



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Pathologie, Toxicologie, Génétique et Nutrition

Présentée et soutenue par :

Mme MARION FAUCHER

le jeudi 8 novembre 2018

Titre :

Le transfert horizontal de gènes chez les mycoplasmes: de l'acquisition de l'antibiorésistance à la dynamique des génomes.

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Interactions Hôtes - Agents Pathogènes (IHAP)

Directeur(s) de Thèse :

MME CHRISTINE CITTI

MME FLORENCE TARDY

Rapporteurs :

M. BENOIT DOUBLET, INRA TOURS

Mme SERGINE EVEN, INRA RENNES

M. OLIVIER TENAILLON, INSERM PARIS

Membre(s) du jury :

Mme CÉCILE BEBEAR, UNIVERSITÉ DE BORDEAUX, Président

Mme CHRISTINE CITTI, INRA TOULOUSE, Membre

Mme FLORENCE TARDY, ANSES, Membre

M. XAVIER NOUVEL, ECOLE NATIONALE VETERINAIRE DE TOULOUSE, Membre

*"La vérité scientifique est toujours plus belle que les créations
de notre imagination et les illusions de notre ignorance."*

Claude Bernard

Résumé :

Les mycoplasmes sont des bactéries atypiques, dépourvues de paroi et souvent considérées comme des cellules minimales du fait de la taille réduite de leur génome. De nombreuses espèces sont pathogènes et ont un impact économique important dans les filières d'élevage, notamment pour les ruminants. Les mycoplasmes n'échappent pas au phénomène mondial de la résistance aux antibiotiques. Contrairement à la plupart des autres bactéries, les mycoplasmes ne contiennent pas de plasmides conjugatifs souvent incriminés dans la dissémination horizontale de gènes de résistance, la base moléculaire principalement décrite étant la mutation chromosomique des gènes cibles. De façon générale, le transfert horizontal de gènes (HGT) chez les mycoplasmes a longtemps été sous-estimé. Récemment, deux mécanismes de HGT ont été décrits chez *Mycoplasma agalactiae* : le transfert d'élément conjugatif et intégratif (ICE), et le transfert non-conventionnel de régions chromosomiques par conjugaison appelé MCT (*Mycoplasma Chromosomal Transfer*). Nos travaux se sont attachés à explorer ce dernier mécanisme et à évaluer son impact sur l'acquisition de la résistance aux antibiotiques. Une analyse de génomique comparative a été conduite à partir du séquençage de nombreux mutants spontanément résistants et de transconjugants générés par des expériences de *mating* et sélectionnés pour leur résistance. Nos résultats montrent que le MCT conduit de façon distributive au transfert simultané de nombreux fragments. En une seule étape de conjugaison impliquant deux souches, ce phénomène génère une population variée de génomes hautement mosaïques. Il accélère la dissémination de l'antibiorésistance, permettant l'acquisition de plusieurs mutations distantes associées à la résistance en un seul événement. De par les multiples possibilités de réassemblage génomique qu'il produit, le MCT pourrait avoir des conséquences importantes sur d'autres processus adaptatifs comme la virulence ou la spécificité d'hôte. Enfin, les modalités distributives et l'ampleur du MCT expliquent l'origine des transferts de gènes précédemment détectés *in silico* entre de nombreux mycoplasmes. Ce phénomène pourrait donc avoir eu des répercussions importantes sur l'évolution de ces bactéries minimales et être un facteur clé de leur persistance et virulence actuelles.

Mots clefs: *Mycoplasma agalactiae*, transfert horizontal de gènes, antibiorésistance, mutations, transfert chromosomique, adaptation.

Remerciements

Ce travail de thèse a été réalisé en cotutelle au sein de l'équipe pathogène des infections à mycoplasmes de l'UMR 1225 ENVT-INRA et de l'UMR Mycoplasmoses des ruminants de l'ANSES de Lyon.

En premier lieu, je tiens à remercier Christine Citti pour m'avoir accueillie au sein de son laboratoire, et pour ses conseils, son soutien et sa bonne humeur au cours de ces trois années. Je remercie également Florence Tardy pour ses conseils et sa disponibilité.

Je souhaite remercier Sergine Even, Olivier Tenaillon et Benoît Doublet d'avoir accepté d'examiner et de juger ce travail et Cécile Bébéar pour avoir accepté de participer au jury de soutenance de cette thèse.

Je remercie sincèrement Xavier Nouvel pour son aide constante, son enthousiasme et sa patience. J'espère ne pas t'avoir trop découragé de l'encadrement d'un doctorant, et que de nombreux autres suivront.

Je remercie Eveline, Marie et Emilie pour vos précieux conseils et votre aide quotidienne au laboratoire, mais aussi pour tous les autres sujets de discussion que nous avons pu avoir.

Je remercie Silvy, Claire, Hanna et Aurélien pour votre joie de vivre, et pour m'avoir fait découvrir le monde des chiens, des chats et des vaches.

Marie, Xavier B, Eric et Faouzi merci d'avoir partagé ces nombreux repas au RU, avec des sujets de conversation très divers mais toujours passionnants.

Un grand merci à Gabriella d'avoir partagé la première moitié de cette thèse et d'avoir partagé bien plus qu'un bureau.

