THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Institut National Polytechnique de Toulouse*
Discipline ou spécialité : *Pathologie, Toxicologie, Génétique & Nutrition*

Présentée et soutenue par *Laurent-Xavier NOUVEL*
Le *26 novembre 2009*

ETUDE DE LA DIVERSITE GENETIQUE DE *MYCOPLASMA AGALACTIAE* : PLASTICITE DES GENOMES, MOBILOME ET DYNAMIQUE DE SURFACE

JURY

M. Henri DABERNAT (Président)	*Professeur à l'Université Paul Sabatier de Toulouse*
Mme Cécile BEBEAR	*Professeur à l'Université Victor Ségalen de Bordeaux*
M. Joachim FREY	*Professeur à l'Université de Berne, Suisse*
Mme Brigitte GICQUEL	*Professeur à l'Institut Pasteur de Paris*
M. Yves LE LOIR	*Directeur de Recherche à l'INRA de Rennes*
M. François THIAUCOURT	*Directeur de Recherche au CIRAD de Montpellier*
Mme Christine CITTI	*Directeur de Recherche à l'INRA de Toulouse*

Ecole doctorale : *Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries*
Unité de recherche : *UMR1225 Interactions Hotes-Agents Pathogènes, INRA, ENVT*
Directeur de Thèse : *Mme Christine CITTI*
Rapporteurs : *Pr. Cécile BEBEAR, Pr. Joachim FREY*

REMERCIEMENTS

Ce travail de thèse a été réalisé au sein de l'équipe «Mycoplasmoses des ruminants, pathologie mammaire et sécurité du lait » rattachée à l'unité mixte de recherche INRA, ENVT, Interactions Hôtes – Agents Pathogènes. Je tiens tout d'abord à remercier Xavier Berthelot de m'avoir accueilli dans son laboratoire et son équipe pédagogique. Merci pour son soutien et ses encouragements côté recherche et côté enseignement, qu'il trouve ici l'expression de ma reconnaissance et de mon profond respect.

Je tiens à remercier chaleureusement Christine Citti pour m'avoir plus qu'aidé et dirigé tout au long de ce travail. Ce fut un soutien quotidien et un partage permanent. Qu'elle trouve ici l'expression de ma sincère gratitude.

Je remercie aussi Brigitte Gicquel qui m'a toujours accompagné, accueilli et soutenu depuis mon DEA dans son laboratoire jusqu'à mes périgrinations vers les ruminants.

Je tiens également à remercier monsieur le professeur Henri Dabernat de me faire l'honneur de présider mon jury de thèse. C'est toujours un plaisir renouvelé qu'un doctorat sous sa présidence.

J'adresse aussi tous mes remerciements à madame le professeur Cécile Bébéar et à monsieur le professeur Joachim Frey pour avoir accepté d'examiner et de juger ce travail.

Merci également à messieurs Yves Le Loir et François Thiaucourt d'avoir accepté de participer au jury de cette thèse.

Je remercie tous ceux qui de près ou de loin ont contribué à ce travail. Un grand merci à Eveline Sagné pour son dévouement et son aide au laboratoire. Merci aussi à Pascal Sirand-Pugnet, Marc Marenda, compagnons lointains rapprochés par la toile bioinformatique. Merci à François Poumarat, Dominique Le Grand et Florence Tardy pour leur appui à la culture des mycoplasmes.

Tous mes plus sincères remerciements à toutes les personnes du laboratoire ou de l'étage de Pathologie de la Reproduction : Christine, Xavier, Eveline, Marie-Claude, El Haadi, Maria, Agnès S., Eric, Dominique, Agnès B., Patricia, Sébastien et Nicole.
Merci aussi à Philippe, bienvenu pour ses lectures et relectures au beau milieu de ses péripéties d'installation.

Sans oublier un grand merci à Mirentxu pour la relecture finale.

Merci à tous mes proches et amis, merci pour votre présence et vos encouragements.
Voici l'aboutissement d'un long cheminement à vos côtés,
c'est l'occasion de vous manifester toute mon affection.
Vous appartenez à une longue liste,
vous répertorier tous pourrait constituer le sujet de ma prochaine thèse,
ne soyez pas fâchés si votre nom n'y apparaît pas, il est écrit à l'encre sympathique :

Gure ama,

Papa et Françoise,
Benoît, Yves, Catherine, Guillaume, Marie, François et Sara,
Tante José, Otto André et toute la tribu Dokhélar,
Tante Marie-Henriette, Oncle Pierre et tous les Nouvel,
Marie-Christine, Catherine, toutes mes cousines et cousins,

Virginie, Vania, Valérie, Gustavo, Brigitte et Olivier, tous ceux de l'UGM,
Caroline, Laurent, Laurent et Louis Pasteur,

Olivier, Corinne, Lucille, Fabien, Guillaume, Pascal et tous les ossalois,
Anne et Sébastien, Henri et tous les Gouffé

Louis, Geneviève, Sophie, Bernard et tous les gais montagnards,

Céline et Sébastien, Annie et Didier,

Céline, Bruno, Aurore et Fabrice, Raphaël, Vincent et Sandrine, Loïc et Fred, Benoît et Isabelle, Jean-Marie, Benoît-Marie, Jean-Claude, Thomas,

Frédéric, Tiphaine, Erwan, Alexandre, Josiane, Stéphanie, tous les Jonckheere - Gimenez,

Mirentxuxeri,

…

Le fait le plus élémentaire qui saisisse la conscience de l'Homme peut être exprimé comme suit : "Je suis vie qui veut vivre parmi d'autres vies qui veulent vivre". (…) Le monde présente le spectacle terrifiant des volontés de vivre qui s'entre-déchirent. Une existence ne survit qu'aux dépens des autres, l'une détruit l'autre. Chaque volonté de vivre n'est volonté que dans l'opposition des autres, en ne connaissant rien d'elles. Mais en moi la volonté de vivre est devenue consciente de l'existence des autres volontés de vivre. En moi apparaît une aspiration à devenir universel et à me fondre dans l'Un.[1]

La seule possibilité de donner un sens à son existence, c'est d'élever sa relation naturelle avec le monde à la hauteur d'une relation spirituelle. (…) L'éthique de l'amour, élargie jusqu'à l'universel.[2]

Albert Schweitzer

[1] "Ethique du respect de la vie" cité dans Albert Schweitzer, *Humanisme et mystique*, Albin Michel, 1995.
[2] Albert Schweitzer, *Ma vie et ma pensée*, Albin Michel, 1960.

TABLE DES MATIERES

INTRODUCTION 15

1. MYCOPLASMES : GENERALITES 15
2. MYCOPLASMES ET VARIABILITE ANTIGENIQUE 19
   - *2.1 Mutations spontanées et variabilité de surface 23*
     - 2.1.1 Mésappariements par glissement des brins d'ADN (*slipped-strand mispairing*) 23
     - 2.1.2 Mutations spontanées par substitution 24
   - *2.2 Conversions de gènes et dynamique de surface 25*
   - *2.3 Réarrangements spécifiques de l'ADN (système "flip/flop") 26*
     - 2.3.1 Interversions des régions promotrices 26
     - 2.3.2 Réarrangements devant un unique locus d'expression 27
3. MYCOPLASMES, FACTEURS DE VIRULENCE ET POUVOIR PATHOGENE 29
   - *3.1 Pouvoir pathogène des mycoplasmes 29*
   - *3.2 Facteurs de virulence 30*
     - 3.2.1 Dommages cellulaires directs 31
     - 3.2.2 Effet apoptotique, mitotique, clastogénique et pouvoir oncogène 33
     - 3.2.3 Adhésion et invasion des cellules hôtes 34
     - 3.2.4 Formation de biofilms 36
4. MYCOPLASMES ET MYCOPLASMOSES DE RUMINANTS 41
   - *4.1 La péripneumonie contagieuse bovine 44*
   - *4.2 Mycoplasma bovis 45*
   - *4.3 Mycoplasma leachii 46*
   - *4.4 La pleuropneumonie contagieuse caprine 47*
   - *4.5 Le syndrome MAKePS 47*
5. L'AGALACTIE CONTAGIEUSE A MYCOPLASMA AGALACTIAE 49
   - *5.1 Epidémiologie 50*
     - 5.1.1 Espèces sensibles 50
     - 5.1.2 Répartition géographique 50
     - 5.1.3 Evolution au sein d'un troupeau atteint 52
   - *5.2 Tableau clinique 52*
     - 5.2.1 Formes cliniques habituellement rencontrées 52
     - 5.2.2 Formes atypiques 54
   - *5.3 Diagnostic 55*
     - 5.3.1 Diagnostic direct 55
       - 5.3.1.1 Prélèvement 55
       - 5.3.1.2 Culture et isolement 56
       - 5.3.1.3 Identification 56
     - 5.3.2 Diagnostic indirect 58
   - *5.4 Traitement 59*
   - *5.5 Prophylaxie 60*
     - 5.5.1 Prophylaxie médicale 60
       - 5.5.1.1 Immunoprévention 60
       - 5.5.1.2 Chimioprévention 60
     - 5.5.2 Prophylaxie sanitaire 60
6. OBJECTIFS DE L'ETUDE 63
   - *6.1 Problématique posée par les mycoplasmoses de ruminants 63*
   - *6.2 Objectifs de l'étude 64*
   - *6.3 Démarche adoptée 64*

REFERENCES 66

PREMIER CHAPITRE : DIVERSITE GENETIQUE ET MARQUEURS EPIDEMIOLOGIQUES CHEZ *M. AGALACTIAE* : UNE SEULE SOUCHE CIRCULE DANS LES FOYERS RECURRENTS DU SUD-OUEST DE LA FRANCE DEPUIS PRES DE 30 ANS. 85

RESUME EN FRANÇAIS 87
1. INTRODUCTION 89
2. MATERIEL ET METHODES 94
3. RESULTATS 97
4. DISCUSSION 115
RÉFÉRENCES 120
DONNEES SUPPLEMENTAIRES 125

DEUXIEME CHAPITRE : PLASTICITE, DISTRIBUTION ET EVOLUTION DES GENES DE LA FAMILLE *VPMA*, UN SYSTEME GENETIQUE CONSACRE A LA VARIATION DE SURFACE A HAUTE FREQUENCE CHEZ *MYCOPLASMA AGALACTIAE*. 137

RESUME EN FRANÇAIS 139
ABSTRACT 142
1. INTRODUCTION 143
2. MATERIALS AND METHODS 146
3. RESULTS AND DISCUSSION 149
4. CONCLUSION 160
REFERENCES 162
SUPPLEMENTARY DATA 167
Supplementary tables 169
Supplementary figures 177

TROISIEME CHAPITRE : ANALYSES GENOMIQUES ET PROTEOMIQUES COMPARATIVES DE DEUX SOUCHES DE *MYCOPLASMA AGALACTIAE* : APPROCHE DES MACRO ET MICRO-EVENEMENTS QUI MODELENT LA DIVERSITE DES MYCOPLASMES 185

RESUME EN FRANÇAIS 187
ABSTRACT 190
1. INTRODUCTION 191
2. MATERIAL AND METHODS 194
3. RESULTS 197
4. DISCUSSION 215
REFERENCES 219
SUPPLEMENTARY DATA 223

DISCUSSIONS GENERALES ET PERSPECTIVES 241

REFERENCES 251

***ANNEXES* 257**

LISTES DES PUBLICATIONS ET COMMUNICATIONS 259
ARTICLE PUBLIE 261

TABLE DES ILLUSTRATIONS : FIGURES

Introduction

Figure 1 : Relations phylogénétiques au sein des *Firmicutes.* 17
Figure 2 : Dynamique de surface et hypervariabilité d'expression des lipoprotéines chez les mycoplasmes, l'exemple des Vpma de *M. agalactiae.* 19
Figure 3 : Représentation schématique de la variation de phase par glissement des brins d'ADN (*slipped-strand mispairing*) chez les mycoplasmes. 21
Figure 4 : Représentation schématique des systèmes de variation d'expression à haute-fréquence par réarrangement d'ADN ou conversion génique. 22
Figure 5 : Phylogénie des *Mollicutes.* 39
Figure 6 : Evolution de l'incidence et de la prévalence de l'agalactie contagieuse dans les Pyrénées-Atlantiques de 1992 à 2008. 51

Premier chapitre

Figure 1 : Position des VNTR et des sondes AGA sur le génome de la souche PG2. 92
Figure 2 : Répartition géographique des isolats du panel 1. 99
Figure 3 : Variations de profils observées pour le VNTR 17. 100
Figure 4 : Variations de profils observées pour la sonde AGA-49. 105
Figure 5 : Dendrogramme représentant la diversité génétique de 16 isolats de *M. agalactiae* d'origines variées basé sur la comparaison des profils AGA. 106
Figure 6 : Profils observés pour la sonde AGA-49 sur les isolats des Pyrénées-Atlantiques. 107
Figure 7 : Analyse du locus *vpma* de *M. agalactiae* PG2 clone 55.5 par *Southern blot*. 109
Figure 8 : Exemples de profils A3F de souches des Pyrénées-Atlantiques. 112
Figure 9 : Carte géographique des isolats ayant un répertoire *vpma* particulier. 114
Figure 10 : Comparaison des profils de restriction des loci *vpma* entre les souches PG2 et 5632. 119

Deuxième chapitre

Figure 1 : Comparison of the *M. agalactiae vpma* loci between the type strain PG2 and the strain 5632. 150
Figure 2 : Structural features and comparison of vpma gene products in *M. agalactiae* strain PG2 and 5632. 152
Figure 3 : Comparison of *vpma* loci or their counterparts in *M. agalactiae* strains PG2 55.5, 5632 c1, 13628, 13375 and 4025..... 157
Figure 4 : Scenario of duplication-insertion of 5632 *vpma* genes, from locus I to locus II via a circular intermediate. 159

Troisième chapitre

Figure 1 : Overall comparison of M. agalactiae genomes from the PG2 and 5632 strains. 198
Figure 2 : Comparison of entire and vestigial ICEs found in *M. agalactiae* strains PG2 and 5632. 201
Figure 3 : Location of insertion sequence and their flanking sequences in *M. agalactiae* strain 5632. 202
Figure 4 : Comparison of *M. agalactiae* PG2 and 5632 revealed that the *drp* loci are a reservoir for strain genetic and surface diversity. 211
Figure 5 : The *M. agalactiae spma* repertoire is extended in strain 5632. 213
Figure 6 : Analysis of the *p48-like* sequence of *M. agalactiae* 5632 suggests a mechanism for phase variation. 214

TABLE DES ILLUSTRATIONS : TABLEAUX

Introduction

Tableau 1 : Taxonomie des genres composant la classe des *Mollicutes*. 17
Tableau 2 : Propriétés générales des génomes de *Mollicutes*. 18
Tableau 3 : Principales espèces de mycoplasmes rencontrées chez les bovins.. 42
Tableau 4 : Principales espèces de mycoplasmes rencontrées chez les ovins et caprins. 43
Tableau 5 : Caractéristiques cliniques générales de l'agalactie contagieuse à *M. agalactiae* chez les ovins et les caprins. 53

Premier chapitre

Tableau 1 : Profils obtenus avec les VNTR sur 40 isolats d'origines variées. 98
Tableau 2 : Répartition des isolats entre les différents profils VNTR observés. . 101
Tableau 3 : Profils obtenus avec 12 sondes AGA sur 16 isolats d'origines variées. 104
Tableau 4 : Détermination du répertoire *vpma* d'isolats d'origines diverses. 111

Deuxième chapitre

Table 1 : Analyses of *abiG* related genes products in mollicutes and other bacteria. 155

Troisième chapitre

Table 1 : General properties of M. agalactiae PG2 and 5632 strains. 197
Table 2 : Restriction/Modification products comparison between strains PG2 and 5632. 206
Table 3 : Lipoproteins and MS/MS detection in Tx-114 phase. 208

TABLE DES ILLUSTRATIONS SUPPLEMENTAIRES

Premier chapitre

Tableau S1 : Profils VNTR des isolats des Pyrénées-Atlantiques de 1977 à 2007. 127
Tableau S2 : Hybridation des 59 sondes AGA avec 16 isolats d'origines variées. 129
Tableau S3 : Détermination du répertoire *vpma* d'isolats d'origines géographiques diverses. 130
Tableau S4 : Détermination du répertoire *vpma* des isolats des Pyrénées-Atlantiques de 1977 à 2007 132

Deuxième chapitre

Figure S1: Multiple sequence alignment of *vpma* promoter regions in PG2 and 5632 loci I and II. 179
Figure S2: AbiGI-like and AbiGII-like phylogenetic trees. 180
Figure S3: Imaging of whole sequence comparison between PG2 *vpma* locus and 5632 *vpma* locus I. 181

Table S1 : Features of strains used in this study. 170
Table S2 : Primers used in this study. 172

Troisième chapitre

Table S1 : Products of 5632 for which more than one specific peptide were detected by LC MS/MS after 1D SDS-PAGE. 225

INTRODUCTION

INTRODUCTION

1. *Mycoplasmes : Généralités*

Au milieu du XIX^{e} siècle, les connaissances sur le caractère transmissible et la prémunition de la "pleuropneumonie épizootique du gros bétail" évoluent grâce aux travaux de Louis Willems, médecin et fils d'éleveur bovin en Belgique (Willems, 1852). Il faudra attendre la fin du siècle pour que débutent les travaux pionniers des microbiologistes Edmond Nocard et Emile Roux sur les "bactéries filtrables", agents infectieux auto-réplicatifs aujourd'hui appelés "mycoplasmes". En 1898, avec leurs collaborateurs A. Borrel, A. Salimbeni et Ed. Dujardin-Baumetz, ils réussissent à cultiver le "microbe de la péripneumonie contagieuse bovine" en milieu inerte (Nocard et Roux, 1898). Ce sera le sujet de la thèse de doctorat en médecine d'Edouard Dujardin-Baumetz soutenue en 1900 : "Etude bactériologique d'un microorganisme à la limite de la visibilité, le microbe de la péripneumonie et sa culture" (Dujardin-Beaumetz, 1900, 1906). En 1923, J. Bridré et A. Donatien isolent pour la première fois *Mycoplasma agalactiae* nommé le "microbe de l'agalactie contagieuse" et démontrent son implication étiologique dans la maladie chez le mouton et la chèvre (Bridré et Donatien, 1923, 1925).

A la suite de ces travaux, de nombreux agents similaires seront découverts chez l'Homme et de nombreuses espèces animales et végétales [pour une revue, voir (Razin et Herrmann, 2002)]. En 1956, Edward et Freundt définiront les bases de leur système de classification avec la création du genre *Mycoplasma*. Puis, au fur et à mesure des découvertes, la classe des *Mollicutes* [Edward et Freundt 1967] ou classe des bactéries à peau molle (L. adj. *mollis*, mou ; L. fem. n. *cutis*, peau ; N.L. fem. pl. n. *Mollicutes*) sera crée pour regrouper l'ensemble des espèces bactériennes apparentées aux mycoplasmes (Edward *et al.*, 1967).

La classe des *Mollicutes* regroupe aujourd'hui plus de 200 espèces dont plus de la moitié appartiennent au genre *Mycoplasma* (Euzéby, 2009) [Tableau 1]. Elle représente un groupe d'organismes dérivant de la branche des *Firmicutes* par une

perte de matériel génétique ayant entraîné une réduction génomique massive (Woese, 1987 ; Sirand-Pugnet *et al.*, 2007a) [Figure 1]. Les caractéristiques spécifiques des espèces de *Mollicutes* découlent grandement de cette évolution régressive : ce sont des bactéries de petite taille (0,15 à 0,8 µm) qui possèdent de petits génomes (580 à 1500 kpb pour les génomes séquencés (Barre *et al.*, 2004)) à faible contenu en G+C (24 à 33%) [Tableau 2]. Les *Mollicutes* disposent d'un nombre réduit de voies de biosynthèses. Pour compenser ce défaut de voies métaboliques majeures comme celles du cholestérol et des acides aminés, ils sont souvent tributaires de leur environnement, d'où l'adoption d'un mode de vie habituellement parasitaire. Ces organismes ont également perdu la capacité de synthétiser le peptidoglycane ou ses précurseurs et sont ainsi totalement dépourvus de paroi. Cette dernière caractéristique les rend constitutivement résistant à tous les antibiotiques ayant pour cible la paroi bactérienne. L'ensemble de ces particularités conduit souvent à les considérer comme des organismes minimalistes et dégénérés. Malgré cette apparente simplicité, beaucoup de *Mollicutes* sont capables de survivre et de s'adapter face aux mécanismes de défenses d'hôtes complexes comme l'Homme et de nombreuses espèces animales ou végétales. Plus particulièrement, le genre *Mycoplasma* compte des pathogènes à l'origine de maladies souvent chroniques et débilitantes chez l'Homme et l'animal (Baseman et Tully, 1997 ; Rosengarten *et al.*, 2001). La surface de leur simple membrane plasmique est considérée comme une interface clef dans leur adaptation et leur survie chez un hôte complexe et immunocompétent (Rosengarten et Wise, 1990 ; Citti *et al.*, 1997 ; Simmons *et al.*, 2004 ; Denison *et al.*, 2005). En effet, l'architecture moléculaire de celle-ci est le lieu d'une grande plasticité induite par des systèmes génétiques sophistiqués. Il apparaît donc plus pertinent d'interpréter cette réduction génomique comme une miniaturisation optimisée plutôt que comme une dégénérescence simplificatrice.

Tableau 1 : Taxonomie des genres composant la classe des *Mollicutes*.

Ordre[a]	Famille	Genre	Nombre d'espèces	Hôtes
Mycoplasmatales	*Mycoplasmataceae*	*Mycoplasma*[b]	107 (10)	Animaux, Homme
		Ureaplasma	7	Animaux, Homme
Entomoplasmatales	*Entomoplasmataceae*	*Entomoplasma*	6	Insectes, Plantes
		Mesoplasma	12	Insectes, Plantes
	Spiroplasmataceae	*Spiroplasma*	34	Insectes, Plantes
Acholeplasmatales	*Acholeplasmataceae*	*Acholeplasma*	14	Animaux, Plantes
		"Phytoplasma"[c]	(6)	Insectes, Plantes
Anaeroplasmatales	*Anaeroplasmataceae*	*Anaeroplasma*	4	Bovins, Ovins
		Asteroleplasma	1	Bovins, Ovins

D'après (Tully *et al.*, 1993 ; Euzéby, 2009 ; Johansson et Pettersson, 2002).
[a] un cinquième ordre incluant la famille d'affiliation incertaine des *Erysipelothrichaceae* est parfois inclus dans la classe des *Mollicutes* comme *Incertae sedis.*
[b] Comprend les espèces hémotropes du genre *Mycoplasma* primitivement incluses dans les genres *Eperythrozoon* et *Haemobartonella (Mycoplasma haemocanis, M. haemofelis, M. haemomuris, M. ovis, M. suis, M. wenyonii, "Candidatus Mycoplasma haematoparvum", "Candidatus Mycoplasma haemodidelphis", "Candidatus Mycoplasma haemolamae", "Candidatus Mycoplasma haemominutum")*.
[c] Genre de taxonomie incertaine composé d'espèces non cultivables *in vitro*. Phylogénétiquement proches des Acholeplasmes, et provisoirement classés dans la famille des *Acholeplasmataceae*.

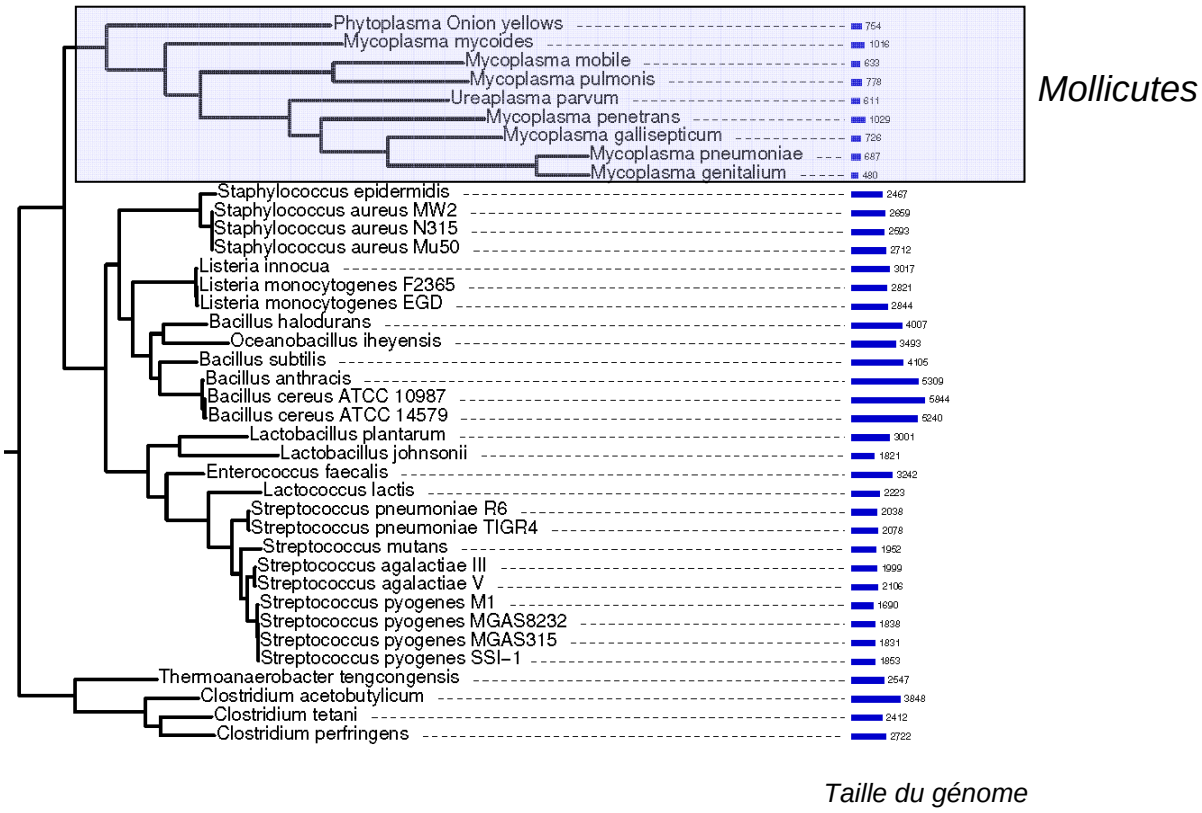


Figure 1 : Relations phylogénétiques au sein des *Firmicutes.*

La taille des génomes est indiquée en kilobases et représentée par une barre de longueur proportionnelle. La branche correspondant à la classe des *Mollicutes* est repérée par un cadre grisé. Cet arbre phylogénétique est basé sur la séquence génomique complète, il a été généré avec iTOL (Letunic et Bork, 2007) : interface interactive accessible à http://itol.embl.de/.

Tableau 2 : Propriétés générales des génomes de *Mollicutes*.

Espèces	Hôtes	Taille du génome (bp)	G+C (%)	Densité de gènes (%)	Nb de CDS	Nb de copie des ARNr	ARNt	n° accession GenBank	Plasmides	ICE[b]	Phages[c]	Transposases[d]
M. mycoides subsp. *mycoides* SC PG1	Ruminant	1 211 703	23,96	81	1016	2	30	BX293980.2	-	-	-	74
M. capricolum subsp. *capricolum* California Kid	Ruminant	1 010 023	23,77	90	827	2	30	CP000123	-	1	-	3
Me. florum L1	Insecte	793 224	27,02	92	683	2	29	AE017263	-	-	-	0
M. arthritidis 158L3-1	Rongeur	820 453	30,70	89	641	1	22	CP001047				
M. mobile 163K	Poisson	777 079	24,95	90	633	1	28	AE017308.1	-	-	-	0
M. hyopneumoniae 232 (pathogenic)	Porc	892 758	28,56	90	691	1	30	AE017332.1	-	-	1	2
M. hyopneumoniae 7448 (pathogenic)	Porc	920 079	28,48	86	663	1	30	AE017244.1	-	1	1	8
M. hyopneumoniae J (non pathogenic)	Porc	897 405	28,52	87	679	1	30	AE017243.1	-	-	1	14
M. pulmonis UAB CTIP	Rongeur	963 879	26,63	90	782	1	29	AL445566.1	-	-	1	8
M. synoviae (Parana)	Oiseau	799 476	28,49	89	694	2+5S	34	AE017245.1	-	-	1	7
***M. agalactiae* PG2**	**Ruminant**	**877 438**	**29,70**	**88**	**752**	**2**	**34**	**CU179680**	**-**	**1(v)**	**-**	**1**
U. urealyticum/parvum serovar 3	Homme	751 719	25,49	91	614	2	30	AF222894.1	-	-	-	3
M. penetrans HF-2	Homme	1 358 633	25,71	88	1037	1	29	BA000026.2	-	-	-	19
M. pneumoniae M129	Homme	816 394	40,00	88	689	1	37	U00089.2	-	-	-	0
M. genitalium G-37	Homme	580 074	31,69	91	484	1	36	L43967.2	-	-	-	0
M. gallisepticum strain R	Oiseau	996 422	31,45	87	726	2	33	AE015450.1	-	-	-	9
A. laidlawii PG8A	Multiples	1 496 992	31,92	90	1380	2	34	CP000896	-			
Ca. Phytoplasma mali AT	Plante	601 943	21,39	76	479	2	32	CU469464	-			
Ca. Phytoplasma australiense	Plante	879 959	27,42	74	839	2	35	AM422018	1			
Ca. Phytoplasma Aster Yellows Witches-Broom strain[a]	Plante/ Insecte	706 569	26,88	73	671	1+23S	31	CP000061.1	4	-	3	18
Ca. Phytoplasma asteris Onion Yellows strain[a]	Plante/ Insecte	860 631	27,73	73	754	2	32	AP006628.1	2	-	7	13

D'après (Sirand-Pugnet *et al.*, 2007a) actualisé en juillet 2009. Données extraites de la base de données MolliGen (Barre *et al.*, 2004) (http://cbi.labri.fr/outils/molligen).

[a] les plasmides ne sont pas pris en compte dans les chiffres rapportés,

[b] Fonctionnelle ou vestigiale (v),

[c] Somme des CDS correspondant au COG5377 “Phage-related protein, predicted endonuclease”,

[d] Somme des COG correspondant à des transposases (COG5421, COG3316, COG4584, COG2801, COG2826, COG3385, COG2963, COG3666, COG3328).

2. *Mycoplasmes et variabilité antigénique*

La surface des *Mollicutes* est d'un genre unique au sein des procaryotes. En l'absence de paroi, elle est l'interface d'interaction directe du mycoplasme avec son hôte, et une des clefs de l'adaptation et de la survie du mycoplasme en milieu hostile. Chez les mycoplasmes et les uréaplasmes, la variabilité de surface constitue une réponse adaptative à ces conditions changeantes voire hostiles. Malgré leur information génétique réduite, les mycoplasmes possèdent un grand nombre de gènes codant des protéines de surface intégrés dans des systèmes génétiques sophistiqués qui, par le biais de mutations ou de recombinaisons de l'ADN, permettent une diversification de l'architecture et de la structure antigénique de la surface cellulaire (Wise, 1993 ; Citti *et al.*, 2005) [Figure 2].

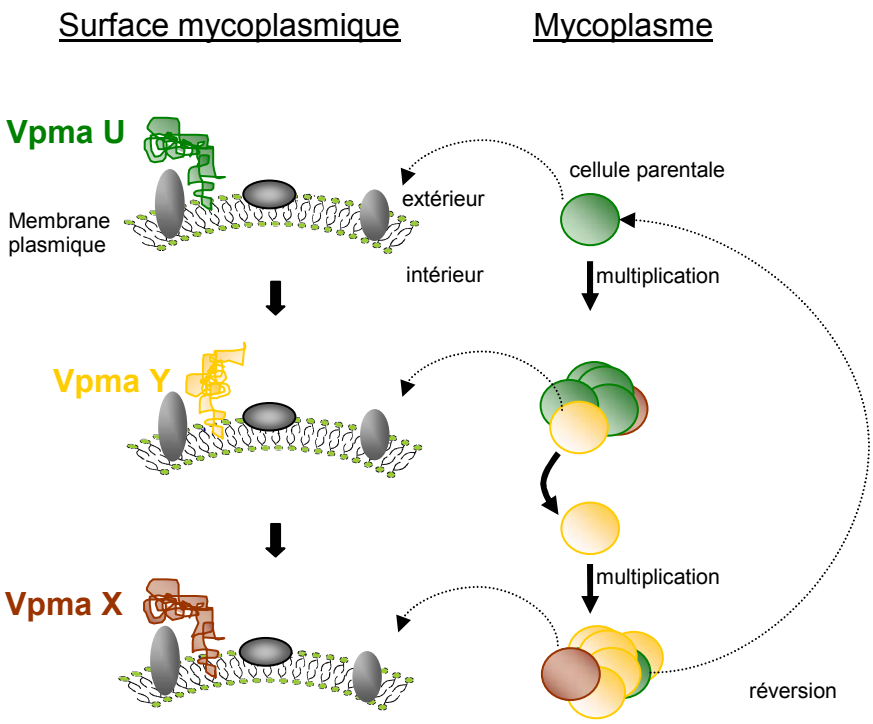


Figure 2 : Dynamique de surface et hypervariabilité d'expression des lipoprotéines chez les mycoplasmes, l'exemple des Vpma de *M. agalactiae*.

Adapté de (Citti, 2006).
Les mycoplasmes sont schématisés, à gauche à l'échelon de leur surface et à droite à celui de la cellule. La membrane plasmique est représentée par sa bicouche lipidique et des protéines membranaires, la lipoprotéine de la famille Vpma exprimée par un clone donné est représentée en vert (Vpma U), en jaune (Vpma Y), en rouge (Vpma X).

Ces systèmes sont habituellement composés d'un groupe de gènes subissant des modifications moléculaires réversibles avec une très haute fréquence (évaluée entre 10^{-2} et 10^{-4} par cellule et par génération *in vitro*) qui entraînent une variation d'expression (ou système ON/OFF ou "tout ou rien") pouvant survenir de manière intraclonale au sein d'une population isogénique. Ces gènes codent souvent des lipoprotéines abondamment exposées à la surface, leur variation d'expression a ainsi pour conséquence une haute variabilité de la structure de la surface cellulaire. La plupart de ces lipoprotéines étant fortement immunogènes, le mycoplasme est vu par l'hôte comme une population exposant des antigènes toujours différents. Plus globalement, la génération de multiples variants phénotypiques au sein d'une population favorise l'émergence de variants mieux adaptés à certaines agressions et changements environnementaux. Selon la fonction des protéines variables, cette capacité permet à la bactérie un échappement de la réponse humorale, mais également une adaptation à de nouvelles niches tissulaires ou écologiques (Citti *et al.*, 2005). Plus rarement, des variations à haute fréquence sont observées pour des gènes codant des protéines cytosoliques : le système Hsd regroupe des enzymes impliquées dans la restriction-modification de l'ADN (Sitaraman et Dybvig, 1997 ; Dybvig *et al.*, 1998), dans cette étude nous nous limiterons à la description de mécanismes liés aux lipoprotéines.

La variation à haute fréquence repose (i) sur des mutations spontanées (généralement mésappariements par glissement des brins d'ADN) [Figure 3], (ii) sur des conversions de gènes, ou (iii) sur des réarrangements spécifiques de l'ADN (système "flip/flop") (Citti *et al.*, 2005) [Figure 4].

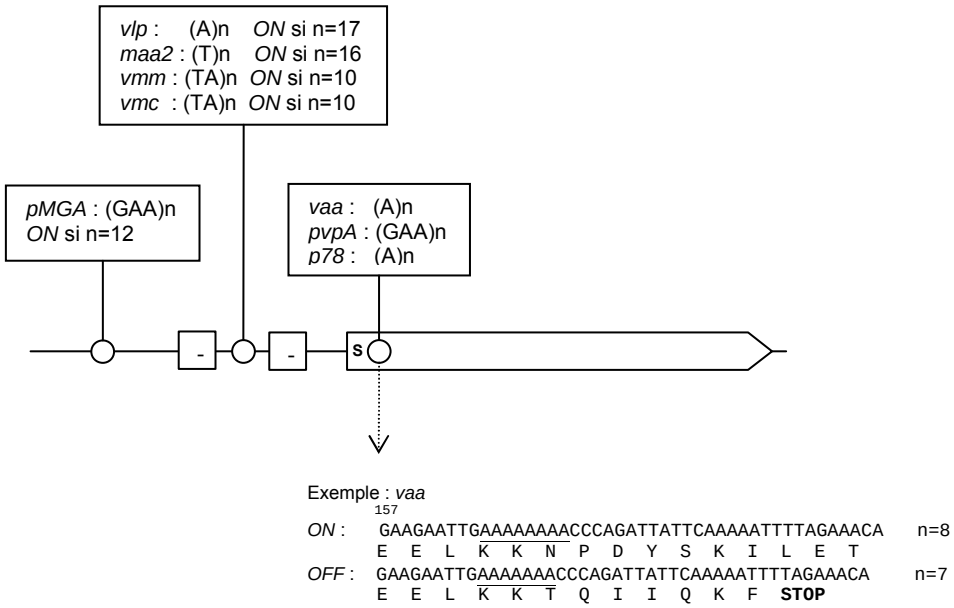
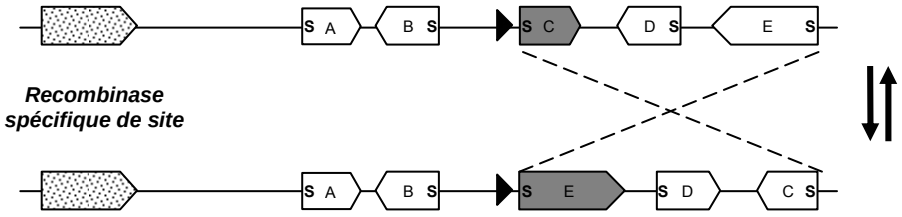


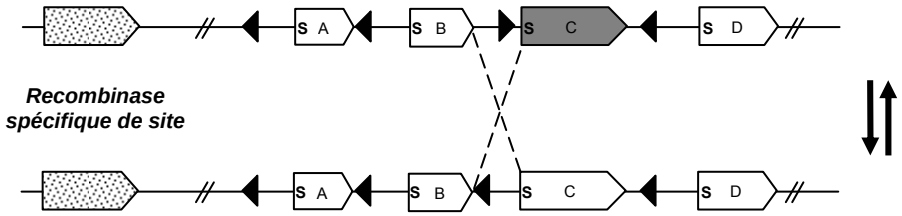
Figure 3 : Représentation schématique de la variation de phase par glissement des brins d'ADN (*slipped-strand mispairing*) chez les mycoplasmes.

D'après (Citti *et al.*, 2005).
Cette figure illustre la variation fréquente et réversible du nombre de répétitions de courts homo- ou hétéro-polymères nucléotidiques liée à la variation d'expression de protéines de surface chez les mycoplasmes.
Les mutations qui surviennent dans la région promotrice (en amont et entre les boites -35 et -10) affectent la transcription du gène situé en aval (flèche large), celles survenant dans la région codante affectent la traduction par l'apparition d'un codon stop causant une interruption prématurée (un exemple est donné pour *vaa*). La lettre ‘s’ symbolise la présence d'une séquence signal.

Famille des gènes *vsa*, *vsp* et *vpma* de *M. pulmonis*, *M. bovis* et *M. agalactiae*



Famille des gènes *mlp* de *M. penetrans*



Famille des gènes *vlh* de *M. synoviae*

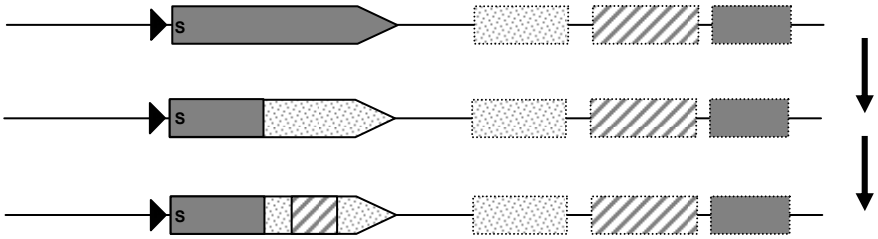


Figure 4 : Représentation schématique des systèmes de variation d'expression à haute-fréquence par réarrangement d'ADN ou conversion génique.

D'après (Citti *et al.*, 2005).
Les séquences promotrices sont représentées par une tête de flèche noire ; les gènes sont représentés par des flèches larges, si un promoteur est juxtaposé en amont, le gène est transcrit (coloration grise) ; la lettre ‘s’ symbolise la présence d'une séquence signal. Pour *M. synoviae,* les rectangles en pointillées représentent les pseudogènes de *vlh*, seul le gène en aval du promoteur est exprimé.

2.1 Mutations spontanées et variabilité de surface

2.1.1 Mésappariements par glissement des brins d'ADN (*slipped-strand mispairing*)

Le premier mécanisme génétique caractérisé comme support de la variation d'expression à haute fréquence des protéines de surface chez les mycoplasmes est le système *vlp* de *M. hyorhinis* (Rosengarten et Wise, 1990, 1991). Ce système est composé d'un ensemble de 3 à 7 gènes codant des lipoprotéines variables. Chacun de ces gènes *vlp* est constitué en unité transcriptionnelle indépendante. Leur transcription est contrôlée par la variation de taille d'un homopolymère nucléotidique (polyA) située entre les boîtes -35 et -10 du promoteur de chaque gène. La synthèse de l'ARN messager d'un gène *vlp* donné est uniquement permise quand la taille du polyA est de 17 adénines, toute insertion ou délétion d'un seul résidu nucléotidique empêche la transcription du gène (Citti et Wise, 1995). Un mécanisme similaire est également rencontré pour le gène *maa2* de *M. arthritidis*, les gènes *vmm* de *M. mycoides* subsp. m*ycoides* SC (Persson *et al.*, 2002 ; Hamsten *et al.*, 2008), les gènes *vmc* de *M. capricolum* subsp. *capricolum* (Wise *et al.*, 2006), les gènes *vlhA* (ou *pMGA*) de *M. gallisepticum* (Glew *et al.*, 1998) et probablement ceux de *M. imitans* (Markham *et al.*, 1999 ; Jan *et al.*, 2001) [Figure 3].

La variation fréquente et réversible du nombre de nucléotides composant de courts homo- ou hétéro-polymères placés cette fois dans la région codante peut également être une source de variabilité antigénique. Chez *M. hominis* (gène *vaa*), *M. fermentans* (gène *p78*), la variation de taille de polyA situés proche de la partie du gène codant l'extrémité N-terminale de la protéine mature entraîne un décalage réversible du cadre de lecture qui peut conduire à l'apparition d'un codon stop et à un arrêt prématuré de la traduction (Zhang et Wise, 1996 ; Theiss et Wise, 1997) [Figure 3]. Le séquençage du génome de *M. pulmonis* (Chambaud *et al.*, 2001), a mis en évidence des polymères nucléotidiques à cette position stratégique dans de nombreuses lipoprotéines, et de même lors de l'analyse des séquences de bien d'autres génomes de mycoplasmes (Westberg *et al.*, 2004 ; Wise *et al.*, 2006). Leur variation via des mésappariements par glissement des brins d'ADN et

leur implication concrète dans la génération de diversité de surface *in vivo* restent toutefois à démontrer.

Si la variation fréquente et réversible du nombre de répétitions composant des polymères est à l'origine de variation d'expression de type "tout ou rien", elle est également à l'origine de diversité structurale des protéines de surface par la génération de variants de taille (Behrens *et al.*, 1994 ; Citti et Wise, 1995 ; Zhang et Wise, 1996 ; Lysnyansky *et al.*, 1999). Ainsi, en plus de la variation de phase, une variation de taille est observée pour un bon nombre de lipoprotéines présentant des blocs d'acides aminés directement répétés. Sans affecter la synthèse de la protéine, cette variation du nombre de blocs répétés entraîne une variation structurale de celle-ci.

2.1.2 Mutations spontanées par substitution

Des gènes d'expression variable présents en simple copie ont été identifiés chez *M. gallisepticum* : *pvpA* (Yogev *et al.*, 1994 ; Boguslavsky *et al.*, 2000) et *gapA (*ou *mgc1)* (Winner *et al.*, 2003) et chez *M. mycoides* subsp. *mycoides* biotype Small Colony : *ptsG* (Gaurivaud *et al.*, 2004). La variabilité d'expression de ces gènes est apparemment dépendante de la présence réversible d'une mutation non-sens interrompant prématurément la phase ouverte de lecture du gène. Cette mutation est due à une substitution de base. C'est la transition réversible d'un C par un T pour *gapA* ou d'un G par un A pour *ptsG* qui entraîne l'apparition d'un codon stop. Aucune des séquences environnantes ne suggère l'intervention de mécanismes connus tels que le glissement des brins d'ADN ou la recombinaison par reconnaissance de séquences spécifiques. De même pour le gène *pvpA,* bien que la mutation se fasse au sein d'un pentamère de GAA, il ne s'agit pas d'une variation du nombre de GAA répétés mais d'une substitution ayant toujours été observée sur le quatrième GAA (remplacé par TAA). Les mécanismes à l'origine de ces mutations apparemment ciblées restent à ce jour inconnus.

2.2 Conversions de gènes et dynamique de surface

La variabilité d'expression des gènes *vlhA* (ou *pMGA*) de *M. gallisepticum* (Glew *et al.*, 1998) et *M. imitans* (Markham *et al.*, 1999 ; Jan *et al.*, 2001) repose sur la variation de taille d'un polymère nucléotidique (GAA_n) dans la région promotrice et de multiples unités transcriptionnelles de gènes *vlhA* sont retrouvées sur le génome. Chez *M. synoviae*, le gène *vlhA* code une hémaglutinine présentant des homologies avec les gènes *vlhA* de *M. gallisepticum* et *M. imitans* (Noormohammadi *et al.*, 1998 ; Markham *et al.*, 1999) mais le mécanisme de contrôle de l'expression de ces protéines diffère. *M. synoviae* possède un seul gène *vlhA* placé devant une unique copie de la région promotrice et de la région codante 5' conservée [Figure 4]. Lui seul est exprimé, d'autres séquences codantes sont sous la forme de multiples pseudogènes répétés en tandem et répartis sur l'ensemble du génome (Allen *et al.*, 2005 ; Vasconcelos *et al.*, 2005 ; Jeffery *et al.*, 2006). L'étude de séries de clones de *M. synoviae* a permis de mettre en évidence des recombinaisons unidirectionnelles entre ces pseudogènes et le gene *vlhA* fonctionnel : la séquence des pseudogènes est dupliquée pour remplacer la région correspondante au locus fonctionnel (Noormohammadi *et al.*, 2000). La variation du gène *vlhA* chez *M. synoviae* est donc contrôlée par des évènements multiples de conversions géniques avec un répertoire de séquences codantes. Le gène exprimé est une chimère issue de ce répertoire, probablement par recombinaison spécifique (Noormohammadi *et al.*, 2000).

Chez *M. bovis*, les lipoprotéines de la famille Vsp sont aussi une source de variation antigénique (Behrens *et al.*, 1994 ; Lysnyansky *et al.*, 1996). Comme pour *M. synoviae*, une seule Vsp est exprimée à la fois par un clone donné car la région promotrice est en une seule copie. En revanche, les différents gènes *vsp* ne sont pas des pseudogènes et contiennent chacun la région codante conservée en 5' contenant la séquence du futur peptide signal. Ils sont de plus regroupés en *cluster* à un endroit du génome : le locus *vsp*. La comparaison des loci de la souche type PG45 et de l'isolat 422 (Nussbaum *et al.*, 2002) met en évidence des séquences répétées spécifiques et des répétitions partagées identiques entre les *vsp* des deux loci. Cela a suggéré une hypothèse de variation par conversion

génique pour le système *vsp* mais il a été démontré qu'il s'agit en fait d'un mécanisme de recombinaison. En effet, l'observation de différentes lignées clonales de *M. bovis* a démontré l'existence de gènes *vsp* qui résulteraient de la recombinaison chimérique entre les régions 3' de différents gènes *vsp* (Lysnyansky *et al.*, 2001a). Cette stratégie pour générer de la diversité antigénique en créant de nouvelles *vsp* chimères du répertoire existant n'est cependant qu'accessoire car elle ne contrôle pas directement l'expression des *vsp.* En effet, cette dernière est régie par un mécanisme de recombinaisons spécifiques ou système "flip/flop" survenant à haute fréquence à l'aide d'une intégrase-recombinase (Lysnyansky *et al.*, 2001b).

2.3 Réarrangements spécifiques de l'ADN (système "flip/flop")

Ces systèmes peuvent être classés en deux grands types selon que les réarrangements intéressent la région promotrice ou la région codante des gènes d'expression variable. La variation de phase repose ainsi, soit sur des interversions de promoteurs devant un ou deux loci codants en orientation opposée, soit sur des réarrangements de gènes silencieux ou de pseudogènes devant un unique locus d'expression. La principale différence entre ces deux modalités est la variation systématique de l'expression à chaque réarrangement ou l'existence de réarrangements phénotypiquement silencieux [Figure 4].

2.3.1 Interversions des régions promotrices

Décrite chez le pathogène humain *M. penetrans*, la famille des gènes *mpl* est un système qui regroupe pas moins de trente-huit gènes répartis en trois *clusters* sur le chromosome (Neyrolles *et al.*, 1999 ; Roske *et al.*, 2001 ; Sasaki *et al.*, 2002 ; Horino *et al.*, 2003). Chacun de ces gènes possède son propre promoteur intervertible (Horino *et al.*, 2003). Chaque gène *mpl* peut être basculé indépendamment en position "ON" ou "OFF" (Sasaki *et al.*, 2002). Ces réarrangements de l'ADN se font entre des séquences inverseés répétées de douze nucléotides reconnues par une tyrosine-recombinase spécifique de site

(Horino *et al.*, 2009). La combinaison des protéines Mpl exprimées à la surface est alors une configuration parmi 2^{38} possibles soit plus de deux cent soixante-dix milliards de possibilités.

Ces systèmes de variation de phase par réarrangements de l'ADN sont rencontrés chez d'autres bactéries (van der Woude et Baumler, 2004 ; van der Woude, 2006). En abondants utilisateurs de la variabilité des composants de surface, les mycoplasmes ne sont pas en reste, certaines espèces sont même pourvues de systèmes d'un genre unique organisés en *cluster* où surviennent des réarrangements spécifiques de gènes silencieux ou de pseudogènes devant un unique site d'expression (Bhugra *et al.*, 1995 ; Shen *et al.*, 2000 ; Lysnyansky *et al.*, 2001b ; Glew *et al.*, 2002 ; Sitaraman *et al.*, 2002 ; Flitman-Tene *et al.*, 2003 ; Chopra-Dewasthaly *et al.*, 2008).

2.3.2 Réarrangements devant un unique locus d'expression

La famille des gènes *vsa* du pathogène murin *M. pulmonis* (Bhugra *et al.*, 1995 ; Shen *et al.*, 2000), des gènes *vsp* du pathogène bovin *M. bovis* (Behrens *et al.*, 1994 ; Lysnyansky *et al.*, 1996) et des gènes *vpma* du pathogène ovin/caprin *M. agalactiae* génèrent de la diversité de surface intraclonale par un mécanisme similaire (Lysnyansky *et al.*, 2001b). Dans les trois systèmes, un gène codant une tyrosine-recombinase est retrouvé à proximité d'un *cluster* de gènes *vsa*, *vsp* ou *vpma* exprimables (Ron *et al.*, 2002) [Figure 4]. Réalisée chez *M. pulmonis* et *M. agalactiae*, l'inactivation de ce gène par transposition donne des mutants figés dans une configuration (Sitaraman *et al.*, 2002 ; Chopra-Dewasthaly *et al.*, 2008). Cette tyrosine-recombinase dirige des inversions des brins d'ADN entre des sites de recombinaisons spécifiques. Elle entraîne ainsi une variation d'expression par juxtaposition devant une région promotrice unique d'un gène exprimé parmi un pool de gènes silencieux [Figure 4].

Si les mécanismes rencontrés sont très similaires, la nature des protéines soumises à variation est spécifique à chacune de ces trois espèces de mycoplasme. Il est alors tentant de voir dans ce singularisme un lien potentiel avec la spécificité d'hôte. Ces systèmes auraient été hérités d'un ancêtre

commun. Au fil de l'évolution, le mécanisme aurait été conservé alors que les protéines exprimées auraient divergé pour une meilleure adaptation à leur hôte (voir aussi le deuxième chapitre de ce mémoire).

3. *Mycoplasmes, facteurs de virulence et pouvoir pathogène*

3.1 Pouvoir pathogène des mycoplasmes

Beaucoup de mycoplasmes ont été isolés chez l'Homme et l'animal. Toutefois pour bon nombre d'entre eux, leur implication directe dans la causalité d'une maladie reste à démontrer et ils sont souvent considérés comme des bactéries opportunistes ou peu pathogènes. Outre l'existence d'espèces commensales (*M. salivarium*, *M. orale*, *M. felis*, M*. bovirhinis*…), l'ambiguïté du pouvoir pathogène de nombreux mycoplasmes est probablement liée à la nature même des infections mycoplasmiques dont la symptomatologie est polymorphe et inconstante. Cependant, les maladies générées par certains donnent lieu à de graves problèmes de santé publique et leurs impacts économiques en santé animale sont loin d'être négligeables. Il s'agit généralement de maladies chroniques à faible mortalité malgré une forte morbidité. L'amélioration des techniques diagnostiques a conduit à la détection de nouvelles espèces de mycoplasmes dont des pathogènes jusqu'alors inconnus ou responsables de manifestations cliniques atypiques, moins classiques que les simples pneumonies, arthrites et mammites (*M. alligatoris*, *M. agassizii*, *M. californicum*, *M. edwardii*, "*Candidatus* Mycoplasma ravipulmonis"…) (Baseman et Tully, 1997 ; Rosengarten *et al.*, 2001). Ainsi, plusieurs espèces de mycoplasmes dont la liste ne cesse d'augmenter (*M. pneumoniae*, *M. hominis*, *M. genitalium*, *M. gallisepticum*, *M. mycoides* subsp. *mycoides* SC, *M. capricolum* subsp. *capripneumoniae*, *M. bovis*, *M. agalactiae*…) sont aujourd'hui considérées comme d'importants pathogènes justifiant la mobilisation de moyens pour améliorer leur contrôle, préventif ou curatif (Rosengarten *et al.*, 2001 ; Blanchard et Bébéar, 2002 ; Frey, 2002). Une solide connaissance des bases moléculaires des infections à mycoplasmes est un pré-requis indispensable à ce progrès.

Au cours des dernières années, grâce aux apports de la biologie moléculaire, d'importants travaux ont été réalisés pour mieux comprendre les interactions entre

le mycoplasme et son hôte. De nouveaux concepts importants au sujet des maladies à mycoplasmes ont été découverts et des outils de génomique fonctionnelle ont été adaptés avec succès, néanmoins beaucoup reste à faire. Les mycoplasmes sont apparus bien plus complexes que leurs petits génomes pouvaient le laisser penser. Malgré la somme d'information génétique accumulée avec le séquençage d'un nombre de plus en plus grand de génomes de mycoplasmes, à quelques rares exceptions, les facteurs moléculaires de leur pouvoir pathogène semblent spécifiques. Présentant peu ou pas de similitude avec ceux déjà décrits chez les autres bactéries, ils restent encore largement inconnus. A l'évidence, au cours de leur évolution régressive, les mycoplasmes ont développés de nombreux systèmes pour compenser leurs faibles capacités métaboliques et s'adapter aux conditions environnementales. Ces systèmes leur confèrent effectivement une virulence et un pouvoir pathogène, incluant la capacité d'attachement, de causer des dommages cellulaires, d'exhiber un phénotype hypervariable, de moduler et échapper à la réponse immunitaire de l'hôte.

3.2 Facteurs de virulence

Les infections mycoplasmiques et leur pathogénie sont des processus complexes qui mettent en jeu de multiples acteurs : le mycoplasme, l'hôte, l'environnement et la présence d'autres microorganismes. Bien qu'un certain nombre de facteurs de virulence aient été suggérés pour plusieurs mycoplasmes, il n'y a pas de relation de cause à effet claire entre ces facteurs et la pathogénie. Les difficultés rencontrées dans le développement de techniques permettant la transformation et la modification génétique de ces organismes *in vitro* pour la génomique fonctionnelle a sûrement freiné l'acquisition de connaissances sur ces déterminants. Cependant, il existe de nombreuses capacités des mycoplasmes qui ont vraisemblablement un lien avec leur pathogénie comme la capacité d'attachement, d'invasion, de créer des dommages cellulaires, de faire varier son phénotype à haute fréquence et de résister ou échapper à la réponse immunitaire de l'hôte. Les études réalisées ont surtout porté sur des lipoprotéines comme les

adhésines et les protéines impliquées dans la variation antigénique qui semblent représenter les facteurs de virulence majeurs des mycoplasmes.

3.2.1 Dommages cellulaires directs

Pendant longtemps, aucune toxine n'avait été caractérisée chez les mycoplasmes, à l'exception d'une toxine inflammatoire chez *M. bovis* (Geary *et al.*, 1981). En l'absence de produit présentant des homologies avec des toxines connues, il paraissait établi que les dommages cellulaires liés à l'infection et les manifestations cliniques qui en découlent, étaient principalement dus à l'exacerbation de la réponse inflammatoire et immunitaire de l'hôte plutôt qu'à un effet toxique direct de composants mycoplasmiques (Cole *et al.*, 1996 ; Rottem et Naot, 1998). Si cette exacerbation de la réponse de l'hôte doit effectivement grandement participer aux dommages cellulaires, des éléments nouveaux ont été récemment apportés.

En effet, des agents de nécrose digérant les tissus ont été mis en évidence (Brown *et al.*, 2004 ; Hunt et Brown, 2007) chez *M. alligatoris*, mycoplasme atypique par le caractère aigu de l'infection qu'il provoque, pouvant entraîner la mort d'un alligator en moins d'une semaine (Clippinger *et al.*, 2000 ; Brown *et al.*, 2001a ; Brown *et al.*, 2001b). Il s'agirait d'une hyaluronidase (NagH) associée à une sialidase (NanI) présentant toute deux des homologies avec des protéines de *Clostridium perfringens*. Si ces enzymes sont sécrétées chez *C. perfringens*, aucune activité hyaluronidase ou sialidase n'est détectée pour les surnageants de culture de *M. alligatoris*, ces enzymes seraient donc cytosoliques chez ce mycoplasme (Brown *et al.*, 2004). *M. crocodyli*, mycoplasme isolé d'alligators comme *M. alligatoris*, présente également une hyaluronidase active non sécrétée, cependant il est dépourvu de sialidase fonctionnelle (Brown *et al.*, 2004). Cette différence pourrait expliquer le plus faible pouvoir pathogène observé pour ce mycoplasme en comparaison à *M. alligatoris*. L'observation d'effets protecteurs d'un inhibiteur des sialidases (acide 2,3-didehydro-2-déoxy-N-acetylneuraminique) vis-à-vis de l'apoptose induite par *M. alligatoris* sur des fibroblastes pulmonaires en co-culture est en faveur de cette hypothèse (Hunt et Brown, 2007). Une

interaction entre hyaluronidase et sialidase serait ainsi nécessaire à l'expression du pouvoir pathogène de *M. alligatoris*.

Plus récemment, un facteur de virulence nommé CARDS TX, pour *community-acquired respiratory distress syndrom toxin*, qui présente des similitudes avec des ADP-ribosyltransférases comme la toxine PTX de *B. pertussis* (sous-unité S1) a été étudié chez *M. pneumoniae* par Kannan *et al.* (Kannan *et al.*, 2005 ; Kannan et Baseman, 2006) et chez *M. penetrans* (Johnson *et al.*, 2009). Les ADP-ribosyltransférases génèrent des modifications protéiques post-traductionnelles et sont suspectées comme mécanisme majeur de l'action de beaucoup d'exotoxines bactériennes (Krueger et Barbieri, 1995). Cette activité potentielle du produit du gène correspondant à CARDS TX avait déjà été repérée *in silico* par Pallen *et al.* en 2001 (Pallen *et al.*, 2001). En 2006, grâce à la production d'une protéine CARDS TX recombinante chez *E. coli*, Kannan et Baseman (Kannan et Baseman, 2006) observent une activité ADP-ribosyltransférase *in vitro* combinée à une vacuolisation et une activité cytotoxique thermosensible en culture cellulaire. La présence d'anticorps dirigés contre CARDS TX dans les serums de patients infectés est en faveur d'une expression *in vivo*. Non détectée dans le surnageant de culture et retrouvée dans les fractions membranaire et cytosolique, CARDS TX ne semble pas sécrétée dans le milieu extérieur, l'hypothèse envisagée serait une action de contact à l'interface membrane-membrane (Kannan et Baseman, 2006). Quoi qu'il en soit, CARDS TX est à ce jour l'un des rares produits mycoplasmiques présentant des homologies avec des toxines bactériennes et la première toxine associant une activité ADP-ribosyltransférase à un effet de vacuolisation des cellules chez les bactéries (Kannan et Baseman, 2006).

Des produits du métabolisme comme des radicaux péroxydes et superoxydes ont également été trouvés comme causes de dommages cellulaires. Un exemple bien caractérisé, en lien avec la métabolisation du glycérol, est le système GlpO de *M. mycoides* subsp. *mycoides* SC. Les observations faites chez ce mycoplasme indiquent que deux gènes d'un transporteur du glycerol (GtsABC) présents chez des souches africaines ne sont pas retrouvés chez les souches européennes considérées comme moins virulentes. L'incapacité de ces souches à assimiler le glycérol pourrait être à l'origine de cette atténuation de virulence (Vilei et Frey,

2001). Une production d'H_2O_2 et une cytotoxicité réduites sont observées pour ces souches européennes déficientes en GtsABC par comparaison aux souches africaines porteuses de GtsABC fonctionnel (Vilei et Frey, 2001 ; Pilo *et al.*, 2005 ; Bischof *et al.*, 2008). L'assimilation du glycérol conduirait à la production de peroxyde d'hydrogène (H_2O_2) par le biais de l'oxydation du glycérol-3-phosphate (G3P) en dihydroxyacétone phosphate (DHAP). La (L-alpha-)glycérophosphate oxydase (GlpO) qui catalyserait cette oxydation productrice d'H_2O_2 a été identifiée comme une protéine membranaire impliquée dans la cytotoxicité de *Mycoplasma mycoides* subsp. *mycoides* SC vis-à-vis des cellules d'épithélium nasal de veau fœtal (Pilo *et al.*, 2005). Cette hypothèse est confortée par l'inhibition de cette cytotoxicité et de la production d'H_2O_2 par des anti-corps anti-GlpO (Pilo *et al.*, 2005 ; Bischof *et al.*, 2009). La phosphorylation du glycérol en G3P nécessaire à sa transformation en DHAP par GlpO serait *a priori* principalement réalisée au cours de son transport par GtsABC et des éléments montrent que l'effet cytotoxique d'H_2O_2 nécessite un contact étroit du mycoplasme avec la cellule eucaryote pour que l'H_2O_2 rentre dans le cytoplasme de cette dernière (Pilo *et al.*, 2005).

3.2.2 Effet apoptotique, mitotique, clastogénique et pouvoir oncogène

Des études *in vitro* ont démontré que des mycoplasmes pouvaient perturber le cycle cellulaire que ce soit dans le sens d'une exacerbation ou d'une inhibition. Selon les conditions expérimentales, il a ainsi été remarqué à la fois des effets inducteurs d'apoptose chez *M. bovis* (Sokolova *et al.*, 1998 ; Vanden Bush et Rosenbusch, 2002), *M. hyorhinis* (Paddenberg *et al.*, 1998), *M. penetrans* (Bendjennat *et al.*, 1999), *M. alligatoris* (Hunt et Brown, 2005, 2007), *M. mycoides* subsp. *mycoides* SC (Dedieu *et al.*, 2005) et des effets anti-apoptotiques via l'induction de NfkB chez *M. fermentans* et *M. penetrans* (Feng *et al.*, 1999) et l'inhibition de caspases chez *M. fermentans* (Gerlic *et al.*, 2004 ; Gerlic *et al.*, 2007). Des effets mitotiques ont été observés chez *M. arginini*, *M. arthritidis* (Stuart *et al.*, 1990), *M. fermentans* et *M. penetrans* (Feng et Lo, 1994).

De plus, certaines infections à mycoplasme peuvent entraîner des aberrations chromosomiques et une instabilité génétique (Fogh et Fogh, 1965). Ce phénomène a été observé avec *M. orale* (Paton *et al.*, 1965 ; Stanbridge *et al.*, 1969), *M. salivarium* (Aula et Nichols, 1967), *M. hominis* et *M. fermentans* (Aula et Nichols, 1967 ; Stanbridge *et al.*, 1969). Chez ce dernier, les effets clastogéniques sont apparemment irréversibles après une infection chronique des cellules en culture (Tsai *et al.*, 1995 ; Feng *et al.*, 1999). Le développement de telles instabilités génétiques et modifications chromosomiques sur des cellules ayant perdu le contrôle apoptotique peut conduire à des néoformations malignes (Macpherson et Russell, 1966 ; Tsai *et al.*, 1995). Cependant il s'agit là uniquement de résultats *in vitro* et la mise en cause de mycoplasmes dans un processus cancéreux *in vivo*, bien que parfois suspectée dans les années soixante (Negroni, 1964 ; Grace *et al.*, 1965 ; Hayflick, 1965 ; Murphy *et al.*, 1965 ; Boden *et al.*, 1983 ; Abo *et al.*, 1988 ; Park *et al.*, 1990 ; Jansson *et al.*, 1991), n'a jamais été clairement démontrée. L'implication des mycoplasmes dans un quelconque cancer reste donc à ce jour uniquement hypothétique (Cimolai, 2001). Toutefois certaines espèces comme *M. penetrans* et *M. fermentans* peuvent constituer un excellent modèle d'oncogénèse pour l'étude des transformations cancéreuses des cellules et des tissus (Tsai *et al.*, 1995 ; Lo, 2002).

3.2.3 Adhésion et invasion des cellules hôtes

Outre la spécificité d'hôte, les mycoplasmes montrent généralement un tropisme tissulaire préférentiel avec une prédilection pour les muqueuses génitales et respiratoires, la glande mammaire, les articulations et les séreuses [pour une revue voir (Citti, 2006)]. La plupart des mycoplasmes adhèrent fermement aux revêtements épithéliaux des voies génitales ou respiratoires et envahissent rarement les tissus. L'adhésion des mycoplasmes à la cellule est une étape indispensable à la colonisation afin d'assurer le succès de l'infection et la survenue de la maladie. En effet, il a été suggéré notamment pour *M. pneumoniae* et *M. gallisepticum* que des mutants déficients pour la cyto-adhésion sont avirulents (Baseman *et al.*, 1982 ; Krause *et al.*, 1982 ; Papazisi *et al.*, 2002). Les composants impliqués dans l'adhésion ont été plus particulièrement étudiés chez *M. pneumoniae et M. genitalium* qui possèdent une structure ou organelle en

pointe appelée "tip" où se concentrent les protéines impliquées dans l'adhésion. Toutes les espèces de mycoplasmes ne possèdent pas de structure "tip" mais sont également capables d'adhésion cellulaire. La cyto-adhésion apparaît comme un processus multifactoriel qui met en jeu différents éléments. Les récepteurs cellulaires identifiés à ce jour comme utilisés par les mycoplasmes pour leur attachement sont des sialoglycoconjugués et des glycolipides sulfatés (Razin, 1985 ; Razin et Jacobs, 1992 ; Zhang *et al.*, 1994). Certaines protéines hypervariables agissent comme des adhésines. La haute variabilité des domaines de liaison aux récepteurs donne d'avantage de chances d'adhérer à différents tissus via des sites de liaison aux récepteurs spécifiques et d'être ainsi adapté à chaque étape de l'infection ou permettre la colonisation de nouvelles niches (Baseman et Tully, 1997 ; Rottem, 2003).

Le dogme des mycoplasmes confinés à la surface des cellules épithéliales a été bousculé par la découverte de la capacité de certains mycoplasmes à entrer dans des cellules hôtes non phagocytaires (Taylor-Robinson *et al.*, 1991 ; Mernaugh *et al.*, 1993 ; Taylor-Robinson *et al.*, 1993 ; Baseman *et al.*, 1995). Plusieurs approches expérimentales ont permis de confirmer cette découverte soit par des techniques d'imageries comme la microscopie électronique combinée au marquage immunologique (Taylor-Robinson *et al.*, 1991 ; Taylor-Robinson *et al.*, 1993 ; Andreev *et al.*, 1995) ou la microscopie confocale (Baseman *et al.*, 1995 ; Winner *et al.*, 2000) soit en culture cellulaire par des essais de résistance à la gentamicine, antibiotique à localisation exclusivement extracellulaire (Andreev *et al.*, 1995 ; Winner *et al.*, 2000).

Cette capacité d'invasion et de survie à l'intérieur des cellules hôtes a été très étudiée chez *M. penetrans*, mycoplasme isolé des voies génitales de patients atteints du SIDA (Lo *et al.*, 1993 ; Andreev *et al.*, 1995). Ce mycoplasme a des propriétés invasives et se localise dans le cytoplasme et la région périnucléaire (Baseman *et al.*, 1995). Cette capacité à résider dans les cellules non phagocytaires dans certaines conditions a également été mise en évidence pour d'autres mycoplasmes comme *M. fermentans*, *M. hominis*, *M. pneumoniae*, *M. genitalium* et *M. gallisepticum* (Citti, 2006).

L'invasion cellulaire est un processus complexe qui implique habituellement une interaction spécifique entre une adhésine ou invasine bactérienne et un récepteur cellulaire. Il semble que les protéines de surface impliquées dans l'adhésion aient une implication dans l'invasion. Cependant, la capacité d'adhésion n'est pas suffisante pour conférer une capacité d'invasion, et les signaux générés par cette dernière interaction restent à étudier (Rottem, 2003). Par analogie aux observations faites chez *Neisseria gonorrhoeae* (van Putten *et al.*, 1998) et *Staphylococcus aureus* (Dziewanowska *et al.*, 1999 ; Dziewanowska *et al.*, 2000) où la fibronectine jouerait un rôle dans l'internalisation par la formation de ponts moléculaires entre des protéines bactériennes liant la fibronectine (FnBP) et des récepteurs cellulaires, l'intervention de FnBP a été suggérée dans le processus d'internalisation chez *M. penetrans* (Giron *et al.*, 1996) et *M. pneumoniae* (Dallo *et al.*, 2002). D'autres facteurs comme la liaison et l'activation du plasminogène, une glycoprotéine plasmatique interagissant avec la surface des cellules eucaryotes, pourraient promouvoir l'invasion de cellules HeLa chez *M. fermentans* (Yavlovich *et al.*, 2001).

L'invasion est inhibée par le taxol ou la vinblastine chez *M. penetrans* (Borovsky *et al.*, 1998) et par le nocodazole chez *M. gallisepticum* (Winner *et al.*, 2000). L'inhibition du processus par ces drogues qui affectent les microtubules suggère l'accompagnement de ces internalisations bactériennes par une déviation du cytosquelette de la cellule eucaryote. La dépolymérisation des microfilaments par la cytochalasine D inhibe l'invasion de *M. penetrans* mais pas celle de *M. gallisepticum*. Il semble donc que les mycoplasmes disposent de différentes stratégies de subversion du cytosquelette pour entrer à l'intérieur des cellules eucaryotes (Winner *et al.*, 2000).

3.2.4 Formation de biofilms

Dans leur environnement naturel, les bactéries se multiplient individuellement, comme libres et baignées dans le milieu extérieur de manière planctonique, ou groupées et inclues dans une matrice extra-cellulaire organisée appelée biofilm (Costerton *et al.*, 1978). Synthétisé par les bactéries, ce biofilm est habituellement constitué de polymères polysaccharidiques (Branda *et al.*, 2005). Il est capable

d'adhérer aux tissus de l'hôte ou à des surfaces inertes telles que des canalisations d'eau ou du matériel médical comme les cathéters de perfusions (Raad *et al.*, 1993 ; Costerton *et al.*, 1995 ; Costerton *et al.*, 1999). Son implication dans des infections nosocomiales a stimulé l'approfondissement des propriétés et des mécanismes de cette différenciation en biofilm (Donlan et Costerton, 2002). Ils ont été incriminés dans des infections chroniques et persistantes, notamment celles ne rétrocédant pas aux traitements antibiotiques malgré une sensibilité bactérienne avérée par antibiogramme (Costerton *et al.*, 1999 ; Parsek et Singh, 2003). Il a été montré que les biofilms constituent une barrière de protection qui protège les bactéries qu'ils renferment en limitant par exemple la diffusion des antibiotiques et l'action des désinfectants. Les biofilms sont résistants à de nombreuses forces physiques (flux sanguin, appareil mucociliaire...), ils permettent aux bactéries de survivre à des agressions extérieures (variations de pH, résistance à la phagocytose, privation de nutriments...) (Donlan et Costerton, 2002).

Récemment, il a été démontré pour *M. bovis*, *M. agalactiae*, *M. putrefaciens*, *M. cottewii*, *M. yeatsii*, *M. mycoides* subsp. *mycoides* biotype Small Colony et *M. pulmonis* que certaines souches sont capables de former des biofilms sur des surfaces inertes (McAuliffe *et al.*, 2006 ; Simmons *et al.*, 2007 ; McAuliffe *et al.*, 2008). La formation de biofilms pourrait expliquer la capacité de certains mycoplasmes à persister dans le milieu extérieur, comme sur du matériel souillé, plus longtemps que le laisserait supposer l'apparente fragilité de ces bactéries dépourvues de paroi. L'utilisation de ce mode de prolifération *in vivo* reste à démontrer, il est toutefois vraisemblable qu'il puisse contribuer à la persistance de l'infection chez l'hôte et à la très faible efficacité des antibiotiques dans le traitement et la prévention des mycoplasmoses.

En conclusion, les données acquises sont encore parcellaires, mais combinées aux récents progrès de la génomique fonctionnelle et au développement de modèles d'infections cellulaires et animaux, *in vitro* et *in vivo*, elles joueront un rôle crucial dans l'avenir de notre compréhension des interactions entre les mycoplasmes et leurs hôtes. Par exemple, les méthodes de mutagénèse par transposition sont prometteuses pour l'identification de nouveaux gènes associés au pouvoir pathogène, même si elles restent limitées aux quelques espèces transformables. Le nombre de génomes de mycoplasmes séquencés ne cesse de s'accroître [Figure 5]. Au fur et à mesure que ces données de génomique abondent, les mycoplasmes apparaissent comme intégrés dans une dynamique complexe d'échanges génétiques (Sirand-Pugnet *et al.*, 2007b). Parmi l'extrême variété d'hôtes des mycoplasmes, un grand nombre d'espèces de *Mollicutes* et de mycoplasmes en particulier sont hébergées par des ruminants (Johansson et Pettersson, 2002 ; Citti, 2006 ; Le Grand *et al.*, 2008). Ces espèces qui partagent une même niche écologique, semblent avoir évolué de concert en échangeant du matériel génétique (Sirand-Pugnet *et al.*, 2007a ; Sirand-Pugnet *et al.*, 2007b).

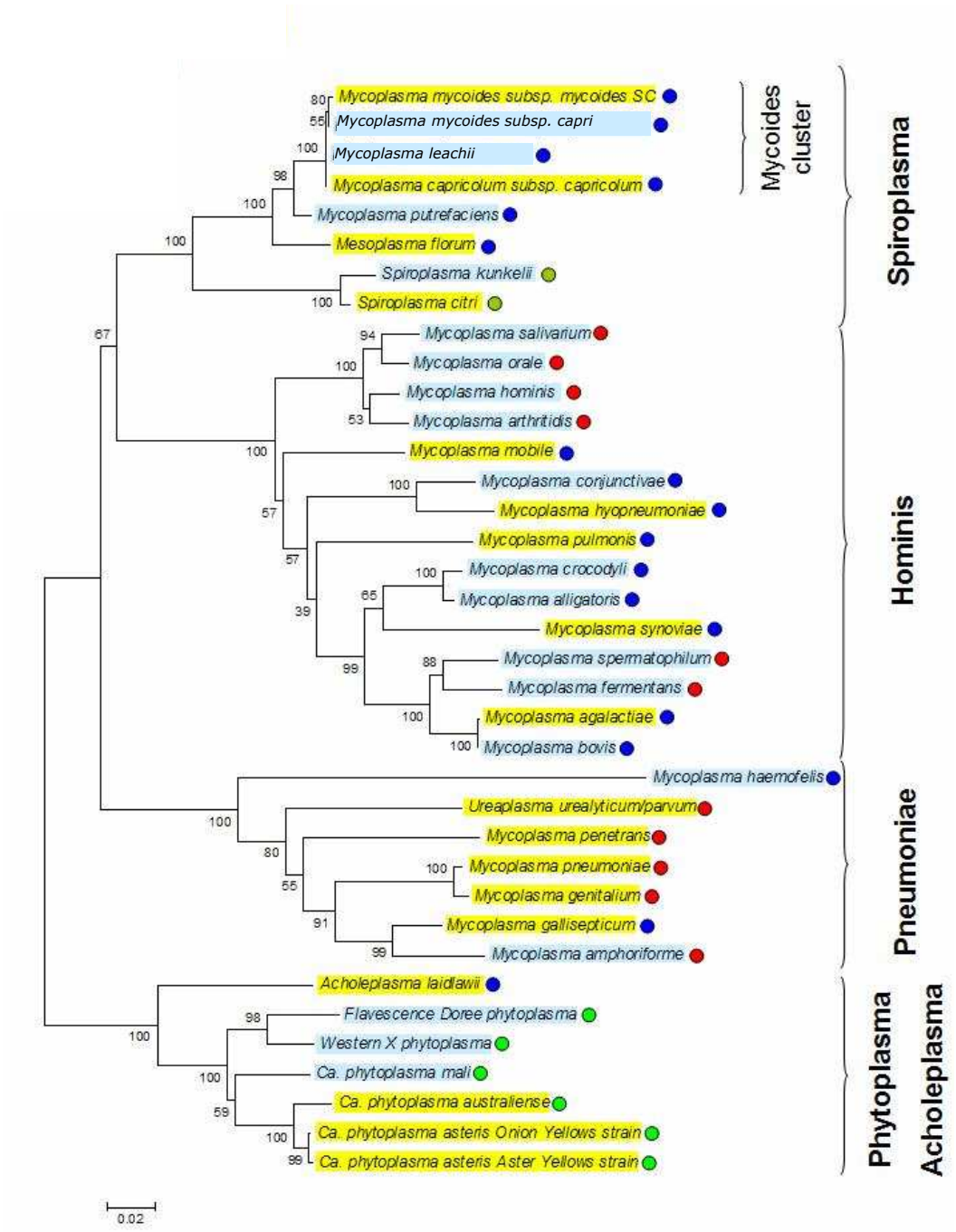


Figure 5 : Phylogénie des *Mollicutes*.

Arbre phylogénétique basé sur l'analyse des séquences de l'ARN 16S des *Mollicutes* dont les génomes sont séquencés (surlignés en jaune) ou en cours de séquençage (surlignés en bleu). Espèces rencontrées chez les végétaux (●), les animaux (●), et l'Homme (●). L'échelle représente le pourcentage de substitution nucléotidique.
Source : http://www.cbib.u-bordeaux2.fr/outils/molligen/INFORMATION/PHYLOGENY/tree.php

4. *Mycoplasmes et mycoplasmoses de ruminants*

Les découvertes scientifiques sur les mycoplasmes ont commencé avec l'étude d'une mycoplasmose de ruminants : la péripneumonie contagieuse bovine (PPCB) (Nocard et Roux, 1898). Depuis, plus de 25 espèces de mycoplasmes ont été isolées chez les ruminants. Leur importance médicale et économique est variable selon les espèces. Il s'agit :

- soit d'espèces rarement isolées pour lesquelles le pouvoir pathogène est mal connu car peu de données sont disponibles comme pour *M. canadense* et *M. alkalescens*,
- soit d'espèces fréquemment rencontrées pour lesquelles aucun pouvoir pathogène n'a été mis en évidence et qui constituent des espèces commensales (*M. bovirhinis*, *A. laidlawi*) ou opportunistes (*M. arginini*),
- soit d'espèces fréquemment isolées qui présentent expérimentalement un certain pouvoir pathogène mais dont l'impact réel lors d'infections naturelles n'est pas clairement défini (*M. dispar*, *U. diversum*, *M. canis*, *M. conjunctivae*, *M. ovis*...)
- soit d'espèces au pouvoir pathogène démontré et à l'origine de pertes économiques importantes comme *M. mycoides* subsp. *mycoides* biotype Small Colony (agent de la péripneumonie contagieuse bovine ou PPCB), *M. capricolum* subsp. *capripneumoniae* (agent de la pleuropneumonie contagieuse caprine ou PPCC), *M. capricolum* subsp. *capricolum*, *M. agalactiae*, *M. bovis*... (Frey, 2002 ; Le Grand *et al.*, 2008)

Dans ce paragraphe sur les mycoplasmoses des ruminants, nous nous limiterons à une brève description des espèces appartenant à cette dernière catégorie, à l'exception de l'agalactie contagieuse qui est l'objet du paragraphe suivant (§5). Les principales mycoplasmoses rencontrées chez les bovins et chez les ovins/caprins sont présentées, respectivement, dans les tableaux 3 et 4 (Citti, 2006).

