commandée est de 250pb. Cette banque est très utile pour remettre en ordre et orienter les contigs obtenus les uns par rapport aux autres (étape de « *scaffolding* »). Toutefois, le prestataire a eu des difficultés à produire des séquences correctes en utilisant le protocole fourni par Illumina. Le prestataire a donc mis au point un protocole qui leur est propre pour réussir à créer une banque correcte. Les détails sur le séquençage de ces fragments est donné en figure 16.

5.9.3. Préparation de l'ADN génomique pour le séquençage haut- débit Illumina :

Une colonie pure est repiquée sur 2ml du milieu M17 liquide et incubée à 30°C pendant 18 heures. L'ADN génomique total de la souche 9B a été extrait par DNeasy blood and tissue kit (DNeasy Blood & Tissue Kit, QIAGEN GmbH, Hilden, Allemagne). L'ADN génomique ainsi purifié et élué dans 50 µl dans du tampon AE du même kit et envoyé dans un Eppendorf au prestataire après avoir vérifié sa concentration et sa qualité au Nanodrop (Peqlab, VWR international GmbH, Erlangen, Allemagne). La concentration de l'ADN génomique à séquencer selon Illumina doit être 1-2µg au minimum (20 ng/µl).

5.9.3.1. Préparation de la librairie :

Procédure expérimentale précédant le séquençage des fragments d'ADN d'intérêt. Varie en fonction de la technologie : L'ADN est fragmenté par cassure mécanique la sonification

Matériels et Méthodes

(Covaris TM S2) et lié à des adaptateurs. Les fragments sont liés à une surface (la flowcell) tapissée de séquences adaptatrices et amplifiés par PCR (Bridge PCR, C-Bot). La qualité de la librairie (la taille et la concentration finale souhaitée) est vérifiée par une électrophorèse et analysée par le Bioanalyzer TM. L'ensemble des fragments (clusters) seront séquencés par synthèse simultanément. La totalité des séquences générées sera assemblée créant une séquence consensus. Les deux adaptateurs utilisés dans ce travail de thèse sont les P5 et P9 dont les séquences sont:

P5 :

5' AATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTC
CGATCT (58 bases).

P9 :

5' GATCGGAAGAGCACACGTCTGAACTCCAGTCACGATCAGATCTCGTATGCC
GTCTTCTGCTG (63 base)

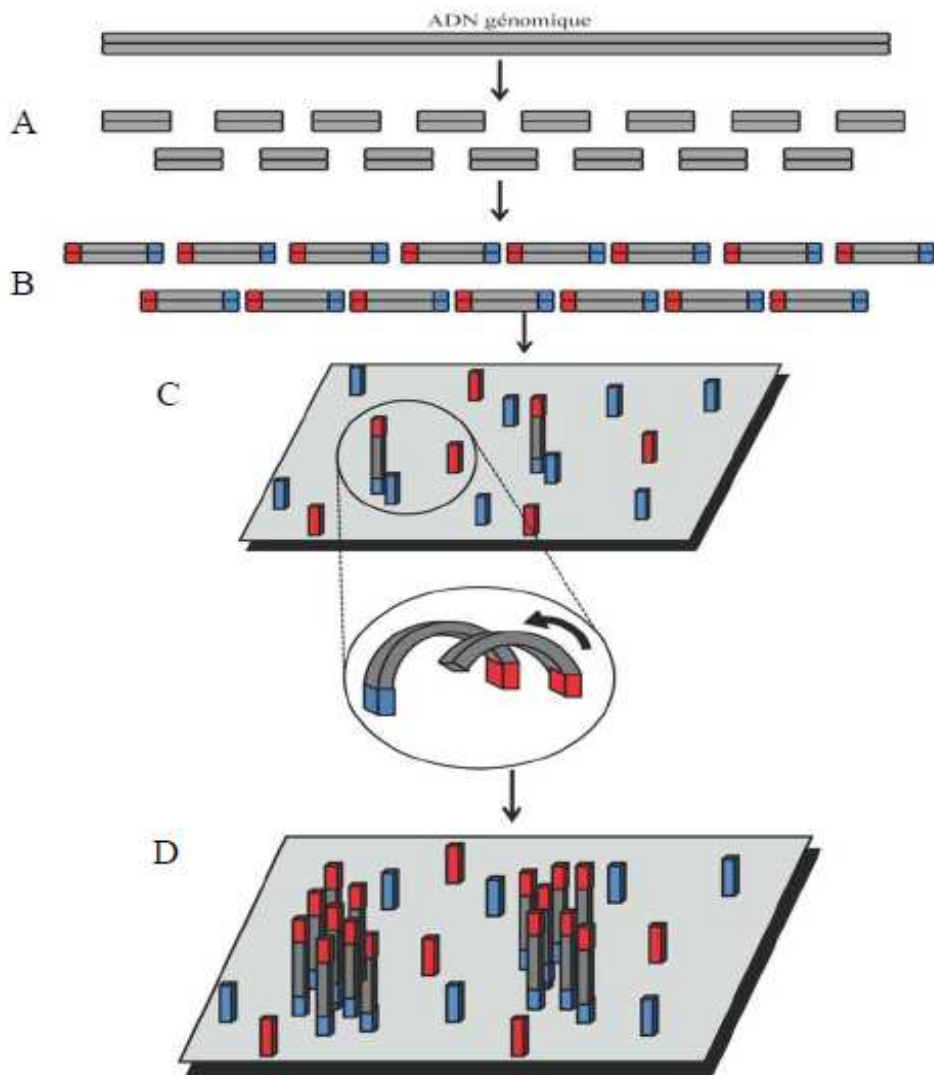


Figure 10 : Protocole de préparation de la librairie « *paired-end* » selon Illumina.

A. L'ADN génomique à séquencer est fragmenté par sonification. B. Des séquences adaptatrices sont ajoutées aux extrémités des fragments, puis le mélange réactionnel est mis à migrer par électrophorèse. Les fragments de taille désirée sont purifiés à partir du gel. La taille des fragments définira la distance séparant les deux lectures séquencées C. Les fragments à séquencer sous forme simple brin sont immobilisés sur une plaque de verre à l'aide d'amorces s'hybridant avec un des adaptateurs. D. Les fragments possédant les bonnes séquences adaptatrices sont sélectionnés par enrichissement spécifique de ces matrices par PCR. L'amplification des fragments se fait par pontage à l'aide de l'ADN polymérase. Cela permet l'amplification localisée d'un fragment unique en plusieurs centaines de millions de fragments identiques, appelés « clusters ».

5.9.4. Principe du séquençage par la technique Illumina (séquençage par terminaison cyclique réversible)

L'entreprise Solexa a commercialisé son premier séquenceur avec succès en 2006 : le génome Analyzer (GA). La spécificité de cette technologie repose sur une amplification en pont (*bridge PCR*) des fragments à séquencer. Elle a lieu sur une surface de verre appelée *flow cell* (FC), similaire à une lame de microscope, divisée en huit lignes (à l'origine, une ligne par échantillon). Les fragments de la librairie à séquencer possèdent des adaptateurs à leurs extrémités. Ceux-ci vont leur permettre de se fixer de façon aléatoire sur la FC, par hybridation sur les amorces qui en couvrent la surface (voir Figure 7). Un nouveau brin est alors synthétisé par une polymérase (Figure 7a) : il est fixé de façon covalente à la FC. Le brin d'origine est alors éliminé par dénaturation (b) et l'extrémité libre du brin restant s'hybride à une amorce adjacente pour former un pont d'ADN double brin (d) puis les deux copies sont libérées par dénaturation (e). Le cycle d'amplification en pont (étapes c à e) recommence pour former à terme un regroupement d'ADN clonal en une zone appelée *cluster* (f). Les brins anti-sens (correspondant aux amorces vertes) sont ensuite clivés (g) : c'est la linéarisation.

L'extrémité 3' libre des fragments d'ADN est bloquée et l'amorce de séquençage s'y hybride (voir Figure 8). Le séquençage s'effectue sur des centaines de millions de *clusters* simultanément, grâce à une chimie de terminateurs réversibles : des nucléotides bloqués marqués par fluorescence sont ajoutés, l'un d'entre eux est incorporé, la fluorescence émise est relevée puis le fluorophore et le bloqueur sont clivés permettant l'ajout d'un nouveau nucléotide. A chaque cycle d'incorporation, une base peut être déterminée.

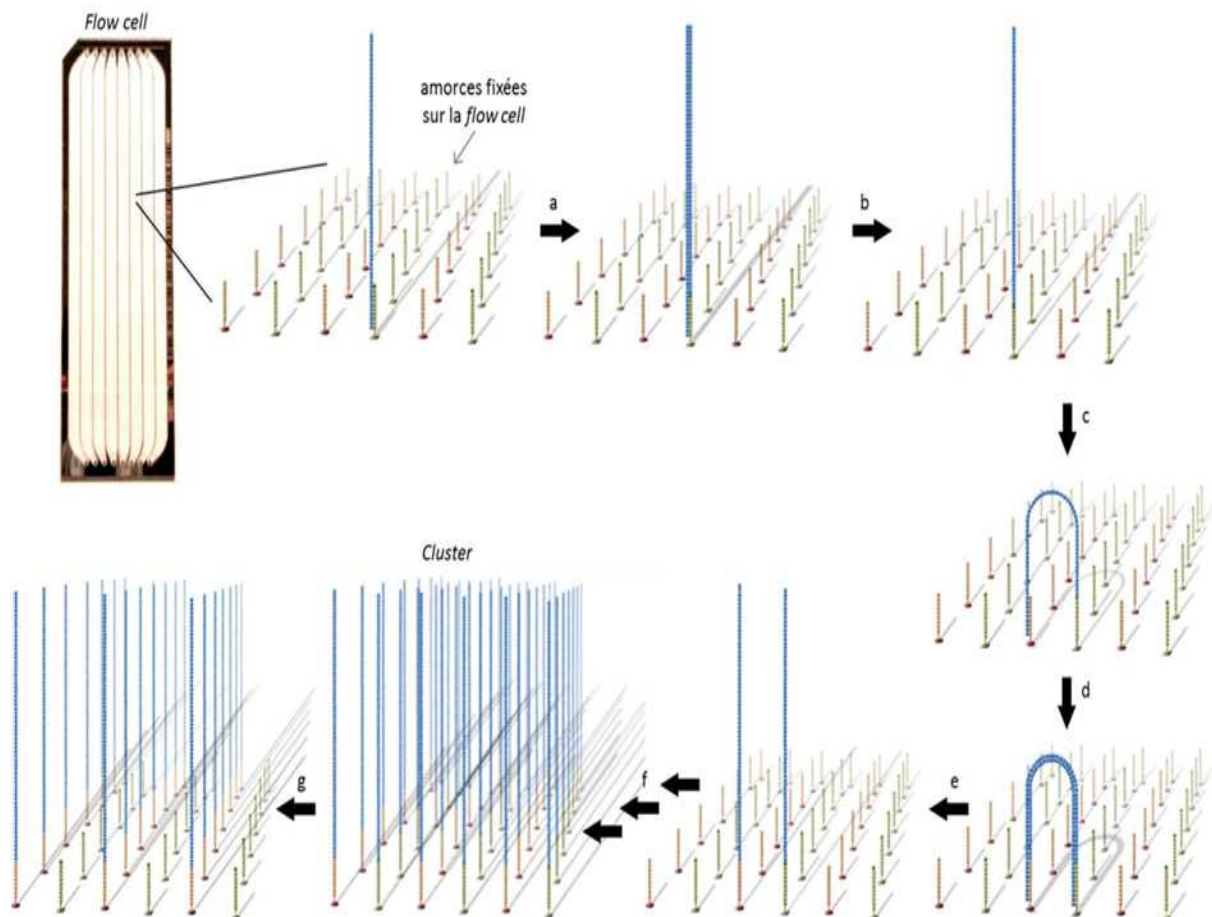


Figure 11 : Amplification en pont de la technologie Illumina

- Synthèse du brin complémentaire par une polymérase.
- Dénaturation.
- Formation d'un pont.
- Synthèse du brin complémentaire par une polymérase.
- Dénaturation.
- Nombreux cycles c à e.
- Linéarisation. Réalisé avec des images issues du site Internet d'Illumina.

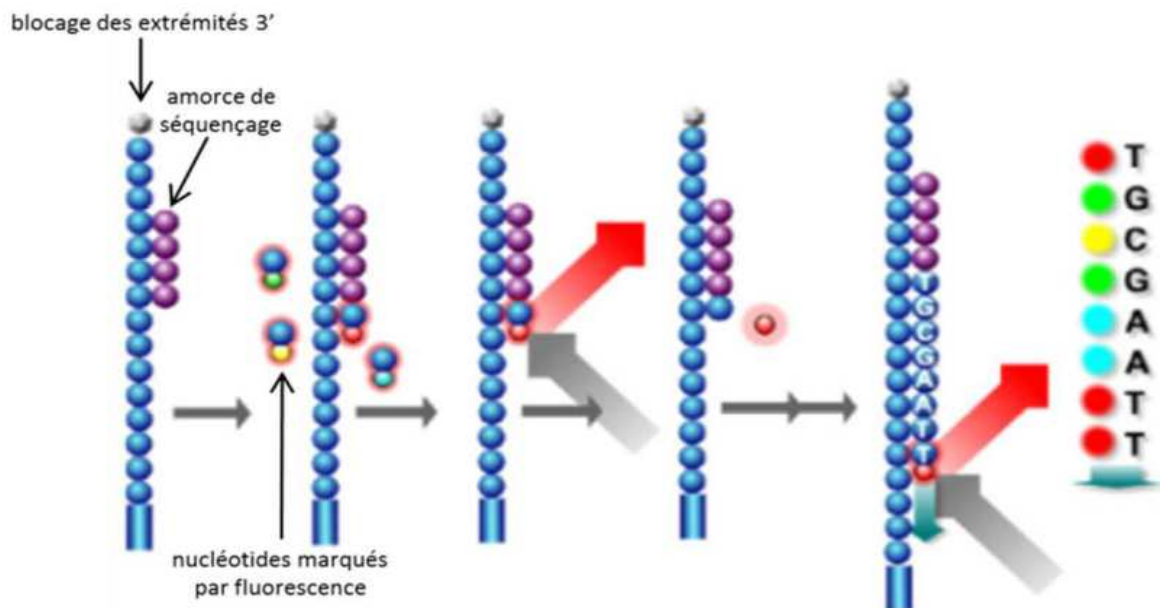


Figure 12: Aperçu de la technologie de séquençage Illumina

Adapté d'une image issue du site Internet d'Illumina

Illumina est maintenant le leader sur le marché du séquençage et possède la plateforme de séquençage la plus largement utilisée à travers le monde, en proposant un coût de séquençage du génome humain de seulement 10000 dollar.

Résumé du principe de séquençage par Illumina :

L'ADN génomique est fragmenté. Des adaptateurs sont fixés aux extrémités des fragments d'ADN. L'ADN est dénaturé puis chaque brin est immobilisé en le fixant à une amorce, elle-même fixée sur un support. Le brin complémentaire est synthétisé puis ce nouvel ADN double-brin est dénaturé. Chaque brin fixé forme alors un pont en s'hybridant localement avec l'amorce complémentaire de l'autre extrémité. Le brin complémentaire est synthétisé. L'opération est répétée un grand nombre de fois : il y a formation d'amas ("*cluster*") du même fragment d'ADN. Un nucléotide marqué par une étiquette fluorescente ("*fluorescently labeled chain-terminating nucleotide*") est incorporé dans la chaîne en cours de synthèse (synthèse du brin complémentaire de celui qui est séquençé). La fluorescence de ce nucléotide est mesurée.

Mais l'incorporation de ce terminateur de chaîne est réversible, ce qui permet que la synthèse continue jusqu'à ce qu'un autre terminateur de chaîne soit incorporé. Avantage d'Illumina: Faible taux d'erreurs d'insertion/délétion (indel). Limite : 1 milliard de bases séquencées (lectures d'environ 30-40 bases - les terminateurs de chaîne réversibles sont incorporés moins efficacement) avec une exactitude moindre que la technologie 454.

5.10. Analyse bio-informatique des données de séquençage

L'automatisation est ce qui conduit les technologies à être pratiques pour un nombre important de personnes. En sciences de la vie, l'automatisation permet d'accélérer les découvertes. De plus, l'automatisation augmente la reproductibilité des processus scientifiques.

En génomique, trois principales composantes technologiques guident les avancées. La première composante est le séquençage à haut débit de l'ADN (Shendure et al. 2005; Shendure and Ji 2008; Fuller et al. 2009; Ng et al. 2009; Hiatt et al. 2010; Marx 2013). Le séquençage de l'ADN permet de lire le contenu de l'ADN. Les systèmes de séquençage sont largement automatisés et permettent de décoder d'immenses quantités de données. La deuxième composante technologique en génomique est l'approvisionnement automatique de ressources informatiques ex : les superordinateurs. La troisième composante importante dans ce triangle est la disponibilité de logiciels adaptatifs pour la génomique. Un logiciel est dit adaptatif s'il peut s'adapter de façon à utiliser les ressources mises à sa disposition. Les nouvelles technologies de séquençage de l'acide désoxyribonucléique analysent des milliards de molécules en parallèles, permettant désormais l'obtention de très volumineux ensembles de données biologiques. De telles quantités de séquences promettent de répondre à d'importantes questions scientifiques. Cependant, il y a un besoin pressant pour des méthodologies parallèles pour transformer les séquences en connaissances.

De manière générale, les personnes utilisant des données de séquençage pour des activités de recherche utilisent majoritairement Microsoft Windows. Les autres systèmes d'exploitation utilisés (en ordre décroissant) sont : Apple Mac OS X, Linux (Ubuntu, Fedora, ou autre) ou autre (OpenBSD, FreeBSD, NetBSD, Minix, ou autre). La plupart des utilisateurs de Microsoft Windows et de Apple Mac OS X n'utilisent pas du tout la ligne de commande. A la quantité grandissante de séquences d'ADN, et à la baisse des prix pour obtenir des séquences d'ADN, des plateformes faciles à utiliser ont été introduites. En effet, plusieurs plateformes nouvelles permettent à un plus grand nombre de personnes (par exemple, les

biologistes) d'expérimenter avec les outils de la bio-informatique sans nécessiter une connaissance approfondie de la ligne de commande. Plusieurs plateformes sont commerciales (Illumina Base Space, DNA nexus) alors que d'autres sont ouvertes (Galaxy).

5.10.1. Types d'analyse

Principalement deux types d'analyse peuvent être réalisés sur des données de séquençage de l'ADN. Ces deux types sont les alignements (ou les analyses utilisant des alignements) et l'assemblage de novo. Une séquence de référence pour le génome de l'espèce étudiée est nécessaire pour aligner des séquences d'ADN. Ce n'est pas le cas pour l'assemblage de novo. En méta-génomique, l'assemblage de novo est souvent nécessaire puisque les génomes de références nécessaires pour aligner des séquences ne sont pas disponibles.

5.10.2. Alignement

Le reséquençage consiste à obtenir de l'information génétique pour un échantillon et à comparer cette information à une information de référence qui est considérée normale. Les différences entre les informations obtenues et le référentiel suggèrent des changements dans l'architecture génétique qui doit être validée afin d'en vérifier l'impact sur le phénotype. Le terme reséquençage est souvent associé implicitement à des méthodes d'analyse utilisant les alignements. Les structures de données et les algorithmes sont différents dans ces aligneurs. Cependant, ce sont des systèmes similaires qui alignent des lectures d'ADN (format : fastq) sur une référence (format : fasta) et produisent des alignements (format : sam (Li et al. 2009)).

5.10.3. Construction de gabarits « *Scaffolding* »

La construction de gabarits (scaffolds) est une tâche bio-informatique. Elle consiste à ordonner et à orienter les séquences assemblées en utilisant des paires de séquences ou en utilisant des cartes optiques de restriction (Nagarajan et al. 2008). Myers et ses collaborateurs ont proposé un algorithme vorace pour la construction de gabarits (Althaus et al. 2002). Les nouveaux outils utilisés dans ce travail et qui permettent de construire des gabarits génomiques, comme SSPACE (Boetzer et al. 2011), sont des meilleures implémentations (vitesse, facilité d'utilisation).

5.10.4. Estimation de la qualité globale des lectures et filtration des lectures

A chaque nucléotide des lectures de Solexa/Illumina est associé un score de qualité. Ce score est dépendant de la qualité du signal détecté en fonction du bruit de fluorescence. Un score élevé indique une probabilité plus faible que le séquenceur se soit trompé. La qualité de la séquence brute est évaluée. Les lectures possédant des nucléotides non-identifiés (représentés par un N dans la séquence) n'ont pas été considérées. Il a été démontré que les

lectures dont la longueur est trop faible ou trop élevée ont un nombre plus important d'erreurs. Les seuils de filtration ont été choisis de manière à conserver les lectures de meilleures qualités tout en limitant la perte d'information.

5.10.5. L'assemblage des données issues du séquençage nouvelle génération

Un assembleur est programme informatique dont la tâche est de combiner les lectures obtenues après séquençage en séquences contigües, aussi appelées contig, en se basant sur l'identité de séquences entre les lectures. De cette façon, deux lectures se chevauchant proviennent vraisemblablement de la même région. L'étape de « *scaffolding* » (échaffaudage) a pour but d'orienter et de placer les contigs les uns par rapport aux autres pour obtenir une séquence unique, bien que possédant des lacunes (soit des trous dans la séquence finale). Ces lacunes seront corrigées lors de l'étape de finition du génome qui vise à combler les trous restant dans l'assemblage par des méthodes expérimentales, comme la PCR ou l'extension d'amorces. Des indices statistiques tels que le N50 (Le N50 est la taille du plus petit contig tel que 50 % de la longueur cumulée de l'ensemble des contigs obtenus après assemblage soit contenue dans des contigs de taille égale ou supérieure), la taille moyenne ou la taille cumulée des contigs permettent d'estimer la qualité d'un assemblage. Le postulat est que plus les contigs formés sont grands, moins il y aura de trous dans la séquence génomique reconstituée. Toutefois, il est nécessaire de vérifier la justesse des contigs formés (la présence de contigs chimériques notamment).

5.10.5.1. Les difficultés de l'assemblage

La principale difficulté lors d'un assemblage est la présence de répétitions dans les génomes séquencés. En effet, pour un assembleur deux lectures se chevauchant parfaitement proviennent de la même région génomique. Or, certaines séquences peuvent se retrouver de multiples fois dans un génome entraînant la formation de contigs chimériques et de trous dans l'assemblage final (figure 13). Pour s'affranchir de ce problème, deux solutions sont possibles. La première est d'utiliser une technique de séquençage produisant des lectures les plus longues possibles. Si une lecture est plus longue que la répétition, la présence de séquence spécifique de la répétition permettra de résoudre le problème. L'autre solution est d'utiliser une banque de séquence dite « *paired-end* » (aussi appelée « *mate-pair* »). Les lectures ainsi produites sont regroupées par paire dont l'orientation et la distance les séparant sont renseignées, ajoutant une contrainte à l'assembleur qui doit respecter ces deux critères supplémentaires lors de la recherche de chevauchements entre séquences. L'autre difficulté de

Matériels et Méthodes

l'assemblage de séquence se situe d'un point de vue matériel. La quantité de données générées par les nouvelles techniques de séquençage augmentant constamment, leur traitement demande l'utilisation de matériel informatique de plus en plus performant. Pour s'affranchir de ces difficultés, les assembleurs mis au point utilisent des méthodes de calcul permettant d'alléger les ressources matérielles demandées mais aussi d'améliorer la prise en compte des répétitions.

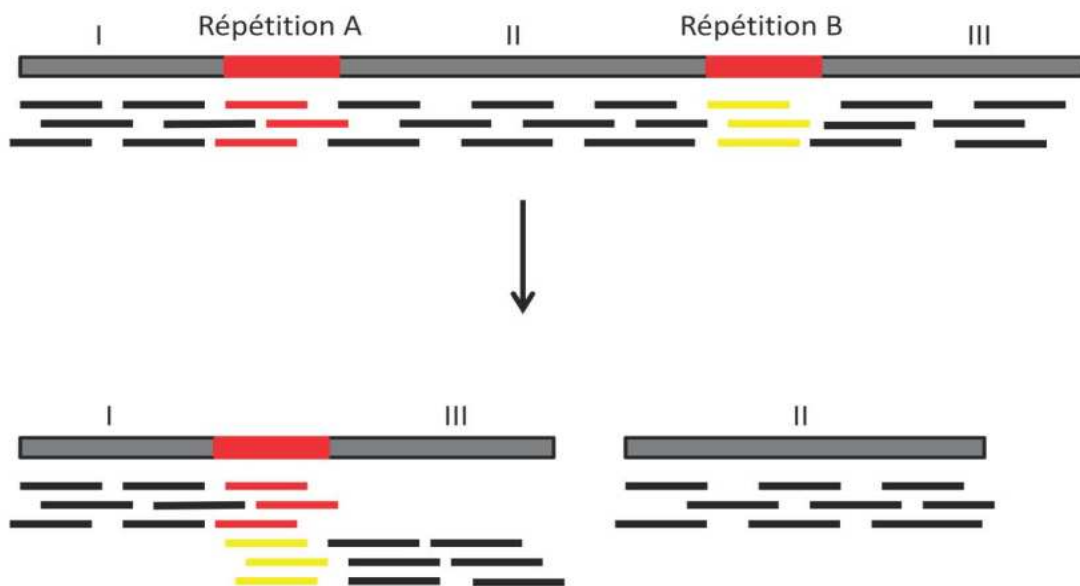


Figure 13 : Schématisation du comportement d'un assembleur vis-à-vis d'une répétition.

Pour l'assembleur, les lectures provenant de la répétition A (lectures rouges) et de la répétition B (lectures jaunes) sont supposées provenir de la même région génomique. Cela peut entraîner la formation de contigs chimériques et l'apparition de trous dans l'assemblage final.

5.10.5.2. Les méthodes d'assemblage des séquences

5.10.5.2.1. La méthode « *overlap-layout-consensus* »

Cette méthode est l'une des premières utilisées avec succès pour assembler une séquence génomique (Miller *et al.* 2010). Le principe général est de fusionner des séquences chevauchantes, en commençant par celles ayant les chevauchements les plus significatifs, jusqu'à formation d'une séquence unique (figure 14). La première étape, dite *overlap*, sert à calculer le meilleur alignement possible entre deux lectures en tenant compte des éventuelles erreurs de séquençage. Chaque alignement est pondéré par un score qui va tenir compte du pourcentage d'identité nucléotidique et de la longueur de cet alignement entre deux lectures.

L'étape de *layout* permet de déterminer l'ordre dans lequel les lectures vont être assemblées. Pour la plupart des assembleurs, des paires de lectures sont sélectionnées en fonction de leur score (les scores les plus élevés en premier) et, si l'alignement est considéré comme cohérent, les deux séquences sont fusionnées pour former un contig. Cette étape est la plus complexe de par la difficulté à déterminer la cohérence de l'alignement entre deux lectures. En effet, il est nécessaire de discriminer un chevauchement réel, où les différences entre les deux séquences sont dues aux erreurs de séquençage, d'un chevauchement dû à une répétition, où les différences sont dues aux mutations ou de vraies variations génétiques. La dernière étape dite de *consensus*, sert, comme son nom l'indique, à établir la séquence consensus des contigs obtenus lors de l'étape de *layout*. La méthode la plus simple est de déterminer à chaque position quelle base est la plus représentée.

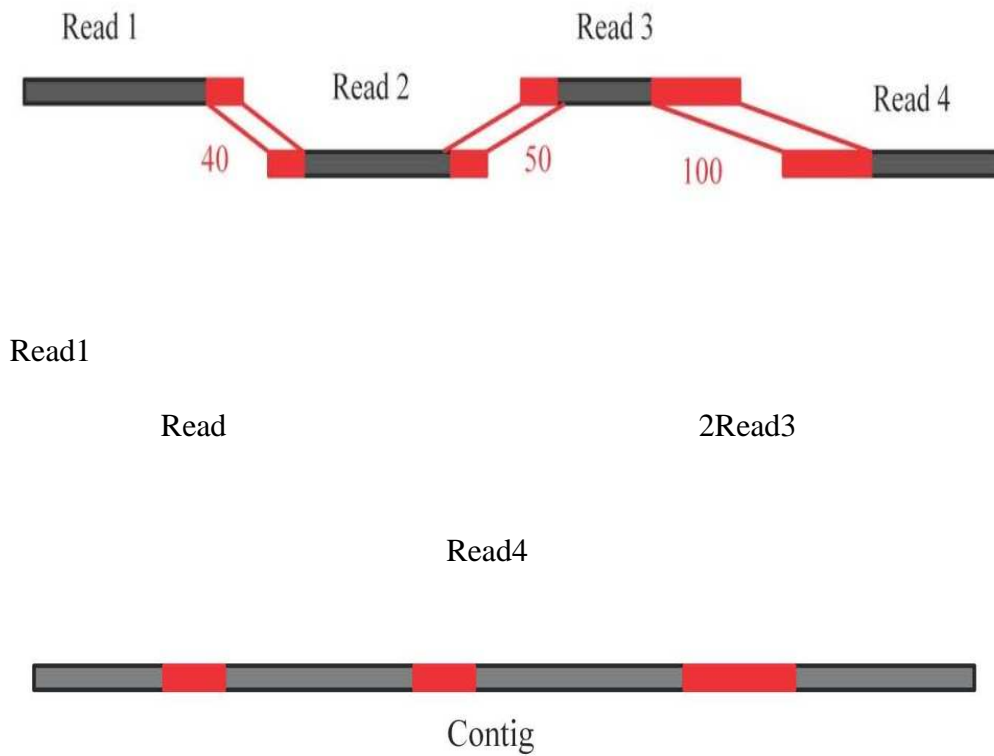


Figure 14 : Schématisation du principe de fonctionnement d'un assembleur utilisant la méthode Overlap-Layout-Consensus.

Dans un premier temps, des chevauchements entre les lectures sont déterminés en se basant sur l'identité de séquence. Chaque alignement est considéré comme significatif selon la longueur de l'alignement et le pourcentage d'identité de séquence (ici pour l'exemple il y a une pondération par un score bien qu'un assembleur puisse utiliser une autre méthode pour définir si un alignement est correct). La séquence génomique est recrée en alignant l'ensemble des lectures qui se chevauchent.

5.10.5.2.2. L'approche du graphique « *De Bruijn* »

L'assemblage de novo de génomes est le processus par lequel une grande quantité de morceaux d'ADN – lesquels sont appelés les lectures – sont mis ensemble afin de reconstruire des morceaux de plus grande taille (Nagarajan et al. 2008; Pop 2009). Ce type d'analyse nécessite des graphes, comme un graphe de *de Bruijn*, pour stocker et travailler sur les données. Cette méthode d'assemblage a été mise au point pour s'affranchir de l'étape complexe de *Layout* utilisée par la méthode « *overlap-layout-consensus* » (Pevzner et al. 2001). Les lectures qui doivent être assemblées sont préalablement fragmentées en sous-séquences chevauchantes de taille constante. Ces sous-séquences sont appelées K-mer (avec K correspondant à la longueur). Deux Kmer vont être assemblés si leurs séquences ne divergent que par le premier de l'un et le dernier nucléotide de l'autre (figure 9). Ainsi une identité parfaite doit être partagée entre les deux séquences qui se chevauchent. De cette façon il est théoriquement possible d'établir un lien entre chacun des K-mer, la présence des répétitions empêchant ce cas idéal dans la pratique, et de reconstituer une séquence génomique (figure 9). Il semble paradoxal de fragmenter des lectures alors que leur longueur est importante pour la résolution des répétitions. Cependant, cette méthode demande peu de ressources computationnelles par rapport au nombre de lectures qu'il est possible d'assembler. En effet, la recherche de chevauchements par la méthode « *overlap-layout-consensus* » nécessite de calculer des alignements entre chaque lectures. Ainsi, plus la quantité de données à analyser augmente, plus le temps et les ressources informatiques nécessaires à l'assemblage augmentent. Pour l'approche de « *de Bruijn* », cette étape de *layout* n'a pas à être opérée pour l'assemblage, car une identité de séquence parfaite doit exister entre deux lectures pour qu'elles se chevauchent. Cette approche d'assemblage est devenue la référence pour assembler des séquences issues de techniques de séquençage produisant un grand nombre de données comme celles d'Illumina/Solexa ou de SOLiD. Toutefois, ce principe d'assemblage est particulièrement sensible aux erreurs car une identité parfaite entre K-mer est requise pour valider un lien. Ainsi en pratique, l'assemblage de séquences par cette méthode demande des lectures possédant peu d'erreurs pour être efficace.

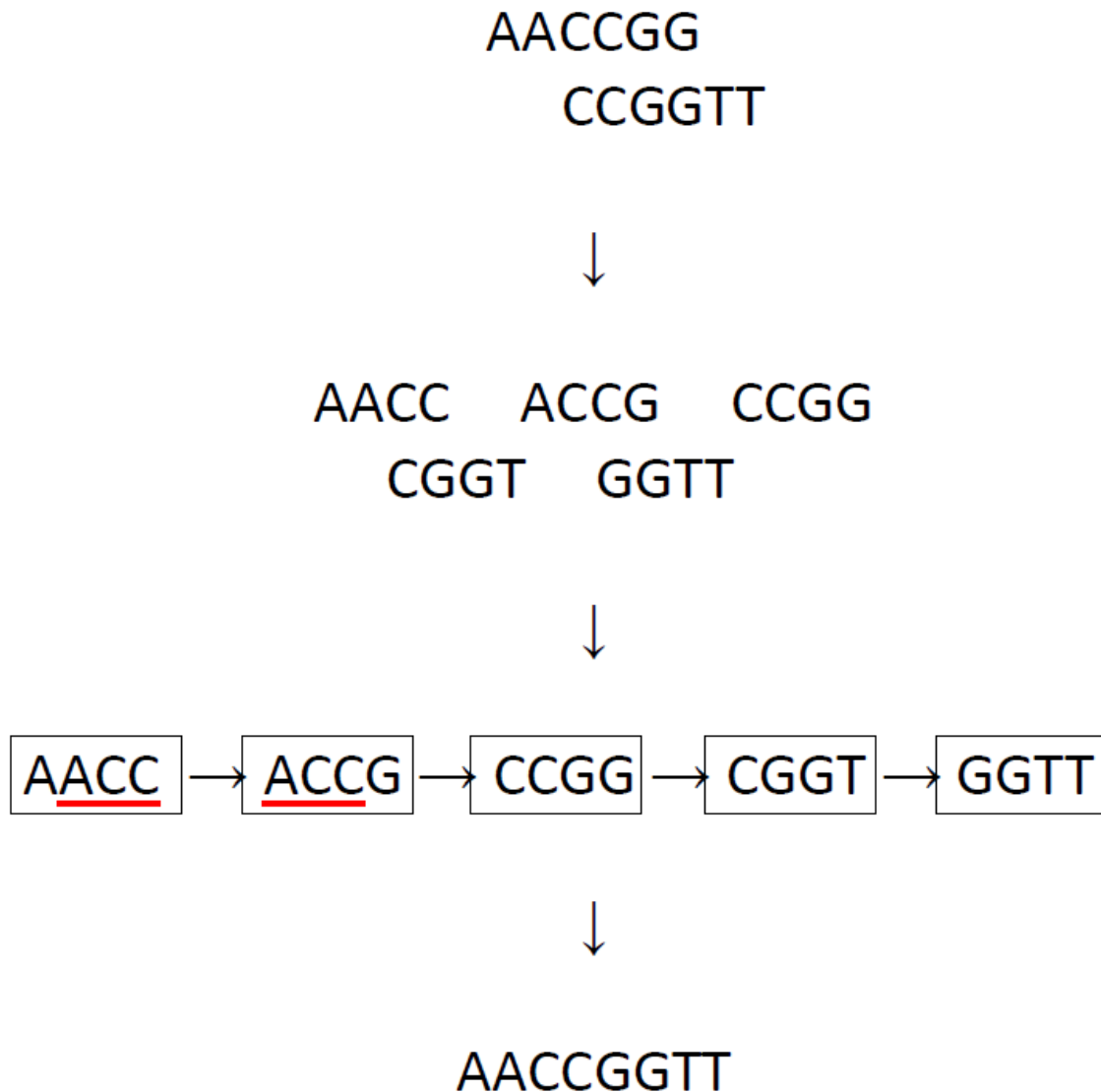


Figure 15 : Représentation schématique de l'assemblage de deux séquences par la méthode graphique « *De Bruijn* ».

Les deux séquences assemblées pour cet exemple sont découpées en 4-mer chevauchants. Il est à noter que le 4-mer CCGG qui est une répétition n'est représenté qu'une fois par la méthode graphique de « De Bruijn ». Une identité de séquence parfaite entre les 3 premiers nucléotides de l'un et les 3 derniers de l'autre est recherchée entre ces sous-séquences. S'il y a concordance exacte (comme par exemple pour les 4-mer soulignés en rouge), l'assembleur va relier les deux fragments. Le consensus obtenu fournit la séquence initiale.

5.10.5.3. Vue d'ensemble des assembleurs et leur utilisation

L'utilisation d'un assembleur particulier sera conditionnée par de nombreux paramètres tels que la stratégie de séquençage utilisée, la présence de « *paired-end* » ou encore la quantité de données. Les assembleurs utilisant la méthode « *overlap-layout-consensus* » sont plutôt utilisés avec des lectures allant de 100 à 800 pb (principalement avec un séquençage Sanger ou un pyroséquençage haut-débit). La méthode graphique « *De Bruijn* » s'emploie plutôt avec des lectures d'une taille allant de 25 à 100 pb. Cette dernière est la méthode d'assemblage la plus utilisée à l'heure actuelle grâce au succès de la plateforme Illumina/Solexa. Toutefois, avec les améliorations des techniques de séquençage, la taille des lectures ne cesse d'augmenter et la méthode « *overlap-layout-consensus* » devrait redevenir la référence dans le futur. La plupart des assembleurs ont démontré leur efficacité sur l'assemblage de génomes bactériens car ils sont plus simples à assembler que les génomes eucaryotes, notamment par la plus faible longueur de leurs génomes et leur plus faible teneur en répétitions. Il faut toutefois noter que certains assembleurs ont été testés sur des données simulées (Butler et al. 2008) alors que d'autres ont montré leur efficacité sur des données réelles (Maccallum et al. 2009), ce dernier critère étant plus déterminant dans le choix de l'utilisation d'un assembleur.

Les assembleurs intégrés dans le pipeline de l'assemblage et utilisés dans ce travail de thèse sont présentés dans le pipeline suivant Figure 16.

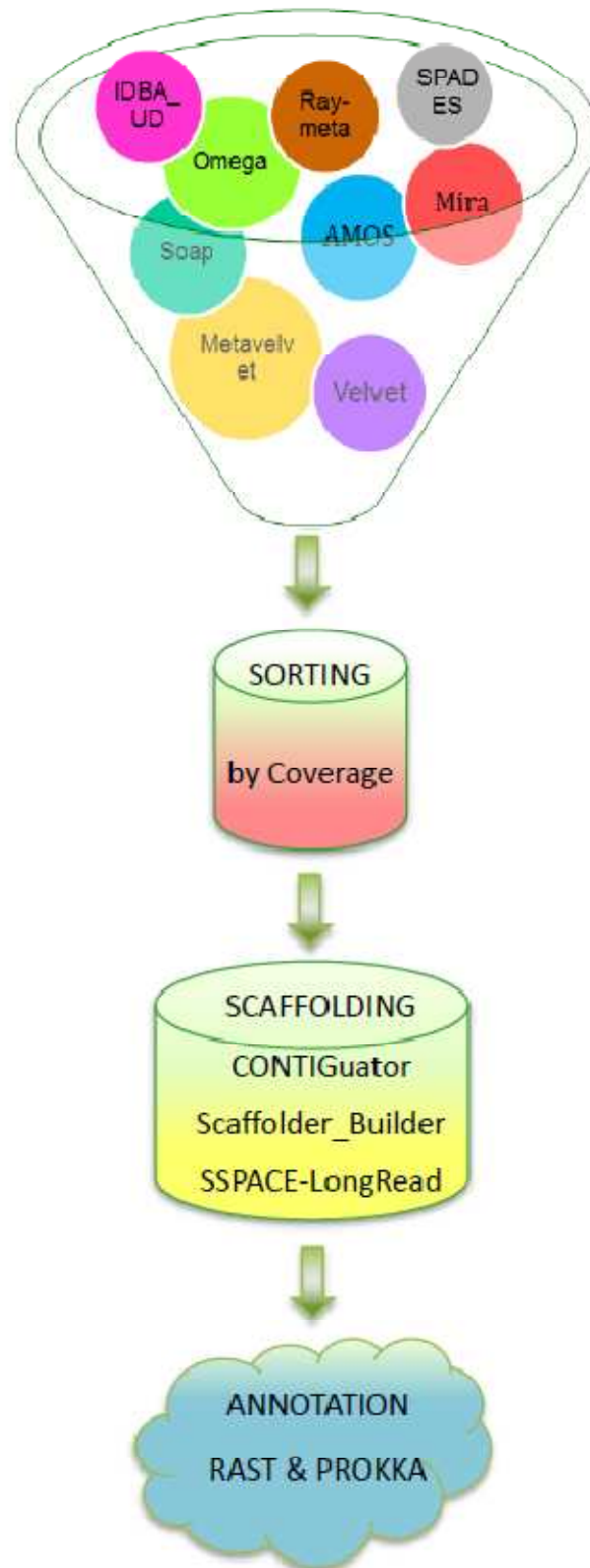


Figure 16: Logiciels intégrés dans le pipeline d'assemblage et l'annotation génomique et méta-génomique de la composante HD9B

5.11. Assemblage génomique et méta-génomique

Une approche d'optimisation « pipeline » est développée pour l'assemblage des séquences obtenues du séquençage (Figure 16). En première étape, des assembleurs du génome total et de la métagénomique ont été triés, car les données obtenues peuvent être issues d'une petite communauté ou une mini-métagénomiques. En comparant la qualité des contigs assemblés par différentes approches, les meilleurs résultats sont choisis à l'aide de BLAST « Basic Local Alignment » (BLAST, (Altschul et al. 1997) selon deux critères importants : le taux de la couverture des lectures et l'identité. Ce classement permet de séparer les contigs appartenant aux différentes souches ou espèces. Ensuite, d'autres programmes sont utilisés pour la construction de gabarits “*scaffolds*” permettent de relier les contigs pour avoir le génome complet. Finalement, le génome peut être annoté par les programmes appropriés.

5.11.1. Assembleurs utilisés :

Les deux types d'assembleurs génomique et méta-génomique sont utilisés. MIRA 4 (*Mimicking Intelligent Read Assembly*) (Chevreux et al. 2004) est un assembleur (assembler/mapper) est particulièrement réputé pour son usage lors de projets de séquençage de génomes possédant beaucoup de répétitions. AMOS (*A Modular Open-Source Assembler*) est une collection des outils d'assemblage des génomes (Treangen et al. 2011). SOAPdenovo est un assembleur de novo utilise pour l'assemblage de courtes séquences nucléotidique en large contigs (Luo et al. 2012a). Velvet (Zerbino and Birney 2008) c'est un assembleur, qui fonctionne selon la méthode de « de *Brujn* », a démontré son efficacité dans de nombreux projets de séquençage de génomes bactériens.

Les assembleurs utilisés pour l'assemblage de méta-génomiques présents dans les échantillons mixte provenant de différents microorganismes sont largement utilisés dans l'étude de petite communauté écologique (Namiki et al. 2012) : Ray Meta (Boisvert et al. 2012) est un assembleur de novo de génomes, de méta-génomiques (Schatz et al. 2010). L'assemblage consiste en la construction de longues séquences à partir de plus courtes séquences. Metavelvet (Namiki et al. 2012). IDBA_UD (Peng et al. 2012) est un assembleur itératif utilisant plusieurs valeurs de k (Yang et al. 2010; Peng et al. 2011; 2012). SPAdes (version 3.5.0, (Bankevich et al. 2012) est un autre assembleur qui fonctionne aussi lorsque les données de séquençage ont été obtenues à partir d'une seule cellule (Pham et al. 2013) , sa version originale peut aussi assembler les lectures issues d'isolats standard et de données

issues d'une seule cellule . IDBA_UD Omega (Haider et al. 2014) est un assembleur de novo se basant sur la méthode de overlap-graph.

5.11.2. Annotation du génome et génomique fonctionnelle

Rapid Annotation using Subsystem Technology (RAST, (Aziz et al. 2008) est un outil largement utilisé pour l'annotation des génomes bactériens. Prokka (Seemann 2014) est un package utilisé aussi pour une annotation rapide des génomes de petite taille. L'annotation des phages est réalisée par PHAge Search Tool (PHAST, (Zhou et al. 2011). En utilisant ces logicielles, l'information fonctionnelle du génome peut être rapidement annoté. L'analyse comparative est basée sur BLAST (Altschul et al. 1990) pour pouvoir comparer le génome de la souche HH9B_GL2 avec ceux d'autres souches de référence complètement séquencées. Les séquences génomiques des souches utilisées pour l'étude comparative sont extraites du GenBank (National Center for Biotechnology information (NCBI, <http://www.ncbi.nlm.nih.gov/>) sous les numéros d'accès suivants:

L. lactis ssp. *lactis* IL1403 (abbr. IL1403, accession no. AE005176), *L. lactis* ssp. *lactis* NCDO2118 (abbr. NCDO2118, accession no. NZ_CP009054.1), *L. lactis* ssp. *lactis* KLDS4.0325 (abbr. KLDS4. 0325, accession no. CP006766), *L. lactis* subsp. *cremoris* MG1363 (abbr. MG1363, accession no. NC_009004.1), KF147 (NC_013656), *Lactococcus lactis* subsp. *lactis* CV56 (NC_017486.1), *Lactococcus lactis* subsp. *cremoris* KW2 (NC_022369.1), and *L. lactis* subsp. *cremoris* SK11 (abbr. SK11, accession no. CP_006766.1)

5.11.3. Analyse méta-génomique

MEGAN (version 5, (Huson et al. 2007)) est un analyseur de la métagénomique qui peut analyser et comparer des données métagénomique et basé sur l'alignement des séquences avec les données de référence dans les bases de données. MEGAN a été utilisé pour analyser les différentes composantes de nos résultats si elles proviennent d'une ou d'un mélange de souches. L'analyse phylogénétique est composée principalement par l'alignement de séquences et la construction d'arbres phylogénétiques. Les séquences sont alignées par DIALIGN (version 2.2, (Morgenstern 1999); un outil établi pour l'alignement multiple de séquences à faible homologie. Les arbres phylogénétiques sont construits par PhyML (version 3.1, (Guindon and Gascuel 2003). PhyML est relativement rapide et basé sur l'approche de maximum likelihood.

5.12. La microscopie électronique à balayage (MEB/SEM *Scanning Electron Microscopy*)

La microscopie électronique à balayage et les microanalyses sont employées très largement dans les secteurs académiques et industriels pour imager, caractériser et quantifier des échantillons solides de toute nature. Pour effectuer des analyses de qualité, précises et reproductibles il n'est souvent pas possible d'utiliser un échantillon tel quel. Une phase préalable de préparation est nécessaire. C'est le but de cet ouvrage d'expliquer aux lecteurs et utilisateurs du MEB et des microanalyses les différents moyens de réaliser cette étape cruciale précédant leur analyse.

5.12.1. Principe de la microscopie électronique à balayage

Le principe du balayage consiste à explorer la surface de l'échantillon par lignes successives et à transmettre le signal du détecteur à un écran cathodique dont le balayage est exactement synchronisé avec celui du faisceau incident. Les microscopes à balayage utilisent un faisceau très fin qui balaie, point par point, la surface de l'échantillon.

5.12.1.1. Interactions du faisceau électronique avec l'échantillon

Sous l'impact du faisceau d'électrons accélérés, des électrons rétrodiffusés et des électrons secondaires émis par l'échantillon (Figure 17) sont recueillis sélectivement par des détecteurs qui transmettent un signal à un écran cathodique dont le balayage est synchronisé avec le balayage de l'objet.

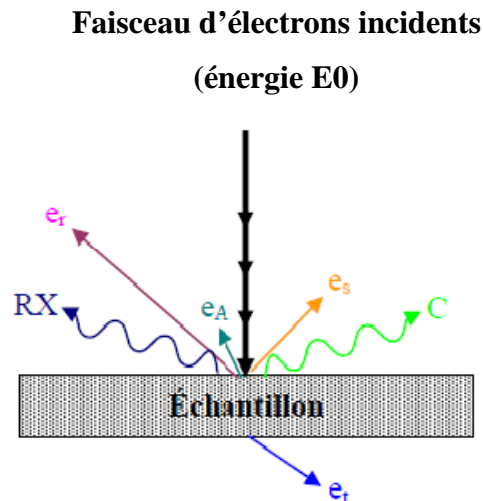


Figure 17 : Représentation schématique de l'interaction entre un faisceau d'électron et la surface d'un échantillon

Légende :

e_r : électrons rétrodiffusés (MEB sans métallisation)

e_s : électrons secondaires (échantillon métallisé)

e_A : électrons Auger

e_t : électrons transmis

C : cathodoluminescence

RX : rayons X

→ : rayonnement électronique

~ : rayonnement électromagnétique

5.12.1.2. Émission d'électrons secondaires

Il s'agit d'électrons arrachés par ionisation. Certains électrons incidents de faible énergie (< 50 eV) sont éjectés de l'échantillon sous l'effet du bombardement. Comme seuls les électrons secondaires produits près de la surface sont détectés, ils formeront des images avec une haute résolution en profondeur (3-5 nm). Le contraste de l'image est surtout donné par le relief de l'échantillon mais on peut également observer un contraste chimique dans le cas de grandes différences de numéros atomiques.

5.12.1.3. Émission d'électrons rétrodiffusés

Les électrons accélérés dans la colonne pénètrent dans l'échantillon. Un parcours plus ou moins important dans la matière leur fait perdre une fraction de leur énergie. La trajectoire suivie est aléatoire, et ils peuvent revenir vers la surface. Ils sont alors détectés après leur sortie de l'échantillon. Du fait de leur plus grande énergie, les électrons rétrodiffusés peuvent provenir d'une profondeur plus importante, et la résolution de l'image sera moins bonne qu'en électrons secondaires (6-10 nm). Suivant le type de détecteur utilisé, les électrons

rétrodiffusés fournissent une image topographique (contraste en fonction du relief) ou une image de composition (contraste en fonction du numéro atomique moyen).

5.12.1.4. Émission de rayons X

Le faisceau d'électrons est suffisamment énergétique pour ioniser les couches profondes des atomes et produire ainsi l'émission de rayons X. La résolution spatiale d'analyse dépend de l'énergie de la raie X détectée, de la nature du matériau et de la fluorescence secondaire. Elle est, en général, supérieure au micromètre.

5.12.1.5. Émission d'électrons Auger

Ce sont des électrons dont la faible énergie est caractéristique de l'élément émetteur (utilisés pour l'analyse élémentaire) et du type de liaison chimique.

5.12.1.6. Cathodoluminescence

Lorsque des matériaux isolants ou semi-conducteurs sont bombardés par le faisceau d'électrons, des photons de grande longueur d'onde (ultraviolet, visible) sont émis. Le spectre obtenu dépend du matériau étudié et de sa pureté.

5.12.1.7. Canalisation d'électrons

La pénétration du faisceau d'électrons dans un cristal est fonction de son incidence par rapport à une famille de plans interréticulaires. Le contraste des images en électrons rétrodiffusés donnera des renseignements sur la structure cristalline du produit.

En pénétrant dans l'échantillon, le fin faisceau d'électrons diffuse peu et constitue un volume d'interaction (poire de diffusion, Figure 18) dont la forme dépend principalement de la tension d'accélération et du numéro atomique de l'échantillon. Dans ce volume, les électrons et les rayonnements électromagnétiques produits sont utilisés pour former des images ou pour effectuer des analyses physico-chimiques. Pour être détectés, les particules et les rayonnements doivent pouvoir atteindre la surface de l'échantillon. La profondeur maximale de détection, donc la résolution spatiale, dépend de l'énergie du faisceau d'électrons incidents.

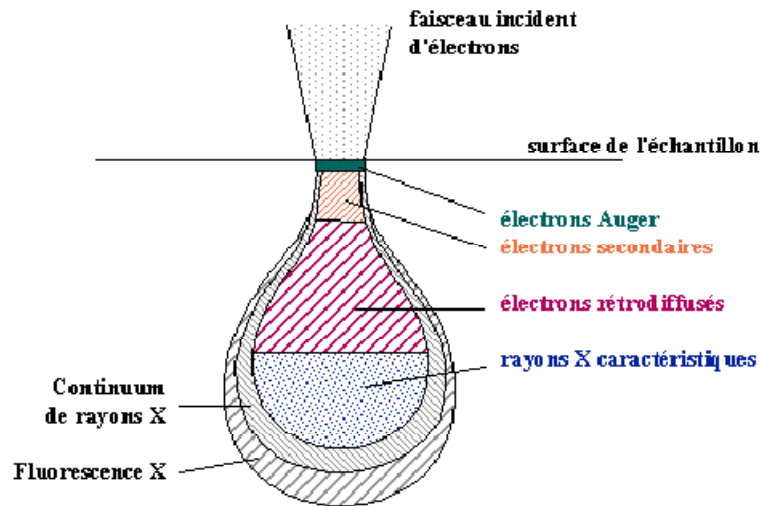


Figure 18 : Poire de diffusion détaille un spectre de distribution d'énergie selon le type d'électrons émis.

La Figure 4 détaille un spectre de distribution d'énergie selon le type d'électrons émis.

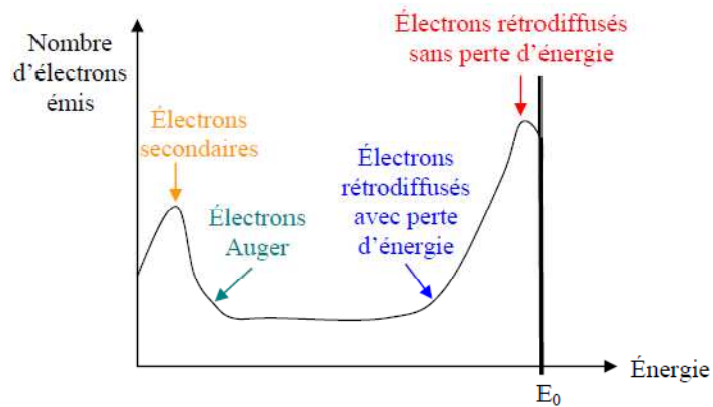


Figure 19: Représentation schématique de l'énergie de distribution des électrons émis par un échantillon

5.12.2. Équipement

Le microscope électronique à balayage utilisé dans le cadre de ce travail est un microscope électronique à balayage JEOL 7500F (SEM- *Scanning electron microscope*). Il comporte les éléments suivants (Figure 20) :

Une colonne :

- une source de rayonnement : canon à filament de tungstène; ce filament, parcouru par un courant électrique, émet spontanément des électrons, qui sont accélérés par un champ électrique leur conférant une certaine énergie;
- une « optique » : diaphragme et lentilles électrostatiques ou magnétiques (réduction du diamètre du faisceau et focalisation sur l'objet);
- un système de balayage : bobines défectrices qui commandent le point d'impact des électrons sur l'échantillon;
- une platine porte-objet : permet le déplacement selon trois directions, la rotation dans son plan et l'inclinaison (variation de l'angle d'incidence);
- des détecteurs d'électrons : reliés à un écran de visualisation et un système de prise de vues photographiques.

Un ensemble électronique :

- des dispositifs d'observation et d'enregistrement : tubes cathodiques à écran rémanent ou non;
- des sources de tensions continues ou variables
- des dispositifs de commande : grandissement, contraste, focalisation, correction d'astigmatisme, vitesse et type de balayage...

Un système d'analyse et de traitement des données :

- un programme d'analyse qualitative : identification des éléments chimiques détectés;
- un programme d'analyse quantitative avec ou sans étalon : calcul de concentration;
- des cartes de répartition des éléments : représentation de la localisation d'éléments par des niveaux de gris ou des couleurs différentes (à chaque couleur un niveau de concentration).

Les paramètres qui influencent la résolution des images sont les suivants :

- la tension d'accélération des électrons : plus la tension est élevée, plus la profondeur d'investigation de la poire de diffusion est grande. La tension est fixée à 20 keV dans le cadre de ce rapport;

Matériels et Méthodes

- le courant de sonde : plus il est élevé, plus le diamètre du faisceau est grand. Le courant est optimisé entre 100 et 120 μA selon les besoins de qualité d'image.
- la distance de travail : c'est la distance entre l'échantillon et la lentille objective. Plus la distance est courte, meilleure est la résolution. La plus grande profondeur de champ est obtenue à grande distance de travail. Elle est fixée à 15 mm dans le cadre de ce rapport;
- le contraste et la brillance sont optimisés aussi selon la qualité d'image requise.

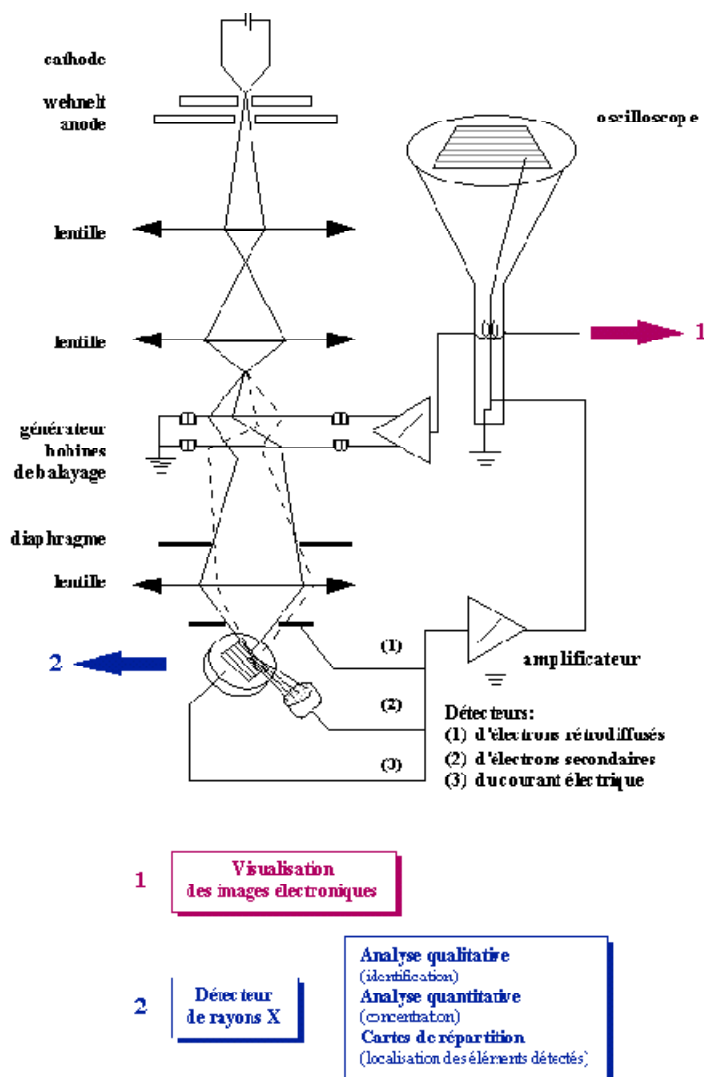


Figure 20: Représentation schématique d'un microscope électronique à balayage équipé d'un système de microanalyse de rayons X

5.12.3. Préparation des échantillons pour la microscopie électronique à balayage MEB selon (Kaletunc et al. 2004):

Le mode de préparation d'un échantillon dépend de l'information recherchée, topographie, microstructure ou composition et bien sûr de la nature de l'échantillon et de la forme dans laquelle il se présente : échantillon dur et massif, échantillon biologique, poudre, etc....

La clé d'une bonne observation ou d'une analyse correcte réside dans une préparation adéquate de son échantillon. Il ne faut pas non plus négliger les risques d'artefacts qui sont fréquents et très importants à toutes les étapes de la préparation.

L'échantillon, placé dans la chambre du microscope, reçoit un flux d'électrons très important. Si les électrons arrivant sur l'échantillon ne sont pas évacués, ils donnent lieu à des phénomènes de charge induisant des déplacements d'images ou des zébrures sur l'image attribuables à des décharges soudaines de la surface. Si la surface est conductrice, les charges électriques sont évacuées par l'intermédiaire du support de l'échantillon. L'observation d'échantillons électriquement isolants se fait grâce à un dépôt préalable (évaporation, pulvérisation cathodique) d'une fine couche conductrice d'or ou de carbone transparente aux électrons. Sans cette préparation, cet objectif est impossible à atteindre à cause d'une interprétation impossible des contrastes dans les images et des artefacts dans les résultats d'analyse.

5.12.3.1. A partir du milieu liquide :

Les souches testées sont mises en culture dans 3 ml du milieu GM17 et incubées à 30°C. 1 ml de la culture bactérienne est centrifugé à 10000 xg pendant 2 minutes. Le culot est ensuite lavé dans le tampon phosphate 50 mM (pH 7,4). 2 µl est resuspendu dans 1 ml du tampon phosphate et filtré sur des filtres de Nitrocellulose de 0,45 µm de porosité (Sartorius, Allemagne) et des grilles de MEB pour les observations en mode microscopie électronique à balayage. L'échantillon est fixé par addition de 1 ml de glutaraldehyde (3%) dans 50 mM du tampon phosphate (pH 7,4) pendant 24 heures à 4°C. L'échantillon est ensuite lavé 5 x 5min avec le tampon phosphate à température ambiante. Une post-fixation est effectuée avec 1% de osmium tetroxide dans le tampon phosphate pendant 1 heure à température ambiante. Une déshydratation est réalisée avec des concentrations croissantes de l'éthanol : 30% pendant 15 minutes, 50% pendant 20 minutes, 70% pendant 30 minutes, 90% pendant 45 minutes, 100% pendant 30 minutes (5 fois), et finalement avec 100% (éthanol absolu) pendant toute la nuit à 4°C.

5.12.3.2. A partir du milieu solide :

A partir des prés cultures, 100 µl de chaque souche est étalée sur le milieu GM17 solide et incubée à 30°C jusqu'à l'obtention d'un tapis bactérien. Les échantillons à partir du milieu solide sont découpés à des petites pièces de quelques millimètres convenablement au volume de la chambre de lecture (specimen chamber). Les échantillons sont fixés avec du glutaraldehyde (6,25%) dans 50 mM du tampon phosphate (pH 7,4) pendant 24 heures à 4°C. Les échantillons fixés sont lavés soigneusement avec du tampon phosphate 5 x 5 min, ensuite ils ont été déshydratés avec une série de concentrations croissante de l'acétone : 30%, 50%, 70%, 90%, pendant 15, 20, 30, 45 minutes respectivement. Ensuite pendant 30 minutes (5 fois) avec 100% de l'acétone et finalement à 100 % pendant toute la nuit à 4°C.

Les échantillons sont séchés dans le BAL-TEC CPD 030 (appareil de séchage par contournement du point critique), ensuite ils sont soumis au recouvrement à l'Or/Palladium dans le métalliseur par pulvérisation cathodique à l'or le sputter coater BAL-TEC SCD 005. Les images sont prises sous le microscope électronique à balayage SEM « scanning electron microscope » JEOL JSM 7500F.

5.12.3.3. Conditions opératoires

Les observations en microscopie électronique à balayage (MEB) ont été effectuées au département de microscopie électronique du Biozentrum à l'université de Würzburg. Le microscope utilisé est un JEOL JSM 7500F équipé d'un système d'analyse EDS Energy Dispersive X-ray Spectrometers. C'est un microscope d'une ultra haute résolution (1 nm) et d'un grossissement jusqu'à x 1 000 000. Les images ont été acquises à 5 kV. Les préparations microbiennes sont déshydratées et recouvertes d'une couche d'Or/Palladium puis observées. L'intérêt de ce traitement est de permettre l'observation, sans déformation structurale, d'un échantillon biologique dans son état natif.



Figure 21: JEOL 7500F

Matériels et Méthodes

5.13. Ressources bioinformatiques en ligne ou logiciels utilisés

Tableau 8 : Logiciels et programmes utilisés dans ce travail de thèse

Logiciels utilisés	Fonction	Références /site web
LAlign	Alignement de séquences	http://www.ch.embnet.org/software/LALIGN_form.html
EMBI - EMBOSS	Alignement de séquences	http://www.ebi.ac.uk/Tools/emboss/align/
Multalign	Alignement de séquences	http://multalin.toulouse.inra.fr/multalin/
DIALIGN-TX (Subramanian et al. 2008)	Alignement de séquences	dialign-tx.gobics.de/ A installer sur l'ordinateur
Clustalw (Thompson et al. 1994)	Alignement multiple	www.genome.jp/tools/clustalw/
Bioedit	Alignement multiple	http://www.mbio.ncsu.edu/BioEdit/bioedit.html
BPROM (Solovyev and Salamov 2011)	Prédiction de promoteurs bactériens	http://linux1.softberry.com/berry.phtml?topic=bprom&group=help&subgroup=gfindb
PePPER (de Jong et al. 2012)	Prédiction de promoteurs bactériens	http://pepper.molgenrug.nl/index.php/prokaryote-promoters
ARNold (Naville et al. 2011)	Prediction de terminateurs Rho-indépendent	http://rna.igmors.u-psud.fr/toolbox/arnold/
TransTermHP	Prediction de terminateurs Rho-indépendent	https://www.cbcb.umd.edu/software/transtermhp A installer sur l'ordinateur
Primer3	Dessin d'amorces	http://frodo.wi.mit.edu/primer3/
OligoCalc	Oligonucleotide properties calculator	http://www.basic.northwestern.edu/biotools/oligocalc.html
BLAST (Altschul et al. 1990)	Alignement de séquences	http://blast.ncbi.nlm.nih.gov/Blast.cgi
blastn et blastp	Comparisons de séquence	(Altschul et al. 1997)
Sequence editor (FR33.net)	Editer les séquences	http://www.fr33.net/seqedit.php
Expasy tools – Swiss prot – Protéomique	Outils bioinformatiques pour l'étude des macromolécules	http://ca.expasy.org/tools/#proteome
Clustal Omega (Sievers et al. 2011)	Alignement multiple des protéines	http://www.ebi.ac.uk/Tools/msa/clustalo/
Mfold (Zuker 2003)	prédiction de la structure secondaire des acides nucléiques	http://unafold.rna.albany.edu/?q=mfold
Psipred (Psi-blast based secondary structure prediction)	prédiction de la structure secondaire des acides nucléiques	http://bioinf.cs.ucl.ac.uk/psipred/ (McGuffin et al. 2000)
PHAST (PHAgeSearch Tool) (Zhou et al. 2011)	Annotation des phages	http://phast.wishartlab.com/
iTOL: Interactive Tree Of Life (Letunic and Bork 2007)	Construction des arbres phylogénétiques	http://itol.embl.de/
Circos (Krzywinski et al. 2009)	Visualisation circulaire des données génomiques	http://mkweb.bcgsc.ca/circos A installer sur l'ordinateur
SMART (Schultz et al. 1998)	Banque dédiée aux domaines "mobiles"	http://smart.embl-heidelberg.de/
Emboss "The European Molecular Biology Open Software Suite"	package de logiciels pour différentes analyses	http://www.ebi.ac.uk/Tools/emboss/ Rice, P. Longden, I. and Bleasby, A. Trends in Genetics 16, (6) pp276-277
Sixpack	Affiche les séquences d'ADN dans les six cadres de lecture	http://www.ebi.ac.uk/Tools/st/emboss_sixpack/

Matériels et Méthodes

Expasy « the Expert Protein Analysis System »	site de collection de plusieurs logiciels pour l'analyse manuelle des protéines	http://www.expasy.org/
SEED-viewer (Overbeek et al. 2007)	Visualisation de données génomiques annotées	rast.nmpdr.org/seedviewer.cgi
FastQC	Contrôler la qualité des séquences	(http://www.bioinformatics.babraham.ac.uk/projects/fastqc).
FLASH (7) “fast length adjustment of short reads” (Magoc and Salzberg 2011)	Fusionner les lectures paire-ends chevauchantes	http://sourceforge.net/projects/flashpage/?source=typ_redirect
cutadapt (Martin M. 2011)	Éliminer les séquences adaptatrices des lectures issues du séquençage	http://code.google.com/p/cutadapt/
Jellyfish (Marcais and Kingsford 2011)	Comptage des <i>k</i> -mers	http://www.cbc.umd.edu/software/jellyfish/ à télécharger et installer
QUAKE (Kelley et al. 2010)	Package de logiciels pour détecter et corriger les erreurs de substitutions issus du séquençage	http://www.cbc.umd.edu/software/quake/
Velvet (Zerbino and Birney 2008)	Assemblage des génomes	https://www.ebi.ac.uk/~zerbino/velvet/
IDBA-UD (Peng et al. 2012)	Assemblage des génomes	http://www.cs.hku.hk/~alse/idba_ud
RAST 4 (Overbeek et al. 2014)	Annotation fonctionnelle des génomes	http://rast.nmpdr.org/
Infernal/Rfam (Nawrocki and Eddy 2013)	Détecter les RNA annotés non-codons	http://infernal.janelia.org .
PhyML (Crisuolo 2011)	Construction des arbres phylogénétiques par la méthode maximum likelihood	http://mobylye.genouest.org/cgi-bin/Mobylye/portal.py#forms::morePhyML
Seaview (Galtier et al. 1996)	Visualisation des arbres phylogénétiques	http://pbil.univ-lyon1.fr/software/seaview3.html
Splitstree 4 (Huson and Bryant 2006)	Construction et visualisation des arbres phylogénétique	http://www.splitstree.org/
MEGAN (version 5, (Huson et al. 2007)	Analyse métagénomique	http://www-ab.informatik.uni-tuebingen.de/software/megan5/
DIALIGN (Morgenstern 1999; Al Ait et al. 2013)	Alignement multiple de séquences	http://dialign.gobics.de/
PhyML (version 3.1, (Guindon and Gascuel 2003)	Construction des arbres phylogénétiques par la méthode maximum likelihood	http://www.atgc-montpellier.fr/phyml/
Mira4 (Chevreux et al. 2004)	Assemblage	http://sourceforge.net/projects/mira-assembler/
KEGG “Kyoto Encyclopedia of Genes and Genomes “	Visualisation des différentes voies métaboliques	www.genome.jp/kegg/
AMOS « A Modular Open-Source Assembler » (Treangen et al. 2011)	Assemblage	ftp://ftp.tigr.org/pub/software/AMOS
Ray Meta (Boisvert et al. 2012)	Assemblage métagénomique	http://denovoassembler.sf.net . http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4056372/
Metavelvet (Namiki et al. 2012)	Assemblage métagénomique	http://metavelvet.dna.bio.keio.ac.jp/ à installer
Omega (Haider et al. 2014)	Assemblage	http://omega.omicsbio.org .

Matériels et Méthodes

SPAdes (version 3.5.0, (Bankevich et al. 2012)	Assemblage- <i>Scaffolding</i>	http://bioinf.spbau.ru/spades à installer
Prokka (Seemann 2014)	Annotation des génomes bactériens	http://www.vicbioinformatics.com/software/prokka.shtml à installer
CONTIGuator (Galardini et al. 2011)	<i>Scaffolding</i>	http://contiguator.sourceforge.net/ A installer sur l'ordinateur
Scaffolder_Builder (Silva et al. 2013)	<i>Scaffolding</i>	http://edwards.sdsu.edu/scaffold_builder/ à installer sur l'ordinateur
SSPACE-LongRead (stand-alone program for scaffolding pre-assembled contigs using long reads) (Boetzer and Pirovano 2014)	Assemblage- <i>Scaffolding</i>	http://www.baseclear.com/genomics/bioinformatics/basetools/SSPACE-longread
SOAPdenovo (Short Oligonucleotide Analysis Package (Luo et al. 2012b)	Assemblage	http://sourceforge.net/projects/soapdenovo2/files/SOAPdenovo2/

CHAPITRE 3
RESULTATS ET DISCUSSION

Rapport-Gratuit.com

Chapitre 3 : Résultats et Discussion

Partie I : Diversité naturelle des lactocoques dans le lait de chamelle d'Algérie

1. Introduction

Les bactéries d'acide lactique (LAB) reconnus généralement comme des organismes (*GRAS*) sont impliqués dans la fabrication d'une grande variété d'aliments fermentés. Les souches LAB produisent des exopolysaccharides (Nyyssola et al. 2005), des agents antibactériens (Arques et al. 2015), des vitamines (Wegkamp et al. 2007) et des métabolites de haute valeur impliqués dans la préservation de la saveur, de la texture, et dans le développement et la préservation des aliments (Pontes et al. 2011). Diverses espèces de LAB agissent comme des probiotiques (Marco et al. 2010) et comme des vecteurs potentielles de composés ayant des effets thérapeutiques ou prophylactiques (Amar et al. 2011). Étant l'un des organismes modèle en métabolisme microbien, *Lactococcus lactis* a été la principale cible de l'ingénierie métabolique chez LAB (Papagianni 2012), il a révélé de nouvelles applications biomédicales comme adjuvants, immunostimulants ou comme des systèmes thérapeutique livrant des enzymes digestives pour compléter l'insuffisance pancréatique chez l'homme ; livraison de l'ADN de plasmide aux cellules eucaryotes ou comme usines à produire des molécules thérapeutiques et plus récemment l'utilisation potentielle de *L. lactis* comme *Gram-positive enhancer matrix* (GEM), comme un système de diffusion des vaccins de la muqueuse (Wyszynska et al. 2015). *Lactococcus lactis*, est une usine cellulaire efficace pour la Production de protéines recombinantes et la sécrétion de diverses applications biotechnologiques (Morello et al. 2008).