Je souhaite remercier l'ensemble des personnes de l'ENVT que j'ai pu côtoyer lors de ces trois années pour leur sympathie et leur bonne humeur quotidienne. Je remercie également Agnès Tricot and Adélie Colin pour leur aide et leur accueil lors de mon séjour à l'ANSES de Lyon.

Je remercie Adeline, présente depuis si longtemps (14 ans déjà), pour sa gaieté et pour m'avoir tenue au courant du monde extérieur au laboratoire. Egalement Paoline et Fabien merci pour ces deux années de master géniales, et courage pour la toute fin de votre thèse.

Merci à tous les membres de ma famille, pour leur soutien sans faille, et pour m'avoir changé les idées les week-ends, Fred, Corinne, Marion, Marine, Aurélien, David.

Merci à Matthis de me montrer tous les jours que le chemin est plus agréable quand on est deux.

Merci à mes parents, pour m'avoir donné le goût de la science et également transmis de tant d'autres choses précieuses. Merci à mon frère, pour être arrivé à me supporter presque quotidiennement pendant ces trois années de thèse, mais surtout pour être toujours présent.

TABLE DES MATIÈRES

INTRODUCTION	13
1. Les mycoplasmes : des bactéries atypiques	15
1.1. Caractéristiques générales (classe des <i>Mollicutes</i>)	15
1.2. Les Mycoplasmes (famille des <i>Mycoplasmataceae</i>)	19
1.2.1. Caractéristiques des mycoplasmes	19
1.2.2. Le point sur les génomes des mycoplasmes séquencés	20
1.2.3. Cellules minimales et biologie synthétique.....	21
1.3. Les mycoplasmes de ruminants et l'agalactie contagieuse	22
1.3.1. Impacts des mycoplasmoses chez les ruminants	22
1.3.2. L'agalactie contagieuse chez les petits ruminants	23
1.3.3. <i>Mycoplasma agalactiae</i> : trois souches séquencées	24
2. La résistance aux antibiotiques chez les mycoplasmes	26
2.1. Les antibiotiques et le phénomène de résistance	26
2.1.1. Les différentes classes d'antibiotiques.....	26
2.1.2. La résistance, un problème de santé publique	27
2.1.3. Les mécanismes moléculaires de la résistance	28
2.1.4. Détermination du niveau de résistance	30
2.2. Mécanismes d'antibiorésistance chez les mycoplasmes	31
2.2.1. Mécanismes de résistance intrinsèque chez les mycoplasmes.....	31
2.2.2. Principal mécanisme de résistance par mutation des gènes codant les cibles	32
2.2.3. Autres exemples de mécanismes de résistance chez les mycoplasmes	36
3. Le Transfert de Gène Horizontal chez les mycoplasmes	38
3.1. Mécanisme et impact du HGT chez les bactéries	38
3.1.1. Définition du HGT.....	38
3.1.2. Les mécanismes de HGT.....	39
3.1.3. Les éléments génétiques mobiles (EGM)	43
3.1.4. Traits adaptatifs ou phénotypes transférés par les EGM.....	43
3.2. Evidence <i>in silico</i> de transferts chromosomiques chez les mycoplasmes	44
3.2.1. Entre pathogènes de ruminants.....	45
3.2.2. Entre pathogènes humains	46
3.2.3. Entre pathogènes de volailles	46
3.2.4. Entre pathogènes de plantes	47
3.3. Les EGM présent chez les mycoplasmes : rôle majeur des ICE	48
3.3.1. Les éléments transposables (transposon et <i>IS</i>).....	48
3.3.2. Les prophages	50
3.3.3. Les plasmides	52
3.3.4. Les Elément Intégratif et Conjugatif (ICE)	54
3.4. Démonstration <i>in vitro</i> de HGT par conjugaison chez <i>M. agalactiae</i>	60
3.4.1. Transfert des ICE de mycoplasmes.....	60
3.4.2. Transfert chromosomique.....	63
4. Objectifs et contexte de l'étude.....	65
5. Références.....	68

CHAPITRE 1 : LES MYCOPLASMES SOUS SÉLECTION ANTIMICROBIENNE EXPÉRIMENTALE: CONTRIBUTION INÉDITE DU TRANSFERT CHROMOSOMIQUE HORIZONTAL DANS L'ACCÉLÉRATION DE L'ADAPTATION À COURT TERME	83
Résumé.....	86
Abstract	87
1. Introduction	89
2. Results.....	91
3. Discussion.....	102
4. Materials and Methods	106
5. References.....	111
6. Supporting information	116
CHAPITRE 2 : LE TRANSFERT CHROMOSOMIQUE HORIZONTAL CONTRIBUE À L'EXTRAORDINAIRE PLASTICITÉ DES GÉNOMES DE MYCOPLASMES	135
Résumé.....	138
Abstract	139
1. Introduction	140
2. Results.....	143
3. Discussion.....	154
4. Materials and Methods	160
5. References.....	163
6. Supplementary material.....	167
CHAPITRE 3 : APPORT DU SÉQUENÇAGE NGS DANS L'ANALYSE DE GÉNOMES DE MYCOPLASMES	211
Résumé.....	214
1. Introduction	215
2. Choix des méthodes de séquençage et d'analyse	216
2.1. Le séquençage NGS	216
2.2. Analyse des données de séquençage	220
3. Résultats et discussion	226
3.1. Détection de variants	226
3.2. Reconstruction des génomes mosaïques des transconjugants	230
3.3. Analyse des variations de couverture.....	233
4. Conclusions	242
5. Matériels et méthodes	244
6. Références.....	246
DISCUSSION ET PERSPECTIVES.....	251
ANNEXES.....	263