Tableau 3 : Principales espèces de mycoplasmes rencontrées chez les bovins.

Mycoplasme[a]	Tropisme	Symptômes associés[b]
M. mycoides subsp. *mycoides* SC	Poumon	Pleuropneumonie
M. bovis	Mamelle Articulation Poumon Appareil génital Oreille interne Œil	Mammite Arthrite Pneumonie (Avortement) (Otite) (Kérato-conjonctivite)
M. leachii[c]	Mamelle Articulation Appareil génital	Mammite (Arthrite) (Avortements)
M. californicum	Mamelle Articulation	(Mammite) *Arthrite*
M. canadense	Mamelle Articulation Poumon Appareil génital	(Mammite)
M. alkalescens	Mamelle Articulation Poumons	(Mammite) (Arthrite)
U. diversum	Appareil génital Poumon	Infertilité Vulovaginite granuleuse *Avortements, salpingites, endométrites* (Broncho-pneumonie)
M. bovigenitalium	Poumon Mamelle Appareil génital	(Mammite) (Infertilité)
M. dispar	Poumon	(Bronchopneumonie)
M. wenyonii[d]	Sang	Anémie
M. bovirhinis	Poumon	non pathogène
M. arginini	Poumon Mamelle	non pathogène

D'après (Citti, 2006).
[a] la liste n'est pas exhaustive, seules les principales espèces sont citées, les espèces en caractère gras ont été isolées de bovins en France en 2007 (Poumarat, 2008).
[b] les symptômes rarement rencontrés sont indiqués entre parenthèses, ceux de fréquence mal connue sont en italique.
[c] nom d'espèce validement publié correspondant à "*Mycoplasma sp.* bovine group 7" (Manso-Silvan *et al.*, 2009).
[d] mycoplasme autrefois nommé *Eperythrozoon* wenyonii.

Tableau 4 : Principales espèces de mycoplasmes rencontrées chez les ovins et caprins.

Mycoplasme[a]	Hôtes	Tropisme	Symptômes associés[b]
M. capricolum subsp. *capripneumoniae*	caprins	Poumon	**Pleuropneumonie**
M. capricolum* subsp. *capricolum	caprins >ovins >>bovins	Mamelle Articulation Poumon	Mammites **Arthrites** **Pneumonie** (Kérato-conjonctivite)
M. mycoides* subsp. *capri [c]	caprins >ovins >>bovins	Mamelle Articulation Poumon Oeil	**Mammite** **Arthrite** **Pneumonie** (Kérato-conjonctivite)
M. agalactiae	ovins, caprins	Mamelle Articulation Poumon Oeil App. génital	**Mammite** Arthrite (Pneumonie) Kérato-conjonctivite
M. putrefaciens	caprins >ovins	Mamelle Articulation Poumon Oeil	**Mammite** **Arthrite** Pneumonie (Kérato-conjonctivite)
M. conjunctivae	ovins	Œil	**Kérato-conjonctivite**
M. ovipneumoniae	ovins	Poumon	**Pneumonie**
M. ovis [d]	ovins	Sang	**Anémie**
M. arginini	carpins, ovins, bovins	Mamelle App. génital	non pathogène

D'après (Citti, 2006).
[a] la liste n'est pas exhaustive, seules les principales espèces sont citées, les espèces en caractère gras ont été isolées d'ovins ou de caprins en France en 2007 (Poumarat, 2008).
[b] les symptômes rares sont indiqués entre parenthèses, ceux majoritairement rencontrés sont en caractères gras, ceux souvent rencontrés sont en police normale.
[c] réunissant les mycoplasmes autrefois distingués comme *M. mycoides* subsp. *capri* et *M. mycoides* subsp. *mycoides* LC
[d] autrefois *Eperythrozoon ovis*

4.1 La péripneumonie contagieuse bovine

Aujourd'hui, la péripneumonie contagieuse bovine (PPCB) est considérée par la FAO (*Food and Agriculture Organisation of the United States*) comme le plus grand fléau de l'élevage bovin dans toute l'Afrique à l'exception de l'Afrique du Nord. Maladie à déclaration obligatoire, inscrite sur la liste des maladies à éradiquer de l'OIE (Organisation Internationale des Epizooties Office International des Epizooties, organisation mondiale de la santé animale), elle est considérée comme la maladie mycoplasmique au plus fort impact économique dans le monde. Après une panzootie mondiale au XIX^{e} siècle, la PPCB est actuellement endémique en Afrique et sévit de manière sporadique en Asie. En Europe, elle a été éradiquée au début du XX^{e} siècle mais plusieurs centaines de foyers ont ré-émergés dans les années 1980 en France et dans le Sud de l'Europe (Lefèvre, 2003 ; Poumarat *et al.*, 2004). Les derniers foyers européens datent de 1999 au Portugal (Office International des Epizooties, 2009) et il est à ce jour trop tôt pour pouvoir affirmer qu'elle est à nouveau éradiquée en Europe (Lefèvre, 2003). Selon les études moléculaires, les souches européennes contemporaines dérivent des souches européennes ancestrales de la fin du XIX^{e} et se différencient nettement des souches africaines actuelles. La récente PPCB européenne serait donc une résurgence d'une PPCB endémique en Europe et non une importation d'Afrique. En outre, cette PPCB européenne semble moins virulente et moins immunogène que celle observée en Afrique (Abdo el *et al.*, 1998). Malgré cette apparente atténuation, les conséquences que la PPCB entraîne pour la circulation des bovins constituent, à elles seules, un très grave problème économique (Le Grand *et al.*, 2008).

La PPCB affecte les bovins, les zébus, les buffles, les élans, les bisons et les yacks. L'agent étiologique, *M. mycoides* subsp. *mycoides* biotype Small Colony, a aussi été isolé d'ovins et de caprins qui peuvent constituer des réservoirs (Brandao, 1995 ; Srivastava *et al.*, 2000). Le tableau clinique est celui d'une grave pneumopathie sans signe spécifique particulier. En revanche, les lésions

observées d'inflammation exsudative du poumon et de la plèvre sont caractéristiques, surtout pour la forme africaine. Elles atteignent préférentiellement les lobes postérieurs, souvent de manière dissymétrique. A la coupe, des lésions hétérogènes d'hépatisation du parenchyme pulmonaire associées à une hypertrophie des espaces interlobulaires sont observées. L'évolution est lente et tous les stades inflammatoires peuvent être observés sur un seul lobule. En phase chronique, des séquestres peuvent être présents (Thiaucourt, 2003 ; Poumarat *et al.*, 2004).

4.2 Mycoplasma bovis

Mycoplasma bovis intervient dans deux affections ayant un impact économique majeur chez les bovins : les mammites, particulièrement fréquentes aux États-Unis, et les bronchopneumonies infectieuses enzootiques des veaux et des jeunes bovins (agés de 6 à 18 mois) (Pfutzner et Sachse, 1996). Il a été isolé pour la première fois en 1961 lors d'un cas de mammite aiguë chez une vache aux Etats-Unis (Hale *et al.*, 1962). Depuis, il a été isolé dans la plupart des pays du monde. Sa fréquente association avec d'autres agents dont l'isolement est plus aisé comme *Mannheimia haemolytica*, *Pasteurella multocida* ou *Haemophilus somnus* conduit sûrement à une sous-estimation de sa prévalence. Cependant il est une cause primaire de maladie et sa présence dans un élevage est invariablement associée à des symptômes de type pneumonie, mammite et/ou arthrite. Il est aussi rencontré lors d'épisodes d'otites, de kérato-conjonctivites, d'infections du tractus génital et d'avortements (Le Grand *et al.*, 1996). Les postulats de Koch ont été vérifiés pour les affections causées par *M. bovis* : des inoculations expérimentales reproduisent les signes cliniques et lésionnels observés chez les bovins, même en l'absence d'autres agents infectieux, et l'agent peut être ré-isolé de ces animaux (Gourlay *et al.*, 1979 ; Illing, 1979 ; LaFaunce et McEntee, 1982 ; Pfutzner *et al.*, 1983 ; Rodriguez *et al.*, 1996 ; Henderson et Ball, 1999 ; Rodriguez *et al.*, 2000). Quelle que soit l'expression clinique observée, la maladie est généralement chronique et invalidante (Gourlay *et al.*, 1989 ; Adegboye *et al.*, 1995 ; Apley et Fajt, 1998 ; Gagea *et al.*, 2006). Des foyers avec une forte et

subite morbidité peuvent survenir et l'impact économique sur les élevages touchés peut être dévastateur (Gourlay *et al.*, 1989 ; Walz *et al.*, 1997 ; Butler *et al.*, 2000). L'impact financier est surtout lié au coût du traitement des animaux atteints. De plus, suite à la faible réponse aux traitements, les animaux peuvent représenter des non-valeurs économiques (Nicholas et Ayling, 2003). L'antibiothérapie donne en effet des résultats décevants. A ce jour, le contrôle de l'infection repose essentiellement sur la maîtrise de facteurs de risques potentiels ou identifiés comme la densité d'animaux, la ventilation… (Ames, 1997 ; Apley, 2006 ; Maunsell et Donovan, 2009)

Il n'existe aucun vaccin ayant une AMM (Autorisation de Mise sur le Marché) en France, des vaccins à base de souches inactivées sont commercialisés par exemple aux Etats-Unis mais leur efficacité est sérieusement contestée (Maunsell *et al.*, 2009). Le développement de stratégies préventives et curatives efficaces est un point critique de la lutte contre les maladies associées à *M. bovis*.

4.3 Mycoplasma leachii

M. leachii autrefois connu comme "*Mycoplasma sp.* bovine group 7" a été initialement identifié en Australie comme responsable d'arthrites chez le veau. Une ré-inoculation expérimentale de la souche isolée a permis de reproduire les symptômes observés (Simmons et Johnston, 1963). Depuis, il a été également isolé lors d'épisodes cliniques combinant des mammites, des polyarthrites ou des pneumonies en Australie (Connole *et al.*, 1967 ; Shiel *et al.*, 1982 ; Alexander *et al.*, 1985). Plus récemment, il a été identifié comme responsable d'un syndrome "mammites, arthrites et avortements" (Hum *et al.*, 2000). Quelques rares cas ont été mis en évidence hors du continent australien (Langford, 1975). Si *M. leachii* est un réel agent pathogène pour les bovins, sa répartition et son importance économique est toutefois sans commune mesure avec *M. bovis*.

4.4 La pleuropneumonie contagieuse caprine

La pleuropneumonie contagieuse caprine (PPCC) est une maladie comparable à la péripneumonie contagieuse bovine qui atteint exclusivement les caprins. Inscrite sur la liste de l'OIE, la PPCC sévit en Afrique, au Moyen-Orient et en Asie. Déjà connue et décrite au XIX^{e} siècle (Thomas, 1873 ; Hutcheon, 1881), le premier isolement de son agent étiologique, *M. capri* subsp. *capripneumoniae* (Leach *et al.*, 1993), date de 1976 (MacOwan et Minette, 1976, 1977). Ce retard est probablement dû aux difficultés posées par sa culture en milieu inerte. Son appartenance aux mycoplasmes du groupe "mycoides" avec lesquels il partage des caractéristiques pathogéniques, antigéniques et génétiques vient également compliquer le tableau. Sur ce point, l'apport des techniques de biologie moléculaire a permis une amélioration du diagnostic (Thiaucourt, 2003).

A la différence des autres mycoplasmes caprins du groupe "mycoides" qui participent à un syndrome polymorphe qu'il convient d'appeler "MAKePS" pour "mammites, arthrite, agalactie, kératites, pneumonies, septicémie" , *M. capri* subsp. *capripneumoniae* est l'agent d'une maladie parfaitement définie au tropisme pulmonaire (Thiaucourt, 2003). À l'autopsie, les lésions sont assez caractéristiques et limitées à la cavité thoracique. En présence d'un liquide fibrineux, des dépôts de fibrine couvrent les poumons et sont souvent adhérents à la plèvre. L'hépatisation pulmonaire n'atteint souvent qu'un seul poumon parfois dans son intégralité. Des abcès purulents de surinfections bactériennes peuvent aussi être observés lors de formes chroniques (Frey, 2002 ; Thiaucourt, 2003).

4.5 Le syndrome MAKePS

Présentes dans le monde entier, les mycoplasmoses du mouton et de la chèvre représentent un problème socio-économique important, particulièrement dans les régions où ces animaux constituent une part importante des ressources alimentaires de lait et de viande pour les populations (Frey, 2002). De nombreuses espèces de mycoplasmes sont rencontrées chez les petits ruminants,

particulièrement chez les caprins où la présence de mélange d'espèces au sein d'un troupeau voire d'un même animal est chose fréquente (Bergonier et Poumarat, 1996). Autrefois source de confusion diagnostique avec la PPCC, un syndrome pouvant être appelé "MAKePS" pour "mammites, agalactie, arthrite, kératites, pneumonies, septicémie" (Thiaucourt, 2003) est également causé par des mycoplasmes du groupe "mycoides" : *M. mycoides* subsp. *capri* (regroupant les taxons historiquement nommés *M. mycoides* subsp. *mycoides* biotype Large Colony et *M. mycoides* subsp. *capri* (Manso-Silvan *et al.*, 2009)), *M. capricolum* subsp. *capricolum*. Les mycoplasmes du groupe "mycoides" sont inclus dans un plus grand groupe phylogénétique appelé "spiroplasma". Outre des pathogènes de plantes, ce groupe contient *M. putrefaciens* qui, sans faire partie du groupe "mycoides", peut être intégré aux mycoplasmes du syndrome MAKePS en tant que pathogène de ruminant phylogénétiquement proche et à la symptomatologie comparable. Ce syndrome présente également des similitudes cliniques avec l'agalactie contagieuse à *M. agalactiae*. Cependant la grande distance phylogénétique qui sépare ces mycoplasmes conduit à considérer ces ressemblances comme le résultat d'une convergence évolutive, en lien notamment avec des transferts génétiques horizontaux et le partage d'une même niche écologique. Il apparaît ainsi plus pertinent de distinguer le syndrome "MAKePS" par une appellation spécifique afin de dédier le terme d'agalactie contagieuse au sens strict à l'infection par *M. agalactiae*.

5. *L'agalactie contagieuse à Mycoplasma agalactiae*

L'agalactie contagieuse au sens large est un syndrome pouvant être causé par quatre mycoplasmes différents. En Europe, le syndrome de l'agalactie contagieuse constitue l'affection mycoplasmique la plus fréquente et la plus grave chez les petits ruminants. Il compte parmi les dominantes pathologiques des troupeaux caprins français. *M. mycoides* subsp. *mycoides* LC (aujourd'hui regroupé au sein du taxon *M. mycoides* subsp. c*apri*), *M. capricolum* subsp. *capricolum*, *M. putrefaciens* et *M. agalactiae* sont les quatre mycoplasmes impliqués (Bergonier et Thiaucourt, 2003). Les trois premiers, plus ou moins apparentés au groupe "mycoides", peuvent être décrits comme les mycoplasmes du syndrome MAKePS (Thiaucourt, 2003) (cf. paragraphe précédent). Ils affectent principalement les caprins alors que *M. agalactie* est retrouvé chez les ovins et les caprins. Nous n'envisagerons ici que l'agalactie contagieuse, au sens strict, causée par *M. agalactiae*.

M. agalactiae est un pathogène important des ovins et des caprins inscrit sur la liste de l'OIE (Office International des Epizooties, organisation mondiale de la santé animale). Bien que des échanges génétiques par transferts horizontaux aient été démontrés avec les mycoplasmes du groupe "mycoides" (Sirand-Pugnet *et al.*, 2007b), *M. agalactiae* est phylogénétiquement éloigné de ce groupe et c'est avec *M. bovis* qu'il a une proche et étroite parenté avec plus de 99,8% de similitude entre les séquences des ARNr 16S (Pettersson *et al.*, 1996). Tout deux appartiennent en effet au groupe phylogénétique "hominis" [Figure 5]. Contrairement au groupe "mycoides" qui ne comporte que des pathogènes de ruminants, le spectre d'hôte des mycoplasmes du groupe "hominis" s'étend de *M. hominis*, pathogène humain, à *M. pulmonis*, pathogène de rongeurs, en passant par *M. crocodyli*, pathogène de reptiles. Avec *M. conjunctivae, M. bovis* et *M. agalactiae* constituent les pathogènes de ruminants de ce groupe.

5.1 Epidémiologie

5.1.1 Espèces sensibles

Les ovins et les caprins sont les principales espèces affectées. Quelques isolements sont rapportés chez les caprins sauvages avec une expression clinique frustre ou atypique comme des pleuropneumonies, notamment chez les bouquetins espagnols (Gonzalez-Candela *et al.*, 2006 ; Verbisck-Bucker *et al.*, 2008) et ceux des Alpes françaises ou italiennes (Auliac, 2008). Globalement, le taux d'atteinte et les signes cliniques sont habituellement plus marqués chez les caprins. Alors que la morbidité atteint généralement moins de 50 % des brebis en lactation, elle peut être supérieure chez les chèvres. Les mâles, les femelles gravides ou taries sont habituellement moins atteints (Bergonier et Thiaucourt, 2003).

5.1.2 Répartition géographique

La répartition de l'agalactie contagieuse est mondiale. Il existe cependant des zones indemnes de *M. agalactiae* malgré une forte concentration de petits ruminants comme le Royaume-Uni et l'Australie. Les régions les plus touchées se situent dans les zones traditionnelles d'élevage de la majorité des pays du pourtour méditerranéen et d'Asie occidentale. La perte de production laitière engendrée par la maladie a un impact majeur sur l'économie des élevages. Dans les régions atteintes, la maladie évolue généralement de manière enzootique, avec la possibilité de flambées épizootiques.

En France, il existait deux zones d'enzootie, la Savoie et les Pyrénées-Atlantiques, bassins d'élevage laitier à dominante caprine pour le premier, ovine pour le second. En Savoie, grâce à des mesures sanitaires sévères, le dernier cheptel infecté remonte à 2002 et cette région peut aujourd'hui être considérée comme assainie et indemne de *M. agalactiae*. Dans les Pyrénées-Atlantiques, la maladie a toujours persisté malgré l'application de mesures semblables à celles appliquées en Savoie.

Mis en place au début des années 1990, le programme de lutte sanitaire des Pyrénées-Atlantiques avait permis de circonscrire l'infection à quelques cantons de Basse-Navarre. Parmi les 3650 troupeaux que compte environ le département, la prévalence de la maladie était ainsi passée de près de 3,5 % en 1992 à 0,54% en 2005. De 2003 à 2005, aucun nouvel élevage n'a déclaré la maladie, et au total, une vingtaine d'élevages infectés était dénombrée en 2005. Mais la maladie refait parler d'elle en 2006 avec 11 nouveaux cheptels infectés, s'amorce alors une recrudescence. Au cours de l'année 2008, 135 foyers d'agalactie dont 95 nouveaux cas sont recensés. La prévalence de l'infection est ainsi de 3,7 % et rejoint le niveau d'avant 1992 et la mise en place des mesures prophylactiques [Figure 6].

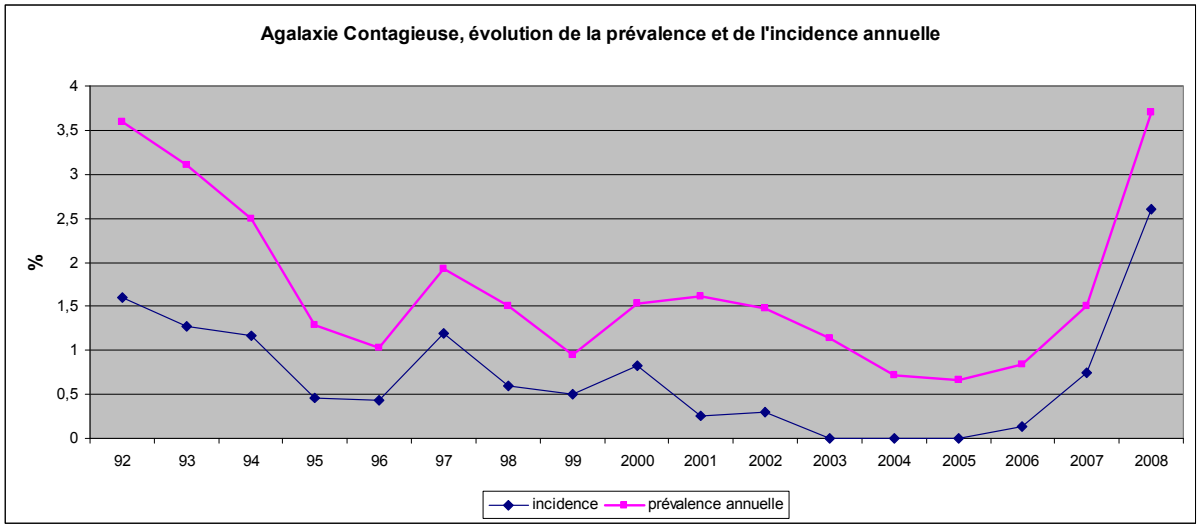


Figure 6 : Evolution de l'incidence et de la prévalence de l'agalactie contagieuse dans les Pyrénées-Atlantiques de 1992 à 2008.

Ce graphique représente, en ordonnée, l'incidence (pourcentage de nouveaux élevages infectés au cours d'une année) et la prévalence (pourcentage d'élevages infectés) en fonction des années, en abscisse. Source : (GDS64, 2009).

5.1.3 Evolution au sein d'un troupeau atteint

Au sein d'un troupeau infecté, l'évolution de l'incidence est variable. La maladie peut apparaître soudainement sous forme d'une "explosion" clinique ou affecter successivement un nombre restreint d'animaux par "bouffées épidémiques". Classiquement, des nouveaux cas échelonnés dans le temps apparaissent régulièrement dans le troupeau. La prévalence au sein du cheptel est fortement influencée par le stade physiologique des animaux, les contacts et mouvements d'animaux liés à la transhumance. De ce fait, la prévalence maximale est généralement observée à trois moments clefs : autour des mises bas, en début de période de traite et durant la transhumance estivale. En l'absence d'intervention (abattage, antibiothérapie, vaccination), la persistance des signes cliniques est habituellement de l'ordre de plusieurs mois avec des récidives possibles aux lactations suivantes. En zone d'enzootie, au fil des années consécutives à l'épisode initial de la maladie, l'atteinte clinique se focalise principalement sur les primipares et les jeunes à la mamelle (Bergonier, 1996 ; Bergonier et Berthelot, 2008).

La persistance sérologique est forte. Une estimation réalisée dans les Pyrénées-Atlantiques montre une négativation de l'indice de troupeau associée à la disparition des symptômes, en moyenne, au bout de 5 ans après l'épisode clinique initial. La persistance microbiologique peut également être longue avec des troupeaux devenus asymptomatiques mais toujours excréteurs 1 à 8 ans après la déclaration de la maladie (Bergonier, 1996).

5.2 Tableau clinique

5.2.1 Formes cliniques habituellement rencontrées

L'expression clinique de l'agalactie contagieuse est en général chronique mais débute souvent par des formes subaiguës à aiguës [Tableau 5]. A l'échelon du troupeau, l'expression clinique est triple : mammaire, articulaire et oculaire. Cette triade est évocatrice de l'agalactie contagieuse mais apparaît souvent de manière

dissociée. L'expression mammaire est souvent la première et parfois la seule constatée sur les animaux en lactation. Cette atteinte est au moins fonctionnelle et va de l'hypogalactie temporaire à l'agalactie totale mais des signes inflammatoires locaux peuvent également être observés. Les symptômes oculaires et articulaires viennent ensuite et atteignent en général moins de 20 à 30 % du troupeau. Il s'agit habituellement de conjonctivites évoluant vers des kérato-conjonctivites, d'arthrites et de polyarthrites plus ou moins sévères (Bergonier et Thiaucourt, 2003).

Tableau 5 : Caractéristiques cliniques générales de l'agalactie contagieuse à *M. agalactiae* chez les ovins et les caprins.

	ovins	**caprins**
Mode d'évolution dominant	subaiguë à chronique	aiguë à chronique
Mortalité		
Adultes (hors lactation)	0-3 %	
Femelles en lactation	0-10 %	0-40 %
Jeunes	1-10 %	3-50 %
Morbidité		
Adultes	0-20 %	
Femelles en lactation	1-50 %	10-90 %
Jeunes	2-60 %	10-70 %
Tropisme habituel		
Adultes	articulaire > oculaire >> respiratoire	
Femelles en lactation	mammaire >> articulaire > oculaire > avortements > respiratoire	
Jeunes	articulaire, oculaire, respiratoire, formes septicémiques possibles	
Formes atypiques	respiratoires, génitales	

D'après (Bergonier *et al.*, 2002).

Des signes pulmonaires, bien que plus rares, peuvent apparaître, ils atteignent alors principalement les jeunes. Sur ces derniers, arthrites et kérato-conjonctivites

dominent mais des formes septicémiques avec peu d'expression clinique sont également possibles.

Des formes frustres, voire asymptomatiques sur l'ensemble du troupeau peuvent exister. Ces cas d'infections subcliniques sont rencontrés dans des régions enzootiques comme dans des régions dépourvues de signes cliniques. Ils sont relativement fréquents et peuvent être détectés par isolement sur lait de tank (Bergonier et Thiaucourt, 2003). Aucune différence entre ces isolats et ceux de formes cliniques n'a été mise en évidence (Bonnefin, 2003). Il ne s'agit pas de souches atténuées, ces mêmes souches peuvent provoquer des formes symptomatiques (Sanchis *et al.*, 2000 ; Galisson, 2001). Ces formes d'emblée subcliniques sont à différencier des récupérations fonctionnelles et de l'atténuation des signes cliniques pouvant aboutir à une absence d'expression clinique au cours du temps après un épisode symptomatique.

5.2.2 Formes atypiques

Outre les formes asymptomatiques, des infections par *M. agalactiae* peuvent être à l'origine de formes atypiques dont l'expression clinique est éloignée de l'habituelle triade "agalactie, arthrite, kérato-conjonctivite". *M. agalactiae* a ainsi été impliqué lors d'épisodes de vulvovaginites granuleuses chez des chèvres en Inde et en Afrique (Singh *et al.*, 1974). Sporadiquement et principalement chez les caprins, des atteintes pulmonaires sans autre signe clinique ont été observées dans divers pays notamment en Inde et en Australie (Bergonier et Poumarat, 1996). Des manifestations pulmonaires de type pneumonie ou pleuropneumonie dues à *M agalactiae* sont aussi observées chez les bouquetins des Alpes françaises (*Capra ibex ibex*). Il semble que cela soit l'expression clinique dominante pour cette espèce de la faune sauvage (Barthe *et al.*, 2006 ; Auliac, 2008). Un portage auriculaire apparemment asymptomatique a été identifié chez des bouquetins espagnols (*Capra pyrenaica hispanica*) (Gonzalez-Candela *et al.*, 2006 ; Gonzalez-Candela *et al.*, 2007 ; Verbisck-Bucker *et al.*, 2008).

5.3 Diagnostic

En présence des trois manifestations cliniques typiques (mammites ou chute de production laitière, kérato-conjonctivites et arthrites), accompagnées de données épidémiologiques, le diagnostic se trouve facilité. Toutefois, selon les troupeaux, l'espèce atteinte et l'ancienneté de l'infection, les symptômes observés peuvent être plus ou moins polymorphes. Le recours au laboratoire doit donc être systématique pour confirmer les suspicions. Le diagnostic différentiel doit avoir au préalable intégré l'ensemble des affections ovines ou caprines ciblant les mêmes organes (arthrite encéphalite virale caprine, virus Maedi-Visna, infections streptococciques ou staphylococciques, arcanobactériose, mannheimiose, érésipèle articulaire...) (Bergonier, 1996).

5.3.1 Diagnostic direct

5.3.1.1 Prélèvement

Il convient de réaliser des prélèvements sur les organes d'animaux non traités présentant une atteinte récente. Compte tenu du fort tropisme pour la mamelle, le lait est un prélèvement de choix, il est de plus de réalisation facile. Le prélèvement de liquide synovial par lavage articulaire peut aussi être intéressant. Plus aisément réalisable *post-mortem,* il donne cependant des résultats inégaux surtout lors d'arthrites chroniques. Le diagnostic direct peut également être établi à partir d'écouvillonnages oculaires, de prélèvements pulmonaires ou génitaux lorsqu'ils sont atteints, ils pourront cependant être contaminés par des bactéries ou des mycoplasmes saprophytes comme *M. arginini* (Bergonier, 1996 ; Bergonier et Berthelot, 2008).

En l'absence de symptômes, le lait de tank ou de mélange constitue un prélèvement facile et instructif pour les troupeaux laitiers. A l'échelon individuel, une analyse de lait ou d'écouvillon auriculaire peut contribuer à la détection d'animaux porteurs et limiter ainsi les risques d'introduction lors d'achats ou de ventes (Lambert, 1987 ; Bergonier et Poumarat, 1996).

5.3.1.2 Culture et isolement

Pour une mise en culture, les prélèvements doivent être rapidement transportés au laboratoire sous couvert du froid positif, la congélation doit être évitée. L'ensemencement est d'abord réalisé en bouillon dans des milieux de culture spécifiques pour mycoplasmes. Ces milieux sont des milieux complexes composés d'au moins trois constituants : i) une poudre spécifique à base d'extrait de viande appelée PPLO pour *pleuropneumoniae-like organisms*, ii) de l'extrait de levure, iii) du sérum (généralement de cheval ou de veau fœtal). L'adjonction d'un ou plusieurs antibiotiques actifs sur les bactéries à paroi, voire d'acétate de thallium permet de rendre le milieu partiellement sélectif (Freundt, 1983 ; Tully, 1995).

Par comparaison aux autres bactéries, la culture des mycoplasmes est lente, et celle de *M. agalactiae* requiert en laboratoire 2 à 5 jours d'incubation en milieu complexe. Des dilutions et passages en milieu frais, réalisés tous les 3 à 4 jours donnent une meilleure sensibilité probablement par la dilution d'inhibiteurs de croissance. Un étalement sur milieu gélosé permet de visualiser les colonies de *M. agalactiae* qui présentent l'aspect typique des colonies de mycoplasmes en "oeuf sur le plat". Après quelques jours à l'étuve, elles forment également un léger film et des concrétions punctiformes appelés "*film and spots*" qui correspondent à la saponification du milieu et à la précipitation de stérols.

5.3.1.3 Identification

L'identification d'espèce basée sur des épreuves biochimiques est d'interprétation délicate pour les mycoplasmes [pour une revue voir (Miles, 1992 ; Pollack, 1995)]. Deux approches dominent le diagnostic courant : l'identification sérologique et par PCR. Historiquement, les méthodes sérologiques font appel à l'inhibition de croissance (Clyde, 1964), l'inhibition du métabolisme (Taylor-Robinson *et al.*, 1966), et à l'immunodétection sur colonies révélée par fluorescence (Rosendal et Black, 1972) ou par la péroxydase (Polak-Vogelzang *et al.*, 1978). Aujourd'hui, un test d'immunoempreinte après filtration sur membrane (*dot blot immunobinding*) (Kotani et McGarrity, 1985 ; Poumarat *et al.*, 1991) est utilisé en routine en France : les mycoplasmes sont retenus sur une membrane par filtration des

cultures, et identifiés au moyen d'antisérums hyperimmuns polyclonaux. La spécificité obtenue est au moins équivalente aux tests précédents et la sensibilité meilleure. Grâce à l'utilisation de cupules en plaques de 96 puits, cette méthode offre un bon rendement avec une grande faisabilité et une standardisation aisée. Outre sa bonne sensibilité, son principal avantage réside dans la possible détection de mélanges d'espèces mycoplasmiques (Poumarat *et al.*, 1991 ; Poumarat *et al.*, 1992). Cette méthode d'identification est employée en première intention dans le cadre du réseau d'épidémio-surveillance de l'ensemble des mycoplasmoses de ruminants, le réseau VIGIMYC (Poumarat *et al.*, 2006). Des réactions croisées d'interprétation complexe peuvent toutefois être observées. Il peut alors être fait appel à la détection de séquences spécifiques d'espèces par PCR et au séquençage partiel de gènes comme l'ARN 16S (Tardy *et al.*, 2009).

Au Royaume Uni, l'équipe Mycoplasmes du *Veterinary Laboratories Agency* a mis au point une méthode d'identification en DGGE (*denaturing gradient gel electrophoresis*) basée sur le polymorphisme des ARN 16S (McAuliffe *et al.*, 2005). Cette méthode visualise en gel gradient dénaturant (DGGE) la différence de migration de produits de PCR intéressant la région V3 du gène de l'ARNr 16S générés à l'aide d'amorces universelles *Mollicutes*. Cette migration différentielle est corrélée aux différences de séquence de cette région de l'ARN entre les espèces. Si cette méthode peut être considérée comme équivalente au *dot blot immunobinding* en première intention, elle donne des résultats d'identification équivoques pour les espèces du groupe "mycoides" (Tardy *et al.*, 2009) et ne permet pas de déceler clairement les mélanges d'espèces mycoplasmiques (Tardy *et al.*, 2007).

La réaction d'amplification par polymérase ou PCR (pour *polymerase chain reaction*) représente une technique prometteuse. Les cibles choisies doivent être à la fois spécifiques inter-espèces et conservées de manière intra-spécifique. De nombreuses cibles peuvent être utilisées comme l'ARN 16S (Chavez Gonzalez *et al.*, 1995), le gène *uvrC* (Subramaniam *et al.*, 1998), le gène *polC* (Marenda *et al.*, 2005), le gène de la protéine P80 (Dedieu *et al.*, 1995 ; Tola *et al.*, 1996 ; Tola *et al.*, 1997 ; Greco *et al.*, 2001 ; Lorusso *et al.*, 2007), le gène *p40* (Oravcova *et al.*, 2009)… Sur culture, la sensibilité peut dépasser 10^2 UFC/ml. La technique est

automatisable et utilisable à grande échelle. A titre d'exemple, le laboratoire d'analyses vétérinaires des Pyrénées-Atlantiques (Laboratoire des Pyrénées, 64150 Lagor) utilise la combinaison systématique de deux PCR en temps réel l'une ciblant l'ARN 16S et l'autre le gène codant la protéine P30 pour l'identification de *M. agalactiae* après une étape d'enrichissement par culture. Cette analyse de routine réalisée pour le dépistage de l'agalactie contagieuse concerne un peu plus de 7000 prélèvements de lait par an.

A ce jour, sûrement à cause de la complexité de spécimens comme le lait, la sensibilité obtenue n'est pas assez satisfaisante pour autoriser la détection directe et s'affranchir de l'étape préalable d'enrichissement en culture. Des résultats prometteurs ont tout de même été obtenus (Tola *et al.*, 1997 ; Lorusso *et al.*, 2007 ; Fitzmaurice *et al.*, 2008 ; Oravcova *et al.*, 2009) et des améliorations sont à l'étude comme l'utilisation d'amorces choisies pour leur sensibilité accrue, l'optimisation de l'étape d'extraction de l'ADN, la concentration des spécimens notamment pour les laits de mélange (GDS64, 2009).

5.3.2 Diagnostic indirect

Diverses techniques ont été testées (ELISA, fixation du complément, immunofluorescence, agglutination sur lame, inhibition de croissance, immunodiffusion…). Aujourd'hui, le diagnostic sérologique repose essentiellement sur des tests ELISA (*enzyme linked immunosorbent assay*) basés sur des antigènes totaux (Lambert *et al.*, 1998 ; Assuncao *et al.*, 2004) ou quelques antigènes recombinants (Fusco *et al.*, 2007). En France, historiquement deux kits commerciaux étaient utilisés (Pepin *et al.*, 2003). Le premier basé sur des antigènes totaux extraits de la souche de référence PG2 n'est plus commercialisé. Le second basé sur un antigène recombinant évalue la réponse sérologique vis-à-vis d'un seul antigène, l'antigène P48, correspondant à une protéine de surface majeure (Rosati *et al.*, 1999 ; Rosati *et al.*, 2000).

Selon l'étude de Pépin *et al.* (Pepin *et al.*, 2003), ce kit présente à l'échelon individuel une bonne spécificité (99%) mais une sensibilité plus faible (82%). A l'échelon du troupeau, la spécificité décroît mais reste acceptable (92%) et la

sensibilité devient maximale (100%). A l'échelon individuel, la sérologie ne constitue donc qu'un outil de dépistage (et non de diagnostic) pour une discrimination floue entre des animaux infectés ou douteux et des animaux présumés sains. L'utilité de ce diagnostic sérologique réside principalement dans le dépistage au niveau collectif. En région infectée, un indice sérologique de troupeau, calculé sur la base des résultats d'un échantillon de sérums individuels, est donc un des outils fondamentaux de la qualification des cheptels pour organiser les programmes de contrôle de la maladie.

5.4 Traitement

Le traitement des mycoplasmoses repose sur l'utilisation d'antibiotiques actifs sur ce type de germes. Il convient donc d'utiliser des molécules actives sur les bactéries dépourvues de paroi, présentant une bonne persistance plasmatique, une bonne diffusion tissulaire et une grande capacité à se concentrer dans la mamelle. Les principales molécules utilisées sont les tétracyclines, les phénicolés, les fluoroquinolones et les macrolides.

L'objectif principal d'un traitement antibiotique est symptomatologique. Il permet une amélioration de l'état clinique des animaux malades et empêche une éventuelle mortalité surtout chez les caprins. Tout au plus il favorise une guérison clinique et une reprise de la lactation mais l'obtention d'une guérison bactériologique est illusoire.

Le traitement systématique de l'ensemble des animaux du troupeau, malades ou non, est régulièrement préconisé (Bergonier, 1996 ; Bergonier et Berthelot, 2008). Compte tenu de l'effectif important des cheptels, les molécules "longue action" ou formes galéniques dite "retard" sont privilégiées. Le traitement des animaux en lactation se heurte au problème des pertes économiques engendrées par les temps d'attente. Globalement, le rapport coût/bénéfice d'une antibiothérapie est décevant.

5.5 Prophylaxie

5.5.1 Prophylaxie médicale

5.5.1.1 Immunoprévention

Aucun vaccin n'est autorisé sur le marché français. Dans le monde, les stratégies vaccinales utilisées reposent principalement sur des vaccins inactivés par la chaleur ou le formol accompagnés ou non d'adjuvants immunogènes. Les cultures utilisées proviennent soit de souches entretenues en laboratoire, soit d'isolats originaires du foyer d'infection à vacciner dans le cas des auto-vaccins. Globalement, les résultats obtenus sont décevants, quelle que soit la souche ou l'adjuvant retenus. L'immunité conférée est partielle et de courte durée obligeant des injections de rappel très rapprochés (jusqu'à un rappel tous les 3 mois). Ces vaccins ne protègent pas contre l'infection, au mieux ils permettent une atténuation des signes cliniques. Leur utilisation favorise ainsi l'apparition d'excréteurs asymptomatiques et fait courir le risque de masquer certaines infections naturelles. Des améliorations sont donc nécessaires dans ce domaine, des vaccins sous-unitaires sont à l'étude. Idéalement, ils devraient diminuer les signes cliniques et l'excrétion par les animaux infectés, protéger les animaux indemnes, tout en autorisant la différenciation entre cheptels infectés et vaccinés.

5.5.1.2 Chimioprévention

La mise en place d'une antibioprophylaxie n'est pas envisageable vu l'insidiosité de la transmission, le caractère chronique et persistent des infections. Une étude de l'antibiothérapie réalisée au tarissement pour accélérer l'assainissement des cheptels infectés est actuellement en cours dans les Pyrénées-Atlantiques (GDS64, 2009).

5.5.2 Prophylaxie sanitaire

En l'absence de moyens médicaux efficaces, les mesures sanitaires restent pour le moment, la seule prophylaxie envisageable pour lutter contre l'agalactie contagieuse. En France, les deux bassins de production où sévissait

historiquement *M. agalactiae* (Pyrénées-Atlantiques, Savoie et Haute-Savoie) ont ainsi choisi d'adopter une prophylaxie purement sanitaire. Dans les Pyrénées-Atlantiques ce programme repose sur i) la déclaration obligatoire des suspicions cliniques et leur confirmation au laboratoire, ii) la qualification des cheptels de l'ensemble du département basée sur un dépistage bisannuel sur lait de tank et un sondage sérologique annuel iii) l'abattage total fortement incité par la subvention du renouvellement, l'interdiction de transhumance et d'échanges, l'isolement strict pour les cheptels infectés, iv) la gestion des transactions et des mouvements d'animaux en fonction des qualifications et la réalisation de sérologies lors d'achats.

Ces mesures permettent de diminuer la prévalence de la maladie. Les éradications réussies en Savoie et dans certaines provinces ou sous-parties des Pyrénées-Atlantiques apportent la preuve de leur efficacité. Cependant, la recrudescence aujourd'hui observée dans les Pyrénées-Atlantiques met en exergue les difficultés et les contraintes qui peuvent faire obstacle au succès d'un plan d'éradication de l'agalactie contagieuse uniquement sanitaire. En effet, malgré le gain de nombreuses batailles, la guerre contre *M. agalactiae* n'a jamais été complètement gagnée dans ce département. L'actuel retour de la maladie prouve les limites de mesures draconiennes dont l'application stricte sur le long terme est difficile. Si la réussite des futurs programmes de lutte contre *M. agalactiae* passe immanquablement par des mesures sanitaires, ils devront également bénéficier d'avancées des mesures médicales pour les appuyer. L'intégration des particularités ou résistances politico-économique propres à chaque foyer d'agalactie contagieuse leur sera aussi indispensable. Un moyen de retracer plus finement les historiques de transmission de l'agent pathogène entre élevages et entre animaux constituerait un atout majeur pour améliorer les programmes de maîtrise de l'agalactie contagieuse et concentrer les efforts sur les points de contamination les plus critiques.

6. *Objectifs de l'étude*

6.1 Problématique posée par les mycoplasmoses de ruminants

Si la PPCB et la PPCC peuvent être considérées comme des fléaux des pays du Sud, il convient tout de même de maintenir une surveillance vis-à-vis de ces maladies dans les zones indemnes comme l'Amérique et l'Europe afin d'y éviter une émergence ou ré-émergence (Lefèvre, 2003 ; Thiaucourt, 2003 ; Poumarat *et al.*, 2004). En tout état de cause, les mycoplasmoses de ruminants sont, dans leur ensemble, des maladies chroniques responsables de pertes économiques significatives pour les productions bovines, ovines et caprines. En Europe et en Amérique du Nord, *M. bovis* est considéré comme l'espèce de mycoplasme responsable des pertes économiques les plus importantes chez les bovins (Gourlay *et al.*, 1989 ; Nicholas et Ayling, 2003 ; Thomas *et al.*, 2003). Les mycoplasmoses caprines représentent une des dominantes pathologiques infectieuses chez la chèvre. Chez les ovins, l'agalactie contagieuse à *M. agalactiae* sévit insidieusement sur tout le bassin méditerranéen. En France, le deuxième bassin ovin laitier (producteur entre autre des fromages d'appellation d'origine contrôlée "Ossau-Iraty") est sévèrement touché par cette maladie avec 95 nouveaux élevages déclarants pour l'année 2008 (GDS64, 2009) [Figure 6]. La lutte contre les mycoplasmes constitue donc un enjeu important pour les filières d'élevage de ruminants et implique des innovations dans les moyens de dépistage, de prévention et de traitement. Il est réellement nécessaire de poursuivre et d'approfondir les recherches et ce, plus particulièrement, sur les antigènes de surface, la diversité génétique et les mécanismes qui y conduisent. Les travaux réalisés au cours de cette thèse sur la diversité génétique de *M. agalactiae* s'inscrivent dans ce contexte.

6.2 Objectifs de l'étude

En réponse à la problématique posée par les mycoplasmoses de ruminants, ce travail a pour objectif de faire avancer les connaissances sur les systèmes d'hypervariabilité des protéines de surface, les échanges et le polymorphisme génétiques chez *M. agalactiae*. Cette description de la diversité génétique constitue un pré-requis indispensable à la définition de cibles moléculaires utilisables pour un suivi épidémiologique des infections par ce mycoplasme.

Au sein des mycoplasmes de ruminants, *M. agalactiae* et *M. bovis* sont phylogénétiquement très proches [Figure 5], la seule différence entre ces deux espèces pourrait presque être réduite à leur spécificité d'hôte, ainsi, ces travaux de nature fondamentale sur le modèle *M. agalactiae* participent également à l'avancée des connaissances sur *M. bovis*.

Plus précisément, les objectifs poursuivis étaient la caractérisation de l'étendue de la diversité génétique par l'étude d'une collection d'isolats disponibles au laboratoire et l'analyse comparative des génomes et des protéomes de deux isolats dont la souche type PG2.

6.3 Démarche adoptée

Dans un premier temps, notre travail s'est appliqué à l'évaluation de la diversité génétique au sein de l'espèce *M. agalactiae* (Premier chapitre). Dans cet objectif, nous avons réalisé une analyse de la variation génétique par PCR (analyse VNTR) et par *Southern blot* en passant au crible différentes sondes sur un panel aussi représentatif que possible de la diversité des souches circulantes et un panel d'isolats issus des Pyrénées-Atlantiques, zone d'enzootie de l'agalactie contagieuse. L'ensemble des 101 isolats étudiés est issu du réseau français d'épidémio-surveillance des mycoplasmoses de ruminants, VIGIMYC, aujourd'hui pilote par l'AFSSA de Lyon (Poumarat *et al.*, 2006).

Ensuite, nous nous sommes attaché à l'étude de l'étendue, la dynamique et l'évolution du système génétique hypervariable codant les Vpma, famille de protéines de surface de *M. agalactiae* (Deuxième chapitre). Pour cela, nous avons réalisé une analyse comparative fine de l'expression des gènes et des loci contenant les séquences d'ADN correspondantes entre la souche type de *M. agalactiae* PG2 clone 55.5 et une souche considérée comme atypique (souche 5632).

Enfin, nous avons entièrement séquencé et annoté la souche 5632. Accompagnées de données protéomiques, les analyses de génomique comparative ont ainsi été élargies à l'ensemble du génome de la souche type PG2 (Sirand-Pugnet *et al.*, 2007b) avec la souche 5632 (Troisième chapitre), choisie pour sa divergence et considérée comme aux antipodes de l'espèce de *M. agalactiae* par rapport à PG2.

Références

Abdo el M., Nicolet J., Miserez R., Goncalves R., Regalla J., Griot C., Bensaide A., Krampe M. et Frey J. (1998). Humoral and bronchial immune responses in cattle experimentally infected with *Mycoplasma mycoides* subsp. *mycoides* small colony type. *Vet Microbiol* **59:**109-122.

Abo W., Sawada Y., Ogawa S., Toyoguchi A. et Chiba S. (1988). Mesothelial cell proliferation and localized pleural pseudotumour associated with *Mycoplasma pneumoniae* infection. *Eur J Pediatr* **148:**62-63.

Adegboye D.S., Hallbur P.G., Cavanaugh D.L., Werdin R.E., Chase C.C., Miskimins D.W. et Rosenbusch R.F. (1995). Immunohistochemical and pathological study of *Mycoplasma bovis*-associated lung abscesses in calves. *J Vet Diagn Invest* **7:**333-337.

Alexander P.G., Slee K.J., McOrist S., Ireland L. et Coloe P.J. (1985). Mastitis in cows and polyarthritis and pneumonia in calves caused by *Mycoplasma species* bovine group 7. *Aust Vet J* **62:**135-136.

Allen J.L., Noormohammadi A.H. et Browning G.F. (2005). The vlhA loci of *Mycoplasma synoviae* are confined to a restricted region of the genome. *Microbiology* **151:**935-940.

Ames T.R. (1997). Dairy calf pneumonia. The disease and its impact. *Vet Clin North Am Food Anim Pract* **13:**379-391.

Andreev J., Borovsky Z., Rosenshine I. et Rottem S. (1995). Invasion of HeLa cells by *Mycoplasma penetrans* and the induction of tyrosine phosphorylation of a 145-kDa host cell protein. *FEMS Microbiol Lett* **132:**189-194.

Apley M. (2006). Bovine respiratory disease: pathogenesis, clinical signs, and treatment in lightweight calves. *Vet Clin North Am Food Anim Pract* **22:**399-411.

Apley M.D. et Fajt V.R. (1998). Feedlot therapeutics. *Vet Clin North Am Food Anim Pract* **14:**291-313.

Assuncao P., De la Fe C., Ramirez A.S., Andrada M. et Poveda J.B. (2004). Serological study of contagious agalactia in herds of goats in the Canary Islands. *Vet Rec* **154:**684-687.

Aula P. et Nichols W.W. (1967). The cytogenetic effects of mycoplasma in human leukocyte cultures. *J Cell Physiol* **70:**281-290.

Auliac P. (2008). Mortalité chez les bouquetins des Alpes, Chasser en Savoie, vol. 93.

Barre A., de Daruvar A. et Blanchard A. (2004). MolliGen, a database dedicated to the comparative genomics of *Mollicutes*. *Nucleic Acids Res* **32:**D307-310.

Barthe M.-N., Revelli P., Game Y., Arcangioli M.-A., Solsona M., Tardy F. et Poumarat F. (2006). *Mycoplasmes et pneumonies chez le bouquetin des Alpes*. Presented at the Coordination Nationale de lutte Agalactie Contagieuse et mycoplasmoses des petits ruminants, Toulouse (France), 21 novembre 2006.

Baseman J.B., Cole R.M., Krause D.C. et Leith D.K. (1982). Molecular basis for cytadsorption of *Mycoplasma pneumoniae*. *J Bacteriol* **151:**1514-1522.

Baseman J.B., Lange M., Criscimagna N.L., Giron J.A. et Thomas C.A. (1995). Interplay between mycoplasmas and host target cells. *Microb Pathog* **19:**105-116.

Baseman J.B. et Tully J.G. (1997). Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. *Emerg Infect Dis* **3:**21-32.

Behrens A., Heller M., Kirchhoff H., Yogev D. et Rosengarten R. (1994). A family of phase- and size-variant membrane surface lipoprotein antigens (Vsps) of *Mycoplasma bovis*. *Infect Immun* **62:**5075-5084.

Bendjennat M., Blanchard A., Loutfi M., Montagnier L. et Bahraoui E. (1999). Role of *Mycoplasma penetrans* endonuclease P40 as a potential pathogenic determinant. *Infect Immun* **67:**4456-4462.

Bergonier D. (1996). Etude de la variabilité intraspécifique de *Mycoplasma agalactiae*. Bases pour l'amélioration du diagnostic de l'Agalactie Contagieuse. Université Claude Bernard, Lyon (France).

Bergonier D. et Berthelot X. (2008). Mycoplasmoses des petits ruminants : Le syndrome de l'agalactie contagieuse. *Bulletin de l'Académie vétérinaire de France* **161:**167-177.

Bergonier D. et Poumarat F. (1996). Agalactie contagieuse des petits ruminants: epidemiologie, diagnostic et contrôle. *Rev Sci Tech* **15:**1431-1475.

Bergonier D. et Thiaucourt F. (2003). Agalactie contagieuse des petits ruminants, p. 809-825. *In* L. P.-C., J. Blancou and R. Chermette (ed.), Principales maladies infectieuses et parasitaires du bétail. Europe et régions chaudes. Maladies bactériennes, mycoses, maladies parasitaires., vol. 2. *Lavoisier Tec et Doc*, Paris (France).

Bergonier D., van de Wiele A., Marenda M. et Berthelot X. (2002). Syndromes mycoplasmiques : L'agalactie contagieuse chez les ovins. *Le Point Vétérinaire* **33:**52-57.

Bhugra B., Voelker L.L., Zou N., Yu H. et Dybvig K. (1995). Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. *Mol Microbiol* **18:**703-714.

Bischof D.F., Janis C., Vilei E.M., Bertoni G. et Frey J. (2008). Cytotoxicity of *Mycoplasma mycoides* subsp. *mycoides* small colony type to bovine epithelial cells. *Infect Immun* **76:**263-269.

Bischof D.F., Vilei E.M. et Frey J. (2009). Functional and antigenic properties of GlpO from *Mycoplasma mycoides* subsp. *mycoides* SC: characterization of a flavin adenine dinucleotide-binding site deletion mutant. *Vet Res* **40:**35.

Blanchard A. et Bébéar C.M. (2002). Mycoplasmas of Humans, p. 45–71. *In* S. Razin and R. Herrmann (ed.), Molecular biology and pathogenicity of *Mycoplasmas*. *Kluwer Academic / Plenum Publishers*, New York (USA).

Boden W.E., Funk E.J., Carleton R.A., Benham I., Khan A.H., Lasser A. et McEnany M.T. (1983). Left ventricular hemangioma masquerading as *Mycoplasma pericarditis*. *Am Heart J* **106:**771-774.

Boguslavsky S., Menaker D., Lysnyansky I., Liu T., Levisohn S., Rosengarten R., Garcia M. et Yogev D. (2000). Molecular characterization of the *Mycoplasma gallisepticum* pvpA gene which encodes a putative variable cytadhesin protein. *Infect Immun* **68:**3956-3964.

Bonnefin N. (2003). Comparaison de la souche 7784 hypovirulente de *Mycoplasma agalactiae* à la souche type virulente PG2 par la technique d'hybridation soustractive suppressive. *Th.Vet.* Université Paul Sabatier - Ecole Nationale Vétérinaire de Toulouse, Toulouse (France).

Borovsky Z., Tarshis M., Zhang P. et Rottem S. (1998). Protein kinase C activation and vacuolation in HeLa cells invaded by *Mycoplasma penetrans*. *J Med Microbiol* **47:**915-922.

Branda S.S., Vik S., Friedman L. et Kolter R. (2005). Biofilms: the matrix revisited. *Trends Microbiol* **13:**20-26.

Brandao E. (1995). Isolation and identification of *Mycoplasma mycoides* subspecies *mycoides* SC strains in sheep and goats. *Vet Rec* **136:**98-99.

Bridré J. et Donatien A. (1923). Le microbe de l'agalactie contagieuse et sa culture *in vitro*. *Compte Rendu de l'Académie des Sciences* **177:**841-843.

Bridré J. et Donatien A. (1925). Le microbe de l'agalaxie du mouton et de la chèvre. *Annales de l'Institut Pasteur* **29:**925-951.

Brown D.R., Farley J.M., Zacher L.A., Carlton J.M., Clippinger T.L., Tully J.G. et Brown M.B. (2001a). Mycoplasma alligatoris sp. nov., from American alligators. *Int J Syst Evol Microbiol* **51:**419-424.

Brown D.R., Nogueira M.F., Schoeb T.R., Vliet K.A., Bennett R.A., Pye G.W. et Jacobson E.R. (2001b). Pathology of experimental mycoplasmosis in American alligators. *J Wildl Dis* **37:**671-679.

Brown D.R., Zacher L.A. et Farmerie W.G. (2004). Spreading factors of Mycoplasma alligatoris, a flesh-eating mycoplasma. *J Bacteriol* **186:**3922-3927.

Butler J.A., Sickles S.A., Johanns C.J. et Rosenbusch R.F. (2000). Pasteurization of discard mycoplasma mastitic milk used to feed calves: thermal effects on various mycoplasma. *J Dairy Sci* **83:**2285-2288.

Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F., Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P. et Blanchard A. (2001). The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. *Nucleic Acids Res* **29:**2145-2153.

Chavez Gonzalez Y.R., Ros Bascunana C., Bolske G., Mattsson J.G., Fernandez Molina C. et Johansson K.E. (1995). In vitro amplification of the 16S rRNA genes from *Mycoplasma bovis* and *Mycoplasma agalactiae* by PCR. *Vet Microbiol* **47:**183-190.

Chopra-Dewasthaly R., Citti C., Glew M.D., Zimmermann M., Rosengarten R. et Jechlinger W. (2008). Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation. *Mol Microbiol* **67:**1196-1210.

Cimolai N. (2001). Do mycoplasmas cause human cancer? *Can J Microbiol* **47:**691-697.

Citti C. (2006). Les mycoplasmes : stratégie d'adaptation et de persistance de bactéries minimales. *Le Nouveau Praticien Vétérinaire élevages et santé* **Août/Septembre/Octobre:**203-209.

Citti C., Browning G.F. et Rosengarten R. (2005). Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas., p. 439–484. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: Molecular biology, pathogenicity and strategies for control. *Horizon Bioscience*, Norfolk (United Kingdom).

Citti C., Kim M.F. et Wise K.S. (1997). Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth-inhibiting host antibodies. *Infect Immun* **65:**1773-1785.

Citti C. et Wise K.S. (1995). *Mycoplasma hyorhinis vlp* gene transcription: critical role in phase variation and expression of surface lipoproteins. *Mol Microbiol* **18:**649-660.

Clippinger T.L., Bennett R.A., Johnson C.M., Vliet K.A., Deem S.L., Oros J., Jacobson E.R., Schumacher I.M., Brown D.R. et Brown M.B. (2000). Morbidity and mortality associated with a new mycoplasma species from captive American alligators (Alligator mississippiensis). *J Zoo Wildl Med* **31:**303-314.

Clyde W.A., Jr. (1964). Mycoplasma Species Identification Based Upon Growth Inhibition by Specific Antisera. *J Immunol* **92:**958-965.

Cole B.C., Knudtson K.L., Oliphant A., Sawitzke A.D., Pole A., Manohar M., Benson L.S., Ahmed E. et Atkin C.L. (1996). The sequence of the *Mycoplasma arthritidis* superantigen, MAM: identification of functional domains and comparison with microbial superantigens and plant lectin mitogens. *J Exp Med* **183:**1105-1110.

Connole M.D., Laws L. et Hart R.K. (1967). Mastitis in cattle caused by a *Mycoplasma sp. Aust Vet J* **43:**157-162.

Costerton J.W., Geesey G.G. et Cheng K.J. (1978). How bacteria stick. *Sci Am* **238:**86-95.

Costerton J.W., Lewandowski Z., Caldwell D.E., Korber D.R. et Lappin-Scott H.M. (1995). Microbial biofilms. *Annu Rev Microbiol* **49:**711-745.

Costerton J.W., Stewart P.S. et Greenberg E.P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science* **284:**1318-1322.

Dallo S.F., Kannan T.R., Blaylock M.W. et Baseman J.B. (2002). Elongation factor Tu and E1 beta subunit of pyruvate dehydrogenase complex act as fibronectin binding proteins in *Mycoplasma pneumoniae*. *Mol Microbiol* **46:**1041-1051.

Dedieu L., Chapey E. et Balcer-Rodrigues V. (2005). *Mycoplasma mycoides* ssp. *mycoides* biotype small colony-secreted components induce apoptotic cell death in bovine leucocytes. *Scand J Immunol* **62:**528-538.

Dedieu L., Mady V. et Lefevre P.C. (1995). Development of two PCR assays for the identification of mycoplasmas causing contagious agalactia. *FEMS Microbiol Lett* **129:**243-249.

Denison A.M., Clapper B. et Dybvig K. (2005). Avoidance of the host immune system through phase variation in *Mycoplasma pulmonis*. *Infect Immun* **73:**2033-2039.

Donlan R.M. et Costerton J.W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* **15:**167-193.

Dujardin-Beaumetz E. (1900). Etude bactériologique d'un microorganisme à la limite de la visibilité, le microbe de la péripneumonie et sa culture. *Th.Med.*, Paris.

Dujardin-Beaumetz E. (1906). Transmission de la Péripneumonie des Bovidés aux espèces Ovine et Caprine. *Annales de l'Institut Pasteur* **20:**449-466.

Dybvig K., Sitaraman R. et French C.T. (1998). A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. *Proc Natl Acad Sci U S A* **95:**13923-13928.

Dziewanowska K., Carson A.R., Patti J.M., Deobald C.F., Bayles K.W. et Bohach G.A. (2000). Staphylococcal fibronectin binding protein interacts with heat shock protein 60 and integrins: role in internalization by epithelial cells. *Infect Immun* **68:**6321-6328.

Dziewanowska K., Patti J.M., Deobald C.F., Bayles K.W., Trumble W.R. et Bohach G.A. (1999). Fibronectin binding protein and host cell tyrosine kinase are required for internalization of *Staphylococcus aureus* by epithelial cells. *Infect Immun* **67:**4673-4678.

Edward D.G., Freundt E.A., Chanock R.M., Fabricant J., Hayflick L., Lemcke R.M., Rezin S., Somerson N.L. et Wittler R.G. (1967). Recommendations on nomenclature of the order *Mycoplasmatales*. *Science* **155:**1694-1696.

Euzéby J.P. (1998). Dictionnaire de Bactériologie Vétérinaire [en ligne]. Mise à jour du 30 juin 2009, consultée le 21 juillet 2009. Société de Bactériologie Systématique et Vétérinaire. Disponible sur: http://www.bacterio.cict.fr

Feng S.H. et Lo S.C. (1994). Induced mouse spleen B-cell proliferation and secretion of immunoglobulin by lipid-associated membrane proteins of *Mycoplasma fermentans* incognitus and *Mycoplasma penetrans*. *Infect Immun* **62:**3916-3921.

Feng S.H., Tsai S., Rodriguez J. et Lo S.C. (1999). Mycoplasmal infections prevent apoptosis and induce malignant transformation of interleukin-3-dependent 32D hematopoietic cells. *Mol Cell Biol* **19:**7995-8002.

Fitzmaurice J., Sewell M., King C.M., McDougall S., McDonald W.L. et O'Keefe J.S. (2008). A real-time polymerase chain reaction assay for the detection of *Mycoplasma agalactiae*. *N Z Vet J* **56:**233-236.

Flitman-Tene R., Mudahi-Orenstein S., Levisohn S. et Yogev D. (2003). Variable lipoprotein genes of *Mycoplasma agalactiae* are activated in vivo by promoter addition via site-specific DNA inversions. *Infect Immun* **71:**3821-3830.

Fogh J. et Fogh H. (1965). Chromosome Changes in Pplo-Infected Fl Human Amnion Cells. *Proc Soc Exp Biol Med* **119:**233-238.

Freundt E.A. (1983). Culture media for classic mycoplasmas., p. 127-135. *In* J. G. Tully (ed.), Methods in mycoplasmology, vol. 1. *Academic Press.*, New York.

Frey J. (2002). Mycoplasmas of Animals, p. 73–90. *In* S. Razin and R. Herrmann (ed.), Molecular biology and pathogenicity of *Mycoplasmas*. *Kluwer Academic / Plenum Publishers*, New York (USA).

Fusco M., Corona L., Onni T., Marras E., Longheu C., Idini G. et Tola S. (2007). Development of a sensitive and specific enzyme-linked immunosorbent assay based on recombinant antigens for rapid detection of antibodies against *Mycoplasma agalactiae* in sheep. *Clin Vaccine Immunol* **14:**420-425.

Gagea M.I., Bateman K.G., Shanahan R.A., van Dreumel T., McEwen B.J., Carman S., Archambault M. et Caswell J.L. (2006). Naturally occurring *Mycoplasma bovis*-associated pneumonia and polyarthritis in feedlot beef calves. *J Vet Diagn Invest* **18:**29-40.

Galisson C. (2001). Evaluation clinique du pouvoir pathogène d'une souche réputée avirulente de *M. agalactiae*. *Th.Vet.* Université Paul Sabatier - Ecole Nationale Vétérinaire de Toulouse, Toulouse (France).

Gaurivaud P., Persson A., Grand D.L., Westberg J., Solsona M., Johansson K.E. et Poumarat F. (2004). Variability of a glucose phosphotransferase system permease in *Mycoplasma mycoides* subsp. *mycoides* Small Colony. *Microbiology* **150:**4009-4022.

GDS64. (2009). Agalactie contagieuse ovine, p. 11-14, Rapport annuel. *Groupement de défense sanitaire du Béarn et du Pays-Basque*, Pau (France).

Geary S.J., Tourtellotte M.E. et Cameron J.A. (1981). Inflammatory toxin from *Mycoplasma bovis*: isolation and characterization. *Science* **212:**1032-1033.

Gerlic M., Horowitz J., Farkash S. et Horowitz S. (2007). The inhibitory effect of *Mycoplasma fermentans* on tumour necrosis factor (TNF)-alpha-induced apoptosis resides in the membrane lipoproteins. *Cell Microbiol* **9:**142-153.

Gerlic M., Horowitz J. et Horowitz S. (2004). *Mycoplasma fermentans* inhibits tumor necrosis factor alpha-induced apoptosis in the human myelomonocytic U937 cell line. *Cell Death Differ* **11:**1204-1212.

Giron J.A., Lange M. et Baseman J.B. (1996). Adherence, fibronectin binding, and induction of cytoskeleton reorganization in cultured human cells by *Mycoplasma penetrans*. *Infect Immun* **64:**197-208.

Glew M.D., Baseggio N., Markham P.F., Browning G.F. et Walker I.D. (1998). Expression of the pMGA genes of *Mycoplasma gallisepticum* is controlled by variation in the GAA trinucleotide repeat lengths within the 5' noncoding regions. *Infect Immun* **66:**5833-5841.

Glew M.D., Marenda M., Rosengarten R. et Citti C. (2002). Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus. *J Bacteriol* **184:**5987-5998.

Gonzalez-Candela M., Cubero-Pablo M.J., Martin-Atance P. et Leon-Vizcaino L. (2006). Potential pathogens carried by Spanish ibex (*Capra pyrenaica hispanica*) in southern Spain. *J Wildl Dis* **42:**325-334.

Gonzalez-Candela M., Verbisck-Bucker G., Martin-Atance P., Cubero-Pablo M.J. et Leon-Vizcaino L. (2007). Mycoplasmas isolated from Spanish ibex (*Capra pyrenaica hispanica*): frequency and risk factors. *Vet Rec* **161:**167-168.

Gourlay R.N., Howard C.J., Thomas L.H. et Wyld S.G. (1979). Pathogenicity of some *Mycoplasma* and *Acholeplasma species* in the lungs of gnotobiotic calves. *Res Vet Sci* **27:**233-237.

Gourlay R.N., Thomas L.H. et Wyld S.G. (1989). Increased severity of calf pneumonia associated with the appearance of *Mycoplasma bovis* in a rearing herd. *Vet Rec* **124:**420-422.

Grace J.T., Jr., Horoszewicz J.S., Stim T.B., Mirand E.A. et James C. (1965). Mycoplasmas (PPLO) and human leukemia and lymphoma. *Cancer* **18:**1369-1376.

Greco G., Corrente M., Martella V., Pratelli A. et Buonavoglia D. (2001). A multiplex-PCR for the diagnosis of contagious agalactia of sheep and goats. *Mol Cell Probes* **15:**21-25.

Hale H.H., Helmboldt C.F., Plastridge W.N. et Stula E.F. (1962). Bovine mastitis caused by a *Mycoplasma species*. *Cornell Vet* **52:**582-591.

Hamsten C., Westberg J., Bolske G., Ayling R., Uhlen M. et Persson A. (2008). Expression and immunogenicity of six putative variable surface proteins in *Mycoplasma mycoides* subsp. *mycoides* SC. *Microbiology* **154:**539-549.

Hayflick L. (1965). Mycoplasmas and human leukemia. *Wistar Inst Symp Monogr* **4:**157-165.

Henderson J.P. et Ball H.J. (1999). Polyarthritis due to *Mycoplasma bovis* infection in adult dairy cattle in Northern Ireland. *Vet Rec* **145:**374-376.

Horino A., Kenri T., Sasaki Y., Okamura N. et Sasaki T. (2009). Identification of a site-specific tyrosine recombinase that mediates promoter inversions of phase-variable mpl lipoprotein genes in *Mycoplasma penetrans*. *Microbiology* **155:**1241-1249.

Horino A., Sasaki Y., Sasaki T. et Kenri T. (2003). Multiple promoter inversions generate surface antigenic variation in *Mycoplasma penetrans*. *J Bacteriol* **185:**231-242.

Hum S., Kessell A., Djordjevic S., Rheinberger R., Hornitzky M., Forbes W. et Gonsalves J. (2000). Mastitis, polyarthritis and abortion caused by *Mycoplasma species* bovine group 7 in dairy cattle. *Aust Vet J* **78:**744-750.

Hunt M.E. et Brown D.R. (2005). Mycoplasma alligatoris infection promotes CD95 (FasR) expression and apoptosis of primary cardiac fibroblasts. *Clin Diagn Lab Immunol* **12:**1370-1377.

Hunt M.E. et Brown D.R. (2007). Role of sialidase in Mycoplasma alligatoris-induced pulmonary fibroblast apoptosis. *Vet Microbiol* **121:**73-82.

Hutcheon D. (1881). Contagious pleuro-pneumonia in angora goats. *Veterinary Journal* **13:**171-180.

Illing K. (1979). [Clinical picture of experimental *Mycoplasma bovis* mastitis in cattle]. *Arch Exp Veterinarmed* **33:**945-948.

Jan G., Le Henaff M., Fontenelle C. et Wroblewski H. (2001). Biochemical and antigenic characterisation of *Mycoplasma gallisepticum* membrane proteins P52 and P67 (pMGA). *Arch Microbiol* **177:**81-90.

Jansson E., Wegelius R., Hakkarainen K. et Miettinen A. (1991). Mycoplasmas and leukaemia. *Res Virol* **142:**333.

Jeffery N., Browning G.F. et Noormohammadi A.H. (2006). Organization of the *Mycoplasma synoviae* WVU 1853T *vlhA* gene locus. *Avian Pathol* **35:**53-57.

Johansson K.E. et Pettersson B. (2002). Taxonomy of *Mollicutes*, p. 1–29. *In* S. Razin and R. Herrmann (ed.), Molecular biology and pathogenicity of Mycoplasmas. *Kluwer Academic / Plenum Publishers*, New York (USA).

Johnson C., Kannan T.R. et Baseman J.B. (2009). Characterization of a unique ADP-ribosyltransferase of Mycoplasma penetrans. *Infect Immun*.

Kannan T.R. et Baseman J.B. (2006). ADP-ribosylating and vacuolating cytotoxin of *Mycoplasma pneumoniae* represents unique virulence determinant among bacterial pathogens. *Proc Natl Acad Sci U S A* **103:**6724-6729.

Kannan T.R., Provenzano D., Wright J.R. et Baseman J.B. (2005). Identification and characterization of human surfactant protein A binding protein of *Mycoplasma pneumoniae*. *Infect Immun* **73:**2828-2834.

Kotani H. et McGarrity G.J. (1985). Rapid and simple identification of mycoplasmas by immunobinding. *J Immunol Methods* **85:**257-267.

Krause D.C., Leith D.K., Wilson R.M. et Baseman J.B. (1982). Identification of *Mycoplasma pneumoniae* proteins associated with hemadsorption and virulence. *Infect Immun* **35:**809-817.

Krueger K.M. et Barbieri J.T. (1995). The family of bacterial ADP-ribosylating exotoxins. *Clin Microbiol Rev* **8:**34-47.

LaFaunce N.A. et McEntee K. (1982). Experimental *Mycoplasma bovis* seminal vesiculitis in the bull. *Cornell Vet* **72:**150-167.

Lambert M. (1987). Contagious agalactia of sheep and goats. *Rev Sci Tech* **6:**699-711.

Lambert M., Calamel M., Dufour P., Cabasse E., Vitu C. et Pepin V. (1998). Detection of false-positive sera in contagious agalactia with a multiantigen ELISA and their elimination with a protein G conjugate. *J Vet Diagn Invest* **10:**326-330.

Langford E.V. (1975). Mycoplasma species recovered from the reproductive tracts of western Canadian cows. *Can J Comp Med* **39:**133-138.

Le Grand D., Arcangioli M.-A., Calavas D., Bézille P. et Poumarat F. (2008). Mycoplasmes et mycoplasmoses bovines : actualités. *Bulletin de l'Académie vétérinaire de France* **161:**159-166.

Le Grand D., Poumarat F. et Bézille P. (1996). Mycoplasmoses bovines à *Mycoplasma bovis*. *Le Point Vétérinaire* **80:**771–778.

Leach R.H., Erno H. et MacOwan K.J. (1993). Proposal for designation of F38-type caprine mycoplasmas as *Mycoplasma capricolum* subsp. *capripneumoniae* subsp. nov. and consequent obligatory relegation of strains currently classified as *M. capricolum* (Tully, Barile, Edward, Theodore, and Erno 1974) to an additional new subspecies, *M. capricolum* subsp. *capricolum* subsp. nov. *Int J Syst Bacteriol* **43:**603-605.

Lefèvre P.-C. (2003). Péripneumonie contagieuse bovine, p. 809-825. *In* L. P.-C., J. Blancou and R. Chermette (ed.), Principales maladies infectieuses et parasitaires du bétail. Europe et régions chaudes. Maladies bactériennes, mycoses, maladies parasitaires., vol. 2. *Lavoisier Tec et Doc*, Paris (France).

Letunic I. et Bork P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23:**127-128.

Lo S.C. (2002). Apoptotic, antiapoptotic, clastogenic and oncogenic effects, p. 403-416. *In* S. Razin and R. Herrmann (ed.), Molecular biology and pathogenicity of Mycoplasmas. *Kluwer Academic / Plenum Publishers*, New York (USA).

Lo S.C., Hayes M.M., Kotani H., Pierce P.F., Wear D.J., Newton P.B., 3rd, Tully J.G. et Shih J.W. (1993). Adhesion onto and invasion into mammalian cells by mycoplasma penetrans: a newly isolated mycoplasma from patients with AIDS. *Mod Pathol* **6:**276-280.

Lorusso A., Decaro N., Greco G., Corrente M., Fasanella A. et Buonavoglia D. (2007). A real-time PCR assay for detection and quantification of *Mycoplasma agalactiae* DNA. *J Appl Microbiol* **103:**918-923.

Lysnyansky I., Ron Y., Sachse K. et Yogev D. (2001a). Intrachromosomal recombination within the *vsp* locus of *Mycoplasma bovis* generates a chimeric variable surface lipoprotein antigen. *Infect Immun* **69:**3703-3712.

Lysnyansky I., Ron Y. et Yogev D. (2001b). Juxtaposition of an active promoter to *vsp* genes via site-specific DNA inversions generates antigenic variation in *Mycoplasma bovis*. *J Bacteriol* **183:**5698-5708.

Lysnyansky I., Rosengarten R. et Yogev D. (1996). Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. *J Bacteriol* **178:**5395-5401.

Lysnyansky I., Sachse K., Rosenbusch R., Levisohn S. et Yogev D. (1999). The *vsp* locus of *Mycoplasma bovis*: gene organization and structural features. *J Bacteriol* **181:**5734-5741.

MacOwan K.J. et Minette J.E. (1976). A mycoplasma from acute contagious caprine pleuropneumonia in Kenya. *Trop Anim Health Prod* **8:**91-95.

MacOwan K.J. et Minette J.E. (1977). The role of Mycoplasma strain F38 in contagious caprine pleuropneumonia (CCCP) in Kenya. *Vet Rec* **101:**380-381.

Macpherson I. et Russell W. (1966). Transformations in hamster cells mediated by mycoplasmas. *Nature* **210:**1343-1345.

Manso-Silvan L., Vilei E.M., Sachse K., Djordjevic S.P., Thiaucourt F. et Frey J. (2009). *Mycoplasma leachii* sp. nov. as a new species designation for *Mycoplasma sp.* bovine group 7 of Leach, and reclassification of *Mycoplasma mycoides* subsp. *mycoides* LC as a serovar of *Mycoplasma mycoides* subsp. *capri*. *Int J Syst Evol Microbiol* **59:**1353-1358.

Marenda M.S., Sagne E., Poumarat F. et Citti C. (2005). Suppression subtractive hybridization as a basis to assess *Mycoplasma agalactiae* and *Mycoplasma bovis* genomic diversity and species-specific sequences. *Microbiology* **151:**475-489.

Markham P.F., Duffy M.F., Glew M.D. et Browning G.F. (1999). A gene family in *Mycoplasma imitans* closely related to the *pMGA* family of *Mycoplasma gallisepticum*. *Microbiology* **145 (Pt 8):**2095-2103.

Maunsell F.P. et Donovan G.A. (2009). *Mycoplasma bovis* Infections in young calves. *Vet Clin North Am Food Anim Pract* **25:**139-177, vii.

Maunsell F.P., Donovan G.A., Risco C. et Brown M.B. (2009). Field evaluation of a *Mycoplasma bovis* bacterin in young dairy calves. *Vaccine* **27:**2781-2788.

McAuliffe L., Ayling R.D., Ellis R.J. et Nicholas R.A. (2008). Biofilm-grown *Mycoplasma mycoides* subsp. *mycoides* SC exhibit both phenotypic and genotypic variation compared with planktonic cells. *Vet Microbiol* **129:**315-324.

McAuliffe L., Ellis R.J., Lawes J.R., Ayling R.D. et Nicholas R.A. (2005). 16S rDNA PCR and denaturing gradient gel electrophoresis; a single generic test for detecting and differentiating *Mycoplasma species*. *J Med Microbiol* **54:**731-739.

McAuliffe L., Ellis R.J., Miles K., Ayling R.D. et Nicholas R.A. (2006). Biofilm formation by mycoplasma species and its role in environmental persistence and survival. *Microbiology* **152:**913-922.

Mernaugh G.R., Dallo S.F., Holt S.C. et Baseman J.B. (1993). Properties of adhering and nonadhering populations of *Mycoplasma genitalium*. *Clin Infect Dis* **17 Suppl 1:**S69-78.

Miles R.J. (1992). Catabolism in mollicutes. *J Gen Microbiol* **138:**1773-1783.

Murphy W.H., Furtado D. et Plata E. (1965). Possible Association between Leukemia in Children and Virus-Like Agents. *Jama* **191:**110-115.

Negroni G. (1964). Isolation of Viruses from Leukaemic Patients. *Br Med J* **1:**927-929.

Neyrolles O., Chambaud I., Ferris S., Prevost M.C., Sasaki T., Montagnier L. et Blanchard A. (1999). Phase variations of the *Mycoplasma penetrans* main surface lipoprotein increase antigenic diversity. *Infect Immun* **67:**1569-1578.

Nicholas R.A. et Ayling R.D. (2003). *Mycoplasma bovis*: disease, diagnosis, and control. *Res Vet Sci* **74:**105-112.

Nocard E. et Roux E. (1898). Le microbe de la Péripneumonie. *Annales de l'Institut Pasteur* **12:**240-262.

Noormohammadi A.H., Markham P.F., Duffy M.F., Whithear K.G. et Browning G.F. (1998). Multigene families encoding the major hemagglutinins in phylogenetically distinct mycoplasmas. *Infect Immun* **66:**3470-3475.

Noormohammadi A.H., Markham P.F., Kanci A., Whithear K.G. et Browning G.F. (2000). A novel mechanism for control of antigenic variation in the haemagglutinin gene family of *Mycoplasma synoviae*. *Mol Microbiol* **35:**911-923.

Nussbaum S., Lysnyansky I., Sachse K., Levisohn S. et Yogev D. (2002). Extended repertoire of genes encoding variable surface lipoproteins in *Mycoplasma bovis* strains. *Infect Immun* **70:**2220-2225.

Office International des Epizooties. (2009). World Animal Health Information Database (WAHID) [en ligne]. Mise à jour du 26 mars 2009, consultée le 21 juillet 2009. Disponible sur: http://www.oie.int/wahis

Oravcova K., Lopez-Enriquez L., Rodriguez-Lazaro D. et Hernandez M. (2009). *Mycoplasma agalactiae p40* gene, a novel marker for diagnosis of contagious agalactia in sheep by real-time PCR: assessment of analytical performance and in-house validation using naturally contaminated milk samples. *J Clin Microbiol* **47:**445-450.