De nouvelles sources alimentaires devraient être exploitées pour faire varier l'alimentation humaine et également de bénéficier de nouveaux ingrédients fonctionnels et d'aliments naturels. Les pays d'Afrique et arabes, où les conditions d'élevage de vaches sont sévères et exigeantes, peuvent franchir cette situation en développant un système de reproduction pour les animaux locaux tels que les chameaux.

La microflore bénéfique du lait de chamelle représentée par LAB est une source potentielle de matériaux biologiques utilisés en technologie laitière (Khedid et al. 2009). A ce jour, le lait cru de chamelle a été négligé comme une source alimentaire potentielle de qualité et n'a pas eu un grand intérêt industriel. Les études sur les produits issus du lait de chamelle sont plutôt

Résultats et discussion

rare, en dépit de leur intérêt nutritionnel. Dans ce contexte, le lait de chamelle doit être étudié d'une manière approfondie. Il y a seulement quelques références sur le lait de chamelle, elles concernent la production (Faye et al. 2005) ou sa composition (Farah 1993). Le lait de chamelle reste encore une source importante de protéines pour les personnes vivant dans les zones arides dans le monde. Ainsi, le lait de chamelle est connu pour ses propriétés médicinales, qui sont largement exploitées pour la santé humaine, comme dans plusieurs pays de l'union ex-Sovietique (Konuspayeva et al. 2009), les pays en développement (Al Katheeri et al. 2006), les zones arides d'Asie centrale (Faye and Konuspayeva 2012; Baubekova et al. 2015) et dans certaines sociétés pastorales arabes (Benmechernene et al. 2014). Le lait de chamelle est considéré d'avoir des propriétés anticancéreuses (Magjeed 2005), anti-allergiques (Shabo et al. 2005) et antidiabétique (Agrawal et al. 2003). Sa teneur élevée en acides gras insaturés contribue à sa qualité alimentaire globale (Konuspayeva et al. 2008). La faible quantité de β -caséine et l'absence de β -lactoglobuline sont liées aux effets hypoallergéniques du lait de chamelle. D'autres composants tels que les immunoglobulines, les lactoferrine, lysozyme ou la vitamine C ont été signalés à jouer un rôle central dans la détermination de ces caractéristiques (Mohamed et al. 2005; Jang et al. 2006; Konuspayeva et al. 2007). En outre, le lait de chamelle est connu pour sa longue durée de vie, ce qui permet son stockage et sa consommation sans danger après plusieurs jours en l'absence de réfrigération. Pour toutes ces raisons, le lait de chamelle et ses produits dérivés sont très populaires en Algérie surtout dans les régions rurales arides. Toutefois, des productions à grande échelle basées sur le lait de chamelle sont assez rares. Les chameaux sont considérés comme des animaux laitiers non conventionnels dans le « *Codex alimentarius* ».

L'identification de la microflore du lait de chamelle fournit une base théorique pour le développement des ferments en utilisant des souches locales pour la production industrielle des produits laitiers traditionnellement fermentés et la majorité des études scientifiques sur les chameaux ont été principalement axée sur leurs caractéristiques anatomiques et sur leur adaptation physiologique aux climats néfastes.

Par conséquent, les informations sur la microflore du lait de chamelle sont très limitées. Des études antérieures sur la caractérisation moléculaire des LAB isolés à partir du lait de chamelle fermenté ont été signalés dans la région du Xinjiang en Chine (Zuo et al. 2014), dans les zones arides d'Asie centrale (Baubekova et al. 2015), à partir du lait cru de chamelle en Algérie (Drici et al. 2010) et à partir de boissons laitiers fermentés « *Leben* » ou « *Raib* » dans

Résultats et discussion

les régions arides en Algérie (Hassaine et al. 2007; Bensalah et al. 2009; Bendimerad et al. 2012).

2. Etude de la diversité des isolats

Les isolats ont été caractérisés par des méthodes phénotypiques et génotypiques.

2.1. Caractéristiques morphologiques

2.1.1. Caractéristiques macroscopiques

Toutes les souches étudiées présentent les caractéristiques morphologiques des Lactocoques. Elles forment sur M17 des colonies uniformes pour chaque souche et identiques dans leurs aspects ; des petites colonies rondes transparentes d'aspect lisse et légèrement bombées (figure 22).



Figure 22 : Aspect macroscopique des colonies de lactocoque
Après 18 heures de culture sur M17 à 30°C.

2.1.2. Caractéristiques microscopiques

L'aspect microscopique des isolats est visualisé après coloration de Gram : Les isolats présentent des formes cocci généralement isolés à Gram-positif (figure 23).

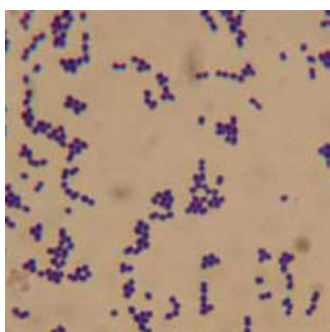


Figure 23 : Caractéristiques microscopiques des isolats étudiés après coloration de Gram.

[Grossissement x100]

2.2. Test de la thermotolérance

Après l'incubation, on remarque que certaines souches ont pu se développer à 45°C, et même à 50°C mais le résultat était négatif pour la totalité des souches à 60 et 65°C.

L'histogramme ci-dessous présente les résultats obtenus dans les trois tests successifs dans le milieu liquide et solide, spécifiques pour chaque série :

Résultats et discussion

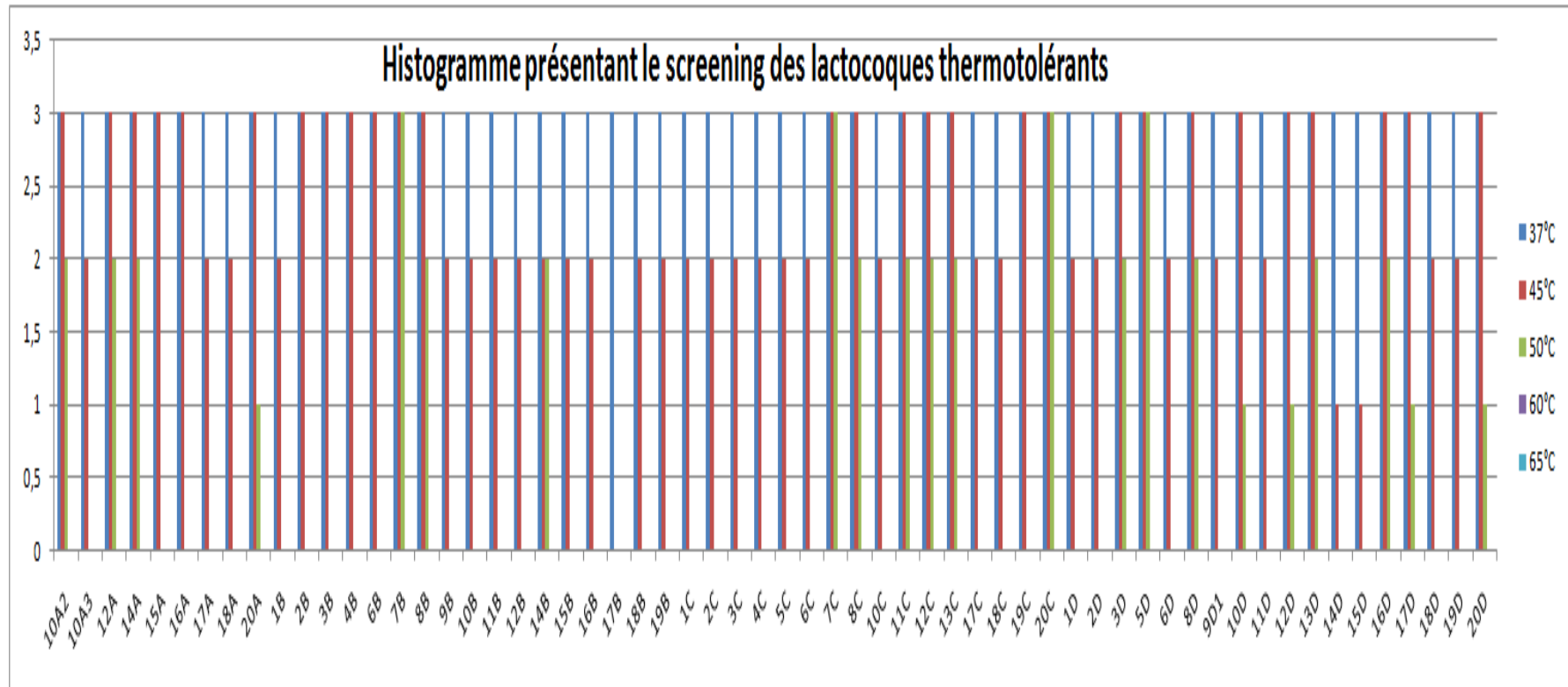


Figure 24 : Histogramme présentant le screening des lactocoques thermotolérants

Résultats et discussion

Résultats et discussion

A l'exception d'une publication qui révèle le caractère thermotolérant de lactocoques issus du leben d'Algérie des zones arides (Bensalah et al. 2009), tous les lactocoques jusqu'ici identifiées et caractérisées sont des mésophiles c'est-à-dire ils sont capables de se développer entre 10°C et 30°C, avec 39 à 40°C comme température maximale de croissance de (Ludwig et al. 1985). La majorité de nos lactocoques issus du lait cru de chamelle d'Algérie testées ont pu tolérer des températures plus élevées allant de 45 à 50°C. Ce caractère est peut être lié à des systèmes de résistance au stress thermique qui est le plus souvent associé à la synthèse de protéines spécifiques appelées les protéines des choc thermique. En effet chez *Lactococcus lactis* 13 protéines HSP ont été identifiées (Richard *et al.*, 1991). Ce nombre de protéines est beaucoup plus élevé chez *Enterococcus faecalis* (34 HSP) (Flahaut et al. 1996) et *Streptococcus mutans* (40HSP) (Svensater et al. 2000).

Le caractère original de la thermotolérance semble être lié à la nature aride de l'écosystème de ce fait nos résultat rejoignent celles décrite par (Bensalah et al. 2009). Cette thermotolérance est un caractère atypique pour les lactocoques.

Les 60 souches testées ainsi que le témoin ne présentent aucune thermotolérance vis-à-vis les températures de 60 et 65°C qui en effet, se rapprochent des températures de pasteurisation (Jones and Dawson 1978; Garvie et al. 1984; Ludwig et al. 1985). Cela nous a permis de faire un classement des souches selon leur intervalle de thermotolérance :

Les souches capables de se développer à 45°C représentent 63% de l'ensemble des souches testées 10A₃, 15A, 16A, 17A, 18A, 1B, 2B, 3B, 4B, 6B, 7B, 10B, 11B, 12B, 16B, 17B, 18B, 19B, 1C, 2C, 3C, 4C, 5C, 7C, 10C, 17C, 18C, 19C, 1D, 9D₁, 10D, 12D, 15D, 17D, 18D, 19D.

Les lactocoques qui apparaissent plus intéressantes, sont en nombre de 22 :

10A₂, 12A, 14A, 20A, 8B, 9B, 14B, 15B, 6C, 8C, 11C, 12C, 13C, 20C, 3D, 5D, 8D, 11D, 13D, 14D, 16D, 20D. Ces lactocoques sont capables de se développer à 50°C.

L'ensemble de ces lactocoques thermotolérant fera l'objet d'une étude qualitative et quantitative selon des approches physiologiques et génétiques.

Résultats et discussion

3. Analyse physiologique

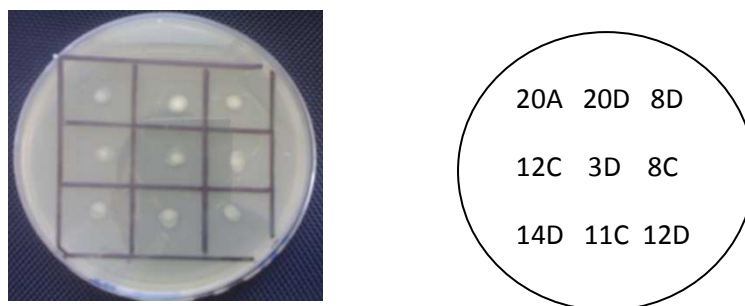


Figure 25: Lactocoques ensemencés par la méthode des patchs

Boite 1 : contenant les 12 souches : de 10A₂ à 3B..... (1-12)

Boite 2 : contenant les 12 souches de 4B à 17 B.... (13-24)

Boite 3 : contenant les 12 souches de 18B à 11C (25-36)

Boite 4 : contenant les 12 souches de 12 C à 8D (37-48)

Boite 5 : contenant les 12 souches de 9D₁ à 20D..... (49-60)

Le lait utilisé était : le lait de vache totalement écrémé « silhouette ».

3.1. Test de Protéolyse en milieu solide Agar-lait

Un exemple de résultat est montré sur la (figure 26). Toutes les souches testées sont protéolytiques à 50°. En effet autour de chaque touche nous remarquons la présence de halos transparents qui traduisent l'hydrolyse des caséines du lait. Ce résultat indique qu'une incubation à 50°C n'affecte pas l'activité protéolytique des lactocoques thermotolérants.

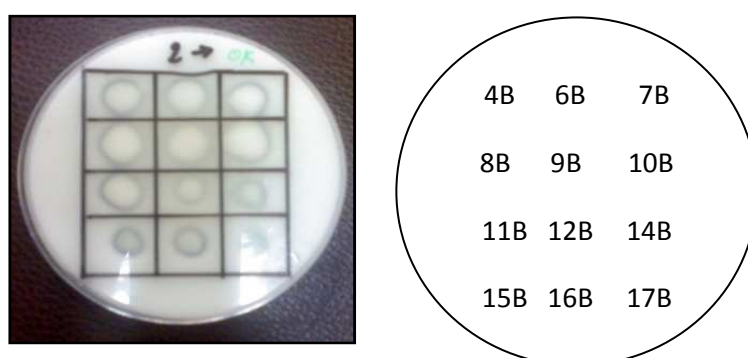


Figure 26: Résultats de la protéolyse

L'activité protéolytique à 50°C suggère que la protéinase de parois PrtP de ces lactocoques reste active en préservant son affinité et son activité catalytique pour les caséines du lait. (Ecroyd et al. 2010; Luan et al. 2010; Zhang et al. 2010; Zhang et al. 2011). La croissance des

Résultats et discussion

souches testées sur milieu Agar-lait à 50°C, nous permet d'envisager ultérieurement une étude de la cinétique de croissance en milieu lait liquide.

3.2. Screening des lactocoques thermotolérants

Les souches 9B, 13C présentent une croissance qui semble la plus rapide comparée aux autres souches thermotolérantes. Ces deux souches sont de bonnes candidates pour une étude cinétique de croissance plus détaillée en présence du témoin 3MD pendant 8h d'incubation dans deux milieux différents et deux températures différentes.

Résultats et discussion

Tableau 9 : Précriblage de souches thermotolérantes à croissance rapide apparente

Les souches	DO à $T_0 = 0$ h	DO à $T_1 = 4$ h	DO à $T_2 = 6$ h
10A2	0.059	0.309	0.664
12A	0.062	0.390	0.550
14A	0.051	0.374	0.528
20A	0.049	0.314	0.633
8B	0.05	0.239	0.302
9B	0.060	0.450	0.818
14B	0.051	0.247	0.556
15B	0.05	0.270	0.603
6C	0.048	0.253	0.517
8C	0.052	0.275	0.575
11C	0.05	0.355	0.639
12C	0.056	0.333	0.656
13C	0.063	0.585	1.065
20C	0.051	0.208	0.630
3D	0.047	0.205	0.513
5D	0.05	0.223	0.548
8D	0.061	0.306	0.621
11D	0.05	0.207	0.631
13D	0.05	0.300	0.629
14D	0.046	0.316	0.540
16D	-	-	-
20D	0.05	0.263	0.549

Résultats et discussion

3.3. Etude de la cinétique de croissance en milieu M17 (à 30°C et 50°C)

Les résultats de la densité optique relevées à 600 nm pour les 3 lactocoques thermotolérants testés « 9B, 13C, 3MD » sont rassemblés dans les tableaux 2, 3, 4,5. Les résultats surlignés en jaune correspondent à la moyenne des 3 tests successifs de prise de DO à différent temps d'incubation. L'ensemble des résultats des cinétiques de croissance en milieu M17 sont représentés sous forme de graphe (Figure 27).

Tableau 10: Résultat de la cinétique de croissance en milieu M17 (30°C)

souche	9B				13C				3MD			
	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne
T ₀ =0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
T ₁ =1	0.552	0.436	0.610	0.532	0.713	0.459	0.693	0.621	0.267	0.372	0.225	0.288
T ₂ =1.5	0.704	0.531	0.721	0.652	0.683	0.735	0.744	0.720	0.298	0.465	0.327	0.363
T ₃ =2	0.782	0.742	0.843	0.789	0.890	0.653	0.733	0.758	0.393	0.621	0.388	0.467
T ₄ =2.5	0.941	0.828	0.857	0.875	0.963	0.760	0.954	0.892	0.388	0.631	0.496	0.505
T ₅ =3	1.023	0.885	0.979	0.962	0.929	0.840	0.965	0.911	0.508	0.769	0.720	0.665
T ₆ =3.5	1.099	0.948	0.999	1.015	0.987	0.928	1.055	0.990	0.642	0.878	0.762	0.760
T ₇ =4	1.166	1.002	1.051	1.073	1.003	0.984	1.086	1.024	0.726	0.936	0.770	0.810
T ₈ =4.5	1.091	1.141	0.943	1.058	1.035	1.034	1.072	1.047	0.919	0.981	0.921	0.940
T ₉ =5	0.973	1.169	1.118	1.086	1.056	1.031	1.013	1.033	1.016	1.019	1.004	1.013
T ₁₀ =5.5	0.971	1.243	1.123	1.112	1.089	1.101	1.079	1.089	1.066	0.129	1.104	1.086
T ₁₁ =6	0.973	1.172	1.245	1.130	1.101	1.131	1.081	1.104	1.079	1.133	1.149	1.120
T ₁₂ =6.5	0.796	1.280	1.156	1.053	1.099	1.119	1.079	1.099	1.172	1.149	1.217	1.179
T ₁₃ =7	1.063	1.164	1.211	1.146	1.103	1.116	1.128	0.782	1.187	1.158	1.239	1.194
T ₁₄ =7.5	1.055	1.198	1.164	1.125	1.169	1.167	1.057	1.131	1.128	1.136	1.305	1.189
T ₁₅ =8	1.152	1.117	1.051	1.106	1.141	1.125	1.185	1.150	1.173	1.171	1.277	1.207

Résultats et discussion

Tableau 11 : Résultat de la cinétique de croissance en milieu M17 (50°C)

souche	9B				13C				3MD			
DO	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne
T ₀ =0	0.093	0.051	0.061	0.068	0.055	0.056	0.081	0.064	0.135	0.052	0.032	0.073
T ₁ =1	0.099	0.099	0.155	0.117	0.091	0.085	0.147	0.107	0.141	0.080	0.045	0.088
T ₂ =1.5	0.076	0.157	0.259	0.164	0.101	0.182	0.172	0.151	0.166	0.080	0.085	0.110
T ₃ =2	0.086	0.240	0.524	0.283	0.087	0.350	0.272	0.236	0.132	0.107	0.139	0.126
T ₄ =2.5	0.109	0.421	0.867	0.452	0.048	0.405	0.420	0.291	0.100	0.172	0.190	0.154
T ₅ =3	0.151	0.788	1.157	0.698	0.294	0.550	0.542	0.462	0.131	0.271	0.307	0.236
T ₆ =3.5	0.111	1.048	1.334	0.831	0.080	0.862	0.674	0.538	0.206	0.483	0.515	0.401
T ₇ =4	0.078	1.275	1.350	0.901	0.139	0.984	0.861	0.661	0.396	0.763	0.780	0.646
T ₈ =4.5	0.159	1.328	1.383	0.956	0.116	1.104	1.005	0.741	0.425	0.836	0.931	0.730
T ₉ =5	0.164	1.066	1.131	0.787	0.173	0.898	0.828	0.633	0.626	0.896	0.862	0.794
T ₁₀ =5.5	0.183	1.103	1.222	0.836	0.200	1.015	0.864	0.693	0.898	0.876	0.774	0.849
T ₁₁ =6	0.162	1.173	1.232	0.855	0.157	0.928	0.886	0.657	0.809	0.828	0.820	0.819
T ₁₂ =6.5	0.211	1.158	1.302	0.820	0.184	0.905	0.908	0.665	0.891	0.862	0.813	0.855
T ₁₃ =7	0.397	1.147	1.244	0.929	0.289	1.029	0.930	0.749	0.974	0.880	0.897	0.917
T ₁₄ =7.5	0.544	1.176	1.250	0.990	0.381	0.969	0.918	0.759	0.961	0.924	0.881	0.922
T ₁₅ =8	0.684	1.159	1.286	1.043	0.705	1.037	0.954	0.774	1.025	0.925	0.906	0.952

Résultats et discussion

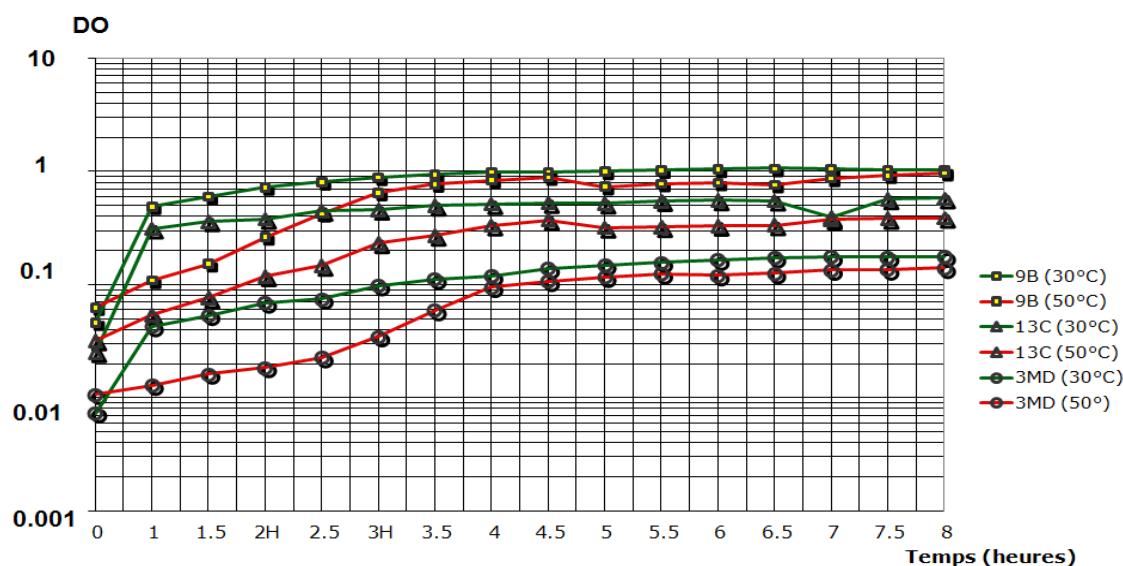


Figure 27 : Comparaison des cinétiques de croissance à 30 et 50°C en milieu M17 des Lactocoques issus du lait cru de chamelle ou du lait de chèvre

Les courbes qui décrivent les cinétiques de croissance des trois lactocoques en milieu M17 « Figure 27 » ne reflètent pas l'allure classique des courbes de croissance des lactocoques déjà étudiées où les différentes phases sont détectées (De Roissart, 1994). Seules les phases exponentielles et stationnaires apparaissent pour les trois souches étudiées (9B, 13C et 3MD). La température de croissance semble influencer la durée des phases exponentielles. En effet, à 30°C les phases exponentielles des trois souches sont très courtes et durent seulement 1 h. Alors qu'à 50°C deux comportements de croissance sont remarqués en fonction des souches des lactocoques étudiées :

- Lactocoques 9B et 13C présentent le même aspect de courbe, vrai semblablement la souche 9B est la plus rapide.
- Lactocoques 3MD présentent une allure différente avec la présence de deux phases exponentielles dont le temps de génération de la 2^{ème} est court.

En conclusion, nous relevons que la croissance de ces trois lactocoques testés en milieu M17 est sous l'influence d'un effet température et origine de souche.

3.4. Etude de la cinétique de croissance en milieu lait (à 30°C et 50°C)

La croissance en milieu lait pour les lactocoques 9B et 13C issus du lait cru de chamelle présentent des courbes en allure de scie, ce type de cinétique rappelle la croissance en milieu

Résultats et discussion

renouvelé. Hors, dans notre cas il s'agit de cinétiques de croissance discontinue. Cette allure de courbe peut être expliquée par le phénomène de diauxie qui est un indice de l'utilisation séquentielle des nutriments disponibles dans un milieu de culture. Nous remarquons que le nombre de phases exponentielles et celui des phases diauxiques varie selon la souche et la température. Par exemple la souche 13C présente 4 phases exponentielles alternées par 3 phases diauxiques à 50°C. Alors que la souche 9B présente seulement une phase diauxique et une succession de 4 phases exponentielles. De plus, c'est la seule souche qui a présenté une phase d'accélération qui est en fait une phase préparative de la phase exponentielle. Il faut noter que le lactocoque 9B se développe mieux à 50°C comparé à 30°C ceci nous conduit à revoir la température optimale de croissance pour cette souche. Il est établi dans la littérature que la cinétique de croissance des lactocoques en milieu lait présente seulement deux phases exponentielles et une seule phase diauxique. Ce cas de figure de cinétique de croissance en milieu lait est original pour ces lactocoques thermotolérants et il peut être expliqué par la présence d'équipement enzymatique propre à ces lactocoques capables d'assimiler d'autres nutriments disponibles dans le lait que les autres lactocoques déjà étudiés ne peuvent pas utiliser.

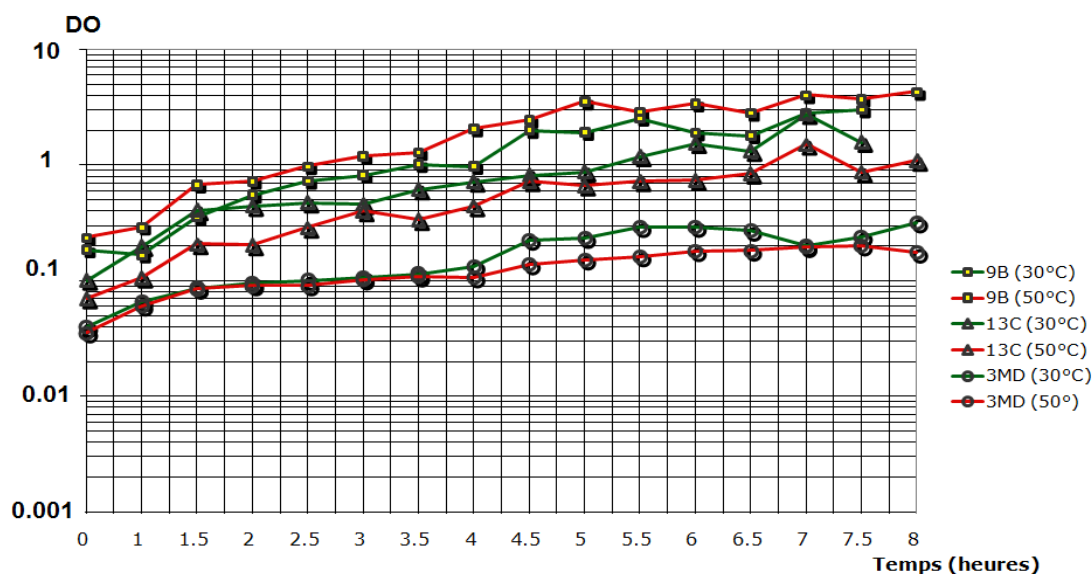


Figure 28 : Comparaison des cinétiques de croissance à 30 et 50°C en milieu lait des lactocoques issus du lait cru de chamelle ou du lait de chèvre

Résultats et discussion

Tableau 12 : Résultat de la cinétique de croissance en milieu lait (30°C)

souche	9B				13C				3MD			
DO	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne
T ₀ =0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
T ₁ =1	0.09	0.032	0.015	0.045	0.090	0.091	0.113	0.098	0.054	0.026	0.051	0.044
T ₂ =1.5	0.041	0.009	0.235	0.095	0.303	0.178	0.120	0.200	0.152	0.010	0.076	0.079
T ₃ =2	0.159	0.120	0.165	0.148	0.305	0.209	0.140	0.218	0.035	0.089	0.058	0.060
T ₄ =2.5	0.174	0.264	0.160	0.196	0.283	0.211	0.205	0.233	0.072	0.113	0.158	0.114
T ₅ =3	0.244	0.215	0.202	0.221	0.330	0.236	0.113	0.226	0.136	0.191	0.168	0.165
T ₆ =3.5	0.308	0.321	0.292	0.307	0.311	0.316	0.283	0.303	0.180	0.196	0.211	0.195
T ₇ =4	0.240	0.257	0.293	0.263	0.460	0.295	0.306	0.353	0.133	0.167	0.224	0.176
T ₈ =4.5	0.390	0.579	0.647	0.538	0.397	0.535	0.286	0.406	0.502	0.389	0.257	0.382
T ₉ =5	0.514	0.461	0.579	0.518	0.393	0.600	0.311	0.434	0.443	0.286	0.456	0.396
T ₁₀ =5.5	0.588	0.523	0.956	0.689	0.444	0.522	0.796	0.587	0.555	0.375	0.549	0.493
T ₁₁ =6	0.749	0.359	0.432	0.513	0.949	0.617	0.723	0.763	0.542	0.396	0.549	0.495
T ₁₂ =6.5	0.491	0.522	0.446	0.486	1.229	1.339	1.393	0.653	0.476	0.287	0.612	0.458
T ₁₃ =7	0.436	0.460	1.370	0.755	1.597	1.840	0.640	1.365	0.426	0.238	0.325	0.329
T ₁₄ =7.5	1.195	0.495	0.747	0.812	0.609	0.577	1.160	0.782	0.328	0.307	0.389	0.341
T ₁₅ =8	-	-	-	-	-	-	-	-	-	-	-	-

On n'a pas pu lire les DO qui correspondent au temps T₁₅= 8h, car le coagulum était très intense.

Résultats et discussion

Tableau 13 : Résultat de la cinétique de croissance en milieu lait (50°C)

souche	9B				13C				3MD			
DO	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne	Test 1	Test 2	Test 3	moyenne
T ₀ =0	0.096	0.041	0.056	0.064	0.048	0.026	0.031	0.035	0.057	0.052	0.072	0.060
T ₁ =1	0.073	0.056	0.108	0.079	0.090	0.031	0.038	0.053	0.108	0.165	0.083	0.118
T ₂ =1.5	0.148	0.187	0.214	0.183	0.200	0.034	0.076	0.103	0.119	0.175	0.139	0.144
T ₃ =2	0.038	0.211	0.268	0.197	0.088	0.102	0.114	0.101	0.130	0.183	0.151	0.154
T ₄ =2.5	0.140	0.234	0.418	0.264	0.127	0.137	0.165	0.143	0.153	0.203	0.159	0.171
T ₅ =3	0.063	0.362	0.555	0.326	0.240	0.180	0.177	0.199	0.058	0.273	0.165	0.165
T ₆ =3.5	0.035	0.503	0.512	0.350	0.109	0.225	0.169	0.167	0.099	0.278	0.176	0.184
T ₇ =4	0.112	0.904	0.672	0.562	0.071	0.290	0.300	0.220	0.236	0.278	0.182	0.232
T ₈ =4.5	0.044	1.090	0.874	0.669	0.087	0.490	0.514	0.363	0.167	0.310	0.229	0.235
T ₉ =5	0.012	1.326	1.556	0.964	0.083	0.400	0.513	0.332	0.174	0.314	0.291	0.259
T ₁₀ =5.5	0.104	1.042	1.213	0.786	0.102	0.345	0.515	0.363	0.194	0.341	0.293	0.276
T ₁₁ =6	0.085	1.233	1.464	0.927	0.120	0.418	0.572	0.370	0.322	0.346	0.252	0.306
T ₁₂ =6.5	0.103	0.918	1.282	0.767	0.186	0.450	0.629	0.421	0.295	0.356	0.276	0.309
T ₁₃ =7	0.298	1.343	1.635	1.092	0.132	0.482	0.632	0.752	0.246	0.364	0.395	0.332
T ₁₄ =7.5	0.638	1.173	1.240	1.017	0.127	0.555	0.622	0.434	0.381	0.401	0.452	0.411
T ₁₅ =8	0.885	1.168	1.454	1.169	0.317	0.703	0.609	0.543	0.595	0.509	0.502	0.535

Seule le lactocoque 3MD se développe de la même manière à 30 et à 50°C du temps 0 jusqu'au temps d'incubation de 3 heures et 30 minutes, où une phase diauxique très longue de 2 heures est observée. C'est uniquement à 30°C qu'on remarque la présence d'une phase d'accélération suivie d'une phase exponentielle courte « 1 heure et demi ».

Résultats et discussion

Tableau 14 : Récapitulation des données de la cinétique

souches	G « temps de génération » en heure			
	UFC		DO	
	30°C	50°C	30°C	50°C
9B	1	2.17	1	2.5
13C	0.91	3.33	1.09	3.44
3MD	1.31	2.43	1.31	2.43

Le calcul du temps de génération « G » des trois lactocoques testés est réalisé, en utilisant les valeurs de la densité optique « DO_{600nm} » comme moyen de quantification de la croissance bactérienne en milieu liquide et UFC en milieu solide.

On remarque une légère différence entre les valeurs des G calculés à partir de la DO ou d'UFC pour les lactocoques issus du lait cru de chamelle. Ceci est peut-être dû à la limite de la technique turbidimétrique qui prend en considération les cellules mortes et les cellules vivantes. Par contre la méthode d'UFC ne prend en considération que les cellules vivantes.

Il est à noter que la souche 3MD a présenté les mêmes valeurs de « G » quelque soit la technique utilisée (méthode turbidimétrique ou dénombrement sur milieu solide). Ce qui renseigne sur un mécanisme de résistance différent qui permet le maintien de la viabilité du lactocoque 3MD issu du lait cru de chèvre comparé aux lactocoques issus du lait cru de chamelle.

Il nous a été impossible d'exploiter mathématiquement ou sur courbe les cinétiques de croissance des trois lactocoques en milieu lait car il est très délicat à ce stade de choisir la phase exponentielle correcte qui permet le calcul des paramètres de croissance crédible.

4. Analyse génétique

4. 1. Test de thermotolérance des variants

On remarque que les variants ont perdu leur capacité à se développer à 50°C, alors que leur développement était normal à 30°C. Ce qui suppose que le caractère de la thermotolérance n'est pas stable car il est perdu après 55 générations. Ce résultat nous permet d'émettre l'hypothèse que le caractère de thermotolérance est peut être porté soit par un ou plusieurs plasmides ou encore localisé sur un transposon conjugatif, sachant que le déterminant génétique des protéines du choc thermique chez les lactocoques est chromosomique (Bolotin et al. 2001).

Ces résultats nous orientent vers l'exploration des protéines du choc thermique HSP chez ces lactocoques pour les comparer à celle déjà décrites et caractérisées. Ce résultat nous incite aussi à déterminer la structure et le mode de régulation des protéines qui s'expriment à 50°C. Ces nouvelles molécules semblent jouer un rôle crucial dans le maintien de l'intégralité des lactocoques thermotolérants.

4. 2. Test d'activité protéolytique et d'hydrolyse du lactose

Les variants obtenus sont également capables d'utiliser les caséines du lait contenues à 10% dans le milieu GMAB, et de ce fait ces variants sont aussi de phénotype Prt⁺.

Les caractères : activité protéolytique et fermentation du lactose ne sont pas perdues chez les variants non thermotolérants obtenus. Il ressort de ce résultat que le phénotype thermotolérance « Therm^T » n'est pas lié au phénotype Prt⁺ Lac⁺, Cependant ces deux derniers caractères sont connus dans la littérature être associés à la présence de plasmides (Siezen et al. 2005; Smid et al. 2005). Le maintien des caractères Prt⁺ Lac⁺ chez nos trois lactocoques testées 9B, 13C, 3MD peut être expliqué soit par un déterminisme génétique chromosomique dont les gènes codent pour la protéinase de paroi et les enzymes impliquées dans la dégradation de lactose ou ces déterminants sont peut être portées sur un grand plasmide dont la perte est difficile.

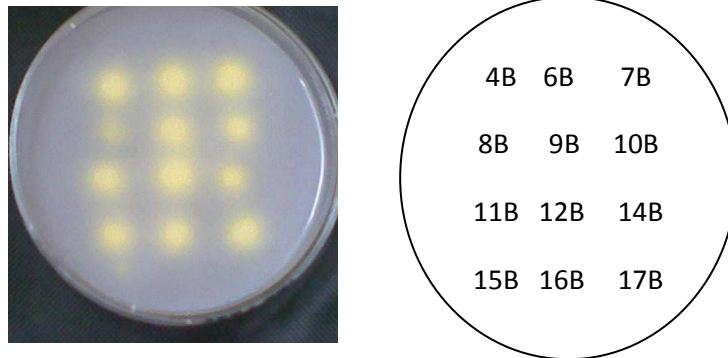


Figure 29: Résultat du test de l'activité protéolytique de souches testées en milieu Lac-MAB

4.3. Expression des gènes de choc thermique par RT-PCR

Les sept gènes étudiés (cinq chaperonnes et deux protéases), sont exprimés en RT-PCR qualitative dans les conditions optimale et stressées. L'expression de ces gènes varie lors de l'adaptation à la température élevée, les chaperonnes sont induites et clp protéase est réprimée. Par analyse du contexte génomique, nous avons recherché les régulateurs qui pourraient contrôler par ces gènes. J'ai basé l'analyse sur (i) la proximité génétique et notamment les gènes transcrits en opéron et (ii) le type de fonctions habituellement contrôlées par le régulateur considéré.

Résultats et discussion

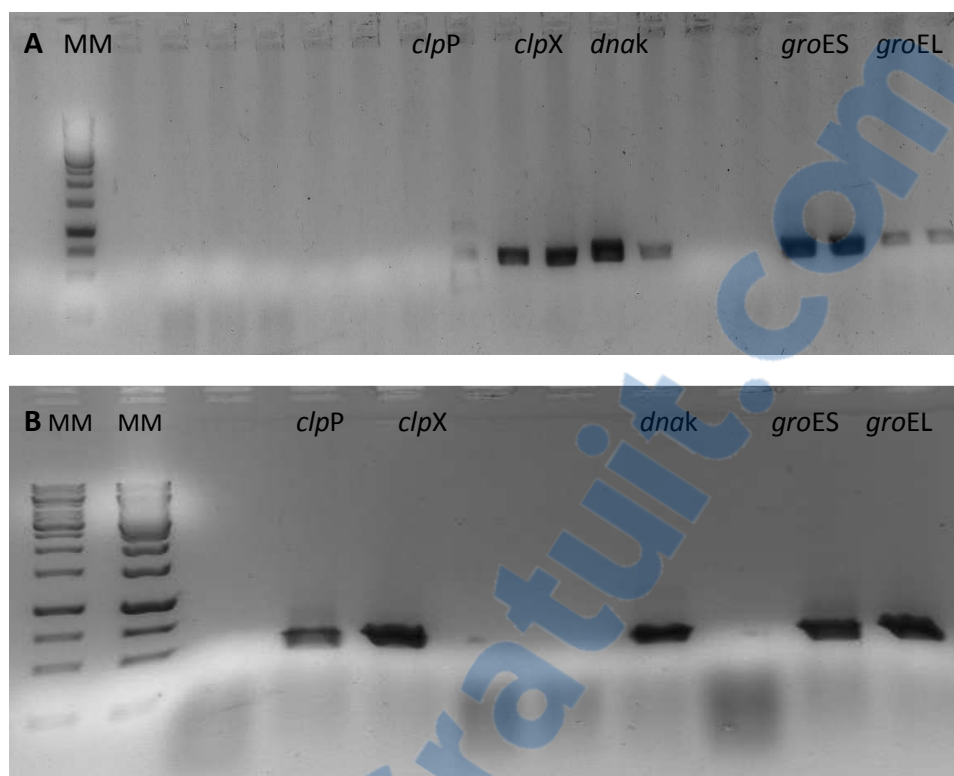


Figure 30 : Expression des principaux gènes de choc thermique appartenant à différents régulons. La souche HD9B incubée à 30 (A) et à 50°C (B). Les RNAs totaux sont extraits et l'expression des gènes *groEL*, *groES*, *dnaK*, *clpX* et *clpP* est évaluée par RT-PCR.

L'induction des protéines chaperons lors de l'adaptation au stress thermique avaient déjà été mise en évidence par protéomique 2D chez plusieurs bactéries lactiques, et notamment dans *Lactococcus lactis*, où les protéines GroES, GroEL et DnaK avaient été identifiées par séquençage N-terminal (Dobric et al. 2000; Lim et al. 2000) ; Guillouard *et al.*, 2004). Dans cette étude, menée sur la souche *L. lactis* HD9B_GL2, nous confirmons l'induction de ces protéines et nous démontrons qu'il s'agit d'une régulation transcriptionnelle de l'opéron *groESL* et des gènes *dnaK* et *grpE*. Les gènes *dnaK* et *grpE* appartenant à l'opéron *hrcA/grpE/dnaK/dnaJ*, il est probable que les gènes *hrcA* et *dnaJ* sont eux aussi induits transcriptionnellement dans ces conditions. La régulation de ces opérons est classiquement effectuée par le répresseur HrcA, mais nous ne pouvons pas affirmer que l'induction de ces gènes est bien dépendante de l'inactivation de ce répresseur dans les conditions testées. Un élément de réponse pourrait être apporté par l'étude transcriptionnelle des autres membres du régulon HrcA défini par analyse *in silico*. Cependant, parmi les produits de ces gènes, ClpP devrait être détecté par les analyses transcriptionnelle, mais il n'a pas été identifié sur le gel. Cette situation est très différente de celle rencontrée pour les gènes *groESL*, *dnaK* et *grpE*, pour lesquels l'induction transcriptionnelle est importante et indéniable Cette différence est en

Résultats et discussion

accord avec deux modes de régulation différents pour *clpL* et le régulon HrcA. Il serait intéressant de déterminer la variation d'expression lors de l'ATR des gènes appartenant au stimulon de choc thermique mais qui ne font pas partie du régulon HrcA. Dans le cas d'une induction transcriptionnelle, se poserait la question d'une régulation coordonnée et du régulateur impliqué. L'induction des protéines de choc thermique lors de l'adaptation au stress thermique expliquerait l'observation que les lactocoques induisent une résistance accrue au choc thermique

Lorsqu'une culture de *L. lactis*, incubée à 30 ou 35 °C est soudainement exposée à une température plus élevée (40 à 42 °C), le profil protéique des cellules qui la compose, analysé par électrophorèse bidimensionnelle après marquage à la (35S) méthionine, révèle la disparition d'un grand nombre de protéines, et l'amplification, par un facteur de 2 à 3, de la synthèse de 13 (Whitaker and Batt 1991) ou 16 (Boutibonnes et al. 1991; Auffray 1992; Boutibonnes et al. 1992) polypeptides. Leur masse moléculaire (MM) s'étage de 20 à 100 kDa. Par immunodétection, sont identifiées 5 protéines interagissant avec des anticorps dirigés contre les hsp majeures d' *E. coli*: DnaK, DnaJ, GroEL, GrpE (Boutibonnes et al. 1992) (Auffray et al, 1992), et Lon (résultat non publié) ; deux autres protéines réagissent avec les anticorps du facteur sigma majeur σ^{43} de *Bacillus subtilis* (Auffray et al, 1992), et de la hsp 104 de *Saccharomyces cerevisiae*, (homologue de la protéase ClpB d'*E. coli*) (Boutibonnes et al. 1995). Bien qu'elles soient immunologiquement semblables à leurs homologues bactériens connus, avec lesquels elles présentent des épitopes communs, elles s'en éloignent quelquefois par leur masse moléculaire: c'est le cas de la protéase Lon (55 kda contre 94 ou 87 chez *E. coli*), de GrpE (94 kda alors qu'elle est chez *E. coli* classée parmi les hsp à faible MM) et du facteur σ^{43} équivalent (ce dernier, estimé à 51-55 kda, riche en résidus chargés ne migre pas sur les gels de polyacrylamide au même niveau que les molécules de référence) (Burton et al. 1981). Sa masse moléculaire réelle, inférée de la séquence génique, est de 39 kda (Gansel et al. 1993a; Gansel et al. 1993b).

Bien qu'elles ne représentent que le tiers des hsp repérées, les protéines identifiées appartiennent à tous les grands groupes (Baird and Georgopoulos 1990; Georgopoulos and Ang 1990; Laminet et al. 1990; Lipinska et al. 1990; Raina and Georgopoulos 1990; Sell et al. 1990; Georgopoulos and Welch 1993; Barrios et al. 1994; Bolliger et al. 1994; Szyperski et al. 1994; van der Vies et al. 1994; Wall et al. 1994; Wu et al. 1994; Zeilstra-Ryalls et al. 1994). Ce sont des protéases endocellulaires (ClpB et Lon) ; des composantes de la machinerie de chaperonnage : chargées d'assurer le repliement régulier des oligopeptides,

Résultats et discussion

elles participent par un mécanisme ATP-dépendant à la désagrégation et à la restauration des protéines fonctionnelles endommagées par la chaleur (Onak et GroEL) (Craig et al. 1993); des protéines d'assistance des chaperons moléculaires (*partner proteins*) (GrpE et OnaJ). Une autre protéine occupe dans la régulation de la réponse au choc thermique une position clé, c'est le facteur σ^{39} . Chaperons moléculaires et protéases endocellulaires appartiennent à des systèmes convergents dont l'activité est strictement coordonnée, au point d'ailleurs que l'une des protéases du groupe de hsp 100, Clp X a, chez *E coli* une activité de chaperonnage (Wawrzynow et al. 1995; Wawrzynow et al. 1996).

Un certain nombre d'agents chimiques, dits «thermomirnétiqes» (Boutibonnes et al. 1995), dont l'éthanol (à 2 ou 4 %) est le représentant le plus exemplaire, peuvent à température ordinaire, induire la synthèse de ces protéines singulières chez les procaryotes. Ces inducteurs ou stressseurs peuvent être répartis en quatre groupes:

- les analogues d'acides aminés, telles l'azétidine analogue de la proline et la canavanine, proche de l'arginine, qui provoquent, par défaut de repliement conforme, la formation des protéines irrégulières;

- la puromycine, qui interrompt l'édification de la chaîne polypeptidique et forme des protéines incomplètes (peptidyl-puromycine);

- les antibiotiques, qui agissent sur l'étape traductionnelle de la synthèse protéique et entraînent, par erreur de lecture ou défaut d'élongation, la production des protéines incorrectes ou tronquées;

- les agents chaotropiques, qui affectent le repliement régulier des protéines ordinaires ou occasionnent la formation d'agrégats; certains d'entre eux, la caféine, la coumarine, l'acétylsalicylate, le chlorure de baryum sont réputés être, chez les cellules eucaryotes en culture des inducteurs efficaces d'hsp. À l'exception de la chaleur elle même, et à un moindre degré, de l'alcool puis de la puromycine et des antibiotiques.

À l'inverse, la majorité d'entre eux déclenchent un processus physiologique, indépendant de l'amplification de quelques protéines privilégiées, qui conduit à l'acquisition de la thermotolérance. De plus, si l'entrée des cellules en phase stationnaire détermine chez *E coli* la mise en place d'un système multigénique complexe, le stimulon de phase stationnaire, qui intègre quelques éléments du régulon de choc thermique, ce qui se traduit par l'amplification de GroEL et de DnaK, il n'en est rien chez *L lactis* alors que, dans cette phase terminale de la croissance, les niveaux de résistance à de nombreuses épreuves létales y compris la chaleur, sont élevées (Hartke et al. 1994).

Résultats et discussion

Les profils d'expression temporelle des principaux gènes de la réponse au stress thermique *groEL*, *groES*, *dnaK*, *clpX* et *clpP* montrent une corrélation élevée au cours de la croissance dans le lait. Ce comportement d'expression similaire a permis de cerner le motif commun l'élément CIRCE. Lorsque *L. lactis* HD9B_GL2 a été inoculée, les gènes de stress ont montré des niveaux relativement identique à 30 et à 50°C (sauf pour les protéases). L'expression des gènes a augmenté lorsque les cellules entrent en l'état de transition de la croissance. Par la suite, l'expression a augmenté encore plus dans le début de la phase stationnaire de croissance. Il semble il n'y aucune corrélation entre les profils d'expression des gènes de stress thermique et l'expression des gènes de la réponse globale pendant la transition entre le début et la fin de la phase stationnaire de croissance.

5. Analyse phylogénétique

Lactococcus lactis est la bactérie la plus caractérisée des BL. Elle est considérée comme l'organisme modèle dans ce groupe, non seulement en raison de son importance économique, mais aussi en raison des caractéristiques suivantes: (i) elle a un génome complètement séquencé ; (ii) il est facile à manipuler génétiquement; et (iii) de nombreux outils génétiques ont été déjà développées pour cette espèce (Pontes et al. 2011). *Lactococcus lactis* n'a jamais été identifié comme une espèce possédant un une voie métabolique de la compétence naturelle permettant l'absorption de l'ADN exogène par d'autres bactéries et son intégration dans le chromosome. Ce processus est actif dans diverses espèces bactériennes et participe au transfert de gènes intra - et probablement inter-espèces (Wydau et al. 2006). Cependant, le lactocoque protéolytique thertmotolerant *L. lactis* HD 20A, isolé du lait de dromadaire (Drici et al. 2010), décrit dans cette étude, est le premier isolat obtenu à partir du lait cru de chamelle et semble d'être une chimère entre subsp *cremoris* et subsp *lactis* , d'après l'arbre phylogénétique construit et basé sur le gène *gdpP* en raison, du phénomène de la recombinaison produite dans ce gène. Ces caractéristiques génétiques de la souche HD20A soulèvent la question de savoir si *L. lactis* peut activer un système de compétence naturelle qui peut contribuer à un transfert génétique dans les écosystèmes laitier. Cette analyse phylogénétique vise à identifier les souches de *lactococcus lactis* au niveau de la sous espèce. Ainsi, cinq lactocoques protéolytiques isolés à partir de lait cru de chamelle dans la région aride au sud d'Algérie ont été caractérisés par des tests phénotypiques, physiologiques et biochimiques (voir résultats et discussion en haut). Toutefois, la caractérisation moléculaire est proposée pour ces isolats afin d'éviter les limites de la caractérisation phénotypique et afin

Résultats et discussion

de réaliser aussi une identification fiable et cohérente. Alors, deux gènes l'ADNr 16S/ *gdpP* sont amplifiés et séquencés au cours de cette partie de travail.

5.1. Séquençage de l'ADNr 16S: le lait de chamelle présente une diversité naturelle parmi les sous espèces de *L. lactis*

Lactococcus lactis contient quatre sous-espèces : *lactis* et *cremoris* isolées de sources laitière, *hordniae* isolée de la cicadelle *Hordnia circellata* (Latorre guzman et al., 1977) et *tractae* isolées des voies intestinales des poissons d'eau douce (Itoi et al. 2008; Itoi et al. 2009). Les deux principales sous-espèces *L. lactis* ssp. *lactis* et *L. lactis* ssp. *cremoris*, diffèrent par quelques caractères phénotypiques seulement (Kelly and Ward 2002) et sont industriellement importantes car elles sont utilisées comme *starters* pour la production de produits laitiers fermentés (Holzapfel 1997).

La sous espèce *Lactococcus lactis* ssp. *lactis* est capable de métaboliser l'arginine, et est tolérant à 40° C et 4 % NaCl, mais elle est sensible à 45° C et 6, 5 % NaCl. En outre, les souches du biovar. *diacetylactis* peuvent fermenter le citrate présent dans le lait, ce qui constitue une source d'énergie secondaire. En revanche, *L. lactis* ssp. *cremoris* est habituellement sensible à 40° C et au 4 % NaCl et la plupart des souches n'ont pas la capacité de cataboliser l'arginine. Occasionnellement, les souches peuvent avoir un génotype *lactis* avec un phénotype *cremoris* ou l'inverse.

La souche MG1363 est l'exemple d'une souche avec le génotype de ssp. *cremoris* et d'un phénotype de ssp. *lactis*. Les deux autres sous-espèce *hordniae* et *tractae* sont rarement rencontrées dans les fermentations laitières et ne seront pas abordés ici. *Lactococcus lactis* ssp. *lactis*, son biovar. *diacetylactis* et *L. lactis* ssp. *cremoris* sont utilisées comme starters dans de nombreuses industries de fermentation laitière (Daly 1983) ce qui exigent le développement rapide des nouvelles cultures présentant de rares phénotypes (van Hylckama Vlieg et al. 2006). Des séquences du fragment de 251-bp de l'ARNr 16S, fournissant une forte discrimination entre ssp. *lactis* et ssp. *cremoris* ont été comparées. Ce fragment contient une région hautement conservée et a fourni des preuves solides pour la discrimination entre les deux sous-espèces de *L. lactis*. Dans cette région, les séquences de ssp. *lactis* et ssp. *cremoris* diffèrent de 10 bp dans les 200 premières bp de la séquence (Ward et al. 1998). Les souches HD9B_GL2, HD11D, HD12C ont été identifiées comme *L. lactis* ssp. *lactis*. Les souches HD20A et HD20D ont été identifiés comme *L. lactis* ssp. *cremoris*.

La discrimination entre ssp. *lactis* et ssp. *cremoris* est particulièrement difficile, mais elle est très importante dans les applications industrielles, car les activités de ces deux sous-espèces

Résultats et discussion

sont différentes dans la production du fromage. En outre, lorsque de nouvelles souches bactériennes sont isolées nouvellement elles sont enregistrées dans des collections de culture publiques, elles doivent être identifiées au niveau de la sous espèce. Du point de vue évolutif, il est raisonnable de classer les sous-espèces selon des gènes de ménage *housekeeping genes* qui sont bien conservées au niveau du genre ou de l'espèce. Le séquençage du gène ARNr 16S est la technique moléculaire la plus utilisée actuellement pour identifier les espèces (Tanigawa et al. 2010).

Résultats et discussion

<i>L.l.lactis</i> IL1403	5'-----GACGAACGCTGGCGGC <u>GTGCCT</u>	50
<i>L.l.hordniae</i> NCD02181	5-----GACGAACGCTGGCGGC <u>GTGCCT</u>	50
<i>L.l.cremoris</i> MG1363	5'-----GACGAACGCTGGCGGC <u>GTGCCT</u>	50

<i>L.l.lactis</i>	<u>AATACATGCAAGTTG</u> AGCGCTGAAGGTTGGTACTTGTACC <u>GACTGGATGA</u>	100
<i>L.l.hordniae</i>	AATACATGCAAGTTGAGCGCTGAAGGTTGGTACTTGTACCAACTGGATGA	100
<i>L.l.cremoris</i>	AATACATGCAAGTTGAGCG <u>ATGAAGATTGGT</u> GCTTGTACCA <u>TTTGAAGA</u>	100

<i>L.l.lactis</i>	GCAGCGAACGGGTGAGTAACGCGTGGGGAATCTGCCTTTGAGCGGGGGAC	150
<i>L.l.hordniae</i>	GCAGCGAACGGGTGAGTAACGCGTGGGGAATCTGCCTTTGAG <u>TGGGGGAC</u>	150
<i>L.l.cremoris</i>	GCAGCGAACGGGTGAGTAACGCGTGGGGAATCTGCCTTTGAGCGGGGGAC	150

<i>L.l.lactis</i>	AACATTTGAAACGAATGCTAATACCGCATAAAAACCTTTAAACACAAGTT	200
<i>L.l.hordniae</i>	AACATTTGAAACGAATGCTAATACCGCATAAAAACCTTTAAACACAAGTT	200
<i>L.l.cremoris</i>	AACATTTGAAACGAATGCTAATACCGCATAAAAACCTTTAAACACAAGTT	200

<i>L.l.lactis</i>	TTAAGTTTGAAAGATGCAATTGCATCACTCAAAGATGATCCCGCTTGTA	250
<i>L.l.hordniae</i>	TTAAGTTTGAAAGATGCAATTGCATCACTCAAAGATGATCCCGCTTGTA	250
<i>L.l.cremoris</i>	TTAAGTTTGAAAGATGCAATTGCATCACTCAAAGATGATCCCGCTTGTA	250

<i>L.l.lactis</i>	TTAGCTAGTTGGTGAGGTAAAGGCTC <u>ACCAAGGCGATGATACATAG</u> CCGA	300
<i>L.l.hordniae</i>	TTAGCTAGTTGGTGAGGTAAAGGCTC <u>ACCAAGGCGATGATACATAG</u> CCGA	300
<i>L.l.cremoris</i>	TTAGCTAGTTGGTGAGGTAAAGGCTC <u>ACCAAGGCGATGATACATAG</u> CCGA	300

Figure 31 : Alignement des séquences partielles des gènes codant pour l'ARNr 16S des souches de référence utilisées pour dessiner les amorces (* identique). Position des amorces internes du gène ADNr 16S employées dans cette étude est soulignée.

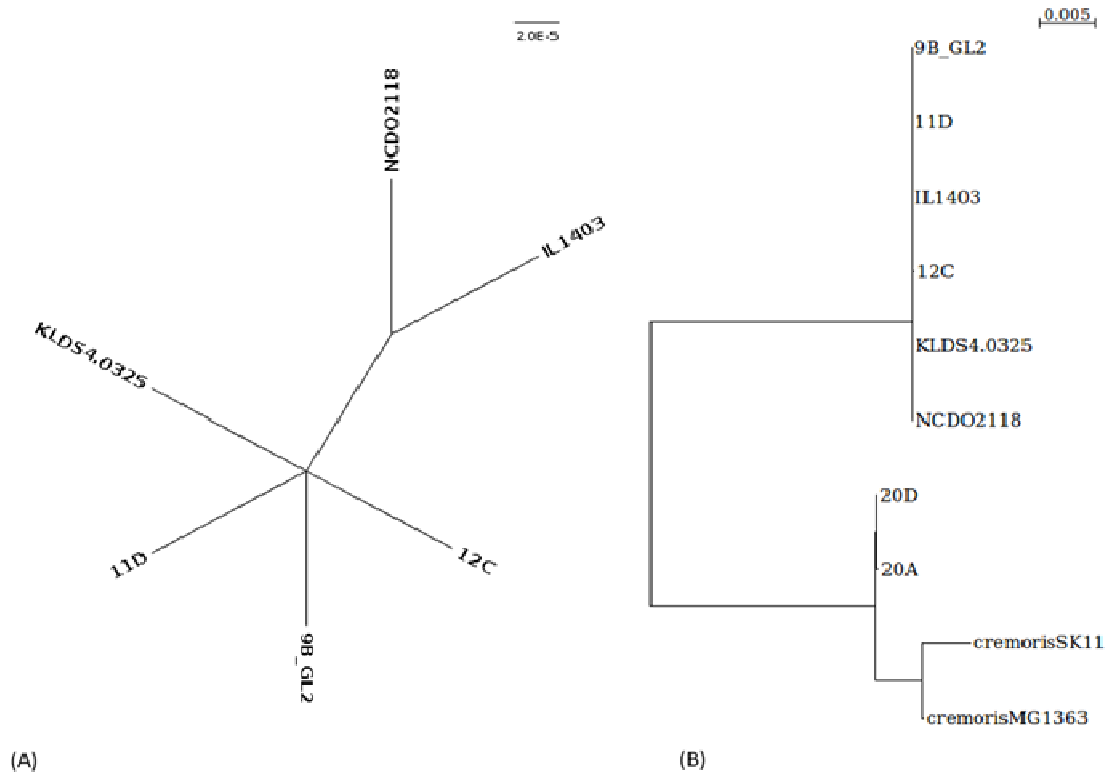


Figure 32 : Arbre phylogénétique base sur 16S DNAr

Figure 32A: décrit la relation phylogénétique entre les souches appartenant à la ssp. *lactis* seulement. HD9B_GL2 est proche à HD11D, HD12C et à KLDS4.3025 et relativement distinct des souches NCDO2118 et IL1403. **Figure 32B** indique que HD20A et HD20D sont plus proches au cluster de la ssp. *cremoris* (*L. lactis* ssp. *cremoris* SK11 et MG1363) en comparaison au cluster de la ssp. *lactis*. Méthode: Les séquences du gène 16S rRNA sont extraites et ont été comparées par Dialign-tx (Subramanian et al. 2008), le résultat est utilisé par PhyML (Criscuolo 2011) pour construire les arbres phylogénétiques selon la méthode de maximum-likelihood. Fig1A est l'arbre phylogénétique généré par les souches de la ssp. *Lactis* seulement. Fig1B l'arbre rassemble toutes les souches des deux ssp. *cremoris* et ssp. *Lactis*. Les séquences nucléotidiques du gène 16S DNAr pour les souches HD11D, HD12C, HD9B_GL2, HD20A et HD20D sont déposées au GenBank sous les numéros d'accès suivants: KR336809, KR336810, EF362617, KR336812, et KR336813 respectivement.

Résultats et discussion

En se basant sur l'arbre phylogénétique du 16S rRNA (Figure 32), la position de deux isolats HD20A et HD20D ont une position intermédiaire entre *L. lactis* subsp *lactis* et subsp *cremoris*, identique à l'arbre phylogénétique du gène *gdpP*. Toutefois, HD20A semble plus proche à *L. lactis* subsp *cremoris* en raison de l'événement de recombinaison dans sa séquence du gène *gdpP* (Figure 33). En outre, nous avons dessiné deux autres paires d'amorces ciblant le gène *gdpP* pour résoudre les limitations du gène de l'ARNr 16S pour identifier ces lactocoques. Le gène *gdpP* dispose certaines régions conservées, qui sont utilisées comme cibles pour les amorces et sont utiles pour décrypter les relations phylogénétiques entre les sous espèces étroitement liée *L. lactis* et ssp. *cremoris*.

5.2. Séquençage de *gdpP* : Variabilité génétique du gène *gdpP* parmi les sous espèces de *lactis* et signes de transfert horizontal de gènes

Le gène *gdpP* (*gdpP* : GGDEF domain protein containing phosphodiesterase) code pour une protéine membranaire de signal sous un stress donné et appartient de la famille de GdpP, dont les membres ont une activité phosphodiesterase spécifique au cyclic dimeric AMP (c-di-AMP). Les protéines de cette famille ont été caractérisées récemment chez *B. subtilis* (anciennement nommée YybT) (Rao et al. 2010) et chez *S. aureus* (Corrigan et al. 2011), qui réduisent les niveaux cellulaires de la molécule messagère secondaire du c-di-AMP. La perte de fonction de la protéine GdpP mène à la tolérance du β -Lactam/Glycopeptide chez *Staphylococcus aureus* (Griffiths and O'Neill 2012). Ces protéines sont liées à la membrane par l'intermédiaire de deux hélices transmembranaires N-terminal (Oppenheimer-Shaanan et al. 2011) et contiennent un domaine PAS (Per-ARNT-Sim) qui se lie à l'hème pour moduler l'activité des domaines DHH/DHHA1 (Rao et al. 2011) et au domaine GGDEF fortement modifié qui hydrolyse l'ATP (Rao et al. 2010). Les domaines DHH et DHHA1 fournissent l'activité phosphodiesterase spécifique du c-di-AMP, fortement inhibée par la molécule signal de la réponse stricte, ppGpp (Rao et al. 2010). Une étude récente a montré que les homologues *llmg_1816* (*gdpP* et *yybT*) sont impliqués dans la résistance ou la tolérance aux antibiotiques active des parois chez *Staphylococcus aureus* et chez *Bacillus subtilis* (Liang et al. 2009; Lo and Wang 2009; Chi et al. 2010; Corrigan et al. 2011; Griffiths and O'Neill 2012). De même, *L. lactis* MG1363 contient une seule protéine (*Llmg_0448*) avec une similitude aux protéines de diadenylate cyclase, qui reste aussi peu étudiée. À ce jour, il a été montré que les mutations au niveau du gène *gdpP* chez les bactéries Gram-positives confèrent

Résultats et discussion

une résistance à l'acide, à la haute température et aux agents de stress de la paroi cellulaire, ce qui suggère un mécanisme général de résistance au stress (Smith et al. 2012).

Pour toutes ces raisons nous avons choisi le séquençage du gène *gdpP* que l'on trouve uniquement dans les deux ssp. *lactis* et ssp. *cremoris*, mais pas dans les autres ssp : *hordniae* et *tructae*, et non pas aussi pas dans les autres firmicutes. Les résultats de séquençage ont confirmé ainsi la relation étroite entre les nouvelles souches HD20A, HD20D et les souches *L. lactis* subsp. *cremoris* SK11 et MG1363 (Figure 33) et entre HD9B_GL2, HD11D, HD12C et ssp. *lactis* KLDS4.3025 et II1403. La souche HD20A, pourrait toutefois se différencier des autres souches sur la base de sa séquence différente au niveau du même gène ; c'est un hybride entre ssp *cremoris* et ssp *lactis* malgré sa ressemblance de la séquence du gène 16s RNA. La séquence *gdpP* de HD20A montre une région de recombinaison supplémentaire dans les dernières 234 bp. Dans son extrémité 5' HD20A est identique à la ssp. *lactis* et à partir de la position nucléotidique 606, elle est identique à la sous-espèce *cremoris* pour tout le reste de la séquence du gène. Ces lactocoques isolés à partir du lait de chamelle diffèrent clairement de ceux des autres lactocoques de référence. Ces lactocoques sont clairement différents des autres lactocoques de référence. Une étude moléculaire profonde pour la souche HD20A doit être réalisée pour avoir une meilleure compréhension du transfert de gènes entre les lactocoques, son rôle dans la nature et leurs relations évolutives (Balcazar et al. 2007).

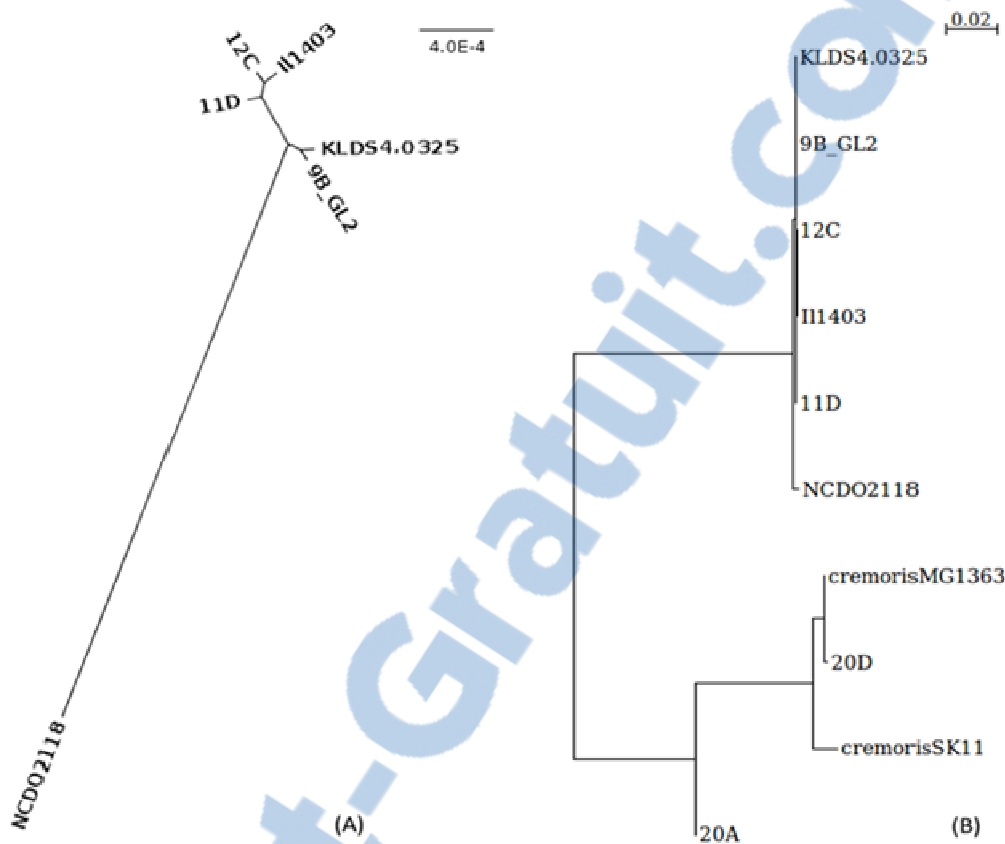


Figure 33: Arbre phylogénétique basé sur le gène *gdpP*

Figure 33A: décrit les relations phylogénétiques entre les souches de la *ssp. lactis*. HD9B_GL2 est proche à la souche KLDS4.3025 et relativement proche au HD11D, HD12C et II1403. Toutes les souches apparaissent distinctes de la souche NCDO2118. **Figure 33B** indique que HD20A et HD20D sont plus distinctes à la *ssp. lactis*. HD20D est très proche à la *ssp. cremoris* MG1363, tandis que HD20A est en quelque sorte différentes du cluster de la *ssp. cremoris*. Méthode: les séquences du gène *gdpP* sont extraites et alignées par Dialign-tx, le profil généré est ensuite utilisé pour calculer les arbres phylogénétiques selon la méthode de maximum-likelihood par PhyML. Figure 2A Les arbres phylogénétiques sont générés avec les souches de la *ssp. lactis* seulement; la figure est générée par Figtree (Vlad I Morariu et al. 2008). Figure 2B est généré par Seaview (Galtier et al. 1996); et contient tous les souches testées. Les séquences nucléotidiques du gène *gdpP* pour les souches HD11D, HD12C, HD9B_GL2, HD20A et HD20D sont déposées au GenBank sous les numéros d'accès suivants: KR904887, KR904888, KR904891, KR904889, KR904890 respectivement.

Résultats et discussion

Les résultats de l'identification des sous-espèces de *Lactococcus lactis* en se basant sur les séquences de l'ARNr 16 s et *gdpP* étaient concordants. Cependant, une attention particulière a été donnée à *L. lactis* ssp. *lactis* biovar. Diacetylactis HD20A après avoir analysé sa séquence du *gdpP* (Figure 34). Cette figure montre les nucléotides du nt 420 au nt 840 seulement, la séquence de HD20A dans son extrémité 5' HD20A est identique à la ssp. *lactis* et à partir de la position nucléotidique 606, elle est identique à la sous-espèce *cremoris* pour tout le reste de la séquence du gène. La séquence de HD20A est une nouvelle variation qui n'a pas été signalée précédemment.

Nos analyses montrent aussi que *gdpP*, présente plusieurs avantages comme marqueurs moléculaires, y compris (i) il n'est présent que chez l'espèce de *L. lactis* et absent chez les autres sous espèces; (ii) absent chez *L. garvieae*, *L. fujiensis*, *L. raffinolactis* et *L. piscium* ; (iii) aucune déclaration n'est possible en raison de l'absence des séquences génomiques complète pour ssp. *tructae* et ssp. *hordniae* ; (iv) pour toutes les autres espèces du genre *Lactococcus*, aucune déclaration ne peut être effectuée parce qu'aucune séquence complète n'est disponible pour ces espèces ; (v) ayant une certaines formes du transfert latéral de gène (en cas de HD20A) n'affecte pas sa résolution phylogénétique ; (vi) considérant la séquence nucléotidique obtenue, le gène *gdpP* montre une plus grande variation dans les taux évolutifs des positions nucléotidiques sur toute sa longueur que par rapport à l'ARNr 16 s (une grande taille de 1935 nt).

Pour toutes ces raisons le *gdpP* donne une meilleure résolution par rapport à l'ARNr 16S pour identifier les souches de *L. lactis* au niveau de la sous-espèce. Toutefois, le séquençage de l'extrémité 5' de l'ARN 16S peut conduire à des résultats ambigus, par exemple sur les six gènes d'ARN 16S trouvés dans le génome de *L. lactis* ssp. *cremoris* SK11, cinq sont identiques et clairement appartiennent à la séquence type de la sous-espèce *cremoris*, tandis que le sixième (2179) montre des substitutions de quatre paires de bases seulement par rapport à la séquence type de la sous-espèce canonique *lactis*.