TABLE DES ILLUSTRATIONS : FIGURES

Introduction

Figure 1 : Place des <i>Mollicutes</i> dans l'arbre phylogénétique de la vie.	17
Figure 2 : Arbre phylogénétique des <i>Mollicutes</i>	18
Figure 3 : Photo de cellules et de colonies de <i>Mycoplasma agalactiae</i>	20
Figure 4 : Les différentes familles d'antibiotiques et leurs cibles cellulaires.	27
Figure 5 : Les principaux mécanismes de résistance aux antibiotiques (ATB).	29
Figure 6 : Méthodes de mesure de la CMI.	31
Figure 7 : La cible des tétracyclines : l'ARNr 16S.	33
Figure 8 : La cible des macrolides : l'ARNr 23S.	34
Figure 9 : Les cibles des fluoroquinolones : les topoisomères de type II.	36
Figure 10 : Principaux mécanismes de HGT : transformation, transduction, conjugaison.	40
Figure 11 : Architecture d'un SSTIV et voie de translocation de l'ADN.	42
Figure 12 : Régions échangées entre <i>M. agalactiae</i> et le cluster Mycoides.	45
Figure 13 : HGT et évolution du clade Mirum des Spiroplasma.	47
Figure 14 : Organisation des transposons et IS.	49
Figure 15 : Organisation génomique et comparaison du prophage de <i>M. agalactiae</i> 14628.	51
Figure 16 : Comparaison des plasmides de mycoplasmes de ruminants.	53
Figure 17 : Organisation et mode de dissémination des ICE.	56
Figure 18 : Organisation génomique et comparaison des ICE chez les mycoplasmes.	58
Figure 19 : Transposon conjugatif comportant <i>tetM</i> chez <i>M. hominis</i> et <i>S. agalactiae</i>	59
Figure 20: Transfert de l'ICE chez <i>M. agalactiae</i>	61
Figure 21 : Organisation et fonctions des CDS des ICE de la souche 5632.	62
Figure 22 : Tout le génome de PG2 peut être transféré chez 5632.	63

Chapitre 1

Fig 1. <i>In vitro</i> selection and genome analysis of spontaneous <i>M. agalactiae</i> Enro ^R mutants.	95
Fig 2. Mycoplasma transconjugant genomes and relative fitness.	97
Fig 3. Features of genes targeted by the enrofloxacin in Enro ^R mutants and transconjugants..	99
Fig 4. Fitness evolution of mutants and transconjugants over passages.....	102

Chapitre 2

Figure 1. Mating experiments create in a single step a diversity of transconjugants with mosaic genomes	145
Figure 2. Number and size of PG2-donor fragments acquired by MCT in mating progeny.....	150
Figure 3. Artificial PEG fusion of mycoplasma cells also results in mosaic genomes.....	152
Figure 4 : Analyses of homologous region (HR) flanking the acquired DNA in transconjugants	153

Chapitre 3

Figure 1 : Profondeur et étendue de la couverture d'un alignement.....	218
Figure 2 : Etapes d'un <i>run</i> de séquençage par la technique Illumina.	219
Figure 3 : Visualisation d'un alignement de <i>reads</i> sur une référence avec IGV ou Artemis.	224
Figure 4 : Etapes réalisées lors d'une analyse de variants vis-à-vis d'une référence.	227
Figure 5 : Visualisation IGV d'alignements des <i>reads</i> sur le génome des 2 souches parentales.	230
Figure 6 : Définition des fragments des transconjugants.	231
Figure 7 : Méthode de tri des <i>reads</i> spécifique de PG2 pour reconstruire les génomes des transconjugants.....	233
Figure 8 : Alignement et analyse de variant chez 5632 ^{G3} , PG2 55-5 et un mutant résistant....	234
Figure 9 : Variation du niveau d'expression des ICE des transconjugants.....	236
Figure 10 : Délétion des ICE due à la recombinaison d'un fragment PG2.	237
Figure 11 : Variation de couverture de l'opéron ribosomique et d'un locus <i>vpma</i>	238
Figure 12 : Détection de l'insertion d'un marqueur de sélection chez 5632 ^{G3}	239
Figure 13 : Duplication du locus <i>ps3</i> chez PG2 55-5.	241

TABLE DES ILLUSTRATIONS : TABLEAUX

Introduction

Tableau 1 : Taxonomie de la classe des <i>Mollicutes</i>	16
Tableau 2 : Caractéristiques des génomes séquencés de la famille des <i>Mycoplasmataceae</i>	21
Tableau 3 : Comparaison des génomes des 3 souches séquencées de <i>M. agalactiae</i>	25
Tableau 4 : Description des prophages présents chez les <i>Mycoplasmes</i>	51
Tableau 5 : Description des plasmides présents chez les <i>Mycoplasmes</i>	54
Tableau 6 : Description des ICE présents chez les <i>Mycoplasmes</i>	60