Paddenberg R., Weber A., Wulf S. et Mannherz H.G. (1998). Mycoplasma nucleases able to induce internucleosomal DNA degradation in cultured cells possess many characteristics of eukaryotic apoptotic nucleases. *Cell Death Differ* **5:**517-528.

Pallen M.J., Lam A.C., Loman N.J. et McBride A. (2001). An abundance of bacterial ADP-ribosyltransferases--implications for the origin of exotoxins and their human homologues. *Trends Microbiol* **9:**302-307; discussion 308.

Papazisi L., Frasca S., Jr., Gladd M., Liao X., Yogev D. et Geary S.J. (2002). GapA and CrmA coexpression is essential for *Mycoplasma gallisepticum* cytadherence and virulence. *Infect Immun* **70:**6839-6845.

Park S.H., Choe G.Y., Kim C.W., Chi J.G. et Sung S.H. (1990). Inflammatory pseudotumor of the lung in a child with mycoplasma pneumonia. *J Korean Med Sci* **5:**213-223.

Parsek M.R. et Singh P.K. (2003). Bacterial biofilms: an emerging link to disease pathogenesis. *Annu Rev Microbiol* **57:**677-701.

Paton G.R., Jacobs J.P. et Perkins F.T. (1965). Chromosome changes in human diploid-cell cultures infected with Mycoplasma. *Nature* **207:**43-45.

Pepin M., Dufour P., Lambert M., Aubert M., Valognes A., Rotis T., Van de Wiele A. et Bergonier D. (2003). Comparison of three enzyme-linked immunosorbent assays for serologic diagnosis of contagious agalactia in sheep. *J Vet Diagn Invest* **15:**281-285.

Persson A., Jacobsson K., Frykberg L., Johansson K.E. et Poumarat F. (2002). Variable surface protein Vmm of *Mycoplasma mycoides* subsp. *mycoides* small colony type. *J Bacteriol* **184:**3712-3722.

Pettersson B., Uhlen M. et Johansson K.E. (1996). Phylogeny of some mycoplasmas from ruminants based on 16S rRNA sequences and definition of a new cluster within the hominis group. *Int J Syst Bacteriol* **46:**1093-1098.

Pfutzner H., Kielstein P., Martin J. et Schimmel D. (1983). [Mycoplasma infection of calves. 2. Experimental infection of calves with *Mycoplasma bovis*]. *Arch Exp Veterinarmed* **37:**445-451.

Pfutzner H. et Sachse K. (1996). *Mycoplasma bovis* as an agent of mastitis, pneumonia, arthritis and genital disorders in cattle. *Rev Sci Tech* **15:**1477-1494.

Pilo P., Vilei E.M., Peterhans E., Bonvin-Klotz L., Stoffel M.H., Dobbelaere D. et Frey J. (2005). A metabolic enzyme as a primary virulence factor of *Mycoplasma mycoides* subsp. *mycoides* small colony. *J Bacteriol* **187:**6824-6831.

Polak-Vogelzang A.A., Hagenaars R. et Nagel J. (1978). Evaluation of an indirect immunoperoxidase test for identification of acholeplasma and mycoplasma. *J Gen Microbiol* **106:**241-249.

Pollack J.D. (1995). Methods for testing metabolic activities in *Mollicutes*, p. 277-286. *In* J. G. Tully (ed.), Molecular and Diagnostic Procedures in Mycoplasmology: Molecular Characterization. *Academic Press.*, San Diego.

Poumarat F. (2008). *Mycoplasmoses des ruminants: bilan du réseau Vigimyc.* Presented at the Coordination Nationale de lutte Agalactie Contagieuse et mycoplasmoses des petits ruminants, Toulouse (France), 20 février 2008, *AFSSA Lyon*.

Poumarat F., Dernburg A., Le Grand D., Arcangioli M.-A. et Calavas D. (2004). La Péripneumonie Contagieuse Bovine, une maladie à ne pas oublier. *Bulletin des GTV* **26:**17-22.

Poumarat F., Le Grand D., Mercier P., Tardy F., Gaurivaud P. et Calavas D. (2006). VIGIMYC : le réseau français d'épidémio-surveillance des mycoplasmoses des ruminants, bilan 2003-2005. *Le Nouveau Praticien Vétérinaire élevages et santé* **Août/Septembre/Octobre:**22-26.

Poumarat F., Longchambon D. et Martel J.L. (1992). Application of dot immunobinding on membrane filtration (MF dot) to the study of relationships within "M. mycoides cluster" and within "glucose and arginine-negative cluster" of ruminant mycoplasmas. *Vet Microbiol* **32:**375-390.

Poumarat F., Perrin B. et Longchambon D. (1991). Identification of ruminant mycoplasmas by dot immunobinding on membrane filtration (MF dot). *Vet Microbiol* **29:**329-338.

Raad I., Costerton W., Sabharwal U., Sacilowski M., Anaissie E. et Bodey G.P. (1993). Ultrastructural analysis of indwelling vascular catheters: a quantitative relationship between luminal colonization and duration of placement. *J Infect Dis* **168:**400-407.

Razin S. (1985). Molecular biology and genetics of mycoplasmas (*Mollicutes*). *Microbiol Rev* **49:**419-455.

Razin S. et Herrmann R. (2002). Molecular biology and pathogenicity of Mycoplasmas. *Kluwer Academic / Plenum Publishers*, New York (USA).

Razin S. et Jacobs E. (1992). Mycoplasma adhesion. *J Gen Microbiol* **138:**407-422.

Rodriguez F., Bryson D.G., Ball H.J. et Forster F. (1996). Pathological and immunohistochemical studies of natural and experimental *Mycoplasma bovis* pneumonia in calves. *J Comp Pathol* **115:**151-162.

Rodriguez F., Sarradell J., Poveda J.B., Ball H.J. et Fernandez A. (2000). Immunohistochemical characterization of lung lesions induced experimentally by *Mycoplasma agalactiae* and *Mycoplasma bovis* in goats. *J Comp Pathol* **123:**285-293.

Ron Y., Flitman-Tene R., Dybvig K. et Yogev D. (2002). Identification and characterization of a site-specific tyrosine recombinase within the variable

loci of *Mycoplasma bovis*, *Mycoplasma pulmonis* and *Mycoplasma agalactiae*. *Gene* **292:**205-211.

Rosati S., Pozzi S., Robino P., Montinaro B., Conti A., Fadda M. et Pittau M. (1999). P48 major surface antigen of *Mycoplasma agalactiae* is homologous to a malp product of *Mycoplasma fermentans* and belongs to a selected family of bacterial lipoproteins. *Infect Immun* **67:**6213-6216.

Rosati S., Robino P., Fadda M., Pozzi S., Mannelli A. et Pittau M. (2000). Expression and antigenic characterization of recombinant *Mycoplasma agalactiae* P48 major surface protein. *Vet Microbiol* **71:**201-210.

Rosendal S. et Black F.T. (1972). Direct and indirect immunofluorescence of unfixed and fixed Mycoplasma colonies. *Acta Pathol Microbiol Scand B Microbiol Immunol* **80:**615-622.

Rosengarten R., Citti C., Much P., Spergser J., Droesse M. et Hewicker-Trautwein M. (2001). The changing image of mycoplasmas: from innocent bystanders to emerging and reemerging pathogens in human and animal diseases. *Contrib Microbiol* **8:**166-185.

Rosengarten R. et Wise K.S. (1990). Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. *Science* **247:**315-318.

Rosengarten R. et Wise K.S. (1991). The Vlp system of *Mycoplasma hyorhinis*: combinatorial expression of distinct size variant lipoproteins generating high-frequency surface antigenic variation. *J Bacteriol* **173:**4782-4793.

Roske K., Blanchard A., Chambaud I., Citti C., Helbig J.H., Prevost M.C., Rosengarten R. et Jacobs E. (2001). Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect Immun* **69:**7642-7651.

Rottem S. (2003). Interaction of mycoplasmas with host cells. *Physiol Rev* **83:**417-432.

Rottem S. et Naot Y. (1998). Subversion and exploitation of host cells by mycoplasmas. *Trends Microbiol* **6:**436-440.

Sanchis R., Abadie G., Lambert M., Cabasse E., Dufour P., Guibert J.M. et Pepin M. (2000). Inoculation of lactating ewes by the intramammary route with *Mycoplasma agalactiae*: comparative pathogenicity of six field strains. *Vet Res* **31:**329-337.

Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K., Yoshino C., Horino A., Shiba T., Sasaki T. et Hattori M. (2002). The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. *Nucleic Acids Res* **30:**5293-5300.

Shen X., Gumulak J., Yu H., French C.T., Zou N. et Dybvig K. (2000). Gene rearrangements in the *vsa* locus of *Mycoplasma pulmonis*. *J Bacteriol* **182:**2900-2908.

Shiel M.J., Coloe P.J., Worotniuk B. et Burgess G.W. (1982). Polyarthritis in a calf associated with a group 7 Mycoplasma infection. *Aust Vet J* **59:**192-193.

Simmons G.C. et Johnston L.A.Y. (1963). Arthritis in calves caused by *Mycoplasma sp. Aust Vet J* **39:**11-14.

Simmons W.L., Bolland J.R., Daubenspeck J.M. et Dybvig K. (2007). A stochastic mechanism for biofilm formation by *Mycoplasma pulmonis*. *J Bacteriol* **189:**1905-1913.

Simmons W.L., Denison A.M. et Dybvig K. (2004). Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. *Infect Immun* **72:**6846-6851.

Singh N., Rajya B.S. et Mohanty G.C. (1974). Granular vulvovaginitis (GVV) in goats associated with *Mycoplasma agalactiae*. *Cornell Vet* **64:**435-442.

Sirand-Pugnet P., Citti C., Barre A. et Blanchard A. (2007a). Evolution of mollicutes: down a bumpy road with twists and turns. *Res Microbiol* **158:**754-766.

Sirand-Pugnet P., Lartigue C., Marenda M., Jacob D., Barre A., Barbe V., Schenowitz C., Mangenot S., Couloux A., Segurens B., de Daruvar A., Blanchard A. et Citti C. (2007b). Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS Genet* **3:**e75.

Sitaraman R., Denison A.M. et Dybvig K. (2002). A unique, bifunctional site-specific DNA recombinase from *Mycoplasma pulmonis*. *Mol Microbiol* **46:**1033-1040.

Sitaraman R. et Dybvig K. (1997). The *hsd* loci of *Mycoplasma pulmonis*: organization, rearrangements and expression of genes. *Mol Microbiol* **26:**109-120.

Sokolova I.A., Vaughan A.T. et Khodarev N.N. (1998). Mycoplasma infection can sensitize host cells to apoptosis through contribution of apoptotic-like endonuclease(s). *Immunol Cell Biol* **76:**526-534.

Srivastava N.C., Thiaucourt F., Singh V.P. et Sunder J. (2000). Isolation of *Mycoplasma mycoides* small colony type from contagious caprine pleuropneumonia in India. *Vet Rec* **147:**520-521.

Stanbridge E., Onen M., Perkins F.T. et Hayflick L. (1969). Karyological and morphological characteristics of human diploid cell strain WI-38 infected with mycoplasmas. *Exp Cell Res* **57:**397-410.

Stuart P.M., Cassell G.H. et Woodward J.G. (1990). Differential induction of bone marrow macrophage proliferation by mycoplasmas involves granulocyte-macrophage colony-stimulating factor. *Infect Immun* **58:**3558-3563.

Subramaniam S., Bergonier D., Poumarat F., Capaul S., Schlatter Y., Nicolet J. et Frey J. (1998). Species identification of *Mycoplasma bovis* and *Mycoplasma agalactiae* based on the *uvrC* genes by PCR. *Mol Cell Probes* **12:**161-169.

Tardy F., Gaurivaud P., Tricot A., Maigre L., Bézille P. et Poumarat F. (2007). *La DGGE, outil de diagnostic ou outil d'expertise pour l'identification des souches de mycoplasmes dans le cadre du réseau vigimyc.* Presented at the 7ème Rencontre Francophone de Mycoplasmologie, Lyon, 17-18 juillet 2007.

Tardy F., Gaurivaud P., Tricot A., Maigre L. et Poumarat F. (2009). Epidemiological surveillance of mycoplasmas belonging to the '*Mycoplasma mycoides*' cluster: is DGGE fingerprinting of 16S rRNA genes suitable? *Lett Appl Microbiol* **48:**210-217.

Taylor-Robinson D., Davies H.A., Sarathchandra P. et Furr P.M. (1991). Intracellular location of mycoplasmas in cultured cells demonstrated by immunocytochemistry and electron microscopy. *Int J Exp Pathol* **72:**705-714.

Taylor-Robinson D., Purcell R.H., Wong D.C. et Chanock R.M. (1966). A colour test for the measurement of antibody to certain mycoplasma species based upon the inhibition of acid production. *J Hyg (Lond)* **64:**91-104.

Taylor-Robinson D., Sarathchandra P. et Furr P.M. (1993). *Mycoplasma fermentans*--HeLa cell interactions. *Clin Infect Dis* **17 Suppl 1:**S302-304.

Theiss P. et Wise K.S. (1997). Localized frameshift mutation generates selective, high-frequency phase variation of a surface lipoprotein encoded by a mycoplasma ABC transporter operon. *J Bacteriol* **179:**4013-4022.

Thiaucourt F. (2003). Pleuropneumonie contagieuse caprine, p. 795-808. *In* L. P.-C., J. Blancou and R. Chermette (ed.), Principales maladies infectieuses et parasitaires du bétail. Europe et régions chaudes. Maladies bactériennes, mycoses, maladies parasitaires., vol. 2. *Lavoisier Tec et Doc*, Paris (France).

Thomas A., Nicolas C., Dizier I., Mainil J. et Linden A. (2003). Antibiotic susceptibilities of recent isolates of *Mycoplasma bovis* in Belgium. *Vet Rec* **153:**428-431.

Thomas P. (1873). Rapport médical sur le Bou Frida. Publication du gouvernement général civil de l'Algérie, p. 35, Publication du gouvernement général civil de l'Algérie. *Jourdan, A. (ed)*, Alger.

Tola S., Angioi A., Rocchigiani A.M., Idini G., Manunta D., Galleri G. et Leori G. (1997). Detection of *Mycoplasma agalactiae* in sheep milk samples by polymerase chain reaction. *Vet Microbiol* **54:**17-22.

Tola S., Idini G., Manunta D., Galleri G., Angioi A., Rocchigiani A.M. et Leori G. (1996). Rapid and specific detection of *Mycoplasma agalactiae* by polymerase chain reaction. *Vet Microbiol* **51:**77-84.

Tsai S., Wear D.J., Shih J.W. et Lo S.C. (1995). Mycoplasmas and oncogenesis: persistent infection and multistage malignant transformation. *Proc Natl Acad Sci U S A* **92:**10197-10201.

Tully J.G. (1995). Culture medium formulation for primary isolation and maintenance of mollicutes., p. 33–39. *In* J. G. Tully (ed.), Molecular and Diagnostic Procedures in Mycoplasmology: Molecular Characterization. *Academic Press.*, San Diego.

Tully J.G., Bové J.M., Laigret F. et Whitcomb R.F. (1993). Revised taxonomy of the class *Mollicutes* : proposed elevation of a monophyletic cluster of arthropod-associated mollicutes to ordinal rank (*Entomoplasmatales ord. nov.*), with provision for familial rank to separate species with nonhelical morphology (*Entomoplasmataceae fam. nov.*) from helical species (*Spiroplasmataceae*), and emended descriptions of the order *Mycoplasmatales*, family *Mycoplasmataceae*. *Int J Syst Bacteriol* **43:**378-385.

van der Woude M.W. (2006). Re-examining the role and random nature of phase variation. *FEMS Microbiol Lett* **254:**190-197.

van der Woude M.W. et Baumler A.J. (2004). Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17:**581-611, table of contents.

van Putten J.P., Duensing T.D. et Cole R.L. (1998). Entry of OpaA+ gonococci into HEp-2 cells requires concerted action of glycosaminoglycans, fibronectin and integrin receptors. *Mol Microbiol* **29:**369-379.

Vanden Bush T.J. et Rosenbusch R.F. (2002). *Mycoplasma bovis* induces apoptosis of bovine lymphocytes. *FEMS Immunol Med Microbiol* **32:**97-103.

Vasconcelos A.T., Ferreira H.B., Bizarro C.V., Bonatto S.L., Carvalho M.O., Pinto P.M., Almeida D.F., Almeida L.G., Almeida R., Alves-Filho L., Assuncao E.N., Azevedo V.A., Bogo M.R., Brigido M.M., Brocchi M., Burity H.A., Camargo A.A., Camargo S.S., Carepo M.S., Carraro D.M., de Mattos Cascardo J.C., Castro L.A., Cavalcanti G., Chemale G., Collevatti R.G., Cunha C.W., Dallagiovanna B., Dambros B.P.,

Dellagostin O.A., Falcao C., Fantinatti-Garboggini F., Felipe M.S., Fiorentin L., Franco G.R., Freitas N.S., Frias D., Grangeiro T.B., Grisard E.C., Guimaraes C.T., Hungria M., Jardim S.N., Krieger M.A., Laurino J.P., Lima L.F., Lopes M.I., Loreto E.L., Madeira H.M., Manfio G.P., Maranhao A.Q., Martinkovics C.T., Medeiros S.R., Moreira M.A., Neiva M., Ramalho-Neto C.E., Nicolas M.F., Oliveira S.C., Paixao R.F., Pedrosa F.O., Pena S.D., Pereira M., Pereira-Ferrari L., Piffer I., Pinto L.S., Potrich D.P., Salim A.C., Santos F.R., Schmitt R., Schneider M.P., Schrank A., Schrank I.S., Schuck A.F., Seuanez H.N., Silva D.W., Silva R., Silva S.C., Soares C.M., Souza K.R., Souza R.C., Staats C.C., Steffens M.B., Teixeira S.M., Urmenyi T.P., Vainstein M.H., Zuccherato L.W., Simpson A.J. et Zaha A. (2005). Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. *J Bacteriol* **187:**5568-5577.

Verbisck-Bucker G., Gonzalez-Candela M., Galian J., Cubero-Pablo M.J., Martin-Atance P. et Leon-Vizcaino L. (2008). Epidemiology of *Mycoplasma agalactiae* infection in free-ranging Spanish ibex (*Capra pyrenaica*) in Andalusia, southern Spain. *J Wildl Dis* **44:**369-380.

Vilei E.M. et Frey J. (2001). Genetic and biochemical characterization of glycerol uptake in *Mycoplasma mycoides* subsp. *mycoides* SC: its impact on H(2)O(2) production and virulence. *Clin Diagn Lab Immunol* **8:**85-92.

Walz P.H., Mullaney T.P., Render J.A., Walker R.D., Mosser T. et Baker J.C. (1997). Otitis media in preweaned Holstein dairy calves in Michigan due to *Mycoplasma bovis*. *J Vet Diagn Invest* **9:**250-254.

Westberg J., Persson A., Holmberg A., Goesmann A., Lundeberg J., Johansson K.E., Pettersson B. et Uhlen M. (2004). The genome sequence of *Mycoplasma mycoides* subsp. *mycoides* SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). *Genome Res* **14:**221-227.

Willems L. (1852). Mémoire sur la pleuro-pneumonie épizootique du gros bétail. *Recueil de Médecine Vétérinaire Pratique* **IX:**401-434.

Winner F., Markova I., Much P., Lugmair A., Siebert-Gulle K., Vogl G., Rosengarten R. et Citti C. (2003). Phenotypic switching in *Mycoplasma gallisepticum* hemadsorption is governed by a high-frequency, reversible point mutation. *Infect Immun* **71:**1265-1273.

Winner F., Rosengarten R. et Citti C. (2000). In vitro cell invasion of *Mycoplasma gallisepticum*. *Infect Immun* **68:**4238-4244.

Wise K.S. (1993). Adaptive surface variation in mycoplasmas. *Trends Microbiol* **1:**59-63.

Wise K.S., Foecking M.F., Roske K., Lee Y.J., Lee Y.M., Madan A. et Calcutt M.J. (2006). Distinctive repertoire of contingency genes conferring mutation- based phase variation and combinatorial expression of surface lipoproteins in *Mycoplasma capricolum* subsp. *capricolum* of the *Mycoplasma mycoides* phylogenetic cluster. *J Bacteriol* **188:**4926-4941.

Woese C.R. (1987). Bacterial evolution. *Microbiol Rev* **51:**221-271.

Yavlovich A., Higazi A.A. et Rottem S. (2001). Plasminogen binding and activation by *Mycoplasma fermentans*. *Infect Immun* **69:**1977-1982.

Yogev D., Menaker D., Strutzberg K., Levisohn S., Kirchhoff H., Hinz K.H. et Rosengarten R. (1994). A surface epitope undergoing high-frequency

phase variation is shared by *Mycoplasma gallisepticum* and *Mycoplasma bovis*. *Infect Immun* **62:**4962-4968.

Zhang Q. et Wise K.S. (1996). Molecular basis of size and antigenic variation of a *Mycoplasma hominis* adhesin encoded by divergent vaa genes. *Infect Immun* **64:**2737-2744.

Zhang Q., Young T.F. et Ross R.F. (1994). Glycolipid receptors for attachment of *Mycoplasma hyopneumoniae* to porcine respiratory ciliated cells. *Infect Immun* **62:**4367-4373.

PREMIER CHAPITRE

PREMIER CHAPITRE : DIVERSITE GENETIQUE ET MARQUEURS EPIDEMIOLOGIQUES CHEZ *M. AGALACTIAE* : UNE SEULE SOUCHE CIRCULE DANS LES FOYERS RECURRENTS DU SUD-OUEST DE LA FRANCE DEPUIS PRES DE 30 ANS.

Laurent-Xavier Nouvel, Marc Marenda, Michelle Glew, Eveline Sagné, Agnès Brouillaud, Renate Rosengarten, François Poumarat et Christine Citti.

Article en préparation pour une soumission à *Veterinary Research*

Résumé en français

Mycoplasma agalactiae est responsable de l'agalactie contagieuse, une maladie du mouton et de la chèvre aux impacts cliniques et économiques importants. Les études précédentes ont montré une assez grande homogénéité des souches et peu de techniques moléculaires permettent une différentiation fine entre les isolats. Pour développer de tels outils, nous avons évalué la diversité génétique de *M. agalactiae* à l'aide de trois techniques moléculaires : les *variable number tandem repeat* (VNTR) mis au point chez *M. agalactiae* par McAuliffe *et al.* [27], les profils de restriction obtenus avec des sondes spécifiques de la souche type PG2 identifiées par Marenda *et al.* [24], et la détermination du répertoire des gènes *vpma* par *Southern blot*. Sur un panel de 40 isolats représentatifs de la diversité des souches de *M. agalactiae*, les résultats obtenus avec les trois approches concordent et mettent en évidence une diversité génétique pour seulement quelques isolats alors que les autres sont tous très similaires à la souche type. Absents des isolats semblables à PG2, les éléments génétiques mobiles détectés parmi les isolats divergents apparaîssent comme un facteur de diversité. Sur un second panel d'isolats provenant du département français des Pyrénées-Atlantiques, zone où l'agalactie contagieuse sévit depuis longue date, les isolats apparaissent identiques avec l'ensemble des méthodes employées, à l'exception d'un seul gène *vpma* non détecté chez 5 isolats sur les 63 étudiés. Ces résultats conduisent à l'hypothèse de circulation d'une seule et même souche dans cette zone d'endémie de 1977 à 2007.

1. *Introduction*

Une disparité des statuts vis-à-vis de *M. agalactiae* existe de par le monde et entre les différentes zones d'élevage avec des régions indemnes et des régions à prévalence sporadique ou endémique [11]. En France, l'agalactie contagieuse à *M. agalactiae* sévissait historiquement de manière endémique dans deux bassins d'élevage : les Savoies (Savoie et Haute-Savoie) et les Pyrénées-Atlantiques (Pays Basque et Béarn) [3,4]. Aujourd'hui, la maladie peut être considérée comme éradiquée des départements savoyards grâce à des mesures sanitaires sévères, le dernier cheptel infecté datant de 2002 [39]. En revanche, la maladie persiste toujours dans les Pyrénées-Atlantiques et le nombre de cas déclarés est en forte augmentation depuis 2006 [15]. En effet, malgré une phase initiale de régression du nombre de troupeaux infectés grâce à l'application de mesures sanitaires mises en place dans ce département dans les années 90, avec 135 élevages infectés en 2008 contre 20 en 2005, on assiste aujourd'hui à une recrudescence du nombre de foyers liée à l'essoufflement des mesures, à la difficulté d'un contrôle efficace des mouvements d'animaux contagieux et au défaut de sensibilité des tests de dépistage.

Des outils d'épidémiologie moléculaire permettraient de retracer l'historique des contaminations et de remonter jusqu'à l'origine des foyers d'agalactie contagieuse par une différenciation fine des souches. Ces outils constitueraient une aide précieuse aux plans de lutte contre la maladie pour une meilleure connaissance de l'origine et des chaînes de transmission. Au préalable, une question se pose : existe-t-il une variabilité génétique entre différentes souches de *M. agalactiae* ? Si oui, quelles sont sa nature et son importance ?

Bien que *M. agalactiae* génère une très grande variabilité antigénique grâce à des systèmes génétiques sophistiqués tel que celui représenté par la famille des gènes *vpma* [5,16], les variations génétiques mises en évidence entre différentes souches à l'ère pré-génomique sont rares [24,25 ,41,44] et n'offrent pas des traceurs moléculaires permettant un suivi épidémiologique. Le séquençage récent du génome de la souche-type PG2 a permis d'envisager la mise au point chez

M. agalactiae de nouvelles techniques de typage très utilisées chez d'autres procaryotes comme l'analyse des variations du nombre de répétitions en tandem (*variable number tandem repeats*, VNTR), le typage par séquençage de multiples loci (*multi locus sequence typing*, MLST), le typage par analyse des répétitions palindromiques groupées, courtes et régulièrement espacées (*Clustered regularly interspaced short palindromic repeats*, CRISPR).

Les VNTR sont comparables aux microsatellites utilisés en génétique eucaryote. Ils ont été rencontrés pour la première fois chez *M. tuberculosis* sous le nom de *Mycobacterial Interspersed Repeat Units* (MIRU) [14,43]. Depuis, des VNTR ont été décrits chez de nombreuses espèces bactériennes de *B. anthracis* [20] à *E. coli* O157H7 [29] en passant par *Staphylococcus aureus* [35] et *Yersinia pestis* [19]. Les VNTR sont des répétitions de séquences d'ADN en tandem dont le nombre de copies varie. Leur pouvoir de discrimination a été démontré comme élevé et donne des informations sur la diversité évolutive des bactéries [46]. Avec l'augmentation du nombre de génomes séquencés, le développement et la croissance de bases de données consacrées aux VNTR, l'analyse des VNTR apparaît comme une technique de plus en plus universelle pour le typage des bactéries. Après une adaptation réussie à *M. mycoides* subsp. *mycoides* SC [26], McAuliffe et collaborateurs ont récemment étudié des VNTR chez *M. agalactiae* [27].

Ces auteurs mettent en évidence que sur un total de trente répétitions en tandem détectées sur le génome de la souche type PG2, beaucoup ne sont présentes qu'en double copie et présentent une conservation inférieure à 80%. Leur étude préliminaire de dix souches choisies pour l'étendue de leurs origines géographiques montre des différences intraspécifiques pour sept répétitions. Quatre ont été retenues comme VNTR. Leur pouvoir de discrimination a été comparé à ceux de l'électrophorèse en champ pulsé (*pulse fiel gel electrophoresis*, PFGE) et de l'amplification au hasard de séquences polymorphes d'ADN (*random amplified polymorphic DNA*, RAPD) sur quatre-vingt-huit souches européennes. Plus performants que les deux autres techniques, les VNTR permettent de décrire un niveau de diversité important avec neuf profils différents

(Index de diversité de Simpson de 0,574 contre seulement 0,324 pour la PFGE) [27].

Une variabilité génétique entre des souches de *M. agalactiae* a aussi été mise en évidence à l'aide de sondes ADN (sondes AGA) identifiées lors d'une étude de comparaison génomique par hybridation soustractive suppressive (ou SSH pour *Suppression Subtractive Hybridization*) réalisée au laboratoire [24]. Utilisée pour des études de génomique comparative intéressant par exemple *Escherichia coli* et *Salmonella typhimurium* [6], *Yersinia pestis* et *Y. pseudotuberculosis* [33], la technique SSH est une approche qui permet de s'affranchir du séquençage génomique complet pour une comparaison globale du génome de bactéries phylogénétiquement proches. Validée au laboratoire pour la comparaison de deux souches de *M. agalactiae* [25], elle a par la suite été utilisée pour comparer deux proches cousins phylogénétiques : *M. agalactiae* et *M. bovis* [24]. Par comparaison à la souche-type PG45 de *M. bovis*, un ensemble de cinquante-neuf séquences d'ADN spécifiques de la souche-type *M. agalactiae* PG2 a été mis en évidence. Des analyses *Southern blot* sur un panel de trente-deux isolats (inclus dans la collection de 101 isolats étudiée ici) ont révélé que 31% des sondes AGA spécifiques de PG2 ne réagissent pas avec la totalité des isolats [24]. Des régions génomiques de PG2 ne sont donc pas partagées par l'ensemble des souches de *M. agalactiae* testées. De plus, les profils obtenus avec certaines sondes diffèrent entre les isolats par la taille ou le nombre de bandes révélées. A l'époque, cette méthode différenciant les isolats n'avait pas été évaluée en comparaison à une autre méthode comme les VNTR. Les résultats des deux techniques sont-ils comparables ?

Les approches VNTR et SSH s'intéressent à des variations réparties sur l'ensemble du génome [Figure 1]. L'existence de zones de plus forte activité appelées *hot spots* a été démontrée chez certaines bactéries [34]. Chez *M. agalactiae*, le locus *vpma* est le siège de nombreux remaniements génétiques. Les Vpma sont des lipoprotéines abondamment exprimées à la surface de *M. agalactiae*. Il s'agit d'une famille de protéine ayant des caractéristiques communes : très immunogènes chez l'hôte et contenant des épitopes qui seraient impliqués dans l'adhésion [18,36], elles possèdent une séquence signal

conservée et peuvent partager des motifs d'acides aminés répétés [16,18]. Leur expression est qualifiée d'hypervariable : il a été démontré chez la souche type PG2 que bien qu'un clone donné exprime un et un seul type de Vpma, le type de Vpma exprimé est quant à lui variable de manière intra-clonale. Cette hypervariabilité d'expression entraîne donc la possible expression d'une Vpma différente entre différents clones issus d'une même cellule parentale [18]. D'après les résultats obtenus par Glew *et al.*, cette hypervariabilité d'expression est due à des réarrangements d'ADN spécifiques survenant au niveau de parties non codantes du locus sous l'action d'une intégrase-recombinase (Xer1). Le nombre et la nature des gènes présents apparaîssent quant à eux stables au sein des cellules filles issues de PG2 [16]. Si l'expression d'une Vpma donnée à la surface d'un mycoplasme est hypervariable, l'hypothèse est donc que le répertoire des gènes *vpma* (c'est-à-dire l'ensemble des Vpma potentiellement exprimables par une souche) est relativement stable pour une souche donnée. Pour la souche-type PG2, six gènes *vpma* regroupés en *cluster* à un unique locus ont été identifiés adjacent à un gène *xer1* [16,40]. A l'aide d'une sonde oligonucléotidique appelée A3F, spécifique de la séquence signal conservée entre les gènes *vpma* chez PG2 [18], et de sondes ADN spécifiques de chacun des gènes correspondants à un type de Vpma connu [16], il semble possible d'évaluer le répertoire *vpma* de différents isolats par *Southern blot*.

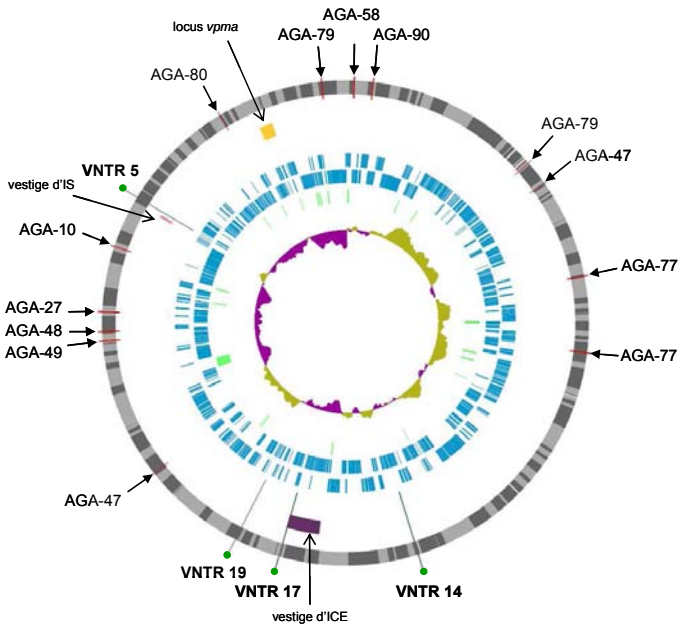


Figure 1 : Position des VNTR et des sondes AGA sur le génome de la souche PG2.

Représentation en cercles concentriques, de l'extérieur vers l'intérieur : VNTR (rayons verts allongés marqué d'un point) ; 12 sondes AGA (rouge) repérées par une flèche (AGA-47, 77 et 79 s'hybrident théoriquement à deux endroits du génome) ; fragments de restriction *Eco*RI (alternance gris clair et foncé) ; locus *vpma* (jaune), vestige d'IS (rouge) et d'ICE (violet) ; génome de PG2 (2 cercles bleus, 1 vert, respectivement : CDS brin +, CDS brin –, ARNs) ; biais en GC.

Dans cette étude, nous avons évalué la diversité génétique de 101 isolats répartis en deux panels de quarante et soixante-trois isolats (les isolats 4206 et 7784 sont communs aux deux panels) pour la mise au point d'éventuels marqueurs épidémiologiques. Le panel 1 est composé d'isolats aussi représentatifs que possible de la diversité des souches rencontrées dans le monde. Le panel 2 est une collection d'isolats provenant des Pyrénées-Atlantiques, département français et bassin d'élevage ovin où l'agalactie contagieuse sévit de manière endémique en système quasi-clos. Pour l'étude de la diversité génétique de ces deux panels, nous avons réalisé l'analyse des profils VNTR, des profils obtenus avec des sondes AGA issues de la SSH et des sondes ciblant le locus *vpma*. Les résultats obtenus avec ces trois approches sont concordants, ils montrent une diversité génétique pour quelques isolats en opposition à une relative constance génétique de nombreux isolats proches de PG2. Dans les Pyrénées-Atlantiques, à l'exception d'un seul gène *vpma* non détecté chez cinq isolats sur soixante-trois, les soixante-trois isolats apparaissent comme identiques avec les trois méthodes employées. Ce résultat conduit à l'hypothèse de la circulation d'une seule et même souche dans cette zone ayant toujours connu l'agalactie de mémoire de berger.

2. *Matériel et méthodes*

Souches bactériennes, conditions de culture et extraction de l'ADN génomique

Le clone 55.5 de la souche type de *M. agalactiae* PG2 utilisé dans cette étude est celui décrit par Glew *et al.* [16]. L'ensemble des isolats étudiés et leurs caractéristiques (année et nature du prélèvement, hôte et pays d'origine) sont présentés dans les tableaux 1 et S1. Ces isolats font partie de la collection du réseau VIGIMYC maintenue par le docteur François Poumarat (AFSSA Lyon). Les mycoplasmes ont été incubés 2 jours à 37℃ en milieu Aluotto [1] ou SP4 [45], étalés sur milieu Aluotto gélosé et incubés à 37℃ pendant 3 jours. Une colonie isolée a été reprise pour chaque isolat, incubée en milieu Aluotto liquide à 37℃ pendant 3 jours et conservée à -80℃. Cette étape de clonage n'a pas été réalisée pour les isolats postérieurs à 2003, les cultures initiales en SP4 ont été directement congelées sans étalement sur milieu solide. Un aliquote des cultures conservées à -80℃ a permis d'ensemencer 20 ml de milieu liquide pour les extractions d'ADN. L'ADN génomique a été extrait au phénol-chloroforme après digestion par la proteinase K [37] ou selon une méthode simplifiée avec du chloroforme [8].

Analyses VNTR

Les analyses VNTR ont été réalisées comme décrit par McAuliffe *et al.* [27]. Les séquences des VNTR ont été amplifiées par PCR avec un thermocycleur MJresearch PT100, dans un volume final de 25µl contenant 0,4 mM de chaque amorce, 200 mM de chaque dNTP, du tampon contenant du MgSO4 fourni par le fabricant (New England Biolabs) à une concentration finale de 1X et 2,5 unités d'ADN polymérase Taq (New England Biolabs). Pour chaque VNTR, un couple d'amorces spécifiques a été utilisé : VNTR 5, 5F (5'-GAAAGAGAAAGGAA GCTGAA-3') et 5R (5'- GGATCATTATCGCTTTTTGA-3') ; VNTR 14, 14F (5'-TTGAAATATCCGCTTAAGAAA-3') et 14R (5'-AATTTGCATTT AATGGTGCT-3') ; VNTR 17, 17F (5'-TTTAGCTTTTGATTCAATACTTTC-3') et 17R (5'-AAAGAATTA TGCGAGCATTT-3') ; VNTR 19, 19F (5'-TTGCTTCTTGTG CTTCTTTT-3') et 19R

(5'-AAGGGGATCAACCAGATAAT-3'). Le programme de PCR utilisé était : 5 min. à 94℃ suivi de 30 cycles de 1 min. à 94℃, 45 s. à 56℃, 1 min. à 72℃ et terminé par une élongation finale de 10 min. à 94℃. Les pr oduits générés ont été colorés au bromure d'éthidium visualisé sous lumière ultra violette après migration par électrophorèse en gel d'agarose à 2,5%. Les dendrogrammes de profils VNTR ont été construits avec le programme Phylip (PHYLogeny Interface Package) accessible en ligne à l'adresse http://pubmlst.org/analysis/ avec la méthode Neighbor-Joining.

Southern blots et PCR

Pour les *Southern blots*, 1 µg d'ADN a été complètement digéré par *Ase*I, *Alw*I ou *EcoR*I (New England Biolabs, Ozyme, France), migré par électrophorèse en gel d'agarose 1% et transféré sur membrane de nylon (Roche). Les sondes spécifiques de chacun des 6 gènes *vpma* du clone 55.5 ont été marquées avec la digoxigénine par PCR en présence de dig-11-dUTP (Roche) comme décrit précédemment [16]. Les réactions PCR ont été réalisées avec le même appareil et des mélanges réactionnels identiques à ceux décrits pour les analyses VNTR. Les couples d'amorces et l'ADN matrice utilisés pour la réalisation de chaque sonde *vpma* sont ceux décrits par Glew *et al.* [16] soit i) pour *vpmaU* : U1F (5'-cgcggatcgGATAAAGAAGATAAGACAGGTAG-3'), U1R (5'-aaactgcagTCAACC TTAGATAAACCACCTAAC-3'), plasmide p5E1.2 ; ii) pour *vpmaV* : V1F (5'-ccggaattcGTTGAGGAAGCAATTAAAACAGC-3'), V1R (5'-gctctagattaTCCAGATG ATGTTTCAACTTC-3'), p5X2.2 ; iii) pour *vpmaW* ; W1F (5'-cgcggatccAATGGCGG AAATAGTAATGGTAAC-3'), W1R (5'-aaactgcagAATAACTTTATCTAGTTCTAT ACC-3'), p5H3.0 ; iv) pour *vpmaX* : X1F (5'-cgcggatccAAAGTAATGAAGGTCAAT TACC-3'), X1R (5'-aaactgcagGCTTAAGGATTTTTTAAAATGATG-3'), p5H1.8 ; v) pour *vpmaY* : Y1F (5'-ccggaattcAATGCAAACGCTGCAGAAAATG-3'), Y1R (5'-gctctagaGATTAAACTTTTTTTACAGTAAATG-3'), ADN génomique de PG2 clone 55-5 ; vi) pour *vpmaZ* : Z1F (5'-cgcggatccCAAACAGATTCAACTC CGTCAAC-3'), Z1R (5'-aaactgcagTTATTCGTATTTAGGTAATAGTCTTC-3'), p5H4.7 (Dans la séquence des amorces, les lettres capitales représentent les séquences spécifiques des portions de gènes amplifiées et les minuscules des nucléotides introduits pour des commodités de clonage plasmidique. Les conditions de PCR utilisées sont : 1 min. à 94℃ suivie de 28 cycles de 20 s. à 94℃, 30 s. à 55℃

(pour *vpmaY*), 60℃ (pour *vpmaW*, *vpmaX* et *vpmaZ*), 65℃ (pour *vpmaV*) ou 70℃ (pour *vpmaU*), et 10 s. (*vpmaV*), 18 s. (*vpmaU*, *vpmaX*, and *vpmaY*), ou 54 s. (*vpmaW* and *vpmaZ*) à 72℃, terminé par une élongation finale de 7 min. à 72℃. Les sondes correspondant à *vpmaY* et *vpmaV* ont été générées à partir de produits PCR purifiés sur gel pour éviter des hybridations croisées avec d'autres gènes *vpma*. Les sondes AGA ont également été obtenues par PCR comme décrit précédemment par Marenda *et al.* [24]. Après dénaturation à 100℃, les sondes ont été incubées avec les membranes sur la nuit dans du tampon Church [10] à 55℃ pour les sondes *vpma*, à 60℃ pour les sondes AGA. Les membranes ont ensuite été lavées 1 heure à 55℃ dans du tampon SSC 0,2X, SDS 0,1% (Le tampon SSC 1X est un tampon aqueux composé de NaCl 0,15 M et Citrate de Sodium 0,015 M). La sonde oligonucléotidique dégénérée A3F (5'-AA(A/G)TG(T/C)GG(A/T)GG(A/T)AC(A/T)A(A/C)(A/T)(A/G)A-3') [18] marquée à la digoxigénine a été incubée avec les membranes sur la nuit à 45℃ dans du tampon Church. Les membranes ont été ensuite lavées deux fois 10 minutes dans du tampon SSC 6X, SDS 0,1%. Les sondes hybridées ont été détectées avec des anticorps dirigés contre la digoxygénine conjugués à la phosphatase alcaline. L'activité phosphatase alcaline a été révélée à l'aide d'un réactif chémiluminescent CDP-star (Roche) visualisé par l'exposition de films radiographiques (Amersham). Les sondes ont été déshybridées des membranes par deux incubations de 20 minutes dans une solution de NaOH 0,2M, SDS 0,1% à température ambiante, puis rincées avec de SSC 2X.

3. *Résultats*

Origines et caractéristiques des collections d'isolats étudiés

Pour caractériser la diversité génétique de *M. agalactiae*, nous avons choisi des isolats parmi la collection du réseau VIGIMYC, réseau français d'épidémio-surveillance des mycoplasmoses de ruminants piloté par l'AFSSA de Lyon. Nous avons ainsi étudié deux panels d'isolats regroupant un total de 101 isolats (les isolats 4206 et 7784 isolats sont communs aux deux panels).

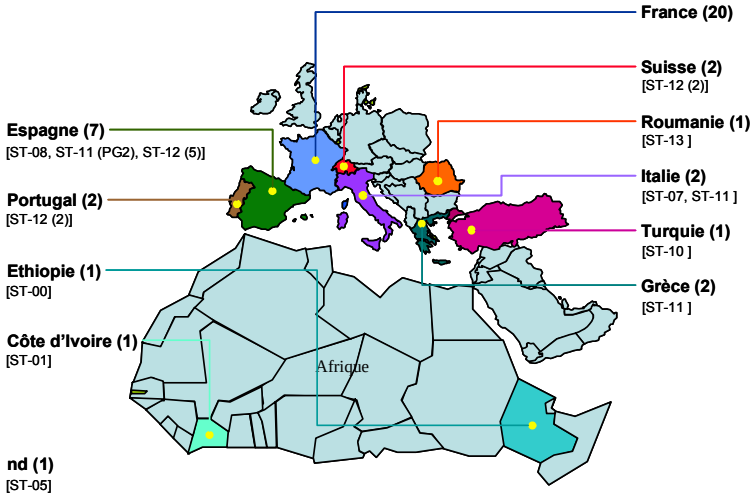
Le panel 1 a été constitué de manière à être aussi représentatif que possible de la diversité des souches rencontrées dans le monde, il comprend quarante isolats initialement identifiés comme appartenant à l'espèce *M. agalactiae* et isolé entre 1951 et 2006 [Tableau 1]. A l'exception de quatre isolats pour lesquels il n'a pas été possible de connaître l'historique de l'hôte d'origine, vingt-et-un isolats proviennent de chèvres, treize de brebis et deux de bouquetins. Le lait est la source majoritaire (27 isolats). Le pays d'origine d'un isolat (3990) n'a pu être retrouvé. A l'exception de deux isolats africains de provenance éloignée de la méditerranée (Ethiopie et Côte d'Ivoire), on compte un isolat turc et trente-six isolats européens : vingt français, sept espagnols, deux suisses, deux portugais, deux italiens, deux grecs et un roumain [Figure 2]. Ce dernier correspond à une souche utilisée pour un vaccin inactivé Agavac (Institut Cantacuzène, Bucarest, Roumanie). Deux isolats (8062 et 8063) avaient été initialement identifiés comme *M. agalactiae* mais appartiennent en réalité à l'espèce *M. bovis* [2,24]. Selon les résultats de Marenda *et al.* [24], ils représentent des souches de *M. bovis* qui sont apparemment très proches de *M. agalactiae*, ils ont été conservés dans la collection étudiée comme éléments de contrôle. Ils permettent également d'évaluer la possibilité d'une transposition des outils utilisés à *M. bovis*.

Tableau 1 : Profils obtenus avec les VNTR sur 40 isolats d'origines variées.

n° isolat	source	hôte	année d'isolement[a]	pays d'origine	Dpt[b]	*IS30-like*[c]	*ISMag1*[d]	P30[e]	P40[f]	VNTR 5	VNTR 14	VNTR 17	VNTR 19	ST[g]
8062[h]	poumon	caprin	1987	Ethiopie		+	-			p0	p0	p0	p0	ST-00
8064	poumon	ovin	1989	C. d'Ivoire		+	-	-	+	p0	p0	p0	p1	ST-01
14628	poumon	bouquetin	2006	France	73					p0	p0	p0	p2	ST-02
4025	lait	ovin	1988	France	12	+	+			p0	p0	p3	p1	ST-03
14668	poumon	caprin	2006	France	7					p0	p0	p6	p1	ST-04
209	lait	caprin	1983	France	31	+	+	-	+	p1	p0	p0	p1	ST-05
3990	nd	ovin	nd	nd		+	+	-	+	p1	p0	p0	p1	ST-05
4055	lait	caprin	1987	France	6	+	+	-	-	p1	p0	p0	p1	ST-05
13387	poumon	bouquetin	2003	France	73					p1	p0	p0	p1	ST-05
14535	lait	caprin	2005	France	26					p1	p0	p0	p1	ST-05
13377	oreille	caprin	2003	France	79					p1	p0	p2	p1	ST-06
9	lait	ovin	nd	Italie		-	-	+	-	p1	p0	p3	p1	ST-07
5632	articulation	caprin	<1991	Espagne		+	+	-	+	p1	p0	p5	p1	ST-08
4054	lait	caprin	1986	France	58	-	+	-	+	p1	p0	p7	p1	ST-09
4212	lait	ovin	1986	France	84	-	+	-	+	p1	p0	p7	p1	ST-09
14340	lait	caprin	2005	France	34					p1	p0	p7	p1	ST-09
8063[h]	nd	nd	1988	Turquie		-	-			p1	p1	p0	p1	ST-10
13375	oreille	caprin	2003	France	79					p1	p1	p0	p1	ST-10
PG2	*nd*	*caprin*	*1952*	*Espagne*		-	-	+	+	*p1*	*p1*	*p1*	*p1*	*ST-11*
4210	lait	ovin	1982	France	73	-	-	+	+	p1	p1	p1	p1	ST-11
4258	lait	caprin	1989	France	79	-	-	+	+	p1	p1	p1	p1	ST-11
6833	lait	caprin	<1993	Italie		-	-	+	+	p1	p1	p1	p1	ST-11
7314	lait	ovin	1986	Grèce		-	-	+	+	p1	p1	p1	p1	ST-11
7327	lait	caprin	1987	Grèce		-	-	+	+	p1	p1	p1	p1	ST-11
9018	lait	caprin	1988	France	73					p1	p1	p1	p1	ST-11
332	articulation	nd	1998	France	nd	-	-			p1	p1	p3	p1	ST-12
4021	lait	ovin	1988	France	12	-	-	+	+	p1	p1	p3	p1	ST-12
4206	lait	ovin	1981	France	64					p1	p1	p3	p1	ST-12
5210	lait	nd	1988	France	81	-	-		+	p1	p1	p3	p1	ST-12
5225	lait	caprin	1991	Espagne		-	-	+	-	p1	p1	p3	p1	ST-12
5670	lait	caprin	<1991	Espagne		-	-	+	+	p1	p1	p3	p1	ST-12
5826	lait	ovin	1992	Espagne		-	-	+	+	p1	p1	p3	p1	ST-12
6968	lait	ovin	1993	Espagne		-	-	+	+	p1	p1	p3	p1	ST-12
7169	lait	caprin	<1993	Suisse		-	-	+	-	p1	p1	p3	p1	ST-12
7375	lait	caprin	<1993	Suisse		-	-	+	-	p1	p1	p3	p1	ST-12
9385	lait	caprin	<1996	Portugal		-	-	+	+	p1	p1	p3	p1	ST-12
9600	lait	caprin	<1996	Portugal		-	-	+	+	p1	p1	p3	p1	ST-12
9875	lait	ovin	<1996	Espagne		-	-			p1	p1	p3	p1	ST-12
7784	lait	ovin	1994	France	64	-	-	+	+	p1	p1	p3	p1	ST-12
190	nd	nd	1951	Roumanie		-	-	+	+	p1	p1	p4	p1	ST-13
N=40						**Nombre de profils différents :**				**2**	**2**	**8**	**3**	**14**

Les isolats sont classés par leur profil VNTR (codage numérique croissant). [a] le sigle < signifie que la date d'isolement précise est inconnue mais au moins antérieure à la date indiquée. [b] code minéralogique des départements : 06, Alpes-Maritimes ; 07, Ardèche ; 12, Aveyron ; 26, Drôme ; 31, Haute-Garonne ; 34, Hérault ; 58, Nièvre ; 64, Pyrénées-Atlantiques ; 73, Savoie ; 79, Deux-Sèvres ; 81, Tarn ; 84, Vaucluse. [c] présence de séquences similaires à l'IS*30*-like de *M. bovis* PG45 déterminée par [25] (si déterminée : détectée, + ; non détectée, -). [d] présence de séquences similaires à l'IS*Mag1* déterminée par [25,32]. [e] détection d'une expression de la protéine P30 par [12]. [f] expression de la protéine P40 détectée par [13]. (g) un numéro, appelé ST pour *Shared Type*, a été attribué à chaque profil VNTR. [h] les souches 8062 et 8063 ont été initialement identifiées comme *M. agalactiae* mais appartiennent en réalité à l'espèce *M. bovis*. nd, non déterminé ; C. d'Ivoire, Côte d'Ivoire ; p1, profil 1 (profil observé pour la souche type de *M. agalactiae* PG2), p2, p3, p4, p5, p6, p7, profils 2, 3, 4, 5, 6, 7 (profils différents de la souche PG2).

(A)



(B)

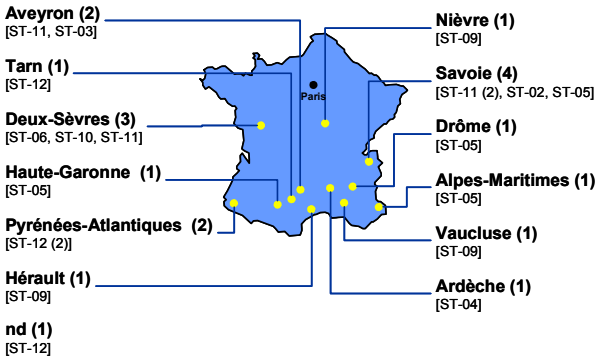


Figure 2 : Répartition géographique des isolats du panel 1.

Ce panel comporte 40 isolats d'origines diverses. Il comprend deux isolats du département français des Pyrénées-Atlantiques, aussi inclus dans le panel 2 (panel constitué uniquement d'isolats de ce département). (A) Répartition géographique des 40 isolats entre les différents pays d'origine. (B) Répartition des isolats français en fonction de leur département d'origine.

Le panel 2 est une collection de soixante-trois isolats sériés dans le temps de 1977 à 2007 provenant des Pyrénées-Atlantiques, département français où l'agalactie contagieuse sévit de manière endémique [Tableau S1]. Du fait d'une forte identité régionale, de la production de fromage d'appellation d'origne contrôlée (AOC) et de l'élevage prédominant de races locales (brebis basco-béarnaises et manechs), les échanges d'animaux avec d'autres zones de production sont limités et ce bassin d'élevage ovin peut être considéré comme un modèle d'étude des souches circulant dans une système quasi-clos. Le département des Pyrénées-Atlantiques est essentiellement un bassin de production de lait de brebis, cet hôte est donc majoritaire avec cinquante-neuf

isolats issus de brebis contre quatre issus de chèvres, tous obtenus à partir de prélèvements de lait.

Description de la diversité génétique de *M. agalactiae* par l'analyse des VNTR

L'étude des quatre VNTR utilisés par McAuliffe *et al.* [27] appliquée à quarante isolats d'origines variées montre une diversité plus ou moins accrue selon les VNTR. Pour chaque VNTR, un numéro de profil est attribué sur la base de sa différence avec les autres isolats : p0 correspond à l'absence d'amplification, p1 à l'observation d'un amplicon de taille identique à celui observé pour PG2, et les profils p2 à p7 à des amplicons de tailles différentes de PG2. Un exemple des profils obtenus pour le VNTR 17 est présenté dans la Figure 3, les résultats de l'ensemble des isolats du panel 1 figurent dans le Tableau 1.

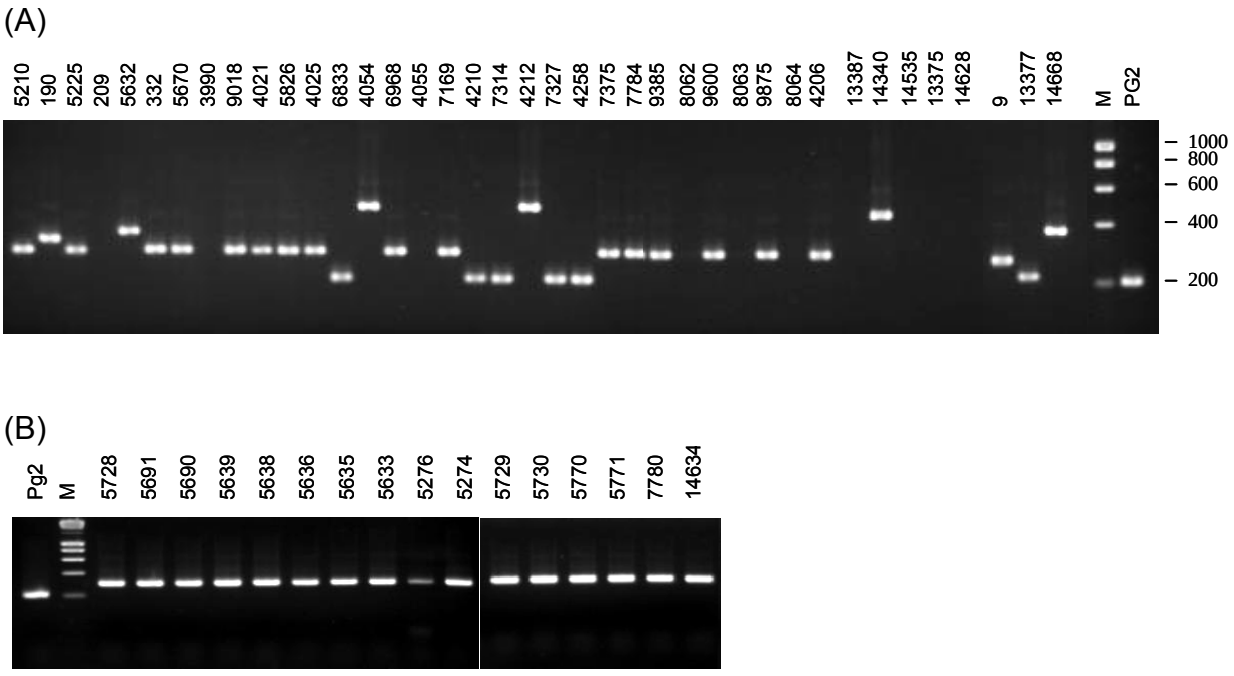


Figure 3 : Variations de profils observées pour le VNTR 17.

(A) panel de 40 isolats d'origines diverses (dont 4206 et 7784 provenant des Pyrénées-Atlantiques) (B) profils de 16 isolats des Pyrénées-Atlantiques. Gel d'agarose 2,5 % coloré au bromure d'éthidium. M ; marqueur de taille (*Smart Ladder* Eurogentec).

Pour les VNTR 5 et 14, il n'y a pas de variation de taille entre les isolats et par rapport à PG2 mais seulement une absence d'amplification pour certains isolats, la variation observée est donc de type p0 ou p1 (absence ou présence) scindant les isolats en deux groupes, l'un avec le VNTR, l'autre sans. Observée pour cinq souches, l'absence du VNTR 5 est toujours liée à l'absence du VNTR 14 mais la réciproque n'est pas observée : onze isolats ne possèdent pas le VNTR 14 et présentent le VNTR 5. Sur les quarante isolats étudiés, le VNTR 19 donne trois profils différents (p0, p1 et p2). Avec un isolat p0 et un p2, le type p1 est largement majoritaire. Le VNTR 17 est le VNTR qui montre le plus fort polymorphisme avec huit profils différents sur les quarante isolats [Figure 3 (A)]. Partagé par seize isolats, le profil p3 est dominant. Les profils p0 et p1 regroupent respectivement huit et sept isolats, le profil p7 trois isolats. Tous les isolats appartenant à ces groupes de profils proviennent de lait à l'exception de deux isolats (un p1 d'origine inconnue et un p3 isolé d'articulation). Les autres profils (p2, p4, p5, p6) sont uniques et rencontrés pour un seul isolat. Ils ne permettent pas de regrouper des isolats et sont appelés profils orphelins.

Tableau 2 : Répartition des isolats entre les différents profils VNTR observés.

Profils	nb
ST-00	1
ST-01	1
ST-02	1
ST-03	1
ST-04	1
ST-05	5
ST-06	1
ST-07	1
ST-08	1
ST-09	3
ST-10	2
ST-11	7
ST-12	14
ST-13	1
Total	*n=40*

En combinant les profils obtenus pour chaque VNTR, les quarante isolats se répartissent parmi quatorze profils différents ou ST (*Shared-Type*) numérotés de 1 à 14 [Tableaux 1 et 2]. Les isolats du panel 1 présentent une diversité supérieure

aux quatre-vingt-huit souches étudiées par McAuliffe *et al.* [27] avec quatorze ST au lieu de neuf. Le profil ST-11 regroupe la souche type originaire d'Espagne, et des souches d'Italie, de Grèce et de France. Le ST-12 est le plus représenté avec quatorze isolats, il comprend des souches isolées en Espagne, au Portugal, en Suisse et en France. Le ST-11 regroupe ensuite sept isolats dont la souche PG2. Le ST-05 contient cinq isolats français mais de départements différents, les ST-09 et 10, respectivement, trois et deux isolats. Enfin les ST-00, 01, 02, 03, 06, 07, 08 et 13 sont orphelins. Par exemple, la souche n°190, souche vaccinale roumaine est la seule représentante du ST-13 qui ne diffère du ST-11 et du ST-12 que par le VNTR 17. Aucune corrélation claire entre l'observation d'un ST donné et l'origine géographique des isolats n'a pu être établie. De même, aucune corrélation franche ne se dégage entre la nature des prélèvements et les profils observés : Sur les cinq isolats de poumons, chacun présente un profil différent, il en est de même pour les deux isolats provenant d'articulation et les deux d'oreilles de chèvre. Les souches isolées à partir du lait semblent plus homogènes avec seulement six profils différents sur vingt-sept isolats.

Stabilité des VNTR des isolats des Pyrénées-Atlantiques de 1977 à 2007

L'analyse des VNTR a été réalisée pour les isolats du panel 2. Les résultats obtenus ne mettent en évidence aucune variation de profil entre les soixante-trois isolats des Pyrénées-Atlantiques même au niveau du VNTR 17 [Figure 3 (B)]. Tous les isolats sont p1 pour les VNTR 5, 14 et 19 et diffèrent de PG2 avec un profil p3 pour le VNTR 17. Ils présentent ainsi tous le type ST-12 [Tableau S1]. Les VNTR ne permettent donc pas une discrimination intraspécifique des isolats des Pyrénées-Atlantiques et il est ainsi tentant de considérer qu'une seule et même souche de *M. agalactiae* circule dans ce département depuis 1977 jusqu'à nos jours.

En regroupant les résultats du panel 2 et ceux obtenus pour les isolats d'origine française du panel 1, les souches des Pyrénées-Atlantiques présentent le même ST que deux isolats ovins de l'Aveyron et du Tarn. Les isolats de Savoie partagent le profil de la souche type PG2 et d'une souche isolée d'un épisode clinique d'agalactie contagieuse à *M. agalactiae* chez la chèvre dans les Deux-Sèvres. Les autres isolats proviennent de portage auriculaire (isolats 13377, 13375), d'animaux asymptomatiques (4025), de niches écologiques jusqu'alors inconnues comme les

bouquetins du Parc Naturel de la Vannoise (13387, 14628), ou de zones où l'agalactie contagieuse sévit de manière sporadique [Figure 2]. Ainsi, si dans les zones endémiques, *M. agalactiae* semble d'une grande constance génétique (ST-11 et 12), des souches atypiques apparaissent ça et là sur l'ensemble du territoire français avec des profils rares ou uniques. Pour mieux définir les similitudes et les différences observées sur l'ensemble de ces isolats, nous avons complété l'analyse des VNTR par des analyses de profils par *Southern blot*.

Des sondes issues de la SSH PG2/PG45 révèlent une diversité génétique entre différents isolats de *M. agalactiae*

Une certaine hétérogénéité intraspécifique a précédemment été démontrée au laboratoire [24] grâce à la comparaison par hybridation soustractive suppressive (ou SSH pour *Suppression Subtractive Hybridization*) des souches types PG2, et de son très proche cousin phylogénétique *M. bovis* PG45. Les résultats indiquent que des régions génétiques semblent ne pas être partagées par toutes les souches de *M. agalactiae*. En plus de l'absence ou de la présence de ces régions détectées par une hybridation positive ou négative de l'ADN des différentes souches avec des sondes spécifiques issues de la SSH, nous avons également recherché la variation des profils de restriction générés par certaines de ces sondes pour tenter d'augmenter le pouvoir de discrimination. Aux isolats présentant une hybridation positive s'ajoute ainsi la possibilité d'une discrimination basée sur la variation du nombre de bandes hybridées et leur variation de taille, ces profils RFLP (*Restriction Length Fragment Polymorphism*) ont été appelés "profils AGA".

Diversité des profils AGA pour les isolats du panel 1

Les cinquante-neuf sondes spécifiques de *M. agalactiae* PG2 [24], appelées sondes AGA, ont été marquées à la digoxygénine et hybridées par *Southern blot* aux digestions *EcoR*I des ADN d'un échantillon de seize isolats du panel 1 [Tableau S2]. Ces seize isolats ont été choisis sur la base de leur répartition entre les différents profils VNTR obtenus sur le panel 1 : six isolats du ST-12, trois isolats du ST-11 (dont PG2), un isolat du ST-05, un du ST-09 et cinq isolats de profils orphelins. [Tableau 3].

Tableau 3 : Profils obtenus avec 12 sondes AGA sur 16 isolats d'origines variées.

n°isolats	source	hôte	année d'isolement[a]	pays d'origine	Dpt[b]	Sondes AGA											Profil AGA[c]	Profil VNTR
						10	27	47	48	49	58	77	79	80	86	90		
4025	lait	ovin	1988	France	12	a2	a0	a3	a0	a2	a1	a2	a3	a1	a1	a1	STaga-01	ST-03
4054	lait	caprin	1986	France	58	a1	a0	a1	a2	a3	a1	a2	a4	a2	a1	a1	STaga-02	ST-09
PG2	*nd*	*caprin*	*1952*	*Espagne*		*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*a1*	*STaga-03*	*ST-11*
190	nd	nd	1951	Roumanie		a1	a1	a1	a1	a1	a2	a1	a1	a1	a1	a1	STaga-04	ST-13
7375	lait	caprin	<1993	Suisse		a1	a1	a1	a1	a1	a3	a1	a2	a1	a1	a2	STaga-05	ST-12
332	articulation	nd	1998	France	nd	a2	a1	a1	a1	a1	a1	a1	a1	a1	a1	a1	STaga-06	ST-12
4210	lait	ovin	1982	France	73	a2	a1	a1	a1	a1	a1	a1	a1	a1	a1	a1	STaga-06	ST-11
5210	lait	nd	1988	France	81	a2	a1	a1	a1	a1	a1	a1	a1	a1	a1	a1	STaga-06	ST-12
5670	lait	caprin	<1991	Espagne		a2	a1	a1	a1	a1	a1	a1	a1	a1	a1	a1	STaga-06	ST-12
9875	lait	ovin	<1996	Espagne		a2	a1	a1	a1	a1	a1	a1	a1	a1	a1	a1	STaga-06	ST-12
7784	lait	ovin	1994	France	64	a2	a1	a1	a1	a1	a2	a1	a1	a1	a1	a1	STaga-07	ST-12
4258	lait	caprin	1989	France	79	a2	a1	a2	a1	a1	a1	a1	a1	a1	a1	a1	STaga-08	ST-11
4055	lait	caprin	1987	France	6	a3	a0	a1	a0	a4	a1	a1	a1	a1	a0	a1	STaga-09	ST-05
5632	articulation	caprin	<1991	Espagne		a4	a0	a4	a3	a1	a4	a3	a5	a1	a1	a1	STaga-10	ST-08
8062[d]	poumon	caprin	1987	Ethiopie		a5	a0	a5	a0	a4	a0	a0	a0	a0	a0	a1	STaga-11	ST-00
8064	poumon	ovin	1989	C.d'Ivoire[d]		a6	a0	a6	a4	a3	a1	a0	a6	a1	a1	a3	STaga-12	ST-01

Les isolats sont classés par leur profil AGA.

[a] le sigle < signifie que la date d'isolement précise est inconnue mais au moins antérieure à la date indiquée. [b] code minéralogique des départements : 06, Alpes-Maritimes ; 12, Aveyron ; 58, Nièvre ; 64, Pyrénées-Atlantiques ; 73, Savoie ; 79, Deux-Sèvres ; 81, Tarn. [c] un numéro de profil AGA, appelé STaga, a été attribué à chaque profil AGA. [d] la souche 8062 a été initialement identifiée comme *M. agalactiae* mais appartient en réalité à l'espèce *M. bovis*. nd, non déterminé ; C. d'Ivoire, Côte d'Ivoire ; p1, profil 1 (profil observé pour la souche type de *M. agalactiae* PG2), p2, p3, p4, p5, p6, profils 2, 3, 4, 5, 6 (profils différents de la souche PG2). Pour faciliter la visualisation des différences : les profils de la souche type sont en italique ; pour chaque sonde AGA, le fond des cases de profils différents de p1 sont de couleurs différentes colorées ; les profils VNTR des isolats (Tableau 2) ont été reportés et les cases des ST retrouvés plusieurs fois (ST-11 et ST-12) ont été colorées (respectivement en jaune et en vert). nd, non déterminé.

Sur la base des profils observés pour ces seize isolats, les douze sondes présentant la variabilité la plus étendue parmi les cinquante-neuf sondes AGA ont été choisies. Parmi celles-ci, la sonde AGA-29 présente une extrême variabilité dans les profils générés, cependant l'analyse de sa séquence indique qu'elle reconnaîtrait une séquence commune aux gènes d'expression hypervariable *vpmaU* et *vpmaV* de PG2. Les réarrangements génétiques à très haute fréquence qui surviennent au locus *vpma* sont intra-clonaux et réversibles [9,16]. Le risque de convergence des profils entre des souches différentes et de différence entre deux clones issus d'un même isolat entraînerait des conclusions erronées, nous avons convenu d'écarter AGA-29 de notre analyse de la diversité génétique au sein des souches de *M. agalactiae*. Un total de onze sondes a donc été utilisé pour les analyses suivantes. L'exemple des profils obtenus avec la sonde AGA-49 est présenté sur la Figure 4.

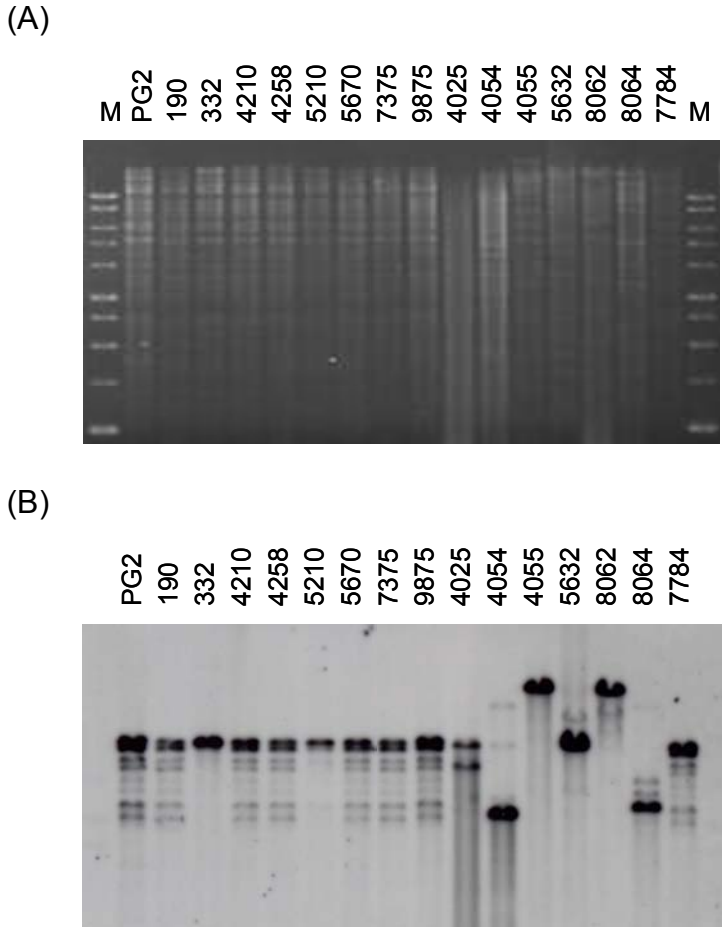


Figure 4 : Variations de profils observées pour la sonde AGA-49.

Profils de restriction *EcoR*I d'un panel de quarante isolats d'origines diverses (A) gel d'agarose 1% et coloration au bromure d'éthidium. (B) hybridation *Southern blot* avec la sonde ADN AGA-49 marquée à la digoxygénine. M, marqueur de taille (*Smart Ladder* Eurogentec).

Pour chaque sonde AGA, les profils ont été désignés a0 en l'absence d'hybridation, a1 lors de l'observation d'un profil identique à PG2 et a2 à a6 pour les autres profils rencontrés. La combinaison des profils obtenus avec ces onze sondes permet de classer les souches en douze groupes notés de STaga-01 à STaga-12 [Tableau 3]. Seul le profil STaga-06 regroupe plusieurs isolats (332, 4210, 5210, 5670 et 9875) [Figure 5]. Ces cinq isolats STaga-06 sont composés de quatre isolats de type VNTR ST-12 et un de type VNTR ST-11. Bien qu'un type VNTR donné puisse se retrouver au sein de plusieurs STaga, la combinaison des deux méthodes de discrimination des souches répartirait les isolats en treize groupes [Tableau 3 et Figure 5].

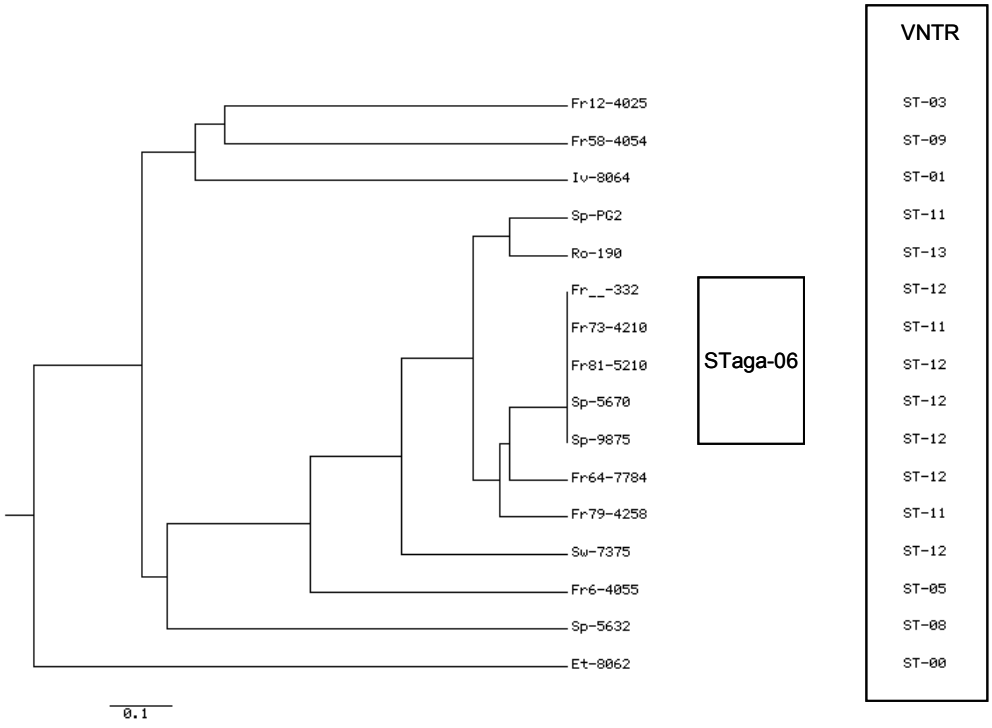


Figure 5 : Dendrogramme représentant la diversité génétique de 16 isolats de *M. agalactiae* d'origines variées basé sur la comparaison des profils AGA.

Le pays d'origine est indiqué avant le numéro d'isolat (Fr, France ; Iv, Côte d'Ivoire ; Ro, Roumanie ; Sp, Espagne ; Sw, Suisse), lorsqu'il est connu pour les isolats français le code minéralogique du département d'origine est également précisé. Arbre construit avec le programme Phylip, méthode UPGMA.

Unicité des profils AGA dans les Pyrénées-Atlantiques de 1977 à 2007

Pour évaluer le pouvoir de discrimination des profils AGA au sein d'une zone d'endémie et leur éventuelle évolution au fil du temps, nous avons étudié les profils AGA sur les isolats provenant des Pyrénées-Atlantiques de 1977 à 2007 (panel 2). Les soixante-trois isolats testés ne présentent aucune variation de profil avec les onze sondes AGA (exemple de la sonde AGA-49, [Figure 6]). Tous ces isolats présentent un profil a1 identique à PG2 avec neuf sondes AGA et un profil a2 pour les sondes AGA-10 et AGA-58. Ils sont ainsi tous de type STaga-07. Comme l'analyse des VNTR, les profils AGA ne permettent donc pas une discrimination des isolats de *M. agalactiae* provenant des Pyrénées-Atlantiques. Ce résultat conforte l'hypothèse de la circulation d'une seule et même souche endémique dans ce département au moins de 1977 jusqu'à nos jours.

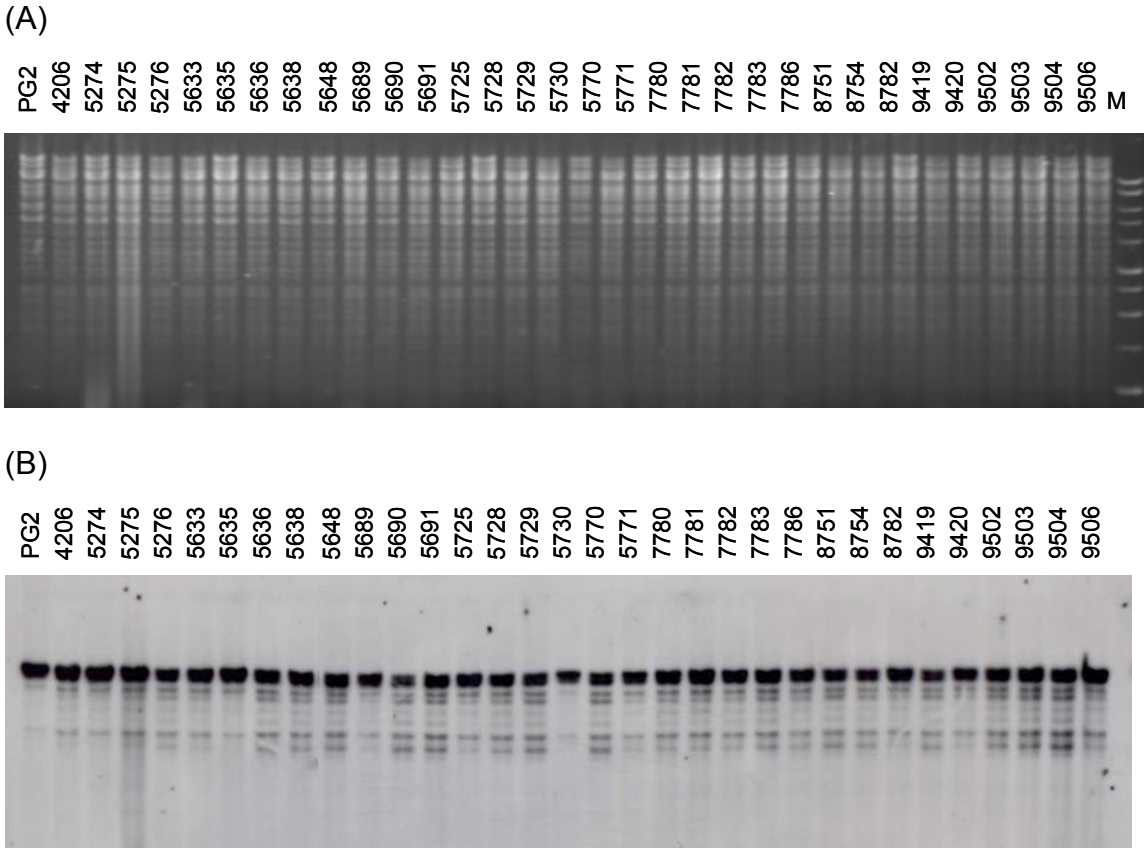


Figure 6 : Profils observés pour la sonde AGA-49 sur les isolats des Pyrénées-Atlantiques.

Profils de restriction *EcoR*I d'un panel de 32 isolats provenant des Pyrénées-Atlantiques (A) gel d'agarose 1% et coloration au bromure d'éthidium. (B) hybridation *Southern blot* avec la sonde ADN AGA-49. M, marqueur de taille (*Smart Ladder* Eurogentec).

Variation du répertoire des gènes *vpma* chez *M. agalactiae*

Si les réarrangements d'ADN spécifiques qui entraînent la variation d'expression sont hautement fréquents à l'intérieur du locus *vpma*, le nombre et la nature des gènes présents semblent quant à eux stables au sein des cellules filles issues d'une même souche [16]. La souche de référence PG2 comprend un répertoire de six gènes au locus *vpma*. Pour compléter les deux approches VNTR et SSH qui observent des sites variables répartis sur l'ensemble du génome, nous avons étudié les variations du répertoire des gènes *vpma* potentiellement exprimables au locus *vpma* des isolats des panels 1 et 2 par *Southern blot*.