5.3. Lait de chamelle, une niche écologique propice au transfert entre les sous espèces *Lactococcus lactis*

Les génomes microbiens sont très flexibles et les échanges d'ADN représentent des sources importantes de diversité intra spécifique. Le transfert de gène joue un rôle important dans l'adaptation de *L. lactis* à l'environnement laitier et peut-être facilité par l'apparition fréquente de transposons, des plasmides et des bactériophages (van Hylckama Vlieg et al. 2006). Le

Résultats et discussion

séquençage récent de quatre plasmides de *L. lactis* subsp. *lactis* ont révélé l'existence au moins de neuf régions qui pourraient avoir été acquis par le biais de transfert horizontal de gènes (Siezen et al. 2005) de streptocoques, des entérocoques et des lactobacilles.

La même étude rapporte la prédiction de plusieurs adaptations à l'environnement laitier. Cette liste contient non seulement des gènes impliqués dans l'utilisation des carbohydrates et de peptides, mais aussi des gènes pour les exigences en cofacteur et en fer, pour le métabolisme énergétique et les réponses au stress (Siezen et al. 2005). D'après nos connaissances, *L. lactis* n'a jamais été identifié comme une espèce possédant une voie naturelle de compétence ainsi que la transformation naturelle n'a jamais été observée chez les souches de *Lactococcus lactis*.

Toutefois, dans une étude récente (Shimizu-Kadota et al. 2013), l'analyse de la séquence génomique complète de *L. lactis* IO-1, une souche bactérienne non laitières isolée de l'eau a révélé la présence de gènes issus d'un transfert horizontal (présence de prophages, présence de transposon conjugatif nisine-saccharose et plusieurs systèmes de restriction-modification), il a été rapporté que la souche *L. lactis* IO-1 est difficile à transformer avec des plasmides préparés à partir d'*E. coli* ou de *B. subtilis*, alors que ceux de IO-1, se sont facilement introduites (Yasui et al. 2009). Boguslawska et coll. (2009), (Boguslawska et al. 2009) ont montré in vivo le transfert d'un transposon de résistance aux antibiotiques semblables à Tn916 de *Lactococcus lactis* à *Enterococcus faecalis*. Ces résultats suggèrent que dans certains cas, les lactocoques pourraient être impliqués dans la propagation des gènes de résistance aux antibiotiques à d'autres bactéries d'acide lactique (Boguslawska et al. 2009).

Une autre étude, (Gabed et al. 2015), l'analyse de la séquence complète du génome de HD9B_GL2, une souche de *L. lactis* ssp *lactis* biovar diactetylactis isolée à partir du même échantillon lait que les souches étudiées dans ce travail (Drici et al. 2010), a révélé la présence d'orthologues de plusieurs gènes impliqués dans le processus de la compétence naturelle dans diverses bactéries (Bolotin et al. 2001; Claverys and Martin 2003). En se basant sur l'homologie de séquence, le gène *comC* et tous les gènes de compétence requis pour la formation de pores pour l'entrée de l'ADN [*comC*, *comEA*, *comEC*, *comGF*, *comGE*, *comGC*, *comGB*, *comGG*, *comFC*, *comGA* et *comGD*] ont été identifiés dans HD9B_GL2. Bien que le transfert de gènes massive entre les procaryotes est suggéré d'être un des facteurs responsables de l'évolution des génomes bactériens (Syvanen 1994; Koonin and Galperin 1997; Doolittle 1998), les gènes codant les composants de la traduction sont hautement

Résultats et discussion

conservés et difficile pour un transfert horizontal en raison de la complexité de leurs interactions (Jain et al. 1999).

Cependant, quelques études récentes ont mis en évidence que le transfert horizontal de gènes a également eu lieu dans l'évolution de certains gènes codant pour les composantes de la traduction 16S rRNA et quelques aminoacyl-tRNA synthétases (Koonin and Galperin 1997; Brown and Doolittle 1999; Ueda et al. 1999; Wolf et al. 1999; Yap et al. 1999).

Nos résultats concordent avec ceux trouvés par Bolotin et al (2004)(Bolotin et al. 2004) ; ils ont révélé que le gène *ycaB*, codant pour une protéine de fonction inconnue a été échangé entre des lactocoques et des entérobactéries. Ils ont suggéré que *L. lactis* est probablement la donatrice et ont révèlent aussi trois moyens du transfert de gène aux entérobactéries. L'analyse du transfert de gène *ycaB* entre les deux sous espèces *L. lactis* subsp. *lactis* et *L. lactis* subsp. *cremoris*, indique que le gène peut être mobilisé, éventuellement par la conjugaison étant donné que l'analyse du gène *ycaB* a révélé que deux souches de *L. lactis* MG1363, QA5 et QA30 contiennent des allèles de IL1403. Cela indique que le gène *ycaB* était horizontalement transféré entre les souches de *L. lactis* (Bolotin et al. 2004). Dans notre cas, il est difficile d'en déduire le mécanisme de la recombinaison à partir des sites de *crossover* parce qu'on n'a pas analysé l'organisation des régions flanquantes le gène *gdpP* chez la souche de HD_20A. Des travaux approfondis devraient permettre le transfert du *gdpP* et de déterminer son rôle.

5.4. *L. lactis* subsp *lactis* biovar *diacetylactis* HD20A est une chimère entre ssp. *cremoris* et ssp. *lactis*

La souche *L. lactis* HD_20A décrite ci-dessus est la première nouvelle chimère provenant de sources naturelles et qui peut être disponible aux scientifiques et aux chercheurs de *Lactococcus*. La disponibilité de nouvelles souches de la sous-espèce de *L. lactis* est très importante pour l'industrie laitière.

Ces isolats pourraient constituer une source de variation génétique naturelle dont les caractéristiques recherchées, y compris la résistance à la chaleur, et peuvent être sélectionnés. En outre, cet isolat peut fournir une ressource pour les études futures de la systématique et la génétique des populations des lactocoques dans les produits laitiers. Ces caractéristiques génétiques de la HD20A soulèvent la question de savoir si *L. lactis* peuvent activer un système de compétence naturelle qui peut contribuer à un transfert génétique dans l'écosystème laitier.

Résultats et discussion

		480
IL1403	TATGTCGCAGAACTTGATTCTGATAAGAATCTTATTTATTTAATTGATGCCACAAAAGAA	
HD9B_GL2	TATGTCGCAGAACTTGATTCTGATAAGAATCTTATTTATTTAATTGATGCCACAAAAGAA	
HD11D	TATGTCGCAGAACTTGATTCTGATAAGAATCTTATTTATTTAATTGATGCCACAAAAGAA	
HD12C	TATGTCGCAGAACTTGATTCTGATAAGAATCTTATTTATTTAATTGATGCCACAAAAGAA	
HD20A	TATGTCGCAGAACTTGATTCTGATAAGAATCTTATTTATTTAATTGATGCCACAAAAGAA	
HD20D	TATGTTGCAGAACTTGATTCTGATAAAAACCTTATTTATTTAATTGATGCCACAAAAGAA	
MG1363	TATGTTGCAGAACTTGATTCTGATAAAAACCTTATTTATTTAATTGATGCCACAAAAGAA	
		540
IL1403	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCGGTTATTGGTGCGATTTCTGTTGAC	
HD9B_GL2	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCGGTTATTGGTGCGATTTCTGTTGAC	
HD11D	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCGGTTATTGGTGCGATTTCTGTTGAC	
HD12C	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCGGTTATTGGTGCGATTTCTGTTGAC	
HD20A	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCGGTTATTGGTGCGATTTCTGTTGAC	
HD20D	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCAAGTTATTGGTGCGATTTCAAGTTGAT	
MG1363	GTTGCTTTTAAATCAGAATTTAATGACAGTCGTGCAAGTTATTGGTGCGATTTCAAGTTGAT	
		600
IL1403	AATTATGATGATGCTACAGATTTAATTACTGACAGCGGTCGAACAGCCATTAATAGTTTT	
HD9B_GL2	AATTATGATGATGCTACAGATTTAATTACTGACAGCGGTCGAACAGCCATTAATAGTTTT	
HD11D	AATTATGATGATGCTACAGATTTAATTACTGACAGCGGTCGAACAGCCATTAATAGTTTT	
HD12C	AATTATGATGATGCTACAGATTTAATTACTGACAGCGGTCGAACAGCCATTAATAGTTTT	
HD20A	AATTATGATGATGCTACAGATTTAATTACTGACAGCGGTCGAACAGCCATTAATAGTTTT	
HD20D	AATTATGATGATGCAACTGATTTGATTACTGATAGTGGTCGAACAGCCATTAATAGTTTT	
MG1363	AATTATGATGATGCAACTGATTTGATTACTGATAGTGGTCGAACAGCCATTAATAGTTTT	
		660
IL1403	ATCGCTTCATTTTTAGAAGAATTTGCTGACAAATATGGAGTTTATCTTCGCCGTATTAAT	
HD9B_GL2	ATCGCTTCATTTTTAGAAGAATTTGCTGACAAATATGGAGTTTATCTTCGCCGTATTAAT	
HD11D	ATCGCTTCATTTTTAGAAGAATTTGCTGACAAATATGGAGTTTATCTTCGCCGTATTAAT	
HD12C	ATCGCTTCATTTTTAGAAGAATTTGCTGACAAATATGGAGTTTATCTTCGCCGTATTAAT	
HD20A	ATCGCATCATTTTTAGAAGAATTTGCTGACAACTATGGAATTTATCTCCGTCGAATAAAC	
HD20D	ATCGCATCATTTTTAGAAGAATTTGCTGACAACTATGGAATTTATCTCCGTCGAATAAAC	
MG1363	ATCGCATCATTTTTAGAAGAATTTGCTGACAACTATGGAATTTATCTCCGTCGAATAAAC	
		720
IL1403	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
HD9B_GL2	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
HD11D	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
HD12C	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
HD20A	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
HD20D	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
MG1363	AGCAGTCGCCATTATTTTTCTGTGATTATCGTATTCTTGAAAAAATGATTAACGATAAA	
		780
IL1403	TTTTCTGTTCTAAAAGAATTTGAGAACTTTCATCTCAAAAAGAAATTCCTTTGACTCTT	
HD9B_GL2	TTTTCTGTTCTAAAAGAATTTGAGAACTTTCATCTCAAAAAGAAATTCCTTTGACTCTT	
HD11D	TTTTCTGTTCTAAAAGAATTTGAGAACTTTCATCTCAAAAAGAAATTCCTTTGACTCTT	
HD12C	TTTTCTGTTCTAAAAGAATTTGAGAACTTTCATCTCAAAAAGAAATTCCTTTGACTCTT	
HD20A	TTCCGCTGTTCTCAAAGAATTTAGAGAACTTTCAGCTCAAAAAGAAATTCCTTTGACTCTT	
HD20D	TTCCGCTGTTCTCAAAGAATTTAGAGAACTTTCAGCTCAAAAAGAAATTCCTTTGACTCTT	
MG1363	TTCCGCTGTTCTCAAAGAATTTAGAGAACTTTCAGCTCAAAAAGAAATTCCTTTGACTCTT	
		840
IL1403	TCAGTTGGTGTGGCTTATGGTTGGAATGATTTTCCAGTTATTGGTAAAGTTGCTCTAAAT	
HD9B_GL2	TCAGTTGGTGTGGCTTATGGTTGGAATGATTTTCCAGTTATTGGTAAAGTTGCTCTAAAT	
HD11D	TCAGTTGGTGTGGCTTATGGTTGGAATGATTTTCCAGTTATTGGTAAAGTTGCTCTAAAT	
HD12C	TCAGTTGGTGTGGCTTATGGTTGGAATGATTTTCCAGTTATTGGTAAAGTTGCTCTAAAT	
HD20A	TCTATCGGTGTGGCTTATGGTTGGAATGATTTTCCATACCATCGGAAAAGTTGCCCTCAAT	
HD20D	TCTATCGGTGTGGCTTATGGTTGGAATGATTTTCCATACCATCGGAAAAGTTGCCCTCAAT	
MG1363	TCTATCGGTGTGGCTTATGGTTGGAATGATTTTCCATACCATCGGAAAAGTTGCCCTCAAT	

Figure 34: Alignement des séquences partielles du gène *gdpP* (*recombination breakpoint region*). Toutes les positions où les souches HD et/ou *L. lactis* ssp. *cremoris* MG1363 s'écartent de la souche de référence *L. lactis* ssp. *lactis* IL1403 sont marqués par des caractères blancs sur fond noir. La figure montre les nucléotides de 420 à 840 seulement. La séquence de HD20A dans son extrémité 5' est identique à la ssp. *lactis* et à partir du nt 606 elle est identique à ssp. *cremoris* pour le reste de la séquence.

Résultats et discussion

Le gène *gdpP* fournit une résolution phylogénétique comparable à celle du gène de l'ARNr 16S au niveau de l'espèce, sauf au niveau de sous-espèces étroitement, pour lequel il fournit une meilleure résolution. Cela est particulièrement pertinent dans le contexte d'un nombre croissant d'études qui portent sur la diversité à l'échelle des sous-espèces, et dans lesquelles les gènes codant une seule copie protéique tels que le gène *gdpP* pourraient compléter les informations fournies par le l'ARNr 16S. L'utilisation du *gdpP* pour l'analyse de *Lactococcus lactis* est une étape importante dans la diversité des BL, car elle pourrait permettre à la mesure exacte de la diversité et des relations phylogénétiques, évitant une perte de résolution phylogénétique de la diversité en raison de la présence d'hétérogénéité intragénomique.

En outre, l'influence de l'hétérogénéité d'intragénomique du 16S rRNA sur la phylogénie bactérienne au niveau de la sous-espèce a été déterminée (Case et al. 2007). Nous suggérons donc que *gdpP* pourrait être utilisé comme un marqueur supplémentaire ou alternatif dans les études de la biodiversité, tel qu'il est capable de déchiffrer les relations phylogénétiques qui passent inaperçues à l'aide du gène l'ARNr 16S. En effet, l'utilisation d'un marqueur comme *gdpP* serait particulièrement importante pour les études qui portent sur la discrimination des souches au niveau de la sous-espèce grâce à son pouvoir de résolution, qu'il fournirait ainsi que l'absence d'hétérogénéité intragénomique comme une source de biais.

Chapitre 3 : Résultats et Discussion

Partie II : Analyse du génome de *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* HD9B_GL2

1. Introduction

Lors de l'initiation de ce travail, peu de données génétiques sur la réponse de *Lactococcus lactis* au stress thermique étaient disponibles et aucune autre analyse génomique complète des gènes heat shock chez les lactocoques n'était encore publiée.

Le génome de la souche *Lactococcus lactis* HD9B_GL2 a fait l'objet d'un projet de séquençage du génome entier au cours de mon stage dans le Biozentrum à l'université de Julius Maximilians à Würzburg en Allemagne, dans le cadre de préparation d'une thèse de Doctorat en science. Ce travail a été co-financé par la Faculté des Sciences de l'Université de Hassiba benbouali de chlef (Algérie) et the German Research Fondation (Grant TR34/Z1, Allemagne). Ce projet génomique était important à plus d'un titre, puisque les données génomiques permettent de connaître le répertoire de gènes et de fonctions présents dans la souche, mais aussi puisqu'elles peuvent servir de base pour explorer les variabilités génomiques de l'espèce et permettre d'accéder aux analyses globales via les approches transcriptomique et protéomique. Cette partie a fait l'objet d'une publication "Draft Genome Sequence of the Moderately Heat-Tolerant *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* Strain GL2 from Algerian Dromedary Milk" dans "Genome Announcement".

Une fois le génome est séquençé, assemblé et l'annotation réalisée, j'ai recherché plus spécifiquement dans le génome les éléments mettant en évidence une possible adaptation de *L. lactis* HD9B_GL2 à son milieu, le milieu lait. Cette recherche a été axée sur l'étude des gènes impliqués dans la réponse de la souche au stress thermique ainsi que sur les systèmes de régulation et les mécanismes de résistance au stress présent. Cette partie a fait l'objet d'une autre publication "Genome wide analysis of heat tolerance-related genes and proteins in *Lactococcus lactis* GL2 isolated from dromedary milk – its *hrcA* gene has an unusual 5'-region and it contains two *recA* genes". (Sous révision dans "Genome").

2. Définitions importantes

Les contigs et l'assemblage, les trous, l'appel de base

Avec les technologies encore les plus courantes dans un grand nombre de laboratoires, chaque séquençage ne permet d'obtenir une lecture que de quelques milliers de paires de base. Il n'est donc pas possible de séquencer en une seule fois des molécules d'ADN aussi grandes que les chromosomes. Pour reconstituer ces immenses séquences, il faut effectuer un grand nombre de séquençages, plusieurs fois supérieur à la taille du chromosome. Ces séquençages redondants permettent de raccorder les séquences les unes aux autres et de s'assurer de la qualité du résultat de chaque lecture. Le séquençage complet d'un génome avec les NGS conduit à un nombre colossal de petits fragments séquencés (un grand nombre de petites séquences ou lectures) que l'on essaye ensuite d'assembler en contigs. La qualité de couverture du séquençage et donc liée à celle des contigs (leur longueur et leur continuité) et donc au nombre de *gaps*.

Lecture ("*read*") : une séquence d'un fragment. **La couverture** d'une séquence est le nombre moyen de fois qu'une base du génome est représentée dans la séquence des lectures.

Profondeur de couverture: nombre de reads chevauchante à une région donnée (en X).

Contigs : séquences continues générées par l'alignement de séquences de fragments qui se chevauchent. **L'assemblage** : La comparaison des séquences permet d'aligner les parties qui se recouvrent partiellement ou chevauchantes.

Les séquences chevauchantes peuvent être reliées en enchaînements plus grands que l'on appelle contigs. En reliant l'ensemble des contigs, on reconstitue des séquences de plusieurs millions à plusieurs dizaines de millions de bases. Ces opérations sont effectuées par des programmes bioinformatiques. **Les trous ou**

"*gap*" : parties du génome non séquencées ou dont les séquences ne chevauchent pas avec d'autres et ne peuvent donc entrer dans un contig. Comme le séquençage est effectué sur des sous-fragments pris de manière aléatoire, même avec un tel niveau de redondance, il reste des parties non assemblées : des trous ("*gap*") qui peuvent être "comblés" par un travail ciblé. Les trous "*gaps*" sont de longueur connue. **Scaffold** : ensemble de contigs orientés et ordonnés.

Mapped scaffold : ensemble de *scaffolds* localisés le long des chromosomes (pas forcément ordonnés ou orientés). Pour déterminer les relations de voisinage des contigs, les liens clones sont considérés, c'est-à-dire les lectures obtenues aux deux extrémités d'un même fragment d'ADN. On recherche parmi ces paires celles qui s'ancrent dans deux contigs différents. Cela permet de jeter un pont entre les deux contigs et de les orienter. De plus, le fragment d'ADN

Résultats et discussion

"à cheval" sur le trou entre les deux contigs peut faire l'objet d'un séquençage supplémentaire, ce qui permet de combler le trou. **La lecture des profils bruts ou "base-calling"** : c'est la détermination de la séquence par appel de bases qui s'effectue en routine par des programmes informatiques qui déterminent l'identité des bases, comparent les séquences et fournissent une plate-forme intuitive de correction. **Régions de faible complexité** : parties du génome dont les séquences sont très peu diversifiées comme les séquences répétées. **Métagénomique** : étude du génome d'un organisme prélevé directement dans un environnement complexe (intestin, océan, sols, ...), à l'inverse d'un organisme de laboratoire. Le but est d'obtenir des informations sur l'incidence de cet environnement. Le préfixe "méta" signifie « après, au-delà de, avec.. ».

Résultats et discussion

séquence A TTAGCGATGATTTCTCGATGCGGTTCCAATTGC
séquence B ATTCGGAATGCATCTTAGCGATGATTTCTCGATGC
contig ATTCGGAATGCATCTTAGCGATGATTTCTCGATGCGGTTCCAATTGC

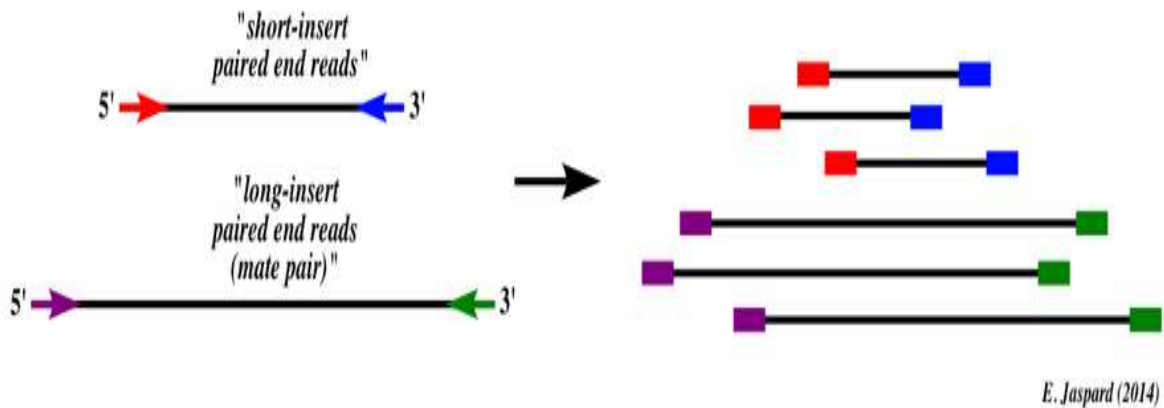
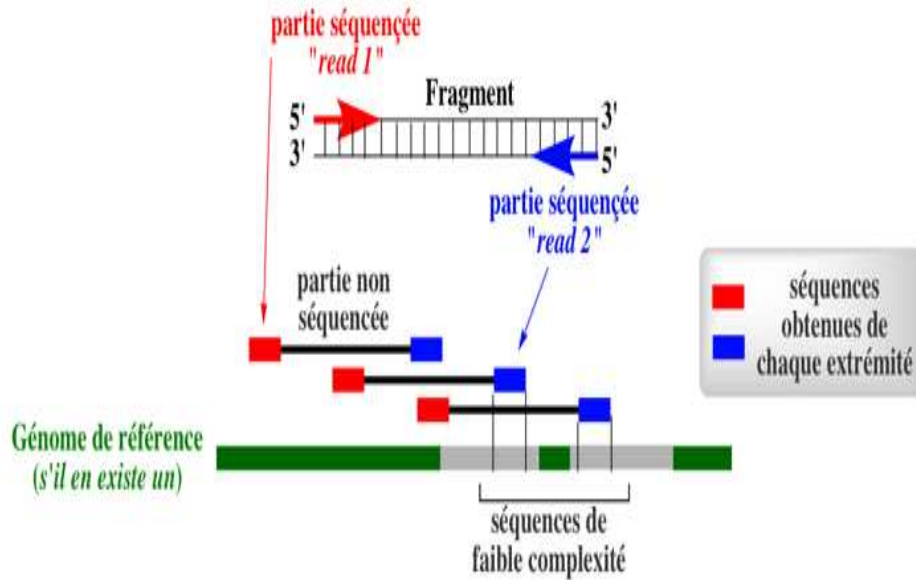


Figure 35: Séquençage en « *paire end reads* » selon Illumina

3. Analyse des données de séquençage

Filtrer les lectures d'un séquençage est une étape indispensable dans l'objectif d'obtenir un assemblage cohérent. La présence d'un trop grand nombre d'erreurs dans les séquences peut conduire à la formation de contigs chimériques et l'obtention d'un mauvais assemblage. Pour les lectures obtenues par la méthode Solexa/Illumina, la filtration des lectures s'est faite en se basant sur les scores de qualité attribués à chaque base par le séquenceur. Ainsi, les séquences

possédant des scores élevés ont été conservées pour l'assemblage. L'analyse de la couverture des lectures le long de séquences de référence a ensuite aussi été utilisée pour estimer la proportion du génome couvert.

3.1. Vérification de la qualité des séquences

Les fichiers que nous avons récupéré sont des séquences issues d'un séquençage par la méthode Illumina. Avant de commencer à les utiliser, nous devons en contrôler la qualité. Pour cela nous avons utilisé le logiciel FASTQC (disponible gratuitement (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>)). Ce logiciel tourne sous toutes les plate-formes et permet de générer un rapport de la qualité moyenne des séquences.

3.2. Suppression des adaptateurs/primers

Les adaptateurs/primers sont utilisés pour former la banque, créer les clusters et les séquencer. En fonction de la taille des séquences initiales (avant formation des clusters) et de la qualité de fabrication de la banque initiale, les adaptateurs peuvent être présents en plus ou moins grande quantité. Pour supprimer ces séquences nous avons utilisé le logiciel CutAdapt, capable de couper les séquences adaptatrices (entières ou non, avec indel/mismatches) et d'éliminer les contaminants résiduels provenant du multiplexage (séquençage de plusieurs individus en même temps) sans risquer d'éliminer trop de vraies séquences. Le seuil de qualité choisi permet aussi d'éliminer les bases de mauvaises qualités situées sur la fin des séquences.

3.3. Filtres des séquences

Afin de conserver uniquement des séquences de bonnes qualités, nous allons les filtrer sur leur qualité moyenne en utilisant FastQ (Quality Control). Pour filtrer les lectures Solexa/Illumina, quatre critères ont été utilisés :

- Si plus de deux nucléotides dans une lecture possèdent un score de qualité inférieur à 10, cette lecture n'est pas considérée. La rigueur de ce filtre permet de ne conserver que les séquences de haute qualité.
- Si il y présence d'un nucléotide non-identifié dans la séquence (un N), la lecture est supprimée. Les assembleurs utilisés remplacent arbitrairement les N dans les lectures par des A, ce qui pourrait provoquer des erreurs d'assemblage.
- S'il y a plus un même nucléotide dans plus de 90% de la lecture, celle-ci est retirée. La séquence de ces lectures est peu informative et leur présence risque de perturber l'assemblage.
- Si le score moyen de la lecture est inférieur à 25, la lecture est éliminée. De façon générale, il n'a jamais été formellement prouvé que ce dernier critère de filtration puisse améliorer de façon notable la qualité des lectures utilisées pour un assemblage. Toutefois,

Résultats et discussion

compte tenu de la grande quantité de séquences reçues du prestataire StarSeq des assemblages sur le matériel informatique utilisé aurait été compliquée du fait de la trop grande quantité de mémoire demandée. Ce critère de filtration a donc permis de diminuer la quantité de séquences sans pour autant diminuer la qualité globale des lectures utilisées lors de l'assemblage. Pour l'ensemble des lectures analysées, le score moyen diminue à l'extrémité 3' de la séquence. Cela traduit globalement une diminution de la fiabilité du séquençage en fin d'extension.

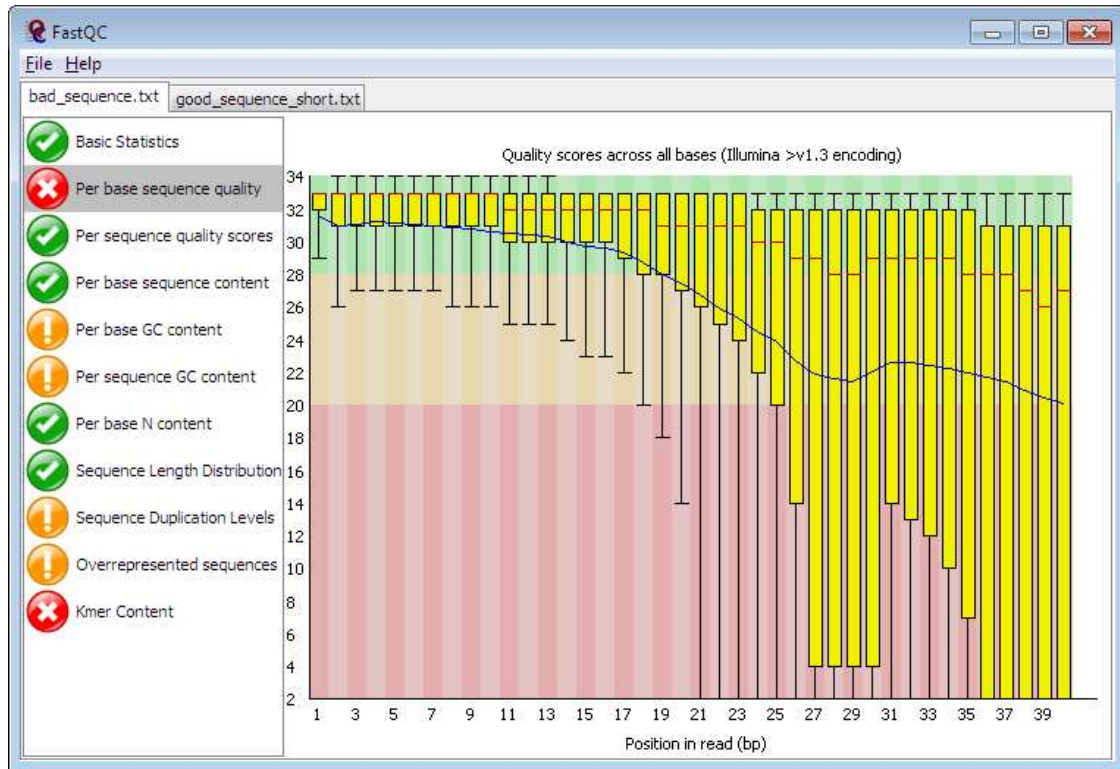


Figure 36 : Interface utilisateur de FastQC. L'interface FastQC permet d'obtenir la qualité des séquences génomiques issues du séquençage ainsi que le % en GC des lectures.

3.4. Assemblage des séquences

Pour obtenir la séquence génomique de la souche HD9B_GL2 de *Lactococcus lactis*, une stratégie de séquençage par terminaison cyclique réversible de Solexa/Illumina, a été mise en place. Une banque de séquences Solexa/Illumina de type « *paired-end* » (une avec des lectures de 63 pb et une autre avec des lectures de 58 pb) a été reçue au cours de ces travaux de la part du prestataire StarSeq (cf : matériel et méthodes). La distance et l'orientation de chaque paire de lectures étant connues, les assembleurs doivent respecter ces deux informations lors de l'assemblage, ce qui en améliore la qualité.

Résultats et discussion

Une fois les filtrations effectuées, nos séquences ont été nettoyées, différents assemblages ont été lancés dans le but d'obtenir le génome de *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* HD9B_GL2. En général, si l'on ne dispose pas de séquence de référence, il convient de créer cette dernière via un assemblage de novo. Les résultats des assemblages sont présentés dans le tableau 5.

3.4.1. Assemblage de novo

Les assembleurs suivants sont utilisés : Soapdenovo-127, Velvet, Metavelvet, Omega, Ray- Meta et IDBA-UD. Pour créer cette séquence, nous avons utilisé les logiciels Velvet et IDBA-UD. Ces deux assembleurs permettent de faire des assemblages quelque soit la technologie utilisé pour générer les séquences. Ils permettent également de faire des assemblages mixtes entre séquences issues de différentes technologies.

3.4.2. Mapping des données sur une référence

Ici, nous n'avons pas besoin de créer une référence, nous allons utiliser la séquence du *Lactococcus lactis* ssp *lactis* KLDS4.0325, dont le génome entier est disponible et bien annoté dans GenBank (abbr. KLDS4. 0325, accession no. CP006766). Pour le mapping nous avons utilisé MIRA et AMOS. Mira permet de faire des assemblages quelque soit la technologie utilisé pour générer les séquences. Il permet également de faire des assemblages mixtes entre séquences issues de différentes technologies. Cet outil allie rapidité et précision, mais est très sensible aux différences entre les reads et la référence.

3.5. Construction de gabarits « Scaffolding »

Elle consiste à ordonner et à orienter les séquences assemblées « contigs ». Les nouveaux outils utilisés dans ce travail et qui permettent de construire des gabarits génomiques, sont SSPACE, ScaffoldBuilder et CONTIGuator.

3.6. Annotation

Nous avons annoté le génome séquencé en rassemblant les résultats d'annotation issus des deux programmes d'annotations : RAST et Prokka. Les caractéristiques générales du génome de la souche HD9B_GL2 sont présentés dans le tableau 16. Toutes les caractéristiques génomiques de la souche HD9B_GL2 issues du séquençage en utilisant la plateforme d'annotation RAST sont détaillées dans la partie Annexe.

3.7. Création de Workflows

Enchaîner toutes ces étapes est fastidieux. C'est pourquoi il est intéressant d'automatiser leur enchaînement. Pour cela nous avons créé un petit workflow qui résume les différents traitements des lectures effectués au cours de ces travaux (Figure 37).

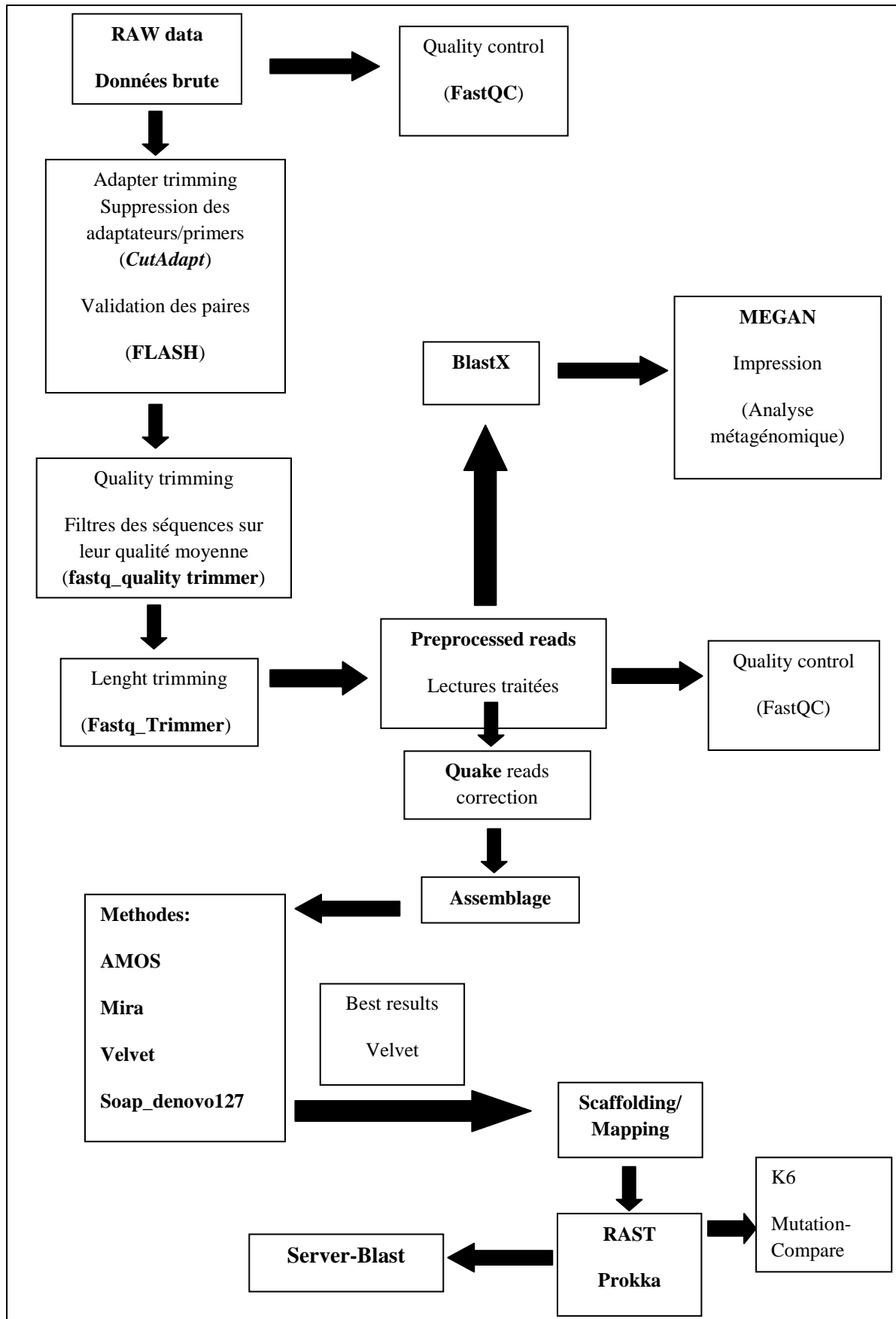


Figure 37 : Workflow : Pipeline de l'assemblage et d'annotation.

Résultats et discussion

La préparation de la librairie d'ADN (TruSeq PE Cluster kit v2, Illumina) a permis de générer des inserts de 250 à 650 bp. Des séquences brutes de 780 Mb paired-end (PE) ont été collectées d'Illumina MiSeq et leur qualité a été vérifiée par FastQC. FLASH (Magoc and Salzberg 2011) est ensuite utilisé pour fusionner les séquences enchevâchées non traitées et les deux adaptateurs (p5 et p9) ont été coupés en utilisant cutadapt (Martin, 2011). Toutes les extrémités queues de faible qualité (<25) ont été éliminées. Les kmers ont été comptés par Jellyfish (Marcais and Kingsford 2011) et les résultats sont utilisés pour corriger les séquences par QUAKE (Kelley et al., 2010). L'assemblage a été fait par Velvet (Zerbino and Birney 2008) et IDBA-UD (Peng et al. 2012). L'annotation du génome a été réalisée par RAST 4 (Overbeek et al. 2014).

4. Analyse des données génomique de *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* HD9B_GL2

Cette souche a été choisie car elle est considérée comme une souche type de la sous espèce *Lactococcus lactis* subsp. *lactis* biovar. *Diacetylactis* et elle a été isolée du lait cru de chamelle d'Algérie (Drici et al. 2010). La séquence génomique obtenue peut servir de base pour explorer les variabilités génomiques de l'espèce et permettre d'accéder aux analyses globales des gènes impliqués dans la réponse au stress thermique via les logiciels de bioinformatique. La séquence génomique totale de la souche HD9B_GL2 issue du projet de séquençage a été déposée dans GenBank/EMBL/DDJB sous les numéros d'accès JNCC02000000 (chromosome), et renfermant les séquences du JNCC02000001 au JNCC02000048.

Les souches de référence utilisées pour la génomique comparative sont: *L. lactis* ssp. *lactis* IL1403 (abbr. IL1403, accès no. AE_005176), *L. lactis* ssp. *Lactis* NCDO 2118 (abbr. NCDO2118, accès no. NZ_CP009054.1), *L. lactis* ssp. *lactis* KLDS 4.0325 (abbr. KLDS 4.0325, accès no. CP_006766), *L. lactis* subsp. *cremoris* MG1363 (abbr. MG1363, accès no. NC_009004.1), KF147 (NC_013656), *Lactococcus lactis* subsp. *lactis* CV56 (NC_017486.1), *Lactococcus lactis* subsp. *cremoris* KW2 (NC_022369.1), et *L. lactis* subsp. *cremoris* SK11 (abbr. SK11, accès no. CP_006766.1).

Résultats et discussion

Tableau 15 : Comparaison des différents assembleurs utilisés

Type	Logiciel	Résultat
Mapping	AMOS	O
	Mira	+
Assemblage de <i>novo</i>	Soapdenovo-127	+
	Velvet	+
	Metavelvet	O
	Omega	+
	Ray-Meta	++
	IDBA-UD	+++

Ce tableau résume les résultats obtenus avec les différents assembleurs utilisés (assemblage de novo ou le mapping). Aucun résultat n'a été obtenu avec les deux assembleurs AMOS et Metavelvet, ce qui peut être dû aux algorithmes utilisés par ces deux assembleurs ne sont pas capables de calculer nos données spéciales. Les assemblages obtenus par Mira, Velvet et Omega, sont de qualité équivalente. Les résultats issus de Ray-Meta sont de bonne qualité par rapport à Omega cependant, les valeurs de N50 ne sont pas supérieures à celles issues de Velvet. IDBA-UD est prouvé que c'est le meilleur assembleur pour notre analyse puisque les valeurs de N50 sont beaucoup plus supérieures que les autres logiciels. En plus la qualité des contigs est bonne. Le nombre de contigs obtenus par l'assembleur IDBA-UD est de 48 (> 5kb), ce qui est supérieur au nombre de contigs créés par les autres assembleurs. Le N50 obtenu pour l'assemblage IDBA-UD est de 72 247 pb, signifiant que la moitié de l'assemblage final de IDBA-UD se trouve dans des contigs ayant une taille supérieure à 72 247 pb. D'après ces résultats, les contigs formés par IDBA-UD sont globalement de plus grande taille. La taille du plus grand contig *slac19_2* est de 152 721 pb contenant 141 CDS et celle du génome total est de 2 245 404 pb avec un contenu en GC% de 35.1%.

Le N50 est la taille du plus petit contig tel que 50 % de la longueur cumulée de l'ensemble des contigs obtenus après assemblage soit contenue dans des contigs de taille égale ou supérieure. En comparant avec les génomes de *L. lactis* ssp. *lactis* IL1403 (Bolotin et al. 2001), ce dernier contient 2 365 589 pb et code 2310 protéines, parmi elles 293 gènes

Résultats et discussion

codants sis prophages et 43 séquences d'éléments (IS). La distribution non non aléatoire de ces éléments IS indique que le chromosome de la souche pourrait être un produit d'une recombinaison récente entre deux génomes relativement proches. Un groupe complet de gènes de compétence naturelle est présent dans les deux souches *Lactococcus lactis* HD9B_GL2 et IL1403 ce qui indique aussi que ces deux souches peuvent être susceptibles d'un phénomène de transformation du matériel génétique. En comparant avec le génome de KLDS4.0325 (Yang et al. 2013). Complete genome accession no. CP_006766), cette souche contient un chromosome circulaire d'une taille de 2 589 261 pb et trois plasmides (plasmide 1 [5.7 kb], plasmide 2 [2.1 kb], et plasmide 3 [2.7 kb]). Le contenu génomique en G+C est de 35.4%, avec 2 662 ORFs (open reading frames) dont 1 310 sont fonctionnelles. 62 tRNA et 6 rRNA sont aussi présents. Le contenu plasmidique en G+C est de 34.7%, contenant 212 ORFs.

Tableau 16: Caractéristiques générales du génome de *Lactococcus lactis* HD9B_GL2

Caractéristique	Valeur
Taille du génome (bp)	2,251,485
Contenu en G+C (%)	35.1
Nombre de contigs	48
Nombre de gènes	2244
Nombre de CDS	2203
Nombre des tRNA	43
Nombre des rRNA	4

L'annotation génomique fonctionnelle a été réalisée par RAST 4 (Aziz et al. 2008) et résulte en 2203 de séquences codantes (CDS), y compris des séquences spécifiques à la souche, 43 tRNAs et 4 rRNAs. En addition, 39 de séquences RNAs non codantes sont annotées par Infernal/Rfam (Nawrocki and Eddy 2013). Une analyse plus approfondie du génome concernant les déterminants génétiques portant les gènes d'intérêt biotechnologique de la souche HD9B_GL2 va être étudiée en détail dans la section suivante.

En utilisant la méthode de maximum-likelihood, l'analyse phylogénétique montre que la souche la plus proche à notre souche *Lactococcus lactis* ssp *lactis* HD9B_GL2 est *L. lactis* subsp. *lactis* KLDS4.0325 (Yang et al. 2013) (Complete genome accession no. CP_006766),

Résultats et discussion

une bacteria probiotique isolée d'un produit traditionnel fermenté le "kumis" dans la région aride Xinjiang, en Chine et dont le génome total a été complètement séquencé. La deuxième souche proche de la HD9B_GL2 est la souche *L. lactis* ssp. *lactis* IL1403 (Bolotin et al. 2001) dont le génome est le premier génome séquencé des bacteries lactique (accession no. AE_005176). Les séquences génomiques des souches de Lactocoques ont été extraites á partir de leurs génomes, et alignées avec Dialign-TX (Subramanian et al. 2008), les résultats sont ensuite analysées par PhyML3 (Criscuolo 2011). L'arbre phylogénétique basé sur « *Maximum Likelihood method* » est montré dans la figure suivante.

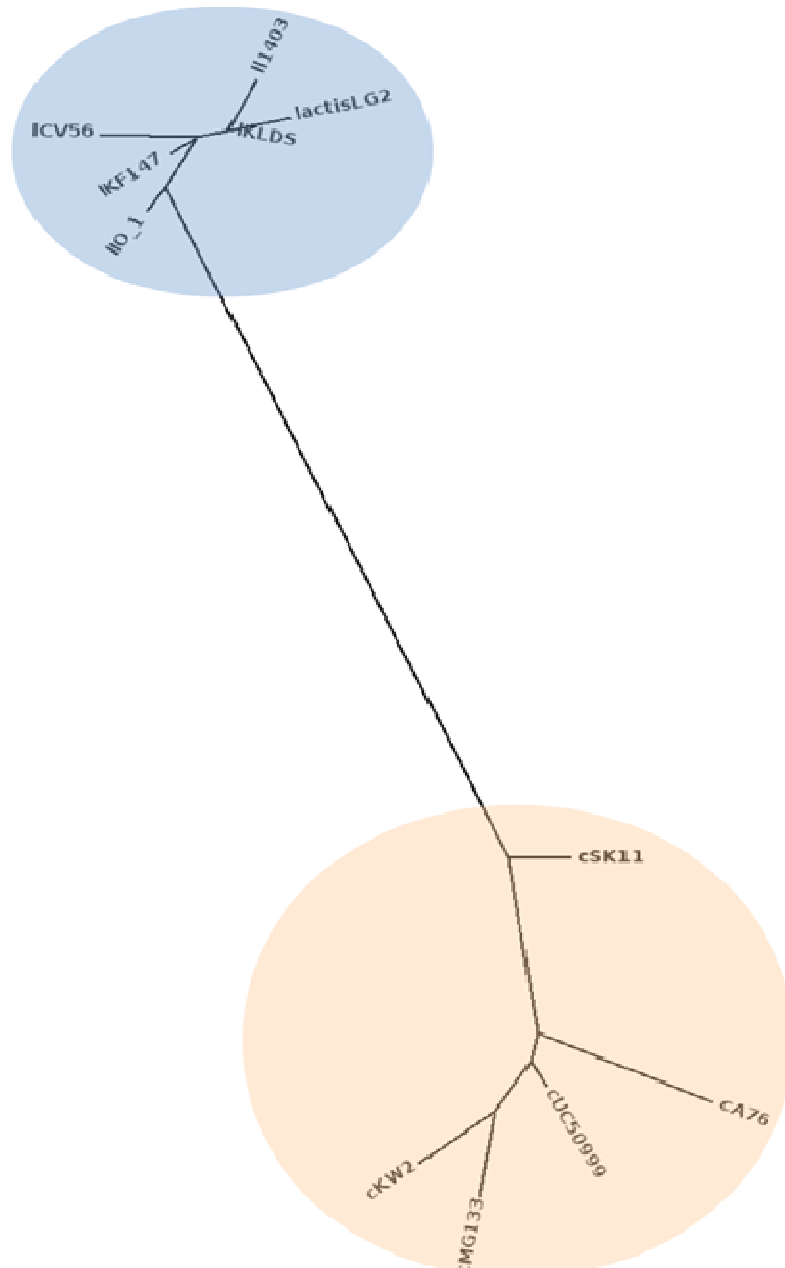


Figure 38 : Arbre phylogénétique basé sur les gènes marqueurs 16s RNA, *rpoA* et la protéine L2. Bleu: *Lactococcus lactis* ssp. *lactis* cluster, Orange: *Lactococcus lactis* ssp. *cremoris* cluster. Les gènes marqueurs hautement conservés sont 16s RNA, *rpoA* et L2. C'est clairement indiqué que la souche HD9B_GL2 est une *Lactococcus Lactis*. Cependant on peut clairement noter, que cette souche reste relativement distincte des autres souches lactocoques.

4.1. Phages et plasmides dans la HD9B_GL2

Il est important de noter, que selon les lactocoques, la taille des plasmides lactose varie de 30 à 100 kb (Neve et al. 1984). Chez les espèces du genre *Leuconostoc*, les gènes qui codent pour le métabolisme du lactose peuvent être chromosomiques ou plasmidiques.

L'analyse des résultats chez le lactocoque HD9B_GL2 montre la présence d'une région de 130 kb contenant un groupe de gènes présentant des homologies avec des protéines plasmidique parmi eux on trouve des gènes responsable du transport et l'hydrolyse du lactose et consiste en un opéron *lacABCDFEGX* et un gène *lacR*, qui sont transcrits de façon divergente. Le gène *lacR* code pour une protéine réprimant l'expression de l'opéron *lacABCDFEGX*. Les gènes *lacABC* et *lacFE*, codent respectivement pour les enzymes de la voie du tagatose-6-P et les protéines (IIAlac IIBC_{lac}) nécessaires au transport du lactose. Quant au gène *lacG*, il code pour la phospho-"-galactosidase qui assure l'hydrolyse du lactose. Enfin, le gène *lacX* code pour une protéine de 35 kDa de fonction inconnue et non essentielle à l'utilisation du lactose par la bactérie. En présence du glucose, l'expression de l'opéron *lac* est réprimée d'un facteur de dix (Dodd et al. 1992; Gasson et al. 1992; Griffin et al. 1992; van Rooijen et al. 1992). Plusieurs publications (Oskouian and Stewart 1990) indiquent que la séquence du "cluster" *lacR-lacABCDFEG* de *Lactococcus lactis* et celle de *Staphylococcus aureus* sont très similaires. Néanmoins, le gène *lacR* de *Staphylococcus aureus* est orienté dans le même sens que l'opéron *lacABCDFEG* (région dépourvue de plus du gène *lacX*).

L'analyse des résultats montre aussi la présence d'un groupe de gènes présentant des homologies avec des protéines phagiques, et notamment des protéines structurant la capsid virale ou des intégrases putatives (Figure 39 et 40). Bien que les scores de BLAST obtenus soient faibles pour certaines de ces protéines, d'après le contexte génomique il est plausible que ces gènes aient une origine phagique.

L'annotation des phages a été réalisée par le programme PHAST (PHAST, (Zhou et al. 2011), les résultats indiquent la présence de trois phages tempérés dans le génome de la HD9B_GL2 et correspondant à trois régions phagiques dont deux sont intactes et une région est contestable. Le premier phage est similaire au phage bIL286 (de la position 1 369 657 au 1 389 342) de 19686 pb. Il contient 21 CDS avec 34.42% en GC. Le deuxième est similaire au phage bII285 (du 1 536 161 au 1 555 153) d'une longueur de 18993 pb, et contient 23 CDS avec un contenu en GC% de 35.81%. Ces deux phages bII285 et bIL286 sont probablement de *Lactococcus lactis* ssp. *Lactis* IL1403. Le troisième phage a une longueur de 25 039 pb (du

Résultats et discussion

1 863 037 au 1 888 075) et contient 36 CDS avec un 35.57% GC, similaire au phage r1t de *Lactococcus lactis* ssp. *cremoris*.

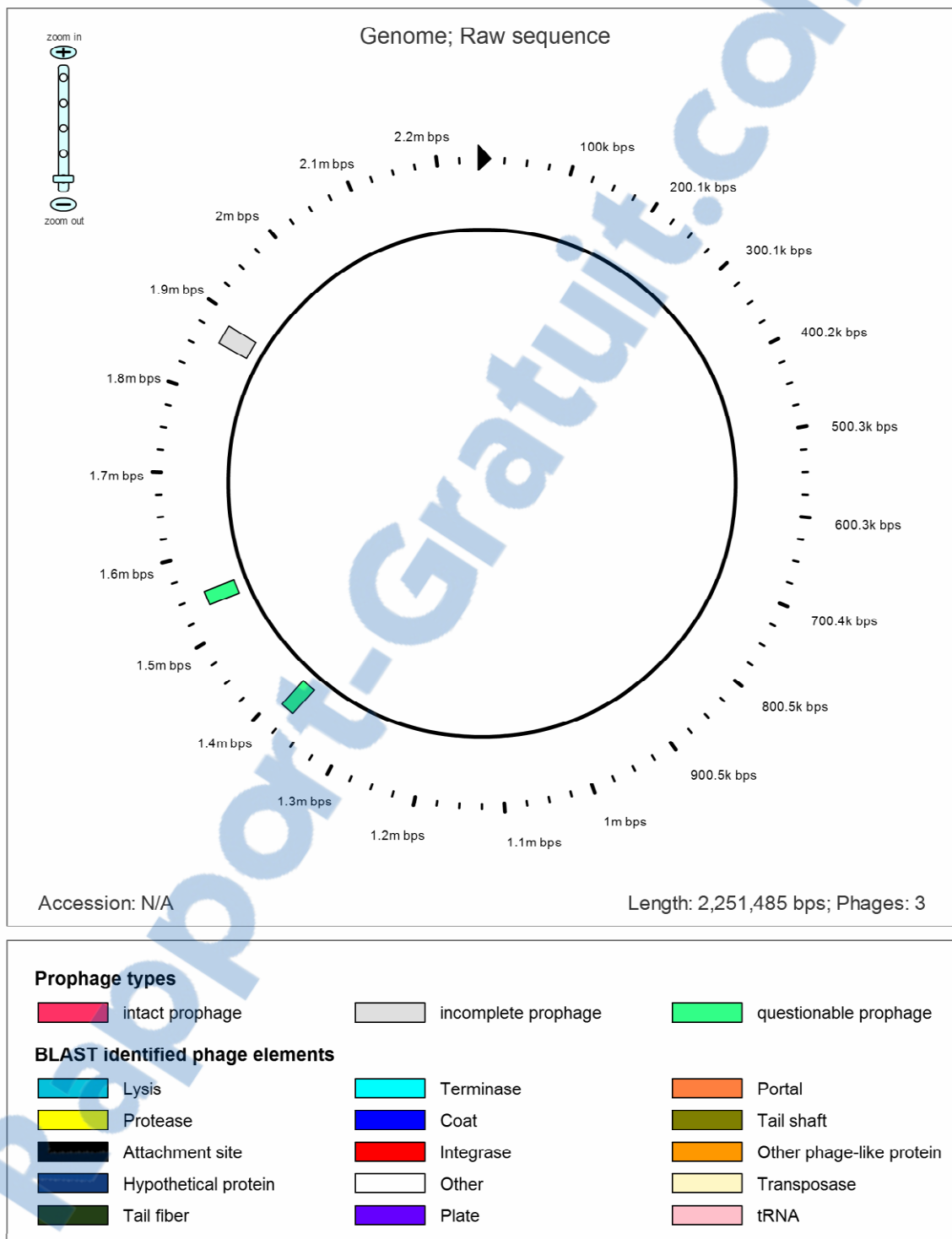


Figure 39 : Phages dans *L. lactis* ssp. *lactis* HD9B_GL2.

Résultats et discussion

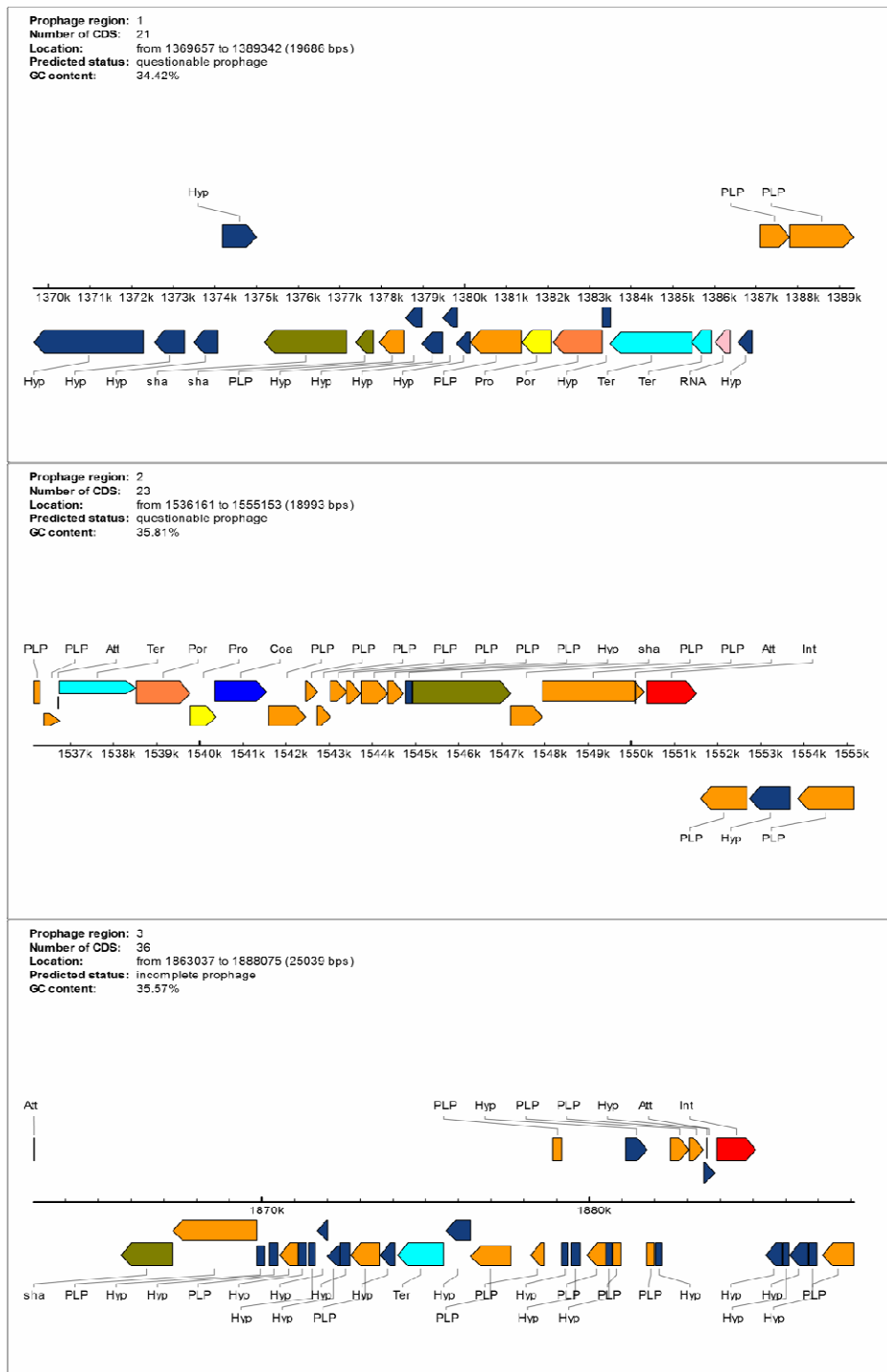


Figure 40 : Organisations génomique des phages tempérés de *L. lactis* ssp. *lactis* HD9B_GL2. L'échelle de la carte est en pb. Chaque fleche represente une ORF. Les fonctions putatives prédites de l'analyse bioinformatique et structurale sont indiquées au dessus des ORF.

Résultats et discussion

La connaissance de la génétique des bacteriophages chez les BL thermophyles est extrêmement nécessaire pour lutter contre les infections phagiques et qui présentent une grande menace sur le processus de la fermentation en industrie laitière. Chez les BL thermotolerants, la première séquence génomique complète obtenue est celle du phage LL-H de *Lactobacillus delbrueckii* subsp. *Lactis* signalée en 1996 (Mikkonen et al. 1996), suivie par celle du phage tempéré ØO1205 de *Streptococcus thermophilus* CNRZ1205 (Stanley et al. 1997).

Lactococcus lactis est largement utilisée pour produire une grande gamme de produits laitiers fermentés. Les phages virulents ubiquitaire en industries laitières sont la principale cause d'échecs de la fermentation du lait conduisant à des pertes économiques et compromettre la qualité du produit final (Madera et al. 2004). Plusieurs stratégies ont été développées pour contrôler les populations de phage dans cette niche écologique à densités élevées d'hôtes. Cependant, des nouveaux phages de lactocoques sont fréquemment isolés et qui peuvent persister pendant une longue période dans cet environnement laitier (Duplessis et al. 2005; Mc Grath et al. 2007). Leur taux d'évolution rapide et leur capacité d'adaptation élevée sont les deux critères responsables de leur vaste diversité observée chez les lactocoques (Labrie and Moineau 2007).

Les phages de lactocoques isolés jusqu'à présent appartiennent à l'ordre des *Caudovirales*, leurs génomes sont composés d'ADN double brin (dsDNA) pourvus d'une capsidie et une queue (Ackermann 1998). Ces phages appartiennent aussi à deux familles ; *Siphoviridae* (queue longue non contractiles) et *Podoviridae* (queue courte non contractiles) et ils sont divisés en 10 groupes génétiques différents. Cette classification est basée sur leur morphology et l'hybridation ADN-ADN (Deveau et al. 2006). Tous ces groupes de phages contiennent des phages virulents à l'exception du groupe P335, qui contient des phages virulents et tempérés. Seulement trois groupes de phage, 936, c2 et P335, sont régulièrement associés aux problèmes de fermentation du lait dans le monde entier (Christiansen et al. 1994; Bissonnette et al. 2000; Madera et al. 2004). Plus de 30 génomes de phages de lactocoques ont été complètement séquencés sont disponibles dans les bases de données publiques dont, douze phages proviennent du groupe 936, deux pour le groupe c2, douze sont du groupe P335 et le reste des phages sont issus de groupes de phages rarement isolés et sont nommés : KSY1, P087, Q54, 1358, 1706, et ascc_28 (groupe P034).

Résultats et discussion

L'Analyse de génomique comparative a confirmé la diversité phagique. Les membres de phages des groupes 936 et c2 sont similaires, tandis que le groupe P335 est considéré comme un groupe polyphylitique (Deveau et al. 2006; Rousseau and Moineau 2009). Le phage Q54 a acquis d'éléments génétiques de phages du type c2 et 936 (Fortier et al. 2006), tandis que KSY1 possède un système de transcription semblable à celui de T7-like phages (Chopin et al. 2007). Les phages de lactocoques P087 (Villion et al. 2009) et 1706 (Garneau et al. 2008) possèdent des similitudes avec ceux des autres firmicutes.

Le phage ascc_28 (P034) (Kotsonis et al. 2008) a une homologie avec les phages de *Streptococcus* et *Bacillus*. Enfin, le phage de lactocoques récemment caractérisée 1358 (Dupuis and Moineau 2010) possède des protéines phagique similaires à celles de *Listeria*. Ainsi, les membres de ces groupes de phages sont probablement issus d'une recombinaison entre différents phages des lactocoques et d'autres phages infectants d'autres bactéries (Dupuis and Moineau 2010). En fait, cette plasticité du génome phagique est probablement nécessaire pour s'adapter aux variations environnementales (Petri and Schmidt-Dannert 2004). Phages de lactocoques communs (comme les phages 936-like phages) sont hautement optimisés et s'adaptent rapidement au cours de la fermentation du lait à grande échelle. Il est supposé que la pression évolutive a conduit à la réplication et l'assemblage de particules phagiques pour permettre cette amplification rapide des phages (Ferguson and Coombs 2000). En revanche, les phages de lactocoques s'adaptent rarement dans cet écosystème laitier (Dupuis and Moineau 2010).

Les mécanismes de résistance naturelle des phages sont utilisés dans l'industrie laitière pour contrôler les populations phagiques, particulièrement les membres des trois principaux groupes de phages (936, c2 et P335) (Labrie et al. 2010). Le phage 949 est le plus grand phage de lactocoque connu jusqu'à présent, il a un génome de 114,768 bp avec un contenu en GC% (32.7%) un pourcentage faible par rapport au GC (35.4%) de son hôte *L. lactis* IL1403 (Bolotin et al. 2001).

Le premier genome phagique séquencé infectant *Lactococcus lactis* est (bIL67) en 1994 (Schouler et al. 1994). Plus que 700 phages de lactocoque isolés ont été cités dans la littérature (Ackermann and Kropinski 2007), ce qui indique que les phages infectants les lactocoques sont parmi les phages les plus étudiés à cause de l'intérêt biotechnologique de ces bactéries (Jarvis et al. 1991; Ackermann 2001; Deveau et al. 2006). Le phage virulent 1358 représentatif du groupe 9 de lactocoque phagique possède un génome à ADN double brin

Résultats et discussion

composé de 36,892 bp et un GC 51.04%. Son contenu en GC est beaucoup plus important que celui de son hôte *L. lactis* (35.3%) (Young and Blasi 1995; Yoong et al. 2004; Makarova et al. 2006) et tous les autres phages déjà caractérisés (Hejnowicz et al. 2009). Le genome à plus faible GC 33% est celui du phage 1706 (Garneau et al. 2008) et assc_28 (Kotsonis et al. 2008), tandis que le plus élevé est 37% chez KSY1 (Chopin et al. 2007).

4.2. Diversité de la robustesse chez *Lactococcus lactis* HD9B_GL2 (robustness related genes)

Lactococcus lactis est utilisée industriellement pour la fabrication de divers produits laitiers fermentés. La méthode la plus rentable pour la préservation des cultures *starter* de *L. lactis* est l'utilisation un spray de séchage, mais au cours de ce processus les cultures de lactocoques sont soumises au stress thermique et oxydative, ce qui entraîne généralement de faible taux de survie. Toutefois, la viabilité des ferments lactiques est essentielle pour leur contribution adéquate à la fermentation du lait, résultant en une amélioration de leurs phénotypes de robustesse.

Dans une étude récente, (Dijkstra et al. 2014) décrit une approche de transcriptome correspondant phénotype dans laquelle *L. lactis* MG1363 était cultivée sous une variété de conditions de différents niveaux d'oxygène et/ou de sel, ainsi que le pH de la fermentation et la température. Dérivé de ces fermentations dans la phase exponentielle de croissance bactérienne des échantillons ont été analysés par transcriptomique complète du génome et l'évaluation de la chaleur et des phénotypes de stress oxydatif. Les bactéries en phase exponentielle de croissance ont été analysées par une étude transcriptomique du génome total et les différents variants sont ensuite évalués pour leurs nouveaux phénotypes de température et de stress oxydatif.

Les variations des conditions de la fermentation ont abouti à une augmentation de la survie 1000 fois plus durant le stress hyperthermique et oxydatif. Plus précisément, l'aération pendant la fermentation induit par la protection contre le stress thermique, considérant qu'une température relativement élevée au cours du processus de la fermentation a donné lieu à une robustesse accrue vis-à-vis le stress oxydatif. Parallèlement, l'oxygène et la température induisaient l'expression différentielle de plusieurs gènes en comparaison avec la fermentation sous de conditions favorables. Analyse de corrélation des phénotypes de la robustesse et des niveaux d'expression des gènes a révélé la présence des signatures transcriptomiques pour les variants survécus du stress oxydatif et thermique, y compris l'expression de l'opéron *metC-cysK* impliqué dans le métabolisme de la méthionine et la cystéine. Pour vérifier cette

Résultats et discussion

association transcriptome-phénotype, *L. lactis* MG1363 était cultivée en l'absence de cystéine qui a conduit à une robustesse accrue vers le stress oxydatif.

En général, Dijkstra et al ont démontré l'importance d'une sélection rigoureuse des paramètres de fermentation avant la transformation industrielle des ferments lactiques. En outre, des gènes établis et impliqués en stress, de nouveaux gènes ont été associés à la robustesse vis-à-vis la chaleur et le stress oxydatif. L'évaluation des niveaux d'expression de ce groupe de gènes pourrait servir d'indicateur pour la sélection des paramètres de fermentation résultant en robustesse améliorée pendant le séchage par pulvérisation. La robustesse accrue après une croissance sans cystéine semble confirmer le rôle de l'expression de l'opéron *metC-cysK* comme un indicateur de solidité et de robustesse et suggère que le métabolisme de l'acide aminés « soufre » joue un rôle essentiel dans la survie sous un stress oxydatif.

Il a été rapporté que les gènes *metC* et *cysK*, codant pour une cystathionine γ -synthase/cystathionine β -lyase et un O-acetylserine sulfhydrylase, respectivement ont été associés à la survie en condition de stress hyperthermique et oxydatif. Ces gènes sont co-transcrits (Fernandez et al. 2000), et que leur opéron est induit en absence du soufre, en particulier en cas de faibles concentration en cystéine, résultant en une augmentation dans l'activité de cystathionine β -lyase (Fernandez et al. 2002).

L'ensemble de ces gènes de robustesse vis-à-vis un stress thermique a été trouvé dans le génome de notre souche *L. lactis* ssp. *lactis* HD9B_GL2 et ils sont présentés dans le tableau 17.

Résultats et discussion

Tableau 17: Gènes associés à la robustesse sous un stress thermique chez *L. lactis* ssp. *lactis* HD9B_GL2

Locus_tag	Fonction
GL2_2237	Hypothetical protein
GL2_2238	GntR family transcriptional regulator
GL2_2289	Hypothetical protein
GL2_2291	Manganese ABC transporter permease
GL2_2425	Sugar ABC transport
GL2_2430	Hypothetical protein
GL2_3330	Beta-glucosidase
GL2_2235	Hypothetical protein
GL2_2654	ABC transporter ATP-binding protein
GL2_1359	Sugar ABC transporter permease
GL2_0021	O-acetylserine sulfhydrylase
GL2_0013	Cysteine synthase (EC 2.5.1.47)
GL2_16	Cysteine desulfurase (EC 2.8.1.7)
GL2_0021	ADA regulatory protein / Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)
GL2_0022	Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)
GL2_0406	S-ribosylhomocysteine lyase (EC 4.4.1.21) / Autoinducer-2 production protein LuxS
GL2_0573	tRNA-Cys-GCA
GL2_0021	Cysteine ABC transporter, substrate-binding protein
GL2_1868	Cysteine ABC transporter, permease protein
GL2_1869	Cystine transport ATP-binding protein
GL2_1035	CysteinyI-tRNA synthetase related protein
GL2_1036	Cystathionine beta-lyase (EC 4.4.1.8)
GL2_1037	5-methyltetrahydropteroyl triglutamate--homocysteine methyltransferase (EC 2.1.1.14)
GL2_1070	Cysteine desulfurase (EC 2.8.1.7), SufS subfamily
GL2_1339	CysteinyI-tRNA synthetase (EC 6.1.1.16)
GL2_1466	5'-methylthioadenosine nucleosidase (EC 3.2.2.16) / S adenosylhomocysteine nucleosidase (EC 3.2.2.9)
GL2_2060	Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)
GL2_2184	Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)
GL2_2192	Cystathionine gamma-synthase (EC 2.5.1.48)
GL2_2215	Cysteine ABC transporter, substrate-binding protein

4.3. Analyse de l'organisation génétique des gènes de choc thermique chez *L. lactis* ssp. *lactis* HD9B_GL2

Chaque fois qu'un service d'annotations répète une erreur systématique, les résultats sont déposés dans la base de données NCBI. Cependant, comme les nouvelles annotations sont produites par le même service, les erreurs précédemment déposées sont utilisées pour valider l'annotation la plus récente, qui contient les mêmes erreurs systématiques. En conséquence, les erreurs systématiques sont utilisées pour valider la répétition des mêmes erreurs, et les bases de données accumulent ces annotations erronées. Pour l'analyse de la séquence génomique des gènes de choc thermique chez *L. lactis* ssp. *Lactis* HD9B_GL2, nous avons identifié manuellement l'organisation génétique de ces gènes: leurs promoteurs, le nombre des ORF, les séquences codantes (CDS), le site de liaison consensus au ribosome spécifique à l'espèce (RBS), les codons start/stop, les séquences inversées répétées (IR) et les terminateurs.

Clustal Omega est utilisé pour les alignements multiples des séquences (Sievers et al. 2011), et est disponible en ligne: <http://www.ebi.ac.uk/Tools/msa/clustalo/>. Les comparaisons de séquences sont réalisées par blastn et blastp disponibles dans la base de données GenBank (Altschul et al. 1997), en utilisant des paramètres par défaut. La prédiction et la visualisation des promoteurs bactériens est effectuée par le logiciel BPROM (Solovyev and Salamov 2011) disponible sur le site : <http://linux1.softberry.com/berry.phtml?topic=bprom&group=help&subgroup=gfindb> et en utilisant aussi PePPER (de Jong et al. 2013) disponible à : <http://server.molgenrug.nl/index.php/prokaryote-promoters>. Le programme ARNold (Naville et al. 2011) est utilisé pour la recherche des terminateurs Rho-independent et est disponible au : <http://rna.igmors.u-psud.fr/toolbox/arnold/>. La prédiction de la structure secondaire d'ARN est faite sur le serveur mfold (Zuker 2003) au : [http://unafold.rna.albany.edu/?q=mfold, using default parameters](http://unafold.rna.albany.edu/?q=mfold,usingdefaultparameters).