Chapitre 1

Table 1: Increase in the level of resistance to enrofloxacin in <i>M. agalactiae</i> mutants correlate with the accumulation of amino acid substitutions in QRDR of <i>parE</i> , <i>parC</i> , <i>gyrA</i> and <i>gyrB</i>	93
---	----

Chapitre 2

Table 1 : Genetic modifications induced MCT in 5632 recipient cells	144
---	-----

Chapitre 3

Tableau 1 : Comparaison des technologies de NGS.....	217
Tableau 2 : Logiciels utilisés pour les analyses de séquençages NGS.	220
Tableau 3 : Variations entre PG2 et 5632 détectées à partir de 3 logiciels	226
Tableau 4 : Analyse de variant chez différents clones.	229
Tableau 5: Liste des oligonucléotides	245

TABLE DES ILLUSTRATIONS : SUPPLEMENTS

Chapitre 1

Figure S1: Loci targeted by specific-PG2 or -5632 PCR assays.	116
Table S1 : Primers used for Sanger sequencing and PCR screening.....	117
Table S2 : Position, nature and impact of mutations in whole genome of Enro ^R mutants.....	119
Table S3 : List of acquired and lost genes in transconjugants.	123
Table S4 : Changes in mutation pattern of Enro ^R mutants and transconjugants during multiplication without selection pressure.	127
Table S5 : Variations in enrofloxacin target enzyme between PG2 and 5632 strains.	130
Table S6 : European Nucleotide Archive (EMBL-EBI) accession numbers.	131

Chapitre 2

Figure S1 : Mosaic genome of the M171.4 transconjugant coming from PG2 vs 4867 mating.	167
Table S1 : <i>M. agalactiae</i> strain features, mating efficiencies and identification of resulting transconjugants.....	168
Table S2 : List of mating partners and chromosomal position of their respective selective markers	169
Table S3 : Oligonucleotides used in this study	170
Table S4 : Extent and size of PG2-donor chromosomal DNA fragments acquired during mating and PEG-cellular fusion	171
Table S5 : PG2-CDSs acquired by transconjugants and PEG-hybrids.	172
Table S6 : Features of micro-complex regions detected <i>M. agalactiae</i> transjugants and hybrids	195
Table S7 : List of SNPs detected in micro-complex region \leq 5nt and their impact in amino acid sequences.....	196
Table S8 : Sizes of the Homologous Region (HR) surrounding the PG2-inherited DNA fragments	200
Table S9 : Comparative genome analysis of <i>M. agalactiae</i> 5632 and 4867 strains.....	207

INTRODUCTION

INTRODUCTION

1. Les mycoplasmes : des bactéries atypiques

1.1. Caractéristiques générales (classe des *Mollicutes*)

Le premier mycoplasme identifié en 1898 est l'agent étiologique de la péripneumonie contagieuse bovine (PPCB), *Mycoplasma mycoides* subsp. *mycoides* (Nocard et Roux 1898). A la suite de la découverte de bactéries similaires (*M. agalactiae*, *M. bovigenitalium*, *M. hominis*, *M. pulmonis* ...), Edward et Freundt proposent de les regrouper au sein de la classe des *Mollicutes* (Edward et Freundt 1956; Edward *et al.* 1967), terme provenant du latin *mollis* (mou) et *cutis* (la peau). Cette classe appartient au phylum des Tenericutes, elle est divisée en 5 ordres (Euzéby 1997; Brown 2011) : *Mycoplasmatales*, *Haloplasmatales*, *Entomoplasmatales*, *Acholeplasmatales* et *Anaeroplasmatales* (Tableau 1). Elle regroupe plus de 200 espèces, avec une grande variété d'hôtes appartenant à tous les domaines du vivant : Homme, animaux, végétaux (Razin et Herrmann 2002). La plupart se regroupent au sein du genre *Mycoplasma*, le terme mycoplasme est ainsi souvent employé pour désigner les *Mollicutes* au sens large.

Initialement considérés comme des virus, les *Mollicutes* sont des bactéries de petite taille (0,15 à 0,45 µm de diamètre), de ce fait, associé à leur pléiomorphisme morphologique, la mesure de la DO (densité optique) n'est pas applicable. Leurs génomes sont de taille réduite, aux alentours de 1 Mb, un des plus petits étant celui de *M. genitalium* avec seulement 580 Kb (Fraser *et al.* 1995). Ces organismes sont dépourvus de paroi cellulaire, ce qui leur confère une certaine plasticité morphologique et une résistance intrinsèque aux antibiotiques agissant sur la paroi.

Tableau 1 : Taxonomie de la classe des *Mollicutes*.