Méthode de détermination du répertoire *vpma* d'une souche de *M. agalactiae*

Les profils de restriction *Ase*I et *Alw*I sont constants sur des variants isogéniques présentant différentes organisations des gènes *vpma* [16,18]. Ces deux enzymes digèrent chaque gène *vpma* de PG2 dans une région conservée en 5' et donnent ainsi des fragments qui contiennent un seul gène entier pour la plupart des gènes *vpma* de PG2 (à quelques exceptions pour certains gènes en orientation opposée) [Figure 7(A)]. Comme cette caractéristique du découpage de l'ADN au locus *vpma* est constante quels que soient les réarrangements génétiques survenant au locus, la détection et l'identification des fragments d'ADN porteurs de *vpma* peuvent être réalisées par *Southern blot* sur des ADN digérés par *Ase*I et *Alw*I en utilisant des sondes spécifiques de chaque gène *vpma*. L'utilisation d'une sonde oligonucléotidique correspondant à la partie 5' conservée des gènes *vpma* de PG2, *a priori* conservée par tout gène *vpma*, permet la détection de l'ensemble des gènes *vpma* d'une souche y compris celle de nouveaux gènes *vpma* inconnus car non présents chez PG2.

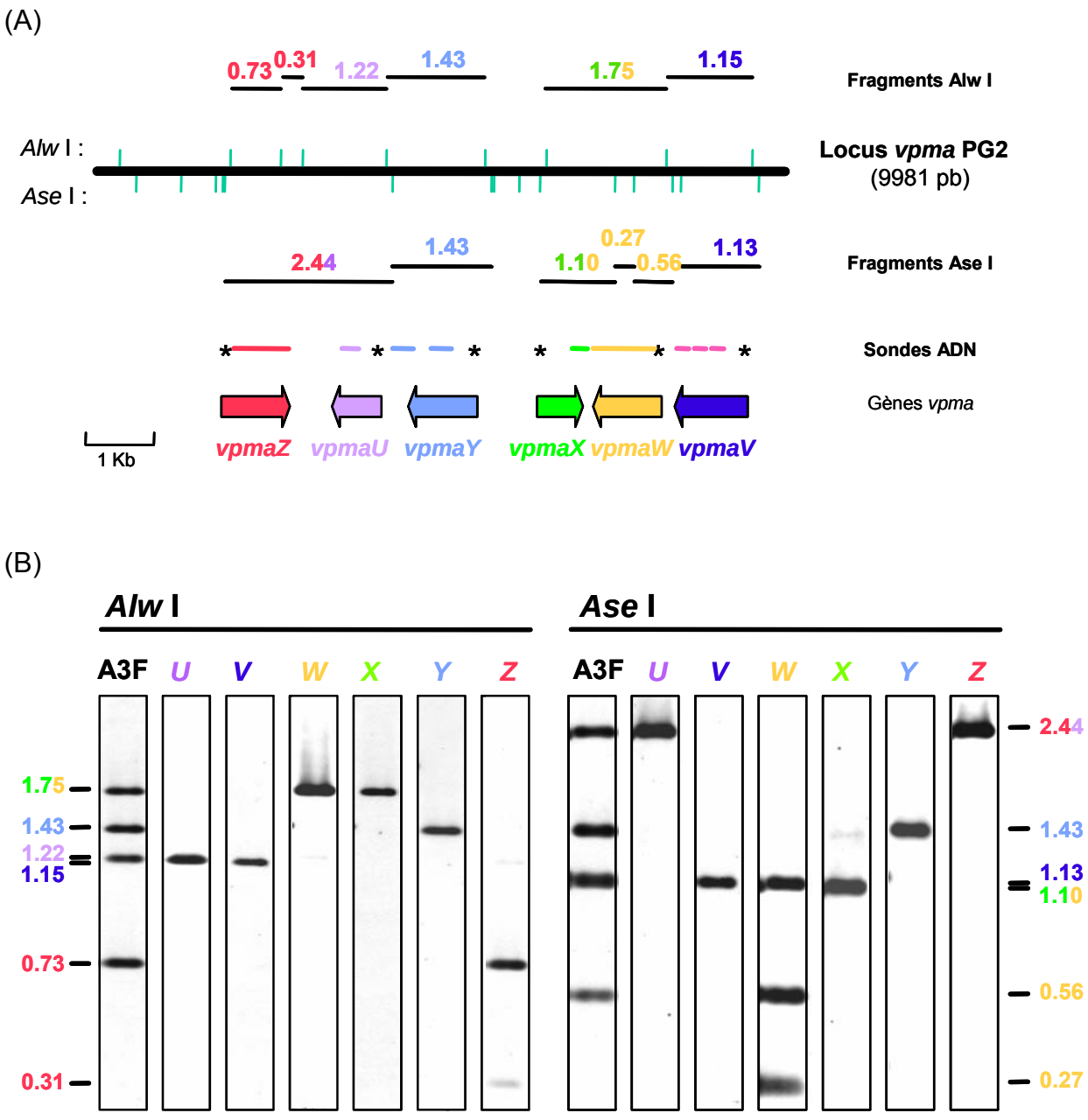


Figure 7 : Analyse du locus *vpma* de *M. agalactiae* PG2 clone 55.5 par *Southern blot*.

(A) Profil de restriction théorique du locus *vpma* de *M. agalactiae* souche PG2 clone 55.5.
La séquence du locus est représentée par un trait gras continu, les sites de coupure reconnus par les enzymes de restriction sont représentés au dessus pour *Alw*I et au dessous pour *Ase*I. Les traits fins disposés en quinconce représentent les différents fragments d'ADN obtenus après digestion enzymatique. Leur taille attendue est mentionnée (en Kb) au dessus de chaque fragment. En bas du schéma, les flèches pleines indiquent la position et l'orientation des gènes *vpma* sur la séquence. Les sondes hybridant avec les séquences spécifiques des gènes sont représentées par les traits au dessus, la sonde A3F hybridant une séquence conservée par l'ensemble des *vpma* par une asterisque. Adapté d'après [17].
(B) Exploration expérimentale du locus *vpma* par *Southern blot*.
ADN de PG2 clone 55.5 digéré par *Alw*I (à droite) ou *Ase*I (à gauche) et hybridé séquentiellement avec la sonde indiquée en haut de piste (A3F et sondes spécifiques de chacun des gènes *vpma*). La taille des fragments attendus (en Kb) est reportée sur les cotés. Adapté d'après [17].

Le locus *vpma* du clone 55.5 de la souche-type PG2, entièrement séquencé [16,40], sert de référence. Son analyse par *Southern blot* confirme la taille attendue des fragments révélés, la spécificité des sondes reconnaissant chaque *vpma* et la capacité de la sonde A3F à détecter l'ensemble des fragments porteurs de gènes *vpma* [Figure 7(B)]. La digestion de l'ADN du clone 55.5 par *Alw*I donne théoriquement six fragments contenant des séquences de gènes *vpma* dont cinq doivent hybrider avec la sonde A3F. Expérimentalement, le profil généré avec la sonde A3F (appelé "profil A3F") met en évidence quatre bandes qui correspondent à (i) un fragment de 0,73 kb contenant *vpmaZ* tronqué en 3', le fragment de 0,31 kb porteur du morceau de *vpmaZ* du côté 3' n'hybridant pas avec A3F, (ii) deux fragments de 1,22 et 1,15 kb superposés dans les conditions de migration utilisées qui contiennent respectivement *vpmaU* et *vpmaV*, (iii) un fragment de 1,43 kb contenant *vpmaY*, (iv) un fragment de 1,75 kb porteur des deux gènes *vpmaW* et *vpmaX* en orientation opposée. La digestion par *Ase*I génère aussi théoriquement six fragments contenant des séquences de gène *vpma* dont cinq doivent hybrider avec A3F. Les résultats obtenus mettent en évidence quatre bandes qui correspondent à (i) un fragment de 2,44 kb porteur des deux gènes *vpmaZ* et *vpmaU* en orientation opposée, (ii) un fragment de 1,46 kb contenant vpmaY, (iii) deux fragments non séparés dans les conditions de migration utilisées, l'un de 1,13 kb qui contient *vpmaX* additionné d'une petite partie de l'extrémité 3' de *vpmaW* et l'autre de 1,10 kb contenant *vpmaV*, (iv) un fragment de 0,56 kb correspondant à la partie 5' de *vpmaW*, le fragment de 0,27 kb contenant une portion 3' de *vpmaW* n'étant pas reconnu par A3F. La combinaison des profils *Alw*I et *Ase*I hybridés avec la sonde A3F et séquentiellement avec chacune des sondes spécifiques des *vpma* de PG2 permet de déterminer la présence de gènes semblables dans un isolat et/ou de mettre en évidence de nouveaux gènes sans similitudes avec les six rencontrés chez PG2.

Tableau 4 : Détermination du répertoire *vpma* d'isolats d'origines diverses.

n° isolats[a]	source	hôte	année d'isolement[b]	pays d'origine	Dpt[c]	*vpma*[d]							Type[f]
						U	*V*	*W*	*X*	*Y*	*Z*	*total*[e]	
PG2	*nd*	*caprin*	*1952*	*Espagne*		*1*	*1*	*1*	*1*	*1*	*1*	*6*	*1*
7327	lait	caprin	1987	Grèce		1	1	1	1	1	1	6	1
7314	lait	ovin	1986	Grèce		1	1	1	1	1	1	6	1
7784	lait	ovin	1994	France	64	1	1	1	1	1	1	6	1
5826	lait	ovin	1992	Espagne		1	1	1	1	1	1	6	1
9600	lait	caprin	<1996	Portugal		1	1	1	1	1	1	6	1
5225	lait	caprin	1991	Espagne		1	1	1	1	1	1	6	1
5210	lait	nd	1988	France	81	1	1	1	1	1	1	6	1
9385	lait	caprin	<1996	Portugal		1	1	1	1	1	1	6	1
5670	lait	caprin	<1991	Espagne		1	1	1	1	1	1	6	1
9875	lait	ovin	<1996	Espagne		1	1	1	1	1	1	6	1
332	articulation	nd	1998	France	nd	1	1	1	1	1	1	6	1
6833	lait	caprin	<1993	Italie		1	1	1	1	1	1	6	1
4021	lait	ovin	1988	France	12	1	1	1	1	1	1	6	1
4206	lait	ovin	1981	France	64	1	1	1	1	1	1	6	1
6968	lait	ovin	1993	Espagne		1	1	1	1	1	1	6	1
190	nd	nd	1951	Romania		1	1	1	1	0	1	5	2
9	lait	ovin	nd	Italie		1	0	1	1	1	1	5	3
7169	lait	caprin	<1993	Suisse		1	0	1	1	1	1	5	3
7375	lait	caprin	<1993	Suisse		1	0	1	1	1	1	5	3
4210	lait	ovin	1982	France	73	0	1	1	1	1	0	4	4
4258	lait	caprin	1989	France	79	0	1	1	1	1	0	4	4
209	lait	caprin	1983	France	31	1	1	0	1	0	0	3	5
4025	lait	ovin	1988	France	12	1	1	0	0	0	1	3	6
8064	poumon	ovin	1989	Côte d'Ivoire		1	0	1	1	0	0	3	7
4054	lait	caprin	1986	France	58	0	0	1	0	0	1	2	8
4212	lait	ovin	1986	France	84	0	0	1	0	0	1	2	8
13387	poumon	bouquetin	2003	France	73	0	0	0	1	0	1	2	9
14628	poumon	bouquetin	2006	France	73	0	0	0	1	0	1	2	9
5632	articulation	caprin	<1991	Espagne		0	0	0	0	1	1	2	10
8063[g]	nd	nd	1988	Turquie		0	1	0	0	0	0	1	11
13377	oreille	caprin	2003	France	79	0	0	1	0	0	0	1	12
14535	lait	caprin	2005	France	26	0	0	1	0	0	0	1	12
3990	nd	ovin	nd	nd		0	0	0	1	0	0	1	13
4055	lait	caprin	1987	France	06	0	0	0	1	0	0	1	13
13375	oreille	caprin	2003	France	79	0	0	0	0	0	1	1	14
8062[g]	poumon	caprin	1987	Ethiopie		0	0	0	0	0	0	0	15

Ce tableau résume les résultats obtenus par *Southern blot* avec les sondes spécifiques des *vpma* de la souche type de *M. agalactiae* PG2. [a] Le panel d'isolats est celui utilisé pour l'analyse VNTR (voir paragraphe "origines et caractéristiques des collections d'isolats étudiés" p.97) à l'exception de trois isolats (9018, 14340 et 14668) pour lesquels des difficultés de culture n'ont pas permis d'obtenir de l'ADN en quantité suffisante pour la réalisation des *Southern blots*. [b] le sigle < signifie que la date d'isolement précise est inconnue mais au moins antérieure à la date indiquée. [c] code minéralogique des départements : 06, Alpes-Maritimes ; 07, Ardèche ; 12, Aveyron ; 26, Drôme ; 31, Haute-Garonne ; 34, Hérault ; 58, Nièvre ; 64, Pyrénées-Atlantiques ; 73, Savoie ; 79, Deux-Sèvres ; 81, Tarn ; 84, Vaucluse. [d] hybridation avec chaque sonde spécifique des *vpma* de PG2 : 0, absence d'hybridation, 1 hybridation positive [e] nombre total de *vpma* de type PG2 détectées. [f] pour nommer chaque répertoire différent, un numéro leur a été attribué [g] initialement identifiée comme *M. agalactiae* mais appartenant en réalité à l'espèce *M. bovis*. nd, non déterminé. Pour le détail des profils *Ase*I / *Alw*I voir Tableau S3 des données supplémentaires.

Diversité des répertoires *vpma* parmi les isolats du panel 1

Appliquée au panel 1, la méthode de détermination du répertoire *vpma* par *Southern blot* permet de mettre en évidence des différences de répertoire entre les isolats [Tableau 4, pour détails voir Tableau S3]. Il n'a pas été possible d'obtenir un résultat pour trois isolats sur les quarante analysés à cause de difficultés de culture qui n'ont pas permis d'obtenir de l'ADN en quantité suffisante pour la réalisation des *Southern blots*. Sur les trente-six profils obtenus, aucun isolat ne présente des profils de réaction identiques à ceux observés pour PG2, cependant pour bon nombre, il s'agit de variation de taille et non de variation de

répertoire au sens strict. En effet, nos résultats indiquent que quinze des souches analysées présentent le même nombre et le même type de gènes *vpma* que PG2. Au total, sur la base du nombre et du type de *vpma*, les trente-sept souches (PG2 comprise) peuvent être séparées en quinze groupes. Le groupe majoritaire contient quinze souches qui disposent de six *vpma* de même nature que PG2. Globalement, par comparaison au répertoire de la souche PG2, quatre grandes catégories de souches se distinguent : (i) Répertoire identique à celui de PG2, (ii) Repertoire avec des gènes *vpma* manquants, (iii) Répertoire avec certains gènes *vpma* manquants et d'autres supplémentaires (iv) Répertoire ne présentant aucun gène *vpma* en commun avec la souche PG2. Pour la souche 5632, les détails du répertoire sont étudiés dans le chapitre suivant.

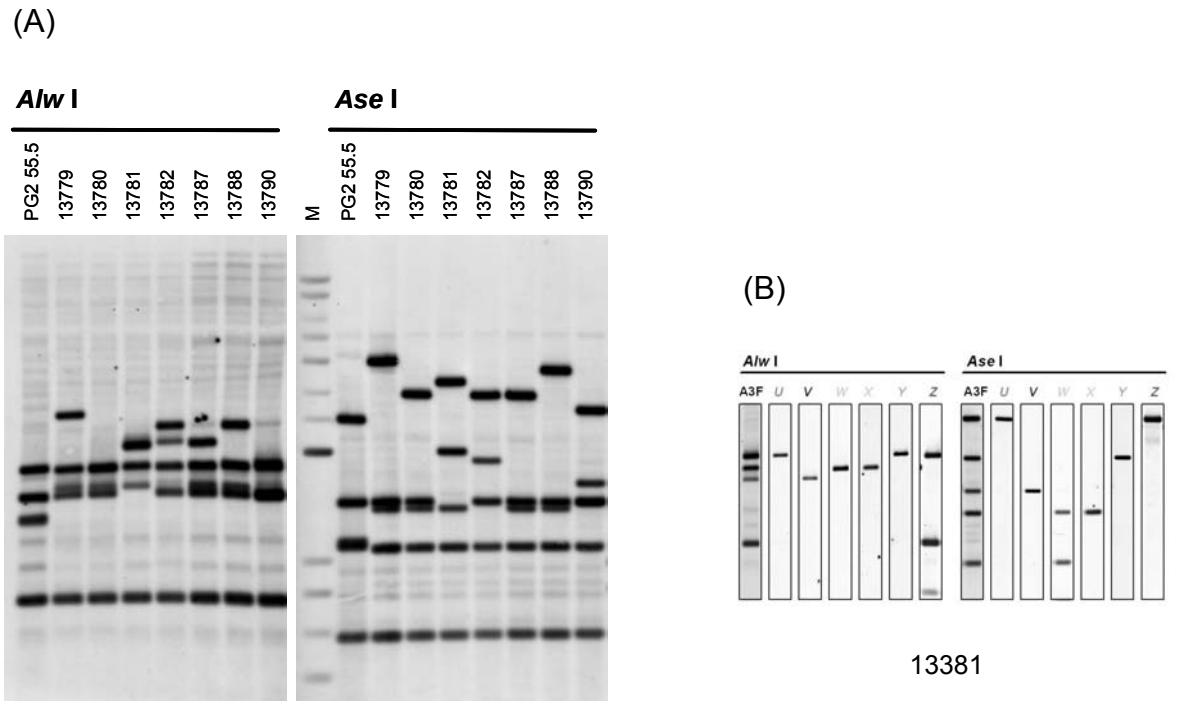


Figure 8 : Exemples de profils A3F de souches des Pyrénées-Atlantiques.

(A) Profils de restriction (*Alw*I à droite, *Ase*I à gauche) hybridés avec la sonde A3F par *Southern blot*. Les hybridations avec les sondes spécifiques des séquences des *vpma* de PG2 démontrent que les variations observées sont des variations de taille du même répertoire de gènes *vpma,* l'encart (B) montre l'exemple de l'isolat 13381.

Stabilité du répertoire *vpma* dans les Pyrénées-Atlantiques de 1977 à 2007

Pour les soixante-trois isolats du panel 2, les résultats indiquent que la majeure partie des isolats (58 sur 63) ne présente aucune variation de répertoire avec six *vpma* détectées [Tableau S4]. Les seules différences observées pour ces cinquante-huit isolats sont des différences de profils [Figure 8] correspondant à des différences de taille des fragments détectés représentant *a priori* des variants d'expression ou des variants de taille. Ces variations survenant de manière intra-clonale et réversible, elles peuvent difficilement être retenues comme traceurs épidémiologiques fiables.

Cinq isolats se distinguent par leur absence d'hybridation avec la sonde spécifique de *vpmaY*. Il s'agit d'un isolat datant de 1977 (isolat n°52 76) et de quatre isolats récents (14634, 14775, 14858, 14891). Il n'a pas été possible de remonter à l'élevage correspondant à l'isolat de 1977, en revanche, les isolats récents correspondent à trois élevages différents (élevages nommés A, B et C). Ainsi, l'isolat 14634 a été isolé en 2006 et l'isolat 14891 l'année suivante dans l'élevage A. Un lien épidémiologique avéré existe entre l'élevage A et l'élevage B par le biais d'un achat d'animaux peu avant que l'élevage A ne déclare la maladie (GDS64, communication personnelle). Ces trois élevages se situent dans la zone considérée comme zone à risque où des mesures renforcées ont été mises en place en 2007 sur les mouvements d'animaux [Figure 9]. Il est donc possible, dans un cas au moins, de corroborer un historique de contamination sur la base du répertoire *vpma*.

A l'exception de cinq isolats pour lesquels une hybridation s'est révélée négative, il semble difficile de différencier la plupart des isolats des Pyrénées-Atlantiques. Les VNTR et les profils AGA ne suggèrent aucune différence et l'analyse du répertoire *vpma* met en évidence des isolats très monomorphes sans différence majeure. L'ensemble de ces résultats est en faveur de la circulation d'une seule et même souche dans les Pyrénées-Atlantiques de 1977 à 2007.

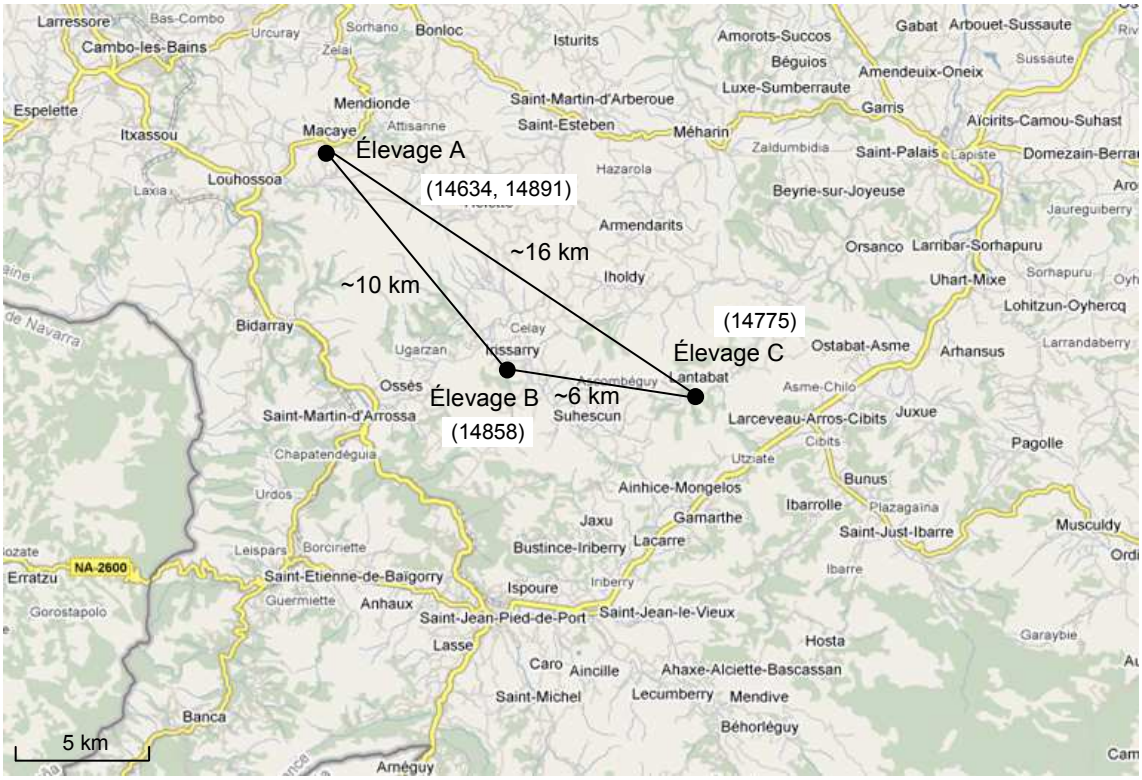


Figure 9 : Carte géographique des isolats ayant un répertoire *vpma* particulier.

Les trois élevages desquels les isolats 14634, 14775, 14858 et 14891 sont issus sont localisés par des ronds noirs, la distance qui les sépare est indiquée approximativement. Les n° des isolats sont placés à proximité des élevages desquels ils ont été isolés.

4. *Discussion*

Pour étudier l'étendue de la diversité de *M. agalactiae*, deux panels de quarante et soixante-trois isolats ont été analysés par trois stratégies : analyse des VNTR, des profils *Southern blot* générés avec des sondes issues de la SSH (sondes AGA) et détermination du répertoire *vpma* par *Southern blot* (à l'aide de sondes spécifiques de chacun de ces six gènes *vpma* de la souche type et d'une sonde commune à l'ensemble de ces six gènes).

Les résultats obtenus avec le premier panel constitué d'isolats d'origines diverses mettent en évidence l'existence d'une diversité au sein de l'espèce *M. agalactiae.* Toutefois, l'étendue de cette diversité ne concerne pas l'ensemble des souches. Il se dessine en effet deux groupes de souches : celles qui, peu polymorphes, ressemblent fortement à la souche de référence PG2 (profils VNTR ST-11 et 12 ; profils AGA STaga-3, 5 et 8 ; Répertoire *vpma* identique à PG2) et les autres, génétiquement plus éloignées et montrant un polymorphisme plus important sans une réelle unicité forte entre elles. Ce résultat confirme les observations réunies jusqu'alors : cette dichotomie a été démontrée sur des isolats inclus dans notre panel par l'expression facultative de certaines protéines immuno-dominantes comme P30 [12] et P40 [13], ou encore la distribution d'éléments génétiques mobiles comme des séquences d'insertion (IS, *Insertion Sequence*) [25,32]. Les résultats de ces études précédentes sont reportés dans le Tableau 1. Des éléments conjugatifs intégratifs (ICE, *Integrative Conjugative Element*) ont également été détectés chez les isolats 3990, 5632 et 8064 [23,24]. La comparaison par SSH réalisée entre *M. agalactiae* et *M. bovis* (PG2 *versus* PG45, et réciproquement) souligne une ressemblance accrue avec *M. bovis* pour les isolats 4025, 4054, 4055, 5632 et 8064 [24]. Nos résultats placent ces isolats parmi des isolats présentant une plus forte variabilité et plutôt éloignés de PG2. Les isolats de *M. bovis* semblent contenir des éléments génétiques mobiles actifs comme des IS et des ICE plus fréquemment que les isolats de *M. agalactiae* [24]. Aucun élément de type IS ou ICE n'a été détecté pour les isolats semblables à PG2 comme ceux de VNTR ST-11 et 12. En revanche, les isolats de *M. agalactiae*

connus comme porteurs de tels éléments font partie des isolats que nos résultats classent comme divergeant de PG2 (isolats de ST-00 (8062), ST-01 (8064), ST-03 (4025), ST-05 (209, 3990, 4055), ST-08 (5632) et ST-09 (4054 et 4212)). Cette plus forte prévalence d'éléments génétiques mobiles est sûrement en lien avec la génération d'une plus grande diversité. La participation des ICE [38] et des IS [21] à des réarrangements et transferts de gènes est bien connue. Par exemple, les comparaisons des génomes chargés d'IS comme *Bordetella* [30], *Burkholderia* [28], *Francisella* [31], *Mycobacterium* [42], *Shigella* [47] et Yersinia [7] ont révélé de multiples réarrangements associés aux IS. Ces réarrangements génomiques peuvent conduire à un accroissement de la virulence des souches et participent de manière générale à une adaptation aux conditions changeantes du milieu [22]. Si la participation des IS à la virulence n'est pas connue à ce jour pour *M. agalactiae* ou *M. bovis*, nos résultats indiquent que les éléments génétiques mobiles y sont pour le moins corrélés à la génération de diversité génétique.

Le second panel analysé est constitué d'isolats provenant exclusivement des Pyrénées-Atlantiques, département français où l'agalactie contagieuse sévit de manière endémique. Dans ce bassin d'élevage ovin, deuxième producteur de lait de brebis en France, sont élevées principalement des brebis de races locales notamment pour la production du fromage d'AOC "Ossau-Iraty". Les isolats de ce bassin peuvent ainsi être considérés comme un modèle d'étude de souches circulant dans une zone endémique ayant peu d'échange avec l'extérieur. Les résultats obtenus mettent en évidence une très faible diversité des isolats de cette région voire une constance génétique quasi-totale. En effet, les analyses VNTR et AGA ne montrent aucune différence entre les soixante-trois isolats et le repertoire *vpma* n'est différent que pour cinq isolats. A l'exception de l'isolat 5276 datant de 1977 pour lequel il n'a pas été possible de retrouver l'historique, quatre de ces cinq isolats présentent des liens épidémiologiques : tous ont été isolés entre 2006 et 2007 dans un espace de moins de 15 km^2, de plus les isolats 14634 et 14891 proviennent d'un même élevage impliqué dans la transmission de la maladie à l'élevage de l'isolat 14858 par une introduction d'animaux. Même si les données épidémiologiques sur l'ensemble des isolats étudiés font défaut, il semble possible, dans un cas au moins, de corroborer un historique de contamination sur la base du répertoire *vpma*. Ces cinq isolats se distinguent par une absence

d'hybridation avec la sonde spécifique de *vpmaY*. Du fait de la superposition de bandes sur le profil A3F, il n'est pas possible de s'assurer de l'absence de ce gène *vpma* ou de sa présence avec une séquence divergente non reconnue par la sonde spécifique de *vpmaY*. Le gène *vpmaY* présente de forte homologie de séquence avec *vpmaX*, la sonde utilisée pour hybrider *vpmaY* s'intéresse à une région divergente entre les deux gènes [16]. Cette région pourrait avoir encore plus divergé ou avoir été perdue par ces isolats. Il serait ainsi intéressant de vérifier par PCR l'existence d'une *vpma* modifiée correspondant à *vpmaY* et de la séquencer au sein de ces cinq isolats pour confirmer leur identité. Néanmoins, cette stratégie pourrait être remise en cause par les réarrangements d'ADN intra-clonaux associés aux nombreuses répétitions retrouvées dans la séquence des *vpma* qui risquent de générer un important bruit de fond et des résultats incertains, une solution pourrait être alors de séquencer entièrement le locus *vpma* de ces isolats.

En conclusion, à l'exception de cinq isolats pour lesquels une hybridation s'est révélée négative, il semble délicat de différencier les isolats des Pyrénées-Atlantiques avec les outils moléculaires actuels. Les VNTR et les profils AGA ne suggèrent aucune différence et l'analyse du répertoire *vpma* met en évidence des isolats très monomorphes. L'ensemble de ces résultats est en faveur de la circulation d'une seule et même souche endémique dans le bassin ovin des Pyrénées-Atlantiques de 1977 à 2007.

Une différence de taille des fragments hybridés avec les sondes spécifiques de chaque *vpma* de PG2 ou la sonde A3F est observée pour bon nombre d'isolats [Figure 8]. Sur la base de nos résultats et de la mécanique du système *vpma*, ces différences ne correspondraient pas à des différences de répertoire *vpma* mais à des différences dans l'arrangement synthénique des gènes *vpma* médié par la recombinase *xer1* ou à des variations de taille des *vpma* par extension ou contraction des motifs répétés qui les composent [voir chapitre Introduction]. Ces variations sont réversibles et surviennent à haute fréquence (estimée à 10^{-2} to 10^{-3} par cellule et par génération *in vitro* [9,18]), elles peuvent donc difficilement être utilisées comme marqueurs épidémiologiques. Cependant dans le contexte des Pyrénées-Atlantiques où circulent apparemment une souche unique, l'utilisation

des variations de profils A3F, même si elles surviennent à haute fréquence pourraient être envisagée pour retracer des contacts directs sur une échelle de temps réduite. Tout en tenant compte des limites liées à la réversibilité de la variation d'expression des Vpma et donc d'une possible convergence erronée des profils, cette méthode permettrait de venir en aide à l'épidémiologie classique dans certains cas particuliers. Il serait bien évidemment nécessaire de suivre des isolats au cours du temps dans un même élevage et/ou des élevages dont l'historique de la chaîne de transmission de la maladie est connu pour évaluer la pertinence de cette méthode. Hélas, les données épidémiologiques sur l'ensemble des isolats étudiés font pour la plupart défaut. A titre d'exemple cependant, les quatre isolats de 2006-2007 qui se distinguent des cinquante-huit autres par une hybridation *vpmaY* négative présentent le même profil A3F. Dans l'attente d'une étude plus exhaustive, ceci constitue un élément en faveur de cette méthode basée sur les profils A3F. En tout état de cause, sous cet angle, la variabilité du profil A3F constitue une sorte "d'adresse IP dynamique" spécifique à chaque isolat à un instant "t" mais résultant d'une combinaison de possibilités réversibles. En connaissant les limites dues à cette réversibilité de code, les variations de profil A3F pourraient donc être envisagées pour un suivi fin de souches extrêmement rapprochées dans l'espace et le temps.

En pratique, des traceurs moléculaires seraient utiles à une meilleure prévention et maîtrise de l'agalactie contagieuse. Le développement de tels outils génétiques passe obligatoirement par une phase de description de la diversité génétique des souches. Faire de l'épidémiologie moléculaire sans avoir étudié au préalable la diversité existante serait hasardeux. La connaissance du polymorphisme génétique et des mécanismes qui y conduisent permet alors de passer à l'étape suivante : l'identification et le choix de marqueurs moléculaires pertinents. Ce travail sur 101 isolats étaye les données acquises sur la diversité génétique de *M. agalactiae*. Il confirme l'existence de souches présentant globalement un faible polymorphisme excepté au niveau des loci hypervariables comme le locus *vpma*, et montre les limites de techniques moléculaires comme les VNTR et des profils *Southern blot* dans de telles conditions. A l'opposé de ces souches qui ressemblent plus particulièrement à la souche type PG2 se dégagent des souches plus éloignées qui, en comparaison, montrent une diversité accrue. Cette étendue

de diversité d'isolats appartenant clairement à l'espèce *M. agalactiae* est plus comparable à celle observée pour les isolats de *M. bovis*. L'étude de souches intermédiaires entre ces deux espèces phylogénétiquement proches pourrait apporter des éléments nouveaux sur leurs particularités et leurs points communs, notamment leur spécificité d'hôte et leurs déterminants pathogéniques.

Pour ces isolats qui diffèrent grandement de PG2, le profil A3F obtenu dans l'analyse du répertoire *vpma* est extrêmement complexe. L'hybridation avec les sondes spécifiques de chaque gène *vpma* de PG2 semble gênée par de trop grandes divergences de séquences avec les *vpma* de ces isolats. De même, le caractère universel de la sonde A3F pour la détection de l'ensemble des fragments de restriction porteurs de *vpma* atteint ses limites [Figure 10]. Ainsi, pour dépasser le pouvoir de résolution de cette méthode *Southern blot*, l'étude du répertoire *vpma* d'une souche éloignée de PG2 pourrait apporter de nouvelles vues sur la dynamique et l'étendue des répertoires *vpma*. Parmi les souches disponibles, la souche 5632 se situe apparemment à l'opposé du spectre de l'espèce *M. agalactiae* et apparaît comme un candidat intéressant pour cette étude comparative des loci *vpma*.

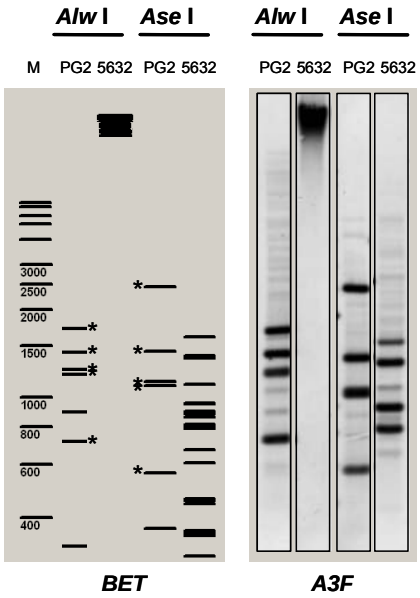


Figure 10 : Comparaison des profils de restriction des loci *vpma* entre les souches PG2 et 5632.

A gauche : migration théorique des fragments de restriction (digestions par *Alw*I et *Ase*I) des loci *vpma* de PG2 et 5632 en gel d'agarose révélés au bromure d'éthidium (BET). Les fragments reconnus par la sonde A3F sont repérés par un astérisque. A droite : analyse de ces mêmes profils de restriction hybridés par *Southern blot* avec la sonde A3F marquée à la digoxygénine. Les deux images ont été normalisées à des échelles comparables.

Références

[1] Aluotto B.B., Wittler R.G., Williams C.O., Faber J.E., Standardized bacteriologic techniques for the characterization of mycoplasma species., Int J Syst Bacteriol 20 (1970) 35-58.

[2] Bashiruddin J.B., Frey J., Konigsson M.H., Johansson K.E., Hotzel H., Diller R., de Santis P., Botelho A., Ayling R.D., Nicholas R.A., Thiaucourt F., Sachse K., Evaluation of PCR systems for the identification and differentiation of *Mycoplasma agalactiae* and *Mycoplasma bovis*: a collaborative trial, Vet J 169 (2005) 268-75.

[3] Bergonier D., Berthelot X., Mycoplasmoses des petits ruminants : Le syndrome de l'agalactie contagieuse, Bulletin de l'Académie vétérinaire de France 161 (2008) 167-177.

[4] Bergonier D., Berthelot X. (2008) L'agalactie contagieuse, principal syndrome mycoplasmique des petits ruminants. In: Académie vétérinaire de France.

[5] Bergonier D., De Simone F., Russo P., Solsona M., Lambert M., Poumarat F., Variable expression and geographic distribution of *Mycoplasma agalactiae* surface epitopes demonstrated with monoclonal antibodies, FEMS Microbiol Lett 143 (1996) 159-65.

[6] Bogush M.L., Velikodvorskaya T.V., Lebedev Y.B., Nikolaev L.G., Lukyanov S.A., Fradkov A.F., Pliyev B.K., Boichenko M.N., Usatova G.N., Vorobiev A.A., Andersen G.L., Sverdlov E.D., Identification and localization of differences between *Escherichia coli* and *Salmonella typhimurium* genomes by suppressive subtractive hybridization, Mol Gen Genet 262 (1999) 721-9.

[7] Chain P.S., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O., Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L., Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C., Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M., Derbise A., Hauser L.J., Garcia E., Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*, Proc Natl Acad Sci U S A 101 (2004) 13826-31.

[8] Chen W.P., Kuo T.T., A simple and rapid method for the preparation of gram-negative bacterial genomic DNA, Nucleic Acids Res 21 (1993) 2260.

[9] Chopra-Dewasthaly R., Citti C., Glew M.D., Zimmermann M., Rosengarten R., Jechlinger W., Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation, Mol Microbiol 67 (2008) 1196-210.

[10] Church G.M., Gilbert W., Genomic sequencing, Proc Natl Acad Sci U S A 81 (1984) 1991-5.

[11] Corrales J.C., Esnal A., De la Fe C., Sanchez A., Assuncao P., Poveda J.B., Contreras A., Contagious agalactia in small ruminants, Small Ruminant Res 68 (2007) 154-166.

[12] Fleury B., Bergonier D., Berthelot X., Schlatter Y., Frey J., Vilei E.M., Characterization and analysis of a stable serotype-associated membrane protein (P30) of *Mycoplasma agalactiae*, J Clin Microbiol 39 (2001) 2814-22.

[13] Fleury B., Bergonier D., Berthelot X., Peterhans E., Frey J., Vilei E.M., Characterization of P40, a cytadhesin of *Mycoplasma agalactiae*, Infect Immun 70 (2002) 5612-21.

[14] Frothingham R., Meeker-O'Connell W.A., Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats, Microbiology 144 (Pt 5) (1998) 1189-96.

[15] GDS64, Agalactie contagieuse ovine.Rapport annuel, Groupement de défense sanitaire du Béarn et du Pays-Basque, Pau (France), 2009, pp. 11-14.

[16] Glew M.D., Marenda M., Rosengarten R., Citti C., Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus, J Bacteriol 184 (2002) 5987-98.

[17] Glew M.D., Marenda M., Bergonier D., Rosengarten R., Citti C., The vpma gene repertoires in *Mycoplasma agalactiae*. in: Poveda, J.B. et al. (Eds.), Mycoplasmas of ruminants: pathogenicity, diagnostics, epidemiology and molecular genetics (COST Action 826 - EUR 19693), Office for Official Publications of the European Communities, 2001, pp.

[18] Glew M.D., Papazisi L., Poumarat F., Bergonier D., Rosengarten R., Citti C., Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*, Infect Immun 68 (2000) 4539-48.

[19] Klevytska A.M., Price L.B., Schupp J.M., Worsham P.L., Wong J., Keim P., Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome, J Clin Microbiol 39 (2001) 3179-85.

[20] Lista F., Faggioni G., Valjevac S., Ciammaruconi A., Vaissaire J., le Doujet C., Gorge O., De Santis R., Carattoli A., Ciervo A., Fasanella A., Orsini F., D'Amelio R., Pourcel C., Cassone A., Vergnaud G., Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis, BMC Microbiol 6 (2006) 33.

[21] Mahillon J., Chandler M., Insertion sequences, Microbiol Mol Biol Rev 62 (1998) 725-74.

[22] Mahillon J., Leonard C., Chandler M., IS elements as constituents of bacterial genomes, Res Microbiol 150 (1999) 675-87.

[23] Marenda M., Barbe V., Gourgues G., Mangenot S., Sagne E., Citti C., A new integrative conjugative element occurs in *Mycoplasma agalactiae* as chromosomal and free circular forms, J Bacteriol 188 (2006) 4137-41.

[24] Marenda M.S., Sagne E., Poumarat F., Citti C., Suppression subtractive hybridization as a basis to assess *Mycoplasma agalactiae* and *Mycoplasma*

bovis genomic diversity and species-specific sequences, Microbiology 151 (2005) 475-89.

[25] Marenda M.S., Vilei E.M., Poumarat F., Frey J., Berthelot X., Validation of the suppressive subtractive hybridization method in *Mycoplasma agalactiae* species by the comparison of a field strain with the type strain PG2, Vet Res 35 (2004) 199-212.

[26] McAuliffe L., Ayling R.D., Nicholas R.A., Identification and characterization of variable-number tandem-repeat markers for the molecular epidemiological analysis of *Mycoplasma mycoides* subspecies *mycoides* SC, FEMS Microbiol Lett 276 (2007) 181-8.

[27] McAuliffe L., Churchward C.P., Lawes J.R., Loria G., Ayling R.D., Nicholas R.A., VNTR analysis reveals unexpected genetic diversity within *Mycoplasma agalactiae*, the main causative agent of contagious agalactia, BMC Microbiol 8 (2008) 193.

[28] Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E., Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C., Davidsen T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R., Mohammoud Y., Nelson W.C., Radune D., Romero C.M., Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N., Zhou L., Fraser C.M., Structural flexibility in the *Burkholderia mallei* genome, Proc Natl Acad Sci U S A 101 (2004) 14246-51.

[29] Noller A.C., McEllistrem M.C., Shutt K.A., Harrison L.H., Locus-specific mutational events in a multilocus variable-number tandem repeat analysis of *Escherichia coli* O157:H7, J Clin Microbiol 44 (2006) 374-7.

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

[31] Petrosino J.F., Xiang Q., Karpathy S.E., Jiang H., Yerrapragada S., Liu Y., Gioia J., Hemphill L., Gonzalez A., Raghavan T.M., Uzman A., Fox G.E., Highlander S., Reichard M., Morton R.J., Clinkenbeard K.D., Weinstock G.M., Chromosome rearrangement and diversification of *Francisella tularensis* revealed by the type B (OSU18) genome sequence, J Bacteriol 188 (2006) 6977-85.

[32] Pilo P., Fleury B., Marenda M., Frey J., Vilei E.M., Prevalence and distribution of the insertion element ISMag1 in *Mycoplasma agalactiae*, Vet Microbiol 92 (2003) 37-48.

[33] Radnedge L., Gamez-Chin S., McCready P.M., Worsham P.L., Andersen G.L., Identification of nucleotide sequences for the specific and rapid detection of *Yersinia pestis*, Appl Environ Microbiol 67 (2001) 3759-62.

[34] Rocha E.P., Blanchard A., Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution, Nucleic Acids Res 30 (2002) 2031-42.

[35] Sabat A., Krzyszton-Russjan J., Strzalka W., Filipek R., Kosowska K., Hryniewicz W., Travis J., Potempa J., New method for typing *Staphylococcus aureus* strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates, J Clin Microbiol 41 (2003) 1801-4.

[36] Sachse K., Helbig J.H., Lysnyansky I., Grajetzki C., Muller W., Jacobs E., Yogev D., Epitope mapping of immunogenic and adhesive structures in repetitive domains of *Mycoplasma bovis* variable surface lipoproteins, Infect Immun 68 (2000) 680-7.

[37] Sambrook J., Fritsch E.F., Maniatis T., Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, N.Y., 1989.

[38] Seth-Smith H., Croucher N.J., Genome watch: breaking the ICE, Nat Rev Microbiol 7 (2009) 328-9.

[39] Simon J.-L. (2006) Prophylaxie de l'agalaxie contagieuse caprine à *Mycoplasma agalactiae* en Savoie et Haute Savoie. 1990-2006. In: Coordination Nationale de lutte Agalactie Contagieuse et mycoplasmoses des petits ruminants, FRGDS Rhône-Alpes, Toulouse (France).

[40] Sirand-Pugnet P., Lartigue C., Marenda M., Jacob D., Barre A., Barbe V., Schenowitz C., Mangenot S., Couloux A., Segurens B., de Daruvar A., Blanchard A., Citti C., Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome, PLoS Genet 3 (2007) e75.

[41] Solsona M., Lambert M., Poumarat F., Genomic, protein homogeneity and antigenic variability of *Mycoplasma agalactiae*, Vet Microbiol 50 (1996) 45-58.

[42] Stinear T.P., Seemann T., Pidot S., Frigui W., Reysset G., Garnier T., Meurice G., Simon D., Bouchier C., Ma L., Tichit M., Porter J.L., Ryan J., Johnson P.D., Davies J.K., Jenkin G.A., Small P.L., Jones L.M., Tekaia F., Laval F., Daffe M., Parkhill J., Cole S.T., Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer, Genome Res 17 (2007) 192-200.

[43] Supply P., Magdalena J., Himpens S., Locht C., Identification of novel intergenic repetitive units in a mycobacterial two-component system operon, Mol Microbiol 26 (1997) 991-1003.

[44] Tola S., Idini G., Manunta D., Casciano I., Rocchigiani A.M., Angioi A., Leori G., Comparison of *Mycoplasma agalactiae* isolates by pulsed field gel electrophoresis, SDS-PAGE and immunoblotting, FEMS Microbiol Lett 143 (1996) 259-65.

[45] Tully J.G., Culture medium formulation for primary isolation and maintenance of mollicutes. in: Tully, J.G. (Ed.) Molecular and Diagnostic

Procedures in Mycoplasmology: Molecular Characterization, Academic Press., San Diego, 1995, pp. 33–39.

[46] van Belkum A., Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA), FEMS Immunol Med Microbiol 49 (2007) 22-7.

[47] Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X., Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S., Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y., Qiang B., Hou Y., Yu J., Jin Q., Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery, Nucleic Acids Res 33 (2005) 6445-58.

Données supplémentaires

Tableau S1 : Profils VNTR des isolats des Pyrénées-Atlantiques de 1977 à 2007.

n° isolat	source	hôte	année d'isolement	pays d'origine	Dpt	VNTR 5	VNTR 14	VNTR 17	VNTR 19
4206	lait	ovin	1981	France	64	p1	p1	p3	p1
5274	lait	ovin	1983	France	64	p1	p1	p3	p1
5275	lait	ovin	1983	France	64	p1	p1	p3	p1
5276	lait	ovin	1977	France	64	p1	p1	p3	p1
5633	lait	ovin	1991	France	64	p1	p1	p3	p1
5635	lait	ovin	1990	France	64	p1	p1	p3	p1
5636	lait	ovin	1991	France	64	p1	p1	p3	p1
5638	lait	ovin	1991	France	64	p1	p1	p3	p1
5648	lait	ovin	1991	France	64	p1	p1	p3	p1
5689	lait	caprin	1991	France	64	p1	p1	p3	p1
5690	lait	caprin	1991	France	64	p1	p1	p3	p1
5691	lait	caprin	1991	France	64	p1	p1	p3	p1
5725	lait	ovin	1990	France	64	p1	p1	p3	p1
5728	lait	ovin	1991	France	64	p1	p1	p3	p1
5729	lait	ovin	1991	France	64	p1	p1	p3	p1
5730	lait	ovin	1991	France	64	p1	p1	p3	p1
5770	lait	ovin	1991	France	64	p1	p1	p3	p1
5771	lait	ovin	1991	France	64	p1	p1	p3	p1
7780	lait	ovin	1994	France	64	p1	p1	p3	p1
7781	lait	ovin	1994	France	64	p1	p1	p3	p1
7782	lait	ovin	1994	France	64	p1	p1	p3	p1
7783	lait	ovin	1994	France	64	p1	p1	p3	p1
7784	lait	ovin	1994	France	64	p1	p1	p3	p1
7786	lait	ovin	1994	France	64	p1	p1	p3	p1
8750	lait	ovin	1995	France	64	p1	p1	p3	p1
8751	lait	ovin	1995	France	64	p1	p1	p3	p1
8754	lait	ovin	1995	France	64	p1	p1	p3	p1
8782	lait	ovin	1995	France	64	p1	p1	p3	p1
9419	lait	ovin	1993	France	64	p1	p1	p3	p1
9420	lait	ovin	1993	France	64	p1	p1	p3	p1
9502	lait	ovin	1993	France	64	p1	p1	p3	p1
9503	lait	ovin	1993	France	64	p1	p1	p3	p1
9504	lait	ovin	1993	France	64	p1	p1	p3	p1
9506	lait	ovin	1994	France	64	p1	p1	p3	p1
9507	lait	ovin	1994	France	64	p1	p1	p3	p1
9508	lait	ovin	1994	France	64	p1	p1	p3	p1
9509	lait	ovin	1994	France	64	p1	p1	p3	p1
9512	lait	ovin	1995	France	64	p1	p1	p3	p1
9523	lait	ovin	1995	France	64	p1	p1	p3	p1
9807	lait	ovin	1995	France	64	p1	p1	p3	p1
9808	lait	ovin	1995	France	64	p1	p1	p3	p1
9809	lait	ovin	1995	France	64	p1	p1	p3	p1
9810	lait	ovin	1995	France	64	p1	p1	p3	p1
9871	lait	ovin	1995	France	64	p1	p1	p3	p1
9872	lait	ovin	1995	France	64	p1	p1	p3	p1
13779	lait	ovin	2004	France	64	p1	p1	p3	p1
13780	lait	ovin	2004	France	64	p1	p1	p3	p1
13781	lait	ovin	2004	France	64	p1	p1	p3	p1
13782	lait	ovin	2004	France	64	p1	p1	p3	p1
13787	lait	ovin	2004	France	64	p1	p1	p3	p1
13788	lait	ovin	2004	France	64	p1	p1	p3	p1
14551	lait	ovin	2005	France	64	p1	p1	p3	p1
14557	lait	ovin	2005	France	64	p1	p1	p3	p1
14558	lait	ovin	2005	France	64	p1	p1	p3	p1
14564	lait	ovin	2005	France	64	p1	p1	p3	p1
14633	lait	caprin	2006	France	64	p1	p1	p3	p1
14634	lait	ovin	2006	France	64	p1	p1	p3	p1
14775	lait	ovin	2006	France	64	p1	p1	p3	p1
14776	lait	ovin	2007	France	64	p1	p1	p3	p1
14858	lait	ovin	2007	France	64	p1	p1	p3	p1
14880	lait	ovin	2007	France	64	p1	p1	p3	p1
14891	lait	ovin	2007	France	64	p1	p1	p3	p1
14899	lait	ovin	2007	France	64	p1	p1	p3	p1

Tous les 63 isolats des Pyrénées-Atlantiques analysés présente le ST-12. Classement par n° d'isolat. p 1, profil 1, p3, profil 3.

Tableau S2 : Hybridation des 59 sondes AGA avec 16 isolats d'origines variées.

	n° isolats															
Sondes	**PG2**	5632	190	332	4210	4258	5210	5670	7375	7784	9875	4025	4054	4055	8064	8062
AGA-1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-4	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-
AGA-7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-12	+	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-
AGA-13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
AGA-17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-18	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-20	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-21	+	-	+	+	+	+	+	+	+	+	+	-	+	-	+	-
AGA-22	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-23	+	-	+	+	+	+	+	+	+	+	+	+	+	-	-	-
AGA-24	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-26	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-27	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-
AGA-29	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-30	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-32	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-33	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-34	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-35	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-
AGA-36	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-38	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-40	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-41	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-42	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-43	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-45	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-47	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-48	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	-
AGA-49	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-52	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-53	+	-	+	+	+	+	+	+	+	+	+	-	-	+	-	-
AGA-58	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-59	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-65	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-66	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-68	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-71	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	-
AGA-72	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
AGA-74	+	-	+	+	+	+	+	+	+	+	+	-	-	+	-	-
AGA-76	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-77	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
AGA-78	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-79	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-80	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-81	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-82	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-
AGA-83	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-
AGA-86	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-
AGA-87	+	-	+	+	+	+	+	+	+	+	+	-	-	-	-	-
AGA-88	+	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-
AGA-89	+	-	+	+	+	+	+	+	+	+	+	-	-	+	-	+
AGA-90	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

+ : hybridation positive, - pas d'hybridation

Tableau S3 : Détermination du répertoire *vpma* d'isolats d'origines géographiques diverses.

n° isolats[a]	source	hôte	année d'isolement[b]	pays d'origine	Dpt	Sondes[c]												A3F bandes surnuméraires		Gènes *vpma* non detectés	Type[d]
						vpmaU		*vpmaV*		*vpmaW*		*vpmaX*		*vpmaY*		*vpmaZ*					
						AlwI	*AseI*	*AlwI*	*AseI*	*AlwI*	*AseI*	*AlwI*	*AseI*	*AlwI*	*AseI*	*AlwI*	*AseI*	*AlwI*	*AseI*		
PG2	*nd*	*caprin*	*1952*	*Espagne*		+	+	+	+	+	+	+	+	+	+	+	+	*0*	*0*	*0*	*1*
7327	lait	caprin	1987	Grèce		+	+	↑	↑	+	+	+	+	+	+	+	+	0	0	0	1
7314	lait	ovin	1986	Grèce		+	+	↑	↑	↑	+	↑	+	+	+	+	+	0	0	0	1
7784	lait	ovin	1994	France	64	↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
5826	lait	ovin	1992	Espagne		↑	+	+	+	↑	+	↑	+	+	2(↓,*↓)	2(↑,↓)	+	0	0	0	1
9600	lait	caprin	<1996	Portugal		↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
5225	lait	caprin	1991	Espagne		↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
5210	lait	nd	1988	France	81	↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
9385	lait	caprin	<1996	Portugal		↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
5670	lait	caprin	<1991	Espagne		↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
9875	lait	ovin	<1996	Espagne		↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
332	articulation	nd	1998	France	nd	↑	↑	↑	↑	+	+	+	+	+	+	2(+,↑)	↑	0	0	0	1
6833	lait	caprin	<1993	Italie		+	+	↑	↑	↓	•+	↓	+	↑	↑	+	+	0	0	0	1
4021	lait	ovin	1988	France	12	↑	↑	↑	↑	+	+	+	+	+	2(↓,*↓)	2(+,↑)	↑	0	0	0	1
4206	lait	ovin	1981	France	64	↑	↑	↑	↑	+	+	+	+	↓	↓	2(+,↑)	↑	0	0	0	1
6968	lait	ovin	1993	Espagne		↑	↑	↑	↑	↑	+	↑	+	↓	↓	2(+,↑)	↑	0	0	0	1
190	nd	nd	1951	Romanie		+	+	↑	↑	+	+	+	+	-	-	+	+	0	0	1	2
9	lait	ovin	nd	Italie		+	+	-	-	+	+	+	+	+	+	+	+	0	0	1	3
7169	lait	caprin	<1993	Suisse		+	+	-	-	+	+	+	+	+	+	+	+	0	0	1	3
7375	lait	caprin	<1993	Suisse		+	+	-	-	+	+	+	+	+	+	+	+	0	0	1	3
4210	lait	ovin	1982	France	73	-	-	+	+	2(+,↓)	+	2(+,↓)	+	+	+	-	-	0	0	2	4
4258	lait	caprin	1989	France	79	-	-	+	+	↓	•+	↓	+	2(+,↓)	2(+,↓)	-	-	0	0	2	4
209	lait	caprin	1983	France	31	M	↓	M	*+f	M	-	M	*+	M	-	M	-	M	3	3	5
4025	lait	ovin	1988	France	12	M	*↓	M	2(*↑,+)	M	-	M	-	M	-	M	*↓	M	3	3	6
8064	poumon	ovin	1989	C. d'Ivoire		2(↑,↑)	↓	-	-	*↓	*•+	*↓	*+	-	-	-	-	3	3	3	7
4054	lait	caprin	1986	France	58	-	-	-	-	*+	*•+f	-	-	-	-	*•↑f	2(↓,↓)	4	5	4	8
4212	lait	ovin	1986	France	84	-	-	-	-	*+	*•+f	-	-	-	-	*•↑f	2(*↓,↓)	4	5	4	8
13387	poumon	bouquetin	2003	France	73	M	-	M	-	M	-	M	+f	M	-	M	2(*↓,↓)	M	2	4	9
14628	poumon	bouquetin	2006	France	73	M	-	M	-	M	-	M	+f	M	-	M	2(↓,↓)	M	2	4	9
5632	articulation	caprin	<1991	Espagne		M	-	M	-	M	-	M	-	M	↓	M	↓	M	3	4	10
8063[e]	nd	nd	1988	Turquie		M	-	M	*↑f	M	-	M	-	M	-	M	-	M	0	5	11
13377	oreille	caprin	2003	France	79	M	-	M	-	M	2(+f,*↓)	M	-	M	-	M	-	M	9	5	12
14535	lait	caprin	2005	France	26	-	-	-	-	*↓	*↓	-	-	-	-	-	-	1	2	5	12
3990	nd	ovin	nd	nd		M	-	M	-	M	-	M	*+	M	-	M	-	M	3	5	13
4055	lait	caprin	1987	France	06	M	-	M	-	M	-	M	+	M	-	M	-	M	3	5	13
13375	oreille	caprin	2003	France	79	-	-	-	-	-	-	-	-	-	-	*+	*+	1	0	5	14
8062[e]	poumon	caprin	1987	Ethiopie		M	-	M	-	M	-	M	-	M	-	M	-	M	0	6	15

Tableau S3 (légende).

Ce tableau résume le détail des profils *Ase*I / *Alw*I obtenus par *Southern blot* avec les sondes spécifiques des *vpma* de la souche type de *M. agalactiae* PG2 sur un panel d'isolats d'origines diverses. [a] Le panel d'isolat est celui utilisé pour l'analyse VNTR (voir paragraphe "origines et caractéristiques des collections d'isolats étudiés" p.97) à à l'exception de 3 isolats (9018, 14340 et 14668) pour lesquels des difficultés de culture n'ont pas permis d'obtenir de l'ADN en quantité suffisante pour la réalisation des *Southern blots*. La référence pour toutes les comparaisons est la souche PG2. [b] le sigle < signifie que la date d'isolement précise est inconnue mais au moins antérieure à la date indiquée. [c] hybridation avec chaque sonde spécifique des *vpma* de PG2 : (+) fragment(s) détecté(s) et de même taille que pour PG2 ; (-) pas de fragment détecté ; (↑) fragment détecté de taille supérieure en comparaison à celui détecté pour PG2 ; (↓) fragment détecté de taille inférieure par rapport à PG2 ; (2) détection d'une bande surnuméraire par rapport à PG2 ; (*) fragment détecté avec une sonde spécifique mais non détecté avec la sonde A3F ; (•) absence d'un fragment en comparaison à PG2 ; (f) hybridation faiblement détectée ; (M) ADN non digéré par *Alw*I en raison de méthylation de type Dam. [d] pour nommer chaque répertoire différent, un numéro leur a été attribué. [e] initialement identifiée comme *M. agalactiae* mais appartenant en réalité à l'espèce *M. bovis*. nd, non déterminé.

Tableau S4 : Détermination du répertoire *vpma* des isolats des Pyrénées-Atlantiques de 1977 à 2007.

n°isolat	source	hôte	année d'isolement	pays d'origine	Dpt	*vpmaU*		*vpmaV*		*vpmaW*		*vpmaX*		*vpmaY*		*vpmaZ*		A3F bandes surnuméraires		Gènes *vpma* non détectés
						AlwI	***AseI***	***AlwI***	***AseI***	***AlwI***	***AseI***	***AlwI***	***AseI***	***AlwI***	***AseI***	***AlwI***	***AseI***	***AlwI***	***AseI***	
PG2	*nd*	*caprin*	*1952*	*Espagne*		↓	↓	↓	↓	+	+	+	+	+	+	+	+	*0*	*0*	*0*
5276	lait	ovin	1977	France	64	+	+	+	+	+	+	+	+	-	-	+	+	0	0	1
4206	lait	ovin	1981	France	64	+	+	+	+	+	+	+	+	↓	↓	+	+	0	0	0
5274	lait	ovin	1983	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5275	lait	ovin	1983	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5635	lait	ovin	1990	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
5725	lait	ovin	1990	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
5633	lait	ovin	1991	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
5636	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5638	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5648	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5689	lait	caprin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5690	lait	caprin	1991	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
5691	lait	caprin	1991	France	64	+,↓	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5728	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5729	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5730	lait	ovin	1991	France	64	+	+	↑	↑	+	+	+	+	+	+	+	+	0	0	0
5770	lait	ovin	1991	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
5771	lait	ovin	1991	France	64	↑	↑	+	+	↑	+	↑	+	+	+	↑	↑	0	0	0
7780	lait	ovin	1994	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
7781	lait	ovin	1994	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
7782	lait	ovin	1994	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
7783	lait	ovin	1994	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
7784	lait	ovin	1994	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
7786	lait	ovin	1994	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
8750	lait	ovin	1995	France	64	↑	↑	+	+	+	+	↑	+	+	+	↑	↑	0	0	0
8751	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
8754	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
8782	lait	ovin	1995	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
9419	lait	ovin	1993	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9420	lait	ovin	1993	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9502	lait	ovin	1993	France	64	+	+	↓	↓	+	+	+	+	+	+	+	+	0	0	0
9503	lait	ovin	1993	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9504	lait	ovin	1993	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9506	lait	ovin	1994	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
9507	lait	ovin	1994	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0

Tableau S4 (suite et fin).

n°isolat	source	hôte	année d'isolement	pays d'origine	Dpt	*vpmaU*		*vpmaV*		*vpmaW*		*vpmaX*		*vpmaY*		*vpmaZ*		A3F bandes surnuméraires		Gènes *vpma* non détectés
						***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	***Alw*I**	***Ase*I**	
9508	lait	ovin	1994	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9509	lait	ovin	1994	France	64	+	+	↑	↑	+	+	+	+	+	+	+	+	0	0	0
9512	lait	ovin	1995	France	64	+	+	↑	↑	+	+	+	+	+	+	+	+	0	0	0
9523	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9807	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9808	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9809	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9810	lait	ovin	1995	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
9871	lait	ovin	1995	France	64	↑	↑	+	+	↑	+	↑	+	+	+	↑	↑	0	0	0
9872	lait	ovin	1995	France	64	+	+	↑	↑	+	+	+	+	+	+	+	+	0	0	0
13779	lait	ovin	2004	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
13780	lait	ovin	2004	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
13781	lait	ovin	2004	France	64	↑	↑	+	+	+	+	+	+	↑	↑	↑	↑	0	0	0
13782	lait	ovin	2004	France	64	+	+	↑	↑	↑	+	↑	+	+	+	+	+	0	0	0
13787	lait	ovin	2004	France	64	+	+	+	+	↑	+	↑	+	+	+	+	+	0	0	0
13788	lait	ovin	2004	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
14551	lait	ovin	2005	France	64	↑	↑	↑	↑	+	+	+	+	+	+	↑	↑	0	0	0
14557	lait	ovin	2005	France	64	↑	↑	+	+	+	+	+	+	+	+	↑	↑	0	0	0
14558	lait	ovin	2005	France	64	+	+	↑	↑	↑	+	↑	+	+	+	+	+	0	0	0
14564	lait	ovin	2005	France	64	↑	↑	+	+	↑	+	↑	+	+	+	↑	↑	0	0	0
14633	lait	caprin	2006	France	64	↑	↑	↑	↑	↑	+	↑	+	+	+	↑	↑	0	0	0
14634	lait	ovin	2006	France	64	+	+	+	+	+	+	+	+	-	-	+	+	0	0	1
14775	lait	ovin	2006	France	64	+	+	+	+	+	+	+	+	-	-	+	+	0	0	1
14776	lait	ovin	2007	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
14858	lait	ovin	2007	France	64	+	+	+	+	+	+	+	+	-	-	+	+	0	0	1
14880	lait	ovin	2007	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0
14891	lait	ovin	2007	France	64	+	+	+	+	+	+	+	+	-	-	+	+	0	0	1
14899	lait	ovin	2007	France	64	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0

Ce tableau résume les résultats obtenus par *Southern blot* avec les sondes spécifiques des *vpma* de la souche type de *M. agalactiae* PG2 sur un panel de 63 isolats des Pyrénées-Atlantiques. La référence pour les comparaisons est le profil de l'isolat 14899 (profil majoritairement observé au sein de ce panel d'isolats). (+) fragment(s) détecté(s) et de même taille que pour 14899 ; (-) pas de fragment détecté ; (↑) fragment détecté de taille supérieure en comparaison celui détecté pour 14899 ; (↓) fragment détecté de taille inférieure par rapport à 14899.

DEUXIEME CHAPITRE

DEUXIEME CHAPITRE : PLASTICITE, DISTRIBUTION ET EVOLUTION DES GENES DE LA FAMILLE *VPMA*, UN SYSTEME GENETIQUE CONSACRE A LA VARIATION DE SURFACE A HAUTE FREQUENCE CHEZ *MYCOPLASMA AGALACTIAE*.

Occurrence, plasticity and evolution of the *vpma* gene family, a genetic system devoted to high frequency surface variation in *Mycoplasma agalactiae*

Laurent-Xavier Nouvel, Marc Marenda, Pascal Sirand-Pugnet, Eveline Sagné, Michelle Glew, Sophie Mangenot, Valérie Barbe, Aurélien Barré, Stéphane Claverol et Christine Citti.

Article publié dans *Journal of Bacteriology*

Accepté le 10 avril 2009. Publié en juillet 2009 (Vol. 191, n°13, pp. 4111-4121) .

Résumé en français

Mycoplasma agalactiae présente une architecture de surface très labile notamment grâce à une hypervariabilité d'expression de lipoprotéines appelées Vpma. Dans la souche type PG2, la variation de phase des protéines Vpma est générée par un cluster de six gènes qui subissent des réarrangements d'ADN spécifiques. Pour étendre nos connaissances sur le degré de diversité pouvant être généré à la surface de *M. agalactiae*, nous avons analysé le répertoire des gènes *vpma* d'une souche atypique, la souche 5632. Cette souche possède un répertoire de vingt-trois gènes *vpma* situés à deux endroits du génome espacés de 250 kilobases. Les loci I et II contiennent respectivement seize et sept gènes *vpma*, et les sept gènes du locus II sont des duplications de gènes rencontrés au locus I. Certaines Vpma présentent une structure chimérique suggérant des recombinaisons homologues entre gène *vpma*. Les analyses de protéomique réalisées indiquent qu'au moins treize des seize Vpma peuvent être exprimées par la souche 5632. A cause de la présence d'un unique promoteur à chaque locus *vpma*, l'expression concomitante de deux Vpma devient possible si une souche contient des loci dupliqués. Par conséquent, le nombre de combinaisons possibles de Vpma exprimées à la surface est beaucoup plus grand pour la souche 5632 en comparaison à la souche type PG2. Enfin, nos données suggèrent que les séquences d'insertion seraient impliquées dans la duplication des loci *vpma* à deux endroits du chromosome chez 5632. L'intervention d'éléments génétiques mobiles dans des réarrangements chromosomiques affectant des gènes codant des composants majeurs de surface pourrait avoir ainsi d'importantes conséquences évolutives et épidémiologiques pour des agents pathogènes sans paroi et au génome réduit comme les mycoplasmes.

Occurrence, plasticity and evolution of the *vpma* gene family, a genetic system devoted to high frequency surface variation in *Mycoplasma agalactiae*

Laurent-Xavier Nouvel[1,2], Marc Marenda[1,2a], Pascal Sirand-Pugnet[3,4], Eveline Sagné[1,2], Michelle Glew[1,2b], Sophie Mangenot[5], Valérie Barbe[5], Aurélien Barré[6], Stéphane Claverol[7], and Christine Citti[2,1]*

[1]Université de Toulouse, ENVT, UMR 1225, F-31076 Toulouse, France
[2]INRA, UMR 1225, F-31076 Toulouse, France
[3]Université de Bordeaux, UB2, UMR1090, F-33076 Bordeaux, France
[4]INRA, UMR 1090, F-33883 Villenave d'Ornon, France
[5]CEA-IG, Genoscope, F-91057 Evry Cedex, France
[6]Centre de Bioinformatique de Bordeaux, UB2, F-33076 Bordeaux, France
[7]Pole Protéomique, Plateforme Génomique Fonctionnelle Bordeaux, UB2, 33076 Bordeaux, France

[a] Current address. School of Veterinary Science, 250 Princes Highway, Werribee, Victoria 3030, Australia.
[b] Current address. Bio21 Institute, Level 3 North, Rm 375, Melbourne Dental School, 30 Flemington Rd, Parkville 3010, Australia.

Corresponding author

Christine Citti
UMR 1225 INRA, ENVT
Ecole Nationale Vétérinaire de Toulouse
23 Chemin des Capelles BP 87614 31076 Toulouse Cedex 3
France
Tel: +33.(0)5.61.19.38.56
Fax: +33.(0)5.61.19.32.73
E-mail: c.citti@envt.fr

Running title: The *vpma* locus: occurrence, plasticity and evolution

Abstract

Mycoplasma agalactiae, an important pathogen of small ruminants, exhibits a very versatile surface architecture by switching ON-OFF multiple, related lipoproteins (Vpmas). In the type strain, PG2, Vpma-phase variation is generated by a cluster of 6 *vpma* genes that undergoes frequent DNA rearrangements via site-specific recombination. To further comprehend the degree of diversity that can be generated at the *M. agalactiae* surface, the *vpma* gene repertoire of a field strain, 5632, was analyzed and shown to contain an extended repertoire of 23 *vpma* genes distributed onto two loci located 250 kbp apart. Loci I and II include 16 and 7 *vpmas*, respectively, with all *vpmas* of locus II being duplicated at locus I. Several Vpmas displayed a chimeric structure suggestive of homologous recombination and a global proteomic analyses further indicate that at least 13 out of the 16 Vpmas can be expressed by the 5632 strain. Because a single promoter is present in each *vpma* locus, concomitant Vpma expression can occur in a strain with duplicated loci. Consequently, the number of possible surface combinations is much higher in 5632 compared to the type strain. Finally, our data suggested that insertion sequences are likely to be involved in 5632 *vpma* locus duplication at a remote chromosomal position. The role of such mobile genetic element in chromosomal shuffling of genes encoding major surface components may have important evolutionary and epidemiological consequences for pathogens such as mycoplasmas that have a reduced genome and no cell wall.

1. *Introduction*

Bacteria of the *Mycoplasma* genus belong to the Class *Mollicutes* and represent a remarkable group of organisms that derived from the Firmicutes lineage by massive genome reduction (41, 51). Consequent to this regressive evolution, modern mycoplasmas have been left with a small genome (580 to 1,400 kb), a limited number of metabolic pathways and no cell wall. Due to these particularities, members of the *Mycoplasma* genus have often been portrayed as "minimal self-replicating organisms". Despite this apparent simplicity, a large number of mycoplasma species are successful pathogens of man and a wide range of animals in which they are known to cause diseases that are often chronic and debilitating (1, 33). The surface of their single membrane is considered as a key interface in mediating adaptation and survival in the context of a complex, immunocompetent host (10, 13, 34, 40). Indeed, mycoplasmas possess a highly versatile surface architecture due to a number of sophisticated genetic systems that promote the intraclonal variation in expression and structure of abundant surface lipoproteins (9, 50). Usually, these systems combine a set of contingency genes with a molecular switch for turning expression ON or OFF that is either based (i) on spontaneous mutation (slipped strand mispairing), (ii) on gene conversion or (iii) on specific DNA rearrangements (9). While high-frequency phenotypic variation using the two first mechanisms has been thoroughly described in other bacteria (47), switching of surface components by shuffling silent genes at a particular single expression locus has mainly been studied in mycoplasmas (3, 8, 14, 16, 23, 39, 43).

Mycoplasma agalactiae, an important pathogen responsible for contagious agalactia in small ruminants (listed by the world organization for animal health), possesses a family of lipoproteins encoded by the *vpma* genes which phase variation in expression is driven by a “cut and paste” mechanism involving a tyrosine site-specific recombinase designated as Xer1 (16). Data previously gathered in the PG2 type strain had identified a single *vpma* cluster (42) composed of 6 *vpma* genes adjacent to one *xer1* gene (Fig. 1A). Based on fine

genetic analyses, Xer1 was further shown to mediate frequent site-specific DNA rearrangements by targeting short DNA sequences located upstream each *vpma* (8, 16). While some *vpma* rearrangements can be phenotypically silent, others result in Vpma ON/OFF switching by linking a silent *vpma* gene sequence immediately downstream to the unique *vpma* promoter. Because site-specific recombination can be reciprocal, the initial *vpma* configuration can be restored without loss of genetic information.

The *vsa* family of the murine pathogen *M. pulmonis* (3, 39), the *vsp* family of the bovine pathogen *M. bovis* (2, 24), and the *vpma* family of *M. agalactiae,* all generate intraclonal surface diversity using a very similar molecular switch (23) although their overall coding sequences seem to be specific of one *Mycoplasma* species. DNA rearrangements also govern phase variation of the 38 *mpl* genes of the human pathogen, *M. penetrans* (27, 35, 38). However, in this mycoplasma species the molecular switch is slightly different since each *mpl* gene possesses its own invertible promoter (19). In *M. penetrans*, the individual expression of each *mpl* gene can then be switched ON/OFF in a combinatory manner resulting in a large number of possible Mpl surface configurations. Since *M. pulmonis*, *M. bovis* and *M. agalactiae* all belong to the *Mycoplasma hominis* phylogenetic cluster (48) and are relatively closely related while *M. penetrans* belongs to the distinct *Mycoplasma pneumoniae* phylogenetic cluster (30, 48), it is tempting to speculate that the *vsa,* the *vsp* and the *vpma* systems have all been inherited from a common ancestor and that the bulk of their coding sequences has evolved independently in their respective host while the molecular switch mechanism has been retained.

In so-called “minimal” bacteria, the occurrence of relatively large genomic portions dedicated to multigene families, each encoding phase variable, related, surface proteins, suggests that they serve an important function(s). Data accumulated over the years in several mycoplasma species tend to indicate that one general purpose of these systems is to provide the mycoplasma with a variable shield that modulates surface accessibility in order to escape the host response and to adapt to rapidly changing environments (10, 11, 13, 40, 50). On the other hand, the sequences of phase variable proteins are relatively conserved within one species

but divergent between species, suggesting a more specific role for these molecules.

The role of the Vpma family of *M. agalactiae* has yet to be elucidated but it was recently shown that Vpma switches in expression occur at a remarkably high rate *in vitro*, (10^{-2} to 10^{-3} per cell per generation (8, 17). The vpma system described in PG2 (16) and in another *M. agalactiae* strain isolated in Israel (in which Vpma were designated as Avg (14)), both revealed a repertoire of 6 *vpma* genes and only one promoter, suggesting that in *M. agalactiae* the number of Vpma configurations is limited to 6. This contrast with the situation commonly found in other *Mycoplasma* variable systems that can offer larger mosaic of surface architecture because of the concomitant switches of several surface related proteins and/or because of a higher number of phase variable genes.

To further understand the degree of diversity that can be generated at the surface of *M. agalactiae*, we analyzed the *vpma* gene content of a field strain, 5632, whose genome has recently been sequenced by our group (unpublished data). The present study shows that 5632 contains a total of 23 *vpma* genes distributed in two distinct loci that both contain a recombinase gene. Further genomic and proteomic analyses indicate that the capacity of 5632 to vary its Vpma surface architecture is far more complex than that described for the type strain. Unlike PG2, both 5632 *vpma* loci are associated with several mobile genetic elements (IS) that could play an evolutionary role in the dynamics of *vpma* repertoires as suggested by data presented here. One 5632 *vpma* locus contains ORFs that are highly conserved in both *M. bovis*, a closely related bovine mycoplasma, and in the phylogenetically distant mycoplasmas of the mycoides cluster that are also important ruminant pathogens. Whether these have been acquired through evolution or through horizontal transfer will be discussed. The present study reveals an additional degree of complexity for the Vpma system and further suggests that some field strains might have a more dynamic genome and a more variable surface than was first estimated (42).

2. *Materials and methods*

Bacterial strains, culture conditions and DNA isolation

M. agalactiae type strain PG2, clone 55.5 (16) and strain 5632, clonal variant C1 (26), used in this study have been previously described. These strains have been independently isolated from goat in Spain. Experiments reported in this manuscript have all been performed with these clonal variants but for simplicity, we will further refer as the PG2 and the 5632 only. *M. agalactiae* field isolates were kindly provided by Dr. Poumarat (AFSSA Lyon, France) (see Table S1 in the supplemental material). Mycoplasmas were propagated in SP4 liquid medium (46) at 37℃ and genomic DNA was extracted as described elsewhere (7, 37).

Identification of *M. agalactiae vpma* and associated loci

Whole genome sequencing of strain 5632 (clonal variant C1) was performed as follows. A library of 3 kb inserts (A) was generated by mechanical shearing of the DNA followed by cloning of the blunt end fragments into the pcDNA2.1 (Invitrogen) *E. coli* vector. Two libraries of 25 kb (B) and 80 kb (C) inserts were generated by *Hin*dIII partial digestion and cloning into the pBeloBAC11 (CALTECH) modified *E. coli* vector. The plasmid inserts of 10752, 3072 and 768 clones picked from the libraries A, B and C, respectively, were end-sequenced by dye-terminator chemistry on ABI3730 sequencers. The PHRED/PHRAP/CONSED software package was used for sequence assemblies. Gap closure and quality assessment were made according to the Bermuda rules with 10307 additional sequences. The *vpma* loci of 5632 were detected by DNA homologies to that previously described of PG2 (positions of the loci on the 5632 genome are given in the text with reference to the first nucleotide of the *dna*A gene as nucleotide 1) and were annotated using the CAAT-Box platform (15) with the aid of the Artemis software (36) and the ACT software (5). The BLAST program suite was used for sequence homology searches in non-redundant databases (http://www.ncbi.nlm.nih.gov/blast/blast.cgi). In order to determine the extent of sequence similarity, alignments between sequences were performed using softwares Needle (Needleman-Wunsch global alignment algorithm) and Water (Smith-Waterman local alignment

algorithm) (http://www.ebi.ac.uk/Tools/emboss/align/). Designation of 5632 *vpma* genes was based on local and global alignment scores using all *vpma* gene sequences including those previously described in PG2. A same name was attributed to *vpma* presenting an amino-acid identity > 50% in global alignment (Needle) and an amino-acid identity >70% in local alignment (Water). Using these rules, strain 5632 was found to contain the *vpmaW* and *vpmaX* genes also present in PG2. As well, a distinct name was given to *vpmaY* and *vpmaF* although they display a global identity >50% because their local identity is <70%. *vpma* products with highly similar sequences that only differed in the number of their C-terminal repeats were considered as allelic versions and designated with a same letter followed by a different number (i. e. VpmaD1 and VpmaD2 and, VpmaF1 and VpmaF2). Finally, *vpmaI* was not considered as an allelic version of *vpmaD* because the C-terminal repeated region of its corresponding product displayed a local identity of 64.7% (<70%) with VpmaD1 and D2 couterparts.

Direct sequencing of PCR products and of genomic DNA (18, 20) was performed using specific primers (see Table S2 in the supplemental material) at the sequencing facility of UMR 5165 (CNRS, UPS, CHU Purpan, Toulouse, France).

Phylogenetic analyses were performed using MEGA 3.1 (21) and the Neighbor-Joining tree method. The reliability of the tree nodes was tested by performing 500 bootstrap replicates.