Rappelons que la séquence génomique de *Lactococcus lactis* ssp. *lactis* HD9B_GL2 a été déposée au GenBank/EMBL/DDBJ sous le numéro d'accèsion no. JNCC02000000. Les locus tags cités ici sont référés aux numéros accessibles et présents dans le projet d'assemblage et d'annotation de *Lactococcus lactis* ssp. *lactis* HD9B_GL2 dans NCBI sous le lien : http://www.ncbi.nlm.nih.gov/genome/proteins/156?genome_assembly_id=248586. Les souches de références utilisées sont : *L. lactis* ssp. *lactis* IL1403 (abbr. IL1403, accession

Résultats et discussion

no. AE005176), *L. lactis* ssp. *lactis* NCDO2118 (abbr. NCDO2118, accession no. NZ_CP009054.1), *L. lactis* ssp. *lactis* KLDS4.0325 (abbr. KLDS4.0325, accession no. CP006766), *L. lactis* subsp. *cremoris* MG1363 (abbr. MG1363, accession no. NC_009004.1) et *L. lactis* subsp. *cremoris* SK11 (abbr. SK11, accession no. CP_006766.1).

Chez les bactéries à Gram positive, la réponse au stress hyperthermique fait intervenir deux familles d'acteurs principaux: les protéines chaperons et les complexes multiprotéiques Clp. Chez *Lactococcus lactis* ssp. *lactis* HD9B_GL2 l'analyse in silico du génome permet de retrouver ces deux grandes familles de protéines mais aussi quelques autres protéines référencées comme «*Heat Shock Protein*» (HSP) chez d'autres micro-organismes. En terme de fonctions potentiellement associées au stress thermique chez *Lactococcus lactis*, le bilan est le suivant.

La résistance bactérienne vis-à-vis le stress physique telle que la thermotolérance, est un phénotype complexe et partiellement connu. Nous réalisons une analyse détaillée et comparative de la séquence génomique de la souche modérément thermotolerante *Lactococcus lactis* ssp. *lactis* GL2 afin de déterminer les gènes ainsi que les protéines impliqués dans sa résistance aux températures élevées. Trois souches de références complètement séquencées *L. lactis* ssp. *lactis* et deux souches appartenant à la ssp. *cremoris* sont incluses dans l'analyse comparative. Les gènes/protéines étudiés comprennent i) les protéines du choc thermique classe I (DnaK, DnaJ, GroESL, GrpE) et leur répresseur transcriptionnel HrcA, ii) les protéines du choc thermique classe III (ClpB, ClpC, ClpE, ClpP, ClpX) et leur régulateur négatif CtsR. Autres gènes/protéines visés incluent iii) RecA, pour laquelle, nous avons identifié exceptionnellement deux gènes différents, iv) FtsH et LexA-like régulateur HdrI, qui interagissent avec RecA, et v) HtrA, GdpP et le système à deux composante CesSR. L'ensemble des gènes de choc thermique chez *Lactococcus lactis* ssp. *lactis* HD9B_GL2 sont présentés dans le tableau suivant.

Résultats et discussion

Tableau 18: Gènes de choc thermique présents chez *Lactococcus lactis*. ssp. *lactis* HD9B_GL2

Nom	Fonction	No.dans GL2-HD9B	Locus tag GenBank
HrcA	Transcriptional regulator	GL2_1052	LGL2_RS04340
GroES	Chaperone	GL2_0545	LGL2_RS01980
GroEL	Chaperone	GL2_0546	LGL2_RS01985
GrpE	Chaperone	GL2_1053	LGL2_RS04345
DnaK	Chaperone	GL2_1054	LGL2_RS04350
DnaJ	Chaperone	GL2_2717	LGL2_RS10750
CtsR	Transcriptional repressor	GL2_0708	LGL2_RS02900
ClpC	Clp protease ATP-binding subunit	GL2_0709	LGL2_RS02905
ClpE	Clp protease ATP-binding subunit	GL2_2254	LGL2_RS02485
ClpP	Clp protease proteolytic subunit	GL2_0726	LGL2_RS02990
ClpX	Clp protease ATP-binding subunit	GL2_1306	LGL2_RS05120
ClpB	Clp protease	GL2_1736	LGL2_RS06880
RecA1	Recombination DNA repair	GL2_1592	LGL2_RS01820
RecA2	Recombination DNA repair	GL2_0609	LGL2_RS05750
FtsH	Protease	GL2_0156	LGL2_RS00115
HdiR	LexA-like transcriptional regulator	GL2_1479	LGL2_RS05910
HtrA	Protease	GL2_2621	LGL2_RS10295
GdpP	Signaling protein	GL2_0787	LGL2_RS03375
CesS	Two-component sensor His-kinase	GL2_0962	LGL2_RS04025
CesR	Two-component response regulator	GL2_0963	LGL2_RS04030
Hsp33	Chaperone	GL2_2318	LGL2_RS09300
Trigger Factor	Chaperone	GL2_0002	LGL2_RS02410

A l'exception de RecA2, qui est présente chez KLDS4.0325 seulement, toutes les protéines ont des homologues identiques ou très similaires dans toutes les souches de références (*L. lactis* ssp. *lactis* IL1403, KLDS4.0325, NCDO 2118 et *L. lactis* ssp. *cremoris* MG1363, SK11)

Résultats et discussion

L'annotation de la séquence génomique a révélé la présence des gènes complets et des protéines telles que : HrcA (locus tag LGL2_RS04340), GrpE (locus tag LGL2_RS04345), DnaK (locus tag LGL2_RS04350), GroES (locus tag LGL2_RS01980), GroEL (locus tag LGL2_RS01985) et DnaJ (locus tag LGL2_RS10750). Leur arrangement sur le chromosome ainsi que leur organisation transcriptionnelle sont identiques aux résultats précédemment publiés ; aucune séquence promotrice n'a été trouvée dans les petites régions intergéniques entre le gène *hrcA* et le gène *grpE* (31 bp), et entre *grpE* et *dnaK* (71 bp) (Eaton et al. 1993) ou entre les deux gènes *groES* et *groEL* (87 bp) (Kim and Batt 1993). On ne sait pas si *hrcA*, *grpE* et *dnaK* sont transcrits comme un mRNA polycistronique chez *L. lactis* tel que le cas chez *B. subtilis* (Homuth et al. 1999). Contrairement au *B. subtilis* (Homuth et al. 1999), il a été montré que chez d'autres souches de *L. lactis* (van Asseldonk et al. 1993), *dnaJ* ne fait pas partie de l'operon *dnaK* et qu'un terminateur transcriptionnel Rho-indépendant a été trouvé en aval du gène *dnaK*. Les différentes séquences codantes ainsi que les séquences protéiques sont comparées aux séquences correspondantes chez les cinq souches de références. Dans tous les cas on a trouvé une identité de 99% avec les références de ssp. *lactis* et de 96 à 98% d'homologie de séquence avec la ssp. *cremoris* (données non mentionnées).

Comme dans le cas chez *B. subtilis* (Zuber and Schumann 1994), les séquences inverses répétées IR ont été trouvées dans *L. lactis* HD9B_GL2 avec une grande similarité significative de l'élément CIRCE (*controlling IR of chaperone expression*) qui la présence d'une séquence consensus 5'-TTAGCACTC-N9-GAGTGCTAA-3', en amont des gènes codants pour les protéines HrcA, GroES et DnaJ (Figure. 1 A). Chez *B. subtilis* on note une interaction directe de HrcA/l'élément CIRCE et la répression transcriptionnelle des gènes en aval du gène *hrcA* a été montrée expérimentalement (Reischl et al. 2002). Il a été apporté aussi que la présence de l'élément CIRCE en amont du gène *dnaJ* est nécessaire pour l'activité de la protéine répresseur HrcA malgré sa présence en amont du promoteur (van Asseldonk et al. 1993). Il est aussi intéressant de noter la présence d'une deuxième séquence inversée répétée en amont du promoteur de *hrcA* (figure 41A), un tel cas n'a jamais été signalé auparavant, et on ne sait pas le rôle exact de ce motif. Bien que les séquences annotées d'acides aminés de GrpE, DnaK, GroES, GroEL et DnaJ ne montrent aucune différence en comparant avec les souches de références, la séquence d'acides aminés prédite pour la protéine HrcA chez *L. lactis* HD9B_GL2 et la souche de référence KLDS4.0325 a montré une longue protéine de 385 acide aminé, c'est-à-dire que cette protéine possède une extension de 38 acide aminé dans sa partie N-terminal, une protéine plus longue par rapport à celles chez

Résultats et discussion

les souches de références de *L. lactis* et chez d'autres firmicutes, par exemple, *S. mutans* et *B. subtilis* (figure 41B). Cela met en question, si ce résultat est issu d'une erreur d'annotation dans la traduction du premier codant start de la protéine HrcA, ou *L. l* HD9B_GL2 et KLDS4.0325 produisent effectivement une telle protéine plus longue. Depuis l'avènement des pipelines d'annotation automatique du génome entier il y'a est une préoccupation importante concernant l'identification correcte des codons start alternatifs de ATG par exemple GTG et TTG (Villegas and Kropinski 2008; DeJesus et al. 2013).

Le logiciel utilisé dans l'annotation de la séquence génomique de la souche HD9B_GL2 est le RAST (Overbeek et al. 2014), ce logiciel a une tendance d'identifier des codons alternatifs pour le codons start ATG comme premier sites d'initiation (Bakke et al. 2009). La figure. 1C montre la partie N-terminale de la protein prédite par RAST montrant le GTG (pos. 111-113) comme codon start. Cependant, ce n'est pas clair s'il s'agit d'une erreur dans l'annotation : d'une part, la partie N-terminale étendue de la protéine HrcA a été déjà identifiée chez d'autres bactéries (Wiegert and Schumann 2003), d'autre part, une séquence présente proche de la séquence génomique en amont du premier codon ATG (pos. 225-227 séquence de ssp. *lactis*) et dans le même cadre de lecture montre des caractéristiques particulières.

Le promoteur du gène *hrcA* est dans la position indiquée clairement dans la figure. 41 C, il n'y à aucune boîte potentielle -35/-10 à l'intérieur ou en aval de IR2. On peut supposer que le site d'initiation transcriptionnel est à quelques paires de bases en aval de la boîte - 10, ceci pourrait expliquer la présence d'un ou même deux sites putatifs de liaison de ribosomes désignés RBS3 et RBS2 (Figure. 41C) dans l'ARNm transcript en amont du codon start GTG, le troisième (RBS1) est en amont du codon ATG codon. La situation est très différente de la région correspondante chez les souches de référence ssp. *cremoris*. Dans toutes ces souches, il y'à une grande délétion de 73 paires de bases dans la région entre IR2 et le codon ATG (pos. 152-154 dans la séquence de ssp. *cremoris*) qui a apporté le codon GTG hors du cadre de lecture et a créé un codon stop (TAG) dans la partie 5'-fin de IR2.

En outre, il n'y a aucuns codons start dans le cadre de lecture (ni ATG ni GTG ou TTG) entre RBS1 et le codon stop TAG dans cette séquence, alors que chez ssp. *cremoris* il n'y à qu'un seul codon start ATG (152-154) et seulement une longue protéine unique HrcA de 347 acides aminés peut être synthétisée. De même, chez *S. mutans* UA159 une seule protéine HrcA canonique (voir figure. 41 B) peut être synthétisée, avec GTG comme codon start (Figure. 41D), une situation semblable se trouve chez *b. subtilis* 168 (non illustré). Étonnamment, la séquence nucléotidique de 227 paires de bases de l'extrémité 5' coté ATG jusqu'au codon

Résultats et discussion

stop TAA du gène précédant s'est avéré absolument être conservée chez toutes les souches de *ssp. lactis* testées, y compris les trois souches de référence. Par conséquent, toutes ces souches de la sous-espèce *lactis* seraient en mesure de produire une longue protéine HrcA d'une extrémité N-terminale étendue de 385 acides aminés ou le transcrit du *hrcA* possède exceptionnellement dans son extrémité 5' une séquence longue leader non traduite.

La conservation de cette séquence dans tous les génomes de *ssp.lactis* testés pointe vers une fonction importante pour cette dernière. Figure. 41E montre que, selon les modèles générés avec le programme mfold (Zuker 2003), l'ARN transcrit de la séquence du IR2 au codon start ATG serait en mesure de former une structure secondaire stable ($\Delta G = -159.9 \text{ kJ mol}^{-1}$).

Dans cette structure hypothétique RBS3 serait librement accessible, RBS2 pourrait être partiellement masqués et l'anticodon GUG serait au début d'une structure en boucle « stem » comprenant IR2 (élément CIRCE). RBS1 et l'anticodon AUG serait accessibles. Chez *B. subtilis*, il a été démontré qu'une structure en « épingle à cheveux » présente en amont du *hrcA* a une influence sur la stabilité du transcrit lorsqu'il se trouve à proximité du RBS (Homuth et al. 1999), cela serait le cas pour RBS3 et en particulier, pour RBS2.

La région la plus courte et correspondante au *hrcA* chez la sous-espèce *cremoris* pourrait constituer une structure secondaire assez simple ($\Delta G = -54.4 \text{ kJ mol}^{-1}$) dans laquelle les RBS et l'anticodon AUG seraient accessibles et éloignés de la structure « épingle à cheveux » de IR2 (Figure. 41E). Depuis quelques années, il est devenu évident que les structures secondaires des séquences leaders non-traduites à l'extrémité 5' chez les bactéries peuvent agir comme des ARN thermosenseurs (RNA thermomètres, RNATs), des éléments cis-régulateurs de l'expression génétique des gènes sous la température (récemment évaluée par Krajewski et Narberhaus 2014(Krajewski and Narberhaus 2014)).

Dans la plupart des RNATs, la structure de RBS et du codon d'initiation de la traduction sont séquestrés dans une région sous forme d'un « épingle à cheveux » qui se dissout progressivement lorsque la température augmente, ce qui permet une traduction efficace de la séquence codante en aval.

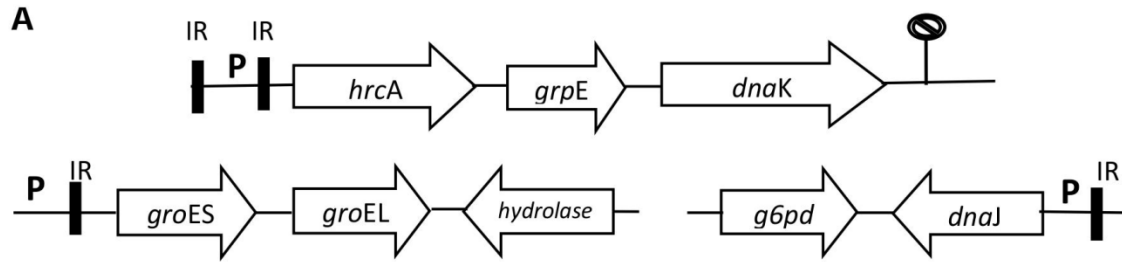
La structure secondaire putative de la protéine « leader » *hrcA* de *ssp. lactis* a une similarité assez distante aux RNATs connus. Il reste donc hypothétique si cette région pourrait agir comme un nouveau élément régulateur, permettant ; l'initiation de la traduction à partir du codon AUG et synthèse d'une longue protéine canonique (347 acides aminés) à des

Résultats et discussion

températures basses ; et une initiation de la traduction à partir du codon à des températures élevées. Pour répondre à ces questions, des expériences telles que la construction de mutans et des analyses protéiques plus sophistiquées sont nécessaires.

Figure 41: A. Organisation transcriptionnelle des opérons *hrcA-grpE-dnaK*, *groES-groEL* et *dnaJ* chez *L. lactis* HD9B_GL2. P: Promoteur; IR: Inverted repeat (element CIRCE); Epingle à cheveux: terminateur Rho-independent. B: Alignement de sequence des acides aminés (Clustal Omega) de l'extrémité N-terminale de la protéine HrcA chez *L. lactis ssp. lactis* et *ssp. cremoris*, *Streptococcus mutans* UA159 et chez *B. subtilis*168 (NCBI). C. Région 5'-en amont du gene *hrcA* chez *L. lactis ssp. lactis* (HD9B_GL2 et trois souches de référence, voir matériels et méthodes) et chez *ssp. cremoris* (MG1363 et SK11). IR: inverted repeats (éléments CIRCE, en italique et souligné), -35 / -10 sont les séquences promotrices. Les sites de fixation des ribosomes (RBS) sont en gras et soulignés, les codons stop (TGA, TAG) et les codons starts (GTG, ATG) sont surlignés. D. Région 5'-en amont du gène *hrcA* chez *Streptococcus mutans* UA159 (GenBank accession no. AE014133.2). Pour explications voir C. E. Structures secondaires predates de la sequence UTR non traduite "leader" de *hrcA* dans *L. lactis ssp. lactis* (137 nt, à gauche) et chez *ssp. cremoris* (64 nt, à droit). Les predictions sont réalisées sur le programme mfold (voir matériels et méthodes). RBS1-3, IR2, Start1-2 sont décrits en C. Les valeurs d'énergie ΔG sont présentées sous les structures correspondantes.

Résultats et discussion



B

<i>L.l.</i> GL2, KLDS	MLKQKKLVDTKQKSKAQEIFELAKEVLPNFLSVEKGVVMI	TERQRQILNLIVSLYA	385aa
<i>L.l.</i> IL1403 etc.	-----	MI TERQRQILNLIVSLYA	347aa
<i>L.l. ssp. cremoris</i>	-----	MI TERQRQILNLIVSLYA	347aa
<i>S. mutans</i> UA159	-----	MITQRQKDILNLIVLFT	347aa
<i>B. subtilis</i> 168	-----	MLTNRQLLILQVIINDFI	343aa

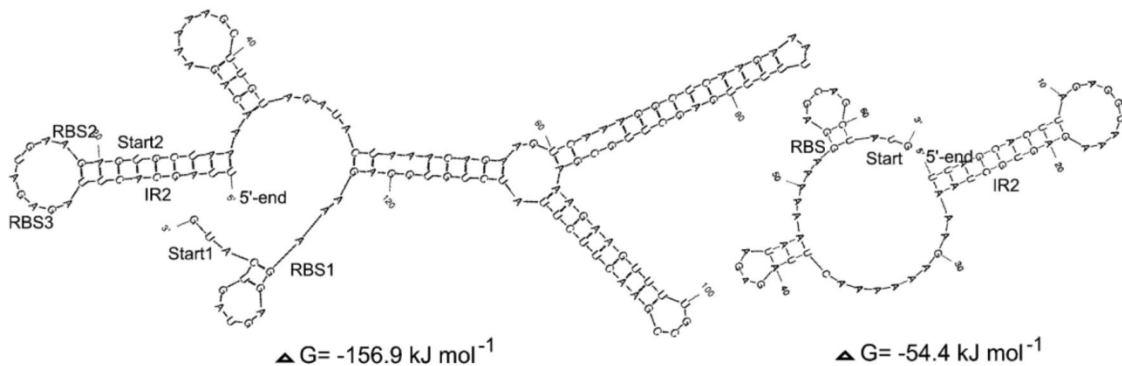
C

		IR1 >	< IR1		-35	
<i>ssp. lactis</i>	A	<u>TTAA</u> ACGAAA	<u>TTAGCACTC</u> TAAACAAAAA	<u>GAGTGCTAA</u> ATTTTTATGCTTTTTTGCT	<u>TTGACA</u>	60
<i>ssp. cremoris</i>	A	<u>TTAA</u> CAGAAA	<u>TTAGCACTC</u> TAATAAAAAA	<u>GAGTGCTAA</u> ATTTTTATGCTTTTTTCT	<u>TTGACA</u>	60
			-10	IR2 >	RBS3	RBS2 < IR2
<i>ssp. lactis</i>	AGGTATAAACAAAGGGT	<u>TATAAT</u> TAAATTTG	<u>TTAGCACTT</u>	<u>AGAGATCA</u>	<u>AGAGCTC</u>	<u>CTAA</u> AAC
					Stop V/Met	L K
<i>ssp. cremoris</i>	AGGTATAAACAAAGGGT	<u>TATAAT</u> TAAACTTG	<u>TTAGCACTT</u>	TAGAGGTAAAG	<u>AGTGCTAA</u>	AAG
					Stop	
<i>ssp. lactis</i>	AGAAAAAGCTTGTAGATACTAAACAGAAGTCAAAGGCTCAAGAAATTTTTGAGCTTGCGA					180
	Q K K L V D T K Q K S K A Q E I F E L A					
<i>ssp. cremoris</i>	A--AAAACTTAGAGATA-----					136
				RBS1		
<i>ssp. lactis</i>	AAGAAGTTTTGCCGAACTTCTTATCTGTGGAGA	<u>AAAGGA</u>	AGTAGTC	<u>CATG</u>	AATTACAGAACGTC	240
	K E V L P N F L S V E K G V V			Met	I T E R	
<i>ssp. cremoris</i>	-----	AAAA	<u>AAAGGAGCAGCT</u>	<u>TATG</u>	AATTACAGAACGTC	167
			RBS1	Met	I T E R	

D

	-35		-10		IR >
<i>S. mutans</i> UA159	CTCT	<u>TTGACAG</u> AAATTAGGAATGGAGT	<u>TATAAT</u> GGAAATTATGAAG	<u>TTAGCAGTC</u>	CGTCTGTTA
	< IR		RBS	Stop	Start <i>hrcA</i>
		<u>GAGTGCTAA</u>	CTCTATTGATTAAAGATGGGAGG	TGGACT	<u>TAGCTC</u> AATTACCCAGCGTCAAA

E



Résultats et discussion

Concernant les gènes et les protéines de choc thermique de la classe III, on a pu identifier les gènes complets pour ClpB (locus tag LGL2_RS06880), ClpC (locus tag LGL2_RS02905), ClpE (locus tag LGL2_RS02485) (Ingmer et al. 1999); ClpP (locus tag LGL2_RS02990) (Frees and Ingmer 1999), ClpX (locus tag LGL2_RS05120) ainsi que la protéine represser CtsR (locus tag LGL2_RS02900) (Varmanen et al. 2000); pour voir les fonctions consulter (Frees et al. 2007).

L'organisation transcriptionnelle du *ctsR* et *clpC* est montrée dans la figure. 42A. Un site de liaison CtsR-binding site, consistant en deux courtes répétitions dans la même orientation, a été trouvé en amont des gènes *ctsR*, *clpB*, *clpE* et *clpP* comme a été décrit précédemment (Varmanen et al. 2000) (Figure. 42C, données non illustrées pour les gènes de Clp). Des homologues pour tous ces gènes, ont été identifiés dans les génomes des souches de référence (non illustré). En ce qui concerne la protéine CtsR, une situation a été observée similaire à celle décrite ci-dessus pour HrcA. La figure. 42 B montre l'annotation d'une protéine prédite de 159 acides aminés chez HD9B_GL2 et KLDS4.0325, alors que pour tous les souches de référence de *L. lactis* ssp. *lactis* et ssp. *cremoris*, une protéine plus longue de 151 acides aminés a été annotée pour CstR. Dans ce cas, une analyse de la région en amont de la séquence codante (Figure. 42C) a révélé que, chez toutes les souches deux codons ATG d'initiation de la traduction sont présents. Les deux codons start sont précédés par un site de liaison "ribosome binding site », entre RBS1 et l'ATG on trouve un codon stop TAG dans le même cadre ouvert de lecture. Ainsi qu'un séquençage de l'extrémité N-terminal de la protéine CtsR pourrait préciser le codon start ATG utilisé in vivo.

Une recherche sur la protéine RecA dans HD9B_GL2 a donné des résultats inattendus et intéressants. Cette souche code pour deux protéines de RecA : RecA1 est de 387 acides aminés (locus tag LGL2_RS01820) et RecA2 avec 346 acides aminés (locus tag LGL2_RS05750). Ces deux protéines montrent 46% et 67% d'homologie de séquences d'acides aminés dans la région du 19 à 342 ainsi que tous les domaines conservés caractéristiques de RecA (Marchler-Bauer et al. 2015) ; par exemple ; les deux protéines RecA1 et RecA2 contiennent les motifs Walker A et Walker B (Walker et al. 1982) à des positions identiques (Figure. 42D). A partir de ces résultats, nous avons conclu que ces deux protéines sont fonctionnelles. L'analyse comparative a révélé qu'il n'y a pas de corrélation entre les sous-espèces et la présence de RecA2 (Figure. 42E).

Résultats et discussion

La présence plus d'une seule copie du gène *recA* a été décrite jusqu'à présent chez quelques autres bactéries seulement, par exemple chez *B. megaterium* (Nahrstedt et al. 2005) et chez la bactérie du désert *Deinococcus deserti* (Dulermo et al. 2009). Dans les deux cas, la présence de deux protéines RecAs chez la bactérie, a été impliquée dans la résistance vis-à-vis la dégradation de l'ADN par les rayons UV. Un rôle similaire a été montré pour la protéine RecA de *L. lactis* (Dywat et al. 1995), tandis que la souche MG1363 de *L. lactis* ssp. *cremoris* utilisée comme référence ne possède qu'une seule protéine RecA1 (Figure. 42 E).

L'étude a montré aussi que RecA1 n'est pas impliquée uniquement dans la recombinaison homologue et la réparation de l'ADN mais aussi elle est impliquée dans la réponse aux stress oxydatif et thermique via l'interaction avec la protéase FtsH (anciennement appelé HflB). Une étude antérieure avait déjà montré le rôle de FtsH dans la thermotolérance (Nilsson et al. 1994). Nous avons identifié des gènes pour la protéine complète de FtsH chez HD9B_GL2 (locus tag LGL2_RS00115) ainsi que chez les souches de référence.

Une autre protéine a été trouvée et capable d'interagir avec RecA et ClpP est HdiR, le LexA-like régulateur négatif transcriptionnel (Savijoki et al. 2003). La protéine correspondante dans la souche HD9B_GL2 (locus tag LGL2_RS05910) a 99% d'identité avec les protéines orthologues chez les souches de référence ssp. *lactis* et 91% d'identité avec les souches de ssp. *cremoris*.

Enfin, nous avons recherché la présence de quatre autres gènes/protéines qui ont été impliqués dans la résistance au stress, en particulier dans la thermotolérance de *L. lactis*. La protéase de surface « de ménage » HtrA joue un rôle dans la défense contre des conditions de stress (Foucaud-Scheunemann and Poquet 2003), un gène *htrA* complet a été retrouvé dans HD9B_GL2 (locus tag LGL2_RS10295) et chez les souches de référence. Des résultats similaires ont été obtenus pour GdpP, une protéine de signalisation du stress lié à la membrane (Smith et al. 2012). Des orthologues identiques ou presque identiques de GdpP de la souche HD9B_GL2 (locus tag LGL2_RS03375) ont été trouvés dans toutes les souches de référence de *L. lactis* ssp. *lactis*, et les protéines correspondantes dans les souches de référence ssp. *cremoris* montrent 95% d'identité, avec 32 acides aminés différents répartis sur toute la longueur de la protéine totale de 654 acides aminés (non illustrée).

Le système à deux composants CesSR contribue à la résistance à la chaleur à travers la régulation transcriptionnelle de deux protéines membranaires de fonction inconnue (Roces et al. 2009). Les deux composantes CesS (locus tag LGL2_RS04025) et CesR (locus tag

Résultats et discussion

LGL2_RS04030) ont été trouvées intactes dans la HD9B_GL2 et chez les souches de référence (données non présentées).

Figure 42: A. Organisation transcriptionnelle de l'opéron *ctsR-clpC*. P: Promoteur; Flèches: CtsR-binding site. B: Alignement (Clustal Omega) de la séquence N-terminale de la protéine CtsR des quatre souches *L. lactis* (les deux premières lignes) et de *Streptococcus mutants* UA159 / *B. subtilis* 168. Les séquences protéiques sont issues de la séquence publiée de HD9B_GL2 ou issues de GenBank. C. Région 5'-en amont du gène *ctsR* chez *L. lactis* ssp. *lactis* (HD9B_GL2 et les trois souches de référence) et chez ssp. *cremoris* (MG1363 et SK11). -35 / -10 régions promotrices, CtsR-binding the putative target site (direct repeats) pour le régulateur négatif (en italique et souligné). Putative ribosome binding sites (RBS) sont en gras et soulignés, le codon stop en aval du RBS1 est surligné. Start1 et Start2 désignent les codons d'initiation de la traduction, # positions de différence nucléotidique entre ssp. *lactis* et ssp. *cremoris*, résultant en deux acides aminés différents dans la séquence protéique. D. Alignement des régions de RecA1 et RecA2 de HD9B_GL2 contenant les motifs conservés Walker A et Walker B. E. Présence ou absence de RecA1 et RecA2 dans différentes souches de *L. lactis* (GenBank accession no. Pour le génome complet de A76: CP003132.1).

Résultats et discussion

A



B

```

GL2/KLDS4.0325  MDERKDGAMAGQKNTSDIIEAYLRKLL..159aa
IL1403/MG1363  -----MAGQKNTSDIIEAYLRKLL..151aa
S.mutans UA159  -----MTS-KNTSDSIEEYIKKIL..154aa
B.subtilis 168  -----M-GH-NISDIIEQYLKRVL..154aa
    
```

C

```

                -35                                -10
ssp.lactis    TTTGGTAAAAATAATATTTTTTCAGGTATAAT
ssp.cremoris TTTGGTAAAAAAGTATTTTTTAGGTATAAT
                CtsR-binding
ssp.lactis    AGTCAGTAATAGTCAAAGATTATTTAGAAA
ssp.cremoris AGTCAGTAATAGTCAAAGATTATTTAGAAA
                RBS1                                Stop
ssp.lactis    ATCGATGAAAGAAAGGATGGAGCGATCGCAG
                Start1 #          RBS2          # Start2
ssp.cremoris ATCGATAAAAGAAAGGATGGATCGATCGCAG
                M D E/K R K D G A/S M A
    
```

D

```

                77      Walker A      91      151      Walker B      161
GL2 RecA1  .EIYGPESSGKTTVA..  ..AVDLVVIDSVA..
GL2 RecA2  .EVFGAESSGKTTVA..  ..ALDLIVIDSVA..
    
```

E

Strain	RecA1	RecA2
<i>L.l.ssp.lactis</i> GL2	+	+
<i>L.l.ssp.lactis</i> IL1403	+	-
<i>L.l.ssp.lactis</i> KLDS4.0325	+	+
<i>L.l.ssp.lactis</i> NCDO2118	+	-
<i>L.l.ssp.cremoris</i> MG1363	+	-
<i>L.l.ssp.cremoris</i> SK11	+	-
<i>L.l.ssp.cremoris</i> A76	+	+

Tableau 19 : Principales caractéristiques des gènes de choc thermique clonés et séquencés chez *Lactococcus lactis* ssp. *lactis* HD9B_GL₂.

Gènes	ORF (nt)	Acides aminés	MM* (Da)	Promoteur de choc thermique	Promoteur végétatif ^a	Terminateur ^a	IR ^{a,b}	RBS ^c	Organisation en opéron	Homologie (%)
<i>dnaJ</i> (chaperone)	1139	379	40725	-	+	-	+	+		97.63 % MG1363m, 99.74 % IL1403, 99.74%KLDS4.032,99.74%NCDO2118
<i>dnaK</i> (chaperone)	1823	607	64987	-	+	+	+	+	hrcA-grpE-dnaK-orf4	98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118
<i>grpE</i> (chaperone)	539	179	20550	-	+	+	+	+		98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118
<i>hrcA</i> (repressor)	1157	385	44190	-	+	+	+	+		98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118
<i>orf4</i> (hypothetical protein)	1763	587	67175	-	-	+	-	+		98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118
<i>groEL</i> (chaperone)	1628	542	57201	-	+	-	+	+		GroES-groEL
<i>groES</i> (chaperone)	284	94	10220	-	+	-	+	+	98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118	
<i>ftsH</i> (protéase)	2087	695	75560	-	+	+	+	+		95.77% MG1363, 100 % IL1403 99.70%KLDS4.032, 100% NCDO2118
<i>rpoD</i> (sigma factor)	1343	447	50741	-	+	+	-	+		98.85 MG1363, 100 % IL1403 100% KLDS4.032, 100% NCDO2118

^aVoir tableau 20. *Lactococcus lactis* subsp. *cremoris* MG1363(abbreviation MG1363, accession number NC_009004.1); *Lactococcus lactis* subsp. *lactis* IL1403 (abbreviation IL1403, accession number AE005176) , *Lactococcus lactis* subsp. *lactis* NCDO 2118 (abbreviation NCDO2118, accession number NZ_CP009054.1), *Lactococcus lactis* subsp. *lactis* KLDS4. 0325 (abbreviation KLDS4. 0325, accession number CP006766). * Molecular mass in Daltons, ^bInverted repeats (IR), ^cRibosomal binding site (RBS)/ Shine Dalgarno Sequence

Résultats et discussion

Tableau 20 : Séquences nucléotidiques remarquables et conservées chez les gènes de choc thermique de *Lactococcus lactis* ssp. *lactis* biovar. *diacetylactis* HD9B_GL₂.

Gènes (a)	Promoteur végétatif -35 -10	Séquence inversée répétée (IR)	Séquence de Shine Dalgarno (RBS)
<i>dnaJ</i>	5' TTGCCA (14pb) TGGTAAAAT 3'	AATTAGCACTCTT (5pb) AAGAGTGCTAATT	AAAGGA
<i>groES-EL</i>	5' TTGACT (14pb) AGATAAAAAT 3'	TTAGCACTC (9pb) GAGTGCTAA	GGAGG
<i>hrcA-grpE-dnaK</i>	5' TTGACA (17 pb) TATAAT 3'	TTAGCACTC (9 pb) GAGTGCTAA + terminateur indépendant de rho GCACTGA (7 pb) GTCAGTGC	AGAGA
<i>orf4</i>	5' TTGCAA (14 pb) TACTATGAT 3'	+ terminateur indépendant de rho AGTGAAGG (5 pb) CCTTCGCT	AGGAGG
<i>ftsH</i>	5' ATTAGCA (18 pb) TATAA 3'	TTAG (4 pb) CAAA + terminateur indépendant de rho TAGTTGTC (4 pb) GACGACTG	AGGA
<i>rpoD</i>	5' TACTGA (13 pb) TATTTAAT 3'	+ terminateur indépendant de rho AAAACCTTACTGA (8 pb) TCAGTAAGTTTT	AGGAGG

(a) Les séquences se rapportent aux gènes décrits dans le tableau 19, isolés à partir de la même souche *Lactococcus lactis* HD9B_GL₂.

Les caractéristiques essentielles de ces gènes figurent dans les tableaux 19 et 20 ; nous résumerons, ci-dessous, leurs particularités les plus saillantes.

À notre connaissance, il s'agit de la première analyse systématique des gènes impliqués dans la thermotolérance et leurs protéines correspondantes dans une même seule souche *L. lactis* dont le génome a été totalement séquencé, combinée avec une analyse comparative des souches de référence largement étudiées appartenantes aux deux sous-espèce importante dans la fermentation des produits laitiers. Nous concluons ainsi que *L. lactis* HD9B_GL₂ possède un potentiel considérable pour la thermotolérance tout en utilisant de mécanismes spécifiques relativement évaluées dans ce travail de thèse.

Les gènes sont isolés ou groupés en petits opérons, qui rappellent ceux d'*E coli* (*groEL-ES*), (Kim and Batt 1993), de *B subtilis* (*grpE-dnaK*) (Barril et al. 1994), ou des deux espèces (*rpoO-dnaE*) (*B subtilis*) ou (*rpoO - dnaG*) (*E coli*). Cependant, chez *L lactis* l'opéron *dnaK* comprend deux autres cadres ouverts de lecture (*orf*) (Eaton et al. 1993), et *dnaJ*, habituellement couplé à *dnaK* fait défaut (van Asseldonk et al. 1993).

De la structure opéronique *rpoO-dnaE*, constituant l'opéron MMS (*macro molecular synthesis*) est absent, en amont, le gène *rpsU* dont le produit est impliqué dans la traduction (Araya et al. 1993).

Résultats et discussion

L'alignement des séquences en acides aminés des produits des gènes clonés, déduites des séquences nucléotidiques, avec celles des protéines homologues de diverses espèces bactériennes (*E coli*, *B subtilis*, *B megaterium*, *Staphylococcus aureus*), démontre une homologie importante (tableau 20). Cependant, chez *L lactis* comme chez les autres bactéries Gram positives on note la délétion de certaines parties: ainsi la région V2 du facteur sigma majeur d'*E coli* (*rpoO*) (Eaton et al. 1993) qui compte chez les bactéries Gram-négatives 250 acides aminés est réduite ici à quelques résidus (Araya et al. 1993; Gansel et al. 1993); de même, la protéine DnaK, est amputée dans sa partie N terminale d'une séquence de 24 acides aminés ce qui justifie une masse moléculaire calculée plus faible que celle de la molécule homologue d'*E coli* (Eaton et al. 1993).

Les cadres ouverts de lecture ou les opérons de *L lactis* sont précédés:

- par un promoteur végétatif présentant des séquences -35 et -10 caractéristiques (TTGACA et TAAAAT) (Eaton et al. 1993) ;
- par une IR, (répétition inversée) de 9 à 12 pb formant une structure en tige-et boucle (*stem-loop*), identique à celle repérée chez 8 *subtilis* et chez d'autres bactéries Gram-positives (Eaton et al. 1993);
- par un site de fixation des ribosomes, (séquence de Shi ne Dalgarno) complémentaire de l'ARN 16S (AGGAG pour *dnaK*) ; chez certains d'entre eux (*rpoO*) sont inclus une séquence TGN, présente dans environ 50 % des gènes des bactéries Gram-positives et une autre répétition inversée faisant office de terminateur potentiel (Araya et al. 1993); aucun ne montre le motif consensuel qui signe l'existence d'un promoteur reconnu par un facteur sigma alternatif, comme *sigma 32* ou *sigma 24* (sigma E) d'*E coli* (Glover and Lindquist 1998) ;
- parmi une soixantaine de marqueurs, quatre de ces gènes sont localisés, (Le Bourgeois et al. 1995) (*rpoO*, *dnaK*, *dnaJ* et *grpE*) sur la carte génétique de *L lactis* 1403, et de *L lactis* subsp *cremoris* MG 1363. Sur la première, *dnaJ* flanque *rpoO* ; sur la seconde *dnaJ*, *rpoO* et *dnaK* divisent le chromosome en trois parties égales.

Le séquençage des gènes isolés ou groupés en opérons et de leurs zones promotrices exclusivement végétatives remet en question l'universalité du seul modèle accepté ces dernières années (Neidhardt et al. 1984; Neidhardt 1987; VanBogelen et al. 1987; Lindquist and Craig 1988), celui d' *E coli*, qui se résume à une commutation transcriptionnelle. Classiquement, la réponse à une stimulation thermique s'accompagne dans les minutes qui suivent par le remplacement du facteur sigma majeur (sigma 70), dégradé à température non permissive, par le facteur sigma alternatif (sigma 32) (Calendar et al. 1988) stable et même activé à haute température et dont la synthèse est par ailleurs amplifiée sous l'effet d'un autre

Résultats et discussion

facteur (sigma 24 ou sigma E) qui reconnaît le promoteur P3 de *rpoH* et son propre promoteur P2, assurant ainsi son autorégulation (Bukau 1993; Yura et al. 1993). L'inflexion de la réponse et le retour à état de repos est sous la dépendance de 4 hsps, DnaK, GroEL, GrpE, HflB ou FtsH (Herman et al. 1995). Le facteur sigma 32 conduit la transcription de la vingtaine de gènes qu'il reconnaît grâce à leur motif *hs* (*heat-shock*) et qui composent ce régulon conventionnel.

La majeure partie des protéines dont les gènes sont inclus dans le régulon dépendant de sigma 32 présentent une affinité pour d'autres protéines cellulaires, cibles de la chaleur, qu'elles assistent, désagrègent, restaurent ou dégradent. Elles participent vraisemblablement à la protection des cellules ou à leur sauvetage lorsqu'elles sont menacées par leur environnement thermique.

Le facteur σ^{24} ou σ^E classé dans le groupe des facteurs ECF (*extracytoplasmic function*) contrôle l'expression d'une dizaine de gènes (dont *rpoH* et *rpoE*) dont les produits sont impliqués dans la croissance aux températures élevées (42-44 °C) ; dans la survie aux épreuves thermiques létales (50 °C) (Raina et al. 1995; Rouviere et al. 1995). Les produits des gènes de ce second régulon interviendraient dans le compartiment extracytoplasmique (Mecenas et al. 1993).

Au contraire chez *B subtilis*, les promoteurs *hs* sont absents; ceux *d'E coli* fusionnés à des gènes rapporteurs sont inactifs (Wetzstein et al. 1992). L'expression de la réponse au choc thermique procède de promoteurs canoniques, reconnus par sigma 43 (homologue de sigma 70 *d'E coli*), qui ont été identifiés en amont d'un groupe de gènes thermorégulés de classe 1, et d'une séquence inversée répétée de 9 pb appelée CIRCE (*controlling IR of chaperone expression*) (Zuber and Schumann 1994), habituellement intercalée entre le signal de transcription et le début du cadre ouvert de lecture. La fixation supposée d'une protéine répresseur sur ce motif récurrent limiterait ou interdirait l'expression du gène. Ce même répresseur constituerait le capteur ou «senseur», l'une des cibles sensibles, qui obéirait à la sollicitation thermique par un changement conformationnel, lequel provoquerait son détachement de l'IR suivi par la transcription du gène correspondant (Wetzstein et al. 1992; Yura et al. 1993; Volker et al. 1994; Zuber and Schumann 1994). À la régulation transcriptionnelle positive s'exprimant, chez *E coli*, par une permutation des facteurs σ^{70} aux facteurs σ^{32} et σ^{24} eux-mêmes emboîtés, se substitue chez *B subtilis*, au moins pour un groupe de gènes, un mécanisme à régulation négative en *cis* faisant intervenir, outre le couple promoteurfacteur sigma végétatif, la reconnaissance d'un élément *hs* (CIRCE), couplé à l'altération d'un hypothétique répresseur.

Résultats et discussion

Deux modèles nous sont ainsi proposés: l'un complexe, étranger semble-t-il aux éléments aujourd'hui reconnus, par l'analyse des gènes séquencés, chez *L. lactis*; l'autre moins élaboré, qui traduit une riposte plus fruste aux sollicitations thermiques.

Ne possédant que des données éparses et des informations lacunaires, nous retiendrons ce dernier modèle pour expliciter le mécanisme global de la réponse chez *L. lactis*.

La présence des 1R ayant été, par ailleurs, reconnue chez une dizaine de bactéries Gram positives et apparentées, appartenant à des groupes très variés, de *Clostridium* sp à *Mycobacterium tuberculosis*, chez lesquelles elles représentent le motif le plus fortement conservé, il apparaît qu'un tel système est plus ancien que le mécanisme régulateur mis en place, chez *E. coli* (Zuber and Schumann 1994). Cependant Duwat et al (1995) proposent une autre voie originale de contrôle impliquant à la fois, par une réaction en cascade, Rec A et HflB.

L'intervention de cette dernière protéine sous-entend la médiation par un facteur σ alternatif du type σ_{32} , qui, jusqu'à ce jour, n'a pas été identifié.

Chez *B. subtilis*, les réponses aux stress ne dépendent pas uniquement de σ^B . Les relations entre le régulon de réponse générale au stress σ^B et la réponse à un stress spécifique peuvent être illustrée par le stimulon (ensemble des gènes induits en réponse à un stimulus) du stress hyperthermique. Les gènes de réponse au choc hyperthermique sont classés en six grandes classes chez *B. subtilis*, mais les mécanismes de régulation sont seulement connus pour les classes I, II, III et V (Wiegert and Schumann 2003). Les gènes de la classe II sont les gènes régulés par σ^B et ne sont donc pas spécifiques du stress hyperthermique.

Présentons maintenant succinctement la régulation des classes I, III et IV. Les gènes de la classe I sont régulés par un régulateur de transcription, HrcA. Chez *B. subtilis*, ils sont organisés en deux opérons, les opérons groES/EL et hrcA/grpE/dnaK/dnaJ/yqeT/yqeU/yqeV (Homuth et al. 1997). HrcA est un répresseur transcriptionnel qui se fixe sur des sites appelés CIRCE (*Controlling Inverted Repeat of Chaperone Expression*) de consensus: TTAGCACTC-N₉-GAGTGCTAA (Zuber et Schumann, 1994). L'activité de HrcA est modulée par la protéine GroEL (Mogk et al. 1997). L'hypothèse est que, pour être active, HrcA a besoin de GroEL. Lors d'un choc hyperthermique, GroEL est titrée par les protéines dénaturées, HrcA est alors inactivé et la répression des deux opérons de chaperons est levée.

Les gènes de la classe III sont au nombre de six et sont régulés par le régulateur transcriptionnel CtsR. Ils sont organisés en trois unités transcriptionnelles, ctsR/mcsA/mcsB/clpC, et les unités monocistroniques clpE et clpP (Kruger and Hecker 1998;

Résultats et discussion

Derre et al. 1999a; Derre et al. 1999b). CtsR réprime la transcription en se fixant sur la CtsR-box dont le consensus est : A/GGTCAAANANA/GGTCAAA. L'activité de CtsR est régulée post-traductionnellement de deux manières. CtsR est dégradée à 37°C par les protéases ATPdépendantes ClpP et ClpX, ce qui permet de garder un niveau basal faible de CtsR (Derre et al. 2000). De plus, CtsR est phosphorylée par la tyrosine kinase McsA/McsB (McsA est l'activateur de la kinase McsB), ce qui inhibe sa fixation sur la CtsR-Box (Kirstein et al. 2005). L'activité de McsB est inhibée par ClpC (codée par le même opéron que CtsR, McsA et McsB). En condition de stress hyperthermique, ClpC serait titrée par les protéines dénaturées, McsB phosphorylerait alors CtsR, provoquant son inactivation ce qui lèverait la répression des gènes de la classe III.

La classe V compte deux membres, htrA et htrB codant des protéases ancrées à la membrane (Darmon et al. 2002). Ces gènes sont contrôlés par le système à deux composants C_{ss}RS, qui répond non seulement au stress hyperthermique mais aussi au stress de sécrétion (induit en laboratoire par surproduction d' α -amylase). C_{ss}S détecterait l'accumulation de protéines dénaturées dans le milieu extérieur et activerait, via la phosphorylation de C_{ss}R et sa fixation sur les opérateurs, l'expression des protéases de surface (Hyyrylainen et al. 2001).

La régulation du stress hyperthermique chez *L. lactis* fait donc intervenir différents types de régulateurs. De plus, certains gènes sont sous double contrôle, comme les gènes de classe III, clpC et clpP sous contrôle de σ^B et de CtsR.

Chez les autres firmicutes, l'organisation observée chez *L. lactis* n'est pas toujours conservée. Par une étude de génomique comparative, Chastanet et coll. ont montré que les sites de liaison de CtsR se retrouvent en amont des opérons groESL des streptocoques (Chastanet et al. 2003) et cette régulation a été démontrée chez *S. pneumoniae* (Chastanet et al. 2001). Chez *Streptococcus salivarius*, clpP est sous contrôle de CtsR et d'HrcA (Chastanet and Msadek 2003). Chez *Staphylococcus aureus*, CtsR contrôle non seulement les gènes *clpC*, *clpB* et *clpP* mais aussi les opérons *dnaK* et *groESL*, qui se retrouvent sous le double contrôle de CtsR et HrcA (Chastanet et al., 2003a). Chez la bactérie lactique *Oenococcus oeni*, seul CtsR est présent, HrcA est absent de la séquence du génome. Chez cette bactérie, les gènes *clp* et les opérons *groESL* et *dnaK* sont sous le contrôle unique de CtsR (Grandvalet et al. 2005).

Résultats et discussion

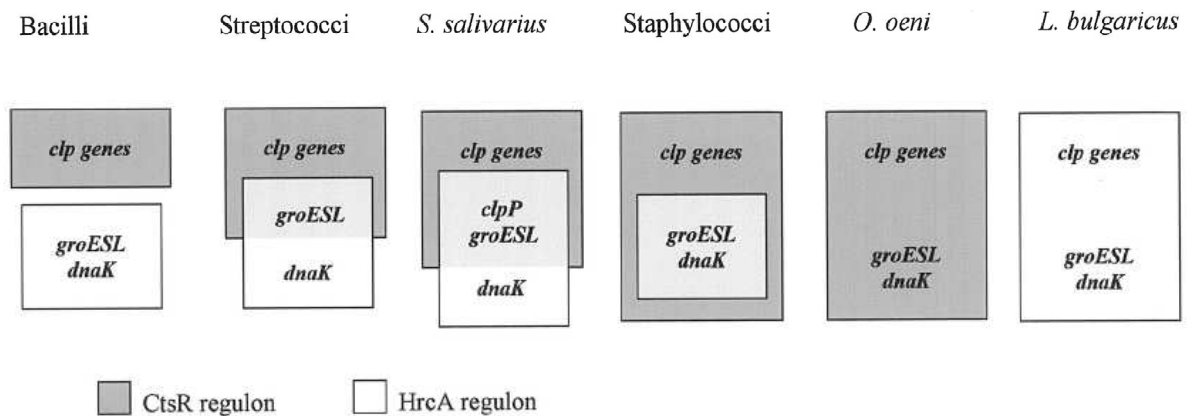


Figure 43: Comparaison des régulons d’HrcA et CtsR chez différents firmicutes (Grandvalet et al. 2005)

Les protéines chaperons ont pour rôle de permettre aux protéines dénaturées par la température de retrouver leur conformation native. Chez *L. lactis*, les deux principaux opérons de protéines chaperons sont codés par les opérons *groES/groEL* et *hrcA/grpE/dnaK/dnaJ*. Chez de nombreux microorganismes, *hrcA* code le régulateur négatif de ces deux opérons. Les autres protéines pouvant être impliquées dans la réponse au stress thermique détectées dans le génome de *L. lactis* HD9B_GL2 sont:

Les complexes multiprotéiques Clp sont la seconde famille de protéines importantes pour la résistance au stress thermique. Les complexes Clp sont composés de deux sous-unités : une sous-unité protéolytique, ClpP et une sous-unité liant l’ATP qui permet la reconnaissance du substrat. Sous forme de complexes, les Clp ont pour fonction de dégrader les protéines qui sont définitivement dénaturées lors du stress. Mais la sous-unité ATPasique peut, sous forme libre, remplir un rôle de protéine chaperon. Citons notamment le cas de ClpB qui ne présente pas le motif (LIV)G(FL) requis pour l’interaction avec ClpP et qui, induite en conditions de stress, aurait un rôle de chaperon protéique (Chastanet et al. 2004). Les gènes *clpP* codant la sous-unité protéolytique et *clpL*, *clpX*, *clpE*, *clpC* codant différentes sous-unités liant l’ATP sont présents dans le génome de *L. lactis* HD9B_GL2. Les gènes *clpY* et *clpQ*, présents eux aussi, codent pour un système de dégradation indépendant du système ClpP et dont il a été montré qu’il a un rôle mineur dans la résistance aux stress de *S. aureus* (Frees et al. 2005).

Chez *B. subtilis*, bactérie modèle des firmicutes, les régulateurs des protéines chaperons et des complexes Clp sont respectivement HrcA et CtsR. Notre analyse de *L. lactis* révèle la

Résultats et discussion

présence d'un homologue de HrcA mais aucun gène susceptible de coder CtsR. Soulignons qu'en amont des *clpP*, *clpC* et *clpE* d'ATCC 11842, aucune séquence comparable au site de fixation de CtsR (A/GGTCAAANANA/GGTCAAAA, deux dégénérescences admises pour la recherche du motif dans le génome d'ATCC 11842) n'est identifiée. Au contraire, une séquence proche du consensus TTAGCACTC-N9-GAGTGCTAA correspondant au site de fixation d'HrcA est présente en amont des opérons *groESL* et *hrcA/grpE/dnaK/dnaJ*. Cette séquence est également retrouvée en amont de *clpP*, *clpE* codant l'homologue de Hsp18 (figure 14) mais pas en amont des gènes codant les autres sous-unités Clp (*clpL*, *clpX* et *clpC*), ni en amont de *GL2_0156* (Hsp15), de Hsp33, *clpY* et *clpQ*. Ces observations nous conduisent à proposer que chez *L. lactis*, le rôle de HrcA soit plus important que chez les autres firmicutes et qu'il pallie en partie l'absence de CtsR. Dans l'état actuel, nous ne pouvons cependant pas affirmer qu'HrcA est le seul régulateur de la réponse au stress thermique chez ATCC 11842. En effet, bien qu'aucune séquence conservée ne soit identifiée en amont de *clpX*, *clpL*, *clpC*, *ldb0364*, *ldb0369*, *clpY* et *clpQ*, il faudrait tester si ces gènes répondent au choc thermique avant de conclure.

Chez *L. bulgaricus*, la régulation des gènes de réponse au stress hyperthermique semble à l'opposé de celle d'*O. oeni* qui ne possède pas HrcA mais dont les opérons codant les chaperons DnaK et GroESL sont sous contrôle de CtsR (Grandvalet et al. 2005).

Nous avons regardé quelle était la situation chez les autres lactobacilles dont les génomes sont séquencés et disponibles. Sur les quatre autres génomes de lactobacilles disponibles, seuls *L. sakei* et *L. plantarum* possèdent CtsR. Les lactobacilles du groupe *L. acidophilus* ne possèdent que HrcA. Chez *L. johnsonii* et *L. acidophilus*, on retrouve des boîtes CIRCE en amont des deux opérons codant les chaperons (dont l'organisation est bien conservée) mais aussi en amont de *clpP* et des deux copies de *clpE*. Les homologues de *ldb0289* (*lba0205* et *lj_0181*) semblent eux aussi régulés par HrcA puisqu'ils possèdent une boîte CIRCE. Aucune autre séquence CIRCE n'est détectée dans ces génomes. Le régulon HrcA semble donc conservé au sein du groupe *L. acidophilus*. Chez *L. plantarum* et *L. sakei* qui n'appartiennent pas au groupe *L. acidophilus*, où *ctsR* est présent, on retrouve des boîtes CIRCE en amont des opérons *groESL* et *hrcA-grpEdnaK-dnaJ* ainsi qu'en amont des homologues de *ldb0289* (*lsa0050*, *lp_3352* et *lp_0129*). Cependant, aucune boîte CIRCE n'est détectée en amont des gènes codant les sous-unités Clp. D'autre part, la boîte de reconnaissance de CtsR est détectée en amont de *ctsR/clpC* et de *clpE* dans ces deux génomes (figure 43).

Chapitre 3 : Résultats et Discussion

Partie III : Analyse métagénomique d'un *consortium* bactérien « *Lactococcus*, *Leuconostoc* et *Enterococcus* » présent dans l'écosystème lait de chamelle

1. Introduction

Le terme écosystème a été défini par Arthur Georges Tansley comme « l'ensemble formé par une association ou communauté d'êtres vivants et son environnement géologique, pédologique et atmosphérique ». Les éléments constituant un écosystème développent un réseau d'interdépendances permettant le maintien et le développement de la vie. Un écosystème est la résultante d'interactions entre les facteurs biotiques et abiotiques (Tansley, 1958).

Dans l'environnement, les bactéries, comme tous les êtres vivants, co-habitent dans des écosystèmes où elles interagissent entre elles et avec le milieu. Des écosystèmes bactériens complexes existaient déjà il y a 3,4 milliards d'années. Dans les écosystèmes, de nombreuses interactions métaboliques et nutritionnelles entre les différents microorganismes ont lieu. Etant donné le nombre de facteurs et d'interactions qui interviennent, l'étude de ces écosystèmes est assez complexe et la résultante (par exemple, en termes de croissance) n'indique pas nécessairement le nombre et le type d'interactions impliquées. Le lait, quel que soit son origine (vache, chèvre, brebis, chamelle ...), constitue une niche écologique importante pour une large diversité de microorganismes, dont les bactéries lactiques. On dénombre 11 genres de bactéries lactiques utilisées dans les produits fermentés alimentaires : *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus*, *Vagococcus* et *Weissella* (Stiles and Holzapfel 1997).

Les bactéries lactiques occupent des niches écologiques ou écosystèmes très variés. Le plus étudié est le lait de vache avec une microflore riche. Dans les régions à climat chaud et dans les zones arides des pays d'Afrique, les chamelles jouent un rôle essentiel en fournissant du lait à la population. Jusqu'à maintenant, très peu d'études ont été menées pour caractériser la flore microbienne du lait de chamelle. Néanmoins, il a été montré que les lactocoques représentent 34,5% de la flore lactique totale issue du lait cru de chamelle d'Algérie (Hassaine et al. 2007). Bensalah et al 2009, (Bensalah et al. 2009) a identifié la présence de différentes espèces d'*Enterococcus* dans « leben », produit fermenté traditionnellement à partir du lait de

Résultats et discussion

chamelle des régions aride en Algérie. Benmechernene et al 2014, (Benmechernene et al. 2014) ont pu aussi identifier des bactéries de genre *Leuconostoc mesenteroides* dans le lait cru de chamelle d'Algérie. L'utilisation industrielle de ces bactéries nécessite leur identification et leur classification qui repose sur des critères phénotypique, physiologique, génétique et moléculaire.

C'est pour cela nous nous sommes intéressés à étudier le lait de chamelle ; un écosystème simple mais riche en même temps, faisant cohabiter un consortium de plusieurs espèces bactériennes appartenant aux différents genres: *Lactococcus*, *Leuconostoc* et *Enterococcus*. Nous disposons de peu de connaissances sur les interactions entre ces trois microorganismes, surtout au niveau des mécanismes et des régulations. Afin de pouvoir analyser la diversité bactérienne au sein de notre échantillon, nous avons fait une analyse métagénomique en utilisant des assembleurs de métagénome (voir Matériels et Méthodes). Pour pouvoir reconnaître l'origine des contigs (i.e. quel gène appartient à quel organisme). Cette analyse métagénomique nous a permis de faire une recherche approfondie des communautés microbiennes et fournit un accès à la diversité génétique présente dans le lait de chamelle.

La métagénomique est un procédé méthodologique qui vise à étudier le contenu génétique d'un échantillon issu d'un environnement complexe, trouvé dans la nature (par opposition à des échantillons cultivés en laboratoire). Le but de cette approche, via le séquençage direct de l'ADN présent dans l'échantillon, est d'avoir non seulement une description génomique du contenu de l'échantillon mais aussi un aperçu du potentiel fonctionnel d'un environnement. L'utilisation du préfixe « méta » fait référence à « ce qui vient après ». Ici, la métagénomique vient après la génomique en étudiant les organismes microbiens directement dans leur environnement sans passer par une étape de culture en laboratoire. L'outil principalement utilisé pour obtenir des données en méta-génomique est le séquençage de l'ADN. Par exemple, il est possible de faire le diagnostic d'infections bactériennes avec la métagénomique (Nakamura et al. 2008). Plusieurs tâches sont associées à la méta-génomique. Une des tâches est la classification taxonomique des séquences obtenues par séquençage de l'ADN (Krause et al. 2008; Brady and Salzberg 2009; Caporaso et al. 2010; Segata et al. 2012). Une autre tâche en méta-génomique est l'assemblage de novo de méta-génomes – c'est-à-dire de construire des séquences longues qui correspondent au matériel génétique des organismes vivant dans un environnement donné à partir des données de séquençage (Laserson et al. 2011; Peng et al. 2011; Boisvert et al. 2012; Namiki et al. 2012). Avec l'assemblage de novo, il est possible d'analyser le contenu d'un échantillon plus efficacement (Cottrell et al. 2005).

Résultats et discussion

En effet, des séquences plus longues sont plus spécifiques pour la méta-génomique (Wommack et al. 2008). Des séquences longues peuvent être obtenues avec l'assemblage de novo (Chaisson et al. 2009) ou avec une technologie de séquençage offrant des séquences plus longues comme les technologies Pacific Biosciences ou Moleculo. Les chercheurs étudient des environnements très diversifiés en méta-génomique, comme les océans en général (Rusch et al. 2007) et l'Arctique (Varin et al. 2012) ou l'Antarctique, en particulier. La méta-génomique comparative consiste à comparer des méta-génomes en utilisant des mesures comme le contenu en gènes (Tringe et al. 2005). Un des objectifs principaux de la méta-génomique est de prédire les réseaux métaboliques et donc les capacités enzymatiques qui sont présentes dans un système pour lequel des données de séquençage de l'ADN sont disponibles.

2. Analyse de la micro-communauté 9B

2.1. Identification des trois espèces du « *consortium* »

Les analyses méta génomiques ont montré la présence de trois génomes appartenant à trois espèces différentes dans l'échantillon, le lait : *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* GL2, et *Enterococcus faecium* GL3. Les projets de séquençage génomique et métagénomique ont été déposés au GenBank/EMBL/DDBJ sous les numéros d'accèsion suivant : LMXE000000000 , JNCC020000000 et LIHM000000000 correspondant aux différentes souches lactiques: *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* HD9B_GL2 et *Enterococcus faecium* GL3 respectivement.

Cette partie a fait l'objet d'un article soumis à la publication (*The isolate HD9B from camel milk is a tightly associated consortium of Lactococcus lactis, Leuconostoc mesenteroides and Enterococcus faecium*).

L'environnement peut contenir de nombreuses espèces bactériennes, cependant, des conditions hautement sélectives réduisent la diversité des espèces. Par exemple, l'augmentation de la salinité du milieu conduit à une augmentation de l'abondance des archéobactéries, tandis que l'eau douce est occupée avec une variété d'espèces procaryotes (Rodriguez-Brito et al. 2010). De même, le lait peut accueillir de nombreuses espèces microbiennes (Delbes et al. 2007). Cependant, les fortes pressions sélectives comme le traitement thermique (Van Hoorde et al. 2008), le processus de fabrication des produits laitiers (Randazzo et al. 2002), les propriétés intrinsèques des LAB et les interactions antagonistes (Lindgren and Dobrogosz 1990), enrichi ce milieu avec quelques espèces de

Résultats et discussion

LAB seulement. La métagénomique a été largement utilisée pour comprendre la structure des communautés microbiennes des systèmes complexes. L'analyse métagénomique d'une communauté simple présente dans le lait révèle la dominance de certains genres de LAB d'espèce différentes : *Lactococcus lactis*, *Leuconostoc mesenteroides* et *Enterococcus faecium*.

Dans l'approche présentée ici, l'identification de ces trois espèces a été améliorée avec la recherche des phages et la quantification des groupes orthologues spécifiques à chaque souche qui ont été déterminés à partir des séquences génomiques issues du projet de séquençage par Illumina « *paire-ends* ». Les trois espèces présentes dans le lait ont été séparées à partir d'un même clone par des méthodes physico-chimiques, ensuite elles ont été caractérisées par des tests physiologiques et biochimiques. Leur caractérisation a été complétée par une étude microscopique au microscope électronique à balayage.

Les analyses effectuées dans cette partie illustrent qu'une communauté bactérienne relativement simple tel que le lait représente un vaste paysage génétique, et les principaux groupes phylogénétiques sont stabilisés au sein de la communauté par la dynamique des phages « *kill-the-winner* » spécifiques à chaque espèce.

Bien que le lait est, en termes de valeur nutritive, un des milieux les plus riches sur terre, il est remarquable qu'en général il contient seulement deux espèces de bactéries lactique, avec une grande dominance de *L. lactis* et d'une petites communauté de *Le. mesenteroides*. Apparemment, LAB créent un environnement monopolistique qui empêche la mise en place d'autres microbes dans l'écosystème lait.

Les lactocoques peuvent être quantifiés et caractérisé par des méthodes classiques telles que le dénombrement de l'unité formant colonies (UFC) à l'aide de divers milieux qui permettent la détection des capacités fonctionnelles spécifiques à chaque souche.

Par exemple, le milieu classique gélosé Reddy (Reddy et al. 1972) permet simultanément de différenciation de *L. lactis* subsp. *cremoris* (absence d'activité d'arginine désiminase, aucune utilisation de citrate), *L. lactis* subsp. *lactis* (activité arginine désiminase, aucune utilisation du citrate) et *L. lactis* subsp. *Lactis* biovar. *diacetylactis* (activité arginine désiminase, utilisation du citrate). En parallèle, les populations de *Leuconostoc mesenteroides* peuvent être énumérées sur géloses sélectives de MRS, additionnés à la vancomycine.

Résultats et discussion

La recherche sur la composition et le fonctionnement de consortiums microbiens complexes s'intensifie et ouvrira de nouvelles voies pour contrôler et améliorer les processus de fermentation alimentaires et développer de nouvelles applications pour des cultures mixtes. La plupart des procédés de fermentation alimentaires dépendent des cultures mixtes qui s'interagissent pour produire les caractéristiques désirées du produit final. En général on peut dire que les consortia microbiens ont des activités plus complexes (polyvalence) et tolèrent plus des variations d'environnement (robustesse) par rapport à des cultures pures.

La polyvalence et la robustesse peuvent s'expliquer par deux caractéristiques (Brenner et al. 2008). Tout d'abord, les membres du consortium communiquent entre eux en échangeant des métabolites ou des signaux moléculaires. Ainsi chaque cellule réagit à la présence d'autres bactéries dans le consortium (Keller and Surette 2006). La deuxième caractéristique est la coopération entre les microorganismes du consortium conduisant à une réponse/production globale qui ne peut s'expliquer en combinant des tâches accomplies par les différentes constitutions ou les sous-populations (Ayad et al. 2001; Keller and Surette 2006).

La robustesse des cultures complexes dans les processus de fermentation est décrite en termes d'activité fiable et la production de produits alimentaires d'une qualité constante. Fait intéressant, les cultures complexes montrent aussi une plus grande résistance contre les bactériophages par rapport à des cultures simple d'une seule souche ou en couple de deux souches seulement. L'explication de ce dernier phénomène est en corrélation avec la présence de bactériophages dans les levains complexes comme membres de la communauté microbienne. Les phages sont effectivement le moteur de la diversité génétique et phénotypique de la population microbienne (Rodriguez-Valera et al. 2009).

La connaissance de la composition des communautés microbiennes est importante pour comprendre ses fonctionnalités et peut être analysée à l'aide du séquençage et analyse métagénomique (Field et al. 2014; Tatusova et al. 2014; Brown et al. 2015; Tatusova et al. 2015). Enfin, les interactions microbe - microbe dans les fermentations complexes peuvent être démontrées, en combinant le profil méta-transcriptomique et les tests de croissance choisis.

Résultats et discussion

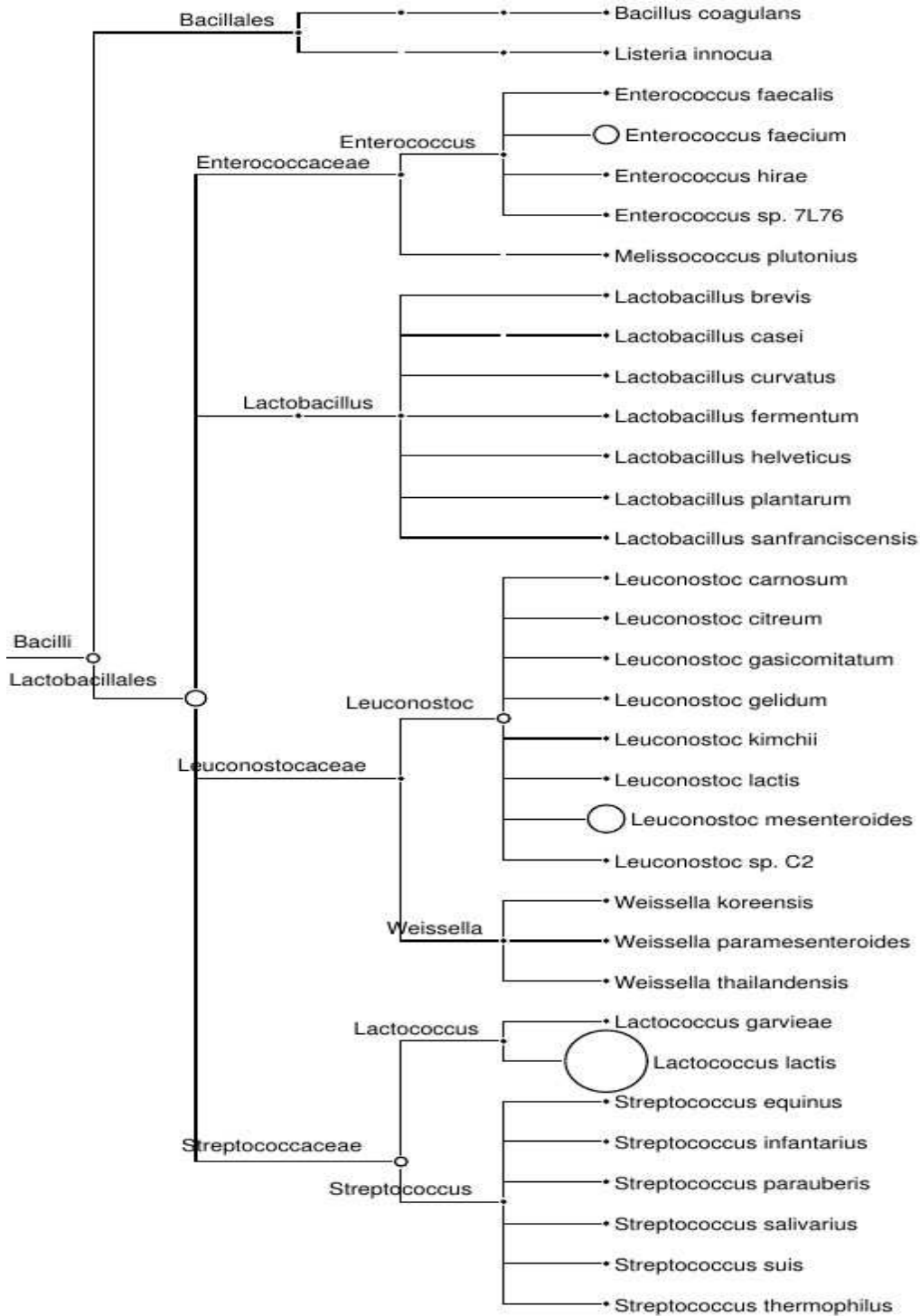


Figure 44: Identification des trois espèces composantes du « consortium » lait. L'arbre phylogénétique est construit par le logiciel Megan5. Chaque cercle représente un taxon et sa taille correspond au pourcentage de contigs dans la séquence génomique spécifique à chaque espèce.

Résultats et discussion

Pour vérifier la présence de trois espèces dans la séquence génomique, on a collecté des lectures du gène 16S rRNA par BLAST et on les a assemblées. Ces « 16S rRNA » pourraient déterminer approximativement de quelle espèce ils proviennent (Figure 45). Ce résultat confirme le résultat obtenu par Megan et montre que l'échantillon contient aussi trois espèces différentes: *Lactococcus lactis*, *Leuconostoc mesenteroides* et *Enterococcus faecium*.

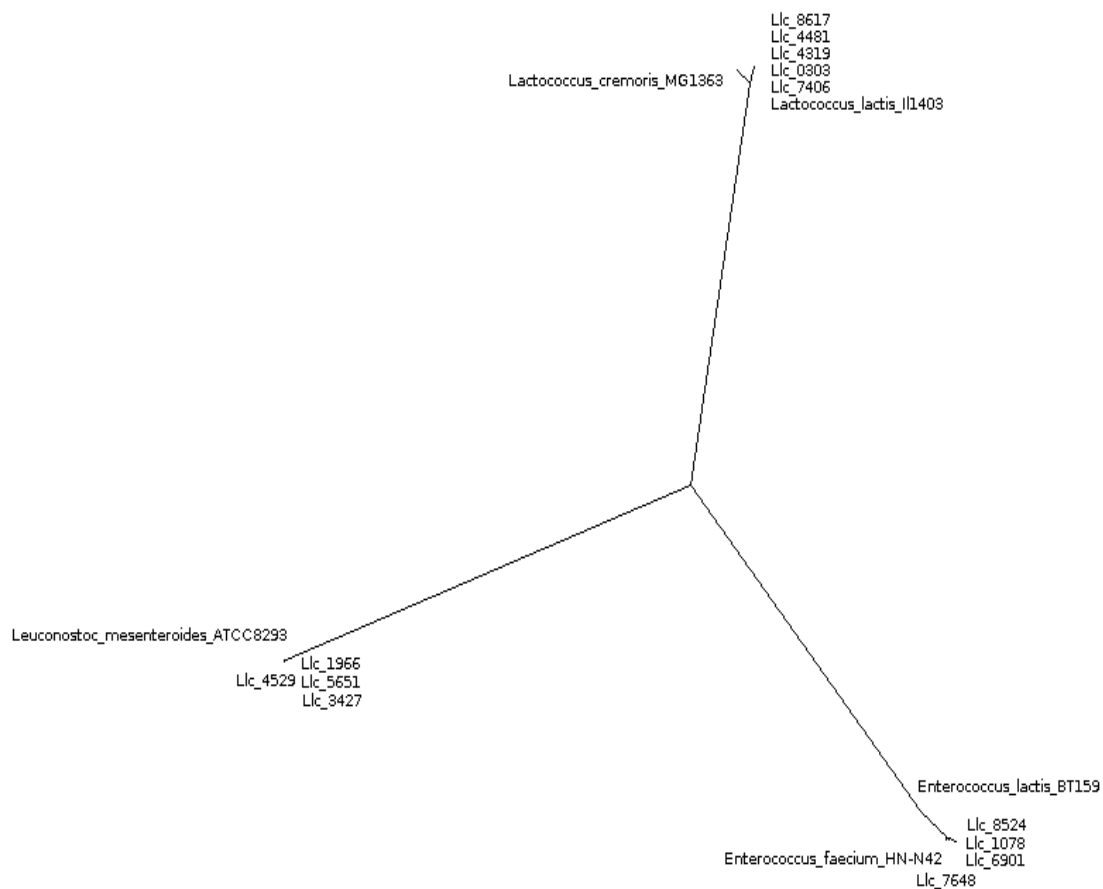


Figure 45: Arbre phylogénétique basé sur 16S rRNA.