Ordre	Famille	Genre	Nombre d'espèce	Hotes
<i>Mycoplasmatales</i>	<i>Mycoplasmataceae</i> ^a	<i>Mycoplasma</i>	127	Animaux, Homme
		<i>Ureaplasma</i>	7	Animaux, Homme
		<i>Eperythrozoon</i> ^b	2	Animaux
		" <i>Candidatus Mycoplasma</i> "	17	Animaux
<i>Entomoplasmatales</i>	<i>Entomoplasmataceae</i>	<i>Entomoplasma</i>	6	Plantes, Insectes
		<i>Mesoplasma</i>	12	Plantes, Insectes
	<i>Spiroplasmataceae</i>	<i>Spiroplasma</i>	38	Plantes, Insectes
<i>Haloplasmatales</i> ^c	<i>Haloplasmataceae</i>	<i>Haloplasma</i>	1	Eau
<i>Acholeplasmatales</i>	<i>Acholeplasmataceae</i>	<i>Acholeplasma</i>	18	Animaux, Plantes
		" <i>Candidatus Phytoplasma</i> " ^d	38	Plantes, Insectes
<i>Anaeroplasmatales</i>	<i>Anaeroplasmataceae</i>	<i>Anaeroplasma</i>	4	Animaux
		<i>Asteroleplasma</i>	1	Animaux

D'après LPSN (www.bacterio.net)(Euzéby 1997; Johansson et Pettersson 2002).

^(a) Dans cette famille, les 3 espèces qui composaient le genre *Haemobartonella* ont été reclassées parmi les *Mycoplasma* (Neimark *et al.* 2002).

^(b) A l'exception d'*Eperythrozoon coccoides* et *E. parvum*, les espèces du genre *Eperythrozoon* ont été reclassées au sein du genre *Mycoplasma* (Neimark *et al.* 2001).

^(c) Nouvel ordre validé en 2008, composé à ce jour d'une seule espèce: *H. contractile*, provenant des eaux de la Mer Rouge (Antunes *et al.* 2008).

^(d) Genre composé d'espèces phylogénétiquement proche des *Acholeplasmes*, non cultivables *in vitro* et provisoirement classées dans la famille des *Acholeplasmataceae*.

Les *Mollicutes* dérivent d'un ancêtre commun aux *Clostridia* et aux *Bacilli* (Firmicutes) (Figure 1), bactéries à Gram positif et à faible teneur en G+C. Leur évolution a longtemps été considérée comme uniquement régressive (Woese *et al.* 1980; Weisburg *et al.* 1989). Leurs génomes ont une faible teneur en G+C, en moyenne de 27 % (Sirand-Pugnet *et al.* 2007a) et le codon UGA, codon stop dans le code génétique universel, correspond au tryptophane chez les *Mollicutes*, excepté chez les *Acholeplasma* et *Phytoplasma* (Yamao *et al.* 1985). Au cours de leur évolution, la réduction de taille des génomes de *Mollicutes* a entraîné la perte de certaines voies de synthèses, comme celle du cholestérol (excepté pour les genres *Mesoplasma*, *Acholeplasma* et *Asteroplasma*), celles de certains acides aminés ou du peptidoglycane, composant essentiel de la paroi (Razin *et al.* 1998). Dépendant de leur hôte pour les nutriments, ces organismes ont adopté un mode de vie parasitaire ou commensal. Leur culture *in vitro* en milieu acellulaire, quand elle est possible, nécessite des milieux riches et complexes.

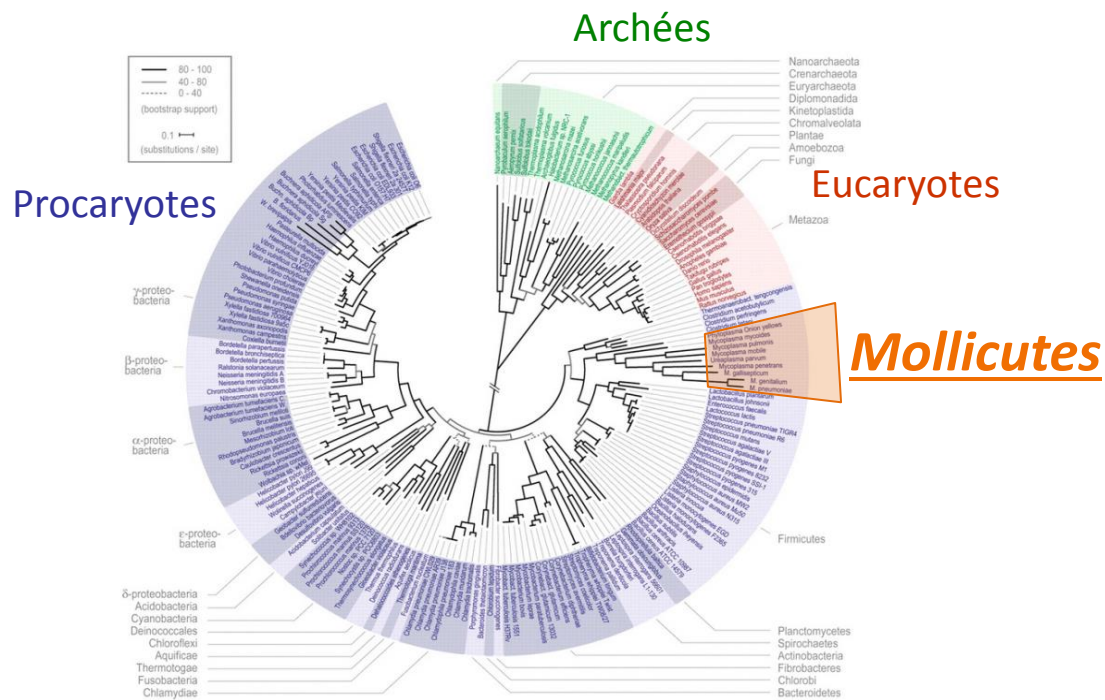


Figure 1 : Place des *Mollicutes* dans l'arbre phylogénétique de la vie.