Accession number for locus I_{5632} and locus II_{5632} are, respectively, FP245515 and FP245514

PCR assays

PCR assays were performed on an Eppendorf Mastercycler ep-Gradient thermocycler using 5 ng *M. agalactiae* DNA as template with specific primers (see Table S2 in the supplemental material). PCR assays were performed in 25 µL reaction mixtures containing 0.4 mM of each primer, PCR reaction buffer (with $MgSO_4$, New England Biolabs [NEB]) at 1X final concentration, 200 mM dNTPs and 2.5 U *Taq* DNA polymerase (NEB). Reaction mixtures were subjected to 2 min at 94℃, 30 cycles of 30s at 94 ℃, 30s at 55℃, 30 s at 72 ℃, and a final elongation step of 5 min at 72℃. All PCR assays we re performed at the unique annealing temperature of 55℃ using primer pair xer F/phydR for specific

amplification of locus I; primer pairs xerF/agpR or Mag2F/ agpR for locus II; primer pairs aip1F/aip1R for the *abiGI* gene; primer pairs aip2F/aip2R for the *abiGII* gene. PCR products were analyzed by gel electrophoresis in 1% agarose.

Colony immunoblotting and proteomic analyses

Colony immunoblottings were performed as previously described (8). Briefly, nitrocellulose membranes were placed on mycoplasma colonies freshly grown on the agar medium, then removed and rinsed three times in TS buffer (10 mM Tris, 154 mM NaCl, pH 7.4). Membranes were then incubated overnight at 4℃ with rabbit Vpma-specific antibodies previously described (8), washed three times in TS buffer containing 0.05% Tween® 20 (Roth) and then incubated for 1 h at 25℃ in 1:2000 dilution of swine anti-rabbit IgG conjugated to horseradish peroxidase (Dako). After three washes, the colony blots were developed for 15–30 min in 4-chloro-1-naphtol and 0.02% hydrogen peroxide. The reaction was stopped by washing the blots in sterile distilled water.

Proteins that partitioned into Triton-Tx114 were extracted from 5632 as previously described (4), precipitated overnight at -70℃ afte r addition of 9 volumes of cold MeOH, centrifugated 10 min at 12,000x*g* and subjected to SDS-PAGE. The gel was sliced into 16 sections which were subjected to trypsin digestion. Peptides were further analyzed by nano liquid chromatography coupled to a MS/MS ion-trap mass spectrometer (LC-MS/MS).

Peptides were identified with SEQUEST through the Bioworks 3.3.1 interface (Thermo-Finnigan, Torrence, CA, USA) against a database consisting of both direct and reverse sense *Mycoplasma agalactiae* strain 5632 entries (1652 entries). Using the following criteria (DeltaCN ≥ 0.1, Xcorr vs Charge State ≥ 1.5 (+1), 2.0 (+2), 2.5 (+3), Peptide Probability ≤ 0.001 and Number of Different Peptides ≥ 2) as validation filters, the False Positive rate is null.

3. *Results and discussion*

Genome sequencing of strain 5632 reveals an extended *vpma* repertoire

The fully sequenced genome of *M. agalactiae* strain 5632 revealed a total of 23 *vpma* genes (Fig. 1). In contrast to the PG2 type strain, this extended repertoire is distributed onto two distinct chromosomal loci, locus I and II, that contain 16 and 7 *vpma* genes, respectively. Blast analyses only identified two having significant similarity with those previously described in PG2 (16, 42), namely *vpmaW* and *vpmaX*, the others (*vpmaA* to *vpmaG*) only shared blocks of high similarity with PG2 *vpmas* (Fig. 2).

More precisely, locus I_{5632} is 19453 bp long and is the counterpart of the PG2 *vpma* locus (Fig. 1A) because of their flanking CDS being nearly identical. In addition to 16 *vpma* genes, locus I_{5632} contains two CDS (MAGa8140 and MAGa8130) that have no homology with the Vpma family but were found to have a high amino-acid similarity with *Streptococcus agalactiae abiGI* and *abiGII* gene products (45). The two corresponding genes are lacking in PG2 and were designated here as *abiGI* and *abiGII*, respectively. Locus I_{5632} further differs from its PG2 counterpart by the presence of an Insertion Sequence (IS) corresponding to the previously described IS*Mag1* (31). This IS element is immediately adjacent to the 3' end of the integrase-recombinase *xer1* gene, whose product is identical to that of PG2.

The second *vpma* locus of 5632, locus II_{5632}, is 8462 kb long and contains only 7 *vpma* genes clustered between two IS*Mag1* copies, themselves flanked by a 16S rRNA gene at the 5' end and by a CDS annotated as a hypothetical protein (HP, MAGa5900) at the 3' end. In PG2, this region is highly conserved except that it does not contain any IS element or *vpma* or any other gene (locus II_{PG2}, Fig. 1C and below). DNA blast analyses of locus II_{5632} demonstrated that all seven *vpma* coding sequences are identical to seven *vpma* genes of locus I_{5632} (see asterisks in Fig. 1B). Interestingly, an identical *xer1* gene also occurs at both *vpma* loci of 5632. Since a single functional gene would be sufficient to generate DNA rearrangements in each locus, the duplication was further confirmed by PCR to rule out any possible artefact of sequence assembling. This was performed using

a *xer1*-specific primer in combination with a primer specific of the 3' region of locus I, phydR, or of locus II, agpR (Fig. 1D).

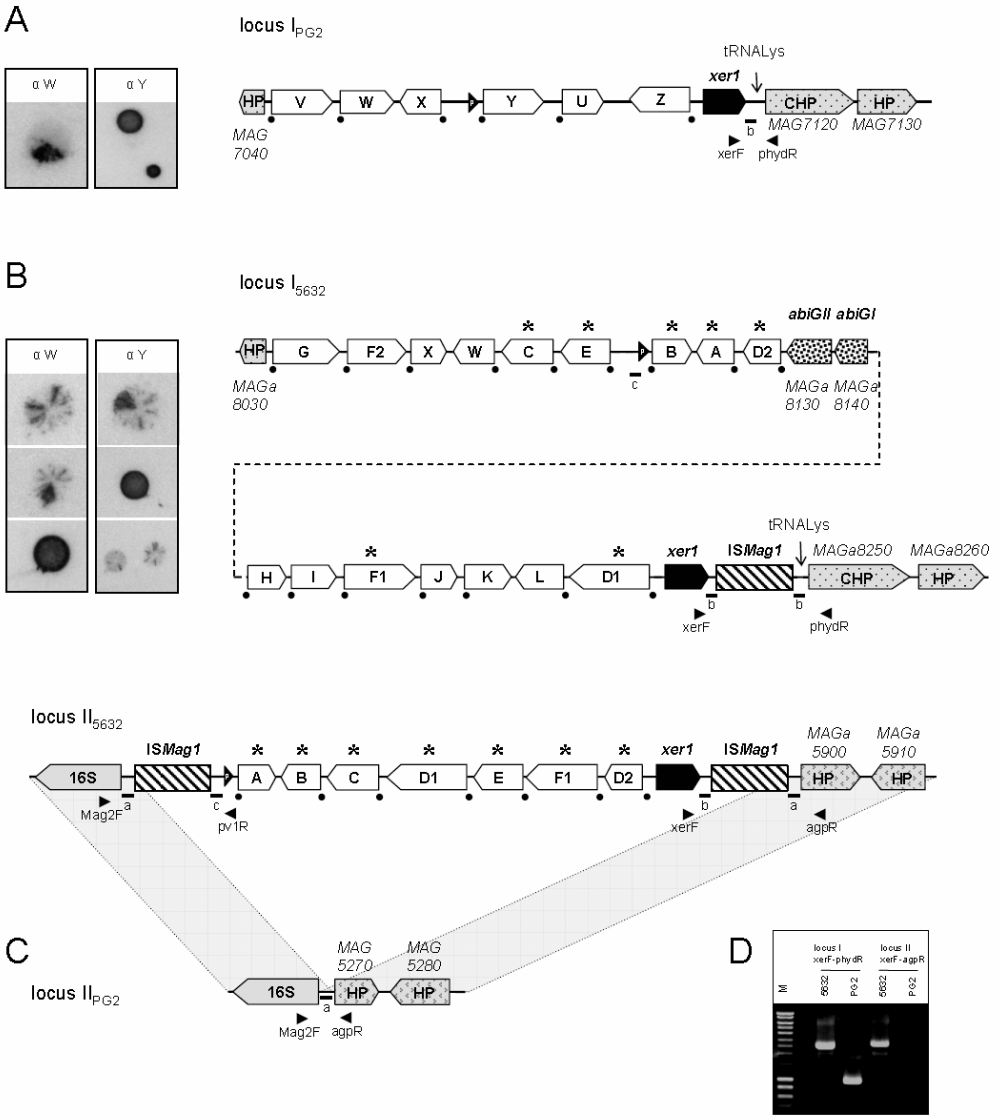


Figure 1: Comparison of the *M. agalactiae vpma* loci between the type strain PG2 and the strain 5632.

Schematics represent the organization of the *vpma* loci in a clonal variant 55.5 derived from PG2 (16, 42) (Panel A), and in a clonal variant c1 derived from strain 5632 (Panel B). Panel C presents the counterpart of locusII$_{5632}$ in PG2 and shows the absence of *vpma* genes in this region. The presence of two distinct loci in 5632 was confirmed by PCR using primers xerF-phydR or xerF-agpR and the resulting amplicons are shown in Panel D. The location of the primers is indicated by arrowheads in panels A and B. White large arrows labelled by a letter represent Vpma CDSs. The position of the promoters is represented by a black arrowhead labelled with a "P". The two non-Vpma related CDSs (*abiGI* and *abiGII*) are indicated by large arrows filled with a dotted pattern. Insertion sequence elements IS*Mag1* are indicated by hatched boxes. Recombination sites downstream of each *vpma* gene are indicated by a black dot. An asterisk (*) indicates that the corresponding *vpma* gene is present at two distinct loci. Schematics were approximately drawn to scale. HP: hypothetical protein; CHP: conserved hypothetical protein. Small letters and bars indicate the position of short particular sequences mentioned in the text and in Fig. 3 and 4.
The pictures on the left side of Panel A and B illustrate the variable surface expression of Vpma as previously described (8, 17). These correspond to colony immunoblot using Vpma-specific pAb recognizing PG2 VpmaW (α W) and VpmaY (α Y) epitopes.

The Vpma repertoire is composed of common and distinct structures

Detailed analyses of the 23 *vpma* coding regions of 5632 showed that (i) two are allelic versions of PG2 *vpma* genes (*vpmaW* and *vpmaX*), (ii) 7 are duplicated (*vpmaA*, *B*, *C*, *D1*, *D2*, *E* and *F1*) and distributed in two *vpma* loci that are 250 kbp apart, and (iii) two are allelic copies of *vpmaD (D1* and *D2) and vpmaF (F1* and *F2)* that mainly vary in the number of their C-terminal repeated motifs. Consequently, strain 5632 contains 12 new, distinct Vpma products when compared to PG2 (Fig. 2). All *vpma* gene products encode a conserved amino-acid signal sequence at their N-termini followed by a lipobox and by a 11 amino-acids conserved sequence (41). Within 5632, some mature *vpma* gene products are unique (VpmaA, G, W) while others share blocks of amino-acids between them (VpmaB, C, D1, D2, E, F1, F2, H, I, J, K, L, X). Also, most Vpmas contains amino-acid blocks that are directly repeated (VpmaA, B, D1, D2, I, F1, F2, G, W). These data indicate that the diversity of the Vpma family is represented by a mosaic of structures that is suggestive of homologous recombination events between *vpma* genes, gene duplication and/or insertion-deletion of repeated motifs. For instance, VpmaF1 and VpmaF2 share a common block of 104 amino-acids with VpmaK and VpmaX and, a common block of 85 amino-acids with VpmaY not found in VpmaK or VpmaX. Some Vpma genes such as VpmaD1 and D2 or VpmaF1 and F2 represent size variants of a same Vpma, with a variable number of repeated motifs at their C-termini (i.e. VpmaD1 and D2 contain respectively 20 and 2 repeated motifs of 15 amino-acids). Additionally, VpmaI may have a common ancestor with VpmaD1 and D2 and can be seen as the result of combined events such as gene duplication followed by sequence drift and/or repeat expansion at the 3' end.

Anti-VpmaW and anti-VpmaY rabbit antibodies previously produced against Vpma products of strain PG2 (8) also reacted with 5632 surface exposed epitopes in colony immunoblotting (Fig. 1) and revealed the presence of sectored colonies characteristic of high-frequency variation in expression. Most likely, 5632 Vpmas recognised by these antibodies correspond to VpmaW and to VpmaF1 and/or F2 that share blocks of amino acids with PG2 VpmaY (Fig. 2).

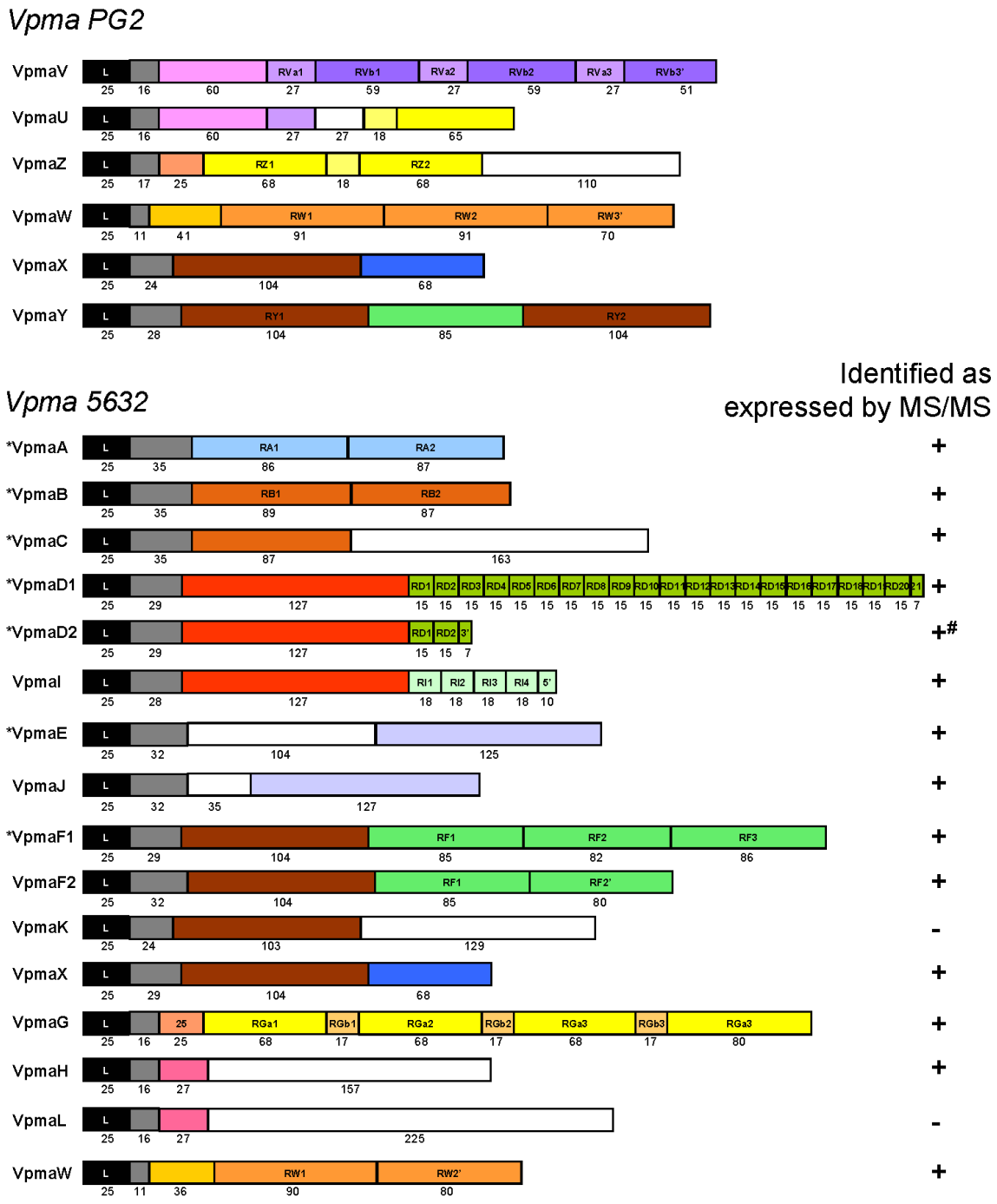


Figure 2: Structural features and comparison of vpma gene products in *M. agalactiae* strain PG2 and 5632.

Predicted Vpma proteins are schematically represented by boxes and begin with a homologous 25-aa leader sequence (black boxes) followed by regions that have homology between *vpma* gene products or that are repeated within a same product (coloured boxes). Two boxes of a same colour display an amino acid identity > 30%. White boxes represent unique sequences. Numbers below boxes indicate the number of amino-acids. For 5632, detection by MS/MS of expressed Vpma specific peptides is noted by +; – indicates that no specific peptides were detected for the corresponding Vpma. For VpmaD2, $+^{\#}$ indicates that VpmaD2 peptides detected are not specific because all are shared with D1 or I (see the list of detected peptides in Table S3 in the supplemental material). An asterisk (*) indicates that the corresponding *vpma* gene is present in 5632 at both *vpma* loci.

To better define the Vpma products expressed by 5632, proteins that partitioned in Triton X-114 detergent were subjected to SDS PAGE followed by in-gel trypsin digestion and by LC-MS/MS. Results showed that at least two specific peptides were detected for 13 Vpma (VpmaA, B, C, D1, E, F1, F2, G, H, I, J, X, and W) indicating that these were expressed by 5632 during propagation (see Table S3 in the supplemental material). While D1-specific peptides were detected, the expression of VpmaD2 could not be assessed because all VpmaD2 trypsin-peptides being also encoded in VpmaD1 or in VpmaI. No peptide specific of VpmaL or VpmaK was detected although each product possesses a long specific sequence of 225 and 129 amino acids, respectively (Fig. 2). Whether VpmaL or VpmaK are expressed at low level or under different conditions can not be ruled out. Finally, all Vpmas encoded by locus II_{5632} are duplicated at locus I_{5632} and consequently, their expression can not be monitored by this method or by any other such as Northern blot analyses because their sequences are also identical at the DNA level.

The global proteomic approach taken above is qualitative and does not reflect the level of expression of each Vpma, two of which are predicted to be expressed at high level by the clonal population while others should only reflect a minority of back-switchers. In PG2, the *vpma* locus contains a unique promoter which drives the expression of the *vpma* localized immediately downstream while other *vpma* genes remain silent (16). This promoter region was more precisely defined by Flitman-Tene *et al.* (14) using primer extension and a similar sequence was detected in each locus of the strain 5632. Therefore, Vpma expression in 5632 is most likely driven at each locus by a single promoter sequence which is highly conserved between the two loci (99% nucleotide identity). In locus I_{5632}, the promoter sequence is located upstream of the *vpmaB* gene while in locus II_{5632} it is upstream of *vpmaA* suggesting that these two Vpmas would be predominantly co-expressed by the clonal population. However, in the absence of specific-antibodies this hypothesis can not be tested. One interesting observation is that the length of a polyT tract, located between the putative -35 and -10 boxes of the *vpma* promoters, varies among the three *vpma* promoters sequenced and is of 14-, 15- and 12-nt respectively for loci I of PG2 and 5632 and for locus II of 5632 (see Fig. S1 in the supplemental material). In other variable mycoplasma systems, i.e.

the Vlp system of *M. hyorhinis*, fluctuation in the length of the polynucleotide tract between the -35 and the -10 boxes has been shown to affect transcription (12). Whether this is also the case for the *M. agalactiae* species and for strain 5632 in particular, cannot be currently tested because of the complexity of the 5632 *vpma* system. In PG2, DNA rearrangements are mediated by Xer1 which one recognises a specific Recombination Site (RS) sequence located upstream of each *vpma*. Indeed, analysis of the 5632 and PG2 *vpma* loci revealed that a rigorously identical RS sequence is located upstream of each of the 29 *vpma* genes (23 in 5632 and 6 in PG2). This suggests that each *vpma* loci of 5632 can independently undergo DNA rearrangements and that the concomitant expression of two distinct Vpmas is most likely to occur frequently in this strain. If so, 5632 could theoretically display 91 different Vpma surface configurations against only 6 in the type strain PG2. In parallel, additional Vpma size variations by insertion or deletion of direct repeats can further participate in altering surface architecture.

Evidence for horizontal gene transfer at the *vpma* locus of 5632

As mentioned earlier, *M. bovis* is a close relative of *M. agalactiae* and generates surface variation using a very similar genetic system designated as the *vsp* gene family. Both, the *vsp* and the *vpma* systems, have several common features: (i) *vsp* or *vpma* related genes are clustered in close proximity to a conserved recombinase gene (32), (ii) the recombination site involved in DNA rearrangements is identical except for a single nucleotide polymorphism that is conserved in each species, and (iii) *vpma* or *vsp* genes code for proteins having a signal peptide that is highly conserved between the two systems. Unlike PG2, *vpma* loci of 5632 are associated with IS elements as in *M. bovis* (23) and one, locus I_{5632}, contains two non-*vpma* genes, *abiGI* and *abiGII*, that have 86.97 and 85.15 % overall DNA identity with two *M. bovis* ORFs, ORF4 and ORF5, (23), located in *vsp* locus. Searching the databases revealed that *abiGI* and *abiGII* from 5632 also match with gene products annotated as conserved hypothetical products in *Mycoplasma mycoides* subsp. *mycoides* biotype LC strain GM12 (22) and *M. mycoides* subsp. *mycoides* SC strain PG1 (49), with more than 50% of amino-acid similarity in global alignment (Table 1). While *abiG* homologs are present in several Gram positive bacteria such as *Streptococcus agalactiae* and *S.*

suis (6, 45), none was found in other Mollicutes sequences available in the databases, which included 21 complete genomes.

The *abiGI* and *abiGII* genes were first described in *Lactococcus lactis* subsp. *cremoris* UC653 (28) where they occur as an operon carried by a conjugative plasmid and confer resistance to phage infection in this species (29). Detection by LC/MS-MS of AbiGII-like specific peptides in the Triton-X114 detergent phase of 5632 indicates that the corresponding *abiGII* gene is expressed although there is no hydrophobic domain that could account for this partitioning. Previous attempts to define the function of *abiGI* and *abiGII* of *Streptococcus* species failed (44) and whether *abiG* homologues found in mycoplasma have a role in protecting them against phage attack remains to be addressed. Nevertheless, comparison of AbiGI and AbiGII amino-acid sequences by constructing Neighbour-Joining trees (500 bootstraps) indicates a strong relationship between *Mycoplasma* and *Streptococcus* sequences suggesting that their corresponding genes might have undergone horizontal gene transfer (see Fig. S2 in the supplemental material).

Table 1: Analyses of *abiG* related genes products in mollicutes and other bacteria.

	***Mycoplasma agalactiae* strain 5632**	
	AbiG I-like (MAGa8140)	**AbiG II-like (MAGa8130)**
***M. agalactiae* PG2**	-	-
***M. bovis* PG45**	+ (84.8% ; 93.6%)	+ (81.1% ; 91.3%)
***M. mycoides mycoides* SC PG1**	+ (40.8% ; 59.2%)	+ (41.8% ; 64.6%)
***M. mycoides mycoides* LC GM12**	+ (41.7% ; 59.2%) (Mmycm_04185)	+ (31.9% ; 51.6%) (Mmycm_04180)
	+ (40.1% ; 58.5%) (Mmycm_00130)	+ (32.6% ; 50.9%) (Mmycm_00135)
Other Mollicutes	-	-
***Streptoccus agalactiae* H36B**	+ (40.2% ; 58.3%)	+ (41.2% ; 65.3%)
***Streptococcus suis* 05ZYH33**	+ (41.2% ; 58.3%)	+ (40.4% ; 66.3%)
Other than *Streptococcus sp.*	+ (<40% ; na)	+ (<40%)

(+): presence of homologous sequence; (-): absence of homologous sequence in public databases (june 2008); identity and similarity percentages between amino-acid sequences are respectively indicated in brackets and were obtained using the Needle alignment program (http://mobyle.pasteur.fr/Mobyle/programs/needle.xml).

To further define the occurrence of the two *abi* genes in *M. agalactiae*, a collection of 92 strains representative of (i) various geographical areas and (ii) various histories (see Table S1 in the supplemental material) was screened by PCR for the presence of *abiG* genes using specific primer pairs, aip1F/aip1R and aip2F/aip2R. PCR products corresponding to each gene were obtained for 11 strains out of 92. Further PCR assays showed that, like in *Streptococcus agalactiae*, the two *abiG* genes always occur next to each other. Interestingly, direct genome sequencing of six *abiG* positive strains using aip2F2 as primer shows that the *abiGII* gene is always associated with *vpma* genes. Because strains displaying *abiG* genes in their *vpma* loci have very different geographical origin (Ivory cost, Spain and France), this results raised the question of whether these genes have been vertically inherited from a *M. agalactiae*/*M. bovis* common ancestor or whether the *vpma* locus is a hot spot for their insertion. Since the *vpma* locus is frequently subjected to DNA rearrangements, deletion of *abiG* genes by intra-recombination, involving for instance *vpma*-RS sequences, may have easily occurred resulting in *abiG* negative strains.

Plasticity of the *M. agalactiae vpma* loci

To define whether the occurrence of the two *vpma* loci observed in 5632 is common among *M. agalactiae* strains, a PCR assay was performed with primer pairs xerF/phydR, xerF/agpR and Mag2F/agpR that yield amplicons corresponding to either locus I or II of PG2 and/or 5632. More specifically, the two pairs xerF/phydR and xerF/agpR target the 3' ends of the loci which are not affected by *vpma* rearrangements while Mag2F/agpR were designed to amplify a more or less empty locus II under our PCR conditions. The assay was conducted with the panel of 92 strains (see above). Results showed that 89 strains produced an amplification profile identical to that obtained with PG2 while three showed new profiles none of which was identical to those obtained with 5632 (Fig. 3). Sequencing of PCR products or selected genomic regions showed that (i) strain 13628, like 5632, contains an IS element at locus I while displaying a locus II similar to PG2, and (ii) strains 13375 and 4025, like PG2, displayed no IS at locus I while locus II contain one or more IS but no *vpma*, with strain 4025 having an IS*30-like* inserted into an IS*Mag1*. These data suggest that the occurrence of a

vpma cluster at locus II is a rare event and has so far only been observed in 5632, while the occurrence of *vpmas* at locus I is shared by all *M. agalactiae* strains tested so far.

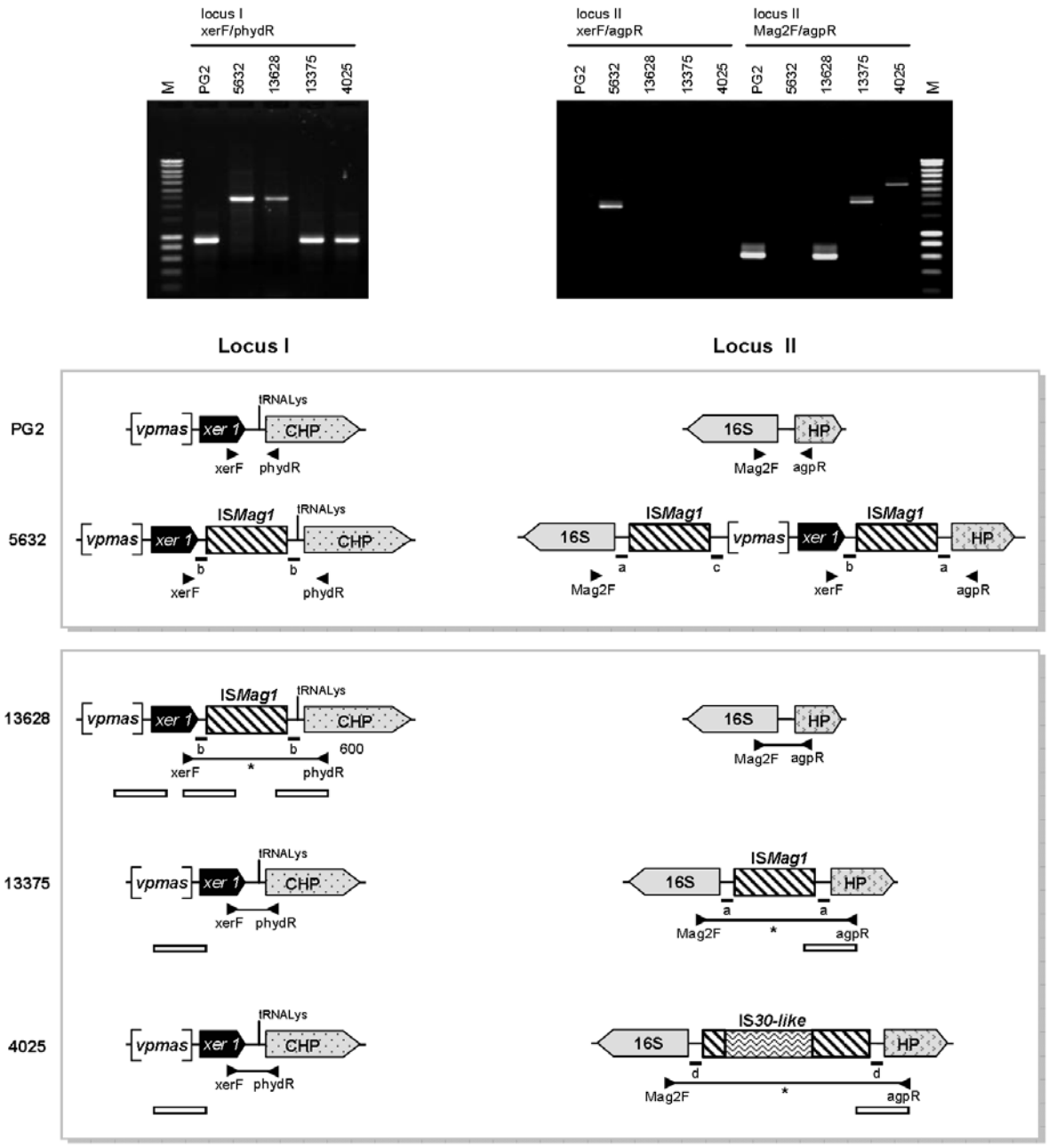


Figure 3: Comparison of *vpma* loci or their counterparts in *M. agalactiae* strains PG2 55.5, 5632 c1, 13628, 13375 and 4025.

Pictures represent the amplification of locus I and/or II using primers xerF/phydR, xerF/agpR and Mag2F/agpR which positions are indicated in the schematics below by arrow heads. Schematics represent the *vpma* locus I and the *vpma* locus II or counterpart. Brackets indicate the *vpma* genes cluster localization next to the *xer1* gene (filled black arrow). Insertion elements (IS*Mag1* or IS*30-like*) are indicated by shaded boxes. Lines traced between primers symbolise PCR amplicons obtained with an asterisk (*) indicating those which have been directly sequenced. White bars below the loci indicate regions that were directly sequenced using genomic DNA. Small letters (a, b, c and d) represent the 14-nt sequences that flanked the IS.

The presence of an IS*Mag1* in close proximity to the *vpma* locus I in 5632 is not an isolated event since it is shared by at least one other strain, 13628, which, unlike 5632, has been isolated in France in 2003 from a goat ear canal. IS*Mag1* belongs to the IS*30* family (31) that generates, when inserting itself into a host genome, a direct repeat that flanks the IS (25). Indeed, IS*Mag1* of locus I is flanked in 5632 and in 13628 by an identical 14-nt directly repeated sequence ("b" in Fig 3) suggesting that its insertion most likely occurred in an ancestor common to the two strains. Interestingly, the 14-nt element found on each side of the IS of locus II_{13375} ("a" in Fig. 3) is also repeated in locus II_{5632} as if flanking the whole locus including the IS. Because no other 14-nt directed repeat is found in the vicinity of the two IS, this suggested the introduction of the *vpma* cluster at locus II_{5632} via a duplication-insertion mechanism involving a circular intermediate carrying a set of *vpma* genes together with an IS*Mag1* element (Fig. 4). In this scenario, the duplicated *vpma* genes would have first clustered at one end of locus I via a series of DNA recombination. In a second step, part of locus I including the *xer1* and IS*Mag1* would have been duplicated and circularized to finally recombine at locus II via a single cross-over between the IS elements. One argument in favour of the occurrence of the circular form is that a PCR product was obtained with 5632 using outwardly oriented primers pv1R and xerF which partial sequencing revealed the presence of an IS*Mag1* next to a sequence "c" usually found next to the promoter (Fig. 4). Whether this result really reflects a duplication-excision event of part of the *vpma* locus I and whether the circular form can re-insert itself elsewhere in the genome have still to be formally demonstrated. Nevertheless, the 100% nucleotide identity observed between locus I and II of 5632 indicates that this duplication event as occurred relatively recently.

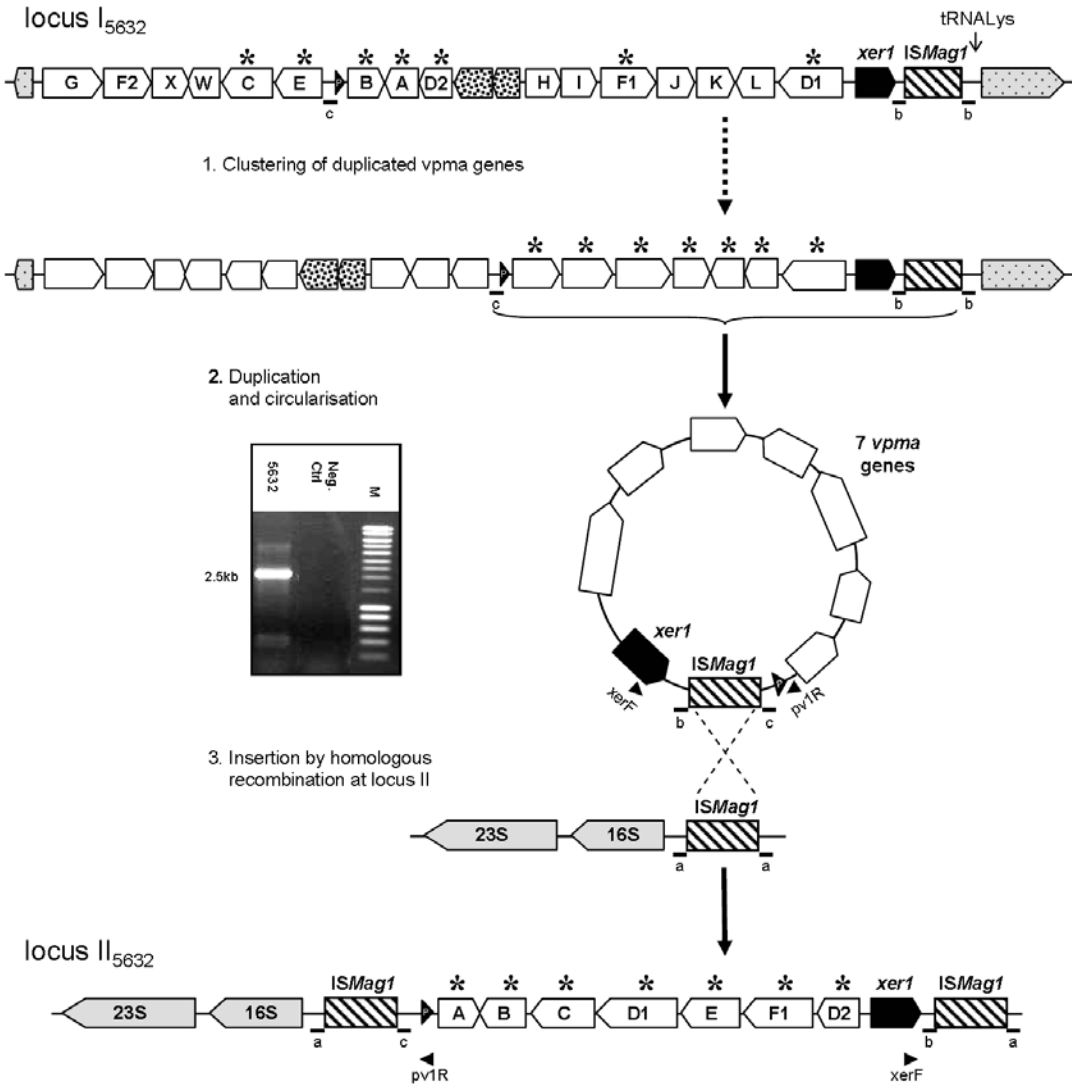


Figure 4: Scenario of duplication-insertion of 5632 *vpma* genes, from locus I to locus II via a circular intermediate.

Schematics illustrate the putative scenario in which the *vpma* cluster of 5632 at locus II originates from locus I by a mechanism of duplication-insertion involving a circular intermediate and IS elements.
Small bars with a letter below represent 14-nucleotides sequences (a, b, c) flanking the IS*Mag1* either after insertion event (a, b) or after homologous recombination event (c). The dotted arrow indicates the clustering process of seven *vpma* gene at one end of the locus I starting from the current configuration. It represents a series of an unknown number of possible DNA rearrangements. Asterisk (*) indicates duplicated *vpmas*.
The picture shows the PCR amplification of the circular intermediate obtained when using primers xerF and pv1R and 5632 crude DNA extract as template.

4. *Conclusion*

A number of sophisticated genetic systems devoted to high frequency surface variation have been described in mycoplasma species. Many of these systems appear as a small set of species-specific genes, which are nearly identical in all the isolates and are clustered at one locus. The clustering of these genes has been proposed to facilitate recombinations which generate diversity, and theoretically a wide surface diversity can be produced with only few genes, particularly as they are prone to internal rearrangements of their sequences. Our characterization of a new set of *vpma* genes in a *M. agalactiae* field isolate (5632) illustrates this point. This new repertoire is far more complex than the one first described in the PG2 type strain and represents a mosaic of structures that is suggestive of homologous recombination events between *vpma* genes, gene duplication and/or insertion-deletion of repeated motifs (see Fig. S3 in the supplemental material). This opens the question of the true diversity in equivalent variable systems from mycoplasmas with apparently few variable genes, most of which have been sequenced from type strains. Field or clinical isolates have the potential to reveal a greater than suspected diversity of variable gene repertoires, which could be useful to design molecular probes for epidemiology studies.

One of the most interesting finding of our study is the occurrence of duplicated *vpma* loci in strain 5632. Because one *vpma* locus can only express a single *vpma* gene per cell, the consequence of such event is tremendous in terms of variability because it allows concomitant expression and, in turns, multiplies the number of possible surface combinations. Our data suggests that this situation is most likely resulting from the duplication of one 5632 *vpmas* region at a remote chromosomal position (250 kbp away) and that it involves IS elements. The role of such mobile genetic elements in chromosomal shuffling of genes encoding major surface components may have important evolutionary and epidemiological consequences for pathogens such as mycoplasmas that have a reduced genome and no cell wall.

Finally, whether a successful vaccine strategy based on surface antigens could be developed for mycoplasmas is not known but this would have to take into account the dynamics and scope of any variable antigen gene repertoires.

References

1. **Baseman, J. B., and J. G. Tully.** 1997. Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. Emerg Infect Dis **3:**21-32.
2. **Behrens, A., M. Heller, H. Kirchhoff, D. Yogev, and R. Rosengarten.** 1994. A family of phase- and size-variant membrane surface lipoprotein antigens (Vsps) of *Mycoplasma bovis*. Infect Immun **62:**5075-84.
3. **Bhugra, B., L. L. Voelker, N. Zou, H. Yu, and K. Dybvig.** 1995. Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. Mol Microbiol **18:**703-14.
4. **Bordier, C.** 1981. Phase separation of integral membrane proteins in Triton X-114 solution. J Biol Chem **256:**1604-7.
5. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics **21:**3422-3.
6. **Chen, C., J. Tang, W. Dong, C. Wang, Y. Feng, J. Wang, F. Zheng, X. Pan, D. Liu, M. Li, Y. Song, X. Zhu, H. Sun, T. Feng, Z. Guo, A. Ju, J. Ge, Y. Dong, W. Sun, Y. Jiang, J. Yan, H. Yang, X. Wang, G. F. Gao, R. Yang, and J. Yu.** 2007. A glimpse of streptococcal toxic shock syndrome from comparative genomics of *S. suis* 2 Chinese isolates. PLoS ONE **2:**e315.
7. **Chen, W. P., and T. T. Kuo.** 1993. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. Nucleic Acids Res **21:**2260.
8. **Chopra-Dewasthaly, R., C. Citti, M. D. Glew, M. Zimmermann, R. Rosengarten, and W. Jechlinger.** 2008. Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation. Mol Microbiol **67:**1196-210.
9. **Citti, C., G. F. Browning, and R. Rosengarten.** 2005. Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas., p. 439–484. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: Molecular biology, pathogenicity and strategies for control. Horizon Bioscience, Norfolk (United Kingdom).
10. **Citti, C., M. F. Kim, and K. S. Wise.** 1997. Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth-inhibiting host antibodies. Infect Immun **65:**1773-85.
11. **Citti, C., and R. Rosengarten.** 1997. Mycoplasma genetic variation and its implication for pathogenesis. Wien Klin Wochenschr **109:**562-8.
12. **Citti, C., and K. S. Wise.** 1995. *Mycoplasma hyorhinis* vlp gene transcription: critical role in phase variation and expression of surface lipoproteins. Mol Microbiol **18:**649-60.
13. **Denison, A. M., B. Clapper, and K. Dybvig.** 2005. Avoidance of the host immune system through phase variation in *Mycoplasma pulmonis*. Infect Immun **73:**2033-9.
14. **Flitman-Tene, R., S. Mudahi-Orenstein, S. Levisohn, and D. Yogev.** 2003. Variable lipoprotein genes of *Mycoplasma agalactiae* are activated in vivo by promoter addition via site-specific DNA inversions. Infect Immun **71:**3821-30.
15. **Frangeul, L., P. Glaser, C. Rusniok, C. Buchrieser, E. Duchaud, P. Dehoux, and F. Kunst.** 2004. CAAT-Box, Contigs-Assembly and Annotation Tool-Box for genome sequencing projects. Bioinformatics **20:**790-7.

16. **Glew, M. D., M. Marenda, R. Rosengarten, and C. Citti.** 2002. Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the vpma multigene locus. J Bacteriol **184:**5987-98.
17. **Glew, M. D., L. Papazisi, F. Poumarat, D. Bergonier, R. Rosengarten, and C. Citti.** 2000. Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*. Infect Immun **68:**4539-48.
18. **Heiner, C. R., K. L. Hunkapiller, S. M. Chen, J. I. Glass, and E. Y. Chen.** 1998. Sequencing multimegabase-template DNA with BigDye terminator chemistry. Genome Res **8:**557-61.
19. **Horino, A., Y. Sasaki, T. Sasaki, and T. Kenri.** 2003. Multiple promoter inversions generate surface antigenic variation in *Mycoplasma penetrans*. J Bacteriol **185:**231-42.
20. **Hudson, P., T. S. Gorton, L. Papazisi, K. Cecchini, S. Frasca, Jr., and S. J. Geary.** 2006. Identification of a virulence-associated determinant, dihydrolipoamide dehydrogenase (lpd), in *Mycoplasma gallisepticum* through in vivo screening of transposon mutants. Infect Immun **74:**931-9.
21. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5:**150-63.
22. **Lartigue, C., J. I. Glass, N. Alperovich, R. Pieper, P. P. Parmar, C. A. Hutchison, 3rd, H. O. Smith, and J. C. Venter.** 2007. Genome transplantation in bacteria: changing one species to another. Science **317:**632-8.
23. **Lysnyansky, I., Y. Ron, and D. Yogev.** 2001. Juxtaposition of an active promoter to vsp genes via site-specific DNA inversions generates antigenic variation in *Mycoplasma bovis*. J Bacteriol **183:**5698-708.
24. **Lysnyansky, I., R. Rosengarten, and D. Yogev.** 1996. Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. J Bacteriol **178:**5395-401.
25. **Mahillon, J., C. Leonard, and M. Chandler.** 1999. IS elements as constituents of bacterial genomes. Res Microbiol **150:**675-87.
26. **Marenda, M. S., E. M. Vilei, F. Poumarat, J. Frey, and X. Berthelot.** 2004. Validation of the suppressive subtractive hybridization method in *Mycoplasma agalactiae* species by the comparison of a field strain with the type strain PG2. Vet Res **35:**199-212.
27. **Neyrolles, O., I. Chambaud, S. Ferris, M. C. Prevost, T. Sasaki, L. Montagnier, and A. Blanchard.** 1999. Phase variations of the *Mycoplasma penetrans* main surface lipoprotein increase antigenic diversity. Infect Immun **67:**1569-78.
28. **O'Connor, L., A. Coffey, C. Daly, and G. F. Fitzgerald.** 1996. AbiG, a genotypically novel abortive infection mechanism encoded by plasmid pCI750 of *Lactococcus lactis* subsp. *cremoris* UC653. Appl Environ Microbiol **62:**3075-82.
29. **O'Connor, L., M. Tangney, and G. F. Fitzgerald.** 1999. Expression, regulation, and mode of action of the AbiG abortive infection system of *Lactococcus lactis* subsp. *cremoris* UC653. Appl Environ Microbiol **65:**330-5.
30. **Pettersson, B., M. Uhlen, and K. E. Johansson.** 1996. Phylogeny of some mycoplasmas from ruminants based on 16S rRNA sequences and definition of a new cluster within the hominis group. Int J Syst Bacteriol **46:**1093-8.
31. **Pilo, P., B. Fleury, M. Marenda, J. Frey, and E. M. Vilei.** 2003. Prevalence and distribution of the insertion element ISMag1 in *Mycoplasma agalactiae*. Vet Microbiol **92:**37-48.

32. **Ron, Y., R. Flitman-Tene, K. Dybvig, and D. Yogev.** 2002. Identification and characterization of a site-specific tyrosine recombinase within the variable loci of *Mycoplasma bovis*, *Mycoplasma pulmonis* and *Mycoplasma agalactiae*. Gene **292:**205-11.
33. **Rosengarten, R., C. Citti, P. Much, J. Spergser, M. Droesse, and M. Hewicker-Trautwein.** 2001. The changing image of mycoplasmas: from innocent bystanders to emerging and reemerging pathogens in human and animal diseases. Contrib Microbiol **8:**166-85.
34. **Rosengarten, R., and K. S. Wise.** 1990. Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. Science **247:**315-8.
35. **Roske, K., A. Blanchard, I. Chambaud, C. Citti, J. H. Helbig, M. C. Prevost, R. Rosengarten, and E. Jacobs.** 2001. Phase variation among major surface antigens of *Mycoplasma penetrans.* Infect Immun **69:**7642-51.
36. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-5.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. ed. Cold Spring Harbor Laboratory Press, N.Y.
38. **Sasaki, Y., J. Ishikawa, A. Yamashita, K. Oshima, T. Kenri, K. Furuya, C. Yoshino, A. Horino, T. Shiba, T. Sasaki, and M. Hattori.** 2002. The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. Nucleic Acids Res **30:**5293-300.
39. **Shen, X., J. Gumulak, H. Yu, C. T. French, N. Zou, and K. Dybvig.** 2000. Gene rearrangements in the vsa locus of *Mycoplasma pulmonis*. J Bacteriol **182:**2900-8.
40. **Simmons, W. L., A. M. Denison, and K. Dybvig.** 2004. Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. Infect Immun **72:**6846-51.
41. **Sirand-Pugnet, P., C. Citti, A. Barre, and A. Blanchard.** 2007. Evolution of mollicutes: down a bumpy road with twists and turns. Res Microbiol **158:**754-66.
42. **Sirand-Pugnet, P., C. Lartigue, M. Marenda, D. Jacob, A. Barre, V. Barbe, C. Schenowitz, S. Mangenot, A. Couloux, B. Segurens, A. de Daruvar, A. Blanchard, and C. Citti.** 2007. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. PLoS Genet **3:**e75.
43. **Sitaraman, R., A. M. Denison, and K. Dybvig.** 2002. A unique, bifunctional site-specific DNA recombinase from *Mycoplasma pulmonis*. Mol Microbiol **46:**1033-40.
44. **Tangney, M., and G. F. Fitzgerald.** 2002. AbiA, a lactococcal abortive infection mechanism functioning in *Streptococcus thermophilus*. Appl Environ Microbiol **68:**6388-91.
45. **Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. O'Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli, and C. M. Fraser.** 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A **102:**13950-5.

46. **Tully, J. G.** 1995. Culture medium formulation for primary isolation and maintenance of mollicutes., p. 33–39. *In* J. G. Tully (ed.), Molecular and Diagnostic Procedures in Mycoplasmology: Molecular Characterization. Academic Press., San Diego.
47. **van der Woude, M. W.** 2006. Re-examining the role and random nature of phase variation. FEMS Microbiol Lett **254:**190-7.
48. **Weisburg, W. G., J. G. Tully, D. L. Rose, J. P. Petzel, H. Oyaizu, D. Yang, L. Mandelco, J. Sechrest, T. G. Lawrence, J. Van Etten, J. Maniloff, and C. R. Woese.** 1989. A phylogenetic analysis of the mycoplasmas: basis for their classification. J Bacteriol **171:**6455-67.
49. **Westberg, J., A. Persson, A. Holmberg, A. Goesmann, J. Lundeberg, K. E. Johansson, B. Pettersson, and M. Uhlen.** 2004. The genome sequence of *Mycoplasma mycoides* subsp. *mycoides* SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). Genome Res **14:**221-7.
50. **Wise, K. S.** 1993. Adaptive surface variation in mycoplasmas. Trends Microbiol **1:**59-63.
51. **Woese, C. R.** 1987. Bacterial evolution. Microbiol Rev **51:**221-71.

Supplementary data

Supplementary tables

Table S1: Features of strains used in this study.

Strain	Source	Host	Date of isolation	Country origin	PCR results			
					AbiGI - like	AbiGII -like	XerF /phydR	Mag2F /agpR
***M.agalactiae* type strain and isolates presenting a same amplification pattern**								
PG2 (type strain)	nk	caprine	1952	Spain	-	-	p1	p1
9	milk	ovine	nk	Italy	-	-	p1	p1
190	nk	nk	1951	Romania	-	-	p1	p1
332	joint	nk	1998	France	-	-	p1	p1
3990	nk	ovine	nk	France	-	-	p1	p1
4021	tank milk	ovine	1988	France	-	-	p1	p1
4055	milk	caprine	1987	France	-	-	p1	p1
4206	milk	ovine	1981	France	-	-	p1	p1
4210	milk	ovine	1982	France	-	-	p1	p1
4258	milk	caprine	before 1989	France	-	-	p1	p1
5210	milk	nk	1988	France	-	-	p1	p1
5225	milk	caprine	1991	Spain	-	-	p1	p1
5274	milk	ovine	1983	France	-	-	p1	p1
5275	milk	ovine	1983	France	-	-	p1	p1
5276	milk	ovine	1977	France	-	-	p1	p1
5633	milk	ovine	nk	France	-	-	p1	p1
5635	milk	ovine	1990	France	-	-	p1	p1
5636	milk	ovine	nk	France	-	-	p1	p1
5638	milk	ovine	nk	France	-	-	p1	p1
5639	milk	ovine	nk	France	-	-	p1	p1
5648	milk	ovine	nk	France	-	-	p1	p1
5670	milk	caprine	before 1991	Spain	-	-	p1	p1
5689	milk	caprine	nk	France	-	-	p1	p1
5690	milk	caprine	nk	France	-	-	p1	p1
5691	milk	caprine	nk	France	-	-	p1	p1
5725	milk	ovine	1990	France	-	-	p1	p1
5728	milk	ovine	nk	France	-	-	p1	p1
5729	milk	ovine	nk	France	-	-	p1	p1
5730	milk	ovine	nk	France	-	-	p1	p1
5770	milk	ovine	nk	France	-	-	p1	p1
5771	milk	ovine	nk	France	-	-	p1	p1
5826	milk	ovine	1992	Spain	-	-	p1	p1
6833	milk	caprine	nk	Italy	-	-	p1	p1
6968	milk	ovine	1993	Spain	-	-	p1	p1
7169	milk	caprine	nk	Switzerland	-	-	p1	p1
7314	milk	ovine	1986	Greece	-	-	p1	p1
7327	milk	caprine	1987	Greece	-	-	p1	p1
7375	milk	caprine	before 1993	Switzerland	-	-	p1	p1
7780	milk	ovine	nk	France	-	-	p1	p1
7781	milk	ovine	nk	France	-	-	p1	p1
7782	milk	ovine	nk	France	-	-	p1	p1
7783	milk	caprine	nk	France	-	-	p1	p1
7784	milk	ovine	before 1994	France	-	-	p1	p1
7786	milk	ovine	nk	France	-	-	p1	p1
8750	tank milk	ovine	nk	France	-	-	p1	p1
8751	milk	ovine	nk	France	-	-	p1	p1
8754	milk	ovine	nk	France	-	-	p1	p1
8782	milk	ovine	nk	France	-	-	p1	p1
9385	milk	caprine	before 1996	Portugal	-	-	p1	p1
9419	milk	ovine	1993	France	-	-	p1	p1

9420	milk	ovine	1993	France	-	-	p1	p1
9502	milk	ovine	1993	France	-	-	p1	p1
9503	milk	ovine	1993	France	-	-	p1	p1
9504	milk	ovine	1993	France	-	-	p1	p1
9506	milk	ovine	1994	France	-	-	p1	p1
9507	milk	ovine	1994	France	-	-	p1	p1
9508	milk	ovine	1994	France	-	-	p1	p1
9509	milk	ovine	1994	France	-	-	p1	p1
9511	milk	ovine	1994	France	-	-	p1	p1
9512	milk	ovine	1995	France	-	-	p1	p1
9523	milk	ovine	1995	France	-	-	p1	p1
9600	nk	caprine	before 1996	Portugal	-	-	p1	p1
9807	milk	ovine	nk	France	-	-	p1	p1
9808	milk	ovine	nk	France	-	-	p1	p1
9809	milk	ovine	nk	France	-	-	p1	p1
9810	milk	ovine	nk	France	-	-	p1	p1
9871	milk	ovine	nk	France	-	-	p1	p1
9872	milk	ovine	nk	France	-	-	p1	p1
9875	milk	ovine	before 1996	Spain	-	-	p1	p1
13373	ear	caprine	2003	France	-	-	p1	p1
13374	ear	caprine	2003	France	-	-	p1	p1
13376	tank milk	caprine	2003	France	-	-	p1	p1
13377	ear	caprine	2003	France	-	-	p1	p1
13381	ear	caprine	2003	France	-	-	p1	p1
13387	lung	ibex	2003	France	-	-	p1	p1
13779	tank milk	ovine	2004	France	-	-	p1	p1
13780	tank milk	ovine	2004	France	-	-	p1	p1
13781	tank milk	ovine	2004	France	-	-	p1	p1
13782	tank milk	ovine	2004	France	-	-	p1	p1
13787	tank milk	ovine	2004	France	-	-	p1	p1
13788	tank milk	ovine	2004	France	-	-	p1	p1
***M.bovis* type strain and *M.agalactiae* isolates presenting a same amplification pattern**								
PG45 (type strain)	milk	bovine	1962	US	+	+	p1	p1
209	milk	caprine	1983	France	+	+	p1	p1
4054	milk	caprine	1986	France	+	+*	p1	p1
4212	milk	ovine	1986	France	+	+*	p1	p1
8064	lung	ovine	1989	Ivory Coast	+	+*	p1	p1
14340	milk	caprine	2005	France	+	+*	p1	p1
14535	milk	caprine	2005	France	+	+	p1	p1
14628	lung	ibex	2006	France	+	+	p1	p1
14668	lung	caprine	2006	France	+	+*	p1	p1
8063	nk	nk	1988	Turkey	+	+	p1	p1
***M.agalactiae* 5632 strain and isolates presenting a particular amplification pattern**								
5632	joint	caprine	before 1991	Spain	+	+	p2	-
4025	tank milk	ovine	1988	France	+	+*	p1	p3
13375	ear	caprine	2003	France	+	+	p1	p2
13628	ear	caprine	2003	France	-	-	p2	p1

nk, not known; +, gene detected by PCR amplification; -, no amplification observed; * direct genomic DNA sequenced strains, attesting that *abiG-like* genes are always associated with *vpma* genes; p1, profile 1 (as observed for *M. agalactiae* type strain PG2), p2, profile 2, p3, profile 3 (other profile than observed with strain PG2 or 5632).

Table S2: Primers used in this study.

Primer	Sequence
xerF	5'-TGCGTAGATCATTTGCTTCG-3'
xerR	5'-CACTTGCCTGTAAAGACGTTG-3'
phydR	5'-TCAGGAATTGAGTGCTGTGC-3'
Mag2F	5'-CTCCGATCTTCTAGCGAAGC-3'
agpR	5'-TCAAAAGGATATTTTACACCATGTC-3'
Pvpma1R	5'-TCTGCCACTGATTACAACTTATGAA-3'
aip1F	5'-TTGCTTTTGTGCTTAAATACTTGT-3'
aip1R	5'-TATAGCCACTGTAAACCGTAAC-3'
aip2F	5'-ATGCAAATGTTTACTACTGTGATAG-3'
aip2R	5'-TCATCTTTCATTCATCAACGGATT-3'
aip2F2	5'-ATCCGTTGATGAATGAAAGATG-3'
ISMag1R	5'-AAAGGCAATAGAAAGCACCGC-3'
ISMag1R2	5'-GATGAATAAGGGTGGCAATG-3'
ISMag1R3	5'-CTTTTTGAATCGTTCCTTGTTTC-3'
ISMag1F0	5'-TAATAACGCGCTCAAAAACG-3'
ISMag1F	5'-CATCAACACAAAATTCCGCACT-3'
ISMag1F2	5'-TGATTGATGAAACACATTGAGATG-3'
ISMag1F3	5'-GACAACGATTATGAAATGGACAC-3'
IS30LF	5'-GTTCCACACTTGGCAGCAG-3'
IS30LF2	5'-TTGTCGCAACAGCTCGATAC-3'

Table S3 : Peptides detected by MS/MS that correspond to Vpma product encoded by *M. agalactiae* strain 5632 after trypsin digestion.

Peptides	Vpma carrying the peptide	
K.DKNSILDALLK.K	W	
K.DKNSILDALLKK.N	W	
K.ILVSFTLNSSATPGDGDATTAEK.K	W	
K.ILVSFTLNSSATPGDGGTTR.-	W	
K.KPAGTPDKTPGGNQNPGTEGGENSNNATTGEK.K	W	
K.KSLSVLISNENQQLGK.I	W	
K.KYLPAIISNENQQLGK.I	W	
K.NSILDALLK.K	W	
K.NSILDALLKK.N	W	
K.SDSVKYEGKILVSFTLNSSATPGDGGTTR.-	W	
K.SLSVLISNENQQLGK.I	W	
K.SQLDVESIENASAVVFAK.S	W	
K.TPGGNQNPGTEGGENSNNATTGEK.K	W	
K.TPGGNQNPGTEGGENSNNATTGEKK.S	W	
K.YLPAIISNENQQLGK.I	W	
K.DGIKKEEVLAVLK.T	G	
K.DGVKKEEVLAVLK.T	G	
K.DIKDEDLATVEFK.N	G	
K.IDQTIQDQLNK.L	G	
K.KEEVLAVLK.T	G	
K.KLVIEAK.E	G	
K.KPAEM*PGGGQADSTPSTSTAVDLSK.I	G	
K.KPAEMPGGGQADSTPSTSTAVDLSK.I	G	
K.LVSGKYEFAAQTDSNSSTSTTVNLGNLK.K	G	
K.LVSGKYEFAAQTDSNSSTSTTVNLGNLKK.D	G	
K.TVEGLKDIKDEDLATVEFK.N	G	
K.YEFAAQTDSNSSTSTTVNLGNLK.K	G	
K.YEFAAQTDSNSSTSTTVNLGNLKK.D	G	
R.MANNTDASGAEIGLR.T	G	
K.IKVGAEELELEFGK.V	F1	F2
K.VGAEELELEFGK.V	F1	F2
K.LGVQAQQK.I	F1	X
K.DKINSIWGLAK.E	F1	
K.DSDTVKAALLAAVK.G	F1	
K.IDELTIDLDPR.E	F1	
K.INSIWGLAK.E	F1	
K.STSYEGEVTVTFTVK.S	F1	
K.TELGEIAKK.D	F1	
K.TPAEGSQNPGTEGGSNQQGSEINIK.A	F1	
K.VSNANTAETNSETTK.Q	F1	
K.DLESVKTALLTVAK.D	F2	
K.DLKVNEVTIELDSK.E	F2	
K.DLKVNEVTIELDSKENK.A	F2	
K.EKIQSIWGLALK.F	F2	

K.IQSIWGLALK.F	F2	
K.TALGEIAK.K	F2	
K.TALGEIAKK.D	F2	
K.TALLTVAK.D	F2	
K.TALLTVAKDLKVNEVTIELDSK.E	F2	
K.TLGLQAQQK.I	F2	
K.VNEVTIELDSK.E	F2	
K.VNEVTIELDSKENK.A	F2	
K.VSNANTTETNSEATAK.K	F2	
K.DEFNESPSYDNIQK.A	E	
K.DLDYIEYLK.A	E	
K.DLLEEYNGITELIK.E	E	
K.EGKDLDYIEYLK.A	E	
K.ELKVETETLIK.D	E	
K.FQGTIILTK.Q	E	
K.KLKDEFNESPSYDNIQK.A	E	
K.NSDFLVQGSSEK.L	E	
K.QSGQPSGGSDVTPPTSNDNGGK.V	E	
K.VETETLIKDLETK.I	E	
K.KPGEGKEPGQNTDQGK.N	A	B
K.EPGQNTDQGKNPETNASPK.E	A	
K.GEFSENPSYDNIVK.E	A	
K.HILVQGSK.D	A	
K.ITATGVGSNPFR.G	A	
K.ITATGVGSNPFR.G	A	
K.TDISKLDEK.T	A	
R.GTLLLSNNNGESSAK.S	A	
	C	
K.ELNVENPSPVTDFEGM*FK.D	C	
K.EPDQNTDQGKNPETDTTPAK.S	C	
K.LETNIQNELEK.L	C	
K.LITGEYEFK.V	C	
K.LTLGDFSAVEFK.D	C	
K.LTLGDFSAVEFKDK.K	C	
K.NGVSPESVLDILK.K	C	
K.NGVSPESVLDILKK.V	C	
K.QVQEQNQEIEK.L	C	
R.STTPPADDPNATPAVPK.K	H	
K.EDGWTNEDIDAVK.E	H	
K.DLDYVSGLNEK.T	H	
K.NPSTENGGNGTPTSPK.T	J	
K.PVDTPAETPGGSQSNPGDNK.N	J	
K.IENENAEDFIAQNK.E	J	
K.DLDYVENLKDEK.T	J	
	X	
K.LGVQAQQK.I	X	F1
K.LASDDDKNNK.L	X	
K.DPNTPAEGSQNPGTEGGSNQQGSEINVK.A	X	
K.ILSIWGLAPK.F	X	

K.KDPNTPAEGSQNPGTEGGSNQQGSEINVK.A	X		
K.VGTEKLELEFGK.V	X		
K.INTNLGYFPK.D	D1	D2	I
K.LLNGRQDFSLAVDYDNHK.A	D1	D2	I
R.QDFSLAVDYDNHK.A	D1	D2	I
K.PGASTQPSGSAQPGKPGASTQPSGSAQ.-	D1	D2	
K.PGASTQPSGSAQPGKPGASTQPSGSAQPGK.P	D1	D2	
K.VTLANGQEITFTFTDER.E	D1	D2	
K.VTLANGQEITFTFTDERER.M	D1	D2	
R.MAQDTDASGADINSGTR.S	D1	D2	
R.SVQPSGATQPGK.P	D1	D2	
R.SVQPSGATQPGKPGASTQPSGSAQPGK.P	D1	D2	
K.TPAEGSMPADK.D	D1		
K.TPAEGSMPADKDK.T	D1		
K.APAEGGM*PADK.D	I	D2	
K.APAEGGM*PADKDK.T	I	D2	
K.APAEGGM*PADKDKTEK.D	I	D2	
K.APAEGGMPADK.D	I	D2	
K.APAEGGMPADKDK.T	I	D2	
K.LQNGQEITFTFTDEK.K	I		
K.PGTSGQPSGTTQPGI.-	I		
K.PGTSGQPSGTTQPGSGAK.P	I		
K.PGTSGQPSGTTQPGSGAKPGTSGQPSGTTQPGI.-	I		
K.PGTSGQPSGTTQPGSGAKPGTSGQPSGTTQPGSGAK.P	I		
R.VANDTDASGADINSGTR.P	I		
R.VANDTDASGADINSGTRPGQPSGTTQPGSGAK.P	I		
K.DVLSEDVIDVLK.K	B		
K.DVLSEDVIDVLKK.A	B		
K.EPGQNTDQGKNPETDTTPAK.S	B		
K.KDVLSEDVIDVLK.K	B		
K.KDVLSEDVIDVLKK.A	B		
K.KLTIEAHK.D	B		
K.LETNIQNELK.K	B		
K.LETNIQNELKK.L	B		
K.LITGKYEFTAQSK.T	B		
K.TNIQNELKK.L	B		
K.VEFNETDKK.L	B		
K.YEFTAQSK.T	B		
R.LGDLSKVEFNETDKK.L	B		

Supplementary figures



Figure S1: Multiple sequence alignment of *vpma* promoter regions in PG2 and 5632 loci I and II.

Promoter region sequences, from -35 box to ATG start codon, of the three PG2 and 5632 *vpma* loci were aligned using ClustalW (http://mobyle.pasteur.fr/Mobyle/programs/clustalw.xml) Blue shading indicates the variable polyT stretches existing between -35 and -10 boxes shaded in dark grey.

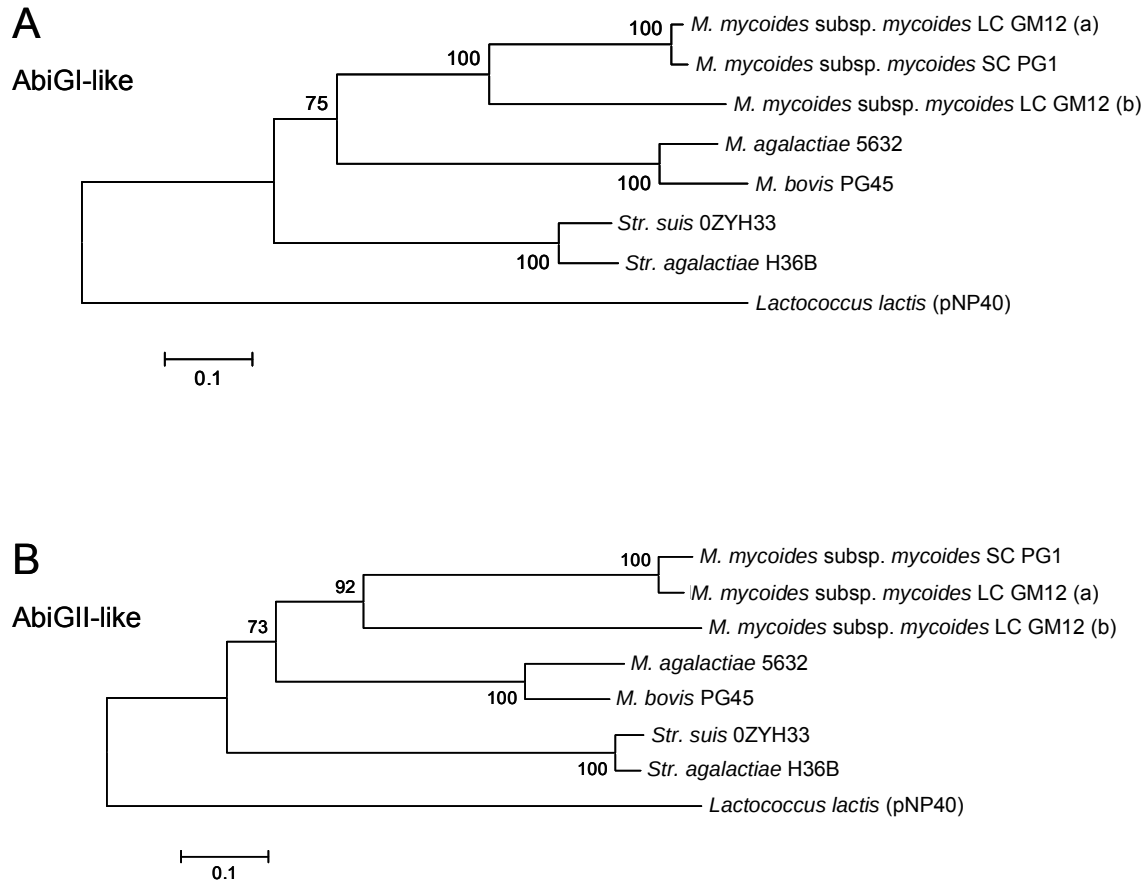


Figure S2: AbiGI-like and AbiGII-like phylogenetic trees.

Neighbour-joining tree based on AbiGI-like (A) and AbiGII-like (B) homologous gene products among Mycoplasma and Streptoccocus genus. AbiEI (A) and AbiEII (B) sequences encoded by pNP40 plasmid genes of Lactococcus lactis serve as the outgroup. Two AbiG-like pair of homologues present in Mycoplasma mycoides subsp. mycoides LC strain GM12, (a) and (b) annotations differentiate each pair. Bootstrap supporting values after 500 replicates are indicated. Sequence accession numbers: (A), NP_975179, ZP_02512914, ZP_02511961, AAK94955, YP_001198332, ZP_00783155, YP_001966453; (B), NP_975178, ZP_02512913, ZP_02511962, AAK94954, YP_001198333, ZP_00783153, YP_001966452.

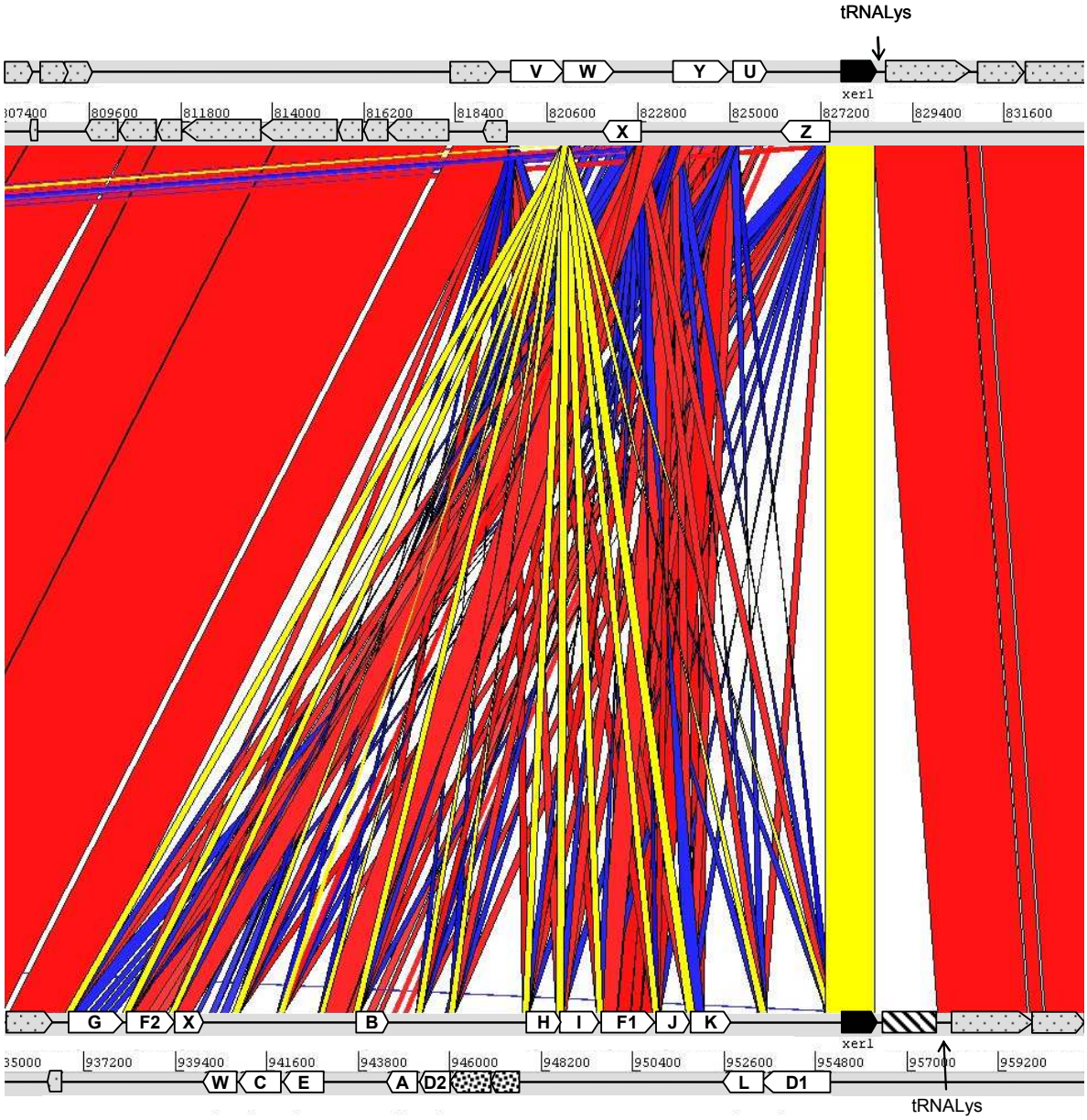


Figure S3: Imaging of whole sequence comparison between PG2 *vpma* locus and 5632 *vpma* locus I.

Screenshot image obtain with Artemis comparison tool. Homologies between the two genome portions are represented by coloured boxes correspond to single match, with red representing a good match between positive strands and blue between negative strands. An example of the redundancy of the recombination site (RS) is illustrated in yellow, and shows the identity of PG2 RS with those of locus I of 5632. As well, the match between the *xer1* genes is also shown in yellow.

TROISIEME CHAPITRE

TROISIEME CHAPITRE : ANALYSES GENOMIQUES ET PROTEOMIQUES COMPARATIVES DE DEUX SOUCHES DE *MYCOPLASMA AGALACTIAE* : APPROCHE DES MACRO ET MICRO-EVENEMENTS QUI MODELENT LA DIVERSITE DES MYCOPLASMES

Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro- and micro-events that are shaping mycoplasma diversity

Laurent-Xavier Nouvel, Pascal Sirand-Pugnet, Marc Marenda, Eveline Sagné, Valérie Barbe, Sophie Mangenot, Daniel Jacobs, Aurélien Barré, Stéphane Claverol, Alain Blanchard, and Christine Citti.

Article à soumettre à BMC Genomics

Résumé en français

Parmi les souches de *Mycoplasma agalactiae*, la souche 5632 apparaît comme éloignée de la souche type PG2 et à l'opposé du spectre génétique rencontré dans l'espèce *M. agalactiae*. Pour mieux comprendre la plasticité des génomes et les mécanismes responsables de la diversité à l'intérieur d'une espèce comme *M. agalactiae*, nous avons séquencé le génome et le protéome de cette souche 5632 pour les comparer à ceux de PG2. Les résultats montrent que cette souche contient 15% de séquence (environ 130 kDa) en plus par rapport à la souche PG2. La composition de ce "supplément" de matériel génétique correspond principalement (i) à des éléments génétiques mobiles et (ii) à l'étendue de familles de gènes qui codent des protéines de surface présentant des caractéristiques de systèmes d'expression hypervariable. Plus précisément, trois copies entières d'un élément d'intégration conjugatif (ICE) précédemment décrit sont retrouvés dans 5632 et correspondent à environ 80 kDa. D'autres éléments génétiques mobiles, absents de PG2 mais présents chez 5632, sont les plus classiques séquences d'insertion (IS). Ces dernières présentent des similitudes de séquence avec des IS retrouvées dans deux autres agents pathogènes de ruminants, *M. bovis* et *M. mycoides* subsp. *mycoides* SC. Chez 5632, les répertoires de familles de gènes codant des protéines de surface sont toujours plus étendus, souvent en raison de duplication de gènes. L'analyse protéomique comparative des deux souches montre que les capacités de codage supplémentaire de 5632 intéressent l'architecture globale de la surface, et suggère la présence de nouveaux systèmes de variation de phase basés sur le polymorphisme d'homopolymères nucléotidiques. Dans l'ensemble, les analyses comparatives des deux souches de

M. agalactiae révèlent une grande dynamique du génome dont la structure a été façonnée par des flux génétiques entre mycoplasmes de ruminants et par l'expansion de répertoires de gènes codant des protéines de surface dont l'expression est soumise à des micro-événements génétiques localisés.

Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro- and micro-events that are shaping mycoplasma diversity

Laurent-Xavier Nouvel[1,2£], Pascal Sirand-Pugnet[3,4£], Marc Marenda[1,2£a], Eveline Sagné[1,2], Valérie Barbe[5], Sophie Mangenot[5], Daniel Jacobs[6], Aurélien Barré[6], Stéphane Claverol[7], Alain Blanchard[3,4,] and Christine Citti[2,1*].

£: theses authors have equally contributed to this work

[1]Université de Toulouse, ENVT, UMR 1225, F-31076 Toulouse, France
[2]INRA, UMR 1225, F-31076 Toulouse, France
[3]Université de Bordeaux, UB2, UMR1090, F-33076 Bordeaux, France
[4]INRA, UMR 1090, F-33883 Villenave d'Ornon, France
[5]CEA-IG, Genoscope, F-91057 Evry Cedex, France
[6]Centre de Bioinformatique de Bordeaux, UB2, F-33076 Bordeaux, France
[7]Pole Protéomique, Plateforme Génomique Fonctionnelle Bordeaux, UB2, 33076 Bordeaux, France

[a] Current address. School of Veterinary Science, 250 Princes Highway, Werribee, Victoria 3030, Australia.

* Corresponding author

Christine Citti
UMR 1225 INRA, ENVT
Ecole Nationale Vétérinaire de Toulouse
23 Chemin des Capelles BP 87614 31076 Toulouse Cedex 3
France
Tel: +33.(0)5.61.19.38.56
Fax: +33.(0)5.61.19.32.73
E-mail: c.citti@envt.fr

Abstract

While the genomic era is accumulating a tremendous amount of data, the question of how genomics can describe a bacterial species remains to be fully addressed. The recent sequencing of the genome of the *Mycoplasma agalactiae* type strain has challenged our general view on mycoplasmas by suggesting that these simple bacteria are able to exchanged significant amount of genetic material via HGT. Yet, events that are shaping mycoplasma genomes and that are underlining diversity within this species have to be fully evaluated. For this purpose, we compared two strains that are representative of the genetic spectrum encountered in this species: the type strain PG2 which genome is already available and a field strain, 5632, which was fully sequenced and annotated in this study.

The two genomes differ by ca. 130 kDa with that of 5632 being the largest (1006 kbps). The make up of this "extra" genetic material mainly corresponds (i) to mobile genetic elements and (ii) to expanded repertoire of gene families that encode putative surface proteins and display features of highly-variable systems. More specifically, three entire copies of a previously described integrative conjugative element (ICE) are found in 5632 that accounts for ca. 80 kDa. Other mobile genetic elements, absent from PG2 but found in 5632, are the more classical IS which sequences are related to that found in two other ruminant pathogens, *M. bovis* and *M. mycoides* subsp. *mycoides* SC. In 5632, repertoires of gene families encoding surface proteins are always larger due to gene duplication. Comparative proteomic analyses of the two strains indicate that the additional coding capacity of 5632 affects the overall architecture of the surface and suggests the occurrence of new phase variable systems based on single nucleotide polymorphisms. Overall, comparative analyses of two *M. agalactiae* strains revealed a very dynamic genome which structure has been shaped by gene flow among ruminant mycoplasmas and expansion-reduction of gene repertoires encoding surface proteins, the expression of which is driven by localized genetic micro-events.

1. *Introduction*

Over the last decade, it has become clear that a single bacterial strain is not always representative of the whole species. Moreover, the range of physiological and virulence properties of a given bacterial pathogen most often relies on a particular subset of genes that are responsible for strain-specific lifestyles and may not be equally distributed within the species [1]. Comparative genomics provide a powerful approach to understand what makes a pathogen but the question of how it can describe a bacterial species is still debated [2]. For a single bacterial species, mathematical models are predicting the discovery of new genes even after sequencing hundreds of different genomes [3].