Résultats et discussion

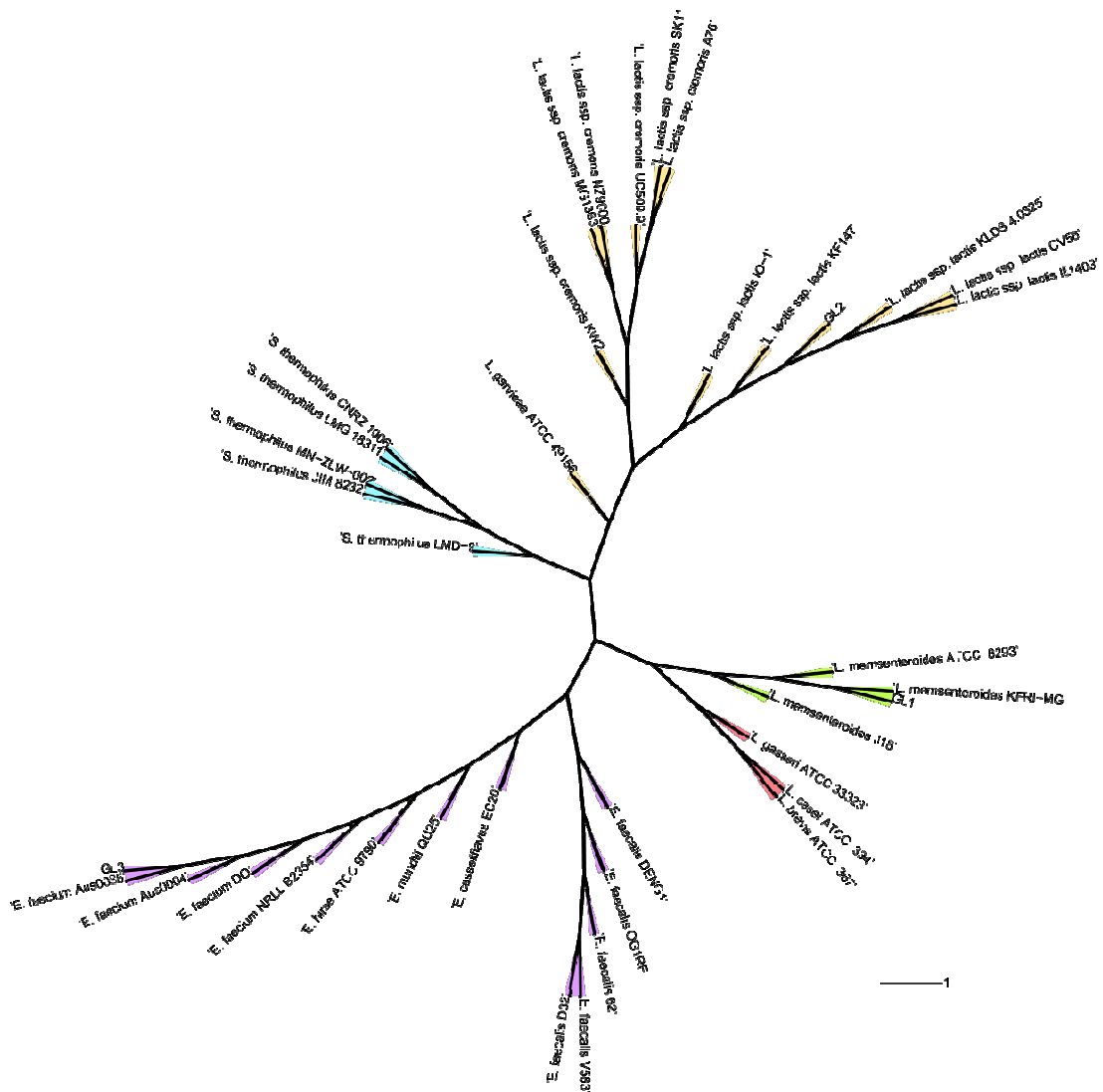


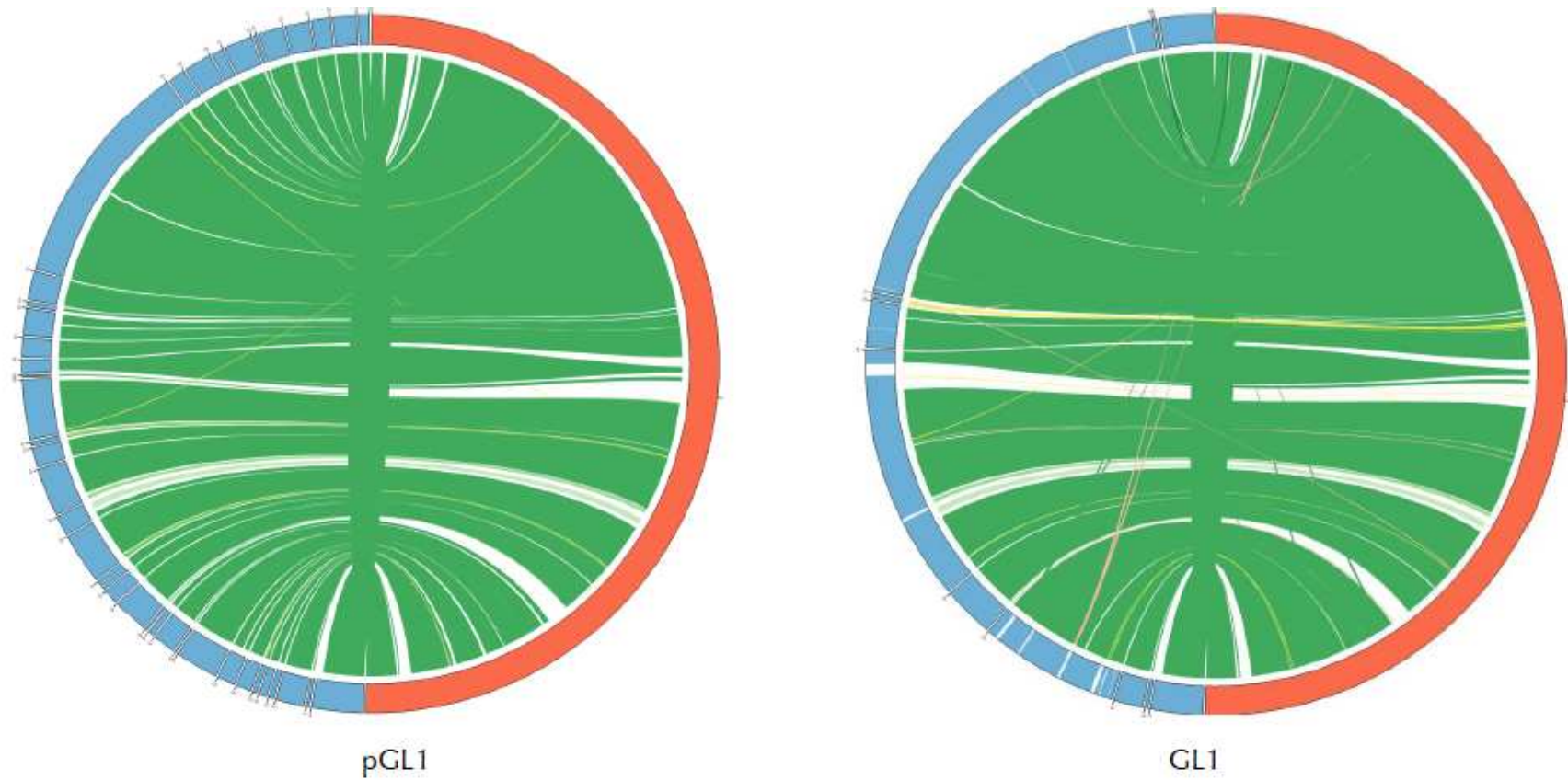
Figure 46: Arbre phylogénétique visualisé par iTOL (Letunic and Bork 2007). Les trois souches composantes du consortium GL1: *Leuconostoc mesenteroides*. GL2: *Lactococcus lactis* ssp. *Lactis*. GL3: *Enterococcus faecium* ont été comparées avec les souches de références : *Leuconostoc mesenteroides* KFRI-MG (Tatusova et al. 2015), *Lactococcus lactis* subsp. *lactis* KLDS 4.0325 (Yang et al. 2013) et *Enterococcus faecium* Aus0085 (Lam et al. 2013) respectivement. Les gènes marqueurs (16S rRNA, *rpoB*, *rpoBb* et *atpD*). GL2 est relativement distincte des autres lactocoques séquencés *L. lactis* subsp. *lactis* et subsp. *cremoris*, et même distincte de subsp. *lactis* bv. *diacetylactis*.

Résultats et discussion

Tableau 21: Informations génomiques générales des trois espèces isolées du *consortium*: *Leuconostoc mesenteriodes* GL1, *Lactococcus lacti* ssp. *Lactis* GL2 et *Enterococcus faecium* GL3

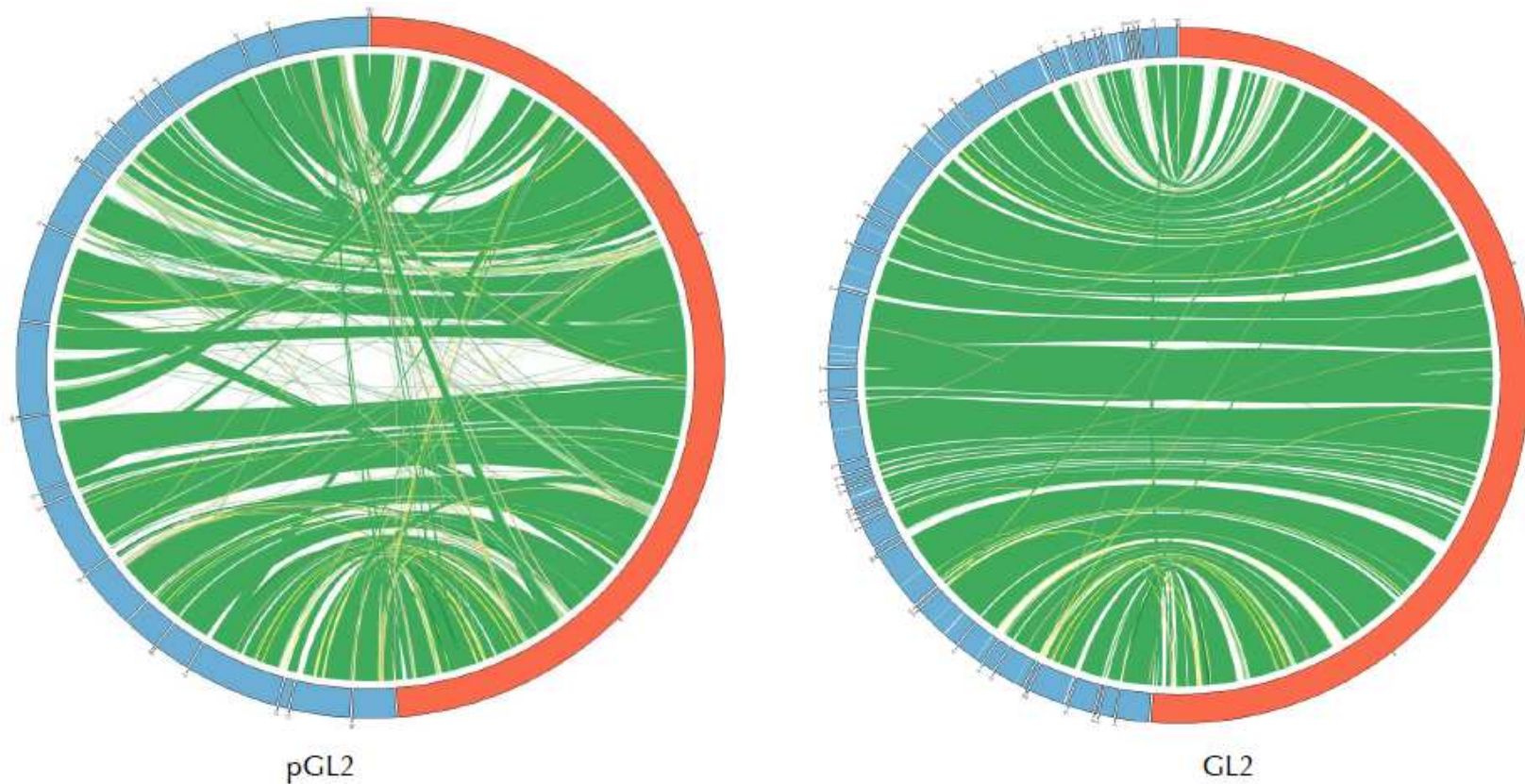
Souche	taille/bp	Sequences codantes	GC%	n50	n90
GL1	1,815,975	1727	37.7	446287	88812
GL2	2,251,485	2201	35.1	72247	25725
GL3	2,507,582	2395	38.1	167313	76444

Le tableau suivant représente la taille et la qualité des trois génomes. La couverture originale (données non présentées) est presque 50 pour les trois génomes. Cependant, le résultat montre que la valeur N50 est au moins 10 fois supérieure dans GL1 et GL3 que dans la GL2, ce qui reflète que la couverture de la lectures dans GL2 est nettement inférieure à 50, c'est pourquoi la taille du génome de GL2 est plus petite en comparaison avec le génome de référence.



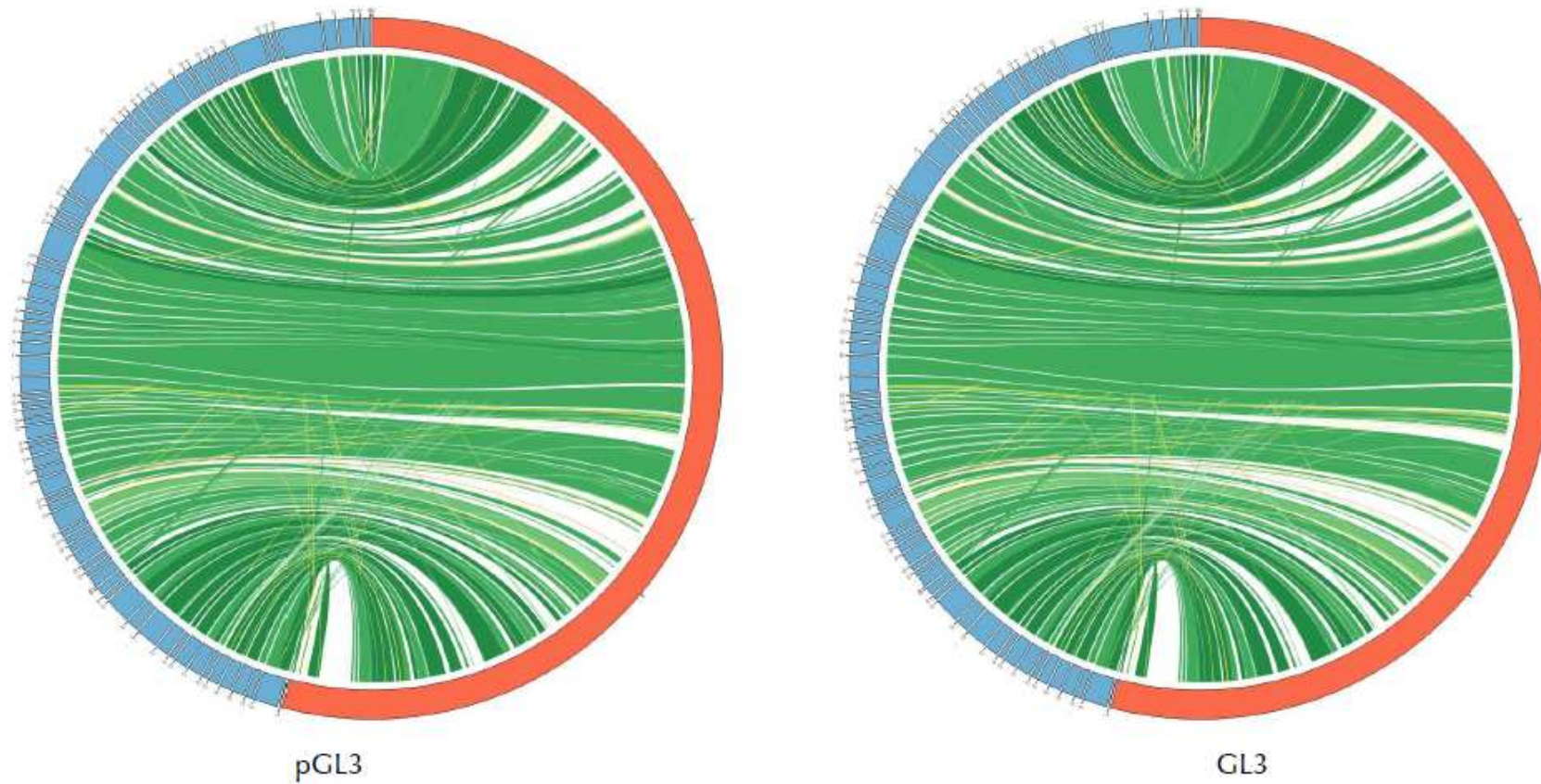
Souche	Taille/pb	Contigs	GC%	le plus long	N50	N90
GL1	1,815,975	11	37.7	702,867	446,287	88,812
pGL1	1,732,246	39	37.7	400,422	59,171	25,672

Figure 47(a) : Comparaison de la qualité de l'assemblage chez *Leuconostoc mesenteroides* GL1



Souche	Taille/pb	Contigs	GC%	le plus long	N50	N90
GL2	2,251,485	48	35.1	178,647	72,247	25,725
pGL2	2,604,375	22	35.3	234,027	188,966	70,424

Figure 47(b) : Comparaison de la qualité de l'assemblage chez *Lactococcus lactis* GL2



Souche	Taille/pb	Contigs	GC%	le plus long	N50	N90
GL3	2,507,582	31	38.1	270,914	167,313	76,444
pGL3	2,184,811	76	38.2	109,186	33,315	13,806

Figure 47(c) : Comparaison de la qualité de l'assemblage chez *Enterococcus faecium* GL3

Résultats et discussion

Figure 47(a) - (c) présente la qualité des trois génomes en détails. Ces cercles ont été visualisés par le logiciel Circos (Krzywinski et al. 2009). pGL1 et pGL3 ont été assemblés par Velvet et pGL2 est assemblé par Mira4 (l'assemblage obtenu par Velvet est d'une qualité inférieure). Il est évident que le résultat de pGL1 et pGL3 n'est pas mal et les assemblages améliorés GL1 et GL3 par la nouvelle méthode ont fortement diminué le nombre de contigs et ont augmenté la valeur N50. Dans GL1, le nombre des contigs est réduit de 39 au 11 et la valeur N50 augmente de 59 171 à 446 287 ; similaire au GL3, le nombre de contigs a été réduit de 76 au 31 contigs et la valeur N50 est levée de 33 315 à 167 313. Le résultat de pGL2 semble mieux que GL2, qui compte 22 contigs comparés à 48 contigs chez GL2 et d'une valeur N50 de 188 966 comparativement à 72 247 chez GL2. Toutefois, des erreurs sont trouvés dans pGL2, tels que l'absence de nouveaux gènes qui n'ont pas été trouvés par BLAST. Le résultat du *mapping* en comparant avec le génome de référence est incompatible, ce qui reflète que ce résultat d'assemblage n'est pas fiable. Le résultat montre également la fonctionnalité de Velvet et Mira : Velvet est comparativement plus précis mais donne des contigs plus courtes et une valeur N50 relativement faible; Mira4 produit des contigs plus longs, cependant il génère des erreurs lorsque la couverture des lectures est assez faible.

2.2. Séparation des trois espèces

Le mélange original (HD9B) est mis en culture sur milieu M17 contenant 100 µg/ml de l'antibiotique la vancomycine (*Leuconostoc* est naturellement résistant aux antibiotiques de classe glycopeptidique) pendant 48 heures à 30°C. Une croissance très faible a été observée. La culture est ensuite centrifugée et le culot est resuspendu dans M17 liquide puis repiquée sur milieu solide et cultivée à 30°C et à 45°C. Différentes colonies sont repiquées et testées par différentes PCR avec des amorces spécifiques aux trois espèces de LAB. Après plusieurs repiquages successifs on a pu séparer *Enterococcus faecium* GL3 du mélange, le résultat est confirmé par l'amplification des clones (Figure 26). Pour séparer *Leuconostoc mesenteroides* GL1 du reste du mélange, on a soumis ce dernier à une ultrasonification. Des dilutions sont effectuées ensuite cultivées sur M17 et incubé à 30°C et à 45°C. Ensuite, les clones thermosensibles sont choisis, isolé les ADNs et les amplifié par PCR, cependant on trouve toujours le mélange (*Leuconostoc* GL1 et *Lactococcus* GL2), pour cela on a jouté la vancomycine car leuconostoc est résistant aux ATB de classe glycopeptidique (Orberg and Sandine 1984) mais le résultat est encore le même on a toujours le mélange donc on a soumis ce mélange traité par la vancomycine à un deuxième traitement à l'ultrasonification.

Résultats et discussion

L'addition de la vancomycine dans le milieu de culture permet l'isolement de la souche désignée *Leuconostoc mesenteroides* GL1 dans une culture mixte de bactéries lactiques.

Le *rpoB* code pour la sous-unité β de l'ARN polymérase, le gène est relativement conservé, en particulier entre les coques LAB : *Lactococcus*, *Leuconostoc* et *Enterococcus*, mais montre une diversité pour dessiner des amorces spécifique pour chaque espèce.

Pour dessiner les amorces spécifiques du gène *rpoB* pour les souches G11, G12 et G13 correspondant aux trois espèces différentes *Leuconostoc mesenteroides*, *Lactococcus lactis* et *Enterococcus faecium*, respectivement, on a utilisé les séquences obtenues du projet de séquençage du génome total du consortium HD9B. Les trois espèces G11, G12 et G13 sont déposées dans GenBank/EMBL/DDBJ sous les numéros d'accès suivants : LMXE000000000 *Leuconostoc mesenteroides* GL1, JNCC020000000 *Lactococcus lactis* et LIHM000000000 *Enterococcus faecium* GL3.

Rapport-Gratuit.com

Résultats et discussion

Région 1

Lactoc_rpoB	GAATGCCGTGTCTTGACACCAGCCGATGTGATTTCAAATATCAGCTACTGGCTTGGTCTT	1185
Enteroc_rpoB	GCTGTAAAAACGGTTCGTCCAGCCGACATCATTGCTTCAATGAGCTATTT-CTTGAACTT	1193
	GCTGTAAAAACGGTTCG> Ent-RpoF	
Leucon_rpoB	GATGATCGTACAGTACGTCCAGCTGATATCTTGGCAGGTATGAACTACTT-CTTGAACTT	1190
Lactoc_rpoB	GCTG-AAGGAATTGGTAAAGTGGACGATATTGACCACTTAGGTAACCGTCGTATCCGTTT	1244
Enteroc_rpoB	GATGGAAGGTATCGGCAATGTGATGATATCGACCACTTAGGAAATCGTCGTATCCGTTT	1253
Leucon_rpoB	GCAAGAAGGTGTTGGTCAATGTTGATGATATTGACCACTTAGGTAACCGTCGTATCCGTTT	1250
Lactoc_rpoB	AGTAGGTGAATTGCTTCAAACCAAGTACGTATCGGACTTTCACGTATGGAACGTGTTAT	1304
Enteroc_rpoB	AGTCGGTGAACCTTTTACAAAACCAATTCCTGATCGGGCTAGCACGTATGGAACGTGTTGT	1313
Leucon_rpoB	TGTTGGTGAATTGTTGCAAAAACCAATTCCTGATTTGGCTTAAACGGTATGGAACGTGTTGT	1310
Lactoc_rpoB	TCGCGAACGTATGTCATCTTCAGAAAATGAAAACATCACACCGCAAGGTTTGATTAACAT	1364
Enteroc_rpoB	TCGTGAAAGAATGTGATCCAAGATACTGAAACATTGACACCTCAACAATTGATCAATAT	1373
Leucon_rpoB	TCGTGAACGCATGTCAATTCAGATGCCAACACGGTTACACCACAACAGTTAATCAATAT	1370
Lactoc_rpoB	TCGTCTGTAAACAGCTTCAATTAAAGAATTCTTTGGTTCATCACAATTGTCACAATTTAT	1424
Enteroc_rpoB	CCGTCCAGTTGTGGCAAGTATCAAAGAATTCTTTGGTTCCTTCAGTTGTCACAGTTTAT	1433
Leucon_rpoB	ACGTCTGTGTAGCCGCTGTCAAAGAATTCTTTGGTTCCTTCACAATTGTCACAGTTTAT	1430
Lactoc_rpoB	GGACCAACACAATCTTTATCTGAGCTTTCACACAAACGTGTTTTTCTGCCTTAGGACC	1484
Enteroc_rpoB	GGATCAAACAAACCCTCTTGGCGAATTGACACACAAACGTGCTTTTCAGCTTAGGACC	1493
	<GAGAACCGCTTAAGTGT Ent-RpoR2	
Leucon_rpoB	GGATCAAACAAACCATTGGGTGAAATGAACCACAAACGTGTTTTGTCAGCTCTTGGACC	1490

Région 2

Lactoc_rpoB	AGTTATGGCTCGTCATACTGGTAACAATATTGAAGTTGAAGCTTCAACAGCAGACTATAT	1844
	Lac-RpoF2 GTTGAAGCTTCAACAGC>	
Enteroc_rpoB	TGTTATGGCTCGTGCTCAAAGTGAACAACTTAGAAGTATCTATTGATAAAGTAGACTACAT	1853
Leucon_rpoB	AGTTATGGCTCGTTTTGGTGAAGAAAACATTGAAACACCAATTGATCGTATTGACTACAT	1850
Lactoc_rpoB	GGACGTTTCACCTAAACAGGTTATCGCCGTTGCTGCTGCATGTATTCCATTCTTGAAAA	1904
Enteroc_rpoB	GGACGTATCACCTAAACAAGTAGTTGCCGTTGCGACAGCATGTATTCTTTCTTGAAAA	1913
Leucon_rpoB	GGACGTTTCGCCAAAGCAGGTTGTTTCTGTTGGAACAGCTGCTATTCCGTTCTTGAAAA	1910
Lactoc_rpoB	CGATGACTCCAACCGTGCCCTCATGGGTGCCAACATGCAACGTCAAGCAGTTCCATTGAT	1964
Enteroc_rpoB	CGATGACTCCAACCGTGCCCTGATGGGTGCCAACATGCAGCGTCAAGCAGTGCCGTTGAT	1973
Leucon_rpoB	CGATGATTCCAACCGTGCTTTGATGGGTGCCAACATGCAACGTCAAGCAGTGCCCTTGGCT	1970
Lactoc_rpoB	TGACCCACATGCACCTTGGATTGGTACTGGTATGGAACATCAAACCTGCTCGTGACTCAGG	2024
	<GTAGTTTGACGAGCACT Lac-RpoR2	
Enteroc_rpoB	CAATCTCAAGCACCATGGGTTGGTACTGGTATGGAATACAAATCAGCGCATGACTCAGG	2033
Leucon_rpoB	TGATCCACATTCACCATTAGTAGGAACAGGAATTGAGTATAAAGCAGCTCACGATTCTGG	2030
Lactoc_rpoB	TGCTGCGCTTATTGCAAAACATGCTGGTGTGTTGAGTATGTTGATGGTAATGAAATTCG	2084
Enteroc_rpoB	AGCTGCTTGTATTGTAACATGACGGTGTGTTGAATACGTTGATGCTTCTGAAATCCG	2093
Leucon_rpoB	TGTTGCGATGATTGCACGCGCATCTGGTGAAGTTGAGTATGTGGACGGCCGCAAAATCCG	2090
Lactoc_rpoB	TGTTTCGTGTAATTCTGGTGAAGTATGATAATTTACAATATTACAAAATATCGTCGTTCAA	2144
Enteroc_rpoB	CGTACGCCGCGACAACGGTGCATTAGATAAATACGATGTAAACAAAATTCGTCGTTCAA	2153
Leucon_rpoB	AGTTTCGTGTAAGATGGTCAATTGGATACTTATGAATTGATGAAATTCGTCGTTCAA	2150
Lactoc_rpoB	CTCAGGAACTTCTTATAACCAACGTCCATTGGCTCGACTTGGAGAAAAAGTTGAGAAAGG	2204
	<CCGAGCTGAACCTCTTT Lac-RpoR4	
Enteroc_rpoB	CTCCGGAACAAGCTACAACCAACGTCCAATCGTTTCAATTTGGGTGAGAAAGTCGAAAAAGG	2213
Leucon_rpoB	CGGTGGTAAAACTATAACCAAAAAGCCAATTGTACATGTCGGTGAGCATATTGAAGCCGA	2210

Résultats et discussion

Région 3

Lactoc_rpoB	CGGACGTACTGGTGAACCATTTGATAACCGAATTTTCAGTCGGTGTGCATGTACATGATTAA	3104
Enteroc_rpoB	CGGACGTACAGGTGAACCATTTGATGGACGTGTGTCTGTCGGTGTGATGTACATGATCAA	3113
Leucon_rpoB	TGGTCGCACTGGTGAAGCTTTTGACAAGCGTGTGGTGTGGTGTGCATGCATTATATGAA	3110
Leu-RpoF3 TGACAAGCGTGTGGTG>		
Lactoc_rpoB	ACTTCACCACATGGTTGATGATAAACTCCATGCTCGTTCAGTTGGTCCTTACTCACTCGT	3164
Enteroc_rpoB	ACTTGCTCACATGGTCGATGACAAAATTGCATGCACGTTCAATCGGACCTTACTCATTGGT	3173
Leucon_rpoB	ACTAGCGCACATGGTTGACGATAAAAATTCATGCACGTTCAATTGGTCCCTTACTCACTTGT	3170
Lactoc_rpoB	TACACAACAACCGCTCGGTGGTAAAGCGCAGTTTCGGTGGACAACGTTTTGGGGAAATGGA	3224
Enteroc_rpoB	TACACAACAACCATTTGGGAGGTAAAGCTCAATTTGGTGGACAACGTTTTGGTGAATGGA	3233
Leucon_rpoB	TACACAACAGCCTTTGGGTGGTAAAGCCCAGTTTGGTGGACAGCGTTTCGGTGAATGGA	3230
Lactoc_rpoB	AGTTTGGGCACTTGAAGCCTATGGTGCCTGCAATGTTCTTCAAGAAATCTTGACTTACAA	3284
Enteroc_rpoB	AGTATGGGCACTGGAAGCTTACGGTGC GGCTTATACTTTACAAGAAATCTTGACATACAA	3293
Leucon_rpoB	AGTTTGGGCTTTGGAAGCTTATGGTGCAGCTTACACATTGCAAGAAATCTTGACGTATAA	3290
Lactoc_rpoB	ATCAGATGACGTGATTGGACGTACGCGCGCTTATGAAGCTATTGTTAAAGGTGAACGTAT	3344
Enteroc_rpoB	ATCAGATGACGTAGTCGGACGTGTGAAAACATACGAAGCAATCGTCAAGGGTGAACCAAT	3353
Leucon_rpoB	GTCAGATGATGTTGCAGGCCGTGTTAAAGTGTATGAATCAATTATTAAGGGTGAACCAAT	3350
<CGTCCGGCACAATTTCA Leu-RpoR2		

Figure 48: Sites des amorces dans *rpoB*. Les séquences sont issues du projet de séquençage total de la composante HD9B. *Lactococcus lactis* GL2, *Leuconostoc mesenteroides* GL1, et *Enterococcus faecium* GL3. Longueurs totales du gène : 3591 nt (*Lactococcus lactis* GL2), 3627 nt (*Enterococcus faecium* GL3), 3609 nt (*Leuconostoc mesenteroides* GL1).

Résultats et discussion

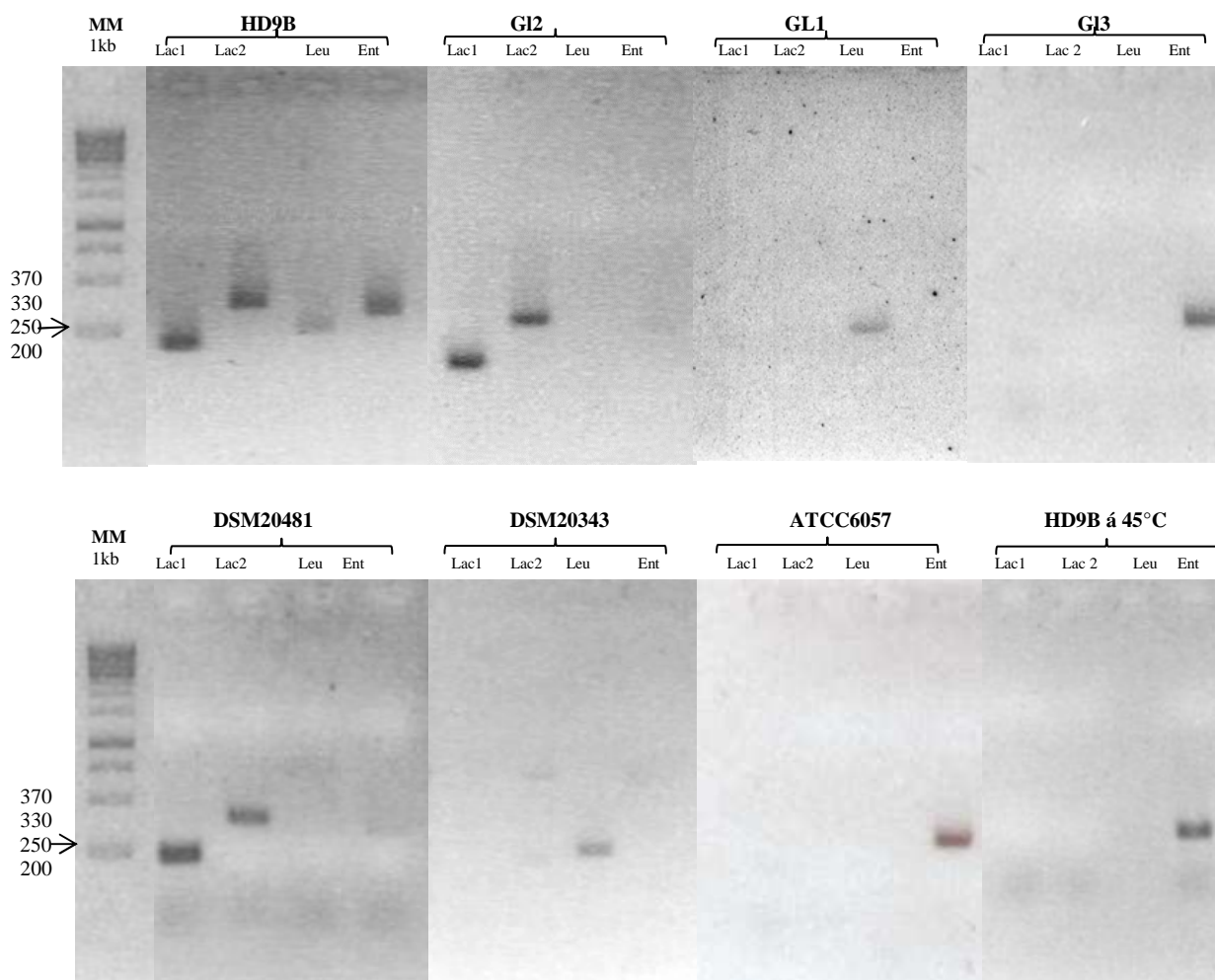


Figure 49 : Amplification du gène *rpoB* chez les différentes souches de LAB

Souches de référence : *Lactococcus lactis* DSM20481, *Enterococcus faecium* ATCC6057 et *Leuconostoc mesenteroides* DSM20343. *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* GL2, *Enterococcus faecium* GL3 : souches composante isolées du consortium HD9B. MM : Marquer moléculaire « Thermoscientific 1kb ». Agarose 0,8% dans 1X TAE. Les différents gènes sont amplifiés avec les couples d'amorces suivants : Lac1 : Lac-RpoF2/ Lac-RpoR2 ; Lac2 : Lac-RpoF2/ Lac-RpoR4 ; Leu : Leu-RpoF3/ Leu-RpoR2 ; Ent : Ent-RpoF/ Ent-RpoR2

2.3. Caractérisation des trois espèces séparées du consortium microbien

2.3.1. Cinétique de croissance à différentes températures

La croissance bactérienne constitue un critère physiologique important d'intérêt taxonomique. Cette croissance a été suivie en milieu M17-glucose liquide (Figure 50) en fonction de trois températures différentes (30°C, 37°C et 45°C). Les résultats obtenus (Figure 50) indiquent que à 30°C, toutes les souches issues du mélange HD9B à l'exception de *Leuconostoc mesenteroides* GL1 et la souche de référence *Lactococcus lactis* DSM20481. Par contre, la capacité, à se développer à 37°C est constatée chez toutes les souches. La souche *Lactococcus lactis* GL2 est capable de se développer à des températures élevées (45°C et 50°C), ce qui est inhabituel chez les lactocoques.

À des températures supérieures à 25° C, *L. lactis* croît nettement plus vite que *Leuconostoc* (Barrette et al. 2000), alors qu'à des températures entre 21° C et 25° C, une croissance plus équilibrée entre ces deux genres est observée (Hemme D, et al 2004). Au cours de la phase initiale de la fabrication de fromage, *Leuconostoc mesenteroides* dépend des souches caséinolytique de *L. lactis* pour la fourniture de l'essentiel des acides aminés libres ou des petits peptides (Cogan et al., 1994). En industrie fromagère, l'interaction entre ces deux bactéries dans une culture mixte est importante pour la formation de l'arôme, depuis que *Leuconostoc mesenteroides* peut produire seulement diacétyl et acétoïne à partir du citrate à pH acide (McSweeney, et al 2000) et l'acidification est produite principalement par la fermentation du lactose par les lactocoques.

Résultats et discussion

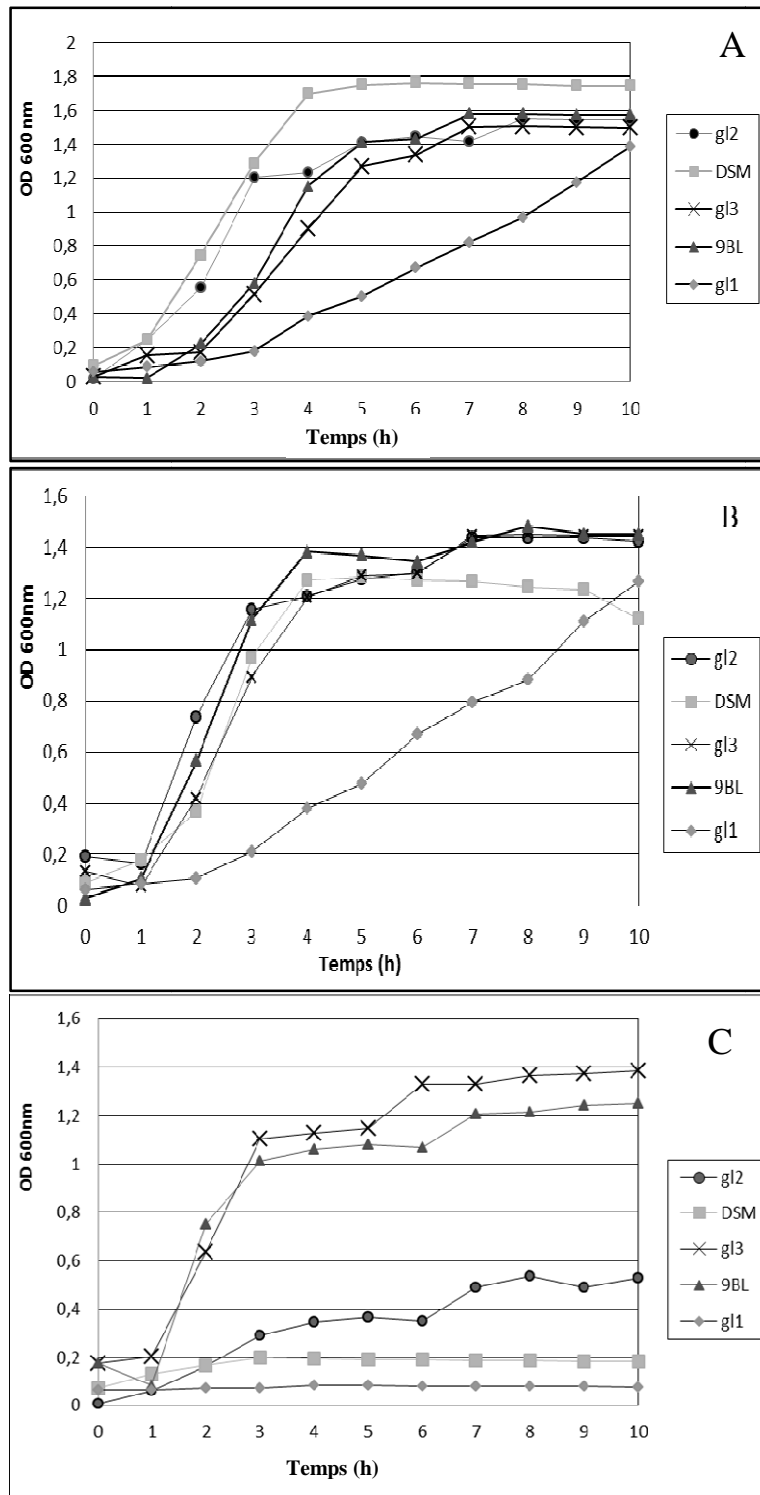


Figure 50 : Croissance en milieu M17-glucose liquide à différentes températures. A: 30°C ; B: 37°C ; C: 45°C ; DSM : DSM : Souche de référence *Lactococcus lactis* DSM20431 ; 9BL : le consortium ; g11 : *Leuconostoc mesenteroides* ; g12 *Lactococcus lactis* ; g13 *Enterococcus faecium* ; OD : densité optique.

2.3.2. Propriétés importantes

Dans cette partie de travail, nous nous sommes intéressés à séparer les trois souches présentes dans la composante HD9B, isolée à partir de lait cru de chamelle (*Camelus dromedarius*) collecté en Algérie. Ces trois souches appartiennent à trois différentes espèces de bactéries lactiques : *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* GL2 et *Enterococcus faecium* GL3. Ces isolats ont été ensuite analysés sur plusieurs types de milieux nutritifs gélosés (Tableau 22). Se sont des coques à Gram positif et catalase négative. Les souches *Lactococcus lactis* DSM20481 et *Leuconostoc mesenteroides* DSM20343 (Prt+ Lac+ Cit-) et *Enterococcus faecium* ATCC6057 (Lait+ Lac+ Cit+), sont utilisées comme souches de références dans les tests d'analyse.

Après incubation à 30°C de 24 heures, la présence d'une activité protéolytique se traduit par la formation d'un halo transparent autour des colonies bactériennes qui résulte de l'hydrolyse des caséines du lait. Dans la figure 51, les isolats protéolytiques préalablement obtenus ont été ensemencés sous forme de traits sur milieu lait glucosé. Cette méthode nous permet de mieux voir les halos transparents sur ce milieu opaque. La capacité de croissance rapide des bactéries liées à l'assimilation de la source d'azote sur les milieux gélosés opaques (lait, lait glucosé et lait additionné de milieu M17) se manifeste par l'apparition d'un halo transparent autour des colonies, due à l'hydrolyse de la caséine du lait.

La caséine représente 80% des protéines totales, et de ce fait, est majoritaire dans le lait (27g/l). Ainsi, la capacité de croissance dans le lait à des densités cellulaires importantes repose sur la présence d'un système protéolytique complexe capable d'utiliser les caséines comme source exogène d'azote (Mills and Thomas, 1981; Kunji *et al.*, 1996). Ce processus est connu sous le nom de caséinolyse et se déroule en trois étapes : (1) la première consiste en l'hydrolyse des caséines par une endoprotéase de paroi appelée PrtP générant des peptides de masse moléculaire variable, (2) certains de ces peptides sont transportés par deux systèmes de transport Opp et Opt, (3) la dernière étape est achevée par l'hydrolyse de ces peptides grâce à l'action de peptidases intracellulaires (Figure 51).

Résultats et discussion



Figure 51 : Activité protéolytique du « consortium » HD9B et ses trois espèces composantes GL1, GL2 et GL3 ensemencées sur milieu Lait glucosé. Le halo clair autour des traits d'ensemencement indique la présence d'une activité protéolytique.

La performance des bactéries lactiques utilisées comme levains dans l'industrie laitière réside dans leur métabolisme azoté. Ce type de métabolisme met en jeu un système protéolytique complexe constitué de trois composantes : une protéinase extracellulaire PrtP (capable d'hydrolyser les caséines, protéines majeures du lait), des systèmes de transport de peptides et des peptidases cytoplasmiques. Le système protéolytique conduit à la libération d'acides aminés, précurseurs d'arômes. En général, chez les lactocoques, la protéinase de paroi et certains systèmes de transport de peptides sont codés par des gènes plasmidiques. La relation entre la présence de plasmide et l'activité protéolytique chez les lactocoques a été prouvée pour la première fois par le travail de McKay en 1974 (McKay and Baldwin 1974b; a). La protéinase PrtP est codée par le gène *prtP* qui est précédé par un autre gène orienté en sens inverse, appelé *prtM*, codant pour une lipoprotéine nécessaire à la maturation de l'enzyme PrtP. Chez les souches Wg2, SK11 et HP de *Lactococcus lactis* subsp. *Cremoris* et *L. lactis* subsp. *Lactis* NCDO763, les gènes *prtP* et *prtM* sont respectivement portés par les plasmides pWV05, pSK11P, pHP003 et pLP763 (Kok et al. 1988a; Kok et al. 1988b; Kiwaki et al. 1989).

Les plasmides pHP003 (13,4 kb) et pSK11P (75,8 kb) sont complètement séquencés et caractérisés (Christensson et al. 2001; Siezen et al. 2005). Selon Christensson et coll. (Christensson et al. 2001), la taille des plasmides protéinase est un critère important pour déterminer le taux d'activité protéolytique. En effet, une souche possédant un plasmide de petite taille dont le nombre de copies est naturellement important montre une activité protéolytique plus élevée comparée à une autre souche dont le plasmide est plus grand.

Résultats et discussion

Les bactéries lactiques sont capables de fermenter les sucres en acide lactique. Ceci représente un outil très important de la technologie industrielle pour la transformation et la conservation de nombreux produits alimentaires d'origine végétale ou animale. Les sucres fermentés par les bactéries lactiques sont nombreux et appartiennent à plusieurs catégories à savoir : les hexoses, les pentoses, les disaccharides et les oligosaccharides. Selon les genres ou les espèces des bactéries lactiques l'une des deux voies suivantes peut être utilisée dans le métabolisme des sucres: (1) la voie homofermentaire est associée aux genres *Lactococcus*, *Streptococcus*, *Enterococcus*, *Pediococcus*, *Vagococcus*, *Tetragenococcus* et *Carnobacterium*. La glycolyse conduit à la formation exclusive d'acide lactique (CH₃-CHOH-COOH) ; (2) la voie hétérofermentaire, appelée aussi voie des pentoses phosphates, est spécifique des genres *Leuconostoc*, *Oenococcus* et *Weissella* et conduit à la formation de quantité équimolaire d'acide lactique, d'acétate, d'éthanol et de CO₂ (Cogan, 1980). Les espèces du genre *Lactobacillus* sont séparées en trois groupes en fonction de leur type fermentaire : homofermentaires, hétérofermentaires stricts et hétérofermentaires facultatifs qui métabolisent les hexoses en acide lactique par la voie homofermentaire, et qui dégradent aussi les pentoses par voie hétérofermentaire.

Le métabolisme des sucres chez les bactéries lactiques requiert d'abord leur transport à l'intérieur de la cellule bactérienne. Selon les espèces, cette internalisation des sucres met en jeu des perméases qui ne modifient pas le soluté au cours du transport, ou un système phosphoénol-transférase (PEP/PTS) qui phosphoryle le sucre au dépend du phosphoénolpyruvate (McKay et al. 1969; McKay et al. 1970; Vaughan et al. 1996). C'est le cas du lactose (C₁₂H₂₂O₁₁) sucre du lait (49 g/L) dont le métabolisme a fait l'objet de plusieurs études approfondies dans des revues (de Vos and Vaughan 1994; Vaughan et al. 1996; Neves et al. 2005). Chez les espèces du genre *Lactococcus*, le lactose phosphorylé est catabolisé par une phospho-β-galactosidase (P-β-gal) alors que chez *Leuconostoc* c'est une β-galactosidase qui hydrolyse le lactose. Ainsi, chez les bactéries lactiques, la différence dans les voies métaboliques (homo- ou hétérofermentaires) du lactose est fonction de l'équipement enzymatique.

Résultats et discussion



Figure 52 : Croissance sur milieu McDonald's HHD (*homofermentative-heterofermentative differential medium*). Ø: témoin control, A: HD9B, B: *Lactococcus lactis* GL2, C: *Lactococcus lactis* DSM20481, D: *Enterococcus faecium* GL3, E: *Enterococcus faecium* ATCC6057, F: *Leuconostoc mesenteroides* GL1, G: *Leuconostoc mesenteroides* DSM20343. Dans ce milieu à base de fructose, les LAB homofermentaire (*Lactococcus*, *Enterococcus*) produisent beaucoup plus d'acide par rapport aux LAB hétérofermentaires (*Leuconostoc*), virant la couleur du vert de bromocrésol dans HHD du bleu au vert.

Les trois isolats se développent sur tous les milieux testés (Tableau 22), comme les souches de référence. Toutes ces souches se développent en lait, et ont donc la capacité de fermenter le lactose et d'hydrolyser les protéines du lait. Ils sont capables de se développer sur des milieux de culture à base de lait (lactose) et présentent donc un phénotype Lac+.

L'étude génétique montre que le transport et l'hydrolyse du lactose chez *Lactococcus lactis* sont associés à la présence de gènes plasmidiques. Les travaux d'identification et de caractérisation des plasmides lactose sont considérables chez les souches de *L. lactis*. En effet, Kozar et coll. étaient les premiers à réaliser des expériences de curage chez des cellules de lactocoques de phénotype lactose-positif (Lac+) (Kozar et al. 1974). Des variants Lac-incapables d'utiliser le lactose ont été obtenus à une fréquence supérieure à 10%. L'ensemble des recherches réalisés jusqu'à l'heure actuelle confirme le déterminisme plasmidique de la fermentation du lactose chez les souches de *Lactococcus lactis* (Siezen et al. 2005). Par exemple le plasmide pSK11L (47,165 kbp) de la souche *L. lactis* subsp. *Cremoris* SK11 est actuellement le plasmide lactose le mieux caractérisé (Figure 52). Il est important de noter, que selon les lactocoques, la taille des plasmides lactose varie de 30 à 100 kb (Neve et al. 1984). Chez les espèces du genre *Leuconostoc*, les gènes qui codent pour le métabolisme du lactose peuvent être chromosomiques ou plasmidiques. Chez *Leuconostoc lactis* NZ6009, la présence d'un plasmide lactose (pNZ63) de 23 kb, a été mise en évidence par David et coll.

Résultats et discussion

(David et al. 1992) alors que chez *Leuconostoc mesenteroides*, le plasmide lactose est caractérisé par une taille de 51 kb.

L'organisation génétique du système *lac* chez *Lactococcus lactis* est différente de celle de *Leuconostoc lactis*. Chez le lactocoque, la région plasmidique *lac* de 10 kb consiste en un opéron *lacABCDFEGX* et un gène *lacR*, qui sont transcrits de façon divergente. Le gène *lacR* code pour une protéine réprimant l'expression de l'opéron *lacABCDFEGX*. Les gènes *lacABC* et *lacFE*, codent respectivement pour les enzymes de la voie du tagatose-6-P et les protéines (IIAlacIIBC_{lac}) nécessaires au transport du lactose. Quant au gène *lacG*, il code pour la phospho- β -galactosidase qui assure l'hydrolyse du lactose. Enfin, le gène *lacX* code pour une protéine de 35 kDa de fonction inconnue et non essentielle à l'utilisation du lactose par la bactérie. En présence du glucose, l'expression de l'opéron *lac* est réprimée d'un facteur de dix (van Rooijen and de Vos, 1990). Plusieurs publications (Oskouian and Stewart 1990) van Rooijen and de Vos, 1990; van Rooijen et al., 1991) indiquent que la séquence du "cluster" *lacR-lacABCDFEG* de *Lactococcus lactis* et celle de *Staphylococcus aureus* sont très similaires. Néanmoins, le gène *lacR* de *Staphylococcus aureus* est orienté dans le même sens que l'opéron *lacABCDFEG* (région dépourvue de plus du gène *lacX*).

Chez *Leuconostoc lactis* NZ6009, le système *lac* est localisé sur le plasmide lactose pNZ63 (Vaughan et al. 1996). Le gène *lacS* code pour la protéine de transport du lactose qui présente une identité de 34.5% et de 35.6%, respectivement avec celles de *Streptococcus thermophilus* et *Lactobacillus delbreuckii* subsp. *bulgaricus*. En amont du gène *lacS*, se trouve un gène *lacZ'* tronqué qui code pour une protéine de 95 résidus dont la séquence amino terminale est similaire à celle de la β -galactosidase I de *Bacillus stearothermophilus* (Vaughan et al. 1996). En aval du gène *lacS*, une séquence IS3 et les gènes *lacL* et *lacM* ont été identifiés. *lacL* et *lacM* sont partiellement chevauchants, transcrits dans le même sens que *lacS* et codent pour deux sous unités de la β -galactosidase.

Certaines bactéries lactiques appartenant aux genres *Lactobacillus*, *Oenococcus*, *Leuconostoc* *Weissella* et le biovar. *diacetylactis* de *Lactococcus lactis* subsp. *lactis*, ont la capacité de fermenter le citrate, mais en le co-métabolisant avec du glucose, car elles sont incapables d'utiliser le citrate comme seule source de carbone (Kempler and McKay 1980). La fermentation de cet acide carboxylique (citrate) présent dans les cellules végétales, animales et dans le lait, aboutit à la production de divers composés d'arômes, comme l'acétoïne, l'acétolactate, l'acétaldéhyde, le 2,3-butanediol et le diacétyle ce dernier étant le plus intéressant dans l'industrie laitière. Ces composés aromatiques contribuent d'une part, à

Résultats et discussion

l'élaboration de la saveur de certains aliments issus de la fermentation lactique et d'autre part, à la diminution de concentrations trop élevées de pyruvate qui sont toxiques pour la cellule bactérienne (Collins, 1972).

Les voies du métabolisme du citrate ont été mises en évidence grâce à plusieurs travaux de recherche. Deux revues font un bilan des connaissances sur le métabolisme du citrate chez les bactéries lactiques sur les plans biochimique, énergétique, génétique et moléculaire (Drider et al. 2004; Mills et al. 2005).

Le transport du citrate dans la cellule bactérienne par un citrate perméase CitP, constitue la première étape (Harvey and Collins 1962) (Figure 52). CitP assure simultanément l'élimination du lactate intracellulaire. Cet antiport citrate/lactate crée un potentiel membranaire. Le citrate internalisé est clivé par l'enzyme citrate lyase en acétate et en oxaloacétate. L'oxaloacétate est ensuite décarboxylé par l'oxaloacétate décarboxylase, donnant du pyruvate et du CO₂. Cette décarboxylation s'accompagne d'une consommation de protons et par conséquent d'une alcalinisation du milieu intracellulaire et d'un gradient de pH (Figure 52).

Les séquences du citrate perméase (CitP) de certaines espèces du genre *Leuconostoc* (*Leu. Mesenteroides* et *Leu. paramesenteroides*) et de *Lactococcus lactis* subsp. *Lactis* biovar. *diacetylactis* sont très proches (99%). Une identité significative de CitP (26 à 50%) est aussi trouvée avec d'autres transporteurs de citrate décrits chez les bactéries à Gram+ ou Gram- (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>).

A ce jour, seul le plasmide citrate perméase pCRL1127 (8278 pb) de la souche *Lactococcus Lactis* subsp. *Lactis* biovar. *Diacetylactis* CRL1127 est complètement séquencé et caractérisé (http://www.ncbi.nlm.nih.gov/genomes/static/eub_p.html). Sur ce plasmide, le Gène *citP* qui code pour la citrate perméase (CitP) appartient à l'opéron *citQRP*, dont les gènes *citQ* et *citR* codent respectivement pour un peptide leader et une protéine régulatrice. L'opéron *citQRP* peut être transcrit à partir de deux promoteurs P1 ou P2 (Drider et al. 2004). Par ailleurs, chez plusieurs souches de *Leuconostoc*, le métabolisme du citrate est lié à la présence de plasmides de taille supérieure (21 à 23 kb) à celle du plasmide présent chez les lactocoques (Bourel *et al.*, 2001). L'organisation en cluster *citI-citMCDEFGRP* du système *cit* plasmidique est aussi différente : *citDEF* codent pour les différentes sous-unités du citrate lyase, *citR* code pour une protéine régulatrice, tandis que *citM* code pour une oxaloacétate décarboxylase. Le gène *citI* est transcrit en sens inverse de l'opéron *citMCDEFGRP*. Il est important de signaler, qu'en présence de citrate, le produit du gène *citI* accroît l'expression de l'opéron *citMCDEFGRP*. Il

Résultats et discussion

s'agit donc d'un régulateur positif de l'utilisation du citrate par *Leuconostoc* (Martin et al. 2004).

La régulation de l'expression du gène *citP* semble différente entre les souches de *Leuconostoc* et *Lactococcus lactis* subsp. *Lactis* biovar. *diacetylactis*. Chez *Leuconostoc*, elle est fortement induite par la simple présence du citrate, mais chez le biovar *diacetylactis*, en plus de la présence du citrate, un pH acide intracellulaire s'avère un signal essentiel pour l'expression de l'opéron *citQRP*. Par contre, chez le biovar *diacetylactis*, la synthèse du citrate lyase CitDEF dont les gènes sont chromosomiques, est constitutive (Drider et al. 2004).

Tableau 22 : Analyse physiologique des trois composantes de l'isolat protéolytique HD9B

Milieux de cultures	Cat	M17-glu	Lait-glu	HHD	M17 X-Gal	KMK	Phénotype
Souches							
HD9B	-	++	++T	Homo	+B	+B	Prt ⁺ Lac ⁺ Cit ⁺
GL1	-	++	+T	Heter	+b	+B	Prt ⁺ Lac ⁻ Cit ⁺
GL2	-	++	++T	Homo	+B	+B	Prt ⁺ Lac ⁺ Cit ⁺
GL3	-	++	++T	Homo	+B	+b	Prt ⁺ Lac ⁺ Cit ⁻
Lact_DSM20481	-	++	++T	Homo	+b	+B	Prt ⁺ Lac ⁻ Cit ⁺
Leuco_DSM20343	-	+	nd	Heter	+b	+B	Lac ⁻ Cit ⁺
Ent_ATCC6057	-	++	nd	Homo	+B	+b	Lac ⁺ Cit ⁻

M17-Glu : M17 additionné de glucose ; KMK : Milieu Kempler et McKay : identification des souches Cit⁺ ; Lait-glu : lait glucosé ; HHD : homofermentative-heterofermentative differential medium; M17 X-Gal : M17 additionné de X-Gal ; + : Croissance ; - : Absence de croissance ou réaction négative ; nd : non déterminé ; T : halo transparent autour des patchs ; Homo : Homofermentaire ; Heter : heterofermentaire ; B : colonie bleue ; b : colonie blanche ; HD9B : représente le consortium (mélange des trois espèces : GL1, GL2 et GL3) ; Souches de référence : Lact_DSM20481 : *Lactococcus lactis* DSM20431 ; Leuco_DSM20343 : *Leuconostoc mesenteroides* DSM20343 ; Ent_ATCC6057: *Enterococcus faecium* ATCC6057 ; GL1 : *Leuconostoc mesenteroides* ; GL2 *Lactococcus lactis* ; GL3 *Enterococcus faecium* ; Cat : catalase négatif ; Prt⁺ : protéinase positif (la souche possède l'enzyme protéinase PrtP active capable d'hydrolyser la caséine du lait) ; Prt⁻ : protéinase négatif (la souche est dépourvue de l'enzyme protéinase PrtP active) ; Lac⁺ : capacité à fermenter le lactose ; Lac⁻ : incapacité à fermenter le lactose ; Cit⁺ : capacité à fermenter le citrate ; Cit⁻ : incapacité à fermenter le citrate.

Résultats et discussion

2.3.3. Observations morphologiques en microscopie électronique à balayage (MEB)

Les observations en MEB effectuées sur les souches *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* GL2 et *Enterococcus faecium* GL3 permettent d'étudier la morphologie bactérienne ainsi que les caractéristiques de surface des cellules. Les images de la microscopie électronique à balayage montrent que la similitude génomique des souches isolées du consortium 9B aux souches de références se reflète également dans leurs morphologies, les images de *Leuconostoc mesenteroides* GL1 et *Enterococcus faecium* GL3 sont très similaires avec les souches de références observées *Leuconostoc mesenteroides* DSM20343 et *Enterococcus faecium* ATCC6057, respectivement. Contrairement à la souche *Lactococcus lactis* GL2, cette dernière présente une structure très particulière (Figure 53-54-55). Les images du consortium bactérien « 9B » avant la séparation physique de ces trois espèces montrent la présence d'une agrégation entre ses espèces ; ceci peut être expliqué soit par la production d'exopolysaccharides EPS produites en abondance par les différentes souches de LAB, soit par les pilis attachés à la surface de *Lactococcus lactis* GL2. En effet, les figures montrent que les bactéries sont difficilement individualisables car toujours engluées dans des épais EPS constituant un gel autour des bactéries. Il faut signaler qu'on a trouvé une difficulté énorme pour séparer les trois souches fortement agrégées dans le consortium « lait ».

De la même manière, les observations réalisées sur les souches cultivées en milieu solide soulignent le fait que les souches *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* GL2 et *Enterococcus faecium* GL3 constituent des micro-colonies. La figure 55 illustre ainsi le fait que le contact entre les différentes cellules individualisées est flou du fait de la présence d'un EPS entourant chacune des bactéries. Par ailleurs, il est important de noter la présence de structures externes aux cellules révélatrices d'interactions entre les bactéries d'espèces différentes GL1, GL2 et GL3. Il n'y a visiblement pas de présence de "coatings" à la surface des cellules mais une production d'EPS toujours importante lorsque LAB est mise en milieu liquide.

Il apparaît que les micro-colonies présentent des contours flous représentatifs de la production d'EPS. De plus, les cellules individuelles de la *Lactococcus lactis* GL2 présentent des structures externes particulières « granulée ».

En résumé, les observations morphologiques des cultures de la *Lactococcus lactis* GL2 réalisées en MEB soulignent la production d'EPS isolée ou en micro-colonie avec les souches *Leuconostoc mesenteroides* GL1 et *Enterococcus faecium* GL3 ainsi qu'une structure externe de type "coating" ou PILI est visible à la surface de la GL2.

Résultats et discussion

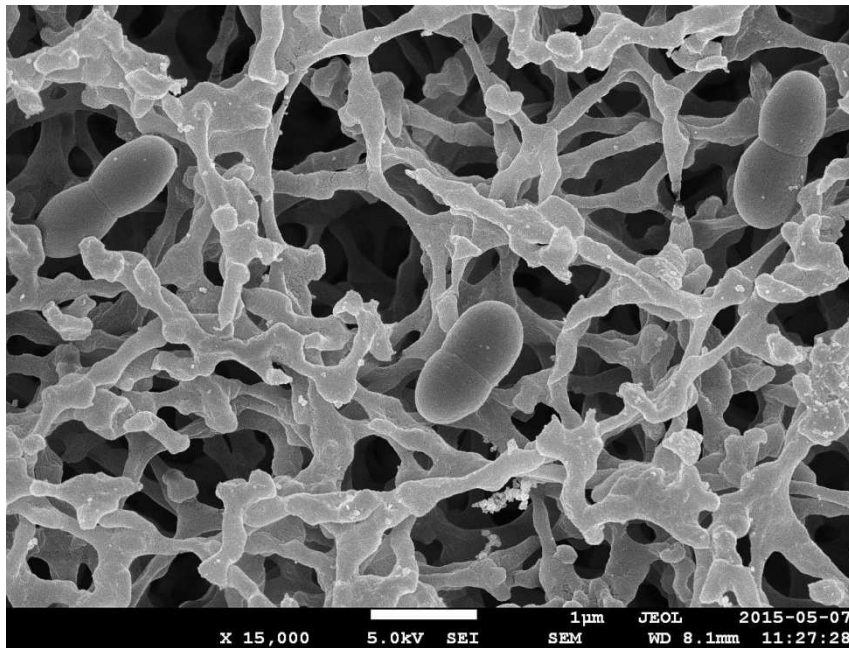
Les images en microscopie électronique à balayage des préparations de la GL2 cultivée séparément ou en présence des autres espèces GL1 et GL3 montrent dans tous les cas la présence d'une structure granulée sur la membrane cellulaire, cette granulation se présente toujours sous forme fibreuse, avec une architecture présentant un réseau dense de très fines aiguilles plus ou moins courbés. Il apparaît ainsi que les réseaux d'aiguilles sont localisés à la membrane cellulaire très fréquemment dans les zones confinées, comme par exemple au contact entre deux organismes, il s'agit des PILI.

De plus, la photo de la figure 54 montre que les EPS peuvent lier plusieurs micro-colonies entre lesquelles les parois cellulaires des organismes ne se touchent pas nécessairement mais dont les précipitations en aiguilles peuvent relier les organismes.

En parallèle des observations réalisées sur des cultures de *Lactococcus lactis* DSM DSM20481 et *Enterococcus faecium* ATCC6057, un protocole exactement identique a été réalisé sur les deux souches de référence. Les images et analyses réalisées sur *Lactococcus lactis* DSM20481 et *Enterococcus faecium* ATCC6057 montrent un comportement très similaire à celui observé sur *Leuconostoc mesenteroides* GL1 et *Enterococcus faecium* GL3 (figures 53-54-55).

Un lien entre les processus d'agrégation, la production d'EPS, de pilis et la formation de production de biofilms est la seule possibilité, toutes ces propriétés constituent des traits biotechnologiques importants pour la souche *Lactococcus lactis* GL2. Enfin, l'une des autres perspectives intéressante ouvertes par les résultats de ce travail porte sur l'étude des processus d'agrégation des trois espèces *Lactococcus lactis*, *Leuconostoc mesenteroides* et *Enterococcus faecium* ainsi que leur rôle dans le lait. Nous proposons d'étudier physiquement et chimiquement les mécanismes de micro-agrégation des bactéries entre elles dans des expérimentations *in vitro*.

A



B : Détail:

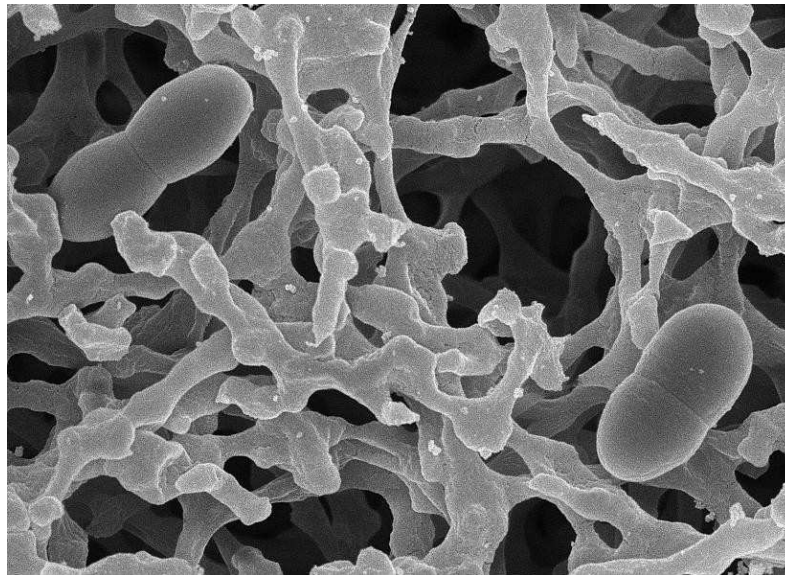
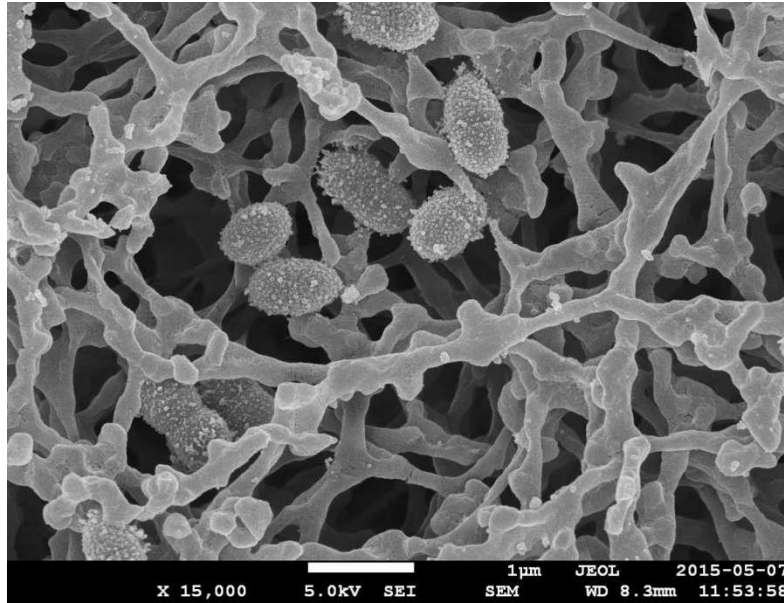


Figure 53: *Lactococcus lactis* subsp. *Lactis* DSM 20481 observée sous le microscope électronique à balayage (sur filtre de Nitrocellulose)

A



B : Détail:

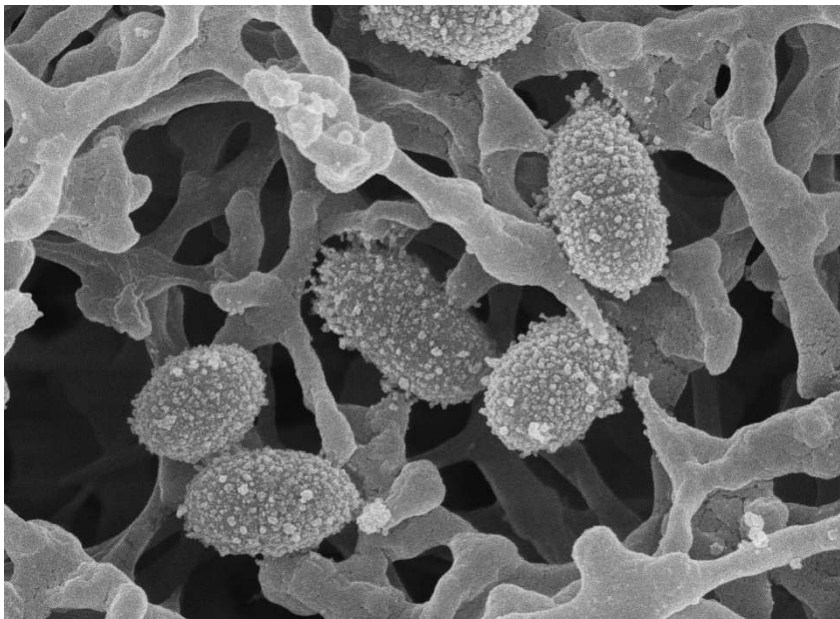


Figure 54: *Lactococcus lactis* subsp. *Lactis* HD9B_GL2 observée sous le microscope électronique à balayage (sur filtre de Nitrocellulose)

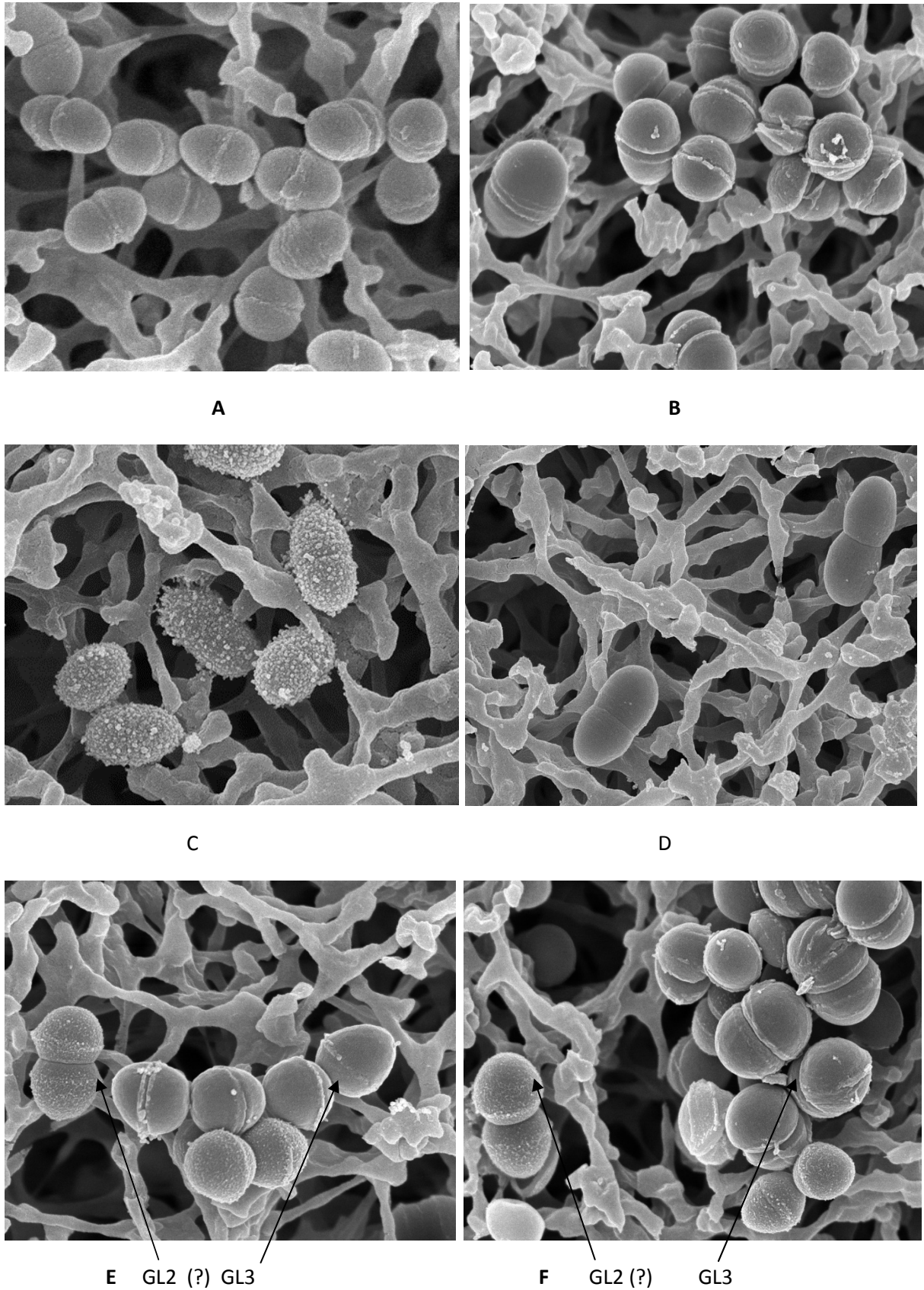


Figure 55: Observation au microscope électronique à balayage des trois constituants du consortium. **A** : *Leuconostoc mesenteroides* GL1. **B** : *Enterococcus faecium* GL3. **C** : *Lactococcus lactis* GL2. **D** : *Lactococcus lactis* DSM 20481. **E** et **F**: HD9B mixture des trois espèces.