Arbre phylogénétique présentant les 3 grands domaines du vivant, en vert les arches, en rouge les eucaryotes et en bleu les procaryotes. La classe des *Mollicutes* est indiquée en orange (adaptée de Ciccarelli *et al.* 2006).

Sur la base de leur séquence d'ARNr 16S, les mycoplasmes peuvent être classés en 5 groupes phylogénétiques (Hominis, Pneumoniae, Phytoplasma, Acholeplasma et Spiroplasma) (Figure 2) (Weisburg *et al.* 1989). Les *Mollicutes* des groupes Hominis et Pneumoniae infectent l'Homme et une grande diversité d'animaux. Ceux appartenant aux groupes Phytoplasma et Acholeplasma sont retrouvés chez les végétaux et leurs insectes vecteurs. Ceux issus du groupe Spiroplasma infectent les végétaux et les animaux. Parmi ces derniers, se trouvent des mycoplasmes de ruminants regroupés au sein du cluster nommé Mycoides (Manso-Silvan *et al.* 2007) comme *M. mycoides* subsp. *mycoides*.

Les avancées dans les techniques de sequenage ont permis l'acces aux sequences de nombreux genomes de *Mollicutes*. L'analyse et la comparaison de ces genomes ont revele la presence de repertoires de lipoproteines etendus, d'elements genetiques mobiles, de rearrangements chromosomiques et de regions mosaiques... soit une plus grande dynamique et plasticite des genomes qu'initialement attendue (Sirand-Pugnet *et al.* 2007a). Ces decouvertes ont renverse le paradigme d'une evolution uniquement regressive. Aujourd'hui, les *Mollicutes* sont ainsi consideres comme des bacteries optimisees specialistes de leur hote mais toujours douees de capacites evolutives ou adaptatives (Citti et Blanchard 2013; Citti *et al.* 2018).