The genus *Mycoplasma* of the class *Mollicutes* includes the smallest self-replicative bacteria, *M. genitalium*, which genome was among the first sequenced [4]. While our genomic era is accumulating a tremendous amount of data with more than 900 microbial genomes currently available in public databases (Microbial Genome Resource, NCIB), only 15 other mycoplasma genomes have been completed since then [5 , 6, 7] with 3 representing a same species, *M. hyopneumoniae* [8, 9]. This is a surprising low number owing the small size of mycoplasma genomes and the significant number of species that are relevant for public and animal health because they are known as pathogenic for man or for a wide range of animals [10]. Recently, sequencing of the *M. agalactiae* genome has changed our view on these simple bacteria by showing that a significant portion of its genome (ca. 18%) has undergone horizontal gene transfer (HGT) with members of the phylogenetically distant "mycoides" cluster [11]. This cluster includes a number of mycoplasma species which are, like *M. agalactiae*, important ruminant pathogens and the nature of the exchanged genes suggests that some may play a role in mycoplasma-host interactions. Thus, a new picture has started to emerge in which mycoplasma evolution is not only driven by gene loss as previously thought but also by gene flow in between organisms sharing a same host [5], offering new means for adaptation.

Based on previous studies, the *M. agalactiae* species is fairly homogeneous with little intra-species genetic variation and with most of the isolates resembling the type strain PG2 [12-14]. One of these studies however pointed out toward a subset of strains having particular genetic features that were found in *M. bovis*, a cattle pathogen closely related to *M. agalactiae*. These features were not found in the PG2 or PG2-like strains [13]. One of these particular strains, namely 5632, turned out to harbor (i) an integrative genetic element ICE of 23 kpbs which occurrence is low in the *M. agalactiae* species but high in *M. bovis* [15], (ii) a different repertoire of genes encoding surface lipoproteins known as the Vpmas [16], and (iii) other genetic elements yet to be characterized [13, 17]. While the 5632 and PG2 strains were both isolated from Spain, all the data accumulated so far tend to indicate that each stands at one end of the genetic spectrum encountered in the *M. agalactiae* species.

Mycoplasma inter-strain whole genome comparison has been carried only once for *M. hyopneumoniae*, an important pathogen of swine, and has provided evidence of intraspecific rearrangements resulting in strain-specific gene clusters as well as clues to factors related to pathogenicity [9]. In this study, to further comprehend the genome plasticity and the mechanisms responsible in mycoplasmas for intra-species genetic diversity, the complete genome of *M. agalactiae* strain 5632 was deciphered and compared with that of the type strain PG2 already sequenced by our group [11]. Althrought *M. agalactiae* is an important pathogen of ruminants [18 , 19] yet little known regarding its virulence or pathogenicity factors. Since all mycoplasmas lack a cell wall, the surface of their membrane acts as the primary interface in the interaction with the host and the environment. For instance, a number of *M. agalactiae* surface components has been shown to stimulate the host humoral response and includes lipoproteins such as the P80 [20], P40 [21], P48 [22], P30 [23] and the Vpma family [24]. Except for P80, all displayed a certain degree of variation in expression either in clonal population as for the phase-variable Vpmas [16, 24, 25] or among strains as shown for the P30 which promoter is mutated in the P30-negative 5632 strain [23]. In the current study, high-throughput identification of proteins expressed under laboratory conditions in *M. agalactiae* strain PG2 and 5632 was performed by a shotgun approach based on protein fractionation by 1D SDS-PAGE followed by proteolyses and nanoLC-

MS/MS. These proteomic data sets were used to validate genome annotation and, by comparative analyses, to further detect rare events that may be responsible for surface diversification. The combination of comparative genomics with comparative proteomics revealed that both large and localized events are shaping the *M. agalactiae* population structure which one might be much more dynamic than first expected because of limited amount of genetic material.

2. *Material and Methods*

Bacterial strains, culture conditions and DNA isolation

M. agalactiae type strain PG2, clone 55.5 [26] and strain 5632, clonal variant C1 [17] used in this study have been previously described. These strains have been independently isolated from goat in Spain. Experiments reported in this manuscript have all been performed with these clonal variants but for simplicity, we will further refer as the PG2 and the 5632 only. Mycoplasmas were propagated in Aluotto [27] or SP4 liquid medium [28] at 37℃ and genomic DNA w as extracted as described elsewhere [29 , 30].

M. agalactiae strain 5632 sequencing and annotation

Whole genome sequencing of strain 5632 was performed as follows. A library of 3 kb inserts (A) was generated by mechanical shearing of the DNA followed by cloning of the blunt end fragments into the pcDNA2.1 (Invitrogen) *E. coli* vector. Two libraries of 25 kb (B) and 80 kb (C) inserts were generated by *Hin*dIII partial digestion and cloning into the pBeloBAC11 (CALTECH) modified *E. coli* vector. The plasmid inserts of 10752, 3072 and 768 clones picked from the libraries A, B and C, respectively, were end-sequenced by dye-terminator chemistry on ABI3730 sequencers. The PHRED/PHRAP/CONSED software package was used for sequence assemblies. Gap closure and quality assessment were made according to the Bermuda rules with 10307 additional sequences. Annotation was performed as previously described using the CAAT-Box platform [31] with an automatic pre-annotation for CDS having a high similarity to PG2 followed by expert validation. The BLAST program suite was used for homology searches in non-redundant databases (http://www.ncbi.nlm.nih.gov/blast/blast.cgi). In order to determine the extent of sequence similarity, alignments between sequences were performed using softwares Needle (Needleman-Wunsch global alignment algorithm) and Water (Smith-Waterman local alignment algorithm) (http://www.ebi.ac.uk/Tools/emboss/align/).

Lipoproteins were detected as previously described [11] based on the presence (i) of the PROSITE Prokaryotic membrane lipoprotein lipid attachment site motif

(PROKAR_LIPOPROTEIN, Acc. Numb. PS00013) which sequence is [DERK](6)-[LIVMFWSTAG](2)-[LIVMFYSTAGCQ]-[AGS]-C and/or (ii) of two motifs previously defined by MEME-MAST that correspond to a charged N-terminal followed by a specific lipobox. After manual expertise, a total of 105 CDSs were annotated as predicted lipoproteins.

The tRNA genes were located on the chromosome using the tRNAscan-SE software [32]; the rRNA genes and the *rnpB* gene from the RNAseP system were searched using BLASTN by homology with *M. agalactiae* strain PG2 [11] and the tmRNA involved in translational surveillance and ribosome rescue was predicted using the ARAGORN software (http://130.235.46.10/ARAGORN/) [33].

Phylogenetic analyses were performed using MEGA 4.0 [34] and the Neighbor-Joining tree method. The reliability of the tree nodes was tested by performing 500 bootstrap replicates.

Comparative genome analysis

Comparative genomic analyses involving *Mollicutes* genomes were performed using a combination of tools available in the MolliGen public database [35] after incorporating the 5632 genome into a private section. Other bioinformatics softwares were used that include (i) Artemis [36], (ii) Artemis Comparison Tool (ACT) [37], (iii) mVISTA [38].

Proteomic analyses

The two *M. agalactiae* strains were respectively grown in Aluotto media and, after incubation at 37℃, the cells were collected by cen trifugation, washed three times in PBS before being resuspended in this buffer. One aliquot was used for defining the total protein content (PG2) while the remaining was subjected to protein partitioning using Triton-Tx114 as previously described [39]. Partitioning resulted in three fractions corresponding to hydrophobic (suspended in Triton-TX114), hydrophilic and insoluble proteins that were further subjected to 1D SDS-PAGE as such except for the hydrophobic fraction which was first precipitated overnight at -70℃ after addition of 9 volumes of cold MeOH and c entrifugated 10 min at 12,000x*g* and resuspended in loading buffer. The gel was sliced into about 15 sections which were subjected to trypsin digestion. Peptides were further analyzed

by nano liquid chromatography coupled to a MS/MS ion-trap mass spectrometer (LC-MS/MS).

Peptides were identified with SEQUEST through the Bioworks 3.3.1 interface (Thermo-Finnigan, Torrence, CA, USA) against a database consisting of both direct and reverse sense *Mycoplasma agalactiae* strain PG2 or 5632 entries. Using the following criteria (DeltaCN ≥ 0.1, Xcorr vs Charge State ≥ 1.5 (+1), 2.0 (+2), 2.5 (+3), Peptide Probability ≤ 0.001 and Number of Different Peptides ≥ 2) as validation filters, the False Positive rate is null. Table S1 summarizes the CDS which at least two specific peptides were detected in at least one of the fraction.

Database submission and web-accessible database

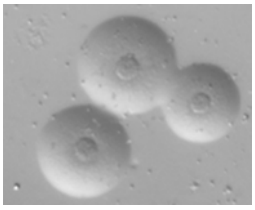
The genome sequence from *M. agalactiae* strain 5632, as well as related features were submitted to the EMBL/GenBank/DDBJ databases. All data were also loaded into the MolliGen database (http://cbi.labri.fr/outils/molligen/) [35].

3. *Results*

Whole genome and proteome comparison

Whole genome sequencing of the *M. agalactiae* strain 5632 revealed that it is composed of 1,006.7 kbp and thus, is ca.130 kpb longer than the genome of the PG2 type strain [11] (see Table 1 for general features).

Table 1: General properties of M. agalactiae PG2 and 5632 strains.

	PG2	**5632**
Isolation date	1952	<1991
Country	Spain	Spain
Source	nk	articulation
Host	caprine	caprine
Genome size (bp)	877,438	1,006,702
G+C (%)	29.70	29.62
Gene density (%)	88.5	88.7
Total number of CDS	752	826
HP	138	148
CHP	186	150
CDS with predicted function	404	505
Pseudogenes	45	23
rRNAs sets	2	2
tRNAs	34	34
GenBank accession number	CU179680	
ICE number	(1 vestigial)	3 (+ 2 vestigial)
Transposase number	1[a] (+ 2 pseudogenes)	15 (+ 2 pseudogenes)
Genomic DNA digested by : *Dpn*I or *Alw*I (sens. to Dam methylation)	Yes	No
*Dpn*II (Dam resistant)	Yes	Yes
Relative colony size[b]	100%	180%
		

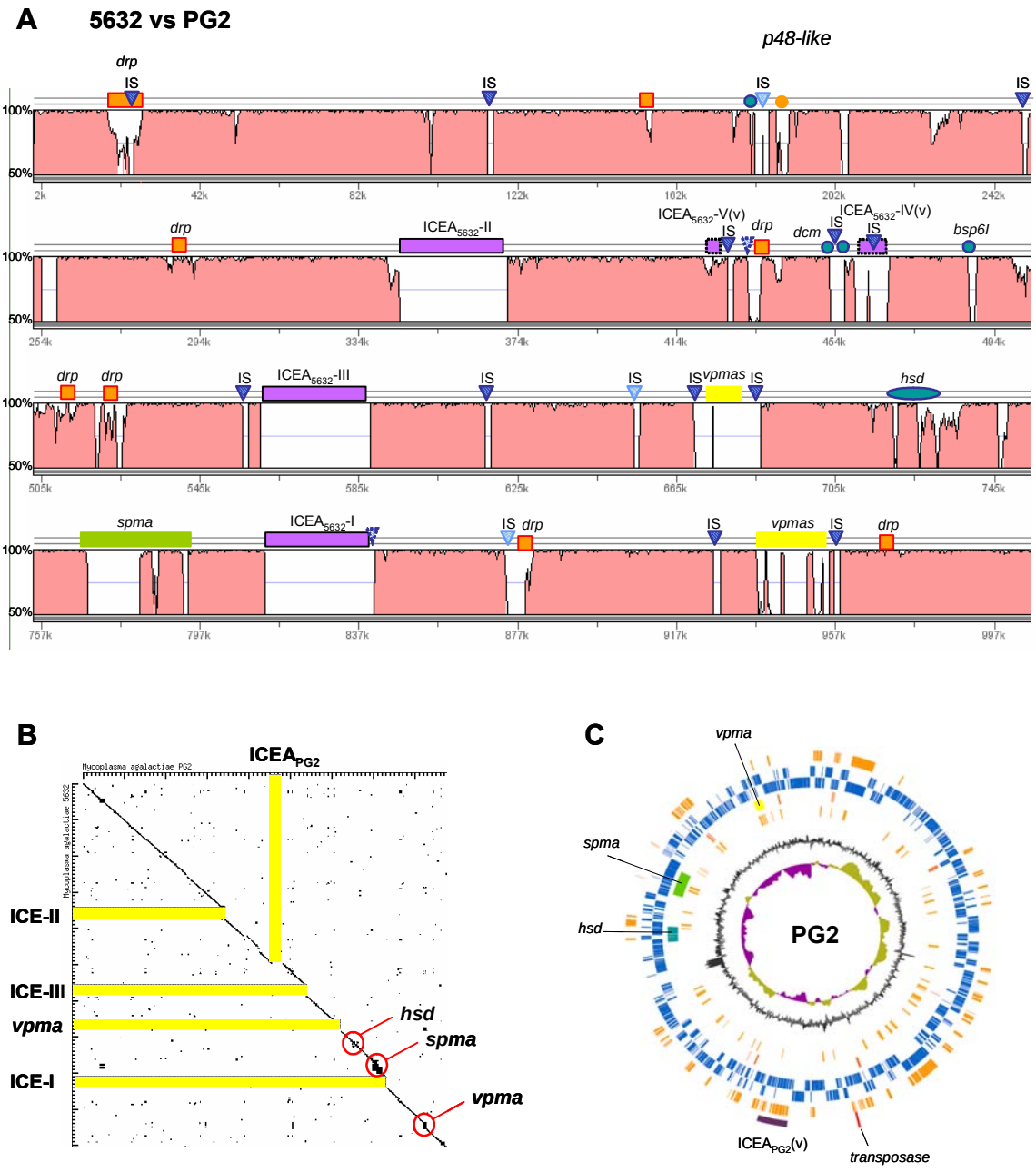
Data were extracted from the MolliGen database (http://cbi.labri.fr/outils/molligen).
[a] MAG3410 was annotated as Transposase, and is detected in proteomics analysis of this study but no IS inverted repeat sequences could be found.
[b] Relative size of colony observed on agar medium is given with PG2 as reference, colonies of 5632 are approximately 1.8 bigger than that of PG2. Image zoom ~200X.
nk, not known

Following annotation, the genome of 5632 displays a total of 826 CDS for only 752 in PG2, most of which are putative. Whole proteome analyses identified a set of 413 *M. agalactiae* putative proteins as being expressed under laboratory conditions in complex, axenic media. Of these proteins, 99 were found in only one strain and 314 in both. Among these, 139 were annotated as hypothetical, 41 were related to hypothetical ABC transporter while most of the remaining corresponded to house keeping genes (Table S1). These data indicate that over 40% of the *M. agalactiae* CDS were not detected by our global proteomic approach. Owing that in vitro cultivation of mycoplasmas is still delicate, it is likely that only a low number of CDS correspond to false ORF and that most are either not detected because of technical limitation of the approach (i.e. low protein quantities, resistance to digestion) or because they are only expressed under different conditions.

Comparison of the two genomes using the Molligen dot plot, VISTA and ACT softwares revealed an almost perfect synteny with no major genome rearrangement but with a number of regions being prominently different (Figure 1). These can be grouped in three main categories which are regions corresponding to (i) mobile genetic elements, (ii) restriction modification systems, and (iii) families of gene encoding surface proteins. As described below, these regions account for most of the difference in CDS content observed in between PG2 and 5632.

Figure 1: Overall comparison of *M. agalactiae* genomes from the PG2 and 5632 strains.

(A) VISTA comparison [38] of 5632 versus PG2. The genome of 5632 is used as a reference and sequence conservation was calculated using a slidding window of 100 bp. Graph represents the sequence percentage of nucleotide identity (in %) between the 5632 sequence versus that of PG2. Features of interest described in this manuscript were highlighted as follows. Orange boxes represent *drp* genes; blue triangles insertion sequence (IS) (dark blue for IS*Mag1*, light blue for IS*Mag2*, dotted lines for pseudogenes of transposase) and purple boxes indicate integrative conjugative elements (ICEs, with dotted lines for those corresponding to vestiges). The p48 lipoprotein gene is indicated by an orange circle and the CDSs related to restriction-modification systems by blue circles. The yellow and the green boxes represent the two *vpma*s loci and the *spma* locus respectively.



(B) CDS sequence comparison between PG2 and 5632 using a dot plot alignment of the genomes available at Molligen [35]. Each dot on the graph represents a blastp hit (threshold 10^{-8}) between a CDS of 5632 (ordinates) and a CDS of PG2 (abscises). On the graph axes, the length between the main ticks corresponds to 100 kbp. **(C) Circular representation of PG2 whole genome using DNAplot of the Artemis suite [40].** The circles represent the following (from outer to inner) : circle1, PG2 mobilome (restricted to one ICE vestige and a unique transposase gene; circle 2, CDS predicted as implicated in HGT with mycoplasmas belonging to the "mycoides" phylogenetic group; circle 3, positive strand annotated CDSs; circle 4, negative strand annotated CDSs; circle 5, CDS of interest discussed in this article (same colors than in panel 1: orange bars, CDS corresponding to *drp* genes; *hsd* locus, petrol blue; *vpma* locus, yellow; *spma* locus, green); circle 6, CDS predicted as lipoproteins; circle 7, percent G+C content (high G+C content in dark grey and low G+C content in light grey); circle 8, GC skew.

Role of the mobilome in *M. agalactiae* genetic diversity and genome plasticity

Analyses the 5632 genome revealed the presence of three large regions (ca. 27 kb) that correspond to an ICE element previously identified in this strain [15] and which was shown in this study to occur in three copies ($ICEA_{5632}$-I, -II and -III, with ICE-I corresponding to that previously published) representing a total of about 80 kbp. In addition, two smaller regions designated $ICEA_{5632}$-IV and -V were detected that relate to the degenerated, single ICE form found in strain PG2 [11] and that appear to be ICEs vestiges as suggested by their reduced size and the presence of insertion sequences and pseudogenes. Predicted proteins encoded by these ICEs were analysed by BlastP and designated according to our previous nomenclature (Figure 2). Interestingly, while phylogenetic and BLAST analyses indicate that $ICEA_{5632}$-I to –III are related to the ICEF of *M. fermentans* PG18 [41], the vestiges $ICEA_{5632}$-IV and –V as well as the degenerated $ICEA_{PG2}$ are somewhat similar to the ICEC of *M. capricolum*; in particular they all contain CDSZ that is not found in $ICEA_{5632}$-I to –III nor in ICEF. $ICEA_{5632}$-IV also contains a CDS, CDS3, which is present as a pseudogene in both the $ICEA_{PG2}$ and ICEC but is absent from the large copies $ICEA_{5632}$-I to III. In *M. fermentans*, the homologous sequence ORF3 is proposed to be a relaxase involved in DNA processing during the conjugative transfer of the ICEs [41]. The ICEA copies I to IV also contain two CDS widely conserved in mycoplasmal ICEs, CDS22 and CDS5 (the copy -V appear to have a degenerated version of CDS22). ICEA left and right borders of the ICEA have been defined for the copy -I [15] and indeed, the copies -II and -III were shown to start by and end with an identical inverted repeat GGAA-TTCC. The high level of sequence conservation between the two *M. agalactiae* genomes allows to define the insertion points of the 3 large ICE copies in the 5632 chromosome which correspond to intergenic regions in PG2. Although ICE insertions do not result in apparent gene disruption, the targeted regions seem to be prone to a certain extend of instability: The copy -I is located next to a conserved insertion sequence (IS), present as a pseudogene in both strains but showing some sequence divergence; the copy -III is also in the vicinity of an IS present only in strain 5632 and which is inserted next to a conserved tRNA gene. The copy -II is inserted next to a predicted lipoprotein gene (MAG2840 or MAGa2970) showing clear sequence divergence as the two predicted proteins

have only 66.4% identities. The three ICE copies I-III are flanked by an almost perfect 9 bp direct repeat which is most likely generated during the integration process. Alignment of ICEA-I, -II and -III DNA sequences using ClustalW from the first G of the GGAA repeat to the last C of the TTCC repeat as the end showed that there are highly similar, presenting only 7 to 8 SNPs. This suggests that the copies originated from subsequent excision and integration events, possibly during chromosomal replication or by exchange within the population. SNPs resulted in generating (i) two pseudogenes corresponding to CDS16 of copy II and CDS27 of copy III, (ii) truncation of CDS22 in copy II and (iii) insertion a codon in a stretch of repeated AAT (poly Asn) in CDS14 of copy III. Overall, 5632 ICEAs account for 21 different CDS that are not present in PG2, two of which were detected by MS/MS in the Triton-X114 fraction of 5632 (CDS14 and 17).

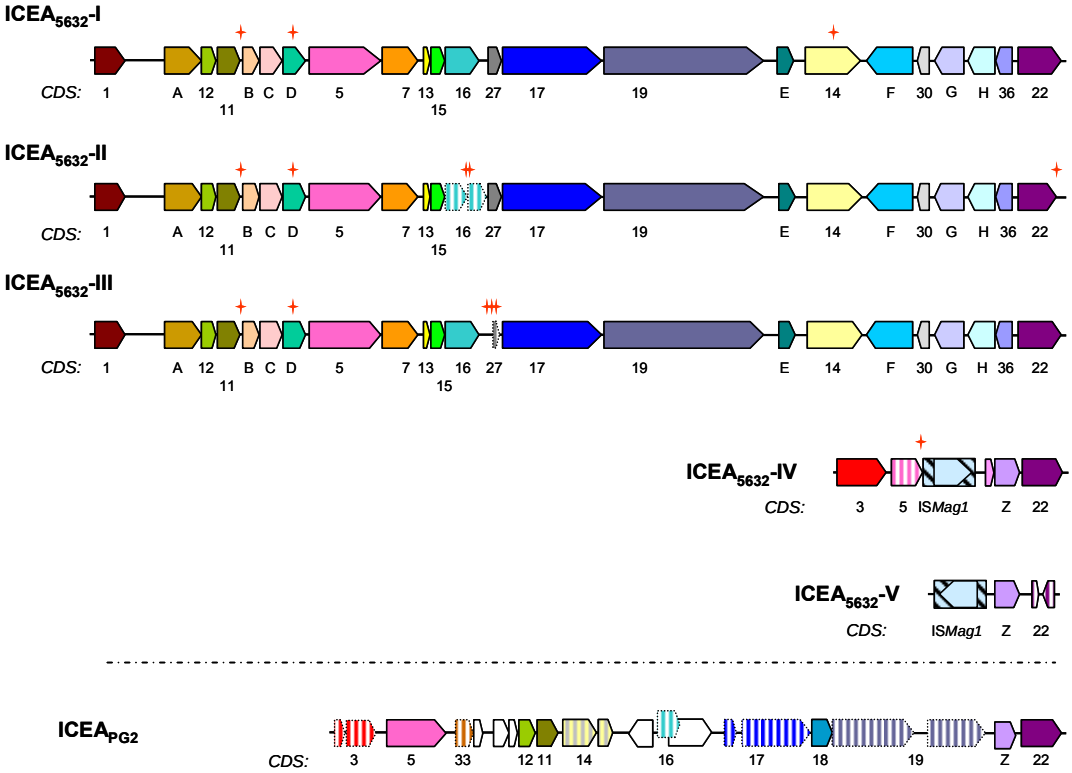


Figure 2: Comparison of entire and vestigial ICEs found in *M. agalactiae* strains PG2 and 5632.

Schematics represent ICEs encountered in 5632 (A) and in PG2 type strain (B). Large arrows represent CDSs, with homologous CDS labelled with the same color. CDS nomenclature indicated below arrows is based on the first ICE study in 5632 [15]. $ICEA_{5632}$-I, -II, -III, -IV, -V extend from MAGa7100 to 6880, MAGa2980 to 3220, MAGa4850 to 5060, MAGa4050 to 4010, MAGa3690 to 3670, respectively. $ICEA_{PG2}$ extend from MAG4060 to 3860. Red crosses indicate SNPs or indels in between ICEs from 5632. Insertion sequence elements (IS*Mag1*) are represented by shaded boxes with transposase CDS in light blue. Pseudogenes are represented by hatched colours with dotted lines.

Other mobile elements were found in the 5632 genome but not in PG2 and correspond to multiple copies of two types of IS elements that both belong to the IS30 family. The location of these elements relative to their flanking CDS is shown in Figure 3.

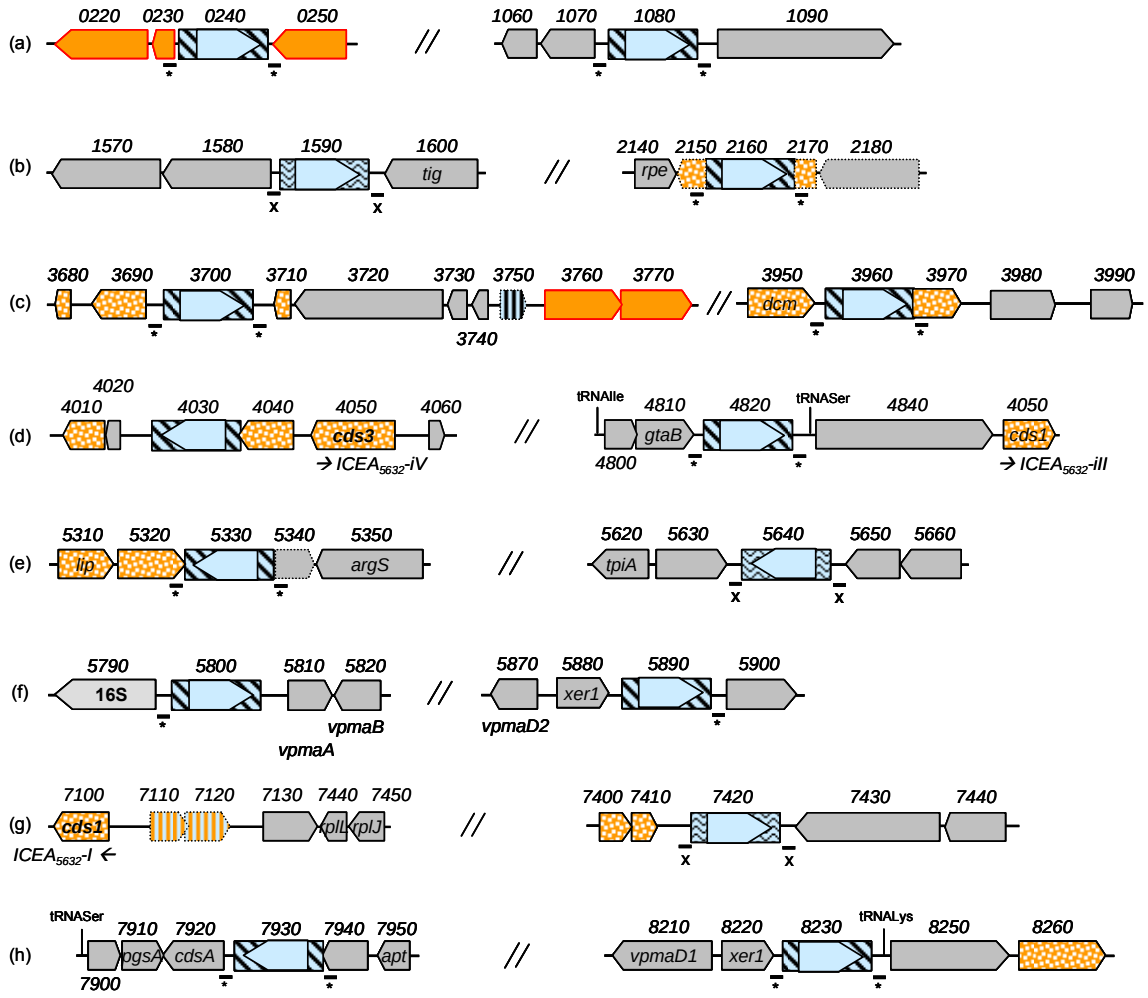


Figure 3: Location of insertion sequence and their flanking sequences in *M. agalactiae* strain 5632.

Schematics representing genomic regions that flank insertion sequence elements in strain 5632. Large arrows represent CDSs. Insertion sequence elements are represented by blue boxes filled with straight or undulating lines for IS*Mag1* and for IS*Mag2*, respectively, with transposase CDS indicated by light blue arrows. CDS predicted as implicated in HGT with mycoplasmas of the "mycoides" group are filled by a dotted orange pattern, except for *drp* genes which are filled in orange with red outlines. MAGa7110 and MAGa7120 are transposase pseudogenes predicted has implicated in HGT with the "mycoides" group, they are hatched in orange and light blue. A hyphen with an asterisk (*) above indicates the presence of a 14-nucleotides direct repeat that flanks IS*Mag1*, a hyphen with letter 'x' above indicates the presence of a 25-nucleotides direct repeat that flanksIS*Mag2*. Pseudogenes are indicated by filled arrows with dotted lines.

The IS element, IS*Mag1*, has previously been described in some *M. agalactiae* strains [42] and an isoform was also described in *M. bovis* (named IS*Mbov1*) [43].

In 5632, this element occurs in 12 copies with 10 that localized either next to genomic islands encoding repertoire of variable surface lipoproteins (see MAGa5980, MAGa5800 and MAGa8230) or to regions associated with HGT (Figure 1C). The second type of IS, IS*Mag2*, resembles IS*Mbov6* recently described [44] and is found only in three copies. None of which seems to truncate or disrupt a CDS.

In silico analyses further indicate that IS*Mag1* occurrence most likely affects gene expression in only three cases. In the first case, the insertion has taken place in between a *dcm* gene encoding a Cytosine-specific DNA-methyltransferase Sau96I (MAGa3950) and a Type II specific deoxyribonuclease *sau96I-like* gene (MAGa3970). Both genes are absent from PG2 but are found next to each other in *M. mycoides* subsp. *mycoides* SC in which they most likely occur as an operon. These genes are highly similar to those found in 5632 with ca. 85% and 78% similarity, respectively, with most difference taking place in the N-terminal of 5632 *sau96I-like* gene. Interestingly, a global proteomic approach using MS/MS (see below and Table S1) detected several specific peptides of the cytosine-specific methyl transferase encoded by MAGa3950 but none corresponding to the *sau96I-like* gene suggesting that the presence of the IS at its 5'end may affect its expression. In the second case, IS insertion at the 3' end of MAGa4040 would result in truncating the product by more than 50% when compared to the situation found in PG2. The third case relates to MAGa5320 and MAGa5350 which are separated by an IS and which have been annotated as two distinct pseudogenes because each present a high similarity with the either N- or the C-terminal part of the *Mycoplasma capricolum* subsp. *capricolum* glycosyl transferase (MCAP0063), respectively. In PG2, homologs to MAGa5320 and MAGa5350 also exist as pseudogenes although no IS is involved.

Finally, two vestiges of transposase having similarities with that of IS*Mmy1* of *M. mycoides* subsp. *mycoides* SC were detected in the 5632 genome, one located next to an ICE element while the other was found next to a hypothetical protein truncated in N-terminal that displays a DUF285 motif (see below) and that is predicted to have undergone HGT with member of the "mycoides" cluster species.

In most cases, IS were flanked by direct repeated sequences of 14 nt for IS*Mag1* and of 25 nt for IS*Mag2* that signed a single IS insertion event. Exception were

found for IS located next to the *vpma* gene family as previously described [16] and was also observed here for the IS insertion located at the 3'end of MAGa4040 suggesting that further genomic rearrangements have occurred in this area. Indeed, this region has been described above as a putative vestige of an ICE integration. Finally, a single transposase gene (MAG3410; see also Figure 4) is found in PG2 but not is 5632 and is detected by proteomic. This transposase has some similarities (46.7%) with an IS*Mmy1* transposase of *M. mycoides* subsp. *mycoides* SC but no flanking repeated sequence could be readily detected.

These data indicate that ca. 76% of the additional genetic material of the 5632 genome is composed of mobile genetic elements when compared to PG2. This represents 10 % of the genome, yet these do not lead to major genome rearrangement. Overall, 5632 has 95 specific additional CDS, 72 of which correspond to CDS of ICE or transposase. Among the remaining, 8 relate to the *spma* family (see below), 11 to hypothetical products and 4 to restriction modification systems (RM).

Protection from DNA degradation and invasion

While transformation of strain PG2 using a tailored-made cloning vector [25, 45] has become a standard protocol in our laboratory, attempts to transform 5632 in the same or modified conditions repeatedly failed. As well, 5632 chromosomal DNA is resistant to two type II restriction enzymes, *Alw*1 (GGATCNNNN↓N) and *Dpn*II (↓GATC), that are sensitive to Dam methylation and that digest DNA extracted from PG2. Conversely 5632 DNA is digested by *Dpn*I (GA↓TC), which cleaves only its recognition site is methylated on the adenine. This suggests the presence in the two strains of a different set of restriction-modification systems and, indeed, 5632 contains four additional CDS that encode for two type II RM systems, each displaying genes for a restriction enzyme and its corresponding methylase. The first one is similar to the *Bacillus sp.* Bsp61I RM system while the other resembles that of the Sau96I-like found in *M. mycoides* subps. *mycoides* SC as already mentioned above. Indeed, phylogenetic tree reconstructions, although not fully demonstrative, suggested that the Bsp61I RM system has most likely been acquired by HGT from the *Firmicutes* while the Sau96I-like system has probably been exchanged with members of the "mycoides" cluster. Further

detailed comparative analyses revealed that 5632 is better equipped than PG2 in terms of RM systems and more specifically in DNA methylase. As indicated in Table 2, 5632 encodes for 11 putative DNA methylase of which 9 were found to be expressed under laboratory conditions while for PG2 this number is only of 8 with the expression of only 3 being detected. Interestingly, one methylase gene seems to have been duplicated in 5632 when compared to PG2. The two paralogs differs from each other and from their PG2 homolog mostly in the central part (ca. aa205-aa400) which is known to contain the N6_N4_Mtase domain (PF01555 in Pfam). Whether this provides the corresponding enzymes with different specificity is not known. Both strains display a locus of six genes with homology to type I restriction-modification (R/M) systems that was designated *hsd* and is composed of (i) two *hsdM* genes coding for two almost identical modification enzyme which would methylate specific adenine residues, (ii) three *hsdS* genes each coding for a distinct R/M specificity subunit (HsdS) that shares homologies with each other (between 50 to 97% similarities) and (iii) one *hsdR* gene encoding a site-specific endonuclease (HsdR). Interestingly, in PG2, the *hsdR* gene is disrupted by the insertion of two nucleotides in a polyA tract localised in the middle of the gene that result in a premature stop codon. This is in agreement with the detection in 5632 of peptides specific of the HsdR enzyme but not in PG2. In mycoplasma, polyA tracts of this length have often been involved in high frequency variation [46]. Finally, the *hsd* locus also contains a hypothetical CDS whose product displays a high similarity with a phage family integrase of *Bifidobacterium longum* [47] and motifs found in molecules involved in DNA recombination and integration. In *M. pulmonis*, the *hsd* locus has been shown to undergo frequent DNA rearrangements but the gene encoding the putatively involved recombinase is located elsewhere on the genome [48]. If the *hsd* locus of *M. agalactiae* is functional then it is worth speculating that recombinase-mediated DNA rearrangements could modulate the specificity of the system by recombining the HsdS-coding sequences, some of which occur in opposite orientation on the chromosome. Attempts to demonstrate DNA rearrangement of the *hsd* using basic molecular approaches failed. Whether this is due to the difficulties in finding specific sequence signatures that would demonstrate recombination is not known. Nevertheless, it is worth noting that they most diverge in their *hsdS* gene

sequences suggesting one more time this system could generate a high diversity in specificity.

Table 2: Restriction/Modification products comparison between strains PG2 and 5632.

MAGa[a]	MAG[b]	Product	Similarity (%)	MS/MS[c] 5632	MS/MS[c] PG2	Comments
MAGa1570	MAG1530	Type III R/M system:Methylase	75.3	+	+	
MAGa1580			77.6	+		
MAGa1770	MAG1790	DNA methylase	97.8	-	-	
MAGa2070	MAG2070	DNA methylase	98.9	+	-	
MAGa2700	*MAG2550*	Adenine-specific DNA methyltransferase	65.8	-[d]	-	Pseudogenes in PG2
	MAG2560				-	
MAGa2710	*MAG2570*	Type II restriction endonuclease **	46.1	-	-	Pseudogenes in PG2
	MAG2580				-	
No homolog	MAG3310	CpG DNA methylase	na	na	-	
No homolog	MAG4030	Conserved hypothetical protein	na	na	-	BBH:*Mmm*SC - Putative C5 methylase (40%)
MAGa4470	MAG4250	Pseudogene of CPG DNA methylase (N-terminal)	83.4	-	-	
MAGa4480	MAG4260	Pseudogene of CPG DNA methylase (C-terminal)	94.7	-	-	
MAGa6280	MAG5640	Type I R/M system specificity subunit	75.0	-	+[d]	Locus *hsd*
MAGa6290	MAG5650	Modification (Methylase) protein of type I restriction-modification system HsdM	98.3	+	-	Locus *hsd*
MAGa6310	MAG5680	Type I R/M system specificity subunit	32.4	-	+[d]	Locus *hsd*
MAGa6330	*MAG5700*	HsdR, R/M enzyme subunitR	95.0	+	-	Pseudogenes in PG2
	MAG5710				-	
MAGa6340	MAG5720	Type I R/M system specificity subunit	30.9	+	-	Locus *hsd*
MAGa6350	MAG5730	Modification (Methylase) protein of type I restriction-modification system HsdM	90.3	+	+	Locus *hsd*
MAGa7650	MAG6680	Modification methylase	97.6	+	-	Modification methylase
MAGa3200 MAGa5050 MAGa6900	No homolog	CDSH	na	-	na	BBH: 92.0% with MCAP0297 - *Mcap* - adenine-specific DNA methylase
MAGa4250	No homolog	Modification methylase Bsp6I	na	+	na	BBH: 81.7 % *Bacillus sp. bsp6*IM Modification methylase Bsp6I
MAGa4260	No homolog	Type II restriction enzyme Bsp6I	na	+	na	BBH:55.1% *Bacillus sp bsp6*IR Type II restriction enzyme Bsp6I
MAGa3950	No homolog	Cytosine-specific methyltransferase	na	+	na	BBH: *Mmm*SC MSC_0216 *dcm* Cytosine-specific DNA-methyltransferase Sau96I
MAGa3970	No homolog	Type II site-specific deoxyribonuclease, *sau*96I-like	na	-	na	BBH: *Mmm*SC MSC_0215 *sau*96I Type II site-specific deoxyribonuclease

[a] CDS of *M. agalactiae* strain 5632 (Molligen Mnemonic)
[b] CDS of PG2 (Molligen Mnemonic), pseudogenes are indicated in italic
[c] Proteomic (see materials and methods) : (+) indicates that peptides were detected by MS/MS for the corresponding CDS, suggesting expression of the corresponding gene, (-) indicates that no specific peptides were detected for the corresponding CDS.
[d] only one peptide detected
[e] MAG5640 and MAG5680 have common peptides.
BBH, Best Blast Hit; *Mmm*SC, Mycoplasma *mycoides* subsp. *Mycoides* SC; *Mcap*, *M. capricolum* subsp. c*apricolum*; na, not applicable; R/M, Restriction / Modification

The flexible gene pool: towards a highly dynamic surface architecture

Comparison of the two *M. agalactiae* genomes further revealed that 5632 contains 105 CDS predicted as lipoproteins for 65 in PG2. Proteomic analyses further confirmed the expression of more than 50% for both strains, with at least 56 being expressed in 5632 for 43 in PG2, all but one being detected in Triton X-114 (see Table 3 for a detailed list). In most cases, these differences are linked to specific genes present in one strain but not in the other (i.e. genes belonging to 5632-ICE and encoding lipoproteins such as CDS14) and to gene family such as the *vpma* that encodes related, phase-variable, lipoproteins and account for 23 CDS in 5632 for only 6 in PG2. As previously described, all *vpma* genes but two (*vpmaK* and *vpmaL*) were shown to be expressed at one point during *in vitro* propagation in 5632 [16].

Table 3: Lipoproteins and MS/MS detection in Tx-114 phase.

MAGa[a]	MAG[b]	Gene name	Product	Tx 5632[c]	Tx PG2[c]	Comments
MAGa0140	MAG0120		Conserved hypothetical protein, predicted lipoprotein, P48	+	+	
MAGa0380	MAG0380	*oppA*	Oligopeptide ABC transporter, substrate-binding protein (OppA), predicted lipoprotein	+	+	
MAGa1090	MAG1000		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa1140	MAG1050		Hypothetical protein, predicted lipoprotein	+	+	
MAGa1490	MAG1450		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa1550	MAG1510		Hypothetical protein, predicted lipoprotein	+	+	
MAGa1620	None		Conserved hypothetical protein, P48-like	**+**	na	No signal peptide and lipobox except if variation of the length of a poly G_{10} (+/-1) upstream the chosen start
MAGa1680	MAG1670		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa1980	MAG1980		Hypothetical protein, predicted lipoprotein	+	+	
MAGa1980	MAG1980		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2000	MAG2000		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2330	MAG2220		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa2500	MAG2340		Conserved hypothetical protein, predicted lipoprotein	+	-	Not predicted as lipoprotein in PG2 due to variation of the length of a poly A (A_6 in PG2, A_7 in 5632)
MAGa2510	MAG2350		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2570	MAG2400		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2580	MAG2410		P40, predicted lipoprotein	+	+	
MAGa2600	MAG2430		Conserved hypothetical protein, predictedlipoprotein, DUF285 family	+	+	
MAGa2670	MAG2510		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2690	***MAG2540 +MAG2530***		Hypothetical protein, Vpma-like, predictedlipoprotein	+	+	Truncated in PG2 (MAG2530=3'end, expression not detected)
MAGa2740	MAG2610		Hypothetical protein, predicted lipoprotein	+	+	
MAGa2820	MAG2690	*phnD*	Alkylphosphonate ABC transporter,substrate-binding protein, predicted lipoprotein	+	+	
MAGa2970	MAG2840		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa3160	None		CDS14	+	na	ICE
MAGa3250	MAG2870		Conserved hypothetical protein, predicted lipoprotein	+	-	None
MAGa3330 +MAG3340	MAG2950		Hypothetical protein, predicted lipoprotein	-	+	Variation of the length of a poly C (C_9 in PG2, C_8 in 5632) downstream of MAGa3330 may be responsible for frameshifting
MAGa3640	MAG3240		Conserved hypothetical protein, predicted lipoprotein	+	+	Not predicted as lipoprotein in PG2
MAGa3820	***MAG3460***		Hypothetical protein, predicted lipoprotein	+	-	Variation of the length of a poly G (G_8 in PG2, G_9 in 5632) upstream of MAG3460 may be responsible for frameshifting
MAGa3830	MAG3470	*p30*	P30, predicted lipoprotein	-	+	Mutation in the *p30* promoter region of 5632 (Fleury *et al.* [23])
MAGa3980	MAG3590		Hypothetical protein, predicted lipoprotein	-	+	
MAGa3990	MAG3600		Hypothetical protein, predicted lipoprotein	+	+	
MAGa4680	MAG4460		Conserved hypothetical protein, predictedlipoprotein	+	+	
MAGa5010	None		CDS14	+	na	ICE
MAGa5110	MAG4640		Conserved hypothetical protein, predictedlipoprotein	-	+	none
MAGa5190	MAG4720		Conserved hypothetical protein, predicted lipoprotein	-	-	Detected in Insoluble pellet for 5632
MAGa5210	MAG4740		Hypothetical protein, predicted lipoprotein	+	+	

MAGa5420	***MAG4960 +MAG4950***		Conserved hypothetical protein, predictedlipoprotein	+	-	MAG4960+MAG4950 previously annotated as pseudogenes and detected in total proteins but not in detergent TX-114 phase
MAGa5490	None[d]		Hypothetical protein, predicted lipoprotein	+	+	CDS missed during annotation of PG2 (nt 586236 to 585832)
MAGa5500	MAG5030		P80, predicted lipoprotein	+	+	
MAGa5510	MAG5040		Conserved hypothetical protein, predicted lipoprotein	+	+	
MAGa5560	MAG5080		Hypothetical protein, predicted lipoprotein	+	+	
MAGa5630	MAG5150		Hypothetical protein, predicted lipoprotein	+	+	Not predicted as lipoprotein in PG2 due to start chosen
MAGa5830	None	*vpmaC*	Variable surface lipoprotein C (VpmaC)	+	na	Duplicated (MAG8080)
MAGa5850	None	*vpmaE*	Variable surface lipoprotein E (VpmaE)	+	na	Duplicated (MAGa8090)
MAGa5860	None	*vpmaF1*	Variable surface lipoprotein F1 (VpmaF1)	+	na	Duplicated (MAGa8170)
MAGa5870	None	*vpmaD2*	Variable surface lipoprotein D2 (VpmaD2)	+	na	Duplicated (MAGa8120)
MAGa6560	MAG5910		5'Nucleotidase, predicted lipoprotein	+	+	
MAGa6940	None		CDS14	+	na	ICE
MAGa7130	MAG6170		Hypothetical protein, predicted lipoprotein	+	+	
MAGa7160	MAG6200		Hypothetical protein, predicted lipoprotein	+	+	
MAGa7470	***MAG6490 +MAG6480***		Hypothetical protein, predicted lipoprotein, DUF285 family	+	-	Variation of the length of a poly A (A_6 in 5632, A_7 in PG2) maybe responsible for frameshift
MAGa7490	MAG6520		Conserved hypothetical protein, predicted lipoprotein	+	+	
MAGa8040	None	*vpmaG*	Variable surface lipoprotein G (VpmaG)	+	na	*vpma* family
MAGa8050	None	*vpmaF2*	Variable surface lipoprotein F2 (VpmaF2)	+	na	*vpma* family
MAGa8060	MAG7070	*vpmaX**	Variable surface lipoprotein X (VpmaX)	+	+	*vpma* family
MAGa8070	MAG7060	*vpmaW**	Variable surface lipoprotein W (VpmaW)	+	+	*vpma* family
MAGa8100	None	*vpmaB*	Variable surface lipoprotein B (VpmaB)	+	na	Duplicated (MAGa8100)
MAGa8110	None	*vpmaA*	Variable surface lipoprotein A (VpmaA)	+	na	Duplicated (MAGa8110)
MAGa8150	None	*vpmaH*	Variable surface lipoprotein H (VpmaH)	+	na	*vpma* family
MAGa8160	None	*vpmaI*	Variable surface lipoprotein I (VpmaI)	+	na	*vpma* family
MAGa8180	None	*vpmaJ*	Variable surface lipoprotein J (VpmaJ)	+	na	*vpma* family
MAGa8210	None	*vpmaD1*	Variable surface lipoprotein D1 (VpmaD1)	+	na	Duplicated (MAGa5840)
MAGa8260	MAG7130		Hypothetical protein, predicted lipoprotein	+	-	Not predicted as lipoprotein in PG2 due to a point mutation: TAA (ochre) ↔TCA (serine))
None	MAG1570		Hypothetical protein	-	+	No signal peptide and lipobox except if variation of the length of a poly G_9 (+/-1) next to the chosen start
None	MAG7050	*vpmaV*	Variable surface lipoprotein V (VpmaV)	na	+	*vpma* family
None	MAG7080	*vpmaY*	Variable surface lipoprotein Y (VpmaY)	na	+	*vpma* family
None	MAG7090	*vpmaU*	Variable surface lipoprotein U (VpmaU)	na	+	*vpma* family
None	MAG7100	*vpmaZ*	Variable surface lipoprotein Z (VpmaZ)	na	+	*vpma* family

[a] CDS of *M. agalactiae* strain 5632 (Molligen Mnemonic).
[b] CDS of PG2 (Molligen Mnemonic), pseudogenes are indicated in italic bold.
[c] Peptides detected by MS/MS in the Triton X114 phase (see materials and methods) : (+) indicates that peptides corresponding to CDS were detected, suggesting expression of the corresponding gene, (-) indicates that no peptides corresponding to CDS were detected.
[d] CDS detected in proteomic but no Mnemonic have been defined because CDS was missed during PG2 original annotation [11].
na, not applicable.

Hypothetical related surface lipoproteins are encoded by two other gene families: the so-called *drp* (for DUF285 related proteins) and the *spma* (surface protein of *M. agalactiae*). Unlike the *vpma*, CDS encoding product with DUF285 motifs are scattered around the chromosome, with both strains having a similar size of the repertoire, which comprises 12 CDS identified as Drp (and one pseudogene) in PG2 and 13 in 5632. One particularity of this family is that it belongs to the gene pool that underwent HGT with members of the "mycoides" cluster. Comparison of 5632 with PG2 revealed that they often localized in regions that most vary in between the two strains. This is illustrated in the Figure 4 that shows the *drp* CDSs and their flanking regions. Except for one locus (MAGa2580 to MAGa2630), all present a different organization when comparing 5632 with PG2 that involves the occurrence of complex DNA rearrangements (i.e. locus MAGa3630 to MAGa3780), the presence of additional IS element or CDS in 5632 (i.e. locus MAGa7410 to MAGa7490), and the occurrence of pseudogenes in PG2 (i.e. MAG4200 and 4210). Interestingly, only two Drp proteins were detected by proteomic LC-MS/MS, with only one being expressed in both strains (MAG2430 and MAGa2600). The other expressed Drp products are MAG4220 in PG2 and MAGa7470 in 5632. These are encoded at two different loci. The homolog of MAGa7470 occurs as a pseudogene in PG2 because of a difference in the length of polyA tract creates a frameshift resulting in a premature stop codon. For MAG4220 of PG2 and its counterpart in 5632, MAGa4450, there is no apparent molecular feature that could account for the difference in expression but these two products slightly differ from the rest of the family in that they have signal peptide but no lipobox.

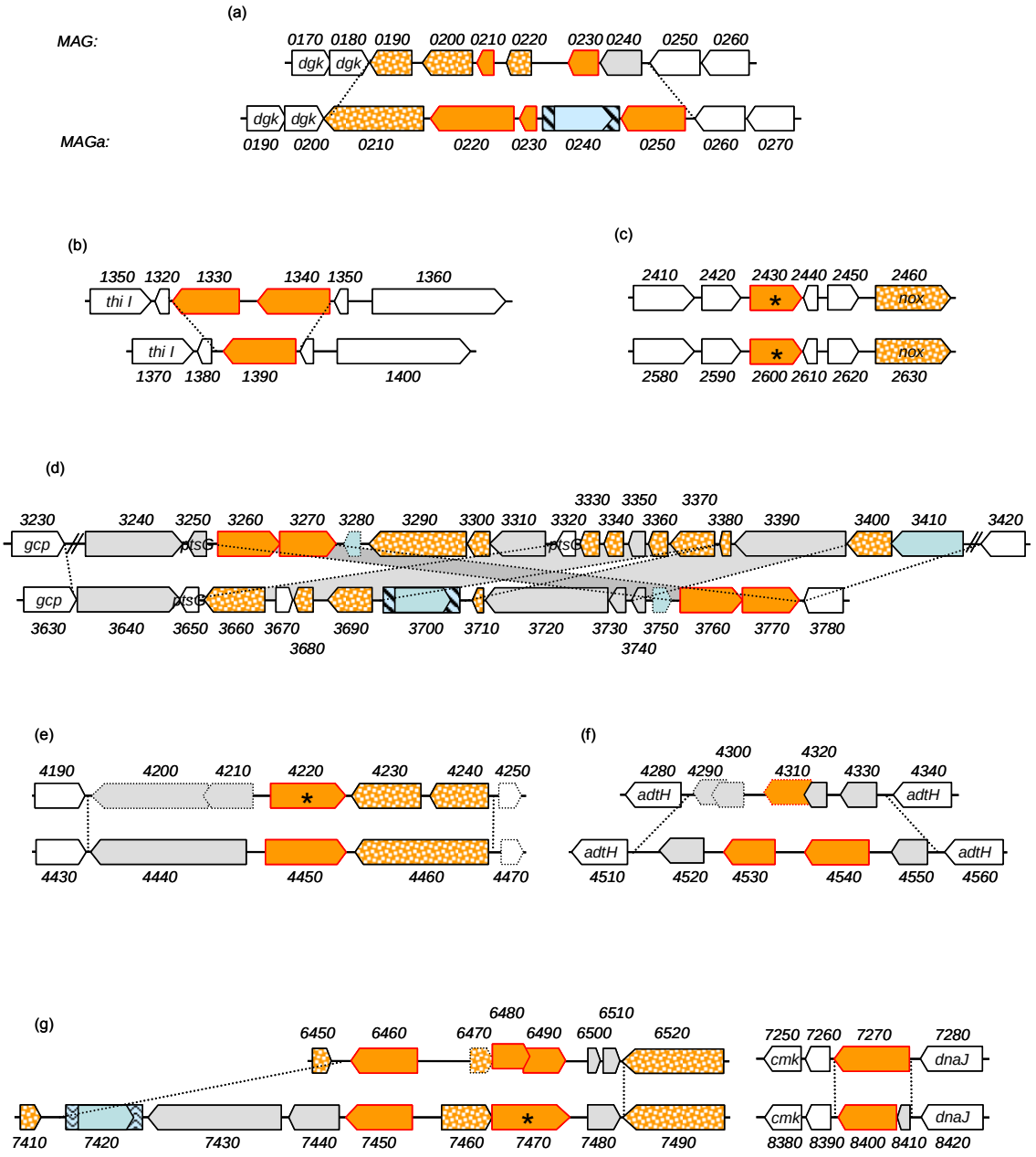


Figure 4: Comparison of *M. agalactiae* PG2 and 5632 revealed that the *drp* loci are a reservoir for strain genetic and surface diversity.

Schematics representing the comparison of genomic regions containing *drp* genes in between strains PG2 and 5632. Large arrows represent CDSs. CDSs corresponding to *drp* genes are filled in orange with red outlines. CDSs others than *drp* and predicted as implicated in HGT with mycoplasmas of the "mycoides" group are filled in dotted orange. White arrows are CDSs conserved between PG2 and 5632. Insertion sequence elements are represented as in Figure 3. Drp detected by LC-MS/MS are labelled by an asterisk (corresponding in PG2 and 5632, respectively, to MAG2430 and MAG4220, MAGa2600 and MAGa7470). Pseudogenes are represented by filled arrows with dotted lines.

Comparison of the 5632 and PG2 genomes revealed one particular locus composed of several putative CDS encoding (i) a similar N-terminal signal peptide followed by a highly conserved lipobox and (ii) particular amino-acid motifs that are repeated with in a particular product. This gene family was further designated as *spma* for “surface protein of *M. agalactiae*” and is larger in 5632 with 8 *spma* genes for only 4 in PG2. Analyses of the two *spma* loci indicate that *spma* genes present in 5632 but not in PG2 have orthologs in the "mycoides" cluster only. More specifically, *M. mycoides* subsp. *mycoides* LC strain GM12 [49] and *M. capricolum* subsp. *capricolum* contain 5 and 1 genes, respectively, that encode putative lipoproteins resembling 5632-Spmas and carrying the motif 3. Although this question cannot be formally addressed by phylogenetic tree reconstruction, this *spma* sequence comparison suggests that these genes are part of the gene pool which has been exchanged in between *M. agalactiae* and members of the "mycoides" cluster. The proteomic approach taken in this study failed to detect any of the Spma products in one or the other *M. agalactiae* strain. Whether these proteins have been missed by this approach or whether they are expressed under different conditions remains to be assessed. A stretch of polyG was found at the 5’untranslated region of each putative *spma* gene (Figure 5). This last feature is unusual in mycoplasmas that have low GC content and is particularly striking in the 5632 *spma* locus which displays 8 polyG tracts with one containing up to 13 G residues. Whether these control or affect the transcription of downstream genes is not known but homopolymeric tracts of residues have often been associated with products whose expression is phase variable in mycoplasmas.

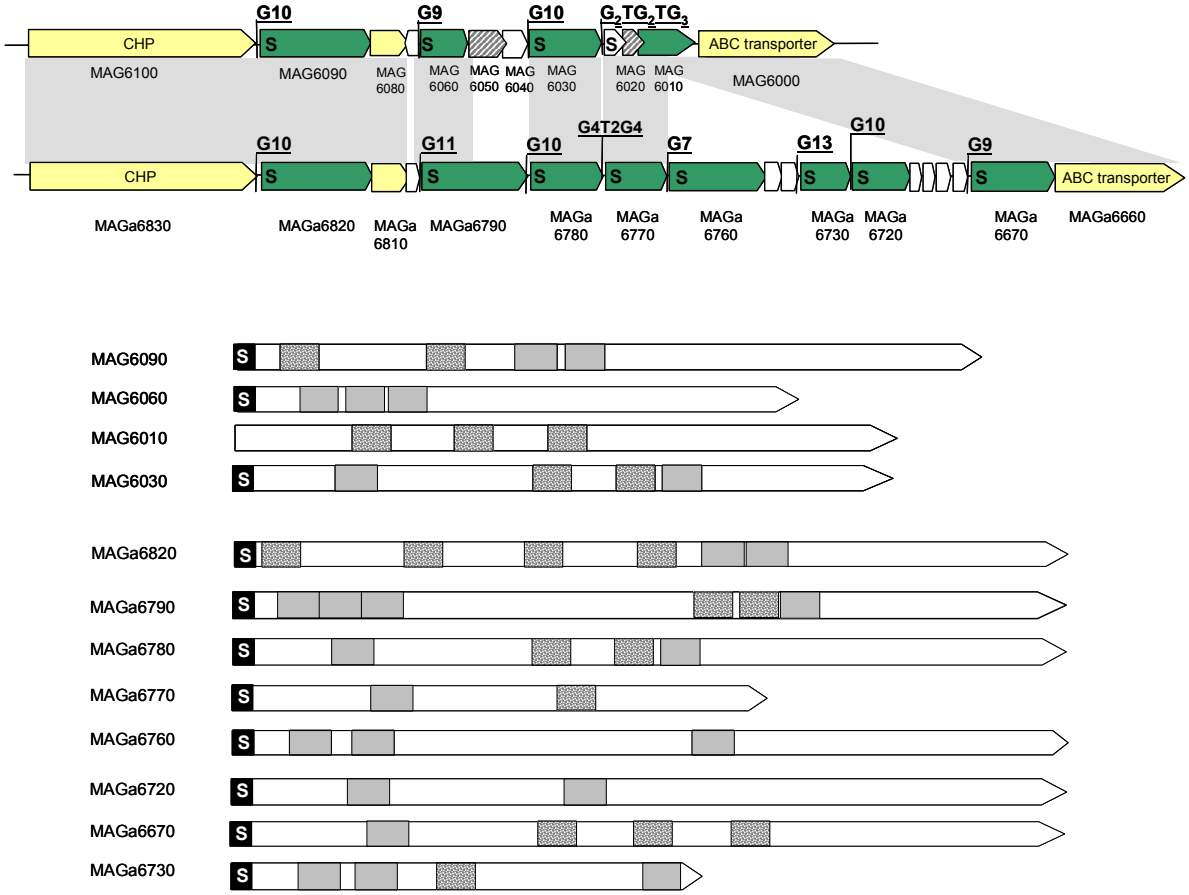


Figure 5: The *M. agalactiae spma* repertoire is extended in strain 5632.

Schematics represent the genomic organization of *spma* locus in comparison between strains PG2 and 5632 (A) and structural features of *spma* gene products in both strains (B). In panel A, CDS corresponding to *spma* genes are represented in green. Letter S in *spma* arrows represent sequence corresponding to a signal peptide. CDSs conserved between PG2 and 5632 are filled in light yellow. Tracks of repeated nucleotides found before *spma* coding sequences are also indicated. In panel B, predicted Spma proteins are represented schematically by large arrows beginning generally with an homologous amino-acid leader sequence (black boxes labelled S) followed by regions that have homology between *spma* gene products or that are repeated within the same product (blue dotted and grey boxes)

Interestingly, polyG tracts were found elsewhere in the genome of 5632, again at the 5'end of gene encoding putative surface protein. For instance, this is the case for the conserved hypothetical product P48-like encoded by MAGa1620 which displays a high similarity with the P48 lipoprotein, is detected by MS/MS in the triton detergent phase but does not contain a proper signal peptide followed by a lipobox. Careful examination of the 5' end of the P48-like coding sequence revealed a stretch of 10 Gs and a ribosomal frameshift at this position or the deletion of one G would generate an in-frame signal peptide followed by a lipobox (Figure 6). These data suggested that a mechanism based on polyG (or C) being prone to ribosomal shifting or to mutation could also account in several cases for the difference between the two strains in lipoproteins detected in Triton X114 by proteomic (see Supplementary Table S1).

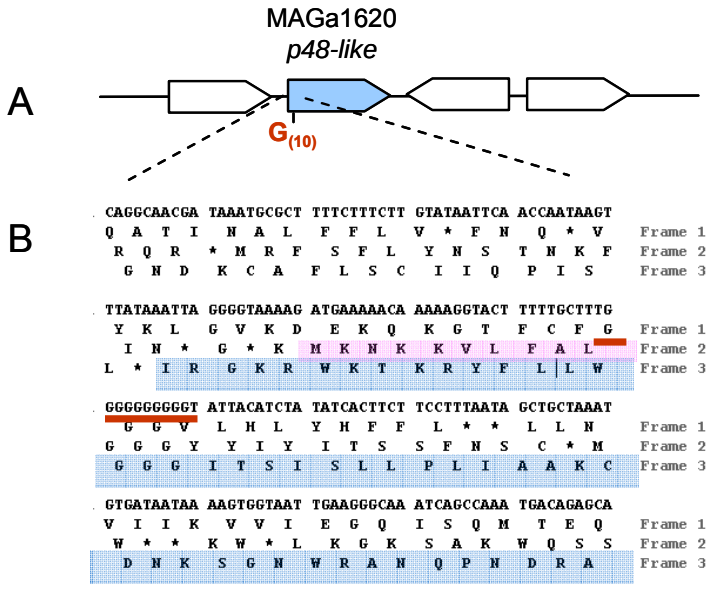


Figure 6: Analysis of the *p48-like* sequence of *M. agalactiae* 5632 suggests a mechanism for phase variation.

Schematic represents *p48-like* genomic region (A). CDSs are represented by large arrows, MAGa1620 corresponding to *p48-like* gene is filled in blue. Translation of the DNA region flanking the polyG track is given in the three frames (B). The polyG suspected to be variable in length (G10 +/-1) is underlined by a bold red bar. The putative beginning of a P48-like lipoprotein with an entire signal peptide sequence is shaded in red while the current annotated MAGa1620 open reading frame is in blue. Global amino-acid alignment results obtained with Needle (program available at http://www.ebi.ac.uk/Tools/emboss/align/) between P48-like of M. agalactiae 5632 and P68 lipoprotein of *M. bovis* PG45 for which a similar polyG was previously described [50], are of 89.3% (identity) and 92.1% (similarity).

4. *Discussion*

Whole genome sequencing of *M. agalactiae* strain 5632 revealed that it contains an additional 130 kbps when compared to the PG2 type strain. For organisms such as mycoplasmas that have a small genome size, this is rather significant. Further analyses indicated that this additional material is mostly made of repeated elements. When translated into gene content and new gene discovery, 5632 displayed 95 additional specific genes, many of which occur as almost perfect multiple copies so that our knowledge of the *M. agalactiae* pan-genome has been enriched by 39 new genes only. A large portion of these, 23, are encoded by the ICE element which is found in three entire copies or by IS. Recent mathematical models by Tettelin *et al.* [51] show that the pan-genome of the Mollicute *Ureaplasma urealyticum* is limited, based on the draft sequences of nine strains. This implies that the sequencing of additional strains might not significantly increase our knowledge on this species unless it is targeting a specific biological question [52]. Although *U. urealyticum* is a human pathogen and has a genome slightly smaller (ca. 750 kbp), the same may apply to the *M. agalactiae* species as indicated by the low number of new genes discovered in our study. Thus, sequencing additional *M. agalactiae* strains might bring little more information on the global coding capacity of this organism.

Overall, data obtained here and elsewhere indicate that about 10% of the 5632 genome is highly dynamic in that large regions corresponding to ICE can excise [15] and, theoretically, relocate elsewhere or be transferred to a recipient cell during conjugation, if occurring. The presence of two ICE's vestiges, ICE IV and V are like scars of past ICE insertions followed by a progressive decay. Interestingly, these more resemble the larger ICE vestige of PG2 or the ICE of *M. capricolum* subsp. *capricolum* than the three entire ICE copies of 5632 suggesting that this strain may have, at one point, hosted two types of ICEs. Whether they could co-exist is not known but these data indicate that the circulation of ICEs in some strains might not be such of a rare event. The presence of ICE circular form in 5632 [15] together with the low number of SNP existing in between the three

copies indicate that multiple ICE insertions are recent. The mechanisms underlying ICE insertion, excision and possibly transfer in mycoplasmas have yet to be investigated but recent studies on ICE elements in Gram-positive bacteria suggest that these events can be under the control of sophisticated regulation systems in response to changing environmental conditions such as stress or population density [53]. The finding of ICE in *M. agalactiae* and member of the "mycoides" cluster together with evidence of HGT in between these species further raised the prospect that these simple bacteria could conjugate. So far, a single report has supported the occurrence of conjugation in mycoplasmas by showing the exchange of genetic material in between *M. pulmonis* cells via a mechanism resistant to DNAse [54]. The idea that this phenomenon might be more common among mycoplasmas than first expected is very exciting because, if occurring, it would change the way we see the evolution of these so called “minimal organisms”.

Although smaller in size than ICE, IS elements as a whole represent a dynamic potential for the genome because of their copy number. In other bacteria, their contribution to genome plasticity and dynamics is well known [55]. Here, no major DNA inversion or rearrangement was detected in between the two *M. agalactiae* genomes that could be associated to IS except for two cases. As previously shown by our group, the first one refers to the duplication in 5632 of the single *vpma* cluster of PG2 that has been most likely driven by IS elements and that resulted in 5632 having extended possibilities for surface diversification when compared to PG2 [16]. The second case refers to a region (Figure 3 d) which organization significantly differs in between PG2 and 5632 and which contains several IS related elements, i.e. IS, transposases and pseudogenes of transposase reflecting the presence of past IS. Events underlying rearrangements in this region can not be exactly retraced but most likely they are ancient and have resulted in duplication of the *ptsG* gene in PG2. Interestingly, this region like many other associated with IS contains several genes or pseudogenes that have undergone HGT suggesting that IS may have a direct contribution in this phenomenon as suggested for other bacteria [55]. Finally, we showed that IS insertion may have an impact on gene expression thus modifying some of the strain properties, such as those associated with restriction-modification in 5632.

Compared to the PG2 type strain, 5632 seems better equipped for DNA exchange. Besides harbouring an impressive mobilome, some of which may be tailored for conjugative transfer, it contains a number of operating restriction-modification (RM) systems. In one hand, these may act as a barrier to DNA invasion [56] and explain why 5632 DNA is resistant to several methylase-sensitive restriction enzymes and to DNA transformation (data not shown). In the other hand, while methylated DNA is protected against degradation, it might be more likely to be accepted by a recipient cell displaying similar RM-systems. Indeed, some of the 5632 specific RM systems not present in PG2 have homologs in members of the "mycoides" cluster (Table 2). Whether the structure of the *M. agalactiae* population is made of a majority of PG2-like strains that are deficient in mobile elements as well as in RM systems with only some strains such as 5632 are being more prone to gene exchange with selected partners, is not yet known. Finally, DNA methylases, whether they belong or not to RM-systems, could play a number function in fitness or virulence such the regulation of numerous physiological processes including chromosome replication, mismatch repair, transposition, and transcription as described in other bacteria [57] or the epigenetic switch of some key factors such as in the Pap of the uropathogenic *E. coli* [58].

Interestingly, a fairly good portion of the flexible gene pool of *M. agalactiae* is dedicated to producing surface proteins, many of which are lipoproteins. Based on *in silico* analysis, 5632 contains ca. 100 lipoproteins with at least 56 expressed under laboratory conditions. Out of these putative and expressed lipoproteins, the Vpma family, which includes 16 different lipoproteins encoded by 23 genes in 5632 against 6 in PG2, are phase variable in expression and probably in size [16]. Phase variation of surface molecules is a common mechanism in mycoplasmas species [46] and is probably a major adaptive strategy for these minimal pathogens. Vpma phase variants are produced at high frequencies and in a reversible manner by site-specific recombination [25, 26] but comparative proteogenomics conducted here suggest that others variable systems may co-exist. For instance, expression of the P48-like protein as a lipoprotein that is soluble in triton X114 may depends on a riboshifting mechanism or on reversible hypermutation in a polyG tract localised at the 5' coding sequence (Figure 6). Indeed, data obtained by Lynyansky *et al.* [50] showed that translation of a full

length P68 lipoprotein in *M. bovis* is associated with the length of a similar polyG tract. The length of this homopolymer varies from 8 to 10 residues when comparing four *M. bovis*, with nine G allowing translation of a complete P68. Indeed, the P48-like and the P68 are two orthologs which expression is most likely phase variable in the two ruminant pathogens. Several other polyG tracts, some containing up to 13 residues, were found in the study that are associated with the 5' end of genes encoding surface lipoprotein suggesting that this may be a common slippage mechanism in *M. agalactiae*. Finally, the *drp* family involves genes that circulate by HGT in between *M. agalactiae* and members of the "mycoides" cluster. Based on comparative proteogenomics, 5632 and PG2 have a same size repertoire each composed of a different set with only 2 out of 12 to 13 *drp* products being expressed, one common to the two strains and one specific. Whether this reflects a mechanism of phase variation is unlikely, but the silent *drp* genes may act as reservoir sequences for the emergence of new Drp pattern of expression. Taken together these results suggest that the two *M. agalactiae* strains might display very different surface architectures which composition is highly dynamic during clonal propagation.

The strain 5632 was initially chosen because of its genetic particular features, several of which were found in its close relative *M. bovis*. This is further confirmed in this study which shows that 5632 unlike PG2 possesses (i) mobile elements such ICE and IS in multiple copies, (ii) a P48-like gene that is expressed, and (iii) two genes related to phage immunity that are also present in *M. bovis* PG45 [16]. The ovine/caprine pathogen, *M. agalactiae* and the cattle pathogen, *M. bovis,* were first classified under the same species and our findings indicate that a continuum of strains might be exist in between this two species. The genome sequence of *M. bovis* has been achieved (Craig Venter Institute, unpublished data and [44]) and its analysis may unravel even more common traits as well as some specificities that may explain the host-specificity of these two species.

References

1. Dobrindt U, Hacker J: **Whole genome plasticity in pathogenic bacteria**. *Curr Opin Microbiol* 2001, **4**(5):550-557.
2. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R: **The microbial pan-genome**. *Curr Opin Genet Dev* 2005, **15**(6):589-594.
3. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS *et al*: **Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome"**. *Proc Natl Acad Sci U S A* 2005, **102**(39):13950-13955.
4. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM *et al*: **The minimal gene complement of *Mycoplasma genitalium***. *Science* 1995, **270**(5235):397-403.
5. Sirand-Pugnet P, Citti C, Barre A, Blanchard A: **Evolution of mollicutes: down a bumpy road with twists and turns**. *Res Microbiol* 2007, **158**(10):754-766.
6. Dybvig K, Zuhua C, Lao P, Jordan DS, French CT, Tu AH, Loraine AE: **Genome of *Mycoplasma arthritidis***. *Infect Immun* 2008, **76**(9):4000-4008.
7. Calderon-Copete SP, Wigger G, Wunderlin C, Schmidheini T, Frey J, Quail MA, Falquet L: **The *Mycoplasma conjunctivae* genome sequencing, annotation and analysis**. *BMC Bioinformatics* 2009, **10 Suppl 6**:S7.
8. Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, Mahairas GG: **The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis**. *J Bacteriol* 2004, **186**(21):7123-7133.
9. Vasconcelos AT, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, Almeida DF, Almeida LG, Almeida R, Alves-Filho L *et al*: **Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae***. *J Bacteriol* 2005, **187**(16):5568-5577.
10. Rosengarten R, Citti C, Much P, Spergser J, Droesse M, Hewicker-Trautwein M: **The changing image of mycoplasmas: from innocent bystanders to emerging and reemerging pathogens in human and animal diseases**. *Contrib Microbiol* 2001, **8**:166-185.
11. Sirand-Pugnet P, Lartigue C, Marenda M, Jacob D, Barre A, Barbe V, Schenowitz C, Mangenot S, Couloux A, Segurens B *et al*: **Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome**. *PLoS Genet* 2007, **3**(5):e75.
12. Solsona M, Lambert M, Poumarat F: **Genomic, protein homogeneity and antigenic variability of *Mycoplasma agalactiae***. *Vet Microbiol* 1996, **50**(1-2):45-58.
13. Marenda MS, Sagne E, Poumarat F, Citti C: **Suppression subtractive hybridization as a basis to assess *Mycoplasma agalactiae* and *Mycoplasma bovis* genomic diversity and species-specific sequences**. *Microbiology* 2005, **151**(Pt 2):475-489.

14. McAuliffe L, Churchward CP, Lawes JR, Loria G, Ayling RD, Nicholas RA: **VNTR analysis reveals unexpected genetic diversity within *Mycoplasma agalactiae*, the main causative agent of contagious agalactia**. *BMC Microbiol* 2008, **8**:193.
15. Marenda M, Barbe V, Gourgues G, Mangenot S, Sagne E, Citti C: **A new integrative conjugative element occurs in *Mycoplasma agalactiae* as chromosomal and free circular forms**. *J Bacteriol* 2006, **188**(11):4137-4141.
16. Nouvel LX, Marenda M, Sirand-Pugnet P, Sagne E, Glew M, Mangenot S, Barbe V, Barre A, Claverol S, Citti C: **Occurrence, plasticity, and evolution of the *vpma* gene family, a genetic system devoted to high-frequency surface variation in *Mycoplasma agalactiae***. *J Bacteriol* 2009, **191**(13):4111-4121.
17. Marenda MS, Vilei EM, Poumarat F, Frey J, Berthelot X: **Validation of the suppressive subtractive hybridization method in *Mycoplasma agalactiae* species by the comparison of a field strain with the type strain PG2**. *Vet Res* 2004, **35**(2):199-212.
18. Frey J: **Mycoplasmas of Animals**. In: *Molecular biology and pathogenicity of Mycoplasmas.* Edited by Razin S, Herrmann R. New York (USA): Kluwer Academic / Plenum Publishers; 2002: 73–90.
19. Corrales JC, Esnal A, De la Fe C, Sanchez A, Assuncao P, Poveda JB, Contreras A: **Contagious agalactia in small ruminants**. *Small Ruminant Res* 2007, **68**(1-2):154-166.
20. Tola S, Crobeddu S, Chessa G, Uzzau S, Idini G, Ibba B, Rocca S: **Sequence, cloning, expression and characterisation of the 81-kDa surface membrane protein (P80) of *Mycoplasma agalactiae***. *FEMS Microbiol Lett* 2001, **202**(1):45-50.
21. Fleury B, Bergonier D, Berthelot X, Peterhans E, Frey J, Vilei EM: **Characterization of P40, a cytadhesin of *Mycoplasma agalactiae***. *Infect Immun* 2002, **70**(10):5612-5621.
22. Rosati S, Robino P, Fadda M, Pozzi S, Mannelli A, Pittau M: **Expression and antigenic characterization of recombinant *Mycoplasma agalactiae* P48 major surface protein**. *Vet Microbiol* 2000, **71**(3-4):201-210.
23. Fleury B, Bergonier D, Berthelot X, Schlatter Y, Frey J, Vilei EM: **Characterization and analysis of a stable serotype-associated membrane protein (P30) of *Mycoplasma agalactiae***. *J Clin Microbiol* 2001, **39**(8):2814-2822.
24. Glew MD, Papazisi L, Poumarat F, Bergonier D, Rosengarten R, Citti C: **Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae***. *Infect Immun* 2000, **68**(8):4539-4548.
25. Chopra-Dewasthaly R, Citti C, Glew MD, Zimmermann M, Rosengarten R, Jechlinger W: **Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation**. *Mol Microbiol* 2008, **67**(6):1196-1210.
26. Glew MD, Marenda M, Rosengarten R, Citti C: **Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus**. *J Bacteriol* 2002, **184**(21):5987-5998.

27. Aluotto BB, Wittler RG, Williams CO, Faber JE: **Standardized bacteriologic techniques for the characterization of mycoplasma species.** *Int J Syst Bacteriol* 1970, **20**:35-58.
28. Tully JG: **Culture medium formulation for primary isolation and maintenance of mollicutes.** In: *Molecular and Diagnostic Procedures in Mycoplasmology: Molecular Characterization.* Edited by Tully JG. San Diego: Academic Press.; 1995: 33–39.
29. Sambrook J, Fritsch EF, Maniatis T: **Molecular cloning: a laboratory manual**, 2nd ed. edn. N.Y.: Cold Spring Harbor Laboratory Press; 1989.
30. Chen WP, Kuo TT: **A simple and rapid method for the preparation of gram-negative bacterial genomic DNA**. *Nucleic Acids Res* 1993, **21**(9):2260.
31. Frangeul L, Glaser P, Rusniok C, Buchrieser C, Duchaud E, Dehoux P, Kunst F: **CAAT-Box, Contigs-Assembly and Annotation Tool-Box for genome sequencing projects**. *Bioinformatics* 2004, **20**(5):790-797.
32. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence**. *Nucleic Acids Res* 1997, **25**(5):955-964.
33. Laslett D, Canback B: **ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences**. *Nucleic Acids Res* 2004, **32**(1):11-16.
34. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0**. *Mol Biol Evol* 2007, **24**(8):1596-1599.
35. Barre A, de Daruvar A, Blanchard A: **MolliGen, a database dedicated to the comparative genomics of *Mollicutes***. *Nucleic Acids Res* 2004, **32**(Database issue):D307-310.
36. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: **Artemis: sequence visualization and annotation**. *Bioinformatics* 2000, **16**(10):944-945.
37. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J: **ACT: the Artemis Comparison Tool**. *Bioinformatics* 2005, **21**(16):3422-3423.
38. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: **VISTA: computational tools for comparative genomics**. *Nucleic Acids Res* 2004, **32**(Web Server issue):W273-279.
39. Bordier C: **Phase separation of integral membrane proteins in Triton X-114 solution**. *J Biol Chem* 1981, **256**(4):1604-1607.
40. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J: **DNAPlotter: circular and linear interactive genome visualization**. *Bioinformatics* 2009, **25**(1):119-120.
41. Calcutt MJ, Lewis MS, Wise KS: **Molecular genetic analysis of ICEF, an integrative conjugal element that is present as a repetitive sequence in the chromosome of *Mycoplasma fermentans* PG18**. *J Bacteriol* 2002, **184**(24):6929-6941.
42. Pilo P, Fleury B, Marenda M, Frey J, Vilei EM: **Prevalence and distribution of the insertion element ISMag1 in *Mycoplasma agalactiae***. *Vet Microbiol* 2003, **92**(1-2):37-48.
43. Thomas A, Linden A, Mainil J, Bischof DF, Frey J, Vilei EM: ***Mycoplasma bovis* shares insertion sequences with *Mycoplasma agalactiae* and**

***Mycoplasma mycoides* subsp. *mycoides* SC: Evolutionary and developmental aspects**. *FEMS Microbiol Lett* 2005, **245**(2):249-255.
44. Lysnyansky I, Calcutt MJ, Ben-Barak I, Ron Y, Levisohn S, Methe BA, Yogev D: **Molecular characterization of newly identified IS3, IS4 and IS30 insertion sequence-like elements in *Mycoplasma bovis* and their possible roles in genome plasticity**. *FEMS Microbiol Lett* 2009, **294**(2):172-182.
45. Baranowski E, Guiral S, Sagne E, Skapski A, Brouillaud A, Citti C: **Conditional essentiality in a minimal genome: the NIF locus is essential for *Mycoplasma agalactiae* survival in cell culture, while dispensable for axenic growth**. 2009: *Manuscript in preparation*.
46. Citti C, Browning GF, Rosengarten R: **Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas.** In: *Mycoplasmas: Molecular biology, pathogenicity and strategies for control.* Edited by Blanchard A, Browning GF. Norfolk (United Kingdom): Horizon Bioscience; 2005: 439–484.
47. Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, Zwahlen MC, Desiere F, Bork P, Delley M *et al*: **The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract**. *Proc Natl Acad Sci U S A* 2002, **99**(22):14422-14427.
48. Sitaraman R, Denison AM, Dybvig K: **A unique, bifunctional site-specific DNA recombinase from *Mycoplasma pulmonis***. *Mol Microbiol* 2002, **46**(4):1033-1040.
49. Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA, 3rd, Smith HO, Venter JC: **Genome transplantation in bacteria: changing one species to another**. *Science* 2007, **317**(5838):632-638.
50. Lysnyansky I, Yogev D, Levisohn S: **Molecular characterization of the *Mycoplasma bovis p68* gene, encoding a basic membrane protein with homology to P48 of *Mycoplasma agalactiae***. *FEMS Microbiol Lett* 2008, **279**(2):234-242.
51. Tettelin H, Riley D, Cattuto C, Medini D: **Comparative genomics: the bacterial pan-genome**. *Curr Opin Microbiol* 2008, **11**(5):472-477.
52. Whitworth DE: **Genomes and knowledge - a questionable relationship?** *Trends Microbiol* 2008, **16**(11):512-519.
53. Churchward G: **Back to the future: the new ICE age**. *Mol Microbiol* 2008, **70**(3):554-556.
54. Teachman AM, French CT, Yu H, Simmons WL, Dybvig K: **Gene transfer in *Mycoplasma pulmonis***. *J Bacteriol* 2002, **184**(4):947-951.
55. Mahillon J, Chandler M: **Insertion sequences**. *Microbiol Mol Biol Rev* 1998, **62**(3):725-774.
56. Thomas CM, Nielsen KM: **Mechanisms of, and barriers to, horizontal gene transfer between bacteria**. *Nat Rev Microbiol* 2005, **3**(9):711-721.
57. Heusipp G, Falker S, Schmidt MA: **DNA adenine methylation and bacterial pathogenesis**. *Int J Med Microbiol* 2007, **297**(1):1-7.
58. Hernday AD, Braaten BA, Low DA: **The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch**. *Mol Cell* 2003, **12**(4):947-957.