3. Caractérisation moléculaire de la Biogenèse des pilis chez *Lactococcus lactis* HD9B_GL2 et leur rôle dans l'agrégation et la Formation de Biofilms

Le génome de la souche *Lactococcus lactis* HD9B_GL2 porte un cluster (Tableau 23) putatif pour la biogenèse des pili consistant : un gène sortase C flanqué par 3 gènes codant des protéines LPxTG (*yhgD*, *yhgE* et *yhhB*), appelés les pilis. Cependant, ces pilis n'ont pas été détecté sous des conditions de croissance optimales. La surexpression de l'opéron pili abouti à la production et l'adsorption des pili sur la surface des lactocoques. Analyse fonctionnelle de la machinerie de biogenèse de ces pilis a indiqué que l'axe de pilis est formé par des oligomères de la piline YhgE, la capsule est formée par la piline YhgD et que YhhB est la piline basale permettant l'attachement des fibres pilis à la paroi cellulaire.

L'oligomérisation des sous-unités de la piline est catalysée par sortase C, tandis que l'ancrage des pilis à la paroi cellulaire est effectué par la sortase A. La souche *L. lactis* HD9B_GL2 présente un phénotype d'auto-agrégation dans des cultures liquides, ce qui a été attribué à la polymérisation de la piline majeur YhgE. Ce lactocoque « pilié » forme des biofilms aériens plus épais par rapport à ceux produits par des bactéries « non-Piliés ». Ce phénotype est attribué à des oligomères de YhgE.

Cette étude fournit la première analyse moléculaire de la biogenèse des pilis chez le lactocoque thermophile isolé du lait de chamelle *L. lactis* HD9B_GL2, cette analyse montre la présence d'un cluster de gènes fonctionnel pour la production de pili dont le rôle est l'agrégation et la formation de biofilms. Cette propriété de la souche *L. lactis* HD9B_GL2, peut être importante pour sa survie en milieu naturel aussi bien que pour d'autres applications.

Résultats et discussion

Tableau 23: Piline chez HD9B_GL2

Locus_tag (GenBank)	Acides aminés	Pilin motif	E-box	LPxTG motif
LGL2_RS03440	1180	-	-	LPFAG
LGL2_RS03445	613	VYPK	YYLVEVAAPTG	LPFTG
LGL2_RS03455	241	IYPK	-	LLKTG

Lactococcus lactis appartient au groupe des bactéries lactiques (LAB), qui vivent généralement dans des niches écologiques riches en nutriments comme les plantes, le mucus de l'intestin et le lait. *L. lactis* est l'espèce la plus largement utilisée dans la fermentation du lait et est également la bactérie de LAB la plus caractérisée et la première dont le génome a été séquencé (Bolotin et al. 2001).

De nombreuses études sur la biologie de cette bactérie ont ouvert les portes à de nouvelles applications biotechnologiques dans lesquelles les lactocoques servent tant « qu'usine » de cellules et comme des vecteurs de molécules bénéfiques. Ce sont des antigènes ou des cytokines pour le développement de vaccins agissant au niveau des muqueuses ou, pour améliorer les thérapies immunomodulatoire (Mercenier et al. 2000; Steidler et al. 2000; Dieye et al. 2003; Bermudez-Humaran et al. 2005; Buccato et al. 2006; Hanniffy et al. 2007), comme des enzymes ou des vitamines pour l'amélioration d'état de santé des consommateurs (Hoefnagel et al. 2002; Hugenholtz et al. 2002; Kleerebezem et al. 2002), comme des antimicrobiens pour améliorer la sécurité alimentaire (Ross et al. 1999) et comme des lysines ou « holins » vis-à-vis des bactériophages d'origine alimentaire (Gaeng et al. 2000; Loessner 2005) ou pour accélérer la maturation des fromages (de Ruyter et al. 1997). Dans toutes ces applications, des interactions entre la surface des lactocoques avec l'environnement physique sont susceptibles d'influencer le comportement de bactéries et donc leur activité (O'Callaghan et al. 2012).

La surface *L. lactis* consiste en une paroi de peptidoglycane associés à des protéines, des acides teichoïques, des acides lipoteichoïques et des polysaccharides (Delcour et al. 1999; Chapot-Chartier et al. 2010). Les protéines de la paroi cellulaire sont des facteurs importants dans les interactions de la bactérie avec différentes surfaces biotiques ou abiotiques (Navarre and Schneewind 1999; Siciliano and Mazzeo 2012). Cela a été étudié chez les agents

Résultats et discussion

pathogènes (Navarre and Schneewind 1999) et dans une moindre mesure chez les lactocoques (Stentz et al. 2004; Habimana et al. 2007). Ces protéines, possédant un domaine d'ancrage dans leur partie C-terminal et présentant un motif de type LPxTG (dans lequel x peut être n'importe quel acide aminé); forment un type important de protéines de surface chez les bactéries Gram-positifs (Cossart and Jonquieres 2000). Ces protéines présentant le motif LPxTG sont sécrétées à travers la membrane par la voie Sec-dépendante et sont ensuite traitées par transpeptidases appelées sortases (Comfort and Clubb 2004; Dramsi et al. 2005; Marraffini and Schneewind 2006; Spirig et al. 2011). Dans certains cas, les substrats au motif LPxTG sont traités par une cystéine transpeptidase ubiquiste appelé sortase de ménage ou sortase de la classe A (SrtA) qui scinde la liaison Thr-Gly dans le motif LPxTG (L : leucine, P : proline, T : Thréonine, G : Glycine) et forme une autre liaison isopeptidique entre le groupe Thr-carboxyle C-terminal résultant et un groupement aminé dans l'intervalle interpeptidique du lipide II précurseur du peptidoglycane (Marraffini et al. 2006). Ce domaine est en position C-terminale de la séquence correspond au domaine d'ancrage de l'enzyme dans la paroi bactérienne. Il est caractérisé par la présence d'un motif consensus « LPxTG » typique de certaines protéines de surface chez les bactéries à Gram positif. La séquence « LPxTG » est clivée par une sortase entre les résidus thréonine et glycine ; puis l'extrémité COOH de la thréonine est liée par la sortase à un groupement amine libre de l'un des acides aminés des ponts peptidiques du peptidoglycane. En effet, il a été montré que la délétion de ce motif consensus conduit à la libération de l'enzyme dans le milieu extracellulaire (Haandrikman et al. 1989). L'archétype de ces substrats est la protéine A de *Staphylococcus aureus* dont le mécanisme d'ancrage de la sortase fut le premier à être caractérisée au niveau moléculaire (Schneewind et al. 1995; Mazmanian et al. 1999; Ton-That et al. 1999; Schneewind and Missiakas 2014). Des SrtA similaires ont été analysées du pont de vue fonctionnel dans plusieurs pathogènes à Gram positifs dans lesquels ils sont impliqués dans l'ancrage des protéines associées à la virulence (Barnett and Scott 2002; Garandeau et al. 2002; Osaki et al. 2002; Igarashi et al. 2003; Lalioui et al. 2005; Nobbs et al. 2007).

Une sortase A fonctionnelle est aussi présent dans *L. lactis* et participe à l'ancrage de plusieurs protéines qui jouent un rôle important dans la biologie de *L. lactis*, par exemple la protéinase ancrée dans la paroi cellulaire qui permet la croissance des lactocoques dans le lait (Dieye et al. 2010). Bien que SrtA mène au mécanisme de l'ancrage des protéines sous forme de monomères, certaines protéines avec le motif LPxTG contenant des acides aminés

Résultats et discussion

supplémentaires définis motifs ont un autre avenir différent et peuvent se polymériser en pilis ancrés à la surface cellulaire des bactéries Gram-positives.

Les pilis des bactéries Gram-positives ont tout d'abord été caractérisés chez les agents pathogènes dans lesquels ils sont impliqués dans l'adhérence et la virulence. Leur formation a été découvert chez *Actinomyces naeslundii* (Yeung and Ragsdale 1997; Yeung et al. 1998) et *Corynebacterium diphtheriae* (Ton-That and Schneewind 2003) et plus tard chez *Bacillus cereus* (Budzik et al. 2007), *Bacillus anthracis* (Budzik et al. 2008b), *Enterococcus faecalis* (Nallapareddy et al. 2006), *Enterococcus faecium* (Hendrickx et al. 2008), et chez des espèces du genre *Streptococcus*, dont *S. agalactiae* (Dramsai et al. 2005; Lauer et al. 2005), *S. pneumoniae* (Barocchi et al. 2006), *S. pyogenes* (Maione et al. 2005; Mora et al. 2005; Raddadi et al. 2005; Nakata et al. 2009) et *S. suis* (Fittipaldi et al. 2010). Les pilis sont des appendices protéiques de 1 à 10 nm de diamètre et de 2 à 3 µm en dehors des cellules bactériennes. Contrairement aux pili trouvés dans les bactéries à Gram-négatives qui sont des polymères de sous-unités liées par des liaisons non-covalente, ceux des bactéries Gram-positives sont le plus souvent hétéropolymères formé par 2 ou 3 protéines LPxTG appelées pilines qui sont polymérisées par des liaisons covalente par des réactions de transpeptidation catalysées par des pili spécifiques de classe sortases C (SrtC) (Proft and Baker 2009). Une fois assemblés, les fibres des pilis peuvent être ancrées au lipide II soit par SrtA (Swaminathan et al. 2007; Konto-Ghiorghi et al. 2009) ou par SrtC (LeMieux et al. 2008). Une de ces pilines est généralement appelée la piline dorsale « *the backbone pilin* » dès que 100 à 200 de ces sous-unités sont assemblées capsule-à-queue pour former l'axe des pilis (Kang et al. 2007). Les autres pilines sont appelées pilines auxiliaires et peuvent être situées exclusivement soit à la capsule et à la base du pili, respectivement ou s'étendent le long de l'axe du pili (Buccato et al. 2006; Telford et al. 2006). Outre leur motif d'ancrage LPxTG-like, les pilines dorsale toujours contenant 2 motifs consensus : un motif de piline WxxxVxVYPK et un motif YxLxETxAPxGY de E-box, appelé ainsi en raison de la présence d'un résidu hautement conservé d'acide glutamique (E) (Ton-That and Schneewind 2003; Ton-That et al. 2004). En ce qui concerne le motif piline, le résidu de lysine (K) est impliqué dans la liaison covalente d'une sous-unité de la piline dorsale avec le résidu de la thréonine (T) dans le motif LPxTG de la sous-unité suivante (Budzik et al. 2008a). La fonction du motif E-box n'est pas indispensable pour l'assemblage des pilis, mais il a été avérée essentiel dans l'incorporation des pilines auxiliaires dans les pilis et la formation de liaison intramoléculaire qui se produisent dans les pilines dorsale (Ton-That et al. 2004; Kang et al. 2007). Contrairement au

Résultats et discussion

pilines dorsale, les pilines et les motifs *E-box* semblent être dispensable pour les pilines auxiliaires soulève la question sur le mécanisme de leur incorporation dans les pilis dorsales. En ce qui concerne la piline de la tête, elle peut être liée par son motif LPxTG au motif de la piline du pili adjacent. Dans le cas de la piline basale, il a été démontré qu'un résidu lysine, quoique non incorporé dans un motif de piline reconnaissable, est impliqué dans une liaison amide avec le LPxTG de la piline dorsale du pili adjacent (Linke et al. 2010).

Jusqu'à présent, les pilis n'ont pas été observés chez les bactéries Gram-positives non pathogènes. Toutefois, la présence de pilis et des propriétés reliées à l'adhérence ont été signalés récemment dans certaines bactéries probiotiques telle que *Lactobacillus rhamnosus* GG et des membres du genre *Bifidobacterium* (Kankainen et al. 2009; Foroni et al. 2011; Gilad et al. 2011; Fanning et al. 2012; Reunanen et al. 2012). Chez *L. lactis*, on a récemment fourni des preuves génétiques et biochimiques d'une sortase A capable d'ancrer au moins 5 protéines au motif LPxTG à sa paroi cellulaire (Dieye et al. 2010).

L'étude a également révélé la présence d'un gène de la sortase C au sein d'un cluster de gènes des pilis évoquant la possibilité de la biogénèse des pilis chez cette bactérie. Bien que les pilis des lactocoques n'ont jamais été signalés, une étude dans laquelle *L. lactis* a été utilisée en tant qu'hôte pour l'expression des composants de l'opéron des pilis de *S. agalactiae*, a montré que *L. lactis* pourrait conduire à la biogénèse des pilis hétérologues quand le gène de la sortase C de *S. agalactiae* gène est exprimé tout comme les composants de l'opéron des pilis chez les streptocoques. En revanche, la surexpression uniquement des gènes de pilines de *S. agalactiae* est impossible d'activer la production de pilis chez *L. lactis* indiquant que les sortases C autochtones des lactocoques n'ont pas pu polymériser les pilines de *S. agalactiae* en pilis (Buccato et al. 2006; Quigley et al. 2010). Cette observation a soulevé la question de fonctionnalité la sortase C chez les lactocoques et/ou de son niveau d'expression dans les conditions expérimentales.

Dans notre analyse, nous avons pu caractériser le cluster de gènes de pilis chez *L. lactis* HD9B_GL2. Ce cluster est probablement fonctionnel car il est identique à 100% à celui étudié chez la souche *L. lactis* IL1403 et tous les motifs essentiels sont présents chez notre souche. La contribution des différents composants du cluster est signalée et nous montrons que ce caractère non déclarée chez les lactocoques fournit à cette bactérie un mode de vie différent en milieu liquide et sur des surfaces solides, induisant à auto-agrégation bactérienne et la formation de biofilms réticulés.

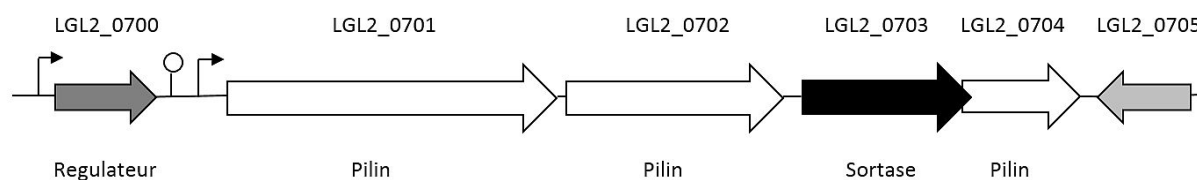


Figure 56: Représentation schématique du cluster de pilis chez *L. lactis* HD9B_GL2. 1) Les cadres de lecture pour LGL2_0703 et LGL2_0704 ont un overlap, le codon start pour LGL2_0704 est TTG. 2) Pour la région en amont de GL2_0701 BPROM a identifié un promoteur avec TTGCAA comme -35 et CAAAAA comme -10, et une distance de 19 pb, les séquences consensus sont TTGACA et TATAAT avec une distance de 17 pb. Par inspection visuelle on a trouvé une autre région avec TTTACA comme -35 et TTAAAT comme -10, avec la bonne distance de 17 bp. C'est-à-dire la similarité avec le consensus et la distance sont mieux pour ce dernier promoteur.

4. Analyse moléculaire d'un nouveau cluster de gènes *eps* chez *L. lactis* HD9B_GL2 et son investigation dans la micro-agrégation

Une autre fonction codée par les plasmides des lactocoques est la production d'exopolysaccharides (EPS) qui sont classés en deux grands groupes selon leur composition chimique : homopolysaccharides ou hétéropolysaccharides. Le premier groupe correspond à des polymères composés d'un seul type de saccharide (glucose, fructose ou galactose). Dans le cas des hétéropolysaccharides, ils correspondent à des polymères résultant de la combinaison de 3 à 8 saccharides (De Vuyst and Degeest 1999).

Les premiers gènes plasmidiques impliqués dans la production des EPS par les bactéries lactiques ont été identifiés et caractérisés chez *Lactococcus Lactis* subsp. *Cremoris* NIZO B40. Ces gènes sont portés par le plasmide pNZ4000 (42 kb) et s'organisent en un opéron *eps* de 12 kb (Hols et al. 1999a; Hols et al. 1999b; Kleerebezem et al. 1999; Looijesteijn et al. 1999; Neves et al. 1999; van Kranenburg et al. 1999a; van Kranenburg et al. 1999b; van Kranenburg et al. 1999c; van Kranenburg et al. 2000). Chez *Pediococcus damnosus* et *Oenococcus oeni* la production d'EPS est liée également à la présence de plasmide (Walling et al. 2005).

Résultats et discussion

Néanmoins, il est important de préciser que chez certaines bactéries lactiques le support génétique des EPS peut être localisé sur le chromosome ; c'est le cas des souches de *Streptococcus thermophilus* et de nombreux lactobacilles (Peant et al. 2005).

Les EPS des bactéries lactiques présentent une grande variabilité de structure au sein d'une même espèce bactérienne. Par exemple *Streptococcus thermophilus* est capable de synthétiser au moins 70 types d'EPS différents (Rallu *et al.*, 2001). De par leur diversité, leurs propriétés (stabilisantes, gélifiantes et épaississantes), les EPS produits par les bactéries lactiques sont d'un grand intérêt industriel. Ils trouvent principalement leur application dans l'industrie laitière. Par exemple l'obtention d'un yaourt brassé avec une texture homogène et stable résulte de la présence d'EPS produit par les souches utilisées dans le ferment (Jolly et al. 2002). Il a aussi été montré que les propriétés de rétention d'eau par les EPS améliorent la texture de fromage de type Mozzarella à faible teneur en matière grasse (Petersen et al. 2000).

Il est intéressant de noter que beaucoup des protéines qu'on ne trouve pas dans *L. l. ssp. lactis* IL1403 ou *L. l. ssp. lactis* KLDS4 ont des homologues dans *L. l. ssp. cremoris* A76 et sur un plasmide pCI658 (58 kb) isolé de *Lactococcus lactis ssp. cremoris* HO2 (O'Driscoll et al. 2006). Sur pCI658 il y a aussi 23 protéines EPS identiques avec les protéines de *L. lactis* HD9B_GL2.

Résultats et discussion

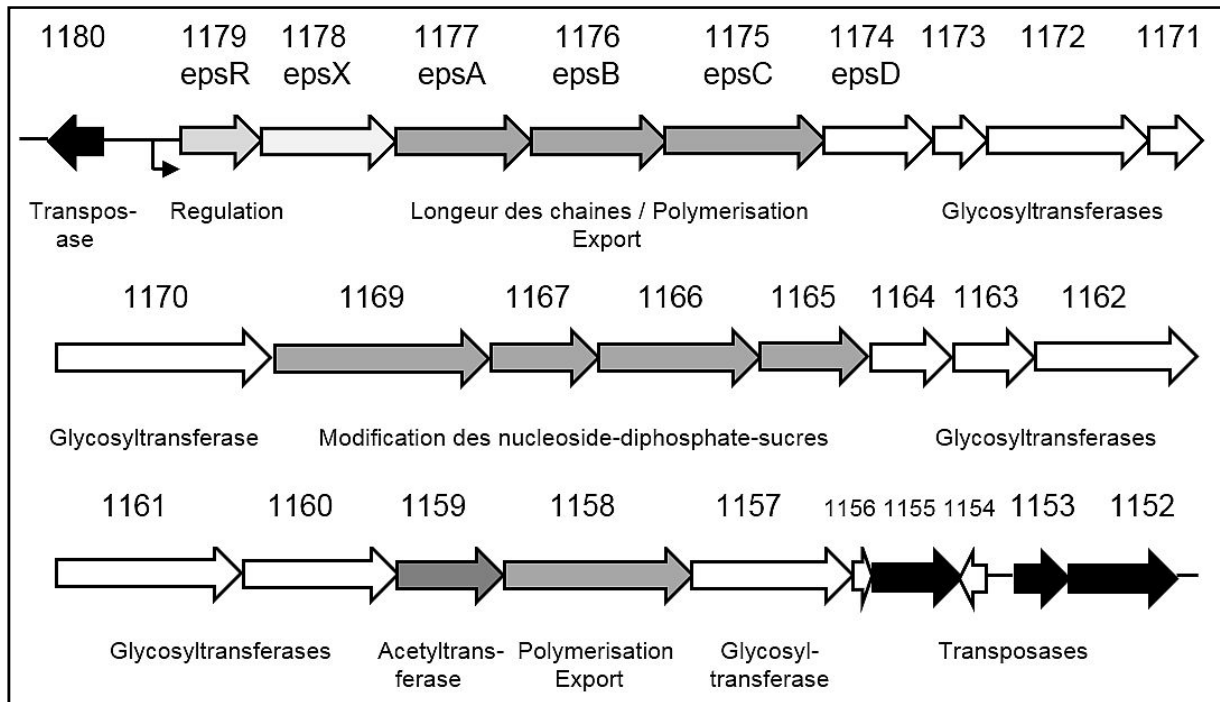


Figure 57: Représentation schématique du cluster de gènes *eps* chez *L. lactis* HD9B_GL2. Les fonctions prédites des gènes de la biosynthèse des *eps* (indiquées au-dessus des gènes) et d'export, de la polymérisation et de détermination de la longueur des chaînes sont indiquées.

Résultats et discussion

Tableau 24: Propriétés des protéines codées par le cluster des gènes EPS chez *L. lactis* ssp. *lactis* GL2

Gène	% GC / gène	Acide aminés/ protéine	% identité/similarité	Fonction de la protéine prédite
GL2_1180	36.8	86	NP_267352 (<i>L.lactis</i> IL1403 chromosome) 100 %	Transposase
GL2_1179 (epsR)	30.2	105	YP_009078166 (<i>L. lactis</i> plasmid pCI658) 99/99 %	Transcriptional regulator (HTH-XRE family)
GL2_1178 (epsX)	30.1	255	YP_009078167 (<i>L. lactis</i> plasmid pCI658) 95/96 %	SGNH-hydrolase, polysacch. biosynthesis protein
GL2_1177 (epsA)	33.8	259	YP_009078168 (<i>L. lactis</i> plasmid pCI658) 96/96 %	Tyrosine-protein kinase transmembrane modulator
GL2_1176 (epsB)	34.2	231	YP_009078169 (<i>L. lactis</i> plasmid pCI658) 95/97 %	Tyrosine-protein kinase
GL2_1175 (epsC)	33.3	254	YP_009078170 (<i>L. lactis</i> plasmid pCI658) 99/99 %	Tyrosine-protein phosphatase
GL2_1174 (epsD)	35.7	228	YP_009078171 (<i>L. lactis</i> plasmid pCI658) 95/97 %	Glycosyltransferase
GL2_1173	32.6	140	EHE93237 (<i>L. lactis</i> ssp. <i>lact.</i> CNM I1631) 97/98 % AAN63788Eps11F(<i>S. thermo-</i> <i>philus</i>) 99/99 %	Hexosyltransferase
GL2_1172	48.6	307	AEU39266 (<i>L. lactis</i> A76) 98/98 %; AAN64789 (<i>S.th.</i>) 349 aa; 96/96 %	Glyosyltransferase, Sugar epimerase

Résultats et discussion

GL2_1171	42.7	113	N-terminus of AEU39267, 416 aa(L. lactis A76) 97/97%; EHE93236, 113aa (L.lactis CNMI-1631 96 %	UDP-glucose dehydrogenase (N-terminal fragment)
GL2_1170	47.6	348	AEU39266 (L. lactis A76) 97/98 %	Glycosyltransferase
GL2_1169	44.7	416	AEU39267 (L. lactis A76) 96/98 %	UDP-glucose dehydrogenase
GL2_1167	51.1	231	AE006257_6 (L. lactis IL1403) 48/63 %	dTDP-glucose dehydratase
GL2_1166	47.6	304	KSU33547 (L. lactis) 95/97 %	dTDP-4-dehydrorhamnose reductase
GL2_1165	49.1	202	AEU39417 (L. lactis A76) 96/99 %	dTDP-4-dehydrorhamnose epimerase
GL2_1164	35.5	154	AF142639_11 EpsG (L. lactis plasmid pCI658) 45/63 %	Glycosyltransferase
GL2_1163	35.8	164	AF1422639_12 EpsH (L. lactis plasmid pCI658) 54/79 %	Glycosyltransferase
GL2_1162	31.4	307	KST89021 (L. lactis) 41/49 %	Glycosyltransferase
GL2_1161	33.2	388	ERK32250 (Clostridium intestinale) 45 % at C-terminus, no homology in Lacto-/Streptococcus	Glycosyltransferase
GL2_1160	31.9	345	AAP32727 EpsN (L. lactis plasmid pCI658) 41/56 %	Glycosyltransferase
GL2_1159	35.2	181	AAP32728 EpsQ (L. lactis plasmid pCI658) 43/60 %	Acetyltransferase
GL2_1158	32.3	420	AAP32729 EpsP (L. lactis plasmid pCI658) 26/45 %	Polymerization/Export of polysaccharide, see foot note
GL2_1157	34.1	321	CAL98283 (L. lactis MG1363) 38/59%	Glycosyltransferase
GL2_1156	30.0	54	No significant homology	
GL2_1155	43.0	185	No significant homology	Transposase (rve type)
GL2_1154	37.1	52	No significant homology	(CDS on reverse strand)

Résultats et discussion

GL2_1153	36.8	86	AAK05294 (L. lactis lactis IL1403) 99/99%	Transposase (fragment)
GL2_1152	39.2	296	AEU40780 (L. lactis A76) 99/99%	Transposase

- 1) GL2_1158 a une homologie significative avec EpsP du plasmide pCI658, les deux protéines ne contiennent aucun domaine conservé, cependant, elles ont diverses (10 chez GL2_1158) hélices transmembranaires selon l'analyse par le logiciel SMART (<http://smart.embl-heidelberg.de/>). Il a été montré (Forde & Fitzgerald 2003) que telles protéines peuvent agir comme des polymérase/ exportatrices des polysaccharides.
- 2) EspL, trouvée dans la plupart des clusters de biosynthèse d'EPS, n'est pas présente ici. Un autre cluster d'EPS fonctionnel sans EpsL a été décrit (Knoshaug et al. 2007), donc EpsL ne semble pas être indispensable.

Résultats et discussion

Pour certains gènes (8 gènes dans le tableau) le GC% est très élevé par rapport à celui du *Lactococcus lactis* HD9B_G12 (35.1%), ça veut dire qu'ils sont d'origine étrangère par un transfert horizontal de gènes peut être par un plasmide ou phage ou d'éléments transposant. L'analyse des régions flanquantes de ce groupe de gènes ont un GC% normale (faible) et il n'y a pas de traces d'éléments mobiles ce qui indique que ces gènes sont acquis par *Lactococcus* il y'a longtemps. L'usage des codants est très différent dans le génome de HD9B_G12 (Figure 25). Ces gènes sont évolués dans *Streptococcus thermophilus* et ensuite ils ont été acquis par *Lactococcus* et on n'a aucune idée de la manière de transfert de ces gènes.

Résultats et discussion

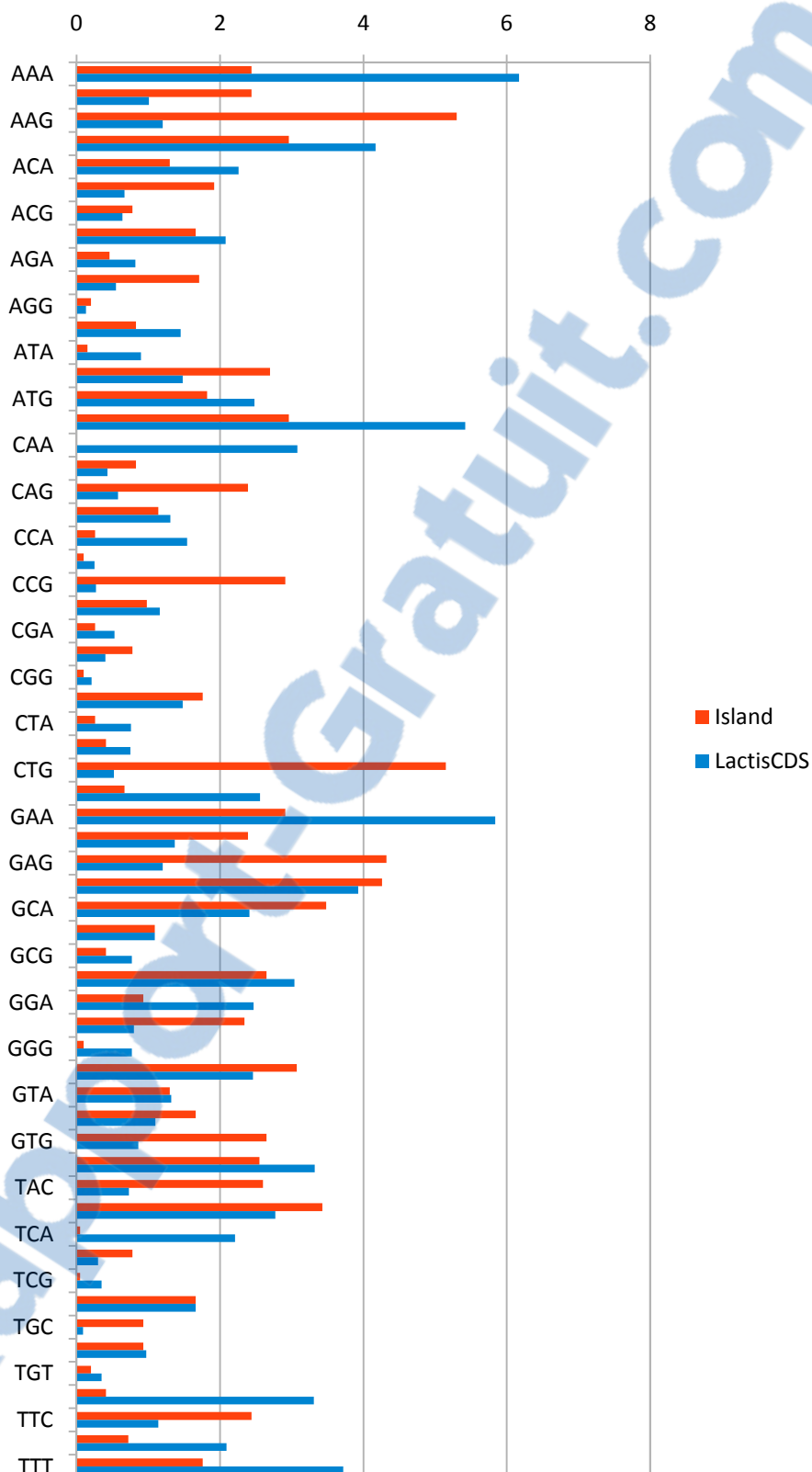


Figure 58 : Comparaison de l'usage de codons dans la séquence génomique de HD9B_GL2. Seul le codon CAA (Gln) n'a aucune correspondance dans cette région « island » en rouge. Cette région n'est pas régulière du contexte génomique de *Lactococcus lactis*. Elle peut provenir d'autres sous-espèces ou espèces différentes.

Résultats et discussion

Les molécules d'EPS produites par différent LAB possèdent une grande variété de propriétés chimiques et biophysiques, y compris les variations dans la composition en sucres, la liaison entre les sucres, longueur de polymères et la ramification des polymères. Ces polymères ont un rôle important dans les propriétés rhéologiques et de la texture des aliments fermentés et sont donc intéressants pour des applications alimentaires comme produit naturels bio-épaississants. En outre, ils pourraient aussi s'appliquer comme des additifs alimentaires naturelles, tout en remplaçant les stabilisateurs et les épaississants actuellement appliqués qui sont produits par d'autres microorganismes non comestible (Becker et al. 1998; Sutherland 1998). Il a été suggéré que certains polysaccharides produites par LAB ont des effets prébiotiques (Gibson and Roberfroid 1995), immunostimulantes (Hosono et al. 1997), antitumoraux (Toba et al. 1991; Yokoi et al. 1991)(Kitazawa et coll., 1991) ou une activité hypocholestérolémiant (Nakajima et al. 1992). Durant la dernière décennie, la production d'EPS par LAB a été particulièrement étudiée (De Vuyst and Degeest 1999), générant des avancées importantes dans la génétique des EPS, la diversité biologique et la distribution, les voies biosynthétiques, les modèles métaboliques et physiques. Cette connaissance peut être appliquée aux études d'ingénierie et du design métabolique afin de modifier la composition et la production d'EPS. La production d'EPS par *L. lactis* est associée aux souches isolées des produits laitiers fermentés hautement visqueux. Les gènes spécifiques *eps* sont encodés sur de grands plasmides qui peuvent être transférés par conjugaison d'un lactocoque à l'autre, introduisant ainsi la capacité de produire les EPS dans la souche réceptrice (Vedamuthu and Neville 1986; von Wright and Tynkkynen 1987; van Kranenburg et al. 1997). Parmi les clusters des gènes *eps* de *L. lactis* le plus caractérisé en détail, est celui de la souche NIZO B40 (Kleerebezem et al. 1999). Cette souche produit un polymère avec une unité de répétition régulière, $\rightarrow 4)[\alpha\text{-L-Rhap-(1}\rightarrow 2)][\alpha\text{-D-Galp-1-PO}_4\text{-3}]\text{-}\beta\text{-D-Galp-(1}\rightarrow 4)\text{-}\beta\text{-D-Glcp-(1}\rightarrow 4)\text{-}\beta\text{-D-Glcp-(1}\rightarrow$, qui est structurellement identique à ce produit par la souche SBT 0495 (Nakajima et al. 1992; van Kranenburg et al. 1997). La souche NIZO B40 possède un plasmide EPS pNZ4000 de 42 180 bp (van Kranenburg et al. 2000), contenant le cluster des gènes *eps* de 12Kb avec 14 gènes exprimés de façon coordonnée, *epsRX-ABCDEFGHIJKL*. En se basant sur des comparaisons de séquences, les fonctions putatives pourraient être assignées à plusieurs gènes *eps*, prédire leur implication dans la biosynthèse des unités répétitive d'oligosaccharides par addition séquentielle de sucres à un transporteur de lipide membranaire ancré et par l'exportation subséquente et la polymérisation de ces lipides-oligosaccharides. Des analyses fonctionnelle des gènes *epsDEFG*, codant des

Résultats et discussion

glycosyltransférases, ont généré un modèle pour la biosynthèse de l'unité de répétition des EPS chez la souche NIZO B40 (van Kranenburg et al. 1999b). Les études génétique des gènes *eps* de *L. lactis* NIZO B40 se concentrent sur l'élucidation du rôle des gènes qui sont prévus pour être impliqués dans l'exportation (*epsK*), la polymérisation (*epsI*) et dans la détermination de la longueur de chaîne (*epsA*, *epsB*) (van Kranenburg et al. 1997).

Ces études permettraient éventuellement d'une manipulation directe de ces fonctions, qui résulte d'une production des polysaccharides de courte ou de longues chaînes par rapport au polymère natif de la souche B40. En outre, la forte disponibilité des séquences de cluster de gènes *eps* divers organismes, y compris plusieurs LAB, a généré une information génétique riche dont un grand nombre de gènes codant pour la glycosyltransferase. L'information génétique, combinée à des données biochimiques concernant la spécificité donatrice et réceptrice d'enzymes de glycosyltransferase, pourraient être exploitées dans une approche combinatoire de glycosyltransferase visant à la construction de la biosynthèse pour la production d'oligo et/ou de polysaccharides chez *L. lactis* qui ont une structure nouvelle et prédéterminée. Cette dernière approche exige que le mécanisme d'export et de la polymérisation ne présente pas une spécificité pour ses polysaccharides correspondante, natif. Il est intéressant à noter qu'une expression hétérologue des gènes du cluster *eps* de *Streptococcus thermophilus* Sfi6 est observée chez *L. lactis* et a abouti à la production d'un polysaccharide contenant une unité de répétition altérée par rapport au polymère produit chez sa bactérie d'origine (Stingele et al. 1999). Ces résultats montrent que ni les glycosyltransférases, ni le mécanisme d'exportation et de la polymérisation codée par ces gènes du cluster *eps* ne montrent aucune spécificité exclusive pour leur substrat native, ce qui suggère que l'approche ci-dessus pourrait être réussie. Par ailleurs, un premier exemple d'une telle approche est la possibilité d'échanger des glycosyltransférases fonctionnelles avec une spécificité de substrat identique provenant de différents genres de bactéries à Gram positif (van Kranenburg et al. 1999a). Les stratégies actuelles et futures d'ingénierie des Eps devraient préciser les possibilités de ces approches vers la production et le 'designer' des polysaccharides. Outre les gènes spécifiques d'*eps* d'origine plasmidique, la production d'EPS chez *L. lactis* nécessite aussi des gènes chromosomiques impliqués dans la biosynthèse des nucléotides du sucre. Plusieurs études suggèrent que la disponibilité de ces blocs principales d'EPS pourrait fonctionner comme un facteur de contrôle important dans la biosynthèse des EPS (Looijesteijn et al. 1999; Drici et al.

Résultats et discussion

2010). Cette notion suggère qu'augmentant le flux vers ces intermédiaires pourrait entraîner une augmentation du niveau de la production d'EPS. Donc, pour être en mesure d'utiliser d'éventuelle ingénierie métabolique comme un moyen d'augmenter la production d'EPS par LAB, il est important d'évaluer la physiologie de la biosynthèse des nucléotides du sucre par rapport aux niveaux de production d'EPS. Une étape clé dans la régulation de la biosynthèse des nucléotides du sucre est l'inter-conversion de l'intermédiaire de la glycolyse le glucose - 6P en le précurseur de nucléotides de sucre glucose-1P, catalysée par la phosphoglucomutase (Vaughan et al. 1996; Ramos et al. 2001). Pour vérifier cette possibilité, le gène *pgm* d'*Escherichia coli* a été exprimé en quantité élevée chez *L. lactis* en utilisant le système NICE. Cette modulation du niveau de l'enzyme phosphoglucomutase a entraîné une forte augmentation des niveaux intracellulaires d'UDP-glucose et UDP-galactose, illustrant le rôle hypothétique de l'enzyme phosphoglucomutase dans le contrôle des niveaux des nucléotides du sucre. Toutefois, même si ces niveaux intermédiaires pourraient nettement augmenter, le niveau de production d'EPS chez B40 n'est pas affecté par cette modification (Boels et al. 2001). Des résultats similaires ont été obtenus pour la surproduction de la phosphoglucomutase chez *S. thermophilus* (Levander and Radstrom 2001). L'intermédiaire centrale, glucose-1P, doit être converti en UDP-glucose, UDP-galactose et dTDP-rhamnose pour répondre aux exigences du précurseur pour la biosynthèse de l'EPS chez B40. Tous les gènes de Lactocoques qui codent pour des enzymes potentiellement impliqués dans la biosynthèse des nucléotides du sucre du glucose-1-P (*galU*, *galE*, *rfaABCD*) ont été clonées et caractérisés (Luesink et al. 1998; Boels et al. 2001), générant des outils génétiques pour cibler les points dans la régulation de la biosynthèse des EPS. Récemment, il a été démontré que la production hétérologue de polysaccharides type 3 chez *Streptococcus pneumoniae* (des polysaccharides contenant du glucose et d'acide glucuronique, codés par le cluster *cps3D* et *cps3S*) chez *L. lactis* pourrait être considérablement augmentée par l'expression simultanée d'un analogue de GalU codé par le gène *cps3U* de pneumocoque (Gilbert et al. 2000). Ces résultats indiquent que le niveau endogène d'UDP-glucose chez *L. Lactis* ne suffit pas à soutenir la production de polysaccharides type 3 chez *S. pneumoniae*. Pour évaluer les effets de la modulation de l'activité de GalU sur la biosynthèse des nucléotides de sucre et les niveaux de production des EPS chez hB40, le niveau de la production du GalU chez les lactocoques a été augmenté jusqu'à 20 fois les niveaux de type sauvage, en utilisant le

Résultats et discussion

système de NICE. En conséquence, les niveaux intracellulaires de UDP-glucose et UDP-galactose se sont révélés être fortement augmentés, soutenant l'idée que l'activité enzymatique de GalU contrôle en effet la production de ces nucléotides de sucre dans les cellules de type sauvage (Boels et al. 2001). Cependant, cette disponibilité augmentée du précurseur n'a pas abouti à des niveaux plus élevés de la production d'EPS chez B40. Une autre enzyme essentielle dans la biosynthèse de l'EPS et des nucléotides de sucre est l'enzyme Leloir GalE. L'évaluation des niveaux nucléotides de sucre et de la production d'EPS dans une souche mutante *galE* de *L. lactis* cultivée sur le glucose comme seule source de carbone a révélé un niveau indétectable d'UDP-galactose et, par conséquent, une abolition complète de la production d'EPS. Cependant, la production d'EPS pourrait être récupérée par l'addition de petites quantités de galactose au milieu (Boels et al. 2001) (Grossiord, 1998 b). L'absence d'effet significatif de la disponibilité de l'UDP-glucose et d'UDP-galactose sur les niveaux de production d'EPS de B40 pourrait être le résultat de la disponibilité inchangée de dTDP-rhamnose, qui est le précurseur de la moitié de la chaîne latérale de B40-EPS et pourrait exercer un niveau élevé de contrôle sur le niveau final de la production d'EPS. Des données supportant l'importance de la disponibilité de dTDP-rhamnose pour la biosynthèse de polymères contenant du rhamnose ont été signalées chez des mutants *rfb* de *S. pneumoniae* (Morona et al. 1997a; b; Paton et al. 1997) et chez *Streptococcus mutans* (Tsukioka et al. 1997).

CONCLUSION ET PERSPECTIVES

Conclusion et perspectives

Rappel

Cette thèse est constituée de cinq chapitres. Le premier chapitre correspond à une synthèse bibliographique ayant pour but de situer le travail dans son contexte scientifique. Elle comporte une étude sur la génomique des bactéries lactiques, la réponse au stress thermique chez les lactocoques, l'impact de la bioinformatique sur l'étude des bactéries lactique, et les techniques du séquençage. Le deuxième chapitre est la présentation du matériel et méthodes utilisés dans cette étude. Le troisième chapitre, correspond aux résultats obtenus. Ce chapitre est divisé en trois sous-chapitres :

Un premier sous-chapitre qui comporte une étude sur la diversité génétique des lactocoques dans le lait de chamelle des régions arides d'Algérie. Dans ce chapitre, nous avons identifié des lactocoques à l'échelle de sous-espèce en se basant sur d'autres gènes « *single protein* » tel que le gène *gdpP* qui a un pouvoir discriminant plus important que le 16r ADN et avec lequel nous avons remarqué la présence de phénomène de transfert horizontal de gène chez certaine souche utilisée dans ce travail de thèse. Le deuxième sous-chapitre, est principalement une étude approfondie de l'analyse de la séquence génomique d'un lactococque thermotolerant afin d'identifier les déterminants génétiques des gènes *heat shock* impliqués dans la réponse au stress thermique chez ce lactococque ainsi que leur organisation génétique dans le chromosome bactérien, avec notamment un inventaire des gènes impliqués dans la réponse au stress thermique est présenté. Le troisième sous-chapitre, est une étude du *consortium* microbien présent dans le lait cru de chamelle d'Algérie, tout en se basant sur l'analyse métagénomique d'une microcommunauté microbienne.

L'impact de la bioinformatique sur les données génomiques des bactéries lactiques et leurs réponses au stress

La constitution d'un inventaire exhaustif des instructions, ou gènes, contenues dans un génome est la principale raison des programmes de séquençage. La connaissance des gènes est en effet une étape indispensable à la compréhension des phénomènes biologiques au niveau moléculaire et cellulaire. Dans les années à venir, les applications seront de plus en plus nombreuses dans les domaines de la médecine et des industries pharmaceutiques,

Conclusion et perspectives

biotechnologiques, agro- alimentaires, ainsi que dans d'autres domaines en prises directes avec les processus biologiques (agriculture, environnement). Pour toutes ces applications, la séquence est le point de départ. Ces méthodologies nouvelles apparaissent, découlant des avancées technologiques issues de l'évolution des connaissances en physique, informatique, chimie et en nanobiotechnologie qui visent d'une part la miniaturisation et donc la réduction des volumes réactionnelles, et le groupage des réactions.

Le séquençage du génome rapide et peu coûteux est devenu un outil magnifique pour la caractérisation des souches. Cela est démontré par le grand nombre de séquences génomiques disponibles ; par exemple l'Institut européen de la bioinformatique « the *European Bioinformatics Institute's homepage* » (<http://www.ebi.ac.uk/genomes/>) contient des séquences génomiques complètes de 100 souches de bactéries lactiques et de *Bifidobacterium* dans sa base de données.

En outre, les fournisseurs de cultures commerciales sont en possession des séquences génomiques d'un grand nombre de souches industrielles de ces mêmes genres. Par exemple pour l'espèce *Lactococcus lactis*, il y'a 14 génomes complètement séquencés, quatre séquences « *Scaffolds* » et environ 50 séquences sous « contigs » disponibles dans la base de donnée GenBank de NCBI. Les comparaisons de séquences génomiques jouent un rôle important dans les relations entre génotypes et phénotypes, surtout quand le phénotype est complexe ou mal compris. Par exemple, les lactocoques sont réputés être des mésophiles alors que la souche HD9B_GL2 est un lactocoque atypique thermotolérant dont nous sommes intéressés à étudier les déterminants génétiques responsables de ce caractère de thermotolerance. Ainsi, les approches de génomiques comparatives fournissent une excellente occasion de comparer les données génomiques afin de comprendre le mécanisme commun de la réponse au stress thermique, et d'identifier les fonctions et les gènes spécifiques à notre souche. La génomique comparative est inévitablement et peut être jumelée à la protéomique comparative afin de mieux comprendre encore plus profondément les fonctions des gènes en cas d'un stress thermique. La plupart des caractéristiques correspondant à un gène donné ont été effectués à l'aide d'une génomique comparative par hybridation d'ADN sur des puces "microarrays" contenant un génome entier. Ceci est utile pour la corrélation entre la présence et l'absence de gènes spécifiques avec des traits spécifiques, mais n'ont pas la résolution nécessaire pour identifier les allèles spécifique du gène avec des traits spécifiques correspondant. Une autre limite est que seuls les gènes présents sur la biopuce peuvent être détectés. Cette limitation peut être éliminée par une atteinte de résolution supérieure à l'aide

Conclusion et perspectives

de séquences de génomes complets. La connaissance du « *pangénome* » d'une espèce donnée peut également servir pour le criblage des souches possédantes des traits désirables ; les chances de succès sont considérablement si les gènes codant ces caractères en question sont déjà connus et présents dans l'espèce considérée. De même, l'amélioration des activités de la souche ne sera réussie que si le matériel génétique nécessaire est présent. En revanche, les souches ayant des propriétés uniques sont susceptibles de disposer d'un matériel génétique unique ou rare dans leur *pangénome*. La comparaison entre les génomes et les *pangenomes* ouvre des possibilités énormes pour sélectionner judicieusement les combinaisons de ferments lactiques fournissant des voies métaboliques pour le développement de saveur au fromage et autres produits fermentés. Le séquençage du génome est utile pour exclure la présence de contenus indésirables tels que les gènes de résistance aux antibiotiques et les facteurs de virulence dans les souches ou dans le *pangénome*. Ces connaissances permettront une détermination rapide si une séquence est une propriété intrinsèque d'une espèce (c'est-à-dire la partie du génome « *core genome* ») ou un caractère acquis qui a une importante régulation. Les fournisseurs de cultures commerciales ont souvent une multitude de souches de la même espèce dans leur gamme de produits. La production de ceux-ci est plus efficace si un processus de production unique et un milieu de culture unique peuvent être utilisés. La connaissance des fonctionnalités communes de toutes les souches d'une espèce provenant du séquençage du génome permet d'optimiser ces processus de fabrication et les milieux de culture.

L'étude de la réponse au stress chez les bactéries lactiques devrait permettre d'accroître la résistance des cellules aux contraintes imposées dans leur environnement en vue d'améliorer la qualité des levains utilisés pour l'ensemencement des milieux extrêmes. Cependant, si les diverses Hsp des autres bactéries lactiques telle que celles de *Bacillus subtilis* sont relativement bien caractérisées, les connaissances acquises sur les Hsp des lactocoques sont encore restreintes. Ce sont les Hsp les moins caractérisées au niveau de leur séquence. Les mécanismes de régulation de leur expression ainsi que leur rôle dans la réponse au stress constituent de larges champs à explorer. La réponse moléculaire au choc thermique est un phénomène tout à fait remarquable tant par son dynamisme que par son caractère ubiquiste.

Cette réponse n'est pas spécifique de l'hyperthermie, d'autres stress chimiques (arsénite, cadmium) et physiques (anoxie, stress hydrique, etc.) induisent parfois les mêmes protéines, poussant même certains auteurs à proposer les séquences *hsp* clonées comme sondes moléculaires permettant de diagnostiquer un état général de stress physiologique. D'autres

Conclusion et perspectives

assimilent le stress thermique à un stress oxydatif, argumentant que bon nombre d'agents inducteurs ont pour point commun de perturber les conditions normales d'oxydo-réduction au sein de la cellule. La signification physiologique de ces protéines reste mal perçue. Des arguments plaident en faveur de leur implication dans la thermotolérance, mais les cas d'acquisition de thermotolérance sans que soient détectée la moindre protéine de choc thermique sont pour le moins intrigants. Seule une connaissance approfondie de leur signification moléculaire, peut-être très diverse comme le sont les protéines elles-mêmes, permettra d'évaluer leur importance dans les différentes conditions physiologiques.

L'efficacité grandissante des techniques de transformation et la grande conservation des gènes HSP permettant un clonage aisé et fondant les espoirs de voir un gène isolé fonctionner en système hétérologue, devraient faciliter l'utilisation de ces gènes par génie génétique. Reste toutefois à identifier les gènes HSP hypothétiquement responsables d'une meilleure tolérance variétale à l'égard de la chaleur, à les isoler et à analyser les conditions de leur expression phénotypique.

Enfin, la réponse au choc thermique n'est pas qu'un remarquable modèle théorique de la dynamique d'expression du génome; elle attire l'attention du biologiste et de l'améliorateur sur les bases moléculaires complexes de l'homéostasie, cette aptitude des organismes à maintenir l'intégrité et l'activité de leurs structures moléculaires et cellulaires dans un environnement en continuelles fluctuations.

La connaissance du déterminisme moléculaire et génétique de l'homéostasie apparaît donc comme essentielle au progrès d'une industrie agroalimentaire plus soucieuse de stabiliser les rendements que de les amener aux plus hauts niveaux potentiels.

Apport technologique

Chez *L. Lactis* les séquences divergent très faiblement. Il serait intéressant de déterminer si les résultats obtenus pour ce lactocoque sont généralisables à d'autres souches de *lactococcus*. L'étude expérimentale complète et plus approfondies de la réponse au stress thermique chez *L. lactis* HD9B_G12 ainsi que l'étude des protéines HSP, leur synthèse et leur régulation nous permettra d'avoir plus des connaissances sur l'adaptation de cette bactérie et de pouvoir mieux rationaliser le choix des souches pour une utilisation industrielle.

Aujourd' hui avec la disponibilité de plusieurs génomes bactériens de diverses communautés microbiennes, il est possible d'appliquer des techniques post-génomiques à des écosystèmes

Conclusion et perspectives

et des études meta-génomiques commencent à émerger. De plus, le développement des programmes informatiques capables de traiter une très grande quantité de données permettra d'aller plus loin dans l'étude de ces écosystèmes.

Enfin, les séquences génomiques de *Leuconostoc mesenteroides* GL1, *Lactococcus lactis* HD9B_GL2 et *Enterococcus faecium* GL3 sont entièrement disponibles, ce qui permettra de comparer les gènes *hsp* présents dans les trois souches et d'avoir une idée de la variabilité possible entre les contenus génétiques *hsp* présents dans les différentes souches de *L. lactis*. Les résultats obtenus lors de l'analyse des données génomiques des gènes *hsp* de *L. lactis* GL2, doivent être complétés par des techniques protéomiques.

Perspectives

Ce travail a apporté une contribution conséquente sur les connaissances de l'écosystème lait de chamelle. De plus, il ouvre de nouvelles voies d'étude qui nous permettront de mieux comprendre l'adaptation des bactéries lactiques impliquées à leur milieu et à leur réponse au stress.

L'étude de *L. lactis*, permettrait d'une part de valider les résultats obtenus dans ce travail et d'autre part, de mieux connaître l'adaptation de cette bactérie dans le lait sous stress chaud. D'après l'analyse *in silico*, la souche *Lactococcus lactis* HD9B_GL2 possède tous les gènes nécessaires à la réponse au stress. Il serait donc intéressant d'avoir les données complètes en transcriptomique et en protéomique pour illustrer le caractère thermotolerant de *L. lactis* HD9B_GL2 sur différents milieux de culture. Ce qui permet l'identification des mécanismes pouvant intervenir dans la thermotolérance et ainsi avoir une idée plus exhaustive de ses mécanismes de régulation.

La disponibilité de la séquence génomique de *L. lactis* HD9B_GL2 ouvre la voie de la transcriptomique par la méthode des puces à ADN. Ces études permettraient d'avoir une vue globale des gènes dont l'expression est modifiée en condition de la thermotolérance et de rechercher une régulation commune à différents gènes par comparaison de leurs séquences promotrices (recherche de motifs conservés).

REFERENCES BIBLIOGRAPHIQUES

References bibliographique

A

Abnizova, I., Leonard, S., Skelly, T., Brown, A., Jackson, D., Gourtovaia, M., Qi, G., Te Boekhorst, R., Faruque, N., Lewis, K. and Cox, T. (2012) Analysis of context-dependent errors for illumina sequencing. *Journal of bioinformatics and computational biology* **10**, 1241005.

Adams, M.R. and Marteau, P. (1995) On the safety of lactic acid bacteria from food. *Int J Food Microbiol* **27**, 263-264.

Aguirre, M. and Collins, M.D. (1993) Lactic acid bacteria and human clinical infection. *The Journal of applied bacteriology* **75**, 95-107.

Ainsworth, S., Zomer, A., de Jager, V., Bottacini, F., van Hijum, S.A., Mahony, J. and van Sinderen, D. (2013) Complete Genome of *Lactococcus lactis* subsp. *cremoris* UC509.9, Host for a Model Lactococcal P335 Bacteriophage. *Genome announcements* **1**.

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M. and Aksoy, S. (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nature genetics* **32**, 402-407.

Allardet-Servent, A., Michaux-Charachon, S., Jumas-Bilak, E., Karayan, L. and Ramuz, M. (1993) Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J Bacteriol* **175**, 7869-7874.

Al Ait, L., Yamak, Z. and Morgenstern, B. (2013) DIALIGN at GOBICS--multiple sequence alignment using various sources of external information. *Nucleic Acids Res* **41**, W3-7.

Althaus, E., Caprara, A., Lenhof, H.P. and Reinert, K. (2002) Multiple sequence alignment with arbitrary gap costs: computing an optimal solution using polyhedral combinatorics. *Bioinformatics* **18 Suppl 2**, S4-S16.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Ackermann, H.W. (1998) Tailed bacteriophages: the order caudovirales. *Advances in virus research* **51**, 135-201.

Références bibliographique

- Ackermann, H.W. (1998)** Tailed bacteriophages: the order caudovirales. *Advances in virus research* **51**, 135-201.
- Ackermann, H.W. (2001)** Frequency of morphological phage descriptions in the year 2000. Brief review. *Archives of virology* **146**, 843-857.
- Ackermann, H.W. and Kropinski, A.M. (2007)** Curated list of prokaryote viruses with fully sequenced genomes. *Res Microbiol* **158**, 555-566.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.
- Araya, T., Ishibashi, N., Shimamura, S., Tanaka, K. and Takahashi, H. (1993)** Genetic and molecular analysis of the rpoD gene from *Lactococcus lactis*. *Bioscience, biotechnology, and biochemistry* **57**, 88-92.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O. (2008)** The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75.
- Ackermann, H.W. (2001)** Frequency of morphological phage descriptions in the year 2000. Brief review. *Archives of virology* **146**, 843-857.
- Ackermann, H.W. and Kropinski, A.M. (2007)** Curated list of prokaryote viruses with fully sequenced genomes. *Res Microbiol* **158**, 555-566.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.
- Araya, T., Ishibashi, N., Shimamura, S., Tanaka, K. and Takahashi, H. (1993)** Genetic and molecular analysis of the rpoD gene from *Lactococcus lactis*. *Bioscience, biotechnology, and biochemistry* **57**, 88-92.
- Arnosti, D.N., Singer, V.L. and Chamberlin, M.J. (1986)** Characterization of heat shock in *Bacillus subtilis*. *J Bacteriol* **168**, 1243-1249.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O. (2008)** The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75.

Références bibliographique

- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formosa, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O. (2008)** The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75.
- Agrawal, R.P., Swami, S.C., Beniwal, R., Kochar, D.K., Sahani, M.S., Tuteja, F.C. and Ghorui, S.K. (2003)** Effect of camel milk on glycemic control, lipid profile and diabetes quality of life in type 1 diabetes: A randomised prospective controlled cross over study. *Indian J Anim Sci* **73**, 1105-1110.
- Al Katheeri, N.A., Wasfi, I.A., Lambert, M., Giuliano Albo, A. and Nebbia, C. (2006)** In vivo and in vitro metabolism of dexamethasone in the camel. *Veterinary journal* **172**, 532-543.
- Amar, J., Chabo, C., Waget, A., Klopp, P., Vachoux, C., Bermudez-Humaran, L.G., Smirnova, N., Berge, M., Sulpice, T., Lahtinen, S., Ouwehand, A., Langella, P., Rautonen, N., Sansonetti, P.J. and Burcelin, R. (2011)** Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: molecular mechanisms and probiotic treatment. *EMBO molecular medicine* **3**, 559-572.
- Auffray, C. (1992)** DNA sequences. *Nature* **355**, 292.
- Aloy, P., Ceulemans, H., Stark, A. and Russell, R.B. (2003)** The relationship between sequence and interaction divergence in proteins. *J Mol Biol* **332**, 989-998.
- 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. and 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.
- Amar, J., Chabo, C., Waget, A., Klopp, P., Vachoux, C., Bermudez-Humaran, L.G., Smirnova, N., Berge, M., Sulpice, T., Lahtinen, S., Ouwehand, A., Langella, P., Rautonen, N., Sansonetti, P.J. and Burcelin, R. (2011)** Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: molecular mechanisms and probiotic treatment. *EMBO molecular medicine* **3**, 559-572.
- Arques, J.L., Rodriguez, E., Langa, S., Landete, J.M. and Medina, M. (2015)** Antimicrobial activity of lactic acid bacteria in dairy products and gut: effect on pathogens. *BioMed research international* **2015**, 584183.
- Azcarate-Peril, M.A., McAuliffe, O., Altermann, E., Lick, S., Russell, W.M. and Klaenhammer, T.R. (2005)** Microarray analysis of a two-component regulatory system

Références bibliographique

involved in acid resistance and proteolytic activity in *Lactobacillus acidophilus*. *Appl Environ Microbiol* **71**, 5794-5804.

Axelsson L. Lactic Acid Bacteria: Classification and Physiology. In: Salminen S, von Wright A, Ouwehand AC, editors. Lactic Acid Bacteria Microbiology and Functional Aspects. 3rd Edition. Marcel Dekker Inc; New York: 2004. pp. 1–66

Andrew Pollack. DNA sequencing caught in deluge of data. New York Times, 1, 2011.

Ayad, E.H., Verheul, A., Engels, W.J., Wouters, J.T. and Smit, G. (2001) Enhanced flavour formation by combination of selected lactococci from industrial and artisanal origin with focus on completion of a metabolic pathway. *J Appl Microbiol* **90**, 59-67.

B

Ben-Zvi, A.P. and Goloubinoff, P. (2001) Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. *Journal of structural biology* **135**, 84-93.

Bakke, P., Carney, N., Deloache, W., Gearing, M., Ingvorsen, K., Lotz, M., McNair, J., Penumetcha, P., Simpson, S., Voss, L., Win, M., Heyer, L.J. and Campbell, A.M. (2009) Evaluation of three automated genome annotations for *Halorhabdus utahensis*. *PLoS one* **4**, e6291.

Barril, J.S., Kim, S.G. and Batt, C.A. (1994) Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis* dnaK gene using a PCR-based approach. *Gene* **142**, 91-96.

Bissonnette, F., Labrie, S., Deveau, H., Lamoureux, M. and Moineau, S. (2000) Characterization of mesophilic mixed starter cultures used for the manufacture of aged cheddar cheese. *J Dairy Sci* **83**, 620-627.

Benachour, A., Muller, C., Dabrowski-Coton, M., Le Breton, Y., Giard, J.C., Rince, A., Auffray, Y. and Hartke, A. (2005) The *Enterococcus faecalis* sigV protein is an extracytoplasmic function sigma factor contributing to survival following heat, acid, and ethanol treatments. *J Bacteriol* **187**, 1022-1035.

Benson, A.K. and Haldenwang, W.G. (1993) The sigma B-dependent promoter of the *Bacillus subtilis* sigB operon is induced by heat shock. *J Bacteriol* **175**, 1929-1935.

Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K.,

Références bibliographique

- O'Neil, S., Rabinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. and Hopwood, D.A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.
- Bentley, S.D. and Parkhill, J. (2004) Comparative genomic structure of prokaryotes. *Annual review of genetics* **38**, 771-792.
- Barnett, T.C. and Scott, J.R. (2002) Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J Bacteriol* **184**, 2181-2191.
- Barocchi, M.A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., Hultenby, K., Taddei, A.R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D.W., Normark, S., Rappuoli, R. and Henriques-Normark, B. (2006) A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* **103**, 2857-2862.
- Barrette, J., Champagne, C.P., Roy, D. and Rodrigue, N. (2000) The production of mixed cultures containing strains of *Lactococcus lactis*, *Leuconostoc cremoris* and *Lactobacillus rhamnosus*, on commercial starter media. *Journal of industrial microbiology & biotechnology* **25**, 288-297.
- Becker, A., Katzen, F., Puhler, A. and Ielpi, L. (1998) Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Appl Microbiol Biotechnol* **50**, 145-152.
- Benmechernene, Z., Fernandez-No, I., Quintela-Baluja, M., Bohme, K., Kihal, M., Calo-Mata, P. and Barros-Velazquez, J. (2014) Genomic and proteomic characterization of bacteriocin-producing *Leuconostoc mesenteroides* strains isolated from raw camel milk in two southwest Algerian arid zones. *BioMed research international* **2014**, 853238.
- Bensalah, F., Delorme, C. and Renault, P. (2009) Characterisation of thermotolerant cocci from indigenous flora of 'leben' in algerian arid area and DNA identification of atypical *Lactococcus lactis* strains. *Curr Microbiol* **59**, 139-146.
- Bermudez-Humaran, L.G., Cortes-Perez, N.G., Lefevre, F., Guimaraes, V., Rabot, S., Alcocer-Gonzalez, J.M., Gratadoux, J.J., Rodriguez-Padilla, C., Tamez-Guerra, R.S., Corthier, G., Gruss, A. and Langella, P. (2005) A novel mucosal vaccine based on live *Lactococci* expressing E7 antigen and IL-12 induces systemic and mucosal immune responses and protects mice against human papillomavirus type 16-induced tumors. *Journal of immunology* **175**, 7297-7302.

Références bibliographique

- Boels, I.C., Ramos, A., Kleerebezem, M. and de Vos, W.M. (2001)** Functional analysis of the *Lactococcus lactis* galU and galE genes and their impact on sugar nucleotide and exopolysaccharide biosynthesis. *Appl Environ Microbiol* **67**, 3033-3040.
- Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F. and Corbeil, J. (2012)** Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* **13**, R122.
- Brady, A. and Salzberg, S.L. (2009)** Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. *Nature methods* **6**, 673-676.
- Brenner, K., You, L. and Arnold, F.H. (2008)** Engineering microbial consortia: a new frontier in synthetic biology. *Trends in biotechnology* **26**, 483-489.
- Brown, G.R., Hem, V., Katz, K.S., Ovetsky, M., Wallin, C., Ermolaeva, O., Tolstoy, I., Buccato, S., Maione, D., Rinaudo, C.D., Volpini, G., Taddei, A.R., Rosini, R., Telford, J.L., Grandi, G. and Margarit, I. (2006)** Use of *Lactococcus lactis* expressing pili from group B *Streptococcus* as a broad-coverage vaccine against streptococcal disease. *The Journal of infectious diseases* **194**, 331-340.
- Budzik, J.M., Marraffini, L.A. and Schneewind, O. (2007)** Assembly of pili on the surface of *Bacillus cereus* vegetative cells. *Mol Microbiol* **66**, 495-510.
- Budzik, J.M., Marraffini, L.A., Souda, P., Whitelegge, J.P., Faull, K.F. and Schneewind, O. (2008a)** Amide bonds assemble pili on the surface of bacilli. *Proc Natl Acad Sci U S A* **105**, 10215-10220.
- Budzik, J.M., Oh, S.Y. and Schneewind, O. (2008b)** Cell wall anchor structure of BcpA pili in *Bacillus anthracis*. *J Biol Chem* **283**, 36676-36686.
- Bischoff, M., Dunman, P., Kormanec, J., Macapagal, D., Murphy, E., Mounts, W., Berger-Bachi, B. and Projan, S. (2004)** Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. *J Bacteriol* **186**, 4085-4099.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997)** The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1462.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. (2012)** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology* **19**, 455-477.

Références bibliographique

- Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D. and Pirovano, W. (2011)** Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578-579.
- Boetzer, M. and Pirovano, W. (2014)** SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC bioinformatics* **15**, 211.
- Boisvert, S., Raymond, F., Godzaridis, E., Laviolette, F. and Corbeil, J. (2012)** Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* **13**, R122.
- Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I.A., Belmonte, M.K., Lander, E.S., Nusbaum, C. and Jaffe, D.B. (2008)** ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res* **18**, 810-820.
- Boekhorst, J., Siezen, R.J., Zwahlen, M.C., Vilanova, D., Pridmore, R.D., Mercenier, A., Kleerebezem, M., de Vos, W.M., Brussow, H. and Desiere, F. (2004)** The complete genomes of *Lactobacillus plantarum* and *Lactobacillus johnsonii* reveal extensive differences in chromosome organization and gene content. *Microbiology* **150**, 3601-3611.
- 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., Burtreau, S., Boutry, M., Delcour, J., Goffeau, A. and Hols, P. (2004)** Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature biotechnology* **22**, 1554-1558.
- Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, D.S. (2004)** Recent genetic transfer between *Lactococcus lactis* and enterobacteria. *J Bacteriol* **186**, 6671-6677.
- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001)** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* **11**, 731-753.
- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001)** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* **11**, 731-753.
- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001)** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* **11**, 731-753.
- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001)** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* **11**, 731-753.

Références bibliographique

- Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001)** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* **11**, 731-753.
- Borodina, I. and Nielsen, J. (2005)** From genomes to in silico cells via metabolic networks. *Current opinion in biotechnology* **16**, 350-355.
- Boutibonnes, P., Gillot, B., Auffray, Y. and Thammavongs, B. (1991)** Heat shock induces thermotolerance and inhibition of lysis in a lysogenic strain of *Lactococcus lactis*. *Int J Food Microbiol* **14**, 1-9.
- Bakke, P., Carney, N., Deloache, W., Gearing, M., Ingvorsen, K., Lotz, M., McNair, J., Penumetcha, P., Simpson, S., Voss, L., Win, M., Heyer, L.J. and Campbell, A.M. (2009)** Evaluation of three automated genome annotations for *Halorhabdus utahensis*. *PloS one* **4**, e6291.
- Barril, J.S., Kim, S.G. and Batt, C.A. (1994)** Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis* *dnaK* gene using a PCR-based approach. *Gene* **142**, 91-96.
- Bissonnette, F., Labrie, S., Deveau, H., Lamoureux, M. and Moineau, S. (2000)** Characterization of mesophilic mixed starter cultures used for the manufacture of aged cheddar cheese. *J Dairy Sci* **83**, 620-627.
- Boutibonnes, P., Tranchard, C., Hartke, A., Thammavongs, B. and Auffray, Y. (1992)** Is thermotolerance correlated to heat-shock protein synthesis in *Lactococcus lactis* subsp. *lactis*? *Int J Food Microbiol* **16**, 227-236.
- Baird, L. and Georgopoulos, C. (1990)** Identification, cloning, and characterization of the *Escherichia coli* *sohA* gene, a suppressor of the *htrA* (*degP*) null phenotype. *J Bacteriol* **172**, 1587-1594.
- Balcazar, J.L., de Blas, I., Ruiz-Zarzuola, I., Vendrell, D., Girones, O. and Muzquiz, J.L. (2007)** Sequencing of variable regions of the 16S rRNA gene for identification of lactic acid bacteria isolated from the intestinal microbiota of healthy salmonids. *Comparative immunology, microbiology and infectious diseases* **30**, 111-118.
- Barrios, C., Georgopoulos, C., Lambert, P.H. and Del Giudice, G. (1994)** Heat shock proteins as carrier molecules: in vivo helper effect mediated by *Escherichia coli* GroEL and DnaK proteins requires cross-linking with antigen. *Clinical and experimental immunology* **98**, 229-233.
- Baubekova, A., Akhmetsadykova, S., Konuspayeva, G., Akhmetsadykov, N., Faye, B. and Loiseau, G. (2015)** Biodiversity Study of the Yeast in Fresh and Fermented Camel and Mare's Milk by Denaturing Gradient Gel Electrophoresis. *J Camel Pract Res* **22**, 91-95.