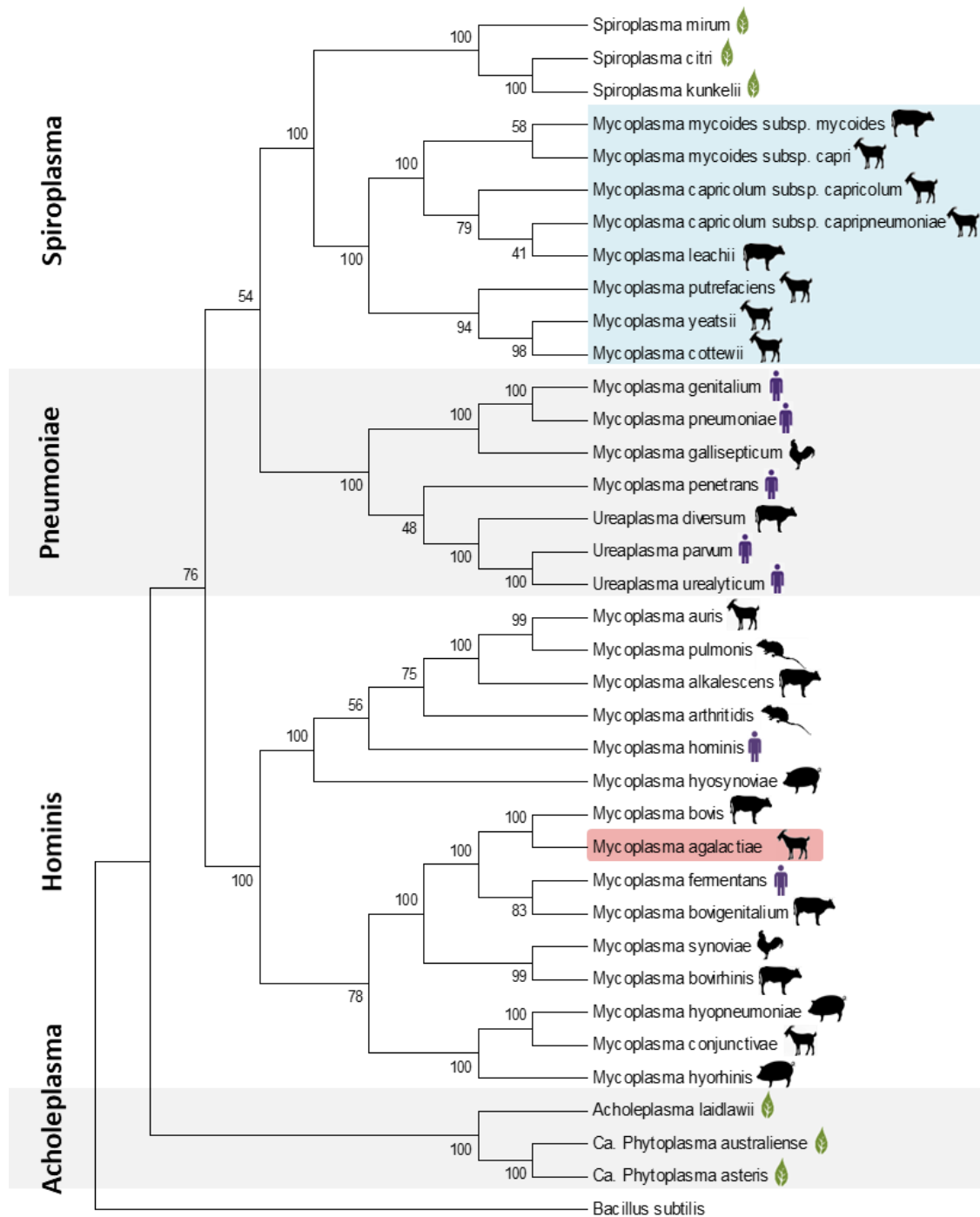


Figure 2 : Arbre phylogénétique des Mollicutes.

Arbre phylogénétique basé sur 37 séquences d'ADN ribosomique 16S alignées avec clustalW (MEGA7), 1340 positions nucléotidiques, méthode Neighbor-Joining, bootstrap de 500 répliquats (pourcentage indiqué à côté de chaque branche). *Bacillus subtilis* utilisé comme extra-groupe. Groupes phylogénétiques indiqués à gauche. Notre modèle d'étude *Mycoplasma agalactiae* surligné en rouge, et le cluster Mycoides au sens large en bleu. Hôte de chaque mollicute symbolisé à côté du nom (végétaux et insectes en vert, animaux en noir, Homme en violet).

1.2. Les Mycoplasmes (famille des *Mycoplasmataceae*)

1.2.1. Caractéristiques des mycoplasmes

Parmi les *Mollicutes*, l'ordre des *Mycoplasmatales* est constitué d'une seule famille, les *Mycoplasmataceae* (Tableau 1). Au sein de cette famille, le genre *Ureaplasma* regroupe des bactéries capables d'hydrolyser l'urée, alors que celles du genre *Mycoplasma* utilisent le glucose, des acides organiques, des alcools ou l'arginine comme source d'énergie. Le genre *Eperythrozoon* se compose d'espèces isolées à partir de prélèvement sanguins qui ne sont pas cultivables (Brown 2011).

Les bactéries appartenant à la famille des *Mycoplasmataceae* sont capables d'infecter une large gamme d'hôtes, comme l'Homme, de nombreux mammifères (ruminant, chat, rongeur...) des insectes, des oiseaux, des poissons et des reptiles. Leur mode de vie est le plus souvent parasitaire ou commensal. Ces mycoplasmes sont associés aux muqueuses respiratoires ou génitales et peuvent être responsables de maladies chroniques à faible mortalité et forte morbidité. Certains mycoplasmes sont pathogènes pour l'Homme (Blanchard et Bebear 2002), comme *M. pneumoniae* responsable d'infections pulmonaires ou *M. hominis*, *M. genitalium* et *U. urealyticum* retrouvés au niveau de l'appareil uro-génital. Des infections à mycoplasmes sont présentes au sein des filières de production animale et sont à l'origine de pertes économiques importantes (Frey 2002). Par exemple, *M. mycoides* subsp. *mycoides* est responsable de la péripneumonie contagieuse bovine (PPCB), *M. gallisepticum* et *M. hyopneumoniae* causent des infections respiratoires dans les filières avicoles et porcines.

Au laboratoire, les mycoplasmes ont une croissance lente en milieu liquide ou gélosé, de quelques jours à quelques semaines. Sur milieu gélosé, les colonies de mycoplasmes ont généralement une morphologie caractéristique dite en "œuf sur le plat" visibles à la loupe binoculaire (Figure 3). Les difficultés de culture de ces bactéries et le peu d'outils génétiques disponibles constituent un frein à la compréhension de leur métabolisme et de leur pouvoir pathogène.

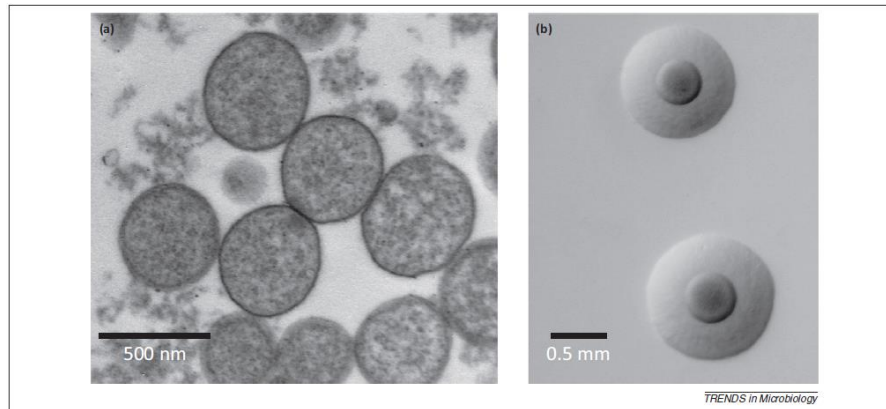


Figure 3 : Photo de cellules et de colonies de *Mycoplasma agalactiae*.