Supplementary data

Table S1: Products of 5632 for which more than one specific peptide were detected by LC MS/MS after 1D SDS-PAGE.

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa0010	Chromosomal replication initiator protein DnaA	+	+
MAGa0020	DNApolymeraseIII,betachain	+	+
MAGa0050	Esterase/lipase	+	-
MAGa0060	Esterase/lipase	-	+
MAGa0070	NADH dependent flavin oxidoreductase	-	-
MAGa0080	Lipoate-proteinligaseA	+	-
MAGa0090	Lipoate-proteinligaseA	+	+
MAGa0100	Conserved hypothetical protein	+	+
MAGa0110	GlycinecleavagesystemHprotein	+	+
MAGa0120	Conserved hypothetical protein	-	-
MAGa0130	Hypothetical protein	+	-
MAGa0140	Conserved hypothetical protein, predictedlipoprotein, P48	+	+
MAGa0150	Hypothetical protein	-	-
MAGa0160	ABC transporter, ATP-binding protein, P59	+	+
MAGa0170	Sugar ABC transporter Permease	+	+
MAGa0180	Sugar ABC transporter permease protein	-	-
MAGa0190	Deoxyguanosine kinase	+	+
MAGa0200	Deoxyguanosine kinase	+	-
MAGa0210	Conserved hypothetical protein, predictedlipoprotein	-	-
MAGa0220	Hypothetical protein, predicted lipoprotein,DUF285 family	-	-
MAGa0230	Conserved hypothetical protein, DUF285 family	-	-
MAGa0240	Transposase	-	-
MAGa0250	Hypothetical protein, predicted lipoprotein,DUF285 family	-	-
MAGa0260	ConservedhypotheticalproteinYeiN	+	+
MAGa0270	ConservedhypotheticalproteinYeiC	+	-
MAGa0280	Hydrolase of the HAD superfamily	+	-
MAGa0290	Conserved hypothetical protein	+	+
MAGa0300	Hypothetical ABC transporter, ATP-bindingprotein	+	-
MAGa0310	Hypothetical protein	-	-
MAGa0320	Hypothetical protein	-	-
MAGa0330	Hypothetical protein	-	-
MAGa0340	Oligopeptide ABC transporter, ATP-bindingprotein (OppF)	+	+
MAGa0350	Oligopeptide ABC transporter, ATP-bindingprotein (OppD)	+	+
MAGa0360	Oligopeptide ABC transporter, permease protein(OppC)	+	-
MAGa0370	Oligopeptide ABC transporter, permeaseprotein(OppB)	+	-
MAGa0380	Oligopeptide ABC transporter, substrate-bindingprotein (OppA), predicted lipoprotein	+	+
MAGa0390	Conserved hypothetical protein	+	+
MAGa0400	Cysteinyl-tRNAsynthetase(Cysteine-tRNAligase)	+	-
MAGa0410	tRNA/rRNA methyltransferase	-	-
MAGa0420	Hypothetical protein	-	-
MAGa0430	50S ribosomal protein L33	-	-
MAGa0440	Hypothetical protein	-	-
MAGa0450	Transcriptionantiterminationprotein(NusG)	+	+
MAGa0460	Conserved hypothetical protein	-	-
MAGa0470	Uridylate kinase	+	+
MAGa0480	Ribosomerecyclingfactor(Ribosomereleasingfactor)	+	+
MAGa0510	Conserved hypothetical protein	-	-
MAGa0520	Conserved hypothetical protein	-	-
MAGa0530	Glycerol-3-phosphate dehydrogenase [NAD(P)+]	+	-
MAGa0540	Dimethyladenosinetransferase (S-adenosylmethionine-6-N',N'-adenosyl(rRNA)dimethyltransferase) (16SrRNAdimethylase) (High level kasugamycin resistance protein ksgA) (Kasugamycindimethyltransferase)	+	-
MAGa0550	Conserved hypothetical protein, Putativedeoxyribonuclease (YabD) Deoxyribonuclease (TatD)	-	+
MAGa0560	TRNAmodificationGTPase(TrmE)	+	+
MAGa0570	Hypothetical protein	-	-
MAGa0580	Glyceraldehyde 3-phosphate dehydrogenase(GAPDH)	+	+
MAGa0590	Seryl-tRNAsynthetase	+	+
MAGa0600	EpsG	+	-
MAGa0610	50S ribosomal protein L34	-	-
MAGa0620	Ribonuclease P protein component	-	-

Table S1: continued (2/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa0630	Conserved hypothetical protein	+	+
MAGa0640	Lipoate-proteinligaseA	+	+
MAGa0650	Esterase/lipase	+	+
MAGa0660	Hypothetical protein	-	-
MAGa0670	Aspartate-ammonialigase(AsnA)	+	-
MAGa0680	DNA polymerase III alpha subunit	+	-
MAGa0710	ATPase,AAAfamily	+	-
MAGa0720	Phenylalanyl-tRNA Synthetase alpha chain (PheS)	+	-
MAGa0730	Uracil-DNA glycosylase (Ung)	-	-
MAGa0740	Phenylalanyl-tRNA synthetase beta chain (PheT)	+	+
MAGa0750	Serinehydroxymethyltransferase(Serinemethylase)(SHMT)	+	+
MAGa0760	Proline iminopeptidase (Pip)	-	-
MAGa0770	NitrogenfixationproteinNifS	+	-
MAGa0780	Nitrogen fixation protein NifU	-	-
MAGa0790	DNA-damagerepairproteinMucB	+	-
MAGa0800	bidomainalprotein	+	+
MAGa0810	RibosomallargesubunitpseudouridinesynthaseB	+	+
MAGa0820	Hypothetical protein, predicted lipoprotein	-	-
MAGa0830	Hypotheticalprotein	+	-
MAGa0840	Fructose-bisphosphatealdolase(Fba)	+	+
MAGa0850	Nuclease	-	-
MAGa0860	50S ribosomal protein L11	+	+
MAGa0870	50S ribosomal protein L1	+	+
MAGa0880	Hypothetical protein	-	-
MAGa0890	Hypothetical protein	+	-
MAGa0900	Hypothetical protein	-	-
MAGa0910	tRNA/rRNAmethyltransferase	+	+
MAGa0920	rRNA methylase	-	+
MAGa0930	Hypothetical protein	-	-
MAGa0940	GTP-binding protein	+	-
MAGa0950	Hypothetical protein	-	-
MAGa0960	HPrkinase/phosphorylase	+	+
MAGa0970	Prolipoprotein diacylglyceryl transferase	+	+
MAGa0980	Thioredoxinreductase	+	+
MAGa0990	Hypothetical protein	+	-
MAGa1000	Pyruvate dehydrogenase E1 component, alphasubunit	+	+
MAGa1010	Pyruvate dehydrogenase E1 component, betasubunit	+	+
MAGa1020	Hypotheticalprotein	+	-
MAGa1030	Dihydrolipoamide acetyltransferase component ofpyruvate deshydrogenase complex	+	+
MAGa1040	Dihydrolipoamidedehydrogenase(E3componentofpyruvatecomplex)	+	+
MAGa1050	50S ribosomal protein L28	+	+
MAGa1060	Conserved hypothetical protein	-	-
MAGa1070	Conserved hypothetical protein	+	+
MAGa1080	Transposase	-	-
MAGa1090	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa1100	Oligopeptide ABC transporter, permease protein(OppB)	+	+
MAGa1110	Oligopeptide ABC transporter system, permeaseprotein (OppC)	+	+
MAGa1120	OligopeptideABCtransporter,ATP-bindingprotein(OppD)	+	+
MAGa1130	Oligopeptide ABC transporter, ATP-bindingprotein (OppF)	+	+
MAGa1140	Hypothetical protein, predicted lipoprotein	+	+
MAGa1150	Hypothetical protein	-	-
MAGa1160	Conservedhypotheticalprotein	+	-
MAGa1170	Hypothetical protein, predicted lipoprotein	-	-
MAGa1180	Ribose-5-phosphateisomerase	+	+
MAGa1190	Hypothetical protein	-	-
MAGa1200	Conserved hypothetical protein	-	-
MAGa1210	Conserved hypothetical protein	-	-
MAGa1220	Cytidine deaminase (Cytidine aminohydrolase)	-	+
MAGa1230	GTP-binding protein era	-	-
MAGa1240	XAA-PRO aminopeptidase	+	+
MAGa1250	Prolyl-tRNA synthetase (Proline-tRNA ligase)	+	+
MAGa1260	Hypothetical protein	-	-

Table S1: continued (3/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa1270	Hypotheticalprotein	+	+
MAGa1280	Conserved hypothetical protein	+	+
MAGa1290	Phosphoketolase	+	+
MAGa1300	Oligoendopeptidase F	+	+
MAGa1310	Spermidine/putrescine ABC transporterATP-binding protein PotA	+	+
MAGa1320	Spermidine/putrescine ABC transporter permeaseprotein PotB	-	-
MAGa1330	Spermidine/putrescine ABC transporter permeaseprotein PotC	-	+
MAGa1340	Conservedhypotheticalprotein	+	+
MAGa1350	50S ribosomal protein L32	-	-
MAGa1360	Conserved hypothetical protein	+	-
MAGa1370	ThiaminebiosynthesisproteinthiI	+	-
MAGa1380	Hypothetical protein	-	+
MAGa1390	Conserved hypothetical protein, DUF285 family	-	-
MAGa1400	Conservedhypotheticalprotein	+	-
MAGa1410	Valyl-tRNAsynthetase(Valine-tRNAligase)(VALRS)	+	-
MAGa1420	FolDbifunctionalprotein(MethylenetetrahydrofolateDehydrogenase;Methenyltetrahydrofolatecyclohydrolase)	+	+
MAGa1430	Phosphateacetyltransferase(Phosphotransacetylase)	+	+
MAGa1440	Acetatekinase(Acetokinase)	+	+
MAGa1450	Phosphopantetheine adenylyltransferase(Pantetheine-phosphate adenylyltransferase) (PPAT)(Dephospho-CoA pyrophosphorylase)	-	-
MAGa1460	GTP-binding protein engB	-	-
MAGa1470	Hypothetical protein	+	+
MAGa1480	Pyruvate kinase (PK)	+	+
MAGa1490	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa1500	Chaperone protein dnaK	+	+
MAGa1510	Glucose-inhibited division protein	-	-
MAGa1520		-	-
MAGa1530	D-lactate dehydrogenase	+	+
MAGa1540	Esterase/lipase	-	-
MAGa1550	Hypothetical protein, predicted lipoprotein	+	+
MAGa1560	Conservedhypotheticalprotein	+	+
MAGa1570	TypeIIIrestriction-modificationsystem:Methylase	+	-
MAGa1580	TypeIIIrestriction-modificationsystem:Methylase	+	+
MAGa1590	IS30-like protein	-	-
MAGa1600	Triggerfactor	+	+
MAGa1610	Hypothetical protein, predicted lipoprotein	-	+
MAGa1620	Conserved hypothetical protein, P48-like	+	-
MAGa1630	Hypothetical protein, predicted lipoprotein	-	-
MAGa1640	ABC transporter, ATP binding protein	-	+
MAGa1650	Conserved hypothetical protein, predictedmembrane protein	-	-
MAGa1660	DNA polymerase I: 5'-3' exonuclease domain	+	-
MAGa1670	DNApolymeraseIII,alphasubunit	+	-
MAGa1680	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa1690	Hypothetical protein	-	-
MAGa1700	Hypothetical protein	-	-
MAGa1710	Hypothetical protein	-	-
MAGa1720	Hypothetical protein	-	-
MAGa1730	GlucoseinhibiteddivisionproteinA	+	+
MAGa1740	Conserved hypothetical protein	-	-
MAGa1750	Conserved hypothetical protein, predictedtransmembrane product	-	-
MAGa1760	Hypothetical protein	+	+
MAGa1770	DNA methylase	-	-
MAGa1780	Formylmethionine deformylase	+	+
MAGa1790	Hypothetical protein	+	+
MAGa1800	Topoisomerase IV subunit B	-	-
MAGa1810	TopoisomeraseIVsubunitA	+	+
MAGa1820	Conserved hypothetical protein	+	+
MAGa1830	Ribonuclease HII (RNase HII)	-	-
MAGa1850	Hypothetical protein	-	-
MAGa1870	Methyltransferase GidB (Glucose inhibiteddivision protein B)	-	-
MAGa1880	Ribose-phosphatepyrophosphokinase(Phosphoribosylpyrophosphatesynthetase)	+	+

Table S1: continued (4/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa1890	Hypothetical protein, predicted lipoprotein	-	-
MAGa1900		-	+
MAGa1910	Hypothetical protein, predicted lipoprotein	-	-
MAGa1920	Ribonuclease	+	-
MAGa1930	TransketolaseI	+	+
MAGa1940	Inorganicpyrophosphatase(Pyrophosphatephospho-hydrolase)	+	+
MAGa1950	30SribosomalproteinS20	+	+
MAGa1960	Hypothetical protein	-	+
MAGa1970	Conserved hypothetical protein	+	+
MAGa1980	Hypothetical protein, predicted lipoprotein	+	+
MAGa1990	HIT-likeprotein(Cellcycleregulation)	+	+
MAGa2000	Hypothetical protein, predicted lipoprotein	+	+
MAGa2010	50S ribosomal protein L33	-	-
MAGa2020	Thiol peroxidase	+	-
MAGa2030	Conserved hypothetical protein	-	-
MAGa2040	CelldivisionproteinFtsY	+	+
MAGa2050	Hypotheticalprotein	+	+
MAGa2060	Methionyl-tRNA synthetase (Methionine-tRNAligase)	+	+
MAGa2070	DNAmethylase	+	-
MAGa2080	RibonucleaseR(VacB-like(Shigellaflexneri)ribonucleaseII)	+	+
MAGa2090	Hypothetical protein	-	-
MAGa2100	Guanylatekinase	+	+
MAGa2110	Protein phosphatase	-	-
MAGa2120	Serine/threonine-protein kinase	+	+
MAGa2130	GTPase EngC	-	-
MAGa2140	Ribulose-phosphate3-epimerase(Pentose-5-phosphate3-epimerase)	+	+
MAGa2150	Pseudogene of Exodeoxyribonuclease V alphachain(RecD) (C-terminal part)	-	-
MAGa2160	Transposase	-	-
MAGa2170	Pseudogene of Exodeoxyribonuclease V alphachain(RecD) (C-terminal part)	-	-
MAGa2180	Pseudogene of Exodeoxyribonuclease V alphachain(RecD) (N-terminal part)	-	-
MAGa2190	Conservedhypotheticalprotein	+	-
MAGa2200	AAAfamilyATPase	+	-
MAGa2210	Conservedhypotheticalprotein	+	-
MAGa2300	CTPsynthase(UTP-ammonialigase)	+	+
MAGa2310	Isoleucyl-tRNAsynthetase(Isoleucine-tRNAligase)	+	+
MAGa2320	Lipoprotein signal peptidase (SPASE II)	+	+
MAGa2330	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa2390	Holliday junction DNA helicase ruvA	-	-
MAGa2400	Holliday junction DNA helicase ruvB	-	-
MAGa2410	Protein-export membrane protein	+	+
MAGa2420	tRNA pseudouridine synthase B	-	+
MAGa2430	Conserved hypothetical protein	+	+
MAGa2440	Riboflavin biosynthesis protein	-	-
MAGa2450	30S ribosomal protein S15	-	+
MAGa2460	Esterase/lipase	+	+
MAGa2470	Glycerol ABC transporter, ATP-bindingcomponent	+	-
MAGa2480	Glycerol ABC transporter, permease component	-	-
MAGa2490	Glycerol ABC transporter, permease component	-	-
MAGa2500	Conserved hypothetical protein, predictedlipoprotein	+	-
MAGa2510	Hypothetical protein, predicted lipoprotein	+	+
MAGa2520	Conserved hypothetical protein	+	+
MAGa2530	50SribosomalproteinL9	+	+
MAGa2540	ReplicativeDNAhelicase	+	+
MAGa2550	Hemolysin-related protein	+	-
MAGa2570	Hypothetical protein, predicted lipoprotein	+	+
MAGa2580	P40, predicted lipoprotein	+	+
MAGa2590	Hypothetical protein, predicted lipoprotein	-	-
MAGa2600	Conserved hypothetical protein, predictedlipoprotein, DUF285 family	+	+
MAGa2610	Hypotheticalprotein	+	-
MAGa2620	Hypothetical protein, predicted lipoprotein	-	-
MAGa2630	NADH oxidase (NOXASE)	-	+
MAGa2640	Proton/glutamate symporter	+	+
MAGa2650	DNA recombination protein	+	+

Table S1: continued (5/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa2660	Hypothetical protein	-	-
MAGa2670	Hypothetical protein, predicted lipoprotein	+	+
MAGa2680	Hypothetical protein, predicted lipoprotein	-	-
MAGa2690	Hypothetical protein, Vpma-like, predictedlipoprotein	+	-
MAGa2700	Adenine-specific DNA methyltransferase	-	-
MAGa2710	Type II restriction endonuclease	-	-
MAGa2720	30S ribosomal protein S2	+	+
MAGa2730	Elongation factor Ts (EF-Ts)	+	+
MAGa2740	Hypothetical protein, predicted lipoprotein	+	+
MAGa2750	Conserved hypothetical protein	-	-
MAGa2760	NADH oxidase (NOXASE)	+	+
MAGa2770	Conservedhypotheticalprotein	+	-
MAGa2780	DNA-directed RNA polymerase sigma factor(Sigma-A)	+	+
MAGa2790	DNAprimase	+	-
MAGa2800	Glycyl-tRNAsynthetase(Glycine-tRNAligase)	+	+
MAGa2810	Hypothetical protein	+	+
MAGa2820	Alkylphosphonate ABC transporter,substrate-binding protein, predicted lipoprotein	+	+
MAGa2830	ABC transporter, ATP-binding protein	-	-
MAGa2840	Alkylphosphonate ABC transporter, permeaseprotein	-	-
MAGa2850	Conservedhypotheticalprotein	+	+
MAGa2860	Preprotein translocase SecA subunit	+	+
MAGa2870	Alcohol dehydrogenase	+	+
MAGa2880	DNA topoisomerase I	+	+
MAGa2890	30S ribosomal protein S6	+	+
MAGa2900	Single-stranded DNA-binding protein	-	+
MAGa2910	30S ribosomal protein S18	+	+
MAGa2920	Conserved hypothetical protein	-	-
MAGa2930	Phosphopentomutase	+	+
MAGa2940	Hypotheticalprotein	+	+
MAGa2950	DNAligase	+	+
MAGa2960	Putative transmembrane protein	-	-
MAGa2970	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa2980	CDS1	-	-
MAGa2990	CDSA	-	-
MAGa3000	CDS12	-	-
MAGa3010		-	-
MAGa3020		-	-
MAGa3030	CDSB	-	-
MAGa3040	CDSC	-	-
MAGa3050	CDSD	-	-
MAGa3060	CDS5	-	-
MAGa3070	CDS7	-	-
MAGa3080	CDS13	-	-
MAGa3090	CDS15	-	-
MAGa3100	CDS16 N-terminal part	-	-
MAGa3110	C-treminal part of CDS16 Previouslycharacterized CDS (tr-Q1IH56); Marenda M., Barbe V.,Gourgues G., Mangenot S., Sagne E., Citti C., A newintegrative conjugative element occurs in Mycoplasmaagalactiae as chromosomal and free circular forms. J.Bacteriol. 188:4137-4141(2006).)	-	-
MAGa3120	Hypothetical protein	-	-
MAGa3130	CDS17	+	-
MAGa3140	CDS19	-	-
MAGa3150	CDSE	-	-
MAGa3160	CDS14	+	-
MAGa3170	CDSF	-	-
MAGa3180	Hypotetical protein	+	-
MAGa3190	CDSG	-	-
MAGa3200	CDSH	-	-
MAGa3210		-	-
MAGa3220	CDS22	-	-
MAGa3230	Conserved hypothetical protein, predictedtransmembrane protein	+	-
MAGa3240	Putative transmembrane protein	-	-

Table S1: continued (6/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa3250	Conserved hypothetical protein, predictedlipoprotein	+	-
MAGa3260	Conserved hypothetical protein, predictedtransmembrane protein	+	-
MAGa3270	Conservedhypotheticalprotein	+	+
MAGa3280	Conserved hypothetical protein	+	-
MAGa3290	Conservedhypotheticalprotein	+	-
MAGa3300	Putative transmembrane protein	+	+
MAGa3310	ATP synthase alpha chain	+	+
MAGa3320	ATPsynthasebetachain	+	+
MAGa3330	Hypothetical protein, predicted lipoprotein	-	+
MAGa3340	Hypothetical protein	-	-
MAGa3350	Conserved hypothetical protein, predictedlipoprotein	-	-
MAGa3360	Conserved hypothetical protein	+	+
MAGa3370	Conserved hypothetical protein	+	-
MAGa3380	ATP-dependenthelicase	+	-
MAGa3390	Conserved hypothetical protein	-	-
MAGa3400	Glycosyltransferase	+	-
MAGa3410	Asparaginyl-tRNA synthetase	+	+
MAGa3420	Hypothetical protein	-	-
MAGa3430	Hypothetical protein	-	-
MAGa3440	Conserved hypothetical protein	+	+
MAGa3450	Mg2+ transport protein (MGTE)	+	+
MAGa3460	Hypothetical protein	+	-
MAGa3470	ElongationfactorP(EF-P)	+	+
MAGa3480	Hypothetical protein	+	+
MAGa3500	ABC transporter ATP-binding protein	+	+
MAGa3510	ABC transporter ATP-binding protein	+	+
MAGa3520	Conserved hypothetical protein	+	-
MAGa3530	Peptide methionine sulfoxide reductase	+	+
MAGa3540	Hypothetical protein	-	-
MAGa3550	Foramidopyrimidine DNA glycosylase	-	-
MAGa3560	Glucose-6-phosphate isomerase	+	-
MAGa3570	Glucose-6-phosphate isomerase	-	-
MAGa3580	30S ribosomal protein S1	-	-
MAGa3590	Enolase	+	+
MAGa3600	Elongation factor Tu (EF-Tu)	+	+
MAGa3610	Conserved hypothetical protein	-	-
MAGa3620	Conservedhypotheticalprotein	+	-
MAGa3630	O-sialoglycoproteinendopeptidase(Glycoprotease)	+	-
MAGa3640	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa3650	PtsG	+	+
MAGa3660	Conserved hypothetical protein	-	-
MAGa3670		-	-
MAGa3680	Hypothetical protein	-	-
MAGa3690	Conserved hypothetical protein, truncated inC-terminal	-	-
MAGa3700	Transposase	-	-
MAGa3710		-	-
MAGa3720	Hypothetical protein, truncated in N-terminal	-	-
MAGa3730	Hypothetical protein, predicted lipoprotein	-	-
MAGa3740	Hypothetical protein	-	-
MAGa3750	Pseudogene of Transposase	-	-
MAGa3760	Conserved hypothetical protein, DUF285 family,truncated in terminal	-	-
MAGa3770	Conserved hypothetical protein, DUF285 family	-	-
MAGa3780	Conserved hypothetical protein	+	-
MAGa3790	Tryptophanyl-tRNA synthetase	-	-
MAGa3800	Threonyl-tRNA synthetase	+	+
MAGa3810	Hypothetical protein	-	-
MAGa3820	Hypothetical protein, predicted lipoprotein	+	-
MAGa3830	P30, predicted lipoprotein	-	+
MAGa3840	Hypotheticalprotein	+	+
MAGa3850	Conserved hypothetical protein	-	-
MAGa3860	ATP synthase A chain	-	-
MAGa3870	ATP synthase C chain	-	-
MAGa3880	ATP synthase B chain	-	+

Table S1: continued (7/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa3890	ATP synthase delta subunit	+	+
MAGa3900	ATP synthase alpha chain	+	+
MAGa3910	ATP synthase gamma chain	-	-
MAGa3920	ATP synthase beta chain	+	+
MAGa3930	ATP synthase epsilon chain	-	-
MAGa3940	Hypothetical protein	+	-
MAGa3950	Cytosine-specificmethyltransferase	+	-
MAGa3960	Transposase	-	-
MAGa3970	Type II site-specific deoxyribonuclease,sau96I-like	-	-
MAGa3980	Hypothetical protein, predicted lipoprotein	-	+
MAGa3990	Hypothetical protein, predicted lipoprotein	+	+
MAGa4000	Conserved hypothetical protein	-	-
MAGa4010	Hypotheticalprotein	+	-
MAGa4020	Hypothetical protein	-	-
MAGa4030	Transposase	-	-
MAGa4040	Conserved hypothetical protein, truncated inC-terminal	-	-
MAGa4050	Hypothetical protein	-	-
MAGa4060	Hypothetical protein	-	-
MAGa4070	Hypothetical protein	-	-
MAGa4080	Conserved hypothetical protein	-	-
MAGa4090	HeatshockATP-dependentprotease	+	+
MAGa4100	Leucyl-tRNAsynthetase(Leucine-tRNAligase)	+	+
MAGa4110	Conserved hypothetical protein	-	-
MAGa4120	Uracilphosphoribosyltransferase	+	+
MAGa4130	Single-strandbindingprotein(Helix-destabilizingprotein)	+	+
MAGa4140	Endopeptidase O	+	+
MAGa4150	Conservedhypotheticalprotein	+	-
MAGa4160	Glycine cleavage system H protein	-	+
MAGa4170	Hypothetical protein	-	-
MAGa4180	Cell division protein ftsZ	+	-
MAGa4190	Conserved hypothetical protein	-	-
MAGa4200	S-adenosyl-methyltransferase	-	-
MAGa4210	Protein MraZ	-	-
MAGa4220	Potassium uptake protein KtrA	+	+
MAGa4230	Potassium uptake protein KtrB	-	-
MAGa4240	Hypothetical protein	-	+
MAGa4250	ModificationmethylaseBsp6I	+	-
MAGa4260	Type II restriction enzyme Bsp6I	+	-
MAGa4270	UvrABCsystemproteinB	+	+
MAGa4280	UvrABC system protein A	+	+
MAGa4290	Conserved hypothetical protein	+	+
MAGa4300	Hypothetical protein	-	-
MAGa4310	GTP-binding protein	-	+
MAGa4330	Hypotheticalprotein	+	-
MAGa4340	Hypoxanthine-guanine phosphoribosyltransferase(HGPRT)	-	-
MAGa4350	TranscriptionelongationfactorgreA	+	+
MAGa4360	Oxidoreductase	+	-
MAGa4370	Conserved hypothetical protein	+	+
MAGa4380	Conserved hypothetical protein	+	+
MAGa4390	Putative Holliday junction resolvase	-	-
MAGa4400	Alanyl-tRNA synthetase (Alanine-tRNA ligase)(ALARS)	+	+
MAGa4410	TRNA(5-methylaminomethyl-2-thiouridylate)-methyltransferase	-	-
MAGa4420	Purine nucleoside phosphorylase (Inosinephosphorylase) (PNP)	+	+
MAGa4430	Hypothetical protein	-	-
MAGa4440	ConservedHypotheticalprotein	+	-
MAGa4450	Conserved hypothetical protein, DUF285 family	-	+
MAGa4460	Hypothetical protein, predicted lipoprotein	-	-
MAGa4470	Pseudogene of CPG DNA methylase (N-terminalpart)	-	-
MAGa4480	Pseudogene of CPG DNA methylase (C-terminalpart)	-	-
MAGa4490	Conserved hypothetical protein	-	-
MAGa4500	Hypothetical protein, predicated lipoprotein	-	-
MAGa4510	Alcohol dehydrogenase	-	+
MAGa4520	Hypothetical protein, predicted lipoprotein	-	-

Table S1: continued (8/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa4530	Hypothetical protein, DUF285 family	-	-
MAGa4540	Conserved hypothetical protein, predictedlipoprotein, DUF285 family	-	-
MAGa4550	Hypothetical protein, predicted lipoprotein	-	-
MAGa4560	Alcohol dehydrogenase	+	+
MAGa4570	DNA processing protein (Smf)	-	-
MAGa4580	Hydrolase	+	+
MAGa4590	ABC transporter permease protein	+	-
MAGa4600	P115-Like (Mycoplasma hyorhinis) ABCtransporter ATP-Binding Protein	+	+
MAGa4610	Ribonuclease III (RNase III)	-	-
MAGa4620	Fatty acid/phospholipid synthesis protein(PlsX)	+	+
MAGa4630	Conserved hypothetical protein	+	+
MAGa4640	Hypothetical protein	-	+
MAGa4650	Conserved hypothetical protein	-	-
MAGa4660	Conserved hypothetical protein (MgpA-like)	+	+
MAGa4670	Conservedhypotheticalprotein(MgpA-like)	+	+
MAGa4680	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa4690	Glycerol kinase	+	+
MAGa4700	Glycerolfacilitatorfactor	+	+
MAGa4710	Aspartyl-tRNA synthetase (Aspartate-tRNAligase) (AspRS)	+	+
MAGa4720	Histidyl-tRNAsynthetase(Histidine-tRNAligase)(HisRS)	+	+
MAGa4730	Conserved hypothetical protein	+	-
MAGa4740	Conserved hypothetical protein	-	-
MAGa4750	Hypothetical protein	+	+
MAGa4760	Conservedhypotheticalprotein	+	-
MAGa4770	50S ribosomal protein L13	+	+
MAGa4780	30S ribosomal protein S9	+	+
MAGa4800	Conserved hypothetical protein	-	-
MAGa4810	UTP-glucose-1-phosphateuridyltransferase	+	+
MAGa4820	Transposase	-	-
MAGa4840	Cation-transporting P-type ATPase	+	+
MAGa4850	CDS1	-	-
MAGa4860	CDSA	-	-
MAGa4870	CDS12	-	-
MAGa4880	CDS11	-	-
MAGa4890	CDSB	-	-
MAGa4900	CDSC	-	-
MAGa4910	CDSD	-	-
MAGa4920	CDS5	-	-
MAGa4930	CDS7	-	-
MAGa4940	CDS13	-	-
MAGa4950	CDS15	-	-
MAGa4960	CDS16	-	-
MAGa4970	Hypothetical product	-	-
MAGa4980	CDS17	+	-
MAGa4990	CDS19	-	-
MAGa5000	CDSE	-	-
MAGa5010	CDS14	+	-
MAGa5020	CDSF	-	-
MAGa5030	Hypotetical protein	+	-
MAGa5040	CDSG	-	-
MAGa5050	CDSH	-	-
MAGa5060	CDS22	-	-
MAGa5070	ABC transporter, permease protein	+	+
MAGa5080	ABC transporter, ATP-binding protein	+	+
MAGa5090	Conserved hypothetical protein	-	-
MAGa5100	Conserved hypothetical protein	-	+
MAGa5110	Conserved hypothetical protein, predictedlipoprotein	-	+
MAGa5120	Phosphomannomutase	+	+
MAGa5130	Conserved hypothetical protein	-	-
MAGa5140	Transcription termination factor (NusB)	+	+
MAGa5150	Acylcarrierproteinphosphodiesterase	+	+
MAGa5160	PhosphocarrierproteinHPr(Histidine-containingprotein)	+	+
MAGa5170	Oligopeptide transport system permease protein(OppB)	-	-

Table S1: continued (9/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa5180	Conserved hypothetical protein	-	-
MAGa5190	Conservedhypotheticalprotein,predictedlipoprotein	+	-
MAGa5200	Conserved hypothetical protein	-	-
MAGa5210	Hypothetical protein, predicted lipoprotein	+	+
MAGa5220	Translation initiation factor IF-3	+	+
MAGa5230	50S ribosomal protein L35	-	-
MAGa5240	50SribosomalproteinL20	+	+
MAGa5250	ATP-binding protein	+	-
MAGa5260	Hypothetical protein	-	+
MAGa5270	LicA	+	-
MAGa5280	GTP-binding protein (LepA)	+	-
MAGa5290	Conservedhypotheticalprotein	+	-
MAGa5300	NADPHflavinoxidoreductase	+	+
MAGa5310	Esterase/lipase	+	-
MAGa5320		-	-
MAGa5330	Transposase	-	-
MAGa5340	Pseudogene of Polysaccharide synthase(C-terminal part)	-	-
MAGa5350	Arginyl-tRNAsynthetase	+	+
MAGa5360	Malate permease	+	+
MAGa5370	L-lactate dehydrogenase (L-LDH)	+	+
MAGa5380	Hypothetical protein	+	+
MAGa5390	Phosphateacetyltransferase(Phosphotransacetylase)	+	+
MAGa5400	Acetatekinase(Acetokinase)	+	+
MAGa5410	Transmembrane protein	-	+
MAGa5420	Conserved hypothetical protein, predictedlipoprotein	+	-
MAGa5430	Hexosephosphate transport protein	+	+
MAGa5440	Conservedhypotheticalprotein	+	-
MAGa5450	EndonucleaseIV	+	+
MAGa5460	Hypothetical protein	-	-
MAGa5470	Thioredoxin (TRX)	-	-
MAGa5480	UvrABC system protein C	-	-
MAGa5490	Hypothetical protein, predicted lipoprotein	+	-
MAGa5500	P80, predicted lipoprotein	+	+
MAGa5510	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa5520	ABC transporter ATP-binding protein	+	+
MAGa5530	ABC transporter permease protein	+	+
MAGa5540	ABC transporter permease protein	+	+
MAGa5550	Hypothetical protein	-	-
MAGa5560	Hypothetical protein, predicted lipoprotein	+	+
MAGa5570	GTP-binding protein	+	+
MAGa5580	Conserved hypothetical protein	+	+
MAGa5590	Purine nucleoside phosphorylase (Inosinephosphorylase)	+	+
MAGa5600	Thymidinephosphorylase(TDRPASE)	+	+
MAGa5610	Deoxyribose-phosphatealdolase(Phosphodeoxyriboaldolase)(Deoxyriboaldolase)	+	+
MAGa5620	Triosephosphateisomerase	+	+
MAGa5630	Hypothetical protein, predicted lipoprotein	+	+
MAGa5640	IS30-like protein	-	-
MAGa5650	ABC transporter permease protein	+	+
MAGa5660	ABC transporter ATP-binding protein	+	+
MAGa5670	ABC transporter ATP-binding protein	+	+
MAGa5680	50S ribosomal protein L17	+	+
MAGa5690	DNA-directed RNA polymerase alpha chain (RNAPalpha subunit) (Transcriptase alpha chain) (RNApolymerase alpha subunit)	+	+
MAGa5700	30S ribosomal protein S11	+	+
MAGa5710	30S ribosomal protein S13	+	+
MAGa5720	TranslationinitiationfactorIF-1	+	+
MAGa5730	Methionineaminopeptidase(MAP)	+	+
MAGa5740	Adenylatekinase(ATP-AMPtransphosphorylase)	+	+
MAGa5750	Preprotein translocase SecY subunit	+	+
MAGa5800	Transposase	-	-
MAGa5810	Variable surface lipoprotein A (VpmaA)	-	-
MAGa5820	Variable surface lipoprotein B (VpmaB)	-	-

Table S1: continued (10/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa5830	Variable surface lipoprotein C (VpmaC)	-	-
MAGa5840	Variable surface lipoprotein D1 (VpmaD1)	-	-
MAGa5850	Variable surface lipoprotein E (VpmaE)	-	-
MAGa5860	Variable surface lipoprotein F1 (VpmaF1)	+	-
MAGa5870	Variable surface lipoprotein D2 (VpmaD2)	-	-
MAGa5880	Integrase-recombinase	-	-
MAGa5890	Transposase	-	-
MAGa5900	Hypothetical protein	-	-
MAGa5910	Hypothetical protein	-	-
MAGa5920	50S ribosomal protein L15	+	+
MAGa5930	30S ribosomal protein S5	+	+
MAGa5940	50S ribosomal protein L18	+	+
MAGa5950	50S ribosomal protein L6	+	+
MAGa5960	30S ribosomal protein S8	+	+
MAGa5970	30S ribosomal protein S14	-	-
MAGa5980	50S ribosomal protein L5	+	+
MAGa5990	50S ribosomal protein L24	+	+
MAGa6000	50S ribosomal protein L14	+	+
MAGa6010	30S ribosomal protein S17	+	+
MAGa6020	50SribosomalproteinL29	+	+
MAGa6030	50SribosomalproteinL16	+	+
MAGa6040	30S ribosomal protein S3	+	+
MAGa6050	50S ribosomal protein L22	+	+
MAGa6060	30S ribosomal protein S19	+	+
MAGa6070	50S ribosomal protein L2	+	+
MAGa6080	50S ribosomal protein L23	+	+
MAGa6090	50S ribosomal protein L4	+	+
MAGa6100	50S ribosomal protein L3	+	+
MAGa6110	30S ribosomal protein S10	+	+
MAGa6120	DNA-binding histone-like protein HU	+	+
MAGa6130	Putative recombinase protein	-	-
MAGa6140	RecAprotein	+	-
MAGa6150	Conserved hypothetical protein	-	+
MAGa6160	Aminopeptidase	+	+
MAGa6170	16S rRNA uridine-516 pseudouridylate synthaseand related pseudouridylate synthase	-	-
MAGa6180	50S ribosomal protein L27	-	+
MAGa6190	50S ribosomal protein L21	+	+
MAGa6200	Conserved hypothetical protein	-	-
MAGa6210	Large-conductance mechanosensitive channel	+	+
MAGa6220	HeatshockproteinGrpE(activationofDnaK)	+	+
MAGa6230	Heat-inducible transcription repressor hrcA	+	-
MAGa6240	50S ribosomal protein L31	-	-
MAGa6250	30S ribosomal protein S4	+	+
MAGa6260	Hypothetical protein	-	-
MAGa6270	DNAgyrasesubunitA	+	+
MAGa6280	Type I R/M system specificity subunit	-	-
MAGa6290	Modification(Methylase)proteinoftypeIrestriction-modificationsystem	+	-
MAGa6300	Conserved hypothetical protein	-	-
MAGa6310	Type I R/M system specificity subunit	-	+
MAGa6320	Phage family integrase	-	-
MAGa6330	HsdR, Restriction-modification enzyme subunitR	+	-
MAGa6340	Type I R/M system specificity subunit	+	-
MAGa6350	Modification (Methylase) protein of type Irestriction-modification system HsdM	+	+
MAGa6360	Hypothetical protein, predicted lipoprotein	-	-
MAGa6370		-	-
MAGa6380	5-formyltetrahydrofolate cyclo-ligase	-	-
MAGa6390	Conserved hypothetical protein	-	-
MAGa6400	Hypothetical protein, predicted lipoprotein	-	-
MAGa6410	Glutamyl-tRNAsynthetase	+	+
MAGa6420	Conserved hypothetical protein	-	-
MAGa6430	Methionine adenosyltransferase	+	-
MAGa6440	Conserved hypothetical protein	-	-

Table S1: continued (11/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa6450	SignalRecognitionParticleProtein	+	+
MAGa6460	Acetyltransferase	+	+
MAGa6470	Hypotheticalprotein	+	-
MAGa6480	Conservedhypotheticalprotein	+	+
MAGa6490	Phosphoglycerate kinase	+	+
MAGa6510	Hypothetical protein	-	-
MAGa6520	Hypothetical protein	-	-
MAGa6530	Hypothetical protein	-	-
MAGa6540	Aminopeptidase	+	-
MAGa6550	Hypothetical protein, predicted lipoprotein	-	-
MAGa6560	5'Nucleotidase, predicted lipoprotein	+	+
MAGa6570	ElongationfactorG	+	+
MAGa6580	30S ribosomal protein S7	+	+
MAGa6590	30S ribosomal protein S12	+	+
MAGa6600	Conservedhypotheticalprotein	+	+
MAGa6610	ABCtransporterATP-BindingProtein	+	+
MAGa6620	Hypothetical protein	-	-
MAGa6650	ABC transporter, ATP-binding protein	+	+
MAGa6660	ABC transporter, ATP-binding protein	+	+
MAGa6670	Hypothetical protein, predicted lipoprotein	-	-
MAGa6680	Hypothetical protein	-	-
MAGa6690	Hypothetical protein	-	-
MAGa6700	Hypothetical protein	-	-
MAGa6710	Hypothetical protein	-	-
MAGa6720	Hypothetical product, predicted lipoprotein	-	-
MAGa6730	Hypothetical protein, predicted lipoprotein	-	-
MAGa6740	Hypothetical protein	-	-
MAGa6750	Hypothetical protein, predicted lipoprotein	-	-
MAGa6760	Hypothetical protein	-	-
MAGa6770	Hypothetical protein, predicted lipoprotein	-	-
MAGa6780	Hypothetical protein, predicted lipoprotein	-	-
MAGa6790	Hypothetical protein	-	-
MAGa6800	Hypothetical protein	-	-
MAGa6810	Aminopeptidase	+	+
MAGa6820	Hypothetical protein, predicted lipoprotein	-	-
MAGa6830	Conservedhypotheticalprotein	+	-
MAGa6840	DNA-directed RNA polymerase beta' chain	+	+
MAGa6850	DNA-directed RNA polymerase beta chain	+	+
MAGa6860	Hypothetical protein, predicted lipoprotein	-	-
MAGa6870	Hypothetical protein	-	-
MAGa6880	CDS22	-	-
MAGa6890		-	-
MAGa6900	CDSH	-	-
MAGa6910	CDSG	-	-
MAGa6920	Hypotetical protein	+	-
MAGa6930	CDSF	-	-
MAGa6940	CDS14	+	-
MAGa6950	CDSE	-	-
MAGa6960	CDS19	-	-
MAGa6970	CDS17	+	-
MAGa6980	Hypothetical protein	-	-
MAGa6990	CDS16	-	-
MAGa7000	CDS15	-	-
MAGa7010	CDS13	-	-
MAGa7020	CDS7	-	-
MAGa7030	CDS5	-	-
MAGa7040	CDSD	-	-
MAGa7050	CDSC	-	-
MAGa7060	CDSB	-	-
MAGa7070	CDS11	-	-
MAGa7080	CDS12	-	-
MAGa7090	CDSA	-	-
MAGa7100	CDS1	-	-

Table S1: continued (12/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa7110	Pseudogene of Transposase	-	-
MAGa7120		-	-
MAGa7130	Hypothetical protein, predicted lipoprotein	+	+
MAGa7140	50SribosomalproteinL7/L12	+	+
MAGa7150	50S ribosomal protein L10	+	+
MAGa7160	Hypothetical protein, predicted lipoprotein	+	+
MAGa7170	Lysyl-tRNAsynthetase	+	+
MAGa7180	Conservedhypotheticalprotein	+	+
MAGa7190	Conservedhypotheticalprotein	+	+
MAGa7200	ClpB	+	+
MAGa7210	30S ribosomal protein S16	+	+
MAGa7220	tRNA (Guanine-N(1)-)-methyltransferase	-	+
MAGa7230	50S ribosomal protein L19	+	+
MAGa7240	Methionyl-tRNAformyltransferase	+	+
MAGa7250	Conserved hypothetical protein	+	-
MAGa7260	Phosphoenolpyruvate-protein phosphotransferase	-	+
MAGa7270	Transcriptional regulator	+	-
MAGa7280	Conserved hypothetical protein	-	-
MAGa7290	SugarisomeraseSGAE	+	+
MAGa7300	Hexulose-6-phosphate isomerase (HUMPI)	-	+
MAGa7310	Hexulose-6-phosphate synthase (HUMPS)	-	+
MAGa7320	Pentitol phosphotransferase enzyme II, Acomponent	-	-
MAGa7330	Pentitol phosphotransferase enzyme II, Bcomponent	-	+
MAGa7340	Transport protein SGAT	+	+
MAGa7350	Conservedhypotheticalprotein	+	+
MAGa7360	Conservedhypotheticalprotein	+	+
MAGa7370	Conserved hypothetical protein	-	-
MAGa7380	Hypothetical protein	-	-
MAGa7390	Hypothetical protein	-	-
MAGa7400	CDSE	-	-
MAGa7410	Conserved hypothetical protein	-	-
MAGa7420	IS30-like protein	-	-
MAGa7430	Hypothetical protein, predicted lipoprotein	-	-
MAGa7440	Hypothetical protein, predicted lipoprotein	-	-
MAGa7450	Conserved hypothetical protein, predictedlipoprotein, DUF285 family, truncated in N-terminal	-	-
MAGa7460	Conserved hypothetical protein, Fic proteinfamily	-	-
MAGa7470	Hypothetical protein, predicted lipoprotein,DUF285 family	+	-
MAGa7480	Hypothetical protein	-	-
MAGa7490	Conserved hypothetical protein, predictedlipoprotein	+	+
MAGa7500	Glycerol ABC transporter, permease component	-	-
MAGa7510	Glycerol transporter subunit B	-	-
MAGa7520	Glycerol ABC transporter, ATP-bindingcomponent	+	-
MAGa7530	Conserved hypothetical protein	-	+
MAGa7540	Conserved hypothetical protein	-	-
MAGa7550	Conserved hypothetical protein	+	+
MAGa7560	NH(3)-dependent NAD(+) synthetase	-	+
MAGa7570	ICEF-IAORF9ortholog	+	-
MAGa7580	Esterase/lipase	+	-
MAGa7590	Hypothetical RNA methyltransferase	-	-
MAGa7600	PutativeDNAhelicase	+	-
MAGa7610	Conserved hypothetical protein	-	-
MAGa7620	Conserved hypothetical protein	-	-
MAGa7630	Conserved hypothetical protein	-	-
MAGa7650	Modificationmethylase	+	-
MAGa7660	Hypothetical protein	+	-
MAGa7670	Conserved hypothetical protein	+	-
MAGa7680	Hypothetical protein	-	-
MAGa7690	Ribonuclease H II	-	-
MAGa7700	Conserved hypothetical protein	+	+
MAGa7710	RibosomallargesubunitpseudouridinesynthaseD	+	+
MAGa7720	Chromate transport protein	-	-
MAGa7730	Chromate transport protein	-	+

Table S1: continued (13/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa7740	tRNA(guanine-N(7)-)-methyltransferase	+	+
MAGa7750	Conserved hypothetical protein	-	+
MAGa7760	Conserved hypothetical protein	-	-
MAGa7770	Conserved hypothetical protein	+	-
MAGa7780	Hypothetical protein	-	-
MAGa7790	Conserved hypothetical protein	-	-
MAGa7800	DNA polymerase III subunit	-	-
MAGa7810	Thymidylate kinase	+	+
MAGa7820	Recombination protein RecR	-	+
MAGa7830	Hypothetical protein	-	-
MAGa7840	DNApolymeraseIIIsubunitgammaandTAU	+	+
MAGa7850	Hypothetical protein	+	-
MAGa7860	2,3-bisphosphoglycerate-independentphosphoglyceratemutase	+	+
MAGa7870	Conserved hypothetical protein	-	-
MAGa7880	Hypothetical protein	+	-
MAGa7900	Conserved hypothetical protein	+	+
MAGa7910	CDP-diacylglycerol-glycerol-3-phosphate3-phosphatidyltransferase	-	-
MAGa7920	Phosphatidate cytidylyltransferase	-	-
MAGa7930	Transposase	-	-
MAGa7940	Conserved hypothetical protein	-	-
MAGa7950	Adenine phosphoribosyltransferase	-	+
MAGa7960	Translation initiation factor IF-2	+	+
MAGa7970	Conserved hypothetical protein	-	-
MAGa7980	Transcription elongation protein NusA	+	+
MAGa7990	Conservedhypotheticalprotein	+	-
MAGa8000	Thymidine kinase	+	+
MAGa8010	Aminopeptidase (Leucine aminopeptidase)	+	+
MAGa8020	Conserved hypothetical protein	+	-
MAGa8030	Hypothetical protein	-	-
MAGa8040	Variable surface lipoprotein G (VpmaG)	+	-
MAGa8050	Variable surface lipoprotein F2 (VpmaF2)	+	-
MAGa8060	Variable surface lipoprotein X (VpmaX)	+	+
MAGa8070	Variable surface lipoprotein W (VpmaW)	+	-
MAGa8080	Variable surface lipoprotein C (VpmaC)	-	+
MAGa8090	Variable surface lipoprotein E (VpmaE)	-	-
MAGa8100	Variable surface lipoprotein B (VpmaB)	+	+
MAGa8110	Variable surface lipoprotein A (VpmaA)	+	-
MAGa8120	Variable surface lipoprotein D2 (VpmaD2)	-	-
MAGa8130	Conserved hypothetical protein	+	-
MAGa8140	Conserved hypothetical protein	-	-
MAGa8150	Variable surface lipoprotein H (VpmaH)	+	-
MAGa8160	Variable surface lipoprotein I (VpmaI)	+	-
MAGa8170	Variable surface lipoprotein F1 (VpmaF1)	+	+
MAGa8180	Variable surface lipoprotein J (VpmaJ)	+	-
MAGa8190	VariablesurfacelipoproteinK(vpmaK)	+	-
MAGa8200	Variable surface lipoprotein L (VpmaL)	-	-
MAGa8210	Variable surface lipoprotein D1 (VpmaD1)	+	-
MAGa8220	Integrase-recombinase	-	-
MAGa8230	Transposase	-	-
MAGa8250	Conserved hypothetical protein	+	+
MAGa8260	Hypothetical protein, predicted lipoprotein	+	-
MAGa8270	Conserved hypothetical protein	-	-
MAGa8280	Hypothetical protein, predicted lipoprotein	-	-
MAGa8290	Aspartyl/glutamyl-tRNA(Asn/Gln)amidotransferasesubunit	+	+
MAGa8300	Glutamyl-tRNA(GLN) amidotransferase subunit A	-	+
MAGa8310	Conserved hypothetical protein	-	-
MAGa8320	RibosomallargesubunitpseudouridinesynthaseC	+	-
MAGa8330	SegregationandcondensationproteinB	+	+
MAGa8340	SegregationandcondensationproteinA	+	-
MAGa8350	1-acyl-SN-glycerol-3-phosphate acyltransferase	+	+
MAGa8360	Holo-[acyl-carrier protein]synthase	-	-
MAGa8370	GTP-binding protein engA	+	+
MAGa8380	Cytidylatekinase(CK)	+	+

Table S1: continued (14/14).

MAGa[a]	Product	Detected by MS/MS in 5632[b]	Homolog detected in PG2[b]
MAGa8390	Hypothetical protein	-	-
MAGa8400	Conserved hypothetical protein, predictedlipoprotein, DUF285 family	-	-
MAGa8410		-	-
MAGa8420	Heat shock protein DNAJ (activation of DNAK)	+	+
MAGa8440	Ribosome-binding factor A	-	-
MAGa8460	Tyrosyl-tRNA synthetase 1	+	+
MAGa8470	Hypothetical protein	-	+
MAGa8480	Conservedhypotheticalprotein	+	+
MAGa8500	Acyl carrier protein homolog	-	-
MAGa8510	Protoporphirogen oxidase HEMK	-	-
MAGa8520	Peptidechainreleasefactor1	+	+
MAGa8530	Conserved hypothetical protein	+	+
MAGa8540	DNA gyrase subunit B	+	+
MAGa8550	Thioredoxin	+	+
MAGa8560	SsrA-binding protein	-	-
MAGa8570	Conserved hypothetical protein	+	+
MAGa8580	Hypothetical protein	-	-
MAGa8600	Cation-transporting P-ATPase	+	+
MAGa8610	ABCtransporterATP-bindingprotein	+	+
MAGa8620	ABC transporter permease protein	+	+
MAGa8630	Cell division protein ftsH homolog	+	+
MAGa8640	Conserved hypothetical protein	-	-
MAGa8650	Peptidyl-tRNA hydrolase	+	+
MAGa8660	Exodeoxyribonuclease V alpha chain	+	-

[a] CDS of *M. agalactiae* strain 5632 (Molligen Mnemonic)

[b] Proteomic (see materials and methods) : (+) indicates that peptides were detected by MS/MS for the corresponding CDS, suggesting expression of the corresponding gene, (-) indicates that no specific peptides were detected for the corresponding CDS.

DISCUSSION GENERALE ET PERSPECTIVES

DISCUSSIONS GENERALES ET PERSPECTIVES

Les mycoplasmoses de ruminants sont des maladies généralement chroniques et débilitantes, responsables de pertes économiques importantes pour les productions bovines, ovines et caprines. La péripneumonie contagieuse bovine et la pleuropneumonie contagieuse caprine affectent les élevages des pays en voie de développement. *Mycoplasma bovis*, agent de la broncho-pneumonie enzootique et de mammites, est un problème important pour les élevages bovins d'Europe et d'Amérique du Nord (Le Grand *et al.*, 2008). L'agalactie contagieuse constitue la mycoplasmose la plus commune et la plus grave chez les petits ruminants en Europe (Bergonier et Berthelot, 2008).

La maîtrise des infections mycoplasmiques passe par une connaissance approfondie de leur épidémiologie, l'amélioration des moyens de dépistage, de prévention et de traitement. Si le développement de vaccins ou de molécules curatives efficaces reste un challenge, des progrès significatifs ont été réalisés ces dernières années grâce à l'acquisition des bases moléculaires des mycoplasmes au plan antigénique et génomique (Citti *et al.*, 2005 ; Sirand-Pugnet *et al.*, 2007a). Ces connaissances ont permis des avancées appliquées sur le terrain, comme entre autres, l'amélioration du diagnostic par le développement de divers essais PCR (Markham et Noormohammadi, 2005). D'un point de vue plus fondamental, ces avancées ont aussi apporté un éclairage nouveau sur la biologie des mycoplasmes. Plus particulièrement, elles ont fait évoluer les connaissances sur les systèmes génétiques à l'origine de la dynamique de surface, mis en évidence une grande plasticité et de nombreux échanges génétiques entre mycoplasmes de ruminants (Sirand-Pugnet *et al.*, 2007b).

Dans la continuité des travaux réalisés en génomique des mycoplasmes, l'objet du travail réalisé ici est une caractérisation de la diversité génétique chez *M. agalactiae*. Cette caractérisation est la base préliminaire indispensable à l'identification et l'utilisation de marqueurs génétiques comme traceurs épidémiologiques. A l'exception des réarrangements à haute fréquence liés à

l'hypervariabilité d'expression des Vpma, les travaux sur une collection de 101 isolats présentés dans le premier chapitre mettent en évidence que *M. agalactiae* est une espèce très homogène : hormis quelques souches très atypiques, la majeure partie des souches ressemble à la souche type PG2. Sur soixante-trois isolats des Pyrénées-Atlantiques, une seule variation de répertoire est observée au niveau du locus *vpma* où surviennent des réarrangements à haute fréquence et concernerait un des six gènes *vpma* caractérisés : le gène *vpmaY* qui n'est pas détecté chez cinq isolats. La méthode utilisée ne permet pas de préciser si ce gène est manquant ou s'il présente une séquence n'hybridant pas avec la sonde utilisée pour reconnaître *vpmaY*. Cette sonde concerne une portion de séquence spécifique à *vpmaY* qui est flanquée de séquences partagées avec *vpmaX*. Sur la base des travaux sur la diversité du répertoire *vpma* de la souche 5632 présentés dans le second chapitre, la portion spécifique de *vpmaY* apparaît répétée respectivement deux et trois fois dans *vpmaF2* et *vpmaF1* ou très divergente dans *vpmaK* [voir 2eme chapitre, Figure 2], il semble ainsi probable que des recombinaisons entre gènes *vpma* aient entraîné la disparition ou la variation de cette séquence chez les cinq isolats. Des explorations plus précises comme le séquençage de ce gène ou du locus *vpma* de ces isolats permettraient de déterminer plus précisément ce point.

A l'exception du la variabilité du gène *vpmaY* pour cinq isolats, les isolats rencontrés dans les Pyrénées-Atlantiques ne semblent pas présenter de variation sur la base des trois méthodes utilisées (VNTR, profils AGA, répertoire *vpma*). Ces résultats sont en faveur de la circulation d'une seule et même souche dans cette région au moins sur la période étudiée qui s'étend de 1977 à 2007. Le bassin d'élevage ovin des Pyrénées-Atlantiques est l'héritage d'une tradition agro-pastorale implantée dans cette région depuis des millénaires. Sans pouvoir précisément dater son apparition, de mémoire de berger, l'agalactie y est connue depuis toujours. A l'image des observations faites sur l'agent de la tuberculose humaine, *M. tuberculosis* (Sreevatsan *et al.*, 1997 ; Brosch *et al.*, 2002 ; Dos Vultos *et al.*, 2008), une hypothèse sur les raisons de la stabilité des isolats étudiés pourrait être une très ancienne co-habitation hôte-parasite entre la souche et l'élevage ovin de cette zone. La souche rencontrée serait assez bien adaptée à son environnement et stable face à une faible variation de ce dernier. Dans ce

contexte, l'épidémiologie moléculaire se retrouve dans une impasse par manque d'outil pour différencier les isolats. Utilisées chez d'autres bactéries, des techniques basées sur de faibles variations comme les polymorphismes nucléotidiques simples ou SNPs (*Single Nucleotide Polymorphisms*) (Schork *et al.*, 2000 ; Gutacker *et al.*, 2002 ; Dos Vultos *et al.*, 2008) ou des techniques s'appuyant sur des "points chauds" ou *hot spots* de réarrangements sur le génome comme certains VNTR ou les loci de type CRISPR (pour *Clustered Regularly Interspaced Short Palindromic Repeats*) (Ishino *et al.*, 1987 ; Goguet de la Salmoniere *et al.*, 1997 ; Kamerbeek *et al.*, 1997 ; Jansen *et al.*, 2002) pourraient être envisagées. Notre travail montre que parmi les VNTR définis chez *M. agalactiae* (McAuliffe *et al.*, 2008), le VNTR 17 différencie les isolats des Pyrénées-Atlantiques étudiés de la souche type PG2 mais qu'aucun VNTR ne varie au sein des isolats des Pyrénées-Atlantiques. Sur la base de la séquence des génomes de PG2 et 5632, la recherche de locus CRISPR à l'aide des programmes actuels de détection (Benson, 1999 ; Grissa *et al.*, 2007) n'a pas abouti. Le locus *vpma* constitue un *hot spot* de recombinaisons spécifiques, cependant ces recombinaisons sont réversibles et peuvent survenir de manière intraclonale (Glew *et al.*, 2002). Ce caractère hypervariable et réversible entraîne un fort risque de convergence des variations observées. Les profils de restriction générés par les enzymes *Ase*I et *Alw*I qui coupent à l'intérieur des gènes *vpma* permettent théoriquement de s'affranchir de cette hypervariabilité. Ces profils hybridés par *Southern blot* avec la sonde A3F qui reconnaît une région des gènes *vpma* conservée proche de la séquence du peptide signal sont appelés profils "A3F". Dans le contexte de souches monomorphes, la comparaison des variations de ces profils pourrait constituer une solution envisageable pour suivre des isolats sur une période de temps réduite, dans des cas particuliers de chaînes de transmission supposées courtes et avec les précautions d'interprétation liées à la possible convergence des profils.

Si l'utilisation des profils "A3F" est envisageable pour des isolats des Pyrénées-Atlantiques, l'étude comparative des loci *vpma* des souches PG2 et 5632 réalisée dans le deuxième chapitre met en évidence une étonnante complexité du répertoire *vpma* chez 5632 expliquant les limites de résolution observées pour certains isolats avec la technique *Southern blot* utilisée. Le profil "A3F" qui révèle

l'ensemble des fragments porteurs de gènes *vpma* est déjà composé de bandes superposées pour la souche type PG2 qui contient six *vpma*. Avec vingt-trois gènes *vpma* découverts dans la souche 5632, il est évident que le pouvoir de discrimination de cette technique se voit dépassé [voir 1er chapitre, Figure 10].

L'étendue du répertoire *vpma* de 5632 représente une structure mosaïque qui suggère des duplications de gènes, des insertions et/ou des délétions de motifs répétés et des événements de recombinaisons homologues entre les gènes *vpma*. Une telle possibilité de dynamique de surface jusqu'alors soupçonnée chez *M. bovis* (Lysnyansky *et al.*, 1999 ; Nussbaum *et al.*, 2002) est donc partagée par 5632 et probablement d'autres isolats de *M. agalactiae*. La duplication des loci *vpma* est, quant à elle, un phénomène uniquement décrit chez 5632 à ce jour. Sa conséquence majeure est la démultiplication du nombre de possibilités des combinaisons exprimées à la surface car elle permet a priori l'expression concomitante de deux Vpma. Vu l'identité des séquences entre les deux loci *vpma*, cet événement de duplication dans 5632 semble récent.

Les séquences d'insertion ou IS (*insertion sequence*) trouvées a proximité ont vraisemblablement joué un rôle dans cette duplication. La participation des IS à des réarrangements et transferts de gènes est connue chez certaines bactéries (Mahillon et Chandler, 1998). Malgré la présence d'une séquence de transposase détectée par les analyses protéomiques, la souche type PG2 ne contient apparemment pas de séquence d'insertion fonctionnelle. L'observation d'une forte proportion de gènes liés à des transferts génétiques horizontaux avec les mycoplasmes du groupe "mycoides" chez cette souche (environ 18% du génome) restait sans explication sur les mécanismes à l'origine de ces transferts. Chez 5632, non seulement la portion des gènes impliqués dans un transfert génétique avec les mycoplasmes du groupe "mycoides" reste une évidence, mais elle arbore aussi trois éléments intégratifs conjugatifs ou ICE (*integrative-conjugative element*) et quinze IS apparemment fonctionnelles (12 IS*Mag1* et 3 IS*Mag2*) fréquemment retrouvées à proximité d'éléments prédits comme transférés. La participation des IS (Mahillon et Chandler, 1998) comme celle des ICE (Seth-Smith et Croucher, 2009) à des réarrangements et transferts de gènes est connue chez certaines bactéries : par exemple, de multiples réarrangements associés aux

IS ont été évoqués lors d'études comparatives de souches d'*Escherischia coli* (Komoda *et al.*, 1991), *Lactococcus lactis* (Daveran-Mingot *et al.*, 1998), *Bordetella* (Parkhill *et al.*, 2003), *Burkholderia* (Nierman *et al.*, 2004), *Yersinia* (Chain *et al.*, 2004), *Shigella* (Yang *et al.*, 2005), *Francisella* (Petrosino *et al.*, 2006) et *Mycobacterium* (Stinear *et al.*, 2007). Ces réarrangements génomiques peuvent conduire à un accroissement de la virulence des souches et participent de manière générale à une adaptation aux conditions changeantes du milieu (Mahillon *et al.*, 1999). Si une différence de virulence entre les souches dépourvues ou porteuses d'éléments génétiques mobiles reste à définir dans le cas de *M. agalactiae* ou *M. bovis*, les éléments génétiques mobiles y apparaissent pour le moins corrélés à la génération de diversité génétique. L'implication de ce mobilome dans des transferts de gènes et des réarrangements chromosomiques pourrait avoir des conséquences évolutives et épidémiologiques chez les mycoplasmes. Ainsi, suite à l'image d'une évolution des mycoplasmes par perte de matériel génétique, une nouvelle vision émerge, celle d'une riche dynamique évolutive liée à des éléments mobiles et des flux de gènes entre les génomes de mycoplasmes partageant un même hôte.

Des études expérimentales sur les flux de gènes entre mycoplasmes apporteraient une connaissance mécanistique de ces phénomènes à ce jour encore très lacunaire. L'identification et la caractérisation récente d'ICE chez *M. agalactiae* 5632 suggère une possibilité d'échange de matériel génétique par un mécanisme de type conjugatif (Marenda *et al.*, 2004 ; Marenda *et al.*, 2006). Chez *M. pulmonis*, quelques éléments expérimentaux sont en faveur de l'existence d'un mécanisme de conjugaison (Teachman *et al.*, 2002). De telles expériences de conjugaison *in vitro* réalisées avec des mélanges de souches permettraient d'étudier les modalités et les déterminants de ces échanges conjugatifs au sein de l'espèce *M. agalactiae* mais également entre différentes espèces de mycoplames de ruminants. De même, pour évaluer la portée de ces échanges de gènes dans la spécificité d'hôte, l'adaptation à de nouvelles niches et l'acquisition de facteurs de virulence, il conviendrait d'étudier la diversité, la prévalence et l'étendue du mobilome de différents isolats de *M. agalactiae* et des autres espèces de mycoplasmes de ruminants.

La comparaison des génomes de plus de souches permettrait d'affiner ces observations sur la plasticité des génomes de mycoplasmes de ruminants. Maintenant que deux génomes de l'espèce *M. agalactiae* sont entièrement séquencés et circularisés, les comparaisons génomiques suivantes pourront s'appuyer sur ces bases et employer de nouvelles technologies de séquençage plus rapides et de moins en moins coûteuses comme les systèmes 454 (Roche Life Sciences) (Margulies *et al.*, 2005) ou Solexa (Ilumina) (Bentley, 2006) [pour une revue voir (Pettersson *et al.*, 2009)]. Des systèmes de cartographie génomique comme l'*optical mapping* (Schwartz *et al.*, 1993 ; Samad *et al.*, 1995) pourraient également être utilisés pour visualiser les réarrangements chromosomiques entre diverses souches et connaître l'ordre d'agencement des contigs obtenus avec les approches 454 et/ou Solexa. L'annotation de ces génomes serait grandement facilitée, au moins en première intention, par l'usage de méthodes bioinformatiques automatisées comme celle utilisée pour *M. agalactiae* 5632 développée par le centre de bioinformatique de Bordeaux sur la plateforme CAAT-Box (Frangeul *et al.*, 2004) ou celle utilisée pour le génome de *M. conjunctivae* (Calderon-Copete *et al.*, 2009).

Notre travail met en évidence quelques souches d'intérêt pour cet objectif de séquençage à haut débit. Il faudrait par exemple comparer les génomes de PG2 et 5632 à la souche 8062, initialement identifiée comme *M. agalactiae* et appartenant en réalité à *M. bovis*. Ces comparaisons permettraient de préciser les différences entre les deux espèces. En effet, si 5632 semble à la bordure du spectre d'espèce de *M. agalactiae*, 8062 peut être vue comme son pendant côté *M. bovis*. L'apport du séquençage de la souche type de *M. bovis* PG45, en cours de réalisation au *Craig Venter Institute*, apportera également des éléments intéressants à ces comparaisons génomiques sur les bases de la spéciation *agalactiae*/*bovis*.

L'émergence de *M. agalactiae* comme pathogène associé à une expression clinique atypique à dominante pulmonaire chez les bouquetins dans les Alpes constitue également un domaine d'étude intéressant (Barthe *et al.*, 2006 ; Auliac, 2008). En effet, l'analyse des génomes des isolats 13387 et 14628 provenant de poumon de bouquetins permettrait d'acquérir des connaissances sur le passage

de la barrière d'espèce et la capacité d'adaptation des mycoplasmes à de nouvelles niches. Un portage de *M. agalactiae* a aussi été rencontré chez des bouquetins espagnols mais sans symptôme clinique apparent (Gonzalez-Candela *et al.*, 2006 ; Gonzalez-Candela *et al.*, 2007 ; Verbisck-Bucker *et al.*, 2008). Il serait intéressant de comparer ces isolats à ceux rencontrés dans les Alpes afin de déterminer s'il se dégage une convergence des mécanismes d'adaptation à un hôte donné. A cela pourraient être ajoutées des souches d'autres espèces de mycoplasmes : des souches d'espèces appartenant au groupe "mycoides" pour préciser l'origine des transferts génétiques horizontaux observés chez *M. agalactiae*, ainsi que des souches à virulence atténuée comme celles de *M. mycoides* subsp. *mycoides* utilisées en vaccination voire des espèces apathogènes comme *M. arginini*.

En conclusion, l'analyse de mycoplasmes qui diffèrent par leur spécificité d'hôte et leur virulence permettrait d'identifier les gènes supports impliqués dans ces valences. Un projet financé par l'Agence Nationale de la Recherche (ANR) s'est fixé cet objectif et des résultats de génomique comparative devraient être obtenus sur vingt génomes de ruminants. Ce projet devrait donner une idée du cœur des génomes de mycoplasmes de ruminants et apporter des éléments sur les bases génétiques des spécificités de chacun. Ces analyses *in silico* serviront de base pour des études de génomique fonctionnelle. Ainsi, une confirmation expérimentale de l'action des gènes identifiés comme probablement impliqués dans la virulence ou la spécificité d'hôte pourra être obtenue par la convergence des analyses *in silico* avec des analyses de banques de mutagénèse de type STM *in vitro* et *in vivo* ou par l'étude spécifique de mutants de ces gènes.

De même, ces analyses *in silico* devraient ouvrir de nouvelles pistes pour la mise au point de stratégies vaccinales innovantes (Kaushik et Sehgal, 2008 ; Bambini et Rappuoli, 2009). Il devient en effet possible de développer des technologies de criblage des peptides candidats présents sur tout un génome appelées "anti-genome". Cette technique intégrative décrite et appliquée chez *Staphylococcus aureus* (Etz *et al.*, 2002) combine le pouvoir de la couverture de l'ensemble de l'information génomique à l'identification d'antigènes par sérologie. Ces auteurs (Etz *et al.*, 2002) ont construit une banque d'expression recombinante par fusion à

une protéine exposée à la surface d'*E. coli* (LambA ou FhuA) de tous les peptides théoriquement codés obtenus par fragmentation du génome. Les clones qui expriment des peptides réellement produits *in vivo* par *S. aureus* sont sélectionnés sur la base de leur affinité avec un mélange de sérums de patients naturellement infectés par *S. aureus* et leur pouvoir opsonisant est mesuré par un test *in vitro*. Les clones sélectionnés ont été séquencés et ont permis d'identifier 60 peptides immunogènes, potentiels candidats vaccins (Etz *et al.*, 2002). Cette méthode a été adaptée notamment à *Helicobacter pylori* (Meinke *et al.*, 2009) ainsi qu'à diverses espèces de Staphylocoques et de Streptocoques (Meinke *et al.*, 2005 ; Giefing *et al.*, 2008). Et des banques ont été crées avec des souches d'*Enterococcus faecalis*, *Chlamydia pneumoniae*, *Escherichia coli* entérotoxinogène, et *Campylobacter jejuni* (Henics *et al.*, 2003). En prenant en compte la particularité du code génétique des mycoplasmes où le codon stop universel UGA code pour le tryptophane, cette méthode permettrait d'identifier les peptides les plus à même de constituer des candidats vaccins. Toujours dans cet objectif vaccinal, il serait également envisageable de sélectionner des candidats sur la base de l'immunoprotéome de *M. agalactiae,* c'est-à-dire des protéines présentes à la fois dans les protéomes de PG2 et 5632 et montrant une réaction avec des sérums de brebis immunisées.

En 2005, Tettelin *et al.* (Tettelin *et al.*, 2005) ont introduit le concept de pan-génome ou supragénome comme le répertoire global des gènes rencontrés dans une espèce donnée par opposition au "cœur de génome" (genome-cœur ou *core genome*) qui correspond à l'ensemble des gènes communs à tous les isolats d'une espèce donnée. Plus le nombre d'isolats séquencés pour une espèce augmente, plus le génome-cœur se resserre tandis que le pan-génome s'agrandit. A la suite de PG2 et 5632, le séquençage de plusieurs isolats de l'espèce *M. agalactiae* permettra de préciser le pan-génome de cette espèce. Cependant, les modélisations mathématiques estiment que le pan-génome de *Mollicutes* comme *Ureaplasma urealyticum* est restreint et que sa couverture est obtenue avec peu de souches (Tettelin *et al.*, 2008). Par extrapolation, le séquençage de nouvelles souches de *M. agalactiae* ne devrait pas apporter pléthore de nouveaux gènes. Toutefois, il apparaît dès à présent que cette vision des gènes constituant une espèce bactérienne est un peu bousculée par les mycoplasmes de ruminants pour

lesquels existe une dynamique inter-espèces avec de nombreux échanges génomiques. Il semble qu'il sera peut-être plus pertinent de définir un pan-génome englobant les espèces de mycoplasmes de ruminants. Il conviendrait même de définir plutôt un "génome accessoire" englobant les gènes retrouvés chez diverses espèces de mycoplasmes de ruminants face à des "cœurs de génomes" qui correspondraient plus à une origine phylogénétique comme les groupes "mycoides" ou "hominis". Ces "cœurs de génomes" reflèteraient diverses voies pour le concept de cellule auto-réplicative minimale. Et le "génome accessoire" définit avec les mycoplasmes de ruminants serait un pool de gènes utiles à la pathogénie et à l'adaptation chez un hôte ruminant. Il est probable que ce "génome accessoire" puisse être scindé plus précisément en fonction de l'espèce hôte (ovins, caprins, bovins, chamois…). La comparaison des cœurs de génomes permettrait d'approcher les combinaisons possibles de gènes essentiels à la construction d'une espèce synthétique douée d'auto-réplication. Des avancées de la biologie synthétique ont déjà été accomplies chez les mycoplasmes avec la transplantation d'une cellule de *M. mycoides* subsp. *capricolum* avec un génome entier de *M. mycoides* subsp. *mycoides* LC (Lartigue *et al.*, 2007). Le génome transplanté peut même être au préalable cloné et modifié dans un système levure (Lartigue *et al.*, 2009). Un chromosome totalement artificiel copiant le génome de *M. genitalium* a aussi parallèlement été constitué avec succès chez la levure (Gibson *et al.*, 2008a ; Gibson *et al.*, 2008b). Le jour où ces techniques réussiront à être combinées, les outils pour la construction d'organismes chimères auto-réplicatifs seront réunis. Pour maîtriser les gènes à intégrer dans un tel modèle de biologie synthétique du chemin reste à faire.

Les mycoplasmes modernes sont le fruit d'une longue évolution à partir de mollicutes ancestraux probablement apparus il y a six cents millions d'années. Ce processus les a conduits au commensalisme ou au parasitisme obligatoire de cellules hôtes (Maniloff, 2002). Les mammifères seraient apparus il y a deux cents millions d'années. La lignée de l'*Homo sapiens* daterait de deux cent mille ans et les animaux auraient commencés à être domestiqués il y a seulement dix mille ans (Diamond, 2000 ; Boyd et Silk, 2003). Ainsi, pour plus de 99,9% de leur histoire évolutive, les mycoplasmes ont interagi avec des animaux non domestiqués pour finir par les conduire aux mycoplasmoses de l'Homme et des

animaux observées aujourd'hui (Brown *et al.*, 2005). Malgré leur petite taille et leur petit génome, ils semblent redoubler de dynamisme avec une grande diversité génétique, fruit de nombreux échanges. Leur évolution pourrait les avoir conduits "malgré eux" vers une biologie artificielle et au service d'outils de biologie synthétique. Il conviendra de définir les objectifs poursuivis par ces mycoplasmes de synthèse. Les champs d'application sont vastes de la vaccination à la détoxification de polluants en passant par le bioterrorisme… Et du mycoplasme sain au mycoplasme équitable, la biologie comme la mycoplasmologie de demain poseront sûrement des questions éthiques.

Références

Auliac P. (2008). Mortalité chez les bouquetins des Alpes, Chasser en Savoie, vol. 93.

Bambini S. et Rappuoli R. (2009). The use of genomics in microbial vaccine development. *Drug Discov Today* **14:**252-260.

Barthe M.-N., Revelli P., Game Y., Arcangioli M.-A., Solsona M., Tardy F. et Poumarat F. (2006). *Mycoplasmes et pneumonies chez le bouquetin des Alpes*. Presented at the Coordination Nationale de lutte Agalactie Contagieuse et mycoplasmoses des petits ruminants, Toulouse (France), 21 novembre 2006.

Benson G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27:**573-580.

Bentley D.R. (2006). Whole-genome re-sequencing. *Curr Opin Genet Dev* **16:**545-552.

Bergonier D. et Berthelot X. (2008). Mycoplasmoses des petits ruminants : Le syndrome de l'agalactie contagieuse. *Bulletin de l'Académie vétérinaire de France* **161:**167-177.

Boyd R. et Silk J. (2003). L'aventure humaine, de la molécule à la culture. *De Boeck*.

Brosch R., Gordon S.V., Marmiesse M., Brodin P., Buchrieser C., Eiglmeier K., Garnier T., Gutierrez C., Hewinson G., Kremer K., Parsons L.M., Pym A.S., Samper S., van Soolingen D. et Cole S.T. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* **99:**3684-3689.

Brown D.R., Zacher L.A., Wendland L.D. et Brown M.B. (2005). Emerging Mycoplasmoses in Wildlife, p. 383-414. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: Molecular biology, pathogenicity and strategies for control. *Horizon Bioscience*, Norfolk (United Kingdom).

Calderon-Copete S.P., Wigger G., Wunderlin C., Schmidheini T., Frey J., Quail M.A. et Falquet L. (2009). The *Mycoplasma conjunctivae* genome sequencing, annotation and analysis. *BMC Bioinformatics* **10 Suppl 6:**S7.

Chain P.S., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O., Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L., Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C., Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M., Derbise A., Hauser L.J. et Garcia E. (2004). Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci U S A* **101:**13826-13831.

Citti C., Browning G.F. et Rosengarten R. (2005). Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas., p. 439–484. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: Molecular biology, pathogenicity and strategies for control. *Horizon Bioscience*, Norfolk (United Kingdom).

Daveran-Mingot M.L., Campo N., Ritzenthaler P. et Le Bourgeois P. (1998). A natural large chromosomal inversion in *Lactococcus lactis* is mediated by

homologous recombination between two insertion sequences. *J Bacteriol* **180:**4834-4842.

Diamond J. (2000). De l'inégalité parmi les sociétés. Essai sur l'homme et l'environnement dans l'histoire. *Gallimard*.

Dos Vultos T., Mestre O., Rauzier J., Golec M., Rastogi N., Rasolofo V., Tonjum T., Sola C., Matic I. et Gicquel B. (2008). Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. *PLoS ONE* **3:**e1538.

Etz H., Minh D.B., Henics T., Dryla A., Winkler B., Triska C., Boyd A.P., Sollner J., Schmidt W., von Ahsen U., Buschle M., Gill S.R., Kolonay J., Khalak H., Fraser C.M., von Gabain A., Nagy E. et Meinke A. (2002). Identification of in vivo expressed vaccine candidate antigens from *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **99:**6573-6578.

Frangeul L., Glaser P., Rusniok C., Buchrieser C., Duchaud E., Dehoux P. et Kunst F. (2004). CAAT-Box, Contigs-Assembly and Annotation Tool-Box for genome sequencing projects. *Bioinformatics* **20:**790-797.

Gibson D.G., Benders G.A., Andrews-Pfannkoch C., Denisova E.A., Baden-Tillson H., Zaveri J., Stockwell T.B., Brownley A., Thomas D.W., Algire M.A., Merryman C., Young L., Noskov V.N., Glass J.I., Venter J.C., Hutchison C.A., 3rd et Smith H.O. (2008a). Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319:**1215-1220.