Références bibliographique

- Bendimerad, N., Kihal, M. and Berthier, F. (2012)** Isolation, identification, and technological characterization of wild leuconostocs and lactococci for traditional Raib type milk fermentation. *Dairy Sci Technol* **92**, 249-264.
- Benmechernene, Z., Fernandez-No, I., Quintela-Baluja, M., Bohme, K., Kihal, M., Calomata, P. and Barros-Velazquez, J. (2014)** Genomic and proteomic characterization of bacteriocin-producing *Leuconostoc mesenteroides* strains isolated from raw camel milk in two southwest Algerian arid zones. *BioMed research international* **2014**, 853238.
- Bensalah, F., Delorme, C. and Renault, P. (2009)** Characterisation of thermotolerant cocci from indigenous flora of 'leben' in algerian arid area and DNA identification of atypical *Lactococcus lactis* strains. *Curr Microbiol* **59**, 139-146.
- Boguslawska, J., Zycka-Krzyszewska, J., Wilcks, A. and Bardowski, J. (2009)** Intra- and interspecies conjugal transfer of Tn916-like elements from *Lactococcus lactis* in vitro and in vivo. *Appl Environ Microbiol* **75**, 6352-6360.
- Bolliger, L., Deloche, O., Glick, B.S., Georgopoulos, C., Jenö, P., Kronidou, N., Horst, M., Morishima, N. and Schatz, G. (1994)** A mitochondrial homolog of bacterial GrpE interacts with mitochondrial hsp70 and is essential for viability. *EMBO J* **13**, 1998-2006.
- Brown, J.R. and Doolittle, W.F. (1999)** Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. *Journal of molecular evolution* **49**, 485-495.
- Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S. and Gross, C.A. (1981)** The nucleotide sequence of the cloned rpoD gene for the RNA polymerase sigma subunit from *E. coli* K12. *Nucleic Acids Res* **9**, 2889-2903.
- Boutibonnes, P., Bisson, V., Thammavongs, B., Hartke, A., Panoff, J.M., Benachour, A. and Auffray, Y. (1995)** Induction of thermotolerance by chemical agents in *Lactococcus lactis* subsp. *lactis* IL1403. *Int J Food Microbiol* **25**, 83-94.
- Boyd, D.A., Peters, G.A., Ng, L. and Mulvey, M.R. (2000)** Partial characterization of a genomic island associated with the multidrug resistance region of *Salmonella enterica* Typhimurium DT104. *FEMS Microbiol Lett* **189**, 285-291.
- Boylan, S.A., Redfield, A.R., Brody, M.S. and Price, C.W. (1993)** Stress-induced activation of the sigma B transcription factor of *Bacillus subtilis*. *J Bacteriol* **175**, 7931-7937.
- Brigulla, M., Hoffmann, T., Krisp, A., Volker, A., Bremer, E. and Volker, U. (2003)** Chill induction of the SigB-dependent general stress response in *Bacillus subtilis* and its contribution to low-temperature adaptation. *J Bacteriol* **185**, 4305-4314.

Références bibliographique

Bukau, B. and Horwich, A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell* **92**, 351-366.

Bukau, B. (1993) Regulation of the Escherichia coli heat-shock response. *Mol Microbiol* **9**, 671-680.

Bunthof, C.J., van den Braak, S., Breeuwer, P., Rombouts, F.M. and Abee, T. (1999) Rapid fluorescence assessment of the viability of stressed Lactococcus lactis. *Appl Environ Microbiol* **65**, 3681-3689.

C

Calendar, R., Erickson, J.W., Halling, C. and Nolte, A. (1988) Deletion and insertion mutations in the rpoH gene of Escherichia coli that produce functional sigma 32. *J Bacteriol* **170**, 3479-3484.

Cavanagh, D., Casey, A., Altermann, E., Cotter, P.D., Fitzgerald, G.F. and McAuliffe, O. (2015a) Evaluation of Lactococcus lactis Isolates from Nondairy Sources with Potential Dairy Applications Reveals Extensive Phenotype-Genotype Disparity and Implications for a Revised Species. *Appl Environ Microbiol* **81**, 3961-3972.

Cavanagh, D., Fitzgerald, G.F. and McAuliffe, O. (2015b) From field to fermentation: the origins of Lactococcus lactis and its domestication to the dairy environment. *Food microbiology* **47**, 45-61.

Chaisson, M.J., Brinza, D. and Pevzner, P.A. (2009) De novo fragment assembly with short mate-paired reads: Does the read length matter? *Genome Res* **19**, 336-346.

Chastanet, A., Derre, I., Nair, S. and Msadek, T. (2004) clpB, a novel member of the Listeria monocytogenes CtsR regulon, is involved in virulence but not in general stress tolerance. *J Bacteriol* **186**, 1165-1174.

Charpentier, E., Novak, R. and Tuomanen, E. (2000) Regulation of growth inhibition at high temperature, autolysis, transformation and adherence in Streptococcus pneumoniae by clpC. *Mol Microbiol* **37**, 717-726.

Chastanet, A., Fert, J. and Msadek, T. (2003) Comparative genomics reveal novel heat shock regulatory mechanisms in Staphylococcus aureus and other Gram-positive bacteria. *Mol Microbiol* **47**, 1061-1073.

Références bibliographique

- Chastanet, A. and Msadek, T. (2003)** ClpP of *Streptococcus salivarius* is a novel member of the dually regulated class of stress response genes in gram-positive bacteria. *J Bacteriol* **185**, 683-687.
- Chastanet, A., Prudhomme, M., Claverys, J.P. and Msadek, T. (2001)** Regulation of *Streptococcus pneumoniae* clp genes and their role in competence development and stress survival. *J Bacteriol* **183**, 7295-7307.
- Chopin, A., Deveau, H., Ehrlich, S.D., Moineau, S. and Chopin, M.C. (2007)** KSY1, a lactococcal phage with a T7-like transcription. *Virology* **365**, 1-9.
- Christiansen, B., Johnsen, M.G., Stenby, E., Vogensen, F.K. and Hammer, K. (1994)** Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. *J Bacteriol* **176**, 1069-1076.
- Case, R.J., Boucher, Y., Dahllorf, I., Holmstrom, C., Doolittle, W.F. and Kjelleberg, S. (2007)** Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol* **73**, 278-288.
- Chi, M.C., Chen, Y.H., Wu, T.J., Lo, H.F. and Lin, L.L. (2010)** Engineering of a truncated alpha-amylase of *Bacillus* sp. strain TS-23 for the simultaneous improvement of thermal and oxidative stabilities. *Journal of bioscience and bioengineering* **109**, 531-538.
- Claverys, J.P. and Martin, B. (2003)** Bacterial "competence" genes: signatures of active transformation, or only remnants? *Trends in microbiology* **11**, 161-165.
- Corrigan, R.M., Abbott, J.C., Burhenne, H., Kaefer, V. and Grundling, A. (2011)** c-di-AMP is a new second messenger in *Staphylococcus aureus* with a role in controlling cell size and envelope stress. *PLoS pathogens* **7**, e1002217.
- Craig, E.A., Gambill, B.D. and Nelson, R.J. (1993)** Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiological reviews* **57**, 402-414.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010)** QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335-336.
- Chapot-Chartier, M.P., Vinogradov, E., Sadovskaya, I., Andre, G., Mistou, M.Y., Trieu-Cuot, P., Furlan, S., Bidnenko, E., Courtin, P., Pechoux, C., Hols, P., Dufrene, Y.F. and Kulakauskas, S. (2010)** Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. *J Biol Chem* **285**, 10464-10471.

Références bibliographique

- Christensson, C., Pillidge, C.J., Ward, L.J. and O'Toole, P.W. (2001)** Nucleotide sequence and characterization of the cell envelope proteinase plasmid in *Lactococcus lactis* subsp. *cremoris* HP. *J Appl Microbiol* **91**, 334-343.
- Comfort, D. and Clubb, R.T. (2004)** A comparative genome analysis identifies distinct sorting pathways in gram-positive bacteria. *Infection and immunity* **72**, 2710-2722.
- Cossart, P. and Jonquieres, R. (2000)** Sortase, a universal target for therapeutic agents against gram-positive bacteria? *Proc Natl Acad Sci U S A* **97**, 5013-5015.
- Cottrell, M.T., Yu, L. and Kirchman, D.L. (2005)** Sequence and expression analyses of Cytophaga-like hydrolases in a Western arctic metagenomic library and the Sargasso Sea. *Appl Environ Microbiol* **71**, 8506-8513.
- Criscuolo, A. (2011)** morePhyML: improving the phylogenetic tree space exploration with PhyML 3. *Molecular phylogenetics and evolution* **61**, 944-948.
- Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Muller, W.E., Wetter, T. and Suhai, S. (2004)** Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* **14**, 1147-1159.
- Costello, M., Pugh, T.J., Fennell, T.J., Stewart, C., Lichtenstein, L., Meldrim, J.C., Fostel, J.L., Friedrich, D.C., Perrin, D., Dionne, D., Kim, S., Gabriel, S.B., Lander, E.S., Fisher, S. and Getz, G. (2013)** Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res* **41**, e67.
- Coucheney, F., Gal, L., Beney, L., Lherminier, J., Gervais, P. and Guzzo, J. (2005)** A small HSP, Lo18, interacts with the cell membrane and modulates lipid physical state under heat shock conditions in a lactic acid bacterium. *Biochim Biophys Acta* **1720**, 92-98.

D

- Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O.P., Devine, K.M. and van Dijl, J.M. (2002)** A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *J Bacteriol* **184**, 5661-5671.
- Daubin, V., Lerat, E. and Perriere, G. (2003)** The source of laterally transferred genes in bacterial genomes. *Genome Biol* **4**, R57.
- Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O.P., Devine, K.M. and van Dijl, J.M. (2002)** A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *J Bacteriol* **184**, 5661-5671.

Références bibliographique

- de Jong, A., Hansen, M.E., Kuipers, O.P., Kilstrup, M. and Kok, J. (2013)** The transcriptional and gene regulatory network of *Lactococcus lactis* MG1363 during growth in milk. *PLoS one* **8**, e53085.
- DeJesus, M.A., Sacchetti, J.C. and Ioerger, T.R. (2013)** Reannotation of translational start sites in the genome of *Mycobacterium tuberculosis*. *Tuberculosis* **93**, 18-25.
- Derre, I., Rapoport, G., Devine, K., Rose, M. and Msadek, T. (1999a)** ClpE, a novel type of HSP100 ATPase, is part of the CtsR heat shock regulon of *Bacillus subtilis*. *Mol Microbiol* **32**, 581-593.
- Derre, I., Rapoport, G. and Msadek, T. (1999b)** CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria. *Mol Microbiol* **31**, 117-131.
- Derre, I., Rapoport, G. and Msadek, T. (2000)** The CtsR regulator of stress response is active as a dimer and specifically degraded in vivo at 37 degrees C. *Mol Microbiol* **38**, 335-347.
- Deveau, H., Labrie, S.J., Chopin, M.C. and Moineau, S. (2006)** Biodiversity and classification of lactococcal phages. *Appl Environ Microbiol* **72**, 4338-4346.
- Dijkstra, A.R., Alkema, W., Starrenburg, M.J., Hugenholtz, J., van Hijum, S.A. and Bron, P.A. (2014)** Fermentation-induced variation in heat and oxidative stress phenotypes of *Lactococcus lactis* MG1363 reveals transcriptome signatures for robustness. *Microbial cell factories* **13**, 148.
- Dodd, H.M., Horn, N., Hao, Z. and Gasson, M.J. (1992)** A lactococcal expression system for engineered nisins. *Appl Environ Microbiol* **58**, 3683-3693.
- Drici, H., Gilbert, C., Kihal, M. and Atlan, D. (2010)** Atypical citrate-fermenting *Lactococcus lactis* strains isolated from dromedary's milk. *J Appl Microbiol* **108**, 647-657.
- Daly, C. (1983) The use of mesophilic cultures in the dairy industry. *Antonie van Leeuwenhoek* **49**, 297-312.
- Dobric, N., Limsowtin, G.K., Hillier, A.J., Dudman, N.P. and Davidson, B.E. (2000)** Identification and characterization of a cystathionine beta/gamma-lyase from *Lactococcus lactis* ssp. cremoris MG1363. *FEMS Microbiol Lett* **182**, 249-254.
- Doolittle, R.F. (1998)** Microbial genomes opened up. *Nature* **392**, 339-342.
- Dulermo, R., Fochesato, S., Blanchard, L. and de Groot, A. (2009)** Mutagenic lesion bypass and two functionally different RecA proteins in *Deinococcus deserti*. *Mol Microbiol* **74**, 194-208.

Références bibliographique

- Duplessis, M., Russell, W.M., Romero, D.A. and Moineau, S. (2005)** Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. *Virology* **340**, 192-208.
- Dupuis, M.E. and Moineau, S. (2010)** Genome organization and characterization of the virulent lactococcal phage 1358 and its similarities to *Listeria* phages. *Appl Environ Microbiol* **76**, 1623-1632.
- Duwat, P., Ehrlich, S.D. and Gruss, A. (1995)** The *recA* gene of *Lactococcus lactis*: characterization and involvement in oxidative and thermal stress. *Mol Microbiol* **17**, 1121-1131.
- de Crecy-Lagard, V., Servant-Moisson, P., Viala, J., Grandvalet, C. and Mazodier, P. (1999)** Alteration of the synthesis of the Clp ATP-dependent protease affects morphological and physiological differentiation in *Streptomyces*. *Mol Microbiol* **32**, 505-517.
- De Virgilio, C., Piper, P., Boller, T. and Wiemken, A. (1991)** Acquisition of thermotolerance in *Saccharomyces cerevisiae* without heat shock protein hsp 104 and in the absence of protein synthesis. *FEBS letters* **288**, 86-90.
- Delcour, J., Ferain, T., Deghorain, M., Palumbo, E. and Hols, P. (1999)** The biosynthesis and functionality of the cell-wall of lactic acid bacteria. *Antonie van Leeuwenhoek* **76**, 159-184.
- Delcour, J., Ferain, T. and Hols, P. (2000)** Advances in the genetics of thermophilic lactic acid bacteria. *Current opinion in biotechnology* **11**, 497-504.
- David, S., Stevens, H., van Riel, M., Simons, G. and de Vos, W.M. (1992)** *Leuconostoc lactis* beta-galactosidase is encoded by two overlapping genes. *J Bacteriol* **174**, 4475-4481.
- de Ruyter, P.G., Kuipers, O.P., Meijer, W.C. and de Vos, W.M. (1997)** Food-grade controlled lysis of *Lactococcus lactis* for accelerated cheese ripening. *Nature biotechnology* **15**, 976-979.
- de Vos, W.M. and Vaughan, E.E. (1994)** Genetics of lactose utilization in lactic acid bacteria. *FEMS Microbiol Rev* **15**, 217-237.
- De Vuyst, L. and Degeest, B. (1999)** Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiol Rev* **23**, 153-177.
- Delbes, C., Ali-Mandjee, L. and Montel, M.C. (2007)** Monitoring bacterial communities in raw milk and cheese by culture-dependent and -independent 16S rRNA gene-based analyses. *Appl Environ Microbiol* **73**, 1882-1891.

Références bibliographique

- Dieye, Y., Hoekman, A.J., Clier, F., Juillard, V., Boot, H.J. and Piard, J.C. (2003)** Ability of *Lactococcus lactis* to export viral capsid antigens: a crucial step for development of live vaccines. *Appl Environ Microbiol* **69**, 7281-7288.
- Dieye, Y., Oxaran, V., Ledue-Clier, F., Alkhalaf, W., Buist, G., Juillard, V., Lee, C.W. and Piard, J.C. (2010)** Functionality of sortase A in *Lactococcus lactis*. *Appl Environ Microbiol* **76**, 7332-7337.
- Dramsi, S., Trieu-Cuot, P. and Bierne, H. (2005)** Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res Microbiol* **156**, 289-297.
- Drider, D., Bekal, S. and Prevost, H. (2004)** Genetic organization and expression of citrate permease in lactic acid bacteria. *Genetics and molecular research : GMR* **3**, 273-281.
- Delmas, F., Pierre, F., Coucheney, F., Divies, C. and Guzzo, J. (2001)** Biochemical and physiological studies of the small heat shock protein Lo18 from the lactic acid bacterium *Oenococcus oeni*. *Journal of molecular microbiology and biotechnology* **3**, 601-610.
- Deutsch, S.M., Molle, D., Gagnaire, V., Piot, M., Atlan, D. and Lortal, S. (2000)** Hydrolysis of sequenced beta-casein peptides provides new insight into peptidase activity from thermophilic lactic acid bacteria and highlights intrinsic resistance of phosphopeptides. *Appl Environ Microbiol* **66**, 5360-5367.
- Dobrindt, U., Hochhut, B., Hentschel, U. and Hacker, J. (2004)** Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**, 414-424.
- Du, Y., Song, L., Feng, W., Pei, G., Zheng, P., Yu, Z., Sun, J. and Qiao, J. (2013)** Draft Genome Sequence of *Lactococcus lactis* subsp. *lactis* Strain YF11. *Genome announcements* **1**.

E

- Earnshaw, R.G., Appleyard, J. and Hurst, R.M. (1995)** Understanding physical inactivation processes: combined preservation opportunities using heat, ultrasound and pressure. *Int J Food Microbiol* **28**, 197-219.
- Eaton, T., Shearman, C. and Gasson, M. (1993)** Cloning and sequence analysis of the *dnaK* gene region of *Lactococcus lactis* subsp. *lactis*. *J Gen Microbiol* **139**, 3253-3264.
- Enright, M.C. and Spratt, B.G. (1999)** Multilocus sequence typing. *Trends in microbiology* **7**, 482-487.
- Ercan, O., Wels, M., Smid, E.J. and Kleerebezem, M. (2015)** Genome-wide transcriptional responses to carbon starvation in nongrowing *Lactococcus lactis*. *Appl Environ Microbiol* **81**, 2554-2561.

Références bibliographique

Ecroyd, H., Thorn, D.C., Liu, Y. and Carver, J.A. (2010) The dissociated form of kappa-casein is the precursor to its amyloid fibril formation. *The Biochemical journal* **429**, 251-260.

Erkus, O., de Jager, V.C., Spus, M., van Alen-Boerrigter, I.J., van Rijswijk, I.M., Hazelwood, L., Janssen, P.W., van Hijum, S.A., Kleerebezem, M. and Smid, E.J. (2013) Multifactorial diversity sustains microbial community stability. *Isme J* **7**, 2126-2136.

F

Falentin, H., Naquin, D., Loux, V., Barloy-Hubler, F., Loubiere, P., Nouaille, S., Lavenier, D., Le Bourgeois, P., Francois, P., Schrenzel, J., Hernandez, D., Even, S. and Le Loir, Y. (2014) Genome Sequence of *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* LD61. *Genome announcements* **2**.

Fan, J.B., Chee, M.S. and Gunderson, K.L. (2006) Highly parallel genomic assays. *Nature reviews Genetics* **7**, 632-644.

Ferain, T., Hobbs, J.N., Jr., Richardson, J., Bernard, N., Garmyn, D., Hols, P., Allen, N.E. and Delcour, J. (1996a) Knockout of the two *ldh* genes has a major impact on peptidoglycan precursor synthesis in *Lactobacillus plantarum*. *J Bacteriol* **178**, 5431-5437.

Ferain, T., Schanck, A.N. and Delcour, J. (1996b) ¹³C nuclear magnetic resonance analysis of glucose and citrate end products in an *ldhL-ldhD* double-knockout strain of *Lactobacillus plantarum*. *J Bacteriol* **178**, 7311-7315.

Farah, Z. (1993) Composition and characteristics of camel milk. *The Journal of dairy research* **60**, 603-626.

Faye, B. and Konuspayeva, G. (2012) The sustainability challenge to the dairy sector – The growing importance of non-cattle milk production worldwide. *Int Dairy J* **24**, 50-56.

Faye, B., Seboussi, R. and Askar, M. (2005) Trace elements and heavy metals in healthy camel blood of United Arab Emirates. *J Camel Pract Res* **12**, 1-6.

Flahaut, S., Hartke, A., Giard, J.C., Benachour, A., Boutibonnes, P. and Auffray, Y. (1996) Relationship between stress response toward bile salts, acid and heat treatment in *Enterococcus faecalis*. *FEMS Microbiol Lett* **138**, 49-54.

Ferdows, M.S. and Barbour, A.G. (1989) Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. *Proc Natl Acad Sci U S A* **86**, 5969-5973.

Fiegler, H., Redon, R., Andrews, D., Scott, C., Andrews, R., Carder, C., Clark, R., Dovey, O., Ellis, P., Feuk, L., French, L., Hunt, P., Kalaitzopoulos, D., Larkin, J., Montgomery, L., Perry, G.H., Plumb, B.W., Porter, K., Rigby, R.E., Rigler, D., Valsesia,

Références bibliographique

- A., Langford, C., Humphray, S.J., Scherer, S.W., Lee, C., Hurles, M.E. and Carter, N.P. (2006)** Accurate and reliable high-throughput detection of copy number variation in the human genome. *Genome Res* **16**, 1566-1574.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M. and et al. (1995)** Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.
- Florez, A.B., Ammor, M.S. and Mayo, B. (2008)** Identification of tet(M) in two *Lactococcus lactis* strains isolated from a Spanish traditional starter-free cheese made of raw milk and conjugative transfer of tetracycline resistance to lactococci and enterococci. *Int J Food Microbiol* **121**, 189-194.
- Fox, G.E., Magrum, L.J., Balch, W.E., Wolfe, R.S. and Woese, C.R. (1977)** Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc Natl Acad Sci U S A* **74**, 4537-4541.
- Fox, P. F. (1993).** Cheese : an overview. In *Cheese : chemistry, physics and microbiology*, pp. 1-36. Edited by P. F. Fox. London: Chapman and Hall.
- Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D., Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A., 3rd and Venter, J.C. (1995)** The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397-403.
- Frees, D. and Ingmer, H. (1999)** ClpP participates in the degradation of misfolded protein in *Lactococcus lactis*. *Mol Microbiol* **31**, 79-87.
- Frees, D., Thomsen, L.E. and Ingmer, H. (2005)** *Staphylococcus aureus* ClpYQ plays a minor role in stress survival. *Archives of microbiology* **183**, 286-291.
- Frees, D., Savijoki, K., Varmanen, P. and Ingmer, H. (2007)** Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol Microbiol* **63**, 1285-1295.
- Ferguson, P.L. and Coombs, D.H. (2000)** Pulse-chase analysis of the in vivo assembly of the bacteriophage T4 tail. *J Mol Biol* **297**, 99-117.
- Fanning, S., Hall, L.J., Cronin, M., Zomer, A., MacSharry, J., Goulding, D., Motherway, M.O., Shanahan, F., Nally, K., Dougan, G. and van Sinderen, D. (2012)** Bifidobacterial

Références bibliographique

surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proc Natl Acad Sci U S A* **109**, 2108-2113.

Field, D., Sterk, P., Kottmann, R., De Smet, J.W., Amaral-Zettler, L., Cochrane, G., Cole, J.R., Davies, N., Dawyndt, P., Garrity, G.M., Gilbert, J.A., Glockner, F.O., Hirschman, L., Klenk, H.P., Knight, R., Kyrpides, N., Meyer, F., Karsch-Mizrachi, I., Morrison, N., Robbins, R., San Gil, I., Sansone, S., Schriml, L., Tatusova, T., Ussery, D., Yilmaz, P., White, O., Wooley, J. and Caporaso, G. (2014) Genomic standards consortium projects. *Standards in genomic sciences* **9**, 599-601.

Fittipaldi, N., Takamatsu, D., de la Cruz Dominguez-Punaro, M., Lecours, M.P., Montpetit, D., Osaki, M., Sekizaki, T. and Gottschalk, M. (2010) Mutations in the gene encoding the ancillary pilin subunit of the *Streptococcus suis* srtF cluster result in pili formed by the major subunit only. *PloS one* **5**, e8426.

Froni, E., Serafini, F., Amidani, D., Turroni, F., He, F., Bottacini, F., O'Connell Motherway, M., Viappiani, A., Zhang, Z., Rivetti, C., van Sinderen, D. and Ventura, M. (2011) Genetic analysis and morphological identification of pilus-like structures in members of the genus *Bifidobacterium*. *Microbial cell factories* **10 Suppl 1**, S16.

Fernandez, M., Kleerebezem, M., Kuipers, O.P., Siezen, R.J. and van Kranenburg, R. (2002) Regulation of the metC-cysK operon, involved in sulfur metabolism in *Lactococcus lactis*. *J Bacteriol* **184**, 82-90.

Fernandez, M., van Doesburg, W., Rutten, G.A., Marugg, J.D., Alting, A.C., van Kranenburg, R. and Kuipers, O.P. (2000) Molecular and functional analyses of the metC gene of *Lactococcus lactis*, encoding cystathionine beta-lyase. *Appl Environ Microbiol* **66**, 42-48.

Fortier, L.C., Bransi, A. and Moineau, S. (2006) Genome sequence and global gene expression of Q54, a new phage species linking the 936 and c2 phage species of *Lactococcus lactis*. *J Bacteriol* **188**, 6101-6114.

Foucaud-Scheunemann, C. and Poquet, I. (2003) HtrA is a key factor in the response to specific stress conditions in *Lactococcus lactis*. *FEMS Microbiol Lett* **224**, 53-59.

Fujii, T., Tomita, Y., Ikushima, S., Horie, A. and Fujiwara, D. (2015) Draft Genome Sequence of *Lactococcus lactis* subsp. *lactis* JCM 5805T, a Strain That Induces Plasmacytoid Dendritic Cell Activation. *Genome announcements* **3**.

Fuller, R. (1991) Probiotics in human medicine. *Gut* **32**, 439-442.

Références bibliographique

Fuller, C.W., Middendorf, L.R., Benner, S.A., Church, G.M., Harris, T., Huang, X., Jovanovich, S.B., Nelson, J.R., Schloss, J.A., Schwartz, D.C. and Vezenov, D.V. (2009) The challenges of sequencing by synthesis. *Nature biotechnology* **27**, 1013-1023.

G

Gabed, N., Yang, M., Bey Baba Hamed, M., Drici, H., Gross, R., Dandekar, T. and Liang, C. (2015) Draft Genome Sequence of the Moderately Heat-Tolerant *Lactococcus lactis* subsp. *lactis* bv. *diacetylactis* Strain GL2 from Algerian Dromedary Milk. *Genome announcements* **3**.

Gao, Y., Lu, Y., Teng, K.L., Chen, M.L., Zheng, H.J., Zhu, Y.Q. and Zhong, J. (2011) Complete genome sequence of *Lactococcus lactis* subsp. *lactis* CV56, a probiotic strain isolated from the vaginas of healthy women. *J Bacteriol* **193**, 2886-2887.

Gilliland, S.E., Nelson, C.R. and Maxwell, C. (1985) Assimilation of cholesterol by *Lactobacillus acidophilus*. *Appl Environ Microbiol* **49**, 377-381.

Glenn, T.C. (2011) Field guide to next-generation DNA sequencers. *Molecular ecology resources* **11**, 759-769.

Goldin, B.R. and Gorbach, S.L. (2008) Clinical indications for probiotics: an overview. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **46 Suppl 2**, S96-100; discussion S144-151.

Gottesman, S. (1996) Proteases and their targets in *Escherichia coli*. *Annual review of genetics* **30**, 465-506.

Gottesman, S. (1999) Regulation by proteolysis: developmental switches. *Current opinion in microbiology* **2**, 142-147.

Grandvalet, C., Coucheney, F., Beltramo, C. and Guzzo, J. (2005) CtsR is the master regulator of stress response gene expression in *Oenococcus oeni*. *J Bacteriol* **187**, 5614-5623.

Guo, Y., Li, J., Li, C.I., Long, J., Samuels, D.C. and Shyr, Y. (2012) The effect of strand bias in Illumina short-read sequencing data. *BMC genomics* **13**, 666.

Gansel, X., Dutreix, M., Hartke, A., Boutibonnes, P. and Auffray, Y. (1993a) Partial characterization of an *rpoD*-like gene of *Lactococcus lactis* subsp. *lactis* ML3 with a polymerase chain reaction-based approach. *Curr Microbiol* **27**, 267-271.

Gansel, X., Hartke, A., Boutibonnes, P. and Auffray, Y. (1993b) Nucleotide sequence of the *Lactococcus lactis* NCDO 763 (ML3) *rpoD* gene. *Biochim Biophys Acta* **1216**, 115-118.

Références bibliographique

- Garneau, J.E., Tremblay, D.M. and Moineau, S. (2008)** Characterization of 1706, a virulent phage from *Lactococcus lactis* with similarities to prophages from other Firmicutes. *Virology* **373**, 298-309.
- Gasson, M.J., Swindell, S., Maeda, S. and Dodd, H.M. (1992)** Molecular rearrangement of lactose plasmid DNA associated with high-frequency transfer and cell aggregation in *Lactococcus lactis* 712. *Mol Microbiol* **6**, 3213-3223.
- Glover, J.R. and Lindquist, S. (1998)** Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* **94**, 73-82.
- Griffin, H.G., Swindell, S.R. and Gasson, M.J. (1992)** Cloning and sequence analysis of the gene encoding L-lactate dehydrogenase from *Lactococcus lactis*: evolutionary relationships between 21 different LDH enzymes. *Gene* **122**, 193-197.
- Galtier, N., Gouy, M. and Gautier, C. (1996)** SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Computer applications in the biosciences : CABIOS* **12**, 543-548.
- Garvie, E.I., Cole, C.B., Fuller, R. and Hewitt, D. (1984)** The effect of yoghurt on some components of the gut microflora and on the metabolism of lactose in the rat. *The Journal of applied bacteriology* **56**, 237-245.
- Georgopoulos, C. and Ang, D. (1990)** The *Escherichia coli* groE chaperonins. *Seminars in cell biology* **1**, 19-25.
- Gaeng, S., Scherer, S., Neve, H. and Loessner, M.J. (2000)** Gene cloning and expression and secretion of *Listeria monocytogenes* bacteriophage-lytic enzymes in *Lactococcus lactis*. *Appl Environ Microbiol* **66**, 2951-2958.
- Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. and Berche, P. (2000)** The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol Microbiol* **35**, 1286-1294.
- Gaillot, O., Bregenholt, S., Jaubert, F., Di Santo, J.P. and Berche, P. (2001)** Stress induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. *Infection and immunity* **69**, 4938-4943.
- Garandeau, C., Reglier-Poupet, H., Dubail, I., Beretti, J.L., Berche, P. and Charbit, A. (2002)** The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence. *Infection and immunity* **70**, 1382-1390.
- Gibson, G.R. and Roberfroid, M.B. (1995)** Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *The Journal of nutrition* **125**, 1401-1412.

Références bibliographique

- Gilad, O., Svensson, B., Viborg, A.H., Stuer-Lauridsen, B. and Jacobsen, S. (2011)** The extracellular proteome of *Bifidobacterium animalis* subsp. *lactis* BB-12 reveals proteins with putative roles in probiotic effects. *Proteomics* **11**, 2503-2514.
- Gilbert, C., Robinson, K., Le Page, R.W. and Wells, J.M. (2000)** Heterologous expression of an immunogenic pneumococcal type 3 capsular polysaccharide in *Lactococcus lactis*. *Infection and immunity* **68**, 3251-3260.
- Georgopoulos, C. and Welch, W.J. (1993)** Role of the major heat shock proteins as molecular chaperones. *Annu Rev Cell Biol* **9**, 601-634.
- Griffiths, J.M. and O'Neill, A.J. (2012)** Loss of function of the gdpP protein leads to joint beta-lactam/glycopeptide tolerance in *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy* **56**, 579-581.
- Galardini, M., Biondi, E.G., Bazzicalupo, M. and Mengoni, A. (2011)** CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source code for biology and medicine* **6**, 11.
- Gerber, S.D. and Solioz, M. (2007)** Efficient transformation of *Lactococcus lactis* IL1403 and generation of knock-out mutants by homologous recombination. *Journal of basic microbiology* **47**, 281-286.
- Goebel, B.M. and Stackebrandt, E. (1994)** Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. *Appl Environ Microbiol* **60**, 1614-1621.
- Gonzalez, F.C., Mans de Marion, M.C., Raffellini, S., Sanahuja, M.C. and Fantuzzi, L. (1990)** [Leuconostocs in raw milk: isolation and identification of species rapidly utilizing lactose]. *Revista Argentina de microbiologia* **22**, 86-89.
- Guindon, S. and Gascuel, O. (2003)** A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology* **52**, 696-704.

H

- Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R. and Goebel, W. (1990)** Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microbial pathogenesis* **8**, 213-225.
- Haller, D., Serrant, P., Granato, D., Schiffrin, E.J. and Blum, S. (2002)** Activation of human NK cells by staphylococci and lactobacilli requires cell contact-dependent

Références bibliographique

costimulation by autologous monocytes. *Clinical and diagnostic laboratory immunology* **9**, 649-657.

Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**, 669-685.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J. and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & biology* **5**, R245-249.

Hansen, M.C., Nielsen, A.K., Molin, S., Hammer, K. and Kilstrup, M. (2001) Changes in rRNA levels during stress invalidates results from mRNA blotting: fluorescence in situ rRNA hybridization permits renormalization for estimation of cellular mRNA levels. *J Bacteriol* **183**, 4747-4751.

Halasz, Lactic Acid Bacteria. In: Lasztity R, editor. Food Quality and Standards (Vol. 3). EOLSS Publishers Co Ltd, UK, (2009).

Harris, T.D., Buzby, P.R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., Causey, M., Colonell, J., Dimeo, J., Efcavitch, J.W., Giladi, E., Gill, J., Healy, J., Jarosz, M., Lapen, D., Moulton, K., Quake, S.R., Steinmann, K., Thayer, E., Tyurina, A., Ward, R., Weiss, H. and Xie, Z. (2008) Single-molecule DNA sequencing of a viral genome. *Science* **320**, 106-109.

Hickman-Brenner, F.W., Huntley-Carter, G.P., Saitoh, Y., Steigerwalt, A.G., Farmer, J.J., 3rd and Brenner, D.J. (1984) *Moellerella wisconsensis*, a new genus and species of Enterobacteriaceae found in human stool specimens. *Journal of clinical microbiology* **19**, 460-463.

Hightower, L.E. (1991) Heat shock, stress proteins, chaperones, and proteotoxicity. *Cell* **66**, 191-197.

Holscher, H.D., Czerkies, L.A., Cekola, P., Litov, R., Benbow, M., Santema, S., Alexander, D.D., Perez, V., Sun, S., Saavedra, J.M. and Tappenden, K.A. (2012) *Bifidobacterium lactis* Bb12 enhances intestinal antibody response in formula-fed infants: a randomized, double-blind, controlled trial. *JPEN Journal of parenteral and enteral nutrition* **36**, 106S-117S.

Holzappel, W.H., Haberer, P., Geisen, R., Bjorkroth, J. and Schillinger, U. (2001) Taxonomy and important features of probiotic microorganisms in food and nutrition. *The American journal of clinical nutrition* **73**, 365S-373S.

Holzappel, W. (1997) Use of starter cultures in fermentation on a household scale. *Food Control* **8**, 241-258.

Références bibliographique

- Hartke, A., Bouche, S., Gansel, X., Boutibonnes, P. and Auffray, Y. (1994)** Starvation-Induced Stress Resistance in *Lactococcus lactis* subsp. *lactis* IL1403. *Appl Environ Microbiol* **60**, 3474-3478.
- Hassaine, O., Zadi-Karam, H. and Karam, N.E. (2007)** Technologically important properties of lactic acid bacteria isolated from raw milk of three breeds of Algerian dromedary (*Camelus dromedarius*). *Afr J Biotechnol* **6**, 1720-1727.
- Haandrikman, A.J., Kok, J., Laan, H., Soemitro, S., Ledeboer, A.M., Konings, W.N. and Venema, G. (1989)** Identification of a gene required for maturation of an extracellular lactococcal serine proteinase. *J Bacteriol* **171**, 2789-2794.
- Habimana, O., Le Goff, C., Juillard, V., Bellon-Fontaine, M.N., Buist, G., Kulakauskas, S. and Briandet, R. (2007)** Positive role of cell wall anchored proteinase PrtP in adhesion of lactococci. *Bmc Microbiol* **7**, 36.
- Hanniffy, S.B., Carter, A.T., Hitchin, E. and Wells, J.M. (2007)** Mucosal delivery of a pneumococcal vaccine using *Lactococcus lactis* affords protection against respiratory infection. *The Journal of infectious diseases* **195**, 185-193.
- Harvey, R.J. and Collins, E.B. (1962)** Citrate transport system of *Streptococcus diacetilactis*. *J Bacteriol* **83**, 1005-1009.
- Hendrickx, A.P., Bonten, M.J., van Luit-Asbroek, M., Schapendonk, C.M., Kragten, A.H. and Willems, R.J. (2008)** Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate. *Microbiology* **154**, 3212-3223.
- Hoefnagel, M.H., Starrenburg, M.J., Martens, D.E., Hugenholtz, J., Kleerebezem, M., Van, S., II, Bongers, R., Westerhoff, H.V. and Snoep, J.L. (2002)** Metabolic engineering of lactic acid bacteria, the combined approach: kinetic modelling, metabolic control and experimental analysis. *Microbiology* **148**, 1003-1013.
- Hols, P., Kleerebezem, M., Schanck, A.N., Ferain, T., Hugenholtz, J., Delcour, J. and de Vos, W.M. (1999a)** Conversion of *Lactococcus lactis* from homolactic to homoalanine fermentation through metabolic engineering. *Nature biotechnology* **17**, 588-592.
- Haider, B., Ahn, T.H., Bushnell, B., Chai, J., Copeland, A. and Pan, C. (2014)** Omega: an overlap-graph de novo assembler for metagenomics. *Bioinformatics* **30**, 2717-2722.
- Hiatt, J.B., Patwardhan, R.P., Turner, E.H., Lee, C. and Shendure, J. (2010)** Parallel, tag-directed assembly of locally derived short sequence reads. *Nature methods* **7**, 119-122.
- Hopfl, P., Ludwig, W., Schleifer, K.H. and Larsen, N. (1989)** The 23S ribosomal RNA higher-order structure of *Pseudomonas cepacia* and other prokaryotes. *European journal of biochemistry / FEBS* **185**, 355-364.

Références bibliographique

- Huson, D.H., Auch, A.F., Qi, J. and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17**, 377-386.
- Huson, D.H. and Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution* **23**, 254-267.
- Hols, P., Ramos, A., Hugenholtz, J., Delcour, J., de Vos, W.M., Santos, H. and Kleerebezem, M. (1999b) Acetate utilization in *Lactococcus lactis* deficient in lactate dehydrogenase: a rescue pathway for maintaining redox balance. *J Bacteriol* **181**, 5521-5526.
- Hosono, A., Lee, J., Ametani, A., Natsume, M., Hirayama, M., Adachi, T. and Kaminogawa, S. (1997) Characterization of a water-soluble polysaccharide fraction with immunopotentiating activity from *Bifidobacterium adolescentis* M101-4. *Bioscience, biotechnology, and biochemistry* **61**, 312-316.
- Hugenholtz, J., Sybesma, W., Groot, M.N., Wisselink, W., Ladero, V., Burgess, K., van Sinderen, D., Piard, J.C., Eggink, G., Smid, E.J., Savoy, G., Sesma, F., Jansen, T., Hols, P. and Kleerebezem, M. (2002) Metabolic engineering of lactic acid bacteria for the production of nutraceuticals. *Antonie van Leeuwenhoek* **82**, 217-235.
- Hejnowicz, M.S., Golebiewski, M. and Bardowski, J. (2009) Analysis of the complete genome sequence of the lactococcal bacteriophage bIBB29. *Int J Food Microbiol* **131**, 52-61.
- Herman, C., Thevenet, D., D'Ari, R. and Bouloc, P. (1995) Degradation of sigma 32, the heat shock regulator in *Escherichia coli*, is governed by HflB. *Proc Natl Acad Sci U S A* **92**, 3516-3520.
- Homuth, G., Masuda, S., Mogk, A., Kobayashi, Y. and Schumann, W. (1997) The dnaK operon of *Bacillus subtilis* is heptacistronic. *J Bacteriol* **179**, 1153-1164.
- Homuth, G., Mogk, A. and Schumann, W. (1999) Post-transcriptional regulation of the *Bacillus subtilis* dnaK operon. *Mol Microbiol* **32**, 1183-1197.
- Hyrylainen, H.L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M., Sarvas, M., Pragai, Z., Bron, S., van Dijl, J.M. and Kontinen, V.P. (2001) A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Mol Microbiol* **41**, 1159-1172.
- Hoper, D., Volker, U. and Hecker, M. (2005) Comprehensive characterization of the contribution of individual SigB-dependent general stress genes to stress resistance of *Bacillus subtilis*. *J Bacteriol* **187**, 2810-2826.
- Huang, X., Decatur, A., Sorokin, A. and Helmann, J.D. (1997) The *Bacillus subtilis* sigma(X) protein is an extracytoplasmic function sigma factor contributing to survival at high temperature. *J Bacteriol* **179**, 2915-2921.

Références bibliographique

Hughes, D. (2000) Evaluating genome dynamics: the constraints on rearrangements within bacterial genomes. *Genome Biol* **1**, REVIEWS0006.

I

Iliopoulos, I., Tsoka, S., Andrade, M.A., Enright, A.J., Carroll, M., Poulet, P., Promponas, V., Liakopoulos, T., Palaios, G., Pasquier, C., Hamodrakas, S., Tamames, J., Yagnik, A.T., Tramontano, A., Devos, D., Blaschke, C., Valencia, A., Brett, D., Martin, D., Leroy, C., Rigoutsos, I., Sander, C. and Ouzounis, C.A. (2003) Evaluation of annotation strategies using an entire genome sequence. *Bioinformatics* **19**, 717-726.

Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M.R. and Steven, A.C. (2001) Translocation pathway of protein substrates in ClpAP protease. *Proc Natl Acad Sci U S A* **98**, 4328-4333.

Ingmer, H., Vogensen, F.K., Hammer, K. and Kilstrup, M. (1999) Disruption and analysis of the *clpB*, *clpC*, and *clpE* genes in *Lactococcus lactis*: ClpE, a new Clp family in gram-positive bacteria. *J Bacteriol* **181**, 2075-2083.

Itoi, S., Abe, T., Washio, S., Ikuno, E., Kanomata, Y. and Sugita, H. (2008) Isolation of halotolerant *Lactococcus lactis* subsp. *lactis* from intestinal tract of coastal fish. *Int J Food Microbiol* **121**, 116-121.

Igarashi, T., Asaga, E. and Goto, N. (2003) The sortase of *Streptococcus mutans* mediates cell wall anchoring of a surface protein antigen. *Oral microbiology and immunology* **18**, 266-269.

Jolly, L., Vincent, S.J., Duboc, P. and Neeser, J.R. (2002) Exploiting exopolysaccharides from lactic acid bacteria. *Antonie van Leeuwenhoek* **82**, 367-374.

Itoi, S., Yuasa, K., Washio, S., Abe, T., Ikuno, E. and Sugita, H. (2009) Phenotypic variation in *Lactococcus lactis* subsp. *lactis* isolates derived from intestinal tracts of marine and freshwater fish. *J Appl Microbiol* **107**, 867-874.

J

James, P.R., Hardman, S.M. and Patterson, D.L. (2000) Osteomyelitis and possible endocarditis secondary to *Lactococcus garvieae*: a first case report. *Postgraduate medical journal* **76**, 301-303.

Jones, P.G., VanBogelen, R.A. and Neidhardt, F.C. (1987) Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol* **169**, 2092-2095.

Références bibliographique

- Jones, G.E. and Dawson, W.O. (1978)** Stability of mutations conferring temperature sensitivity on tobacco mosaic virus. *Intervirology* **9**, 149-155.
- Junemann, S., Sedlazeck, F.J., Prior, K., Albersmeier, A., John, U., Kalinowski, J., Mellmann, A., Goesmann, A., von Haeseler, A., Stoye, J. and Harmsen, D. (2013)** Updating benchtop sequencing performance comparison. *Nature biotechnology* **31**, 294-296.
- Jain, R., Rivera, M.C. and Lake, J.A. (1999)** Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci U S A* **96**, 3801-3806.
- Jang, A., Szabo, J., Hosni, A.A., Coughlin, M. and Bishop, P.L. (2006)** Measurement of chlorine dioxide penetration in dairy process pipe biofilms during disinfection. *Appl Microbiol Biotechnol* **72**, 368-376.
- Jarvis, A.W., Fitzgerald, G.F., Mata, M., Mercenier, A., Neve, H., Powell, I.B., Ronda, C., Saxelin, M. and Teuber, M. (1991)** Species and type phages of lactococcal bacteriophages. *Intervirology* **32**, 2-9.
- Jarvis, A.W., Fitzgerald, G.F., Mata, M., Mercenier, A., Neve, H., Powell, I.B., Ronda, C., Saxelin, M. and Teuber, M. (1991)** Species and type phages of lactococcal bacteriophages. *Intervirology* **32**, 2-9.

K

- Karp, P.D., Paley, S. and Romero, P. (2002)** The Pathway Tools software. *Bioinformatics* **18 Suppl 1**, S225-232.
- Kasprzyk, A., Keefe, D., Smedley, D., London, D., Spooner, W., Melsopp, C., Hammond, M., Rocca-Serra, P., Cox, T. and Birney, E. (2004)** EnsMart: a generic system for fast and flexible access to biological data. *Genome Res* **14**, 160-169.
- Kato, H., Shiwa, Y., Oshima, K., Machii, M., Araya-Kojima, T., Zendo, T., Shimizu-Kadota, M., Hattori, M., Sonomoto, K. and Yoshikawa, H. (2012)** Complete genome sequence of *Lactococcus lactis* IO-1, a lactic acid bacterium that utilizes xylose and produces high levels of L-lactic acid. *J Bacteriol* **194**, 2102-2103.
- Kazmierczak, M.J., Mithoe, S.C., Boor, K.J. and Wiedmann, M. (2003)** *Listeria monocytogenes* sigma B regulates stress response and virulence functions. *J Bacteriol* **185**, 5722-5734.
- Kelly, W. and Ward, L. (2002)** Genotypic vs. phenotypic biodiversity in *Lactococcus lactis*. *Microbiology* **148**, 3332-3333.

Références bibliographique

- Kelly, W.J., Altermann, E., Lambie, S.C. and Leahy, S.C. (2013)** Interaction between the genomes of *Lactococcus lactis* and phages of the P335 species. *Frontiers in microbiology* **4**, 257.
- Kelley, D.R., Schatz, M.C. and Salzberg, S.L. (2010)** Quake: quality-aware detection and correction of sequencing errors. *Genome Biol* **11**, R116.
- Khedid, K., Faid, M., Mokhtari, A., Soulaymani, A. and Zinedine, A. (2009)** Characterization of lactic acid bacteria isolated from the one humped camel milk produced in Morocco. *Microbiological research* **164**, 81-91.
- Konuspayeva, G., Faye, B. and Loiseau, G. (2009)** The composition of camel milk: A meta-analysis of the literature data. *J Food Compos Anal* **22**, 95-101.
- Konuspayeva, G., Faye, B., Loiseau, G. and Levieux, D. (2007)** Lactoferrin and immunoglobulin contents in camel's milk (*Camelus bactrianus*, *Camelus dromedarius*, and hybrids) from Kazakhstan. *Journal of Dairy Science* **90**, 38-46.
- Konuspayeva, G., Lemarie, E., Faye, B., Loiseau, G. and Montet, D. (2008)** Fatty acid and cholesterol composition of camel's (*Camelus bactrianus*, *Camelus dromedarius* and hybrids) milk in Kazakhstan. *Dairy Sci Technol* **88**, 327-340.
- Koonin, E.V. and Galperin, M.Y. (1997)** Prokaryotic genomes: the emerging paradigm of genome-based microbiology. *Current opinion in genetics & development* **7**, 757-763.
- Keiler, K.C., Waller, P.R. and Sauer, R.T. (1996)** Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990-993.
- Keseler, I.M., Collado-Vides, J., Gama-Castro, S., Ingraham, J., Paley, S., Paulsen, I.T.,
- Peralta-Gil, M. and Karp, P.D. (2005)** EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Res* **33**, D334-337.
- Kieser, H.M., Kieser, T. and Hopwood, D.A. (1992)** A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. *J Bacteriol* **174**, 5496-5507.
- Kilstrup, M., Jacobsen, S., Hammer, K. and Vogensen, F.K. (1997)** Induction of heat shock proteins DnaK, GroEL, and GroES by salt stress in *Lactococcus lactis*. *Appl Environ Microbiol* **63**, 1826-1837.
- Kim, J., Gordon, L., Dehal, P., Badri, H., Christensen, M., Groza, M., Ha, C., Hammond, S., Vargas, M., Wehri, E., Wagner, M., Olsen, A. and Stubbs, L. (2001a)** Homology-driven assembly of a sequence-ready mouse BAC contig map spanning regions related to the 46-Mb gene-rich euchromatic segments of human chromosome 19. *Genomics* **74**, 129-141.

Références bibliographique

- Kim, Y.I., Levchenko, I., Fraczkowska, K., Woodruff, R.V., Sauer, R.T. and Baker, T.A. (2001b)** Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat Struct Biol* **8**, 230-233.
- Kim, S.G. and Batt, C.A. (1993)** Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis* groESL operon. *Gene* **127**, 121-126.
- Kirstein, J., Zuhlke, D., Gerth, U., Turgay, K. and Hecker, M. (2005)** A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *EMBO J* **24**, 3435-3445.
- Kotsonis, S.E., Powell, I.B., Pillidge, C.J., Limsowtin, G.K., Hillier, A.J. and Davidson, B.E. (2008)** Characterization and genomic analysis of phage ascphi28, a phage of the family Podoviridae infecting *Lactococcus lactis*. *Appl Environ Microbiol* **74**, 3453-3460.
- Krajewski, S.S. and Narberhaus, F. (2014)** Temperature-driven differential gene expression by RNA thermosensors. *Biochim Biophys Acta* **1839**, 978-988.
- Kruger, E. and Hecker, M. (1998)** The first gene of the *Bacillus subtilis* clpC operon, ctsR, encodes a negative regulator of its own operon and other class III heat shock genes. *J Bacteriol* **180**, 6681-6688.
- Kirk, N. and Piper, P.W. (1991)** The determinants of heat-shock element-directed lacZ expression in *Saccharomyces cerevisiae*. *Yeast* **7**, 539-546.
- Klaenhammer, T.R., Barrangou, R., Buck, B.L., Azcarate-Peril, M.A. and Altermann, E. (2005)** Genomic features of lactic acid bacteria effecting bioprocessing and health. *FEMS Microbiol Rev* **29**, 393-409.
- Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O.P., Leer, R., Tarchini, R., Peters, S.A., Sandbrink, H.M., Fiers, M.W., Stiekema, W., Lankhorst, R.M., Bron, P.A., Hoffer, S.M., Groot, M.N., Kerkhoven, R., de Vries, M., Ursing, B., de Vos, W.M. and Siezen, R.J. (2003)** Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci U S A* **100**, 1990-1995.
- Kleerebezemab, M., Hols, P. and Hugenholtz, J. (2000)** Lactic acid bacteria as a cell factory: rerouting of carbon metabolism in *Lactococcus lactis* by metabolic engineering. *Enzyme and microbial technology* **26**, 840-848.
- Kleerebezem, M., Boels, I.C., Groot, M.N., Mierau, I., Sybesma, W. and Hugenholtz, J. (2002)** Metabolic engineering of *Lactococcus lactis*: the impact of genomics and metabolic modelling. *Journal of biotechnology* **98**, 199-213.
- Kleerebezem, M., van Kranenburg, R., Tuinier, R., Boels, I.C., Zoon, P., Looijesteijn, E., Hugenholtz, J. and de Vos, W.M. (1999)** Exopolysaccharides produced by *Lactococcus*

Références bibliographique

lactis: from genetic engineering to improved rheological properties? *Antonie van Leeuwenhoek* **76**, 357-365.

Kang, H.J., Coulibaly, F., Clow, F., Proft, T. and Baker, E.N. (2007) Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* **318**, 1625-1628.

Kankainen, M., Paulin, L., Tynkkynen, S., von Ossowski, I., Reunanen, J., Partanen, P., Satokari, R., Vesterlund, S., Hendrickx, A.P., Lebeer, S., De Keersmaecker, S.C., Vanderleyden, J., Hamalainen, 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., Salusjarvi, T., Auvinen, P. and 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.

Keller, L. and Surette, M.G. (2006) Communication in bacteria: an ecological and evolutionary perspective. *Nat Rev Microbiol* **4**, 249-258.

Kempler, G.M. and McKay, L.L. (1980) Improved Medium for Detection of Citrate-Fermenting *Streptococcus lactis* subsp. *diacetylactis*. *Appl Environ Microbiol* **39**, 926-927.

Kaletunc, G., Lee, J., Alpas, H. and Bozoglu, F. (2004) Evaluation of structural changes induced by high hydrostatic pressure in *Leuconostoc mesenteroides*. *Appl Environ Microbiol* **70**, 1116-1122.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-1645.

Kiwaki, M., Ikemura, H., Shimizu-Kadota, M. and Hirashima, A. (1989) Molecular characterization of a cell wall-associated proteinase gene from *Streptococcus lactis* NCD0763. *Mol Microbiol* **3**, 359-369.

Kok, J., Hill, D., Haandrikman, A.J., de Reuver, M.J., Laan, H. and Venema, G. (1988a) Deletion analysis of the proteinase gene of *Streptococcus cremoris* Wg2. *Appl Environ Microbiol* **54**, 239-244.

Kok, J., Leenhouts, K.J., Haandrikman, A.J., Ledebøer, A.M. and Venema, G. (1988b) 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., Dumenil, G., Caliot, E., Trieu-Cuot, P. and Dramsi, S. (2009) Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS pathogens* **5**, e1000422.

Références bibliographique

- Kozar, W., Rajchert-Trzpil, M. and Dobrzanski, W.T. (1974) The effect of proflavin, ethidium bromide and an elevated temperature on the appearance of nisin-negative clones in nisin-producing strains of *Streptococcus lactis*. *J Gen Microbiol* **83**, 295-302.
- Krause, L., Diaz, N.N., Goesmann, A., Kelley, S., Nattkemper, T.W., Rohwer, F., Edwards, R.A. and Stoye, J. (2008) Phylogenetic classification of short environmental DNA fragments. *Nucleic Acids Res* **36**, 2230-2239.
- Korber, P., Stahl, J.M., Nierhaus, K.H. and Bardwell, J.C. (2000) Hsp15: a ribosome-associated heat shock protein. *EMBO J* **19**, 741-748.
- Koski, L.B., Morton, R.A. and Golding, G.B. (2001) Codon bias and base composition are poor indicators of horizontally transferred genes. *Molecular biology and evolution* **18**, 404-412.
- Kovacs, T., Hargitai, A., Kovacs, K.L. and Mecs, I. (1998) pH-dependent activation of the alternative transcriptional factor sigmaB in *Bacillus subtilis*. *FEMS Microbiol Lett* **165**, 323-328.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Cordani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denziot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Togoni, A., Tosato, V., Uchiyama, S., Vandebol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt,

Références bibliographique

R., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and Danchin, A. (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249-256.

L

Ladero, V., Del Rio, B., Linares, D.M., Fernandez, M., Mayo, B., Martin, M.C. and Alvarez, M.A. (2014) Genome Sequence Analysis of the Biogenic Amine-Producing Strain *Lactococcus lactis* subsp. *cremoris* CECT 8666 (Formerly GE2-14). *Genome announcements* **2**.

Ladero, V., Del Rio, B., Linares, D.M., Fernandez, M., Mayo, B., Martin, M.C. and Alvarez, M.A. (2015) Draft Genome Sequence of the Putrescine-Producing Strain *Lactococcus lactis* subsp. *lactis* 1AA59. *Genome announcements* **3**.

Lambie, S.C., Altermann, E., Leahy, S.C. and Kelly, W.J. (2014) Draft Genome Sequence of *Lactococcus lactis* subsp. *cremoris* HPT, the First Defined-Strain Dairy Starter Culture Bacterium. *Genome announcements* **2**.

Lancefield, R.C. and Perlmann, G.E. (1952) Preparation and properties of type-specific M antigen isolated from a group A, type 1 hemolytic streptococcus. *The Journal of experimental medicine* **96**, 71-82.

Lander, E.S. (2011) Initial impact of the sequencing of the human genome. *Nature* **470**, 187-197.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D.,

Références bibliographique

- Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglu, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J. and International Human Genome Sequencing, C. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.
- Laserson, J., Jojic, V. and Koller, D. (2011) Genovo: de novo assembly for metagenomes. *Journal of computational biology : a journal of computational molecular cell biology* **18**, 429-443.
- Lalioui, L., Pellegrini, E., Dramsi, S., Baptista, M., Bourgeois, N., Doucet-Populaire, F., Rusniok, C., Zouine, M., Glaser, P., Kunst, F., Poyart, C. and Trieu-Cuot, P. (2005) The

Références bibliographique

- SrtA Sortase of *Streptococcus agalactiae* is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. *Infection and immunity* **73**, 3342-3350.
- Lam, M.M., Seemann, T., Tobias, N.J., Chen, H., Haring, V., Moore, R.J., Ballard, S., Grayson, L.M., Johnson, P.D., Howden, B.P. and Stinear, T.P. (2013)** Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. *BMC genomics* **14**, 595.
- Lauer, P., Rinaudo, C.D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A.R., Mora, M., Rappuoli, R., Grandi, G. and Telford, J.L. (2005)** Genome analysis reveals pili in Group B *Streptococcus*. *Science* **309**, 105.
- LeMieux, J., Woody, S. and Camilli, A. (2008)** Roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J Bacteriol* **190**, 6002-6013.
- Letunic, I. and Bork, P. (2007)** Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127-128.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009)** The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079.
- Lamiet, A.A., Ziegelhoffer, T., Georgopoulos, C. and Pluckthun, A. (1990)** The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the beta-lactamase precursor. *EMBO J* **9**, 2315-2319.
- Liang, W.C., Lin, M.G., Chi, M.C., Hu, H.Y., Lo, H.F., Chang, H.P. and Lin, L.L. (2009)** Deletion analysis of the C-terminal region of a molecular chaperone DnaK from *Bacillus licheniformis*. *Archives of microbiology* **191**, 583-593.
- Lim, E.M., Ehrlich, S.D. and Maguin, E. (2000)** Identification of stress-inducible proteins in *Lactobacillus delbrueckii* subsp. *bulgaricus*. *Electrophoresis* **21**, 2557-2561.
- Lipinska, B., Zylicz, M. and Georgopoulos, C. (1990)** The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. *J Bacteriol* **172**, 1791-1797.
- Lo, W.T. and Wang, C.C. (2009)** Community-associated methicillin-resistant *Staphylococcus aureus* disease in children in Taiwan: the importance of antibiotic stewardship. *Journal of the Formosan Medical Association = Taiwan yi zhi* **108**, 751-753.
- Luan, J., Zhuang, Z., Liu, Y., Yun, K., Chen, M. and Wang, P.G. (2010)** Expression of EspA in *Lactococcus lactis* NZ9000 and the detection of its immune effect in vivo and vitro. *Immunopharmacology and immunotoxicology* **32**, 133-140.

Références bibliographique

- Ludwig, W., Seewaldt, E., Kilpper-Balz, R., Schleifer, K.H., Magrum, L., Woese, C.R., Fox, G.E. and Stackebrandt, E. (1985) The phylogenetic position of Streptococcus and Enterococcus. *J Gen Microbiol* **131**, 543-551.
- Luo, C., Tsementzi, D., Kyrpides, N.C. and Konstantinidis, K.T. (2012a) Individual genome assembly from complex community short-read metagenomic datasets. *Isme J* **6**, 898-901.
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D.W., Yiu, S.M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.W. and Wang, J. (2012b) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 18.
- Levander, F. and Radstrom, P. (2001) Requirement for phosphoglucomutase in exopolysaccharide biosynthesis in glucose- and lactose-utilizing Streptococcus thermophilus. *Appl Environ Microbiol* **67**, 2734-2738.
- Lindgren, S.E. and Dobrogosz, W.J. (1990) Antagonistic activities of lactic acid bacteria in food and feed fermentations. *FEMS Microbiol Rev* **7**, 149-163.
- Linke, C., Young, P.G., Kang, H.J., Bunker, R.D., Middleditch, M.J., Caradoc-Davies, T.T., Proft, T. and Baker, E.N. (2010) Crystal structure of the minor pilin FctB reveals determinants of Group A streptococcal pilus anchoring. *J Biol Chem* **285**, 20381-20389.
- Loessner, M.J. (2005) Bacteriophage endolysins--current state of research and applications. *Current opinion in microbiology* **8**, 480-487.
- Looijesteijn, P.J., Boels, I.C., Kleerebezem, M. and Hugenholtz, J. (1999) Regulation of exopolysaccharide production by Lactococcus lactis subsp. cremoris By the sugar source. *Appl Environ Microbiol* **65**, 5003-5008.
- Luesink, E.J., van Herpen, R.E., Grossiord, B.P., Kuipers, O.P. and de Vos, W.M. (1998) Transcriptional activation of the glycolytic las operon and catabolite repression of the gal operon in Lactococcus lactis are mediated by the catabolite control protein CcpA. *Mol Microbiol* **30**, 789-798.
- Latorreguzman, B.A., Kado, C.I. and Kunkee, R.E. (1977) Lactobacillus-Hordniae, a New Species from Leafhopper (Hordnia-Circellata). *Int J Syst Bacteriol* **27**, 362-370.
- Lawrence, J.G. and Hartl, D.L. (1992) Inference of horizontal genetic transfer from molecular data: an approach using the bootstrap. *Genetics* **131**, 753-760.
- Lawrence, J.G. and Ochman, H. (2002) Reconciling the many faces of lateral gene transfer. *Trends in microbiology* **10**, 1-4.

Références bibliographique

- Le Bourgeois, P., Lautier, M., van den Berghe, L., Gasson, M.J. and Ritzenthaler, P. (1995)** Physical and genetic map of the *Lactococcus lactis* subsp. *cremoris* MG1363 chromosome: comparison with that of *Lactococcus lactis* subsp. *lactis* IL 1403 reveals a large genome inversion. *J Bacteriol* **177**, 2840-2850.
- Labrie, S.J. and Moineau, S. (2007)** Abortive infection mechanisms and prophage sequences significantly influence the genetic makeup of emerging lytic lactococcal phages. *J Bacteriol* **189**, 1482-1487.
- Labrie, S.J., Samson, J.E. and Moineau, S. (2010)** Bacteriophage resistance mechanisms. *Nat Rev Microbiol* **8**, 317-327.
- Lindquist, S. and Craig, E.A. (1988)** The heat-shock proteins. *Annual review of genetics* **22**, 631-677.
- Lemaux, P.G., Herendeen, S.L., Bloch, P.L. and Neidhardt, F.C. (1978)** Transient rates of synthesis of individual polypeptides in *E. coli* following temperature shifts. *Cell* **13**, 427-434.
- Leong, K.H., Chen, Y.S., Pan, S.F., Chen, J.J., Wu, H.C., Chang, Y.C. and Yanagida, F. (2014)** Diversity of lactic acid bacteria associated with fresh coffee cherries in Taiwan. *Curr Microbiol* **68**, 440-447.
- Linares, D.M., Kok, J. and Poolman, B. (2010)** Genome sequences of *Lactococcus lactis* MG1363 (revised) and NZ9000 and comparative physiological studies. *J Bacteriol* **192**, 5806-5812.
- Liou, M.R., Chen, Y.R. and Liou, R.F. (2004)** Complete nucleotide sequence and genome organization of a Cactus virus X strain from *Hylocereus undatus* (Cactaceae). *Archives of virology* **149**, 1037-1043.
- Lipson, D., Raz, T., Kieu, A., Jones, D.R., Giladi, E., Thayer, E., Thompson, J.F., Letovsky, S., Milos, P. and Causey, M. (2009)** Quantification of the yeast transcriptome by single-molecule sequencing. *Nature biotechnology* **27**, 652-658.
- Liu, M., van Enckevort, F.H. and Siezen, R.J. (2005)** Genome update: lactic acid bacteria genome sequencing is booming. *Microbiology* **151**, 3811-3814.
- Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. and Pallen, M.J. (2012)** Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology* **30**, 434-439.
- Lorenz, P. and Eck, J. (2005)** Metagenomics and industrial applications. *Nat Rev Microbiol* **3**, 510-516.

M

- Madhumalar, A. and Bansal, M. (2005)** Sequence preference for BI/BII conformations in DNA: MD and crystal structure data analysis. *Journal of biomolecular structure & dynamics* **23**, 13-27.
- Madera, C., Monjardin, C. and Suarez, J.E. (2004)** Milk contamination and resistance to processing conditions determine the fate of *Lactococcus lactis* bacteriophages in dairies. *Appl Environ Microbiol* **70**, 7365-7371.
- Magoc, T. and Salzberg, S.L. (2011)** FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957-2963.
- Marcais, G. and Kingsford, C. (2011)** A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764-770.
- Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C. and Bryant, S.H. (2015)** CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222-226.
- Mc Grath, S., Fitzgerald, G.F. and van Sinderen, D. (2007)** Bacteriophages in dairy products: pros and cons. *Biotechnology journal* **2**, 450-455.
- Mecas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J. and Gross, C.A. (1993)** The activity of sigma E, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes & development* **7**, 2618-2628.
- Mikkonen, M., Raisanen, L. and Alatossava, T. (1996)** The early gene region completes the nucleotide sequence of *Lactobacillus delbrueckii* subsp. *lactis* phage LL-H. *Gene* **175**, 49-57.
- Mogk, A., Homuth, G., Scholz, C., Kim, L., Schmid, F.X. and Schumann, W. (1997)** The GroE chaperonin machine is a major modulator of the CIRCE heat shock regulon of *Bacillus subtilis*. *EMBO J* **16**, 4579-4590.
- Maccallum, I., Przybylski, D., Gnerre, S., Burton, J., Shlyakhter, I., Gnirke, A., Malek, J., McKernan, K., Ranade, S., Shea, T.P., Williams, L., Young, S., Nusbaum, C. and Jaffe, D.B. (2009)** ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. *Genome Biol* **10**, R103.
- Marx, V. (2013)** Genomics in the clouds. *Nature methods* **10**, 941-945.
- McDonald, L.C., McFeeters, R.F., Daeschel, M.A. and Fleming, H.P. (1987)** A differential medium for the enumeration of homofermentative and heterofermentative lactic Acid bacteria. *Appl Environ Microbiol* **53**, 1382-1384.

Références bibliographique

- McGuffin, L.J., Bryson, K. and Jones, D.T. (2000)** The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405.
- Maione, D., Margarit, I., Rinaudo, C.D., Massignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E.T., Rosini, R., D'Agostino, N., Miorin, L., Buccato, S., Mariani, M., Galli, G., Nogarotto, R., Nardi-Dei, V., Vegni, F., Fraser, C., Mancuso, G., Teti, G., Madoff, L.C., Paoletti, L.C., Rappuoli, R., Kasper, D.L., Telford, J.L. and Grandi, G. (2005)** Identification of a universal Group B streptococcus vaccine by multiple genome screen. *Science* **309**, 148-150.
- Marraffini, L.A., Dedent, A.C. and Schneewind, O. (2006)** Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol Mol Biol Rev* **70**, 192-221.
- Marraffini, L.A. and Schneewind, O. (2006)** Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol Microbiol* **62**, 1402-1417.
- Martin, M.G., Sender, P.D., Peiru, S., de Mendoza, D. and Magni, C. (2004)** Acid-inducible transcription of the operon encoding the citrate lyase complex of *Lactococcus lactis* Biovar diacetylactis CRL264. *J Bacteriol* **186**, 5649-5660.
- Mazmanian, S.K., Liu, G., Ton-That, H. and Schneewind, O. (1999)** *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760-763.
- McKay, L., Miller, A., 3rd, Sandine, W.E. and Elliker, P.R. (1970)** Mechanisms of lactose utilization by lactic acid streptococci: enzymatic and genetic analyses. *J Bacteriol* **102**, 804-809.
- McKay, L.L. and Baldwin, K.A. (1974a)** Altered metabolism in a *Streptococcus lactis* C2 mutant deficient in lactic dehydrogenase. *J Dairy Sci* **57**, 181-186.
- McKay, L.L. and Baldwin, K.A. (1974b)** Simultaneous loss of proteinase- and lactose-utilizing enzyme activities in *Streptococcus lactis* and reversal of loss by transduction. *Appl Microbiol* **28**, 342-346.
- McKay, L.L., Walter, L.A., Sandine, W.E. and Elliker, P.R. (1969)** Involvement of phosphoenolpyruvate in lactose utilization by group N streptococci. *J Bacteriol* **99**, 603-610.
- Mercenier, A., Muller-Alouf, H. and Grangette, C. (2000)** Lactic acid bacteria as live vaccines. *Current issues in molecular biology* **2**, 17-25.
- Mills, D.A., Rawsthorne, H., Parker, C., Tamir, D. and Makarova, K. (2005)** Genomic analysis of *Oenococcus oeni* PSU-1 and its relevance to winemaking. *FEMS Microbiol Rev* **29**, 465-475.

Références bibliographique

- Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G., Maggi, T., Taddei, A.R., Grandi, G. and 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.
- Morona, J.K., Morona, R. and Paton, J.C. (1997a)** Characterization of the locus encoding the Streptococcus pneumoniae type 19F capsular polysaccharide biosynthetic pathway. *Mol Microbiol* **23**, 751-763.
- Morona, J.K., Morona, R. and Paton, J.C. (1997b)** Molecular and genetic characterization of the capsule biosynthesis locus of Streptococcus pneumoniae type 19B. *J Bacteriol* **179**, 4953-4958.
- Morgenstern, B. (1999)** DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. *Bioinformatics* **15**, 211-218.
- Mullis, K.B. and Faloona, F.A. (1987)** Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in enzymology* **155**, 335-350.
- Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D.M., Hawkins, T., Plengvidhya, V., Welker, D., Hughes, J., Goh, Y., Benson, A., Baldwin, K., Lee, J.H., Diaz-Muniz, I., Dosti, B., Smeianov, V., Wechter, W., Barabote, R., Lorca, G., Altermann, E., Barrangou, R., Ganesan, B., Xie, Y., Rawsthorne, H., Tamir, D., Parker, C., Breidt, F., Broadbent, J., Hutkins, R., O'Sullivan, D., Steele, J., Unlu, G., Saier, M., Klaenhammer, T., Richardson, P., Kozyavkin, S., Weimer, B. and Mills, D. (2006)** Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**, 15611-15616.
- Makarova, K.S. and Koonin, E.V. (2007)** Evolutionary genomics of lactic acid bacteria. *J Bacteriol* **189**, 1199-1208.
- Marco, M.L., de Vries, M.C., Wels, M., Molenaar, D., Mangell, P., Ahrne, S., de Vos, W.M., Vaughan, E.E. and Kleerebezem, M. (2010)** Convergence in probiotic Lactobacillus gut-adaptive responses in humans and mice. *Isme J* **4**, 1481-1484.
- Mohamed, S.A., Fahmy, A.S. and Mohamed, T.M. (2005)** Carbohydrases in camel (*Camelus dromedarius*) pancreas. Purification and characterization of glucoamylase. *Comparative biochemistry and physiology Part B, Biochemistry & molecular biology* **140**, 73-80.

Références bibliographique

- Morello, E., Bermudez-Humaran, L.G., Llull, D., Sole, V., Miraglio, N., Langella, P. and Poquet, I. (2008) *Lactococcus lactis*, an efficient cell factory for recombinant protein production and secretion. *Journal of molecular microbiology and biotechnology* **14**, 48-58.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bembien, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380.
- Marty-Teyssset, C., de la Torre, F. and Garel, J. (2000) Increased production of hydrogen peroxide by *Lactobacillus delbrueckii* subsp. *bulgaricus* upon aeration: involvement of an NADH oxidase in oxidative stress. *Appl Environ Microbiol* **66**, 262-267.
- Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc Natl Acad Sci U S A* **74**, 560-564.
- McCulloch, J.A., de Oliveira, V.M., de Almeida Pina, A.V., Perez-Chaparro, P.J., de Almeida, L.M., de Vasconcelos, J.M., de Oliveira, L.F., da Silva, D.E., Rogez, H.L., Cretenet, M., Mamizuka, E.M. and Nunes, M.R. (2014) Complete Genome Sequence of *Lactococcus lactis* Strain AI06, an Endophyte of the Amazonian Acai Palm. *Genome announcements* **2**.
- McLean, M.J., Wolfe, K.H. and Devine, K.M. (1998) Base composition skews, replication orientation, and gene orientation in 12 prokaryote genomes. *Journal of molecular evolution* **47**, 691-696.
- Medigue, C., Rouxel, T., Vigier, P., Henaut, A. and Danchin, A. (1991) Evidence for horizontal gene transfer in *Escherichia coli* speciation. *J Mol Biol* **222**, 851-856.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R.A. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics* **9**, 386.
- Mira, A., Ochman, H. and Moran, N.A. (2001) Deletional bias and the evolution of bacterial genomes. *Trends in genetics : TIG* **17**, 589-596.