Gibson D.G., Benders G.A., Axelrod K.C., Zaveri J., Algire M.A., Moodie M., Montague M.G., Venter J.C., Smith H.O. et Hutchison C.A., 3rd. (2008b). One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proc Natl Acad Sci U S A* **105:**20404-20409.

Giefing C., Meinke A.L., Hanner M., Henics T., Bui M.D., Gelbmann D., Lundberg U., Senn B.M., Schunn M., Habel A., Henriques-Normark B., Ortqvist A., Kalin M., von Gabain A. et Nagy E. (2008). Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. *J Exp Med* **205:**117-131.

Glew M.D., Marenda M., Rosengarten R. et Citti C. (2002). Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus. *J Bacteriol* **184:**5987-5998.

Goguet de la Salmoniere Y.O., Li H.M., Torrea G., Bunschoten A., van Embden J. et Gicquel B. (1997). Evaluation of spoligotyping in a study of the transmission of *Mycobacterium tuberculosis*. *J Clin Microbiol* **35:**2210-2214.

Gonzalez-Candela M., Cubero-Pablo M.J., Martin-Atance P. et Leon-Vizcaino L. (2006). Potential pathogens carried by Spanish ibex (*Capra pyrenaica hispanica*) in southern Spain. *J Wildl Dis* **42:**325-334.

Gonzalez-Candela M., Verbisck-Bucker G., Martin-Atance P., Cubero-Pablo M.J. et Leon-Vizcaino L. (2007). Mycoplasmas isolated from Spanish ibex (*Capra pyrenaica hispanica*): frequency and risk factors. *Vet Rec* **161:**167-168.

Grissa I., Vergnaud G. et Pourcel C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35:**W52-57.

Gutacker M.M., Smoot J.C., Migliaccio C.A.L., Ricklefs S.M., Hua S., Cousins D.V., Graviss E.A., Shashkina E., Kreiswirth B.N. et Musser J.M. (2002). Genome-wide analysis of synonymous single nucleotide Polymorphisms in *Mycobacterium tuberculosis* complex organisms: Resolution of genetic relationships among closely related microbial strains. *Genetics* **162:**1533-1543.

Henics T., Winkler B., Pfeifer U., Gill S.R., Buschle M., von Gabain A. et Meinke A.L. (2003). Small-fragment genomic libraries for the display of putative epitopes from clinically significant pathogens. *Biotechniques* **35:**196-202, 204, 206 passim.

Ishino Y., Shinagawa H., Makino K., Amemura M. et Nakata A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169:**5429-5433.

Jansen R., Embden J.D., Gaastra W. et Schouls L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43:**1565-1575.

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

Kaushik D.K. et Sehgal D. (2008). Developing antibacterial vaccines in genomics and proteomics era. *Scand J Immunol* **67:**544-552.

Komoda Y., Enomoto M. et Tominaga A. (1991). Large inversion in *Escherichia coli* K-12 1485IN between inversely oriented IS3 elements near lac and cdd. *Genetics* **129:**639-645.

Lartigue C., Glass J.I., Alperovich N., Pieper R., Parmar P.P., Hutchison C.A., 3rd, Smith H.O. et Venter J.C. (2007). Genome transplantation in bacteria: changing one species to another. *Science* **317:**632-638.

Lartigue C., Vashee S., Algire M.A., Chuang R.Y., Benders G.A., Ma L., Noskov V.N., Denisova E.A., Gibson D.G., Assad-Garcia N., Alperovich N., Thomas D.W., Merryman C., Hutchison C.A., 3rd, Smith H.O., Venter J.C. et Glass J.I. (2009). Creating Bacterial Strains from Genomes That Have Been Cloned and Engineered in Yeast. *Science*.

Le Grand D., Arcangioli M.-A., Calavas D., Bézille P. et Poumarat F. (2008). Mycoplasmes et mycoplasmoses bovines : actualités. *Bulletin de l'Académie vétérinaire de France* **161:**159-166.

Lysnyansky I., Sachse K., Rosenbusch R., Levisohn S. et Yogev D. (1999). The *vsp* locus of *Mycoplasma bovis*: gene organization and structural features. *J Bacteriol* **181:**5734-5741.

Mahillon J. et Chandler M. (1998). Insertion sequences. *Microbiol Mol Biol Rev* **62:**725-774.

Mahillon J., Leonard C. et Chandler M. (1999). IS elements as constituents of bacterial genomes. *Res Microbiol* **150:**675-687.

Maniloff J. (2002). Phylogeny and Evolution, p. 31–43. *In* S. Razin and R. Herrmann (ed.), Molecular biology and pathogenicity of Mycoplasmas. *Kluwer Academic / Plenum Publishers*, New York (USA).

Marenda M., Barbe V., Gourgues G., Mangenot S., Sagne E. et Citti C. (2006). A new integrative conjugative element occurs in *Mycoplasma agalactiae* as chromosomal and free circular forms. *J Bacteriol* **188:**4137-4141.

Marenda M.S., Vilei E.M., Poumarat F., Frey J. et Berthelot X. (2004). Validation of the suppressive subtractive hybridization method in *Mycoplasma agalactiae* species by the comparison of a field strain with the type strain PG2. *Vet Res* **35:**199-212.

Margulies M., Egholm M., Altman W.E., Attiya S., Bader J.S., Bemben 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. et Rothberg J.M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437:**376-380.

Markham P.F. et Noormohammadi A.H. (2005). Diagnosis of Mycoplasmosis in Animals, p. 355-382. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: Molecular biology, pathogenicity and strategies for control. *Horizon Bioscience*, Norfolk (United Kingdom).

McAuliffe L., Churchward C.P., Lawes J.R., Loria G., Ayling R.D. et Nicholas R.A. (2008). VNTR analysis reveals unexpected genetic diversity within *Mycoplasma agalactiae*, the main causative agent of contagious agalactia. *BMC Microbiol* **8:**193.

Meinke A., Henics T., Hanner M., Minh D.B. et Nagy E. (2005). Antigenome technology: a novel approach for the selection of bacterial vaccine candidate antigens. *Vaccine* **23:**2035-2041.

Meinke A., Storm M., Henics T., Gelbmann D., Prustomersky S., Kovacs Z., Minh D.B., Noiges B., Stierschneider U., Berger M., von Gabain A., Engstrand L. et Nagy E. (2009). Composition of the ANTIGENome of *Helicobacter pylori* defined by human serum antibodies. *Vaccine* **27:**3251-3259.

Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E., Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C., Davidsen T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R., Mohammoud Y., Nelson W.C., Radune D., Romero C.M., Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N., Zhou L. et Fraser C.M. (2004). Structural flexibility in the *Burkholderia mallei* genome. *Proc Natl Acad Sci U S A* **101:**14246-14251.

Nussbaum S., Lysnyansky I., Sachse K., Levisohn S. et Yogev D. (2002). Extended repertoire of genes encoding variable surface lipoproteins in *Mycoplasma bovis* strains. *Infect Immun* **70:**2220-2225.

Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N., Harris D.E., Holden M.T., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A., Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I., Chillingworth

T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C., Rabbinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G. et Maskell D.J. (2003). Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet* **35:**32-40.

Petrosino J.F., Xiang Q., Karpathy S.E., Jiang H., Yerrapragada S., Liu Y., Gioia J., Hemphill L., Gonzalez A., Raghavan T.M., Uzman A., Fox G.E., Highlander S., Reichard M., Morton R.J., Clinkenbeard K.D. et Weinstock G.M. (2006). Chromosome rearrangement and diversification of *Francisella tularensis* revealed by the type B (OSU18) genome sequence. *J Bacteriol* **188:**6977-6985.

Pettersson E., Lundeberg J. et Ahmadian A. (2009). Generations of sequencing technologies. *Genomics* **93:**105-111.

Samad A., Huff E.F., Cai W. et Schwartz D.C. (1995). Optical mapping: a novel, single-molecule approach to genomic analysis. *Genome Res* **5:**1-4.

Schork N.J., Fallin D. et Lanchbury J.S. (2000). Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin Genet* **58:**250-264.

Schwartz D.C., Li X., Hernandez L.I., Ramnarain S.P., Huff E.J. et Wang Y.K. (1993). Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. *Science* **262:**110-114.

Seth-Smith H. et Croucher N.J. (2009). Genome watch: breaking the ICE. *Nat Rev Microbiol* **7:**328-329.

Sirand-Pugnet P., Citti C., Barre A. et Blanchard A. (2007a). Evolution of mollicutes: down a bumpy road with twists and turns. *Res Microbiol* **158:**754-766.

Sirand-Pugnet P., Lartigue C., Marenda M., Jacob D., Barre A., Barbe V., Schenowitz C., Mangenot S., Couloux A., Segurens B., de Daruvar A., Blanchard A. et Citti C. (2007b). Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS Genet* **3:**e75.

Sreevatsan S., Pan X., Stockbauer K.E., Connell N.D., Kreiswirth B.N., Whittam T.S. et Musser J.M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A* **94:**9869-9874.

Stinear T.P., Seemann T., Pidot S., Frigui W., Reysset G., Garnier T., Meurice G., Simon D., Bouchier C., Ma L., Tichit M., Porter J.L., Ryan J., Johnson P.D., Davies J.K., Jenkin G.A., Small P.L., Jones L.M., Tekaia F., Laval F., Daffe M., Parkhill J. et Cole S.T. (2007). Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Res* **17:**192-200.

Teachman A.M., French C.T., Yu H., Simmons W.L. et Dybvig K. (2002). Gene transfer in *Mycoplasma pulmonis*. *J Bacteriol* **184:**947-951.

Tettelin H., Masignani V., Cieslewicz M.J., Donati C., Medini D., Ward N.L., Angiuoli S.V., Crabtree J., Jones A.L., Durkin A.S., Deboy R.T., Davidsen T.M., Mora M., Scarselli M., Margarit y Ros I., Peterson J.D., Hauser C.R., Sundaram J.P., Nelson W.C., Madupu R., Brinkac L.M.,

Dodson R.J., Rosovitz M.J., Sullivan S.A., Daugherty S.C., Haft D.H., Selengut J., Gwinn M.L., Zhou L., Zafar N., Khouri H., Radune D., Dimitrov G., Watkins K., O'Connor K.J., Smith S., Utterback T.R., White O., Rubens C.E., Grandi G., Madoff L.C., Kasper D.L., Telford J.L., Wessels M.R., Rappuoli R. et Fraser C.M. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc Natl Acad Sci U S A* **102:**13950-13955.

Tettelin H., Riley D., Cattuto C. et Medini D. (2008). Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* **11:**472-477.

Verbisck-Bucker G., Gonzalez-Candela M., Galian J., Cubero-Pablo M.J., Martin-Atance P. et Leon-Vizcaino L. (2008). Epidemiology of *Mycoplasma agalactiae* infection in free-ranging Spanish ibex (*Capra pyrenaica*) in Andalusia, southern Spain. *J Wildl Dis* **44:**369-380.

Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X., Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S., Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y., Qiang B., Hou Y., Yu J. et Jin Q. (2005). Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Res* **33:**6445-6458.

ANNEXES

Listes des publications et communications

Articles publiés dans des revues internationales à comité de lecture

- Nouvel LX, Marenda M, Sirand-Pugnet P, Sagné E, Glew M, Mangenot S, Barbe V, Barré A, Claverol et Citti C.
  Occurrence, plasticity and evolution of the vpma gene family, a genetic system devoted to high frequency surface variation in Mycoplasma agalactiae.
  Journal of Bacteriology. July 2009, p. 4111–4121 191 13

Communications orales

- Nouvel LX, Sagne E, Marenda M, Poumarat F et Citti C.
  Etude de la diversité génétique d'isolats de *Mycoplasma agalactiae*. In : *7ème Rencontres Francophones de Mycoplasmologie*, ENVL, Lyon, France. 17-18 Juillet 2007.
- Nouvel LX.
  Diversité génétique des souches de *Mycoplasma agalactiae* à l'échelle des Pyrénées-Atlantiques. In : *Coordination nationale des recherches sur l'agalactie contagieuse et les mycoplasmoses des petits ruminants*, ENVT, Toulouse, France. 21 Novembre 2006.
- Citti C, Nouvel LX, Marenda M, Barbe V, Mangenot S, Claverol S, Sirand-Pugnet P et Blanchard A.
  Comparing the genome and proteome of two *Mycoplasma agalactiae* strains: clues to the macro- and micro-events that are shaping mycoplasma diversity. In : *The 17th International Organisation for Mycoplasmology Congress*, Tianjin, Chine. 6-11 juillet 2008.
- C. Citti, Nouvel LX, Baranowski E, Sirand-Pugnet P et Blanchard A.
  Mycoplasma agalactiae decoded. Mycoplasmology: A review of developments over the last decade. 10-12th June, Gran Canaria, Spain.
- C. Citti, Nouvel LX, Baranowski E, Sirand-Pugnet P et Blanchard A.
  Occurrence of genomic islands in minimal, wall-less pathogens: a means for preventing evolutionary dead-end and for enhancing adaptation? 3rd Congress of European Microbiologists (FEMS 2009), Gothenburg, Sweden, June 28-July 2, 2009.

Communications écrites (posters)

- Nouvel LX, Marenda M, Sirand-Pugnet P, Barbe V, Claverol S et Citti C
  Anatomy of a hyper-variable genetic system in the *Mycoplasma agalactiae* species: extent, dynamics and evolution In : *The 17th International Organisation for Mycoplasmology congress*, Tianjin, Chine. 6-11 Juillet 2008.

Article publié

JOURNAL OF BACTERIOLOGY, July 2009, p. 4111–4121 Vol. 191, No. 13
0021-9193/09/$08.00+0 doi:10.1128/JB.00251-09
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

Occurrence, Plasticity, and Evolution of the *vpma* Gene Family, a Genetic System Devoted to High-Frequency Surface Variation in *Mycoplasma agalactiae*▽†

Laurent-Xavier Nouvel,[1,2] Marc Marenda,[1,2]‡ Pascal Sirand-Pugnet,[3,4] Eveline Sagné,[1,2] Michelle Glew,[1,2]§ Sophie Mangenot,[5] Valérie Barbe,[5] Aurélien Barré,[6] Stéphane Claverol,[7] and Christine Citti[1,2]*

Université de Toulouse, ENVT, UMR 1225, F-31076 Toulouse, France[1]; INRA, UMR 1225, F-31076 Toulouse, France[2]; Université de Bordeaux, UB2, UMR1090, F-33076 Bordeaux, France[3]; INRA, UMR 1090, F-33883 Villenave d'Ornon, France[4]; CEA-IG, Genoscope, F-91057 Evry Cedex, France[5]; Centre de Bioinformatique de Bordeaux, UB2, F-33076 Bordeaux, France[6]; and Pole Protéomique, Plateforme Génomique Fonctionnelle Bordeaux, UB2, 33076 Bordeaux, France[7]

Received 25 February 2009/Accepted 10 April 2009

***Mycoplasma agalactiae*, an important pathogen of small ruminants, exhibits a very versatile surface architecture by switching multiple, related lipoproteins (Vpmas) on and off. In the type strain, PG2, Vpma phase variation is generated by a cluster of six *vpma* genes that undergo frequent DNA rearrangements via site-specific recombination. To further comprehend the degree of diversity that can be generated at the *M. agalactiae* surface, the *vpma* gene repertoire of a field strain, 5632, was analyzed and shown to contain an extended repertoire of 23 *vpma* genes distributed between two loci located 250 kbp apart. Loci I and II include 16 and 7 *vpma* genes, respectively, with all *vpma* genes of locus II being duplicated at locus I. Several Vpmas displayed a chimeric structure suggestive of homologous recombination, and a global proteomic analysis further indicated that at least 13 of the 16 Vpmas can be expressed by the 5632 strain. Because a single promoter is present in each *vpma* locus, concomitant Vpma expression can occur in a strain with duplicated loci. Consequently, the number of possible surface combinations is much higher for strain 5632 than for the type strain. Finally, our data suggested that insertion sequences are likely to be involved in 5632 *vpma* locus duplication at a remote chromosomal position. The role of such mobile genetic elements in chromosomal shuffling of genes encoding major surface components may have important evolutionary and epidemiological consequences for pathogens, such as mycoplasmas, that have a reduced genome and no cell wall.**

Bacteria of the *Mycoplasma* genus belong to the class *Mollicutes* and represent a remarkable group of organisms that derived from the *Firmicutes* lineage by massive genome reduction (41, 51). Consequent to this regressive evolution, modern mycoplasmas have been left with small genomes (580 to 1,400 kb), a limited number of metabolic pathways, and no cell wall. Due to these particularities, members of the *Mycoplasma* genus have often been portrayed as "minimal self-replicating organisms." Despite this apparent simplicity, a large number of mycoplasma species are successful pathogens of humans and a wide range of animals, in which they are known to cause diseases that are often chronic and debilitating (1, 33). The surface of their single membrane is considered a key interface in mediating adaptation and survival in the context of a complex, immunocompetent host (10, 13, 34, 40). Indeed, mycoplasmas possess a highly versatile surface architecture due to a number of sophisticated genetic systems that promote intraclonal variation in the expression and structure of abundant surface lipoproteins (9, 50). Usually, these systems combine a set of contingency genes with a molecular switch for turning expression on or off that is based on either (i) spontaneous mutation (slipped-strand mispairing), (ii) gene conversion, or (iii) specific DNA rearrangements (9). While high-frequency phenotypic variation using the two first mechanisms has been described thoroughly for other bacteria (47), switching of surface components by shuffling of silent genes at a particular single expression locus has been studied mainly in mycoplasmas (3, 8, 14, 16, 23, 39, 43).

Mycoplasma agalactiae, an important pathogen responsible for contagious agalactia in small ruminants (listed by the World Organisation for Animal Health), possesses a family of lipoproteins encoded by the *vpma* genes for which phase variation in expression is driven by a "cut-and-paste" mechanism involving a tyrosine site-specific recombinase designated Xer1 (16). Data previously gathered with the PG2 type strain identified a single *vpma* cluster (42) composed of six *vpma* genes adjacent to one *xer1* gene (Fig. 1A). Based on fine genetic analyses, Xer1 was further shown to mediate frequent site-specific DNA rearrangements by targeting short DNA sequences located upstream of each *vpma* gene (8, 16). While some *vpma* rearrangements can be phenotyp-

* Corresponding author. Mailing address: UMR 1225 INRA, ENVT, Ecole Nationale Vétérinaire de Toulouse, 23 Chemin des Capelles, BP 87614, 31076 Toulouse Cedex 3, France. Phone: 33.(0)5.61.19.38.56. Fax: 33.(0)5.61.19.32.73. E-mail: c.citti@envt.fr.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: School of Veterinary Science, 250 Princes Highway, Werribee, Victoria 3030, Australia.

§ Present address: Bio21 Institute, Melbourne Dental School, 30 Flemington Rd., Parkville 3010, Australia.

▽ Published ahead of print on 17 April 2009.

Downloaded from jb.asm.org at INRA on June 16, 2009

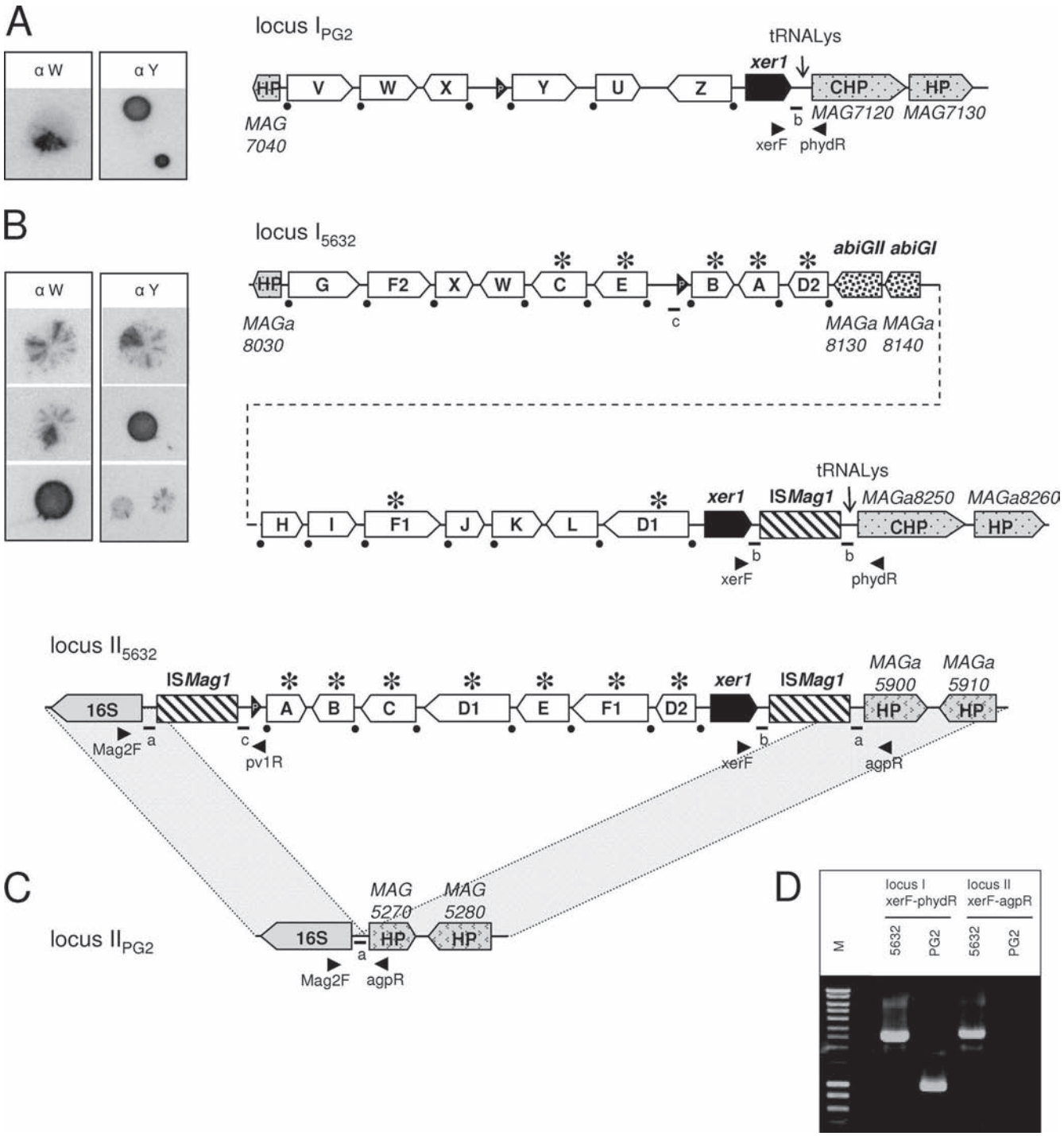


FIG. 1. Comparison of *M. agalactiae vpma* loci between the PG2 type strain and strain 5632. Schematics represent the organization of the *vpma* loci in clonal variant 55.5 derived from PG2 (16, 42) (A) and in clonal variant c1 derived from strain 5632 (B). (C) Counterpart of locus II_{5632} in PG2 showing the absence of *vpma* genes in this region. (D) The presence of two distinct loci in 5632 was confirmed by PCR, using the primer pair xerF-phydR or xerF-agpR, and the resulting amplicons are shown. The locations of the primers are indicated by arrowheads in panels A, B, and C. Large white arrows labeled with letters represent Vpma CDSs. The positions of the promoters are represented by black arrowheads labeled "P." The two non-Vpma-related CDSs ($abiG_I$ and $abiG_{II}$) are indicated by large arrows filled with a dotted pattern. IS*Mag1* elements are indicated by hatched boxes. Recombination sites downstream of each *vpma* gene are indicated by black dots. An asterisk indicates that the corresponding *vpma* gene is present at two distinct loci. Schematics were drawn approximately to scale. HP, hypothetical protein; CHP, conserved hypothetical protein. Small letters and bars indicate the positions of short particular sequences mentioned in the text and in Fig. 3 and 4. The pictures on the left side of panels A and B illustrate the variable surface expression of Vpma, as previously described (8, 17). These correspond to colony immunoblots using Vpma-specific polyclonal antibodies recognizing PG2 VpmaW (α W) and VpmaY (α Y) epitopes.

Downloaded from jb.asm.org at INRA on June 16, 2009

ically silent, others result in Vpma on-off switching by linking a silent *vpma* gene sequence immediately downstream of the unique *vpma* promoter. Because site-specific recombination can be reciprocal, the initial *vpma* configuration can be restored without a loss of genetic information.

The *vsa* family of the murine pathogen *M. pulmonis* (3, 39), the *vsp* family of the bovine pathogen *M. bovis* (2, 24), and the *vpma* family of *M. agalactiae* all generate intraclonal surface diversity by using very similar molecular switches (23), although their overall coding sequences seem to be specific to the *Mycoplasma* species. DNA rearrangements also govern phase variation of the 38 *mpl* genes of the human pathogen *M. penetrans* (27, 35, 38). However, in this mycoplasma species the molecular switch is slightly different, since each *mpl* gene possesses its own invertible promoter (19). In *M. penetrans*, the individual expression of each *mpl* gene can then be switched on and off in a combinatory manner, resulting in a large number of possible Mpl surface configurations. Since *M. pulmonis*, *M. bovis*, and *M. agalactiae* all belong to the *Mycoplasma hominis* phylogenetic cluster (48) and are relatively closely related, while *M. penetrans* belongs to the distinct *Mycoplasma pneumoniae* phylogenetic cluster (30, 48), it is tempting to speculate that the *vsa*, *vsp*, and *vpma* systems were all inherited from a common ancestor and that the bulk of their coding sequences evolved independently in their respective hosts while the molecular switch mechanism was retained.

In so-called "minimal" bacteria, the occurrence of relatively large genomic portions dedicated to multigene families, with genes encoding phase-variable, related surface proteins, suggests that they serve an important function(s). Data accumulated over the years for several mycoplasma species tend to indicate that one general purpose of these systems is to provide the mycoplasma with a variable shield that modulates surface accessibility in order to escape the host response and to adapt to rapidly changing environments (10, 11, 13, 40, 50). On the other hand, the sequences of phase-variable proteins are relatively conserved within one species but divergent between species, suggesting a more specific role for these molecules.

The role of the Vpma family of *M. agalactiae* has yet to be elucidated, but it was recently shown that Vpma switches in expression occur at a remarkably high rate in vitro (10^{-2} to 10^{-3} per cell per generation) (8, 17). The *vpma* systems described for PG2 (16) and another *M. agalactiae* strain, isolated in Israel (in which Vpmas were designated Avg proteins [14]), both revealed a repertoire of six *vpma* genes and only one promoter, suggesting that in *M. agalactiae* the number of Vpma configurations is limited to six. This contrasts with the situation commonly found in other *Mycoplasma* variable systems, which can offer a larger mosaic of surface architecture because of the concomitant switches of several related surface proteins and/or because of a larger number of phase-variable genes.

To further understand the degree of diversity that can be generated at the surface of *M. agalactiae*, we analyzed the *vpma* gene content of a field strain, 5632, whose genome was recently sequenced by our group (unpublished data). The present study shows that 5632 contains a total of 23 *vpma* genes distributed in two distinct loci that both contain a recombinase gene. Further genomic and proteomic analyses indicated that the capacity of 5632 to vary its Vpma surface architecture is far more complex than that described for the type strain. Unlike the case for PG2, both 5632 *vpma* loci are associated with several mobile genetic insertion elements (IS) that could play an evolutionary role in the dynamics of *vpma* repertoires, as suggested by data presented here. One 5632 *vpma* locus contains open reading frames (ORFs) that are highly conserved in both *M. bovis*, a closely related bovine mycoplasma, and the phylogenetically distant mycoplasmas of the *M. mycoides* cluster, which are also important ruminant pathogens. Whether these were acquired through evolution or through horizontal transfer is discussed. The present study reveals an additional degree of complexity for the Vpma system and further suggests that some field strains might have more dynamic genomes and a more variable surface than was first estimated (42).

MATERIALS AND METHODS

Bacterial strains, culture conditions, and DNA isolation. *M. agalactiae* type strain PG2, clone 55.5 (16), and strain 5632, clonal variant C1 (26), have been described previously. These strains were isolated independently from goats in Spain. The experiments reported here were all performed with these clonal variants, but for simplicity, we refer to them as strains PG2 and 5632 only. *M. agalactiae* field isolates were kindly provided by F. Poumarat (AFSSA, Lyon, France) (see Table S1 in the supplemental material). Mycoplasmas were propagated in SP4 liquid medium (46) at 37°C, and genomic DNAs were extracted as described elsewhere (7, 37).

Identification of *M. agalactiae vpma* and associated loci. Whole-genome sequencing of strain 5632 (clonal variant C1) was performed as follows. A library of 3-kb inserts (library A) was generated by mechanical shearing of the DNA followed by cloning of the blunt-ended fragments into the pcDNA2.1 (Invitrogen) *Escherichia coli* vector. Two libraries of 25-kb (library B) and 80-kb (library C) inserts were generated by HindIII partial digestion and cloning into the pBeloBAC11 (Caltech) modified *E. coli* vector. The plasmid inserts of 10,752, 3,072, and 768 clones picked from libraries A, B, and C, respectively, were end sequenced by dye terminator chemistry on ABI3730 sequencers. The PHRED/PHRAP/CONSED software package was used for sequence assemblies. Gap closure and quality assessment were performed according to the Bermuda rules, with 10,307 additional sequences. The *vpma* loci of strain 5632 were detected by DNA homology to those previously described for PG2 (positions of the loci on the 5632 genome are given with reference to the first nucleotide of the *dnaA* gene as nucleotide [nt] 1) and were annotated using the CAAT-Box platform (15), with the aid of Artemis software (36) and ACT software (5). The BLAST program suite was used for sequence homology searches of nonredundant databases (http://www.ncbi.nlm.nih.gov/blast/blast.cgi). In order to determine the extent of sequence similarity, alignments between sequences were performed using Needle (Needleman-Wunsch global alignment algorithm) and Water (Smith-Waterman local alignment algorithm) (http://www.ebi.ac.uk/Tools/emboss/align/) software. Designations of 5632 *vpma* genes were based on local and global alignment scores, using all *vpma* gene sequences, including those previously described for PG2. The same name was attributed to a *vpma* gene presenting an amino acid sequence identity of >50% in the global alignment (Needle) and an amino acid sequence identity of >70% in the local alignment (Water). Using these rules, strain 5632 was found to contain the *vpmaW* and *vpmaX* genes also present in PG2. Also, distinct names were given to *vpmaY* and *vpmaF*, although they displayed global identities of >50%, because their local identities were <70%. *vpma* products with highly similar sequences that differed only in the number of C-terminal repeats were considered allelic versions and were designated with the same letter followed by different numbers (i.e., VpmaD1 and VpmaD2 or VpmaF1 and VpmaF2). Finally, *vpmaI* was not considered an allelic version of *vpmaD* because the C-terminal repeated region of its corresponding product displayed a local identity of 64.7% (<70%) with the VpmaD1 and -D2 counterparts.

Direct sequencing of PCR products and of genomic DNA (18, 20) was performed using specific primers (see Table S2 in the supplemental material) at the sequencing facility of UMR 5165 (CNRS, UPS, CHU Purpan, Toulouse, France).

Phylogenetic analyses were performed using MEGA 3.1 (21) and the neighbor-joining tree method. The reliability of the tree nodes was tested by performing 500 bootstrap replicates.

Downloaded from jb.asm.org at INRA on June 16, 2009

PCR assays. PCR assays were performed on an Eppendorf Mastercycler ep-Gradient thermocycler, using 5 ng *M. agalactiae* DNA as a template and using specific primers (see Table S2 in the supplemental material). PCR assays were performed with 25-μl reaction mixtures containing a 0.4 mM concentration of each primer, PCR buffer (with $MgSO_4$; New England Biolabs) at a 1× final concentration, a 200 mM concentration of each deoxynucleoside triphosphate, and 2.5 U *Taq* DNA polymerase (New England Biolabs). Reaction mixtures were subjected to 2 min at 94°C; 30 cycles of 30 s at 94°C, 30 s at 55°C, and 30 s at 72°C; and a final elongation step of 5 min at 72°C. All PCR assays were performed at the unique annealing temperature of 55°C, using primer pair xerF-phydR for specific amplification of locus I, primer pair xerF-agpR or Mag2F-agpR for locus II, primer pair aip1F-aip1R for the *abiGI* gene, and primer pair aip2F-aip2R for the *abiGII* gene. PCR products were analyzed by gel electrophoresis on 1% agarose gels.

Colony immunoblotting and proteomic analyses. Colony immunoblotting was performed as previously described (8). Briefly, nitrocellulose membranes were placed on mycoplasma colonies freshly grown on agar medium and then removed and rinsed three times in TS buffer (10 mM Tris, 154 mM NaCl, pH 7.4). Membranes were then incubated overnight at 4°C with previously described rabbit Vpma-specific antibodies (8), washed three times in TS buffer containing 0.05% Tween 20 (Roth), and then incubated for 1 h at 25°C in a 1:2,000 dilution of swine anti-rabbit immunoglobulin G conjugated to horseradish peroxidase (Dako). After three washes, the colony blots were developed for 15 to 30 min in 4-chloro-1-naphthol and 0.02% hydrogen peroxide. The reaction was stopped by washing the blots in sterile distilled water.

Proteins that partitioned into Triton X-114 were extracted from strain 5632 as previously described (4), precipitated overnight at −70°C after the addition of 9 volumes of cold methanol, centrifuged for 10 min at 12,000 × *g*, and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The gel was sliced into 16 sections which were subjected to trypsin digestion. Peptides were further analyzed by nano-liquid chromatography coupled to ion-trap tandem mass spectrometry (LC-MS/MS).

Peptides were identified with SEQUEST through the Bioworks 3.3.1 interface (Thermo-Finnigan, Torrance, CA), using a database consisting of both direct and reverse-sense *Mycoplasma agalactiae* strain 5632 entries (1,652 entries). Using the following criteria as validation filters, the false-positive rate was null: $\Delta CN \geq 0.1$; Xcorr versus charge state, ≥1.5 (+1), 2.0 (+2), or 2.5 (+3); peptide probability, ≤ 0.001; and number of different peptides, ≥2.

Nucleotide sequence accession numbers. The GenBank accession numbers for locus I_{5632} and locus II_{5632} are FP245515 and FP245514, respectively.

RESULTS AND DISCUSSION

Genome sequencing of strain 5632 reveals an extended *vpma* repertoire. The fully sequenced genome of *M. agalactiae* strain 5632 revealed a total of 23 *vpma* genes (Fig. 1). In contrast to the case for the PG2 type strain, this extended repertoire is distributed on two distinct chromosomal loci, loci I and II, that contain 16 and 7 *vpma* genes, respectively. BLAST analyses identified only two genes having significant similarity with those previously described for PG2 (16, 42), namely, *vpmaW* and *vpmaX*, with the others (*vpmaA* to *vpmaG*) only sharing blocks of high similarity with PG2 *vpma* genes (Fig. 2).

More precisely, locus I_{5632} is 19,453 bp long and is the counterpart of the PG2 *vpma* locus (Fig. 1A) because of their flanking coding sequences (CDSs) being nearly identical. In addition to 16 *vpma* genes, locus I_{5632} contains two CDSs (MAGa8140 and MAGa8130) that have no homology with the Vpma family but were found to have high amino acid similarity with *Streptococcus agalactiae* $abiG_I$ and $abiG_{II}$ gene products (45). The two corresponding genes are lacking in PG2 and were designated $abiG_I$ and $abiG_{II}$, respectively. Locus I_{5632} further differs from its PG2 counterpart by the presence of an IS element corresponding to the previously described IS*Mag1* (31). This IS element is immediately adjacent to the 3′ end of the integrase-recombinase *xer1* gene, whose product is identical to that in PG2.

The second *vpma* locus of 5632, locus II_{5632}, is 8,462 kb long and contains only seven *vpma* genes clustered between two IS*Mag1* copies, themselves flanked by a 16S rRNA gene at the 5′ end and by a CDS annotated as a hypothetical protein (HP) gene (MAGa5900) at the 3′ end. In PG2, this region is highly conserved, except that it does not contain any IS element or *vpma* or any other gene (locus II_{PG2}) (Fig. 1C and below). DNA BLAST analyses of locus II_{5632} demonstrated that all seven *vpma* coding sequences are identical to the seven *vpma* genes of locus I_{5632} (Fig. 1B). Interestingly, an identical *xer1* gene also occurs at both *vpma* loci of strain 5632. Since a single functional gene would be sufficient to generate DNA rearrangements in each locus, the duplication was further confirmed by PCR to rule out any possible artifact of sequence assembling. This was performed using a *xer1*-specific primer in combination with a primer specific to the 3′ region of locus I (phydR) or locus II (agpR) (Fig. 1D).

The Vpma repertoire is composed of common and distinct structures. Detailed analyses of the 23 *vpma* coding regions of strain 5632 showed that (i) 2 are allelic versions of PG2 *vpma* genes (*vpmaW* and *vpmaX*), (ii) 7 are duplicated (*vpmaA*, *-B*, *-C*, *-D1*, *-D2*, *-E*, and *-F1*) and distributed in two *vpma* loci that are 250 kbp apart, and (iii) 2 are allelic copies of *vpmaD* (*vpmaD1* and *-D2*) and *vpmaF* (*vpmaF1* and *-F2*) that vary mainly in the number of C-terminal repeat motifs. Consequently, strain 5632 contains 12 new, distinct Vpma products compared to PG2 (Fig. 2). All *vpma* gene products contain a conserved amino acid signal sequence at the N terminus, followed by a lipobox and an 11-amino-acid conserved sequence (41). Within 5632, some mature *vpma* gene products are unique (VpmaA, -G, and -W), while others share blocks of amino acids between them (VpmaB, -C, -D1, -D2, -E, -F1, -F2, -H, -I, -J, -K, -L, and -X). Also, most Vpmas contain amino acid blocks that are directly repeated (VpmaA, -B, -D1, -D2, -I, -F1, -F2, -G, and -W). These data indicate that the diversity in the Vpma family is represented by a mosaic of structures that is suggestive of homologous recombination events between *vpma* genes, gene duplication, and/or insertion-deletion of repeated motifs. For instance, VpmaF1 and VpmaF2 share a common block of 104 amino acids with VpmaK and VpmaX and a common block of 85 amino acids with VpmaY that is not found in VpmaK or VpmaX. Some Vpmas, such as VpmaD1 and -D2 or VpmaF1 and -F2, represent size variants of the same Vpma with a variable number of repeated motifs at the C terminus (e.g., VpmaD1 and -D2 contain 20 and 2 repeated motifs of 15 amino acids, respectively). Additionally, VpmaI may have a common ancestor with VpmaD1 and -D2 and can be seen as the result of combined events, such as gene duplication followed by sequence drift and/or repeat expansion at the 3′ end.

Anti-VpmaW and anti-VpmaY rabbit antibodies previously produced against Vpma products of strain PG2 (8) also reacted with 5632 surface-exposed epitopes in colony immunoblots (Fig. 1) and revealed the presence of sectored colonies characteristic of high-frequency variation in expression. Most likely, 5632 Vpmas recognized by these antibodies correspond to VpmaW and to VpmaF1 and/or -F2, which share blocks of amino acids with PG2 VpmaY (Fig. 2).

To better define the Vpma products expressed by strain 5632, proteins that partitioned in Triton X-114 detergent were

Downloaded from jb.asm.org at INRA on June 16, 2009

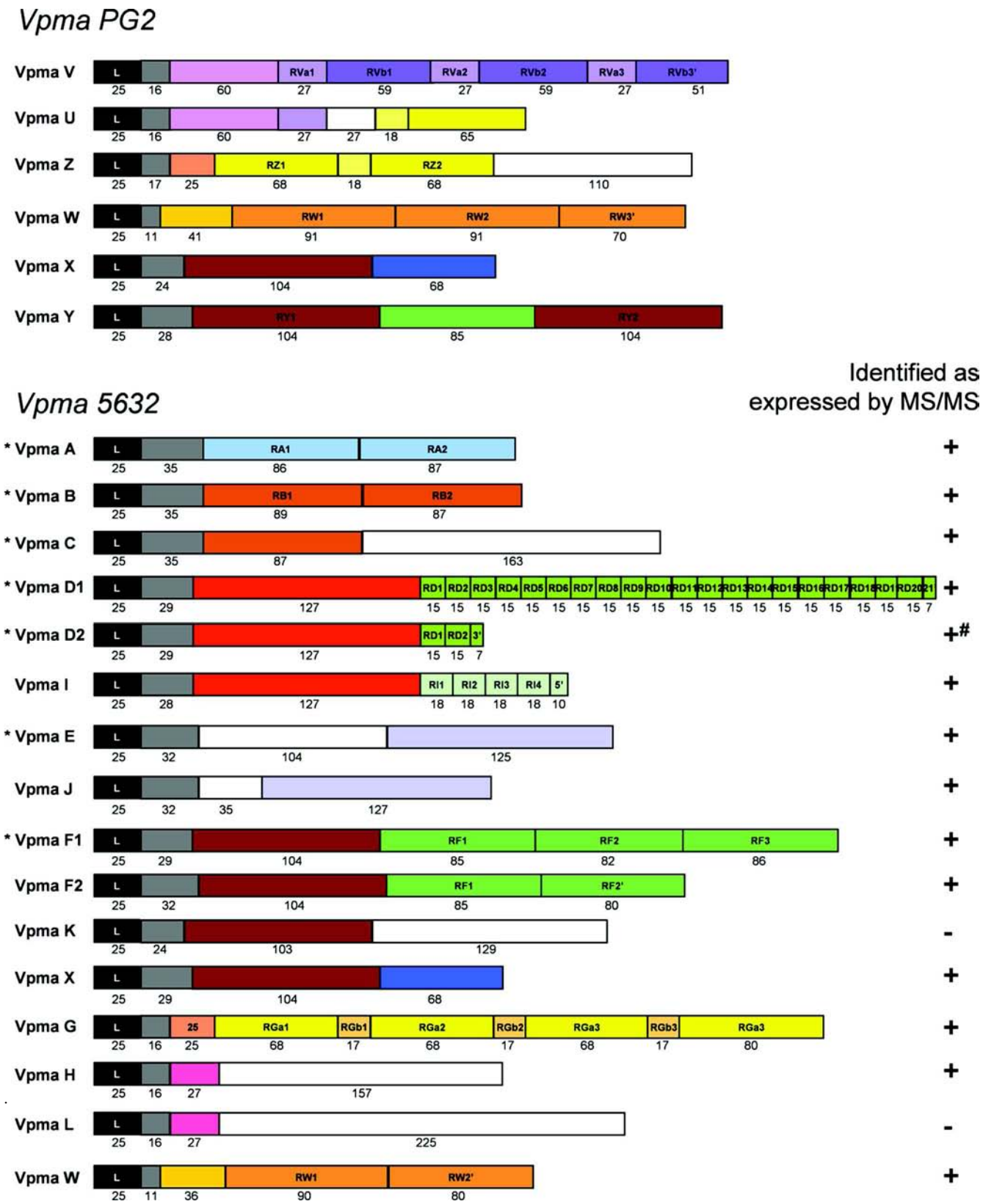


FIG. 2. Structural features and comparison of *vpma* gene products in *M. agalactiae* strains PG2 and 5632. Predicted Vpma proteins are represented schematically by boxes and begin with a homologous 25-amino-acid leader sequence (black boxes) followed by regions that have homology between *vpma* gene products or that are repeated within the same product (colored boxes). Two boxes of the same color display an amino acid identity of >30%. White boxes represent unique sequences. Numbers below boxes indicate the numbers of amino acids. For strain 5632, detection by MS/MS of expressed specific Vpma peptides is noted (+). −, no specific peptides were detected for the corresponding Vpma; $+^{\#}$, VpmaD2 peptides detected are not specific because all are shared with VpmaD1 or -I (see the list of detected peptides in Table S3 in the supplemental material). An asterisk indicates that the corresponding *vpma* gene is present in 5632 at both *vpma* loci.

Downloaded from jb.asm.org at INRA on June 16, 2009

subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by in-gel trypsin digestion and by LC-MS/MS. The results showed that at least two specific peptides were detected for 13 Vpmas (VpmaA, -B, -C, -D1, -E, -F1, -F2, -G, -H, -I, -J, -X, and -W), indicating that these were expressed by 5632 during propagation (see Table S3 in the supplemental material). While VpmaD1-specific peptides were detected, the expression of VpmaD2 could not be assessed because all VpmaD2 trypsin peptides were also carried in VpmaD1 or VpmaI. No peptides specific to VpmaL and VpmaK were detected, although they possess long specific sequences of 225 and 129 amino acids, respectively (Fig. 2). Whether VpmaL and VpmaK are expressed at low levels or under different conditions cannot be determined. Finally, all Vpmas encoded by locus II_{5632} are duplicated at locus I_{5632}, and consequently, their expression cannot be monitored by this method or by any other method, such as Northern blot analysis, because their sequences are also identical at the DNA level.

The global proteomic approach taken above is qualitative and does not reflect the level of expression of each Vpma, two of which are predicted to be expressed at high levels by the clonal population, while others should reflect only a minority of back-switchers. In PG2, the *vpma* locus contains a unique promoter which drives the expression of the *vpma* gene localized immediately downstream, while the other *vpma* genes remain silent (16). This promoter region was more precisely defined by primer extension by Flitman-Tene et al. (14), and a similar sequence was detected in each locus of strain 5632. Therefore, Vpma expression in 5632 is most likely driven at each locus by a single promoter sequence which is highly conserved between the two loci (99% nucleotide identity). In locus I_{5632}, the promoter sequence is located upstream of the *vpmaB* gene, while in locus II_{5632} it is upstream of *vpmaA*, suggesting that these two Vpmas would be coexpressed predominantly by the clonal population. However, in the absence of specific antibodies, this hypothesis cannot be tested. One interesting observation is that the length of a poly(T) tract located between the putative −35 and −10 boxes of the *vpma* promoters varies among the three *vpma* promoters sequenced and is 14, 15, and 12 nt long for loci I of PG2 and 5632 and for locus II of 5632, respectively (see Fig. S1 in the supplemental material). In other variable mycoplasma systems, i.e., the Vlp system of *M. hyorhinis*, fluctuation in the length of the polynucleotide tract between the −35 and −10 boxes has been shown to affect transcription (12). Whether this is also the case for the *M. agalactiae* species, and for strain 5632 in particular, cannot currently be tested because of the complexity of the 5632 *vpma* system. In PG2, DNA rearrangements are mediated by Xer1, which recognizes a specific recombination site (RS) sequence located upstream of each *vpma* gene. Indeed, analysis of the 5632 and PG2 *vpma* loci revealed that a rigorously identical RS sequence is located upstream of each of the 29 *vpma* genes (23 in 5632 and 6 in PG2). This suggests that each *vpma* locus of strain 5632 can independently undergo DNA rearrangements and that the concomitant expression of two distinct Vpmas is most likely to occur frequently in this strain. If so, 5632 could theoretically display 91 different Vpma surface configurations, compared to only 6 for the type strain PG2. In parallel, additional Vpma size variations by insertion or deletion of direct repeats can further participate in altering surface architecture.

TABLE 1. *abiG*-related gene products in *Mollicutes* and other bacteria

Strain	Presence of sequence (% identity, % similarity) related to *Mycoplasma agalactiae* strain 5632 sequence[a]	
	AbiGI-like (MAGa8140)	AbiGII-like (MAGa8130)
M. agalactiae PG2	−	−
M. bovis PG45	+ (84.8, 93.6)	+ (81.1, 91.3)
M. mycoides mycoides SC PG1	+ (40.8, 59.2)	+ (41.8, 64.6)
M. mycoides mycoides LC GM12	+ (41.7, 59.2) (Mmycm_04185)	+ (31.9, 51.6) (Mmycm_04180)
	+ (40.1, 58.5) (Mmycm_00130)	+ (32.6, 50.9) (Mmycm_00135)
Other *Mollicutes* strains	−	−
Streptoccus agalactiae H36B	+ (40.2, 58.3)	+ (41.2, 65.3)
Streptococcus suis 05ZYH33	+ (41.2, 58.3)	+ (40.4, 66.3)
Non-*Streptococcus* spp.	+ (<40, not available)	+ (<40)

[a] +, presence of homologous sequence; −, absence of homologous sequence in public databases (as of June 2008). Identity and similarity percentages between amino acid sequences were obtained using the Needle alignment program (http://mobyle.pasteur.fr/cgi-bin/portal.py).

Evidence of horizontal gene transfer at the *vpma* locus of 5632. As mentioned earlier, *M. bovis* is a close relative of *M. agalactiae* and generates surface variation by using a very similar genetic system, designated the *vsp* gene family. Both the *vsp* and *vpma* systems have several common features, including (i) *vsp* or *vpma* related genes are clustered in close proximity to a conserved recombinase gene (32); (ii) the recombination site involved in DNA rearrangements is identical, except for a single nucleotide polymorphism that is conserved in each species; and (iii) *vpma* or *vsp* genes code for proteins having a signal peptide that is highly conserved between the two systems. Unlike the PG2 *vpma* loci, the *vpma* loci of 5632 are associated with IS elements, as in *M. bovis* (23), and one, locus I_{5632}, contains two non-*vpma* genes, $abiG_I$ and $abiG_{II}$, that have 86.97 and 85.15% overall DNA identity with two *M. bovis* ORFs, ORF4 and ORF5 (23), respectively, located in the *vsp* locus. A search of the databases revealed that $abiG_I$ and $abiG_{II}$ from strain 5632 also match with gene products annotated as conserved hypothetical products in *Mycoplasma mycoides* subsp. *mycoides* biotype LC strain GM12 (22) and *M. mycoides* subsp. *mycoides* SC strain PG1 (49), with >50% amino acid similarity in the global alignment (Table 1). While *abiG* homologs are present in several gram-positive bacteria, such as *Streptococcus agalactiae* and *Streptococcus suis* (6, 45), none was found in other *Mollicutes* sequences available in the databases, which included 21 complete genomes.

The $abiG_I$ and $abiG_{II}$ genes were first described for *Lactococcus lactis* subsp. *cremoris* UC653 (28), where they occur as an operon carried by a conjugative plasmid and confer resistance to phage infection (29). Detection by LC/MS-MS of AbiGII-like specific peptides in the Triton X-114 detergent phase of 5632 indicates that the corresponding $abiG_{II}$ gene is expressed, although there is no hydrophobic domain that could account for this partitioning. Previous attempts to define the functions of $abiG_I$ and $abiG_{II}$ of *Streptococcus* species failed (44), and whether *abiG* homologues found in mycoplasmas have a role in protecting them against phage attack remains to be addressed. Nevertheless, comparison of AbiGI and AbiGII amino acid sequences by constructing

Downloaded from jb.asm.org at INRA on June 16, 2009

neighbor-joining trees (500 bootstraps) indicated a strong relationship between *Mycoplasma* and *Streptococcus* sequences, suggesting that their corresponding genes might have undergone horizontal gene transfer (see Fig. S2 in the supplemental material).

To further define the occurrence of the two *abi* genes in *M. agalactiae*, a collection of 92 strains representative of (i) various geographical areas and (ii) various histories (see Table S1 in the supplemental material) was screened by PCR for the presence of *abiG* genes, using the specific primer pairs aip1F-aip1R and aip2F-aip2R. PCR products corresponding to each gene were obtained for 11 strains of 92 total strains. Further PCR assays showed that like the case for *Streptococcus agalactiae*, the two *abiG* genes always occur next to each other. Interestingly, direct genome sequencing of six *abiG*-positive strains, using aip2F2 as a primer, showed that the $abiG_{II}$ gene is always associated with *vpma* genes. Because strains displaying *abiG* genes in their *vpma* loci have very different geographical origins (Ivory Coast, Spain, and France), this result raised the question of whether these genes were inherited vertically from an *M. agalactiae*/*M. bovis* common ancestor or whether the *vpma* locus is a hot spot for their insertion. Since the *vpma* locus is frequently subjected to DNA rearrangements, deletion of *abiG* genes by intrarecombination involving, for instance, *vpma*-RS sequences, may easily have occurred, resulting in *abiG*-negative strains.

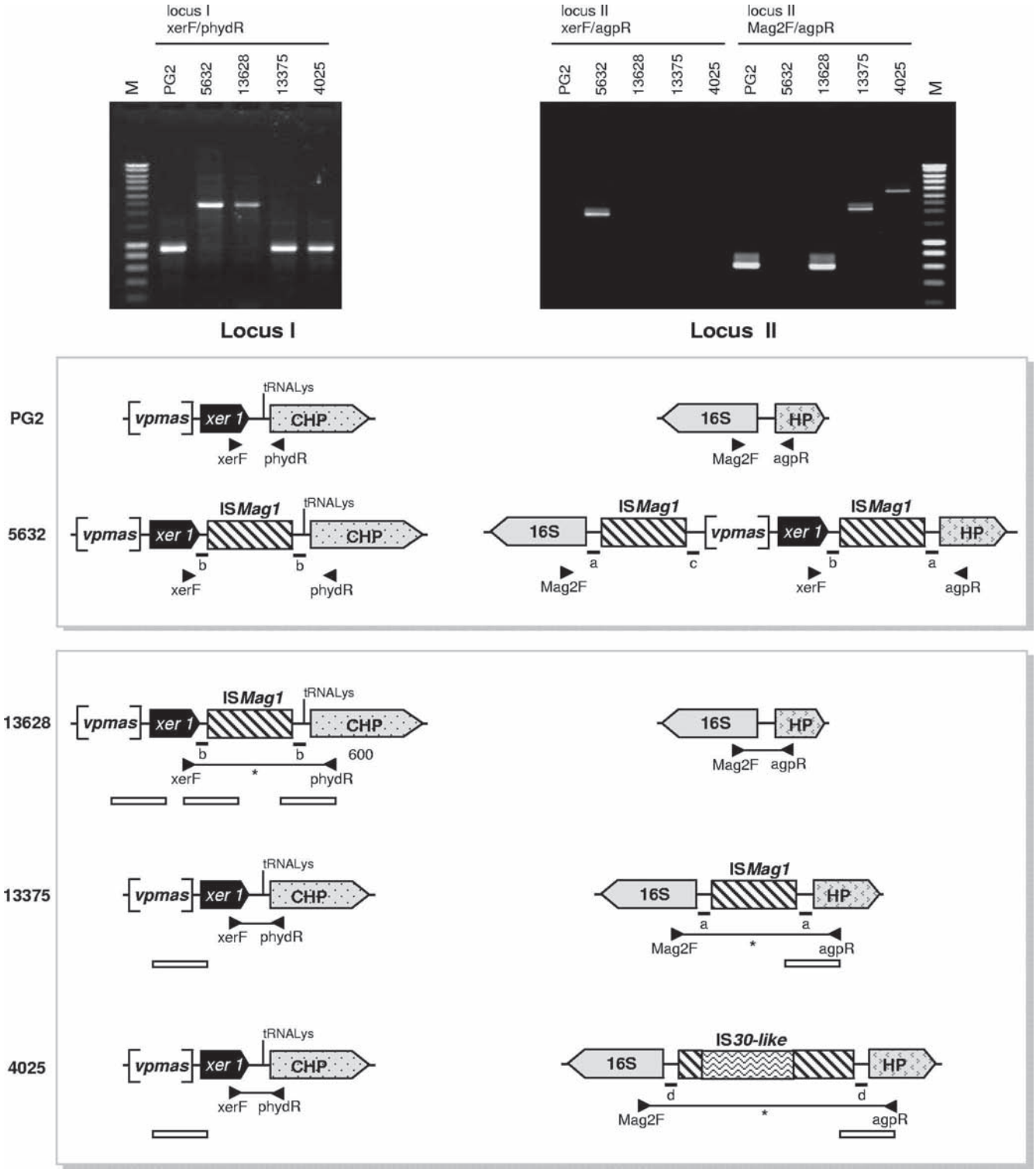
Plasticity of *M. agalactiae vpma* loci. To define whether the occurrence of the two *vpma* loci observed in strain 5632 is common among *M. agalactiae* strains, a PCR assay was performed with primer pairs xerF-phydR, xerF-agpR, and Mag2F-agpR, which yield amplicons corresponding to either locus I or II of PG2 and/or 5632. More specifically, the primer pairs xerF-phydR and xerF-agpR target the 3′ ends of the loci, which are not affected by *vpma* rearrangements, while Mag2F and agpR were designed to amplify a more or less empty locus II under our PCR conditions. The assay was conducted with the panel of 92 strains (see above). The results showed that 89 strains produced an amplification profile identical to that obtained with PG2, while 3 showed new profiles, none of which was identical to that obtained with 5632 (Fig. 3). Sequencing of PCR products or selected genomic regions showed that (i) strain 13628, like 5632, contains an IS element at locus I while displaying a locus II similar to that of PG2; and (ii) strains 13375 and 4025, like PG2, display no IS at locus I while carrying a locus II containing one or more IS but no *vpma* genes, with strain 4025 having an IS*30*-like insert in an IS*Mag1* element. These data suggest that the occurrence of a *vpma* cluster at locus II is a rare event and has so far been observed only in strain 5632, while the occurrence of *vpma* genes at locus I is shared by all *M. agalactiae* strains tested so far.

The presence of an IS*Mag1* element in close proximity to *vpma* locus I in 5632 is not an isolated event, since it is shared by at least one other strain, 13628, which unlike 5632, was isolated in France in 2003 from a goat ear canal. IS*Mag1* belongs to the IS*30* family (31), which generates, when inserting itself into a host genome, a direct repeat that flanks the IS (25). Indeed, IS*Mag1* of locus I is flanked in 5632 and in 13628 by an identical 14-nt directly repeated sequence (Fig. 3), suggesting that its insertion most likely occurred in an ancestor common to the two strains. Interestingly, the 14-nt element found on each side of the IS of locus II_{13375} (Fig. 3) is also repeated in locus II_{5632} as if flanking the whole locus, including the IS. Because no other 14-nt directed repeat is found in the vicinity of the two IS, this suggested the introduction of the *vpma* cluster at locus II_{5632} via a duplication-insertion mechanism involving a circular intermediate carrying a set of *vpma* genes together with an IS*Mag1* element (Fig. 4). In this scenario, the duplicated *vpma* genes would have first clustered at one end of locus I via a series of DNA recombination events. In a second step, part of locus I, including *xer1* and IS*Mag1*, would have been duplicated and circularized to finally recombine at locus II via a single crossover between the IS elements. One argument in favor of the occurrence of the circular form is that a PCR product was obtained with strain 5632, using outwardly oriented primers pv1R and xerF, and its partial sequencing revealed the presence of an IS*Mag1* element next to a sequence "c," usually found next to the promoter (Fig. 4). Whether this result really reflects a duplication-excision event of part of *vpma* locus I and whether the circular form can reinsert itself elsewhere in the genome have still to be demonstrated formally. Nevertheless, the 100% nucleotide identity observed between loci I and II of 5632 indicates that this duplication event occurred relatively recently.

Conclusions. A number of sophisticated genetic systems devoted to high-frequency surface variation have been described for mycoplasma species. Many of these systems appear as small sets of species-specific genes, which are nearly identical in all of the isolates and are clustered at one locus. The clustering of these genes has been proposed to facilitate recombination events which generate diversity, and theoretically, a wide range of surface diversity can be produced with only a few genes, particularly as they are prone to internal rearrangements of their sequences. Our characterization of a new set of *vpma* genes in an *M. agalactiae* field isolate (5632) illustrates this point. This new repertoire is far more complex than that first described for the PG2 type strain and represents a mosaic of structures that is suggestive of homologous recombination events between *vpma* genes, gene duplication, and/or insertion-deletion of repeated motifs (see Fig. S3 in the supplemental material). This opens the question of the true diversity in equivalent variable systems from mycoplasmas with apparently few variable genes, most of which have been sequenced from type strains. Field or clinical isolates have the potential to reveal greater-than-suspected diversity in variable gene repertoires, which could be useful to design molecular probes for epidemiology studies.

One of the most interesting findings of our study was the occurrence of duplicated *vpma* loci in strain 5632. Because one *vpma* locus can express only a single *vpma* gene per cell, the consequences of such an event are tremendous in terms of variability because it allows concomitant expression and, in turn, multiplies the number of possible surface combinations. Our data suggest that this situation most likely resulted from the duplication of one 5632 *vpma* region at a remote chromosomal position (250 kbp away) and that it involves IS elements. The role of such mobile genetic elements in chromosomal shuffling of genes encoding major surface components may have important evolutionary and epidemiological

Downloaded from jb.asm.org at INRA on June 16, 2009


locus I
xerF/phydR
locus II
xerF/agpR
locus II
Mag2F/agpR
M
PG2
5632
13628
13375
4025
Locus I
Locus II
PG2
5632
13628
13375
4025
tRNALys
vpmas
xer 1
CHP
HP
16S
ISMag1
ISMag1
ISMag1

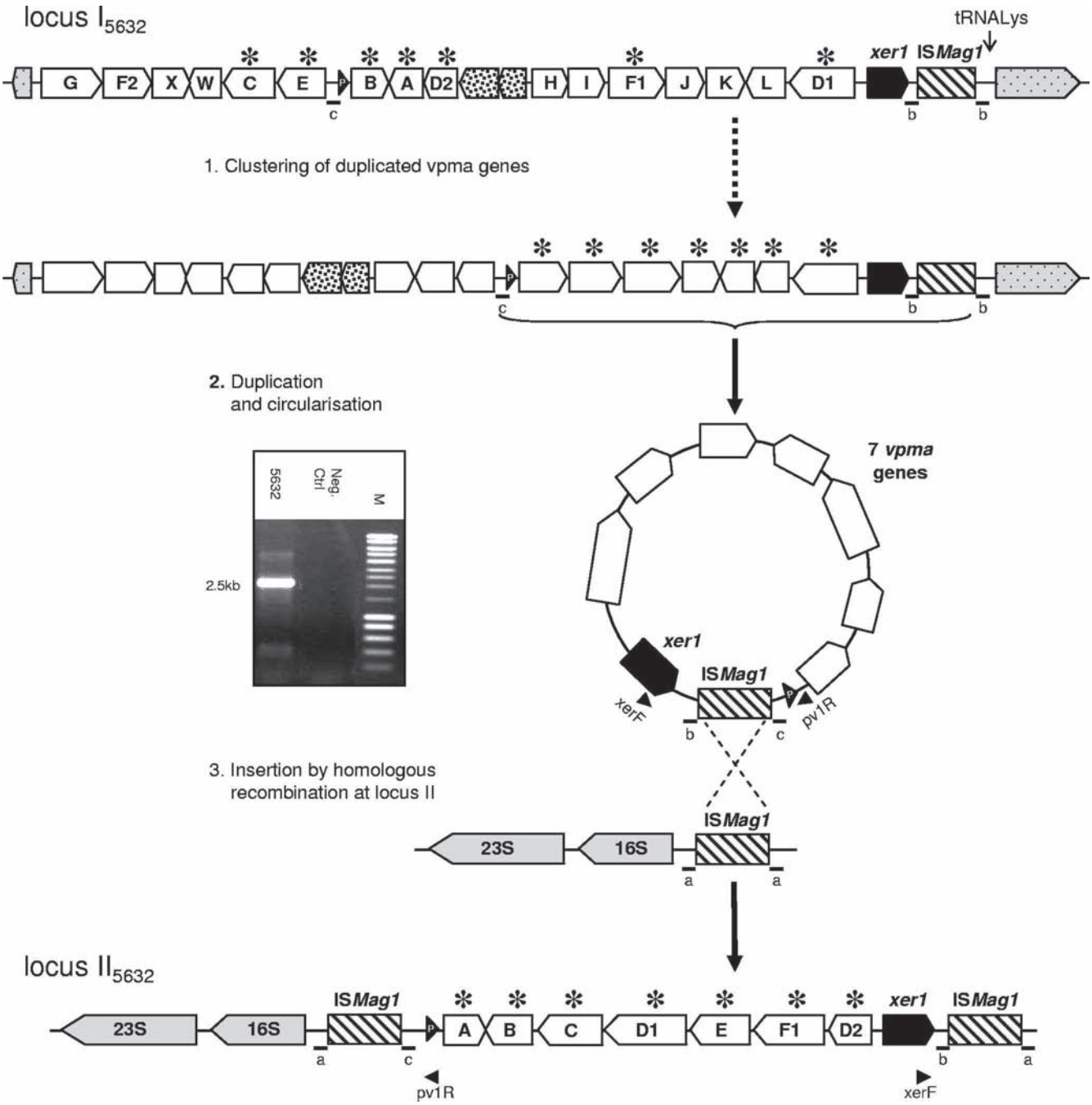


FIG. 4. Scenario of duplication-insertion of strain 5632 *vpma* genes from locus I to locus II via a circular intermediate. Schematics illustrate a putative scenario in which the *vpma* cluster of 5632 at locus II originates from locus I by a mechanism of duplication-insertion involving a circular intermediate and IS elements. Small bars with a letter below represent 14-nt sequences (a, b, and c) flanking the IS*Mag1* element either after an insertion event (a and b) or after a homologous recombination event (c). The dotted arrow indicates the clustering process for seven *vpma* genes at one end of locus I, starting from the current configuration. It represents a series of an unknown number of possible DNA rearrangements. Asterisks indicate duplicated *vpma* genes. The gel picture shows PCR amplification of the circular intermediate obtained using primers xerF and pv1R, with strain 5632 crude DNA extract as the template.

consequences for pathogens, such as mycoplasmas, that have a reduced genome and no cell wall.

Finally, whether a successful vaccine strategy based on surface antigens could be developed for mycoplasmas is not known, but this would have to take into account the dynamics and scope of any variable antigen gene repertoires.

ACKNOWLEDGMENTS

This work was supported by the French Ministry of Agriculture and Fisheries and by the National Institute for Agricultural Research (INRA).

We thank F. Poumarat and the French National Network for Surveillance of Mycoplasmosis (AFSSA, Lyon, France, and VIGIMYC,

Downloaded from jb.asm.org at INRA on June 16, 2009

France) for providing the strains. We also thank A. Blanchard for helpful discussions.

REFERENCES

1. **Baseman, J. B., and J. G. Tully.** 1997. Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. Emerg. Infect. Dis. **3:**21–32.
2. **Behrens, A., M. Heller, H. Kirchhoff, D. Yogev, and R. Rosengarten.** 1994. A family of phase- and size-variant membrane surface lipoprotein antigens (Vsps) of *Mycoplasma bovis*. Infect. Immun. **62:**5075–5084.
3. **Bhugra, B., L. L. Voelker, N. Zou, H. Yu, and K. Dybvig.** 1995. Mechanism of antigenic variation in Mycoplasma pulmonis: interwoven, site-specific DNA inversions. Mol. Microbiol. **18:**703–714.
4. **Bordier, C.** 1981. Phase separation of integral membrane proteins in Triton X-114 solution. J. Biol. Chem. **256:**1604–1607.
5. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423.
6. **Chen, C., J. Tang, W. Dong, C. Wang, Y. Feng, J. Wang, F. Zheng, X. Pan, D. Liu, M. Li, Y. Song, X. Zhu, H. Sun, T. Feng, Z. Guo, A. Ju, J. Ge, Y. Dong, W. Sun, Y. Jiang, J. Yan, H. Yang, X. Wang, G. F. Gao, R. Yang, and J. Yu.** 2007. A glimpse of streptococcal toxic shock syndrome from comparative genomics of S. suis 2 Chinese isolates. PLoS ONE **2:**e315.
7. **Chen, W. P., and T. T. Kuo.** 1993. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. Nucleic Acids Res. **21:**2260.
8. **Chopra-Dewasthaly, R., C. Citti, M. D. Glew, M. Zimmermann, R. Rosengarten, and W. Jechlinger.** 2008. Phase-locked mutants of Mycoplasma agalactiae: defining the molecular switch of high-frequency Vpma antigenic variation. Mol. Microbiol. **67:**1196–1210.
9. **Citti, C., G. F. Browning, and R. Rosengarten.** 2005. Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas, p. 439–484. *In* A. Blanchard and G. F. Browning (ed.), Mycoplasmas: molecular biology, pathogenicity and strategies for control. Horizon Bioscience, Norfolk, United Kingdom.
10. **Citti, C., M. F. Kim, and K. S. Wise.** 1997. Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth-inhibiting host antibodies. Infect. Immun. **65:**1773–1785.
11. **Citti, C., and R. Rosengarten.** 1997. Mycoplasma genetic variation and its implication for pathogenesis. Wien Klin. Wochenschr. **109:**562–568.
12. **Citti, C., and K. S. Wise.** 1995. Mycoplasma hyorhinis vlp gene transcription: critical role in phase variation and expression of surface lipoproteins. Mol. Microbiol. **18:**649–660.
13. **Denison, A. M., B. Clapper, and K. Dybvig.** 2005. Avoidance of the host immune system through phase variation in *Mycoplasma pulmonis*. Infect. Immun. **73:**2033–2039.
14. **Flitman-Tene, R., S. Mudahi-Orenstein, S. Levisohn, and D. Yogev.** 2003. Variable lipoprotein genes of *Mycoplasma agalactiae* are activated in vivo by promoter addition via site-specific DNA inversions. Infect. Immun. **71:**3821–3830.
15. **Frangeul, L., P. Glaser, C. Rusniok, C. Buchrieser, E. Duchaud, P. Dehoux, and F. Kunst.** 2004. CAAT-box, contigs-assembly and annotation tool-box for genome sequencing projects. Bioinformatics **20:**790–797.
16. **Glew, M. D., M. Marenda, R. Rosengarten, and C. Citti.** 2002. Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus. J. Bacteriol. **184:**5987–5998.
17. **Glew, M. D., L. Papazisi, F. Poumarat, D. Bergonier, R. Rosengarten, and C. Citti.** 2000. Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*. Infect. Immun. **68:**4539–4548.
18. **Heiner, C. R., K. L. Hunkapiller, S. M. Chen, J. I. Glass, and E. Y. Chen.** 1998. Sequencing multimegabase-template DNA with BigDye terminator chemistry. Genome Res. **8:**557–561.
19. **Horino, A., Y. Sasaki, T. Sasaki, and T. Kenri.** 2003. Multiple promoter inversions generate surface antigenic variation in *Mycoplasma penetrans*. J. Bacteriol. **185:**231–242.
20. **Hudson, P., T. S. Gorton, L. Papazisi, K. Cecchini, S. Frasca, Jr., and S. J. Geary.** 2006. Identification of a virulence-associated determinant, dihydrolipoamide dehydrogenase (*lpd*), in *Mycoplasma gallisepticum* through in vivo screening of transposon mutants. Infect. Immun. **74:**931–939.
21. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. **5:**150–163.
22. **Lartigue, C., J. I. Glass, N. Alperovich, R. Pieper, P. P. Parmar, C. A. Hutchison, III, H. O. Smith, and J. C. Venter.** 2007. Genome transplantation in bacteria: changing one species to another. Science **317:**632–638.
23. **Lysnyansky, I., Y. Ron, and D. Yogev.** 2001. Juxtaposition of an active promoter to *vsp* genes via site-specific DNA inversions generates antigenic variation in *Mycoplasma bovis*. J. Bacteriol. **183:**5698–5708.
24. **Lysnyansky, I., R. Rosengarten, and D. Yogev.** 1996. Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. J. Bacteriol. **178:**5395–5401.
25. **Mahillon, J., C. Leonard, and M. Chandler.** 1999. IS elements as constituents of bacterial genomes. Res. Microbiol. **150:**675–687.
26. **Marenda, M. S., E. M. Vilei, F. Poumarat, J. Frey, and X. Berthelot.** 2004. Validation of the suppressive subtractive hybridization method in Mycoplasma agalactiae species by the comparison of a field strain with the type strain PG2. Vet. Res. **35:**199–212.
27. **Neyrolles, O., I. Chambaud, S. Ferris, M. C. Prevost, T. Sasaki, L. Montagnier, and A. Blanchard.** 1999. Phase variations of the *Mycoplasma penetrans* main surface lipoprotein increase antigenic diversity. Infect. Immun. **67:**1569–1578.
28. **O'Connor, L., A. Coffey, C. Daly, and G. F. Fitzgerald.** 1996. AbiG, a genotypically novel abortive infection mechanism encoded by plasmid pCI750 of *Lactococcus lactis* subsp. *cremoris* UC653. Appl. Environ. Microbiol. **62:**3075–3082.
29. **O'Connor, L., M. Tangney, and G. F. Fitzgerald.** 1999. Expression, regulation, and mode of action of the AbiG abortive infection system of *Lactococcus lactis* subsp. *cremoris* UC653. Appl. Environ. Microbiol. **65:**330–335.
30. **Pettersson, B., M. Uhlen, and K. E. Johansson.** 1996. Phylogeny of some mycoplasmas from ruminants based on 16S rRNA sequences and definition of a new cluster within the hominis group. Int. J. Syst. Bacteriol. **46:**1093–1098.
31. **Pilo, P., B. Fleury, M. Marenda, J. Frey, and E. M. Vilei.** 2003. Prevalence and distribution of the insertion element ISMag1 in Mycoplasma agalactiae. Vet. Microbiol. **92:**37–48.
32. **Ron, Y., R. Flitman-Tene, K. Dybvig, and D. Yogev.** 2002. Identification and characterization of a site-specific tyrosine recombinase within the variable loci of Mycoplasma bovis, Mycoplasma pulmonis and Mycoplasma agalactiae. Gene **292:**205–211.
33. **Rosengarten, R., C. Citti, P. Much, J. Spergser, M. Droesse, and M. Hewicker-Trautwein.** 2001. The changing image of mycoplasmas: from innocent bystanders to emerging and reemerging pathogens in human and animal diseases. Contrib. Microbiol. **8:**166–185.
34. **Rosengarten, R., and K. S. Wise.** 1990. Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. Science **247:**315–318.
35. **Roske, K., A. Blanchard, I. Chambaud, C. Citti, J. H. Helbig, M. C. Prevost, R. Rosengarten, and E. Jacobs.** 2001. Phase variation among major surface antigens of *Mycoplasma penetrans*. Infect. Immun. **69:**7642–7651.
36. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38. **Sasaki, Y., J. Ishikawa, A. Yamashita, K. Oshima, T. Kenri, K. Furuya, C. Yoshino, A. Horino, T. Shiba, T. Sasaki, and M. Hattori.** 2002. The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans. Nucleic Acids Res. **30:**5293–5300.
39. **Shen, X., J. Gumulak, H. Yu, C. T. French, N. Zou, and K. Dybvig.** 2000. Gene rearrangements in the *vsa* locus of *Mycoplasma pulmonis*. J. Bacteriol. **182:**2900–2908.
40. **Simmons, W. L., A. M. Denison, and K. Dybvig.** 2004. Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. Infect. Immun. **72:**6846–6851.
41. **Sirand-Pugnet, P., C. Citti, A. Barre, and A. Blanchard.** 2007. Evolution of mollicutes: down a bumpy road with twists and turns. Res. Microbiol. **158:**754–766.
42. **Sirand-Pugnet, P., C. Lartigue, M. Marenda, D. Jacob, A. Barre, V. Barbe, C. Schenowitz, S. Mangenot, A. Couloux, B. Segurens, A. de Daruvar, A. Blanchard, and C. Citti.** 2007. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. PLoS Genet. **3:**e75.
43. **Sitaraman, R., A. M. Denison, and K. Dybvig.** 2002. A unique, bifunctional site-specific DNA recombinase from Mycoplasma pulmonis. Mol. Microbiol. **46:**1033–1040.
44. **Tangney, M., and G. F. Fitzgerald.** 2002. AbiA, a lactococcal abortive infection mechanism functioning in *Streptococcus thermophilus*. Appl. Environ. Microbiol. **68:**6388–6391.
45. **Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. O'Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli, and C. M. Fraser.** 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome." Proc. Natl. Acad. Sci. USA **102:**13950–13955.
46. **Tully, J. G.** 1995. Culture medium formulation for primary isolation and maintenance of mollicutes, p. 33–39. *In* J. G. Tully (ed.), Molecular and diagnostic procedures in mycoplasmology: molecular characterization. Academic Press, San Diego, CA.

Downloaded from jb.asm.org at INRA on June 16, 2009

47. **van der Woude, M. W.** 2006. Re-examining the role and random nature of phase variation. FEMS Microbiol. Lett. **254:**190–197.
48. **Weisburg, W. G., J. G. Tully, D. L. Rose, J. P. Petzel, H. Oyaizu, D. Yang, L. Mandelco, J. Sechrest, T. G. Lawrence, J. Van Etten, J. Maniloff, and C. R. Woese.** 1989. A phylogenetic analysis of the mycoplasmas: basis for their classification. J. Bacteriol. **171:**6455–6467.
49. **Westberg, J., A. Persson, A. Holmberg, A. Goesmann, J. Lundeberg, K. E. Johansson, B. Pettersson, and M. Uhlen.** 2004. The genome sequence of Mycoplasma mycoides subsp. mycoides SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). Genome Res. **14:**221–227.
50. **Wise, K. S.** 1993. Adaptive surface variation in mycoplasmas. Trends Microbiol. **1:**59–63.
51. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.

Downloaded from jb.asm.org at INRA on June 16, 2009

AUTHOR: Laurent-Xavier NOUVEL

TITLE:
Study of *Mycoplasma agalactiae* genetic diversity: Genomic plasticity, mobilome and dynamic of surface components.

SUMMARY

Mycoplasma agalactiae is responsible of contagious agalactia, a disease of small ruminants that is still difficult to control and is listed by the OIE. In order to evaluate the genetic diversity of this pathogen, 101 isolates were compared using three techniques (VNTR, RFLP, *vpma* repertoire). Results revealed a high genetic homogeneity with the PG2 type strain as representative. Some isolates however diverged such as the 5632 which was sequenced and analysed here. Whole comparative genomic and proteomic analyses of the 5632 and PG2 strains indicate that their genomic plasticity resides in important genes flux and in the presence of several mobile genetic elements (10% of the genome). These analyses also revealed that specific loci encoding repertoire of surface proteins are highly dynamic. For these minimal bacteria that lack a cell-wall, these events have most likely played a major role in their survival and adaptation to complex hosts.

KEY-WORDS
mycoplasmas, ruminants, *Mycoplasma agalactiae*, genetic diversity, molecular epidemiology, mobile genetic elements, integrative conjugative elements (ICE), insertion sequences (IS).

AUTEUR : Laurent-Xavier NOUVEL

TITRE :
Etude de la diversité génétique de *Mycoplasma agalactiae* : Plasticité des génomes, mobilome et dynamique de surface.

DIRECTEUR DE THESE : Christine CITTI

LIEU ET DATE DE SOUTENANCE :
Ecole Nationale Vétérinaire, Jeudi 26 novembre 2009

RESUME

Mycoplasma agalactiae est responsable de l'agalactie contagieuse, maladie des petits ruminants difficilement contrôlée et figurant sur la liste de l'OIE. Afin d'évaluer la diversité génétique de ce pathogène, 101 isolats ont été comparés par trois techniques (VNTR, RFLP, répertoire *vpma*). Les résultats révèlent une grande homogénéité génétique dont la souche type PG2 est représentative. Quelques isolats font exception telle la souche 5632 que nous avons séquencée et analysée ici. La comparaison des génomes et des protéomes entre 5632 et PG2 indiquent que la plasticité de ces génomes est liée à d'importants échanges d'ADN et à la présence de nombreux éléments génétiques mobiles (10% du génome). Ces analyses révèlent également une forte dynamique au sein de répertoires de gènes codant des protéines de surfaces. Pour les mycoplasmes, bactéries minimales dépourvues de paroi, ces évènements ont certainement joués un rôle dans leur survie et leur adaptation à des hôtes complexes.

MOTS-CLES
mycoplasmes, ruminants, *Mycoplasma agalactiae*, diversité génétique, marqueurs épidémiologiques, éléments génétiques mobiles, éléments intégratifs conjugatifs (ICE), séquences d'insertion (IS).

DISCIPLINE ADMINISTRATIVE Pathologie, Toxicologie, Génétique et Nutrition

---

INTITULE ET ADRESSE DU LABORATOIRE

Equipe "Mycoplasmoses des ruminants, pathologie mammaire et sécurité du lait"
Unité Mixte de Recherche 1225, "Interactions Hôtes – Agents Pathogènes"
INRA, ENVT
Ecole Nationale Vétérinaire de Toulouse
Pathologie de la Reproduction- 23 chemin des Capelles - B.P. 87614
31 076 TOULOUSE CEDEX 3