Références bibliographique

- Mofredj, A., Bahloul, H. and Chanut, C. (2007)** [Lactococcus lactis: an opportunistic bacterium?]. *Medecine et maladies infectieuses* **37**, 200-207.
- Molenaar, D., Bringel, F., Schuren, F.H., de Vos, W.M., Siezen, R.J. and Kleerebezem, M. (2005)** Exploring Lactobacillus plantarum genome diversity by using microarrays. *J Bacteriol* **187**, 6119-6127.
- Morel-Deville, F., Fauvel, F. and Morel, P. (1998)** Two-component signal-transducing systems involved in stress responses and vancomycin susceptibility in Lactobacillus sakei. *Microbiology* **144** (Pt 10), 2873-2883.
- Moszer, I., Rocha, E.P. and Danchin, A. (1999)** Codon usage and lateral gene transfer in Bacillus subtilis. *Current opinion in microbiology* **2**, 524-528.
- Mrazek, J. and Karlin, S. (1999)** Detecting alien genes in bacterial genomes. *Annals of the New York Academy of Sciences* **870**, 314-329.
- Msadek, T. (1999)** When the going gets tough: survival strategies and environmental signaling networks in Bacillus subtilis. *Trends in microbiology* **7**, 201-207.
- Msadek, T., Dartois, V., Kunst, F., Herbaud, M.L., Denizot, F. and Rapoport, G. (1998)** ClpP of Bacillus subtilis is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. *Mol Microbiol* **27**, 899-914.
- Msadek, T., Kunst, F. and Rapoport, G. (1994)** MecB of Bacillus subtilis, a member of the ClpC ATPase family, is a pleiotropic regulator controlling competence gene expression and growth at high temperature. *Proc Natl Acad Sci U S A* **91**, 5788-5792.

N

- Nakamura, S., Maeda, N., Miron, I.M., Yoh, M., Izutsu, K., Kataoka, C., Honda, T., Yasunaga, T., Nakaya, T., Kawai, J., Hayashizaki, Y., Horii, T. and Iida, T. (2008)** Metagenomic diagnosis of bacterial infections. *Emerging infectious diseases* **14**, 1784-1786.
- Nakajima, H., Hirota, T., Toba, T., Itoh, T. and Adachi, S. (1992)** Structure of the extracellular polysaccharide from slime-forming Lactococcus lactis subsp. cremoris SBT 0495. *Carbohydrate research* **224**, 245-253.
- Nakata, M., Koller, T., Moritz, K., Ribardo, D., Jonas, L., McIver, K.S., Sumitomo, T., Terao, Y., Kawabata, S., Podbielski, A. and Kreikemeyer, B. (2009)** Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in Streptococcus pyogenes serotype M49. *Infection and immunity* **77**, 32-44.

Références bibliographique

- Nallapareddy, S.R., Singh, K.V., Sillanpaa, J., Garsin, D.A., Hook, M., Erlandsen, S.L. and Murray, B.E. (2006)** Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *The Journal of clinical investigation* **116**, 2799-2807.
- Namiki, T., Hachiya, T., Tanaka, H. and Sakakibara, Y. (2012)** MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* **40**, e155.
- Nagarajan, N., Read, T.D. and Pop, M. (2008)** Scaffolding and validation of bacterial genome assemblies using optical restriction maps. *Bioinformatics* **24**, 1229-1235.
- Nair, S., Poyart, C., Beretti, J.L., Veiga-Fernandes, H., Berche, P. and Trieu-Cuot, P. (2003)** Role of the *Streptococcus agalactiae* ClpP serine protease in heat-induced stress defence and growth arrest. *Microbiology* **149**, 407-417.
- Nair, S., Milohanic, E. and Berche, P. (2000)** ClpC ATPase is required for cell adhesion and invasion of *Listeria monocytogenes*. *Infection and immunity* **68**, 7061-7068.
- Nair, S., Frehel, C., Nguyen, L., Escuyer, V. and Berche, P. (1999)** ClpE, a novel member of the HSP100 family, is involved in cell division and virulence of *Listeria monocytogenes*. *Mol Microbiol* **31**, 185-196.
- Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A. and Gautheret, D. (2011)** ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA biology* **8**, 11-13.
- Nahrstedt, H., Schroder, C. and Meinhardt, F. (2005)** Evidence for two *recA* genes mediating DNA repair in *Bacillus megaterium*. *Microbiology* **151**, 775-787.
- Nawrocki, E.P. and Eddy, S.R. (2013)** Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933-2935.
- Neidhardt, F.C. (1987)** What the bacteriologists have learned about heat shock. *Genes & development* **1**, 109-110.
- Neidhardt, F.C., VanBogelen, R.A. and Vaughn, V. (1984)** The genetics and regulation of heat-shock proteins. *Annual review of genetics* **18**, 295-329.
- Neve, H., Geis, A. and Teuber, M. (1984)** Conjugal transfer and characterization of bacteriocin plasmids in group N (lactic acid) streptococci. *J Bacteriol* **157**, 833-838.
- Neves, A.R., Pool, W.A., Kok, J., Kuipers, O.P. and 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.

Références bibliographique

- Neves, A.R., Ramos, A., Nunes, M.C., Kleerebezem, M., Hugenholtz, J., de Vos, W.M., Almeida, J. and Santos, H. (1999) In vivo nuclear magnetic resonance studies of glycolytic kinetics in *Lactococcus lactis*. *Biotechnology and bioengineering* **64**, 200-212.
- Nilsson, D., Lauridsen, A.A., Tomoyasu, T. and Ogura, T. (1994) A *Lactococcus lactis* gene encodes a membrane protein with putative ATPase activity that is homologous to the essential *Escherichia coli* ftsH gene product. *Microbiology* **140** (Pt 10), 2601-2610.
- Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., Bamshad, M., Nickerson, D.A. and Shendure, J. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272-276.
- Navarre, W.W. and 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.
- Nobbs, A.H., Vajna, R.M., Johnson, J.R., Zhang, Y., Erlandsen, S.L., Oli, M.W., Kreth, J., Brady, L.J. and Herzberg, M.C. (2007) Consequences of a sortase A mutation in *Streptococcus gordonii*. *Microbiology* **153**, 4088-4097.
- Nover, L. and Hightower, L. (1991) Heat shock and development. Introduction. *Results and problems in cell differentiation* **17**, 1-4.
- Nyysola, A., Pihlajaniemi, A., Palva, A., von Weymarn, N. and Leisola, M. (2005) Production of xylitol from D-xylose by recombinant *Lactococcus lactis*. *Journal of biotechnology* **118**, 55-66.

O

- O'Connell-Motherway, M., van Sinderen, D., Morel-Deville, F., Fitzgerald, G.F., Ehrlich, S.D. and Morel, P. (2000) Six putative two-component regulatory systems isolated from *Lactococcus lactis* subsp. *cremoris* MG1363. *Microbiology* **146** (Pt 4), 935-947.
- Ochman, H. (2002) Bacterial evolution: chromosome arithmetic and geometry. *Current biology : CB* **12**, R427-428.
- Oliveira, L.C., Saraiva, T.D., Soares, S.C., Ramos, R.T., Sa, P.H., Carneiro, A.R., Miranda, F., Freire, M., Renan, W., Junior, A.F., Santos, A.R., Pinto, A.C., Souza, B.M., Castro, C.P., Diniz, C.A., Rocha, C.S., Mariano, D.C., de Aguiar, E.L., Folador, E.L., Barbosa, E.G., Aburjaile, F.F., Goncalves, L.A., Guimaraes, L.C., Azevedo, M., Agresti, P.C., Silva, R.F., Tiwari, S., Almeida, S.S., Hassan, S.S., Pereira, V.B., Abreu, V.A., Pereira, U.P., Dorella, F.A., Carvalho, A.F., Pereira, F.L., Leal, C.A., Figueiredo, H.C.,

Références bibliographique

- Silva, A., Miyoshi, A. and Azevedo, V. (2014)** Genome Sequence of *Lactococcus lactis* subsp. *lactis* NCDO 2118, a GABA-Producing Strain. *Genome announcements* **2**.
- Oppenheimer-Shaanan, Y., Wexselblatt, E., Katzhendler, J., Yavin, E. and Ben-Yehuda, S. (2011)** c-di-AMP reports DNA integrity during sporulation in *Bacillus subtilis*. *EMBO reports* **12**, 594-601.
- Orla-Jensen, S. (1921)** The Main Lines of the Natural Bacterial System. *J Bacteriol* **6**, 263-273.
- Overbeek, R., Bartels, D., Vonstein, V. and Meyer, F. (2007)** Annotation of bacterial and archaeal genomes: improving accuracy and consistency. *Chemical reviews* **107**, 3431-3447.
- Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F. and Stevens, R. (2014)** The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* **42**, D206-214.
- Oskouian, B. and Stewart, G.C. (1990)** Repression and catabolite repression of the lactose operon of *Staphylococcus aureus*. *J Bacteriol* **172**, 3804-3812.
- Ozsolak, F. and Milos, P.M. (2011)** RNA sequencing: advances, challenges and opportunities. *Nature reviews Genetics* **12**, 87-98.
- O'Callaghan, J., Butto, L.F., MacSharry, J., Nally, K. and O'Toole, P.W. (2012)** Influence of adhesion and bacteriocin production by *Lactobacillus salivarius* on the intestinal epithelial cell transcriptional response. *Appl Environ Microbiol* **78**, 5196-5203.
- O'Driscoll, J., Glynn, F., Fitzgerald, G.F. and van Sinderen, D. (2006)** Sequence analysis of the lactococcal plasmid pNP40: a mobile replicon for coping with environmental hazards. *J Bacteriol* **188**, 6629-6639.
- Orberg, P.K. and Sandine, W.E. (1984)** Common occurrence of plasmid DNA and vancomycin resistance in *Leuconostoc* spp. *Appl Environ Microbiol* **48**, 1129-1133.
- Osaki, M., Takamatsu, D., Shimoji, Y. and Sekizaki, T. (2002)** Characterization of *Streptococcus suis* genes encoding proteins homologous to sortase of gram-positive bacteria. *J Bacteriol* **184**, 971-982.
- Oskouian, B. and Stewart, G.C. (1990)** Repression and catabolite repression of the lactose operon of *Staphylococcus aureus*. *J Bacteriol* **172**, 3804-3812.
- Ozsolak, F., Platt, A.R., Jones, D.R., Reifenberger, J.G., Sass, L.E., McInerney, P., Thompson, J.F., Bowers, J., Jarosz, M. and Milos, P.M. (2009)** Direct RNA sequencing. *Nature* **461**, 814-818.

P

- Parsell, D.A., Kowal, A.S. and Lindquist, S. (1994a)** Saccharomyces cerevisiae Hsp104 protein. Purification and characterization of ATP-induced structural changes. *J Biol Chem* **269**, 4480-4487.
- Parsell, D.A., Kowal, A.S., Singer, M.A. and Lindquist, S. (1994b)** Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* **372**, 475-478.
- Passerini, D., Coddeville, M., Le Bourgeois, P., Loubiere, P., Ritzenthaler, P., Fontagne-Faucher, C., Daveran-Mingot, M.L. and Cocaïgn-Bousquet, M. (2013)** The carbohydrate metabolism signature of lactococcus lactis strain A12 reveals its sourdough ecosystem origin. *Appl Environ Microbiol* **79**, 5844-5852.
- Peant, B., LaPointe, G., Gilbert, C., Atlan, D., Ward, P. and Roy, D. (2005)** Comparative analysis of the exopolysaccharide biosynthesis gene clusters from four strains of Lactobacillus rhamnosus. *Microbiology* **151**, 1839-1851.
- Paton, J.C., Morona, J.K. and Morona, R. (1997)** Characterization of the capsular polysaccharide biosynthesis locus of Streptococcus pneumoniae type 19F. *Microbial drug resistance* **3**, 89-99.
- Peng, Y., Leung, H.C., Yiu, S.M. and Chin, F.Y. (2011)** Meta-IDBA: a de Novo assembler for metagenomic data. *Bioinformatics* **27**, i94-101.
- Peng, Y., Leung, H.C., Yiu, S.M. and Chin, F.Y. (2012)** IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420-1428.
- Petri, R. and Schmidt-Dannert, C. (2004)** Dealing with complexity: evolutionary engineering and genome shuffling. *Current opinion in biotechnology* **15**, 298-304.
- Pevzner, P.A., Tang, H. and Waterman, M.S. (2001)** An Eulerian path approach to DNA fragment assembly. *Proc Natl Acad Sci U S A* **98**, 9748-9753.
- Pevzner, P. and Shamir, R. (2009)** Computing has changed biology--biology education must catch up. *Science* **325**, 541-542.
- Pevzner, P.A. (2004)** Educating biologists in the 21st century: bioinformatics scientists versus bioinformatics technicians. *Bioinformatics* **20**, 2159-2161.
- Pevzner, P.A., Kim, S. and Ng, J. (2008)** Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry". *Science* **321**, 1040; author reply 1040.

Références bibliographique

- Pham, S.K., Antipov, D., Sirotkin, A., Tesler, G., Pevzner, P.A. and Alekseyev, M.A. (2013)** Pathset graphs: a novel approach for comprehensive utilization of paired reads in genome assembly. *Journal of computational biology : a journal of computational molecular cell biology* **20**, 359-371.
- Pop, M. (2009)** Genome assembly reborn: recent computational challenges. *Briefings in bioinformatics* **10**, 354-366.
- Petersen, B.L., Dave, R.I., McMahon, D.J., Oberg, C.J. and Broadbent, J.R. (2000)** Influence of capsular and ropy exopolysaccharide-producing *Streptococcus thermophilus* on Mozzarella cheese and cheese whey. *J Dairy Sci* **83**, 1952-1956.
- Proft, T. and Baker, E.N. (2009)** Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cellular and molecular life sciences : CMLS* **66**, 613-635.
- Pedersen, A.G., Jensen, L.J., Brunak, S., Staerfeldt, H.H. and Ussery, D.W. (2000)** A DNA structural atlas for *Escherichia coli*. *J Mol Biol* **299**, 907-930.
- Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J.D., Volker, U. and Hecker, M. (2001)** Global analysis of the general stress response of *Bacillus subtilis*. *J Bacteriol* **183**, 5617-5631.
- Pfeiler, E.A. and Klaenhammer, T.R. (2007)** The genomics of lactic acid bacteria. *Trends in microbiology* **15**, 546-553.
- Plumed-Ferrer, C., Gazzola, S., Fontana, C., Bassi, D., Cocconcelli, P.S. and von Wright, A. (2015)** Genome Sequence of *Lactococcus lactis* subsp. *cremoris* Mast36, a Strain Isolated from Bovine Mastitis. *Genome announcements* **3**.
- Pochart, P., Marteau, P., Bouhnik, Y., Goderel, I., Bourlioux, P. and Rambaud, J.C. (1992)** Survival of bifidobacteria ingested via fermented milk during their passage through the human small intestine: an in vivo study using intestinal perfusion. *The American journal of clinical nutrition* **55**, 78-80.
- Pontes, D.S., de Azeyedo, M.S.P., Chatel, J.M., Langella, P., Azeyedo, V. and Miyoshi, A. (2011)** *Lactococcus lactis* as a live vector: Heterologous protein production and DNA delivery systems. *Protein Express Purif* **79**, 165-175.
- Papagianni, M. (2012)** Metabolic engineering of lactic acid bacteria for the production of industrially important compounds. *Computational and structural biotechnology journal* **3**, e201210003.
- Poquet, I., Saint, V., Sez nec, E., Simoes, N., Bolotin, A. and Gruss, A. (2000)** HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing. *Mol Microbiol* **35**, 1042-1051.

Références bibliographique

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

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

Ptitsyn, A., Temanni, R., Bouchard, C. and Anderson, P.A. (2015) Semantic Assembly and Annotation of Draft RNAseq Transcripts without a Reference Genome. *PloS one* **10**, e0138006.

Pushkarev, D., Neff, N.F. and Quake, S.R. (2009) Single-molecule sequencing of an individual human genome. *Nature biotechnology* **27**, 847-850.

Q

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P. and Gu, Y. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics* **13**, 341.

Quigley, B.R., Hatkoff, M., Thanassi, D.G., Ouattara, M., Eichenbaum, Z. and Scott, J.R. (2010) A foreign protein incorporated on the Tip of T3 pili in *Lactococcus lactis* elicits systemic and mucosal immunity. *Infection and immunity* **78**, 1294-1303.

R

Rademaker, J.L., Herbet, H., Starrenburg, M.J., Naser, S.M., Gevers, D., Kelly, W.J., Hugenholtz, J., Swings, J. and van Hylckama Vlieg, J.E. (2007) Diversity analysis of dairy and nondairy *Lactococcus lactis* isolates, using a novel multilocus sequence analysis scheme and (GTG)₅-PCR fingerprinting. *Appl Environ Microbiol* **73**, 7128-7137.

Ragan, M.A. (2001) On surrogate methods for detecting lateral gene transfer. *FEMS Microbiol Lett* **201**, 187-191.

Raina, S., Missiakas, D. and Georgopoulos, C. (1995) The *rpoE* gene encoding the sigma E (sigma 24) heat shock sigma factor of *Escherichia coli*. *EMBO J* **14**, 1043-1055.

Références bibliographique

- Raina, S. and Georgopoulos, C. (1990)** A new *Escherichia coli* heat shock gene, *htrC*, whose product is essential for viability only at high temperatures. *J Bacteriol* **172**, 3417-3426.
- Reischl, S., Wiegert, T. and Schumann, W. (2002)** Isolation and analysis of mutant alleles of the *Bacillus subtilis* HrcA repressor with reduced dependency on GroE function. *J Biol Chem* **277**, 32659-32667.
- Roces, C., Campelo, A.B., Veiga, P., Pinto, J.P., Rodriguez, A. and Martinez, B. (2009)** Contribution of the CesR-regulated genes *llmg0169* and *llmg2164-2163* to *Lactococcus lactis* fitness. *Int J Food Microbiol* **133**, 279-285.
- Rousseau, G.M. and Moineau, S. (2009)** Evolution of *Lactococcus lactis* phages within a cheese factory. *Appl Environ Microbiol* **75**, 5336-5344.
- Rouviere, P.E., De Las Penas, A., Mecsas, J., Lu, C.Z., Rudd, K.E. and Gross, C.A. (1995)** *rpoE*, the gene encoding the second heat-shock sigma factor, sigma E, in *Escherichia coli*. *EMBO J* **14**, 1032-1042.
- Rao, F., Ji, Q., Soehano, I. and Liang, Z.X. (2011)** Unusual heme-binding PAS domain from YybT family proteins. *J Bacteriol* **193**, 1543-1551.
- Rao, F., See, R.Y., Zhang, D., Toh, D.C., Ji, Q. and Liang, Z.X. (2010)** YybT is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. *J Biol Chem* **285**, 473-482.
- Rallu, F., Gruss, A., Ehrlich, S.D. and Maguin, E. (2000)** Acid- and multistress-resistant mutants of *Lactococcus lactis* : identification of intracellular stress signals. *Mol Microbiol* **35**, 517-528.
- Reed, J.L., Famili, I., Thiele, I. and Palsson, B.O. (2006a)** Towards multidimensional genome annotation. *Nature reviews Genetics* **7**, 130-141.
- Reed, J.L., Patel, T.R., Chen, K.H., Joyce, A.R., Applebee, M.K., Herring, C.D., Bui, O.T., Knight, E.M., Fong, S.S. and Palsson, B.O. (2006b)** Systems approach to refining genome annotation. *Proc Natl Acad Sci U S A* **103**, 17480-17484.
- Reid, B.G., Fenton, W.A., Horwich, A.L. and Weber-Ban, E.U. (2001)** ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc Natl Acad Sci U S A* **98**, 3768-3772.
- Reid, G., Anand, S., Bingham, M.O., Mbugua, G., Wadstrom, T., Fuller, R., Anukam, K. and Katsivo, M. (2005)** Probiotics for the developing world. *Journal of clinical gastroenterology* **39**, 485-488.
- Reid, G., Jass, J., Sebulsky, M.T. and McCormick, J.K. (2003)** Potential uses of probiotics in clinical practice. *Clinical microbiology reviews* **16**, 658-672.

Références bibliographique

- Richmond, C.S., Glasner, J.D., Mau, R., Jin, H. and Blattner, F.R. (1999)** Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res* **27**, 3821-3835.
- Ritossa, F. (1996)** Discovery of the heat shock response. *Cell Stress Chaperones* **1**, 97-98.
- Ronaghi, M. (2001)** Pyrosequencing sheds light on DNA sequencing. *Genome Res* **11**, 3-11.
- Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. (1996)** Real-time DNA sequencing using detection of pyrophosphate release. *Analytical biochemistry* **242**, 84-89.
- Ronaghi, M., Uhlen, M. and Nyren, P. (1998)** A sequencing method based on real-time pyrophosphate. *Science* **281**, 363, 365.
- Ross, M.G., Russ, C., Costello, M., Hollinger, A., Lennon, N.J., Hegarty, R., Nusbaum, C. and Jaffe, D.B. (2013)** Characterizing and measuring bias in sequence data. *Genome Biol* **14**, R51.
- Raddadi, N., Cherif, A., Mora, D., Brusetti, L., Borin, S., Boudabous, A. and Daffonchio, D. (2005)** The autolytic phenotype of the *Bacillus cereus* group. *J Appl Microbiol* **99**, 1070-1081.
- Ramos, A., Boels, I.C., de Vos, W.M. and Santos, H. (2001)** Relationship between glycolysis and exopolysaccharide biosynthesis in *Lactococcus lactis*. *Appl Environ Microbiol* **67**, 33-41.
- Randazzo, C.L., Torriani, S., Akkermans, A.D., de Vos, W.M. and Vaughan, E.E. (2002)** Diversity, dynamics, and activity of bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. *Appl Environ Microbiol* **68**, 1882-1892.
- Reddy, M.S., Vedomuthu, E.R., Washam, C.J. and Reinbold, G.W. (1972)** Agar medium for differential enumeration of lactic streptococci. *Appl Microbiol* **24**, 947-952.
- Reunanen, J., von Ossowski, I., Hendrickx, A.P., Palva, A. and de Vos, W.M. (2012)** Characterization of the SpaCBA pilus fibers in the probiotic *Lactobacillus rhamnosus* GG. *Appl Environ Microbiol* **78**, 2337-2344.
- Rodriguez-Brito, B., Li, L., Wegley, L., Furlan, M., Angly, F., Breitbart, M., Buchanan, J., Desnues, C., Dinsdale, E., Edwards, R., Felts, B., Haynes, M., Liu, H., Lipson, D., Mahaffy, J., Martin-Cuadrado, A.B., Mira, A., Nulton, J., Pasic, L., Rayhawk, S., Rodriguez-Mueller, J., Rodriguez-Valera, F., Salamon, P., Srinagesh, S., Thingstad, T.F., Tran, T., Thurber, R.V., Willner, D., Youle, M. and Rohwer, F. (2010)** Viral and microbial community dynamics in four aquatic environments. *Isme J* **4**, 739-751.

Références bibliographique

- Rodriguez-Valera, F., Martin-Cuadrado, A.B., Rodriguez-Brito, B., Pasic, L., Thingstad, T.F., Rohwer, F. and Mira, A. (2009) Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* **7**, 828-836.
- Ross, R.P., Galvin, M., McAuliffe, O., Morgan, S.M., Ryan, M.P., Twomey, D.P., Meaney, W.J. and Hill, C. (1999) Developing applications for lactococcal bacteriocins. *Antonie van Leeuwenhoek* **76**, 337-346.
- Rainey, F.A., Fritze, D. and Stackebrandt, E. (1994) The phylogenetic diversity of thermophilic members of the genus *Bacillus* as revealed by 16S rDNA analysis. *FEMS Microbiol Lett* **115**, 205-211.
- Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.H., Falcon, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Neilson, K., Friedman, R., Frazier, M. and Venter, J.C. (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology* **5**, e77.
- Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., Hoon, J., Simons, J.F., Marran, D., Myers, J.W., Davidson, J.F., Branting, A., Nobile, J.R., Puc, B.P., Light, D., Clark, T.A., Huber, M., Branciforte, J.T., Stoner, I.B., Cawley, S.E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J.A., Namsaraev, E., McKernan, K.J., Williams, A., Roth, G.T. and Bustillo, J. (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**, 348-352.
- Rothberg, J.M. and Leamon, J.H. (2008) The development and impact of 454 sequencing. *Nature biotechnology* **26**, 1117-1124.
- Ruiz-Barba, J.L., Piard, J.C. and Jimenez-Diaz, R. (1991) Plasmid profiles and curing of plasmids in *Lactobacillus plantarum* strains isolated from green olive fermentations. *The Journal of applied bacteriology* **71**, 417-421.
- Rouquette, C., Ripio, M.T., Pellegrini, E., Bolla, J.M., Tascon, R.I., Vazquez-Boland, J.A. and Berche, P. (1996) Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes*. *Mol Microbiol* **21**, 977-987.

Références bibliographique

Rouquette, C., de Chastellier, C., Nair, S. and Berche, P. (1998) The ClpC ATPase of *Listeria monocytogenes* is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Mol Microbiol* **27**, 1235-1245.

Robertson, G.T., Ng, W.L., Foley, J., Gilmour, R. and Winkler, M.E. (2002) Global transcriptional analysis of *clpP* mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence. *J Bacteriol* **184**, 3508-3520.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. and Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944-945.

S

Salama, M., Sandine, W. and Giovannoni, S. (1991) Development and application of oligonucleotide probes for identification of *Lactococcus lactis* subsp. *cremoris*. *Appl Environ Microbiol* **57**, 1313-1318.

Sanchez, Y. and Lindquist, S.L. (1990) HSP104 required for induced thermotolerance. *Science* **248**, 1112-1115.

Sanders, J.W., Venema, G. and Kok, J. (1999) Environmental stress responses in *Lactococcus lactis*. *Fems Microbiology Reviews* **23**, 483-501.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.

Schroeter, J. and Klaenhammer, T. (2009) Genomics of lactic acid bacteria. *FEMS Microbiol Lett* **292**, 1-6.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O. and Huttenhower, C. (2012) Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature methods* **9**, 811-814.

Schneewind, O., Fowler, A. and Faull, K.F. (1995) Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science* **268**, 103-106.

Schneewind, O. and Missiakas, D. (2014) Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochim Biophys Acta* **1843**, 1687-1697.

Siciliano, R.A. and Mazzeo, M.F. (2012) Molecular mechanisms of probiotic action: a proteomic perspective. *Current opinion in microbiology* **15**, 390-396.

Siezen, R.J., Renckens, B., van Swam, I., Peters, S., van Kranenburg, R., Kleerebezem, M. and de Vos, W.M. (2005) Complete sequences of four plasmids of *Lactococcus lactis*

Références bibliographique

subsp. cremoris SK11 reveal extensive adaptation to the dairy environment. *Appl Environ Microbiol* **71**, 8371-8382.

Siezen, R.J. and Teusink, B. (2005) Metabolic models for rational improvement of lactic acid bacteria as cell factories. *J Appl Microbiol* **98**, 1326-1331.

Siezen, R.J., Bayjanov, J., Renckens, B., Wels, M., van Hijum, S.A., Molenaar, D. and van Hylckama Vlieg, J.E. (2010) Complete genome sequence of *Lactococcus lactis* subsp. *lactis* KF147, a plant-associated lactic acid bacterium. *J Bacteriol* **192**, 2649-2650.

Siezen, R.J., van Enckevort, F.H., Kleerebezem, M. and Teusink, B. (2004) Genome data mining of lactic acid bacteria: the impact of bioinformatics. *Current opinion in biotechnology* **15**, 105-115.

Sell, S.M., Eisen, C., Ang, D., Zylicz, M. and Georgopoulos, C. (1990) Isolation and characterization of dnaJ null mutants of *Escherichia coli*. *J Bacteriol* **172**, 4827-4835.

Shabo, Y., Barzel, R., Margoulis, M. and Yagil, R. (2005) Camel milk for food allergies in children. *The Israel Medical Association journal : IMAJ* **7**, 796-798.

Shimizu-Kadota, M., Kato, H., Shiwa, Y., Oshima, K., Machii, M., Araya-Kojima, T., Zendo, T., Hattori, M., Sonomoto, K. and Yoshikawa, H. (2013) Genomic features of *Lactococcus lactis* IO-1, a lactic acid bacterium that utilizes xylose and produces high levels of L-lactic acid. *Bioscience, biotechnology, and biochemistry* **77**, 1804-1808.

Smid, E.J., van Enckevort, F.J., Wegkamp, A., Boekhorst, J., Molenaar, D., Hugenholtz, J., Smith, W.M., Pham, T.H., Lei, L., Dou, J., Soomro, A.H., Beatson, S.A., Dykes, G.A. and Turner, M.S. (2012) Heat resistance and salt hypersensitivity in *Lactococcus lactis* due to spontaneous mutation of llmg_1816 (gdpP) induced by high-temperature growth. *Appl Environ Microbiol* **78**, 7753-7759.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B. and Hood, L.E. (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674-679.

Subramanian, A.R., Kaufmann, M. and Morgenstern, B. (2008) DIALIGN-TX: greedy and progressive approaches for segment-based multiple sequence alignment. *Algorithms for molecular biology : AMB* **3**, 6.

Savijoki, K., Ingmer, H., Frees, D., Vogensen, F.K., Palva, A. and Varmanen, P. (2003) Heat and DNA damage induction of the LexA-like regulator HdiR from *Lactococcus lactis* is mediated by RecA and ClpP. *Mol Microbiol* **50**, 609-621.

Schouler, C., Ehrlich, S.D. and Chopin, M.C. (1994) Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. *Microbiology* **140** (Pt 11), 3061-3069.

Références bibliographique

- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D. and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539.
- Stanley, E., Fitzgerald, G.F., Le Marrec, C., Fayard, B. and van Sinderen, D. (1997) Sequence analysis and characterization of phi O1205, a temperate bacteriophage infecting *Streptococcus thermophilus* CNRZ1205. *Microbiology* **143** (Pt 11), 3417-3429.
- Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution* **4**, 406-425.
- Schatz, M.C., Langmead, B. and Salzberg, S.L. (2010) Cloud computing and the DNA data race. *Nature biotechnology* **28**, 691-693.
- Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-5864.
- Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069.
- Shendure, J. and Ji, H. (2008) Next-generation DNA sequencing. *Nature biotechnology* **26**, 1135-1145.
- Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D. and Church, G.M. (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728-1732.
- Silva, G.G., Dutilh, B.E., Matthews, T.D., Elkins, K., Schmieder, R., Dinsdale, E.A. and Edwards, R.A. (2013) Combining de novo and reference-guided assembly with scaffold_builder. *Source code for biology and medicine* **8**, 23.
- Svensater, G., Sjogreen, B. and Hamilton, I.R. (2000) Multiple stress responses in *Streptococcus mutans* and the induction of general and stress-specific proteins. *Microbiology* **146** (Pt 1), 107-117.
- Syvanen, M. (1994) Horizontal gene transfer: evidence and possible consequences. *Annual review of genetics* **28**, 237-261.
- Szyperski, T., Pellecchia, M., Wall, D., Georgopoulos, C. and Wuthrich, K. (1994) NMR structure determination of the *Escherichia coli* DnaJ molecular chaperone: secondary structure and backbone fold of the N-terminal region (residues 2-108) containing the highly conserved J domain. *Proc Natl Acad Sci U S A* **91**, 11343-11347.

Références bibliographique

- Spirig, T., Weiner, E.M. and Clubb, R.T. (2011)** Sortase enzymes in Gram-positive bacteria. *Mol Microbiol* **82**, 1044-1059.
- Steidler, L., Hans, W., Schotte, L., Neiryck, S., Obermeier, F., Falk, W., Fiers, W. and Remaut, E. (2000)** Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* **289**, 1352-1355.
- Stentz, R., Jury, K., Eaton, T., Parker, M., Narbad, A., Gasson, M. and Shearman, C. (2004)** Controlled expression of CluA in *Lactococcus lactis* and its role in conjugation. *Microbiology* **150**, 2503-2512.
- Stiles, M.E. and Holzappel, W.H. (1997)** Lactic acid bacteria of foods and their current taxonomy. *Int J Food Microbiol* **36**, 1-29.
- Stiles, M.E. (1996)** Biopreservation by lactic acid bacteria. *Antonie van Leeuwenhoek* **70**, 331-345.
- Stingle, F., Vincent, S.J., Faber, E.J., Newell, J.W., Kamerling, J.P. and Neeser, J.R. (1999)** Introduction of the exopolysaccharide gene cluster from *Streptococcus thermophilus* Sfi6 into *Lactococcus lactis* MG1363: production and characterization of an altered polysaccharide. *Mol Microbiol* **32**, 1287-1295.
- Sutherland, I.W. (1998)** Novel and established applications of microbial polysaccharides. *Trends in biotechnology* **16**, 41-46.
- Swaminathan, A., Mandlik, A., Swierczynski, A., Gaspar, A., Das, A. and Ton-That, H. (2007)** Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. *Mol Microbiol* **66**, 961-974.
- Spiess, C., Beil, A. and Ehrmann, M. (1999)** A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**, 339-347.
- Stackebrandt, E., Ludwig, W., Weizenegger, M., Dorn, S., McGill, T.J., Fox, G.E., Woese, C.R., Schubert, W. and Schleifer, K.H. (1987)** Comparative 16S rRNA oligonucleotide analyses and murein types of round-spore-forming bacilli and non-spore-forming relatives. *J Gen Microbiol* **133**, 2523-2529.
- Stuart, M.R., Chou, L.S. and Weimer, B.C. (1999)** Influence of carbohydrate starvation and arginine on culturability and amino acid utilization of *Lactococcus lactis* subsp. *lactis*. *Appl Environ Microbiol* **65**, 665-673.
- Sun, E., Zhao, L., Ren, F., Liu, S., Zhang, M. and Guo, H. (2015)** Complete genome sequence of *Bifidobacterium animalis* subsp. *lactis* A6, a probiotic strain with high acid resistance ability. *Journal of biotechnology* **200**, 8-9.

T

- Taddese, B., Upton, G.J., Bailey, G.R., Jordan, S.R., Abdulla, N.Y., Reeves, P.J. and Reynolds, C.A. (2014)** Do plants contain G protein-coupled receptors? *Plant physiology* **164**, 287-307.
- Tanskanen, E.I., Tulloch, D.L., Hillier, A.J. and Davidson, B.E. (1990)** Pulsed-Field Gel Electrophoresis of SmaI Digests of Lactococcal Genomic DNA, a Novel Method of Strain Identification. *Appl Environ Microbiol* **56**, 3105-3111.
- Teixeira, P., Castro, H., Mohacsi-Farkas, C. and Kirby, R. (1997)** Identification of sites of injury in *Lactobacillus bulgaricus* during heat stress. *J Appl Microbiol* **83**, 219-226.
- Teuber, M. and Perreten, V. (2000)** Role of milk and meat products as vehicles for antibiotic-resistant bacteria. *Acta veterinaria Scandinavica Supplementum* **93**, 75-87; discussion 111-117.
- Tanigawa, K., Kawabata, H. and Watanabe, K. (2010)** Identification and typing of *Lactococcus lactis* by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Appl Environ Microbiol* **76**, 4055-4062.
- Teusink, B., van Enckevort, F.H., Francke, C., Wiersma, A., Wegkamp, A., Smid, E.J. and Siezen, R.J. (2005)** In silico reconstruction of the metabolic pathways of *Lactobacillus plantarum*: comparing predictions of nutrient requirements with those from growth experiments. *Appl Environ Microbiol* **71**, 7253-7262.
- Tissieres, A., Mitchell, H.K. and Tracy, U.M. (1974)** Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs. *J Mol Biol* **84**, 389-398.
- Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter, J.C., Bork, P., Hugenholtz, P. and Rubin, E.M. (2005)** Comparative metagenomics of microbial communities. *Science* **308**, 554-557.
- Tansley, A. G. (1958)** Sir Arthur Georges Tansley, F.R.S. *The Journal of Ecology*. **46** (1), 1-8.
- Tatusova, T., Ciufu, S., Federhen, S., Fedorov, B., McVeigh, R., O'Neill, K., Tolstoy, I. and Zaslavsky, L. (2015)** Update on RefSeq microbial genomes resources. *Nucleic Acids Res* **43**, D599-605.
- Tatusova, T., Ciufu, S., Fedorov, B., O'Neill, K. and Tolstoy, I. (2014)** RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res* **42**, D553-559.

Références bibliographique

- Tatusova, T., Pruitt, K.D., Maglott, D.R. and Murphy, T.D. (2015)** Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res* **43**, D36-42.
- Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R. and Grandi, G. (2006)** Pili in gram-positive pathogens. *Nat Rev Microbiol* **4**, 509-519.
- Toba, T., Kotani, T. and Adachi, S. (1991)** Capsular polysaccharide of a slime-forming *Lactococcus lactis* ssp. *cremoris* LAPT 3001 isolated from Swedish fermented milk 'langfil'. *Int J Food Microbiol* **12**, 167-171.
- Ton-That, H., Liu, G., Mazmanian, S.K., Faull, K.F. and Schneewind, O. (1999)** Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc Natl Acad Sci U S A* **96**, 12424-12429.
- Ton-That, H., Marraffini, L.A. and Schneewind, O. (2004)** Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol Microbiol* **53**, 251-261.
- Ton-That, H. and Schneewind, O. (2003)** Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* **50**, 1429-1438.
- Tsukioka, Y., Yamashita, Y., Oho, T., Nakano, Y. and Koga, T. (1997)** Biological function of the dTDP-rhamnose synthesis pathway in *Streptococcus mutans*. *J Bacteriol* **179**, 1126-1134.
- Terzaghi, B.E. and Sandine, W.E. (1975)** Improved medium for lactic streptococci and their bacteriophages. *Appl Microbiol* **29**, 807-813.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994)** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.
- Treangen, T.J., Sommer, D.D., Angly, F.E., Koren, S. and Pop, M. (2011)** Next generation sequence assembly with AMOS. *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]* **Chapter 11**, Unit 11 18.
- Tschoeke, D.A., Moreira, A.P., Chimetto Tonon, L.A., de Mesquita, M.M., Gregoracci, G.B., Gomez-Gil, B., Valle, R., Thompson, C.C. and Thompson, F.L. (2014)** Exploring the Genome of Cheese Starter Lactic Acid Bacterium *Lactococcus lactis* subsp. *lactis* CECT 4433. *Genome announcements* **2**.

U

- Ulrich, H.D. (2002)** Natural substrates of the proteasome and their recognition by the ubiquitin system. *Current topics in microbiology and immunology* **268**, 137-174.

Références bibliographique

Ueda, K., Seki, T., Kudo, T., Yoshida, T. and Kataoka, M. (1999) Two distinct mechanisms cause heterogeneity of 16S rRNA. *J Bacteriol* **181**, 78-82.

V

van Asseldonk, M., Simons, A., Visser, H., de Vos, W.M. and Simons, G. (1993) Cloning, nucleotide sequence, and regulatory analysis of the *Lactococcus lactis* dnaJ gene. *J Bacteriol* **175**, 1637-1644.

Van der Meulen, R., Grosu-Tudor, S., Mozzi, F., Vaningelgem, F., Zamfir, M., de Valdez, G.F. and De Vuyst, L. (2007) Screening of lactic acid bacteria isolates from dairy and cereal products for exopolysaccharide production and genes involved. *Int J Food Microbiol* **118**, 250-258.

van de Guchte, M., Ehrlich, S.D. and Maguin, E. (2001) Production of growth-inhibiting factors by *Lactobacillus delbrueckii*. *J Appl Microbiol* **91**, 147-153.

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., Bessieres, P., Weissenbach, J., Ehrlich, S.D. and 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.

van de Guchte, M., Serror, P., Chervaux, C., Smokvina, T., Ehrlich, S.D. and Maguin, E. (2002) Stress responses in lactic acid bacteria. *Antonie van Leeuwenhoek* **82**, 187-216.

van Rooijen, R.J., Gasson, M.J. and de Vos, W.M. (1992) Characterization of the *Lactococcus lactis* lactose operon promoter: contribution of flanking sequences and LacR repressor to promoter activity. *J Bacteriol* **174**, 2273-2280.

VanBogelen, R.A., Acton, M.A. and Neidhardt, F.C. (1987a) Induction of the heat shock regulon does not produce thermotolerance in *Escherichia coli*. *Genes & development* **1**, 525-531.

VanBogelen, R.A., Kelley, P.M. and Neidhardt, F.C. (1987b) Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. *J Bacteriol* **169**, 26-32.

van der Vies, S.M., Gatenby, A.A. and Georgopoulos, C. (1994) Bacteriophage T4 encodes a co-chaperonin that can substitute for *Escherichia coli* GroES in protein folding. *Nature* **368**, 654-656.

Références bibliographique

- van Hylckama Vlieg, J.E., Rademaker, J.L., Bachmann, H., Molenaar, D., Kelly, W.J. and Siezen, R.J. (2006) Natural diversity and adaptive responses of *Lactococcus lactis*. *Current opinion in biotechnology* **17**, 183-190.
- Varmanen, P., Ingmer, H. and Vogensen, F.K. (2000) *ctsR* of *Lactococcus lactis* encodes a negative regulator of *clp* gene expression. *Microbiology* **146** (Pt 6), 1447-1455.
- Villegas, A. and Kropinski, A.M. (2008) An analysis of initiation codon utilization in the Domain Bacteria - concerns about the quality of bacterial genome annotation. *Microbiol-Sgm* **154**, 2559-2561.
- Villion, M., Chopin, M.C., Deveau, H., Ehrlich, S.D., Moineau, S. and Chopin, A. (2009) P087, a lactococcal phage with a morphogenesis module similar to an *Enterococcus faecalis* prophage. *Virology* **388**, 49-56.
- Volker, U., Engelmann, S., Maul, B., Riethdorf, S., Volker, A., Schmid, R., Mach, H. and Hecker, M. (1994) Analysis of the induction of general stress proteins of *Bacillus subtilis*. *Microbiology* **140** (Pt 4), 741-752.
- Van Hoorde, K., Verstraete, T., Vandamme, P. and Huys, G. (2008) Diversity of lactic acid bacteria in two Flemish artisan raw milk Gouda-type cheeses. *Food microbiology* **25**, 929-935.
- van Kranenburg, R., Boels, I.C., Kleerebezem, M. and de Vos, W.M. (1999a) Genetics and engineering of microbial exopolysaccharides for food: approaches for the production of existing and novel polysaccharides. *Current opinion in biotechnology* **10**, 498-504.
- van Kranenburg, R., Kleerebezem, M. and de Vos, W.M. (2000) Nucleotide sequence analysis of the lactococcal EPS plasmid pNZ4000. *Plasmid* **43**, 130-136.
- van Kranenburg, R., Marugg, J.D., van, S., II, Willem, N.J. and de Vos, W.M. (1997) Molecular characterization of the plasmid-encoded *eps* gene cluster essential for exopolysaccharide biosynthesis in *Lactococcus lactis*. *Mol Microbiol* **24**, 387-397.
- van Kranenburg, R., van, S., II, Marugg, J.D., Kleerebezem, M. and de Vos, W.M. (1999b) Exopolysaccharide biosynthesis in *Lactococcus lactis* NIZO B40: functional analysis of the glycosyltransferase genes involved in synthesis of the polysaccharide backbone. *J Bacteriol* **181**, 338-340.
- van Kranenburg, R., Vos, H.R., van, S., II, Kleerebezem, M. and de Vos, W.M. (1999c) Functional analysis of glycosyltransferase genes from *Lactococcus lactis* and other gram-positive cocci: complementation, expression, and diversity. *J Bacteriol* **181**, 6347-6353.

Références bibliographique

- Vaughan, E.E., David, S. and de Vos, W.M. (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.
- Vedamuthu, E.R. and Neville, J.M. (1986)** Involvement of a Plasmid in Production of Ropiness (Mucoïdness) in Milk Cultures by *Streptococcus cremoris* MS. *Appl Environ Microbiol* **51**, 677-682.
- von Wright, A. and Tynkkynen, S. (1987)** Construction of *Streptococcus lactis* subsp. *lactis* Strains with a Single Plasmid Associated with Mucoïd Phenotype. *Appl Environ Microbiol* **53**, 1385-1386.
- van Hijum, S.A., Zomer, A.L., Kuipers, O.P. and Kok, J. (2005)** Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies. *Nucleic Acids Res* **33**, W560-566.
- Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. and Swings, J. (1996)** Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological reviews* **60**, 407-438.
- Varin, T., Lovejoy, C., Jungblut, A.D., Vincent, W.F. and Corbeil, J. (2012)** Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Appl Environ Microbiol* **78**, 549-559.
- Venkatesan, B.M. and Bashir, R. (2011)** Nanopore sensors for nucleic acid analysis. *Nature nanotechnology* **6**, 615-624.

W

- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982)** Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**, 945-951.
- Ward, L.J., Brown, J.C. and Davey, G.P. (1998)** Two methods for the genetic differentiation of *Lactococcus lactis* ssp. *lactis* and *cremoris* based on differences in the 16S rRNA gene sequence. *FEMS Microbiol Lett* **166**, 15-20.
- Wall, D., Zylicz, M. and Georgopoulos, C. (1994)** The NH₂-terminal 108 amino acids of the *Escherichia coli* DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. *J Biol Chem* **269**, 5446-5451.
- Wawrzynow, A., Banecki, B. and Zylicz, M. (1996)** The Clp ATPases define a novel class of molecular chaperones. *Mol Microbiol* **21**, 895-899.

Références bibliographique

- Wawrzynow, A., Wojtkowiak, D., Marszalek, J., Banecki, B., Jonsen, M., Graves, B., Georgopoulos, C. and Zylicz, M. (1995) The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. *EMBO J* **14**, 1867-1877.
- Wegkamp, A., van Oorschot, W., de Vos, W.M. and Smid, E.J. (2007) Characterization of the role of para-aminobenzoic acid biosynthesis in folate production by *Lactococcus lactis*. *Appl Environ Microbiol* **73**, 2673-2681.
- Whitaker, R.D. and Batt, C.A. (1991) Characterization of the Heat Shock Response in *Lactococcus lactis* subsp. *lactis*. *Appl Environ Microbiol* **57**, 1408-1412.
- Wolf, Y.I., Aravind, L., Grishin, N.V. and Koonin, E.V. (1999) Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* **9**, 689-710.
- Wu, B., Ang, D., Snavely, M. and Georgopoulos, C. (1994) Isolation and characterization of point mutations in the *Escherichia coli* *grpE* heat shock gene. *J Bacteriol* **176**, 6965-6973.
- Wydau, S., Dervyn, R., Anba, J., Dusko Ehrlich, S. and Maguin, E. (2006) Conservation of key elements of natural competence in *Lactococcus lactis* ssp. *FEMS Microbiol Lett* **257**, 32-42.
- Wyszynska, A., Kobierecka, P., Bardowski, J. and Jagusztyn-Krynicka, E.K. (2015) Lactic acid bacteria-20 years exploring their potential as live vectors for mucosal vaccination. *Appl Microbiol Biot* **99**, 2967-2977.
- Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738.
- Wegmann, U., O'Connell-Motherway, M., Zomer, A., Buist, G., Shearman, C., Canchaya, C., Ventura, M., Goesmann, A., Gasson, M.J., Kuipers, O.P., van Sinderen, D. and Kok, J. (2007) Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J Bacteriol* **189**, 3256-3270.
- Welch, R.A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L., Sonnenberg, M.S. and Blattner, F.R. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **99**, 17020-17024.
- Whittam, T.S. and Bumbaugh, A.C. (2002) Inferences from whole-genome sequences of bacterial pathogens. *Current opinion in genetics & development* **12**, 719-725.

Références bibliographique

- Wetzstein, M., Volker, U., Dedio, J., Lobau, S., Zuber, U., Schiesswohl, M., Herget, C., Hecker, M. and Schumann, W. (1992)** Cloning, sequencing, and molecular analysis of the dnaK locus from *Bacillus subtilis*. *J Bacteriol* **174**, 3300-3310.
- Wiegert, T. and Schumann, W. (2003)** Analysis of a DNA-binding motif of the *Bacillus subtilis* HrcA repressor protein. *FEMS Microbiol Lett* **223**, 101-106.
- Winter, J., Linke, K., Jatzek, A. and Jakob, U. (2005)** Severe oxidative stress causes inactivation of DnaK and activation of the redox-regulated chaperone Hsp33. *Mol Cell* **17**, 381-392.
- Wise, A.A. and Price, C.W. (1995)** Four additional genes in the sigB operon of *Bacillus subtilis* that control activity of the general stress factor sigma B in response to environmental signals. *J Bacteriol* **177**, 123-133.
- Woese, C.R. (1987)** Bacterial evolution. *Microbiological reviews* **51**, 221-271.
- Wommack, K.E., Bhavsar, J. and Ravel, J. (2008) Metagenomics: read length matters. *Appl Environ Microbiol* **74**, 1453-1463.
- Walling, E., Gindreau, E. and Lonvaud-Funel, A. (2005)** A putative glucan synthase gene dps detected in exopolysaccharide-producing *Pediococcus damnosus* and *Oenococcus oeni* strains isolated from wine and cider. *Int J Food Microbiol* **98**, 53-62.
- Wetzstein, M., Volker, U., Dedio, J., Lobau, S., Zuber, U., Schiesswohl, M., Herget, C., Hecker, M. and Schumann, W. (1992)** Cloning, sequencing, and molecular analysis of the dnaK locus from *Bacillus subtilis*. *J Bacteriol* **174**, 3300-3310.
- Wiegert, T. and Schumann, W. (2003)** Analysis of a DNA-binding motif of the *Bacillus subtilis* HrcA repressor protein. *FEMS Microbiol Lett* **223**, 101-106.
- Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F.D. and Lewis, J.D. (2011)** Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105-108.

Y

- Yamamori, T., Ito, K., Nakamura, Y. and Yura, T. (1978)** Transient regulation of protein synthesis in *Escherichia coli* upon shift-up of growth temperature. *J Bacteriol* **134**, 1133-1140.
- Yang, X., Wang, Y. and Huo, G. (2013)** Complete Genome Sequence of *Lactococcus lactis* subsp. *lactis* KLDS4.0325. *Genome announcements* **1**.

Références bibliographique

- Yang, B., Peng, Y., Leung, H.C., Yiu, S.M., Chen, J.C. and Chin, F.Y. (2010)** Unsupervised binning of environmental genomic fragments based on an error robust selection of l-mers. *BMC bioinformatics* **11 Suppl 2**, S5.
- Yap, W.H., Zhang, Z. and Wang, Y. (1999)** Distinct types of rRNA operons exist in the genome of the actinomycete *Thermomonospora chromogena* and evidence for horizontal transfer of an entire rRNA operon. *J Bacteriol* **181**, 5201-5209.
- Yasui, K., Kano, Y., Tanaka, K., Watanabe, K., Shimizu-Kadota, M., Yoshikawa, H. and Suzuki, T. (2009)** Improvement of bacterial transformation efficiency using plasmid artificial modification. *Nucleic Acids Res* **37**, e3.
- Yoong, P., Schuch, R., Nelson, D. and Fischetti, V.A. (2004)** Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant *Enterococcus faecalis* and *Enterococcus faecium*. *J Bacteriol* **186**, 4808-4812.
- Young, R. and Blasi, U. (1995)** Holins: form and function in bacteriophage lysis. *FEMS Microbiol Rev* **17**, 191-205.
- Yura, T., Nagai, H. and Mori, H. (1993)** Regulation of the heat-shock response in bacteria. *Annu Rev Microbiol* **47**, 321-350.
- Yeung, M.K., Donkersloot, J.A., Cisar, J.O. and Ragsdale, P.A. (1998)** Identification of a gene involved in assembly of *Actinomyces naeslundii* T14V type 2 fimbriae. *Infection and immunity* **66**, 1482-1491.
- Yeung, M.K. and Ragsdale, P.A. (1997)** Synthesis and function of *Actinomyces naeslundii* T14V type 1 fimbriae require the expression of additional fimbria-associated genes. *Infection and immunity* **65**, 2629-2639.
- Yokoi, H., Watanabe, T., Fujii, Y., Mukai, T., Toba, T. and Adachi, S. (1991)** Some taxonomical characteristics of encapsulated *Lactobacillus* sp. KPB-167B isolated from kefir grains and characterization of its extracellular polysaccharide. *Int J Food Microbiol* **13**, 257-264.

Z

- Zellmeier, S., Zuber, U., Schumann, W. and Wiegert, T. (2003)** The absence of FtsH metalloprotease activity causes overexpression of the sigmaW-controlled pbpE gene, resulting in filamentous growth of *Bacillus subtilis*. *J Bacteriol* **185**, 973-982.

Références bibliographique

- Zeilstra-Ryalls, J., Fayet, O. and Georgopoulos, C. (1994)** Two classes of extragenic suppressor mutations identify functionally distinct regions of the GroEL chaperone of *Escherichia coli*. *J Bacteriol* **176**, 6558-6565.
- Zhang, J., Wang, F., Xu, H., Yu, J., Liu, W., Bao, Q., Sun, Z. and Zhang, H. (2010)** [Comparison of six molecular methods for differentiation of *Lactococcus lactis* subsp. *lactis* and *cremoris*]. *Wei sheng wu xue bao = Acta microbiologica Sinica* **50**, 1670-1676.
- Zhang, W., Wang, C., Huang, C., Yu, Q., Liu, H., Zhang, C. and Pei, X. (2011)** Construction and expression of food-grade beta-galactosidase gene in *Lactococcus Lactis*. *Curr Microbiol* **62**, 639-644.
- Zuo, F.L., Feng, X.J., Chen, L.L. and Chen, S.W. (2014)** Identification and partial characterization of lactic acid bacteria isolated from traditional dairy products produced by herders in the western Tianshan Mountains of China. *Letters in applied microbiology* **59**, 549-556.
- Zheng, Y., Roberts, R.J. and Kasif, S. (2002)** Genomic functional annotation using co-evolution profiles of gene clusters. *Genome Biol* **3**, RESEARCH0060.
- Zhou, Y., Gottesman, S., Hoskins, J.R., Maurizi, M.R. and Wickner, S. (2001)** The RssB response regulator directly targets sigma(S) for degradation by ClpXP. *Genes & development* **15**, 627-637.
- Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J. and Wishart, D.S. (2011)** PHAST: a fast phage search tool. *Nucleic Acids Res* **39**, W347-352.
- Zerbino, D.R. and Birney, E. (2008)** Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821-829.
- Zuber, U. and Schumann, W. (1994)** CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*. *J Bacteriol* **176**, 1359-1363.
- Zuker, M. (2003)** Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-3415.

ANNEXE

Annexe

Milieux de cultures et conditions de croissance

1-1- Milieu M17 (Terzaghi and Sandine, 1975)

Ce milieu est destiné à la culture des lactocoques. Nous avons utilisé le milieu M17 (BiokarDiagnostics, Beauvais, France) déshydraté prêt à l'emploi dont la composition est la suivante (pour 1 L): tryptone 5,0 g, peptone de soja 5,0 g, infusion de viande 5,0 g, extrait de levure 2,5 g, glycérohydrogénophosphate de sodium 19,0 g, lactose 5,0 g, acide ascorbique 0,5 g, sulfate de magnésium 0,25 g, pH = 7,1.

Le milieu M17 est stérilisé à 120°C pendant 20 min à l'autoclave.

Dans le cas de bactéries lactiques Lac- (incapable d'utiliser le lactose), 25 ml d'une solution de glucose à 20% sont ajoutés à ce milieu M17. La forme solide du milieu M17 est obtenue par ajout d'agar avant autoclavage à la concentration de 15 g/L.

1-2- Milieux à base de lait

Ces milieux sont utilisés dans l'identification de souches bactériennes à activité protéolytique et à pouvoir acidifiant. La forme liquide du milieu lait est utilisée dans l'étude des cinétiques de croissance chez les lactocoques. La forme gélosée de ces différents milieux comprend de l'agar « Difco » à 12 g/L de concentration finale.

1-2-1- Milieu lait

Ce milieu est reconstitué à 250 g/L avec la poudre de « lait Magermilchpulver » (SaliterGmbH D-87634Obergünzburg) dans de l'eau déminéralisée stérile. La solution de lait ne subit aucun autoclavage.

1-2-2- Milieu lait citraté

8 g de « lait Magermilchpulver » sont dissous dans 100 mL d'eau déminéralisée stérile, puis 1,3 mL de citrate de sodium stérile (1 M, PH 6,5) sont ajoutés. Après agitation puis incubation à température ambiante pendant environ 20 min, la préparation prend un aspect jaune translucide. 100 mL d'agar (24 g/L) en surfusion sont mélangés au milieu de culture. Le mélange est homogénéisé, puis réparti dans des boîtes de Petri à raison de 25 mL par boîte.

1-2-3- Milieu lait glucose

20 g de la poudre « lait Magermilchpulver » sont dissous dans 100mL d'eau déminéralisée stérile, puis 5 mL de solution de glucose à 20% sont ajoutés. La solution de glucose est stérilisée par filtration. 100 mL d'agar (15 g/L) en surfusion sont mélangés au milieu de

culture. Le mélange est homogénéisé, puis réparti dans des boîtes de Petri à raison de 25 mL par boîte.

1-2-4- Milieu lait X-Gal

C'est un milieu qui permet de différencier les souches de bactéries lactiques capables de fermenter le lactose de celles qui ne le fermentent pas. Il est composé du milieu M17 solide (15 g/l d'agar) contenant 0,25% de lactose préparé et autoclavé préalablement, ensuite 2 ml/l de X-Gal (5-Bromo-4chloro-3-indolyl- β -D-galactopyranoside) sont ajoutés au milieu. Une solution stock du X-Gal est préparée (20 mg/ml) dans diméthylsulfoxyde (DMSO).

1-3- Milieu HHD « *homofermentative-heterofermentative differential medium* » (McDonald et al. 1987)

Ce milieu est utilisé sous forme liquide ou solide, il permet la mise en évidence des voies fermentaires des bactéries étudiées (la voie des acides mixtes ou la voie du butane-2,3-diol). Il est composé de : fructose 2,5 g/l, KH_2PO_4 2,5 g/l, Trypticase peptone 10 g/l, phytonepeptone 1,5 g/L, casaminoacids 3g/L, extrait de levure 1 g/L, Tween 80 1 g/l, vert de Bromocrésol 20 ml. Après dissolution des ingrédients, le milieu est ajusté à pH à 7,0, autoclavé pendant 15 minutes, puis réparti dans des tubes stériles (10 mL de milieu par tube). Une solution stock est préparée par addition de 0,1 g du vert de bromocrésol dans 30 ml du NaOH 0,01N.

1-6- Milieu KMK (Kempner and McKay, 1980)

C'est un milieu qui permet de différencier les souches de bactéries lactiques capables de fermenter le citrate de celles qui ne le fermentent pas. Le milieu de base est composé de : lait écrémé 10 g/L, 2,5 g/L d'hydrolysate de caséine et 5 g/L de glucose. Le pH du milieu est ajusté à 6,6, puis stérilisé à 115°C pendant 15 min à l'autoclave.

Deux solutions A et B sont préparées et stérilisées par filtration sur une membrane millipore de 0,45 μm :

A : 10% (m/v) de ferricyanure de potassium

B : 1 g de citrate de fer et 1 g de citrate de sodium dans 40 mL d'eau distillée.

Dix millilitres de chaque solution sont ajoutés stérilement dans un litre du milieu de base, après homogénéisation ; une solution d'agar en surfusion est ajoutée au milieu KMK à une concentration finale de 15g/L.

2- Conditions de croissance

2-1- Cultures

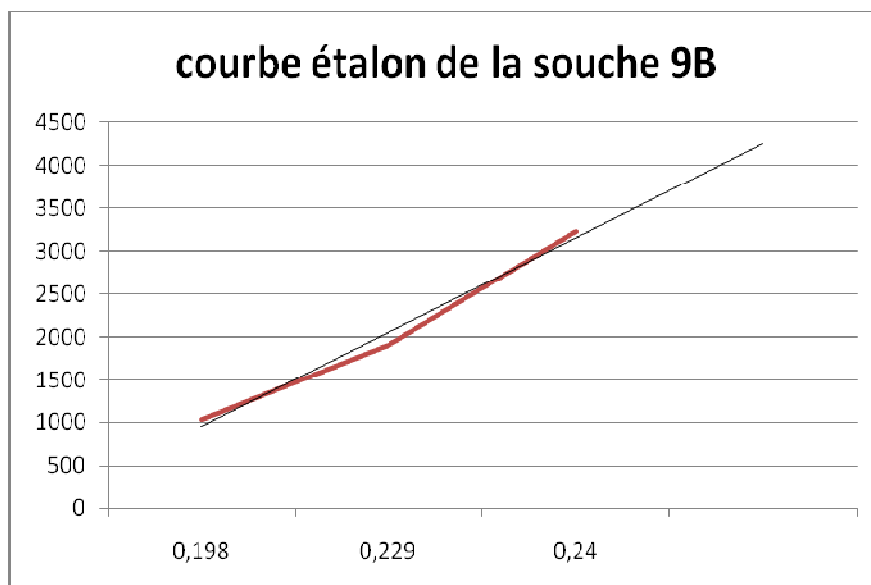
Des pré-cultures de 10 mL de milieu liquide M17 incubées pendant une nuit à 30°C ont servi pour inoculer 50 ou 100 mL de milieu M17 ou milieu lait liquide. Les cultures sont incubées sans agitation à 30°C, 37°C, 45°C ou 50°C. Ces milieux sont également utilisés pour l'étude des cinétiques de croissance bactériennes pendant 10 h. En milieu solide les cultures sont incubées à 30°C, en aérobiose car les lactocoques sont microaérophiles et tolèrent ainsi de petites quantités d'oxygène atmosphérique.

2-2- Mesure de la densité optique

La densité cellulaire d'une suspension bactérienne de lactocoques est mesurée par turbidimétrie à une longueur d'onde de 600 nm. En milieu M17 la DO est directement relevée (lecture contre du M17), alors qu'en milieu lait, une transparasitaion du milieu est nécessaire par l'ajout de 3,5 ml d'EDTA (10 g/L pH12) à 1mL de culture. La lecture se fait contre un témoin composé de 3,5 ml d'EDTA (10 g/L pH12) additionné de 1mL de milieu lait stérile.

3- Établissement de la courbe d'étalonnage de la croissance bactérienne

À l'aide de ces courbes étalons on a pu évaluer le nombre UFC correspondant aux DO prélevées lors des cinétiques de croissance et puis tracer les courbes de croissance. Le principe d'étalonnage est de faire des dilution de $1/6^{\text{ème}}$ $1/8^{\text{ème}}$ et $1/10^{\text{ème}}$ à partir de culture de nuit. L'étalonnage est propre à la même souche. La longueur d'onde utilisée en turbidimétrie = 600 nm.



Annexe

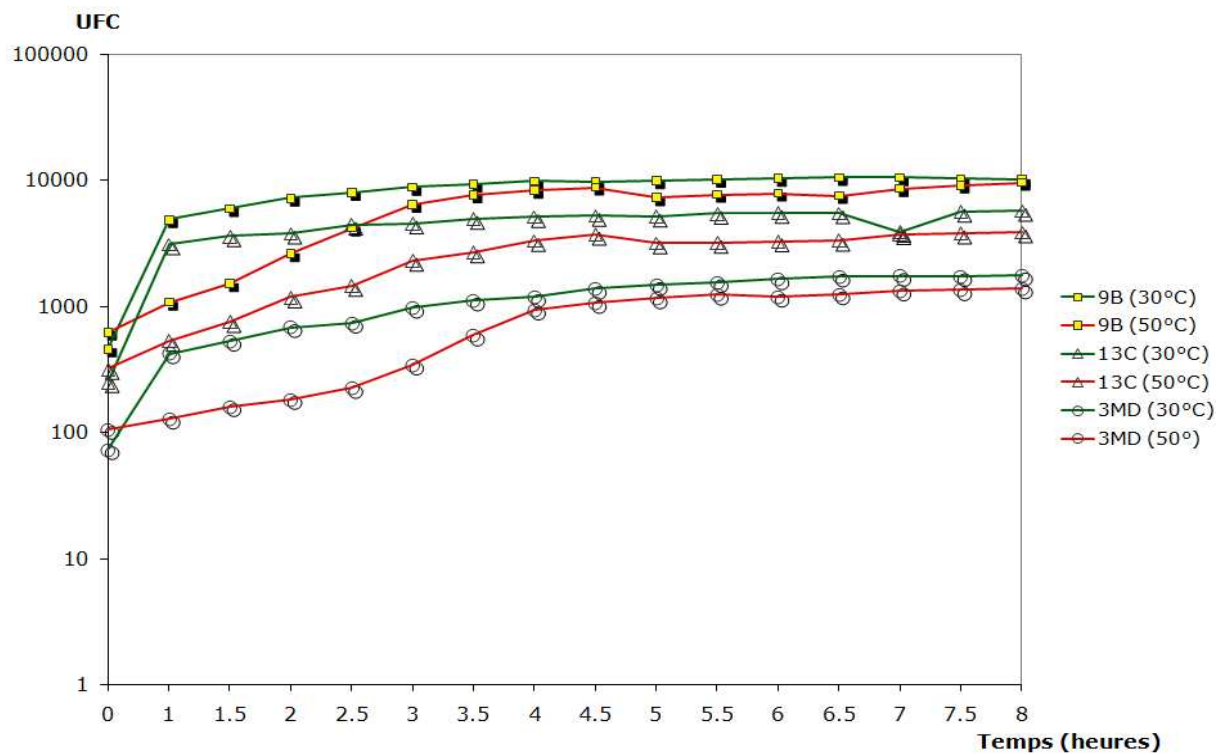


Figure A : comparaison des cinétiques de croissance à 30 et 50°C en milieu M17 des Lactocoques issus du lait cru de chamelle ou du lait de chèvre

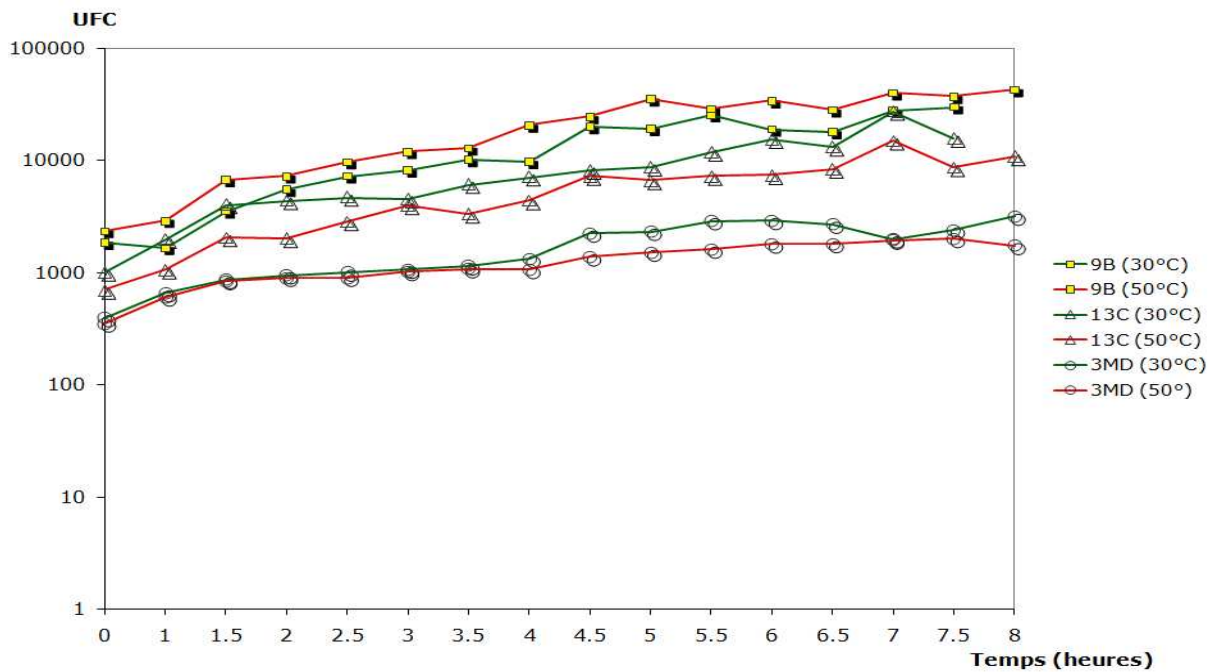


Figure B : comparaison des cinétiques de croissance à 30 et 50°C en milieu Lait des Lactocoques issus du lait cru de chamelle ou du lait de chèvre

Annexe

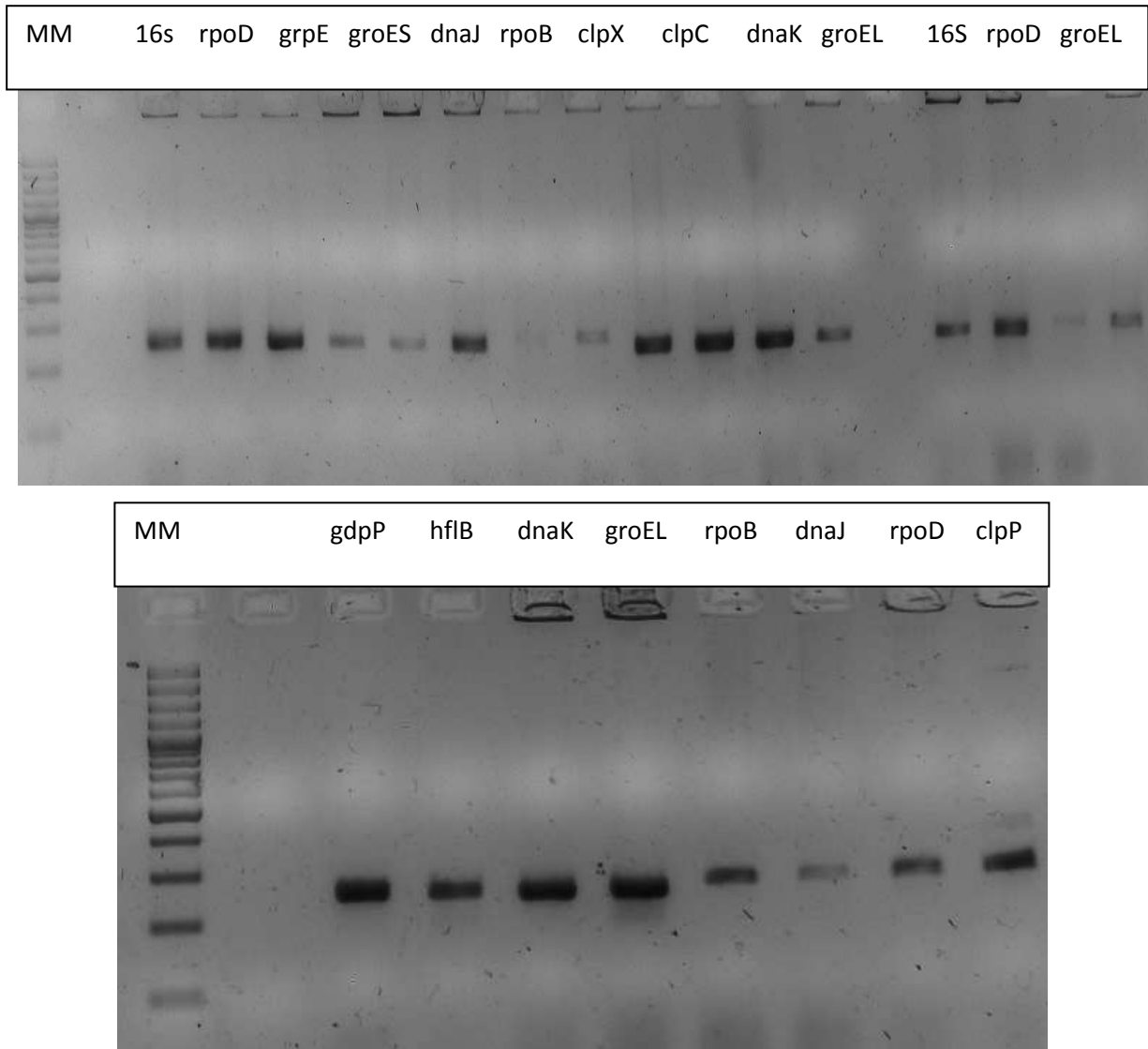


Figure : Amplification par PCR des différents gènes de choc thermique étudiés chez *Lactococcus lactis* HD_9B