(a) Photo de cellule de *M. agalactiae* par microscope électronique à transmission.

(b) Photo de colonies ayant la forme typique d'un œuf sur le plat. (Citti et Blanchard 2013).

En l'absence de paroi, la membrane plasmique des mycoplasmes est en contact direct avec leur environnement. Elle joue un rôle clé dans les interactions avec le milieu environnant et l'hôte, notamment dans l'échappement au système immunitaire (Rosengarten et Wise 1990). Un grand nombre de lipoprotéines sont soumises à des variations de phase (on/off) et des variations antigéniques (modification de la taille ou de la structure) qui sont gouvernées par des mécanismes génétiques complexes (Citti *et al.* 2010). Ces protéines associées à la membrane sont parfois regroupées au sein de loci appelés *vpma* chez *M. agalactiae* (Glew *et al.* 2000); *vsp* chez *M. bovis* (Lysnyansky *et al.* 1999); *vsa* chez *M. pulmonis* (Shen *et al.* 2000); *vlp* chez *M. hyorhinis* (Citti *et al.* 2000)... La plasticité de la surface des mycoplasmes provient d'une hyper-variabilité de l'expression de ces protéines (Citti *et al.* 2010).

1.2.2. Le point sur les génomes des mycoplasmes séquencés

A sein du genre *Mycoplasma*, le nombre d'espèces dont au moins une souche est séquencée a considérablement augmenté passant de seulement 12 en 2009 à 80 aujourd'hui (Tableau 2). Au sein d'une même espèce, les séquences de plusieurs souches sont parfois disponibles, aboutissant à un total de 286 génomes. Certains organismes sont surreprésentés, ce sont les mycoplasmes ayant un intérêt en médecine humaine (*M. pneumoniae* 48 souches ; *M. hominis* 14 souches) ou vétérinaire (*M. bovis* 31 souches ; *M. gallisepticum* 21 souches). Sur les 286 génomes de *Mycoplasma*, seulement la moitié sont entièrement assemblés et circularisés, les autres sont sous la forme de *contigs* ou de *scaffolds*. Les génomes du genre *Mycoplasma* ont

un pourcentage en G+C de 27,7 % et une taille moyenne de 886 Kb (Tableau 2). Le plus grand génome est celui de *M. moatsii* (souche ATCC 27625, 1481 Kb) isolé chez le singe et le plus petit est celui de *M. genitalium* (souche M6282, 580 Kb) retrouvé chez l'Homme. Deux autres espèces ont un génome de taille encore plus réduite, *Candidatus M. haemominutum* (514 Kb) et *Eperythrozoon parvum* (564 Kb), hémoparasites respectivement isolés chez le chat et le cochon.

Tableau 2 : Caractéristiques des génomes séquencés de la famille des *Mycoplasmataceae*.

Genres	<i>Mycoplasma</i>	<i>Ureaplasma</i>	<i>Eperythrozoon</i>	<i>Candidatus Mycoplasma</i>	Total*
Nombre d'espèces	127	7	2	17	153
Nombre d'espèces séquencées	80	4	1	4	89
Nombre de souches séquencées (génomes circularisés)	286 (145)	28 (10)	1 (1)	5 (3)	320 (159)
Taille moyenne des génomes (Kb)	886	858	564	708	873
Génome plus grand (Kb)	1481	992	564	936	1481
	<i>M. moatsii</i>	<i>U. diversum</i>	<i>E. parvum</i>	<i>Candidatus M. haemobos</i>	
Génome plus petit (Kb)	580	741	564	514	514
	<i>M. genitalium</i>	<i>U. parvum</i>	<i>E. parvum</i>	<i>Candidatus M. haemominutum</i>	
Pourcentage en G+C moyen (%)	27,7	27,9	27,0	33,5	27,9

D'après NCBI, en Juin 2018. (www.ncbi.nlm.nih.gov/Taxonomy/Browser).

* le total correspond à la somme ou à la moyenne obtenue chez la famille des *Mycoplasmataceae*.

1.2.3. Cellules minimales et biologie synthétique

Les génomes de *M. genitalium* (1995) et *M. pneumoniae* (1996) ont été parmi les premiers génomes bactériens entièrement séquencés. *M. genitalium* est un des plus petits organismes capables de réplication autonome. Son génome est composé de seulement 580 Kb et contient 482 gènes (Fraser *et al.* 1995), dont 382 caractérisés comme essentiels lors d'expériences de mutagenèse par transposition (Glass *et al.* 2006). Du fait de leur génome réduit, les mycoplasmes constituent un modèle pour appréhender le concept de cellule minimale (Peterson et Fraser 2001).

Ils constituent ainsi un modèle pour des études de biologie de synthèse (ou biologie synthétique) qui a pour but de synthétiser à façon de nouveaux organismes avec comme objectifs l'amélioration de notre compréhension du vivant et la production d'organismes utiles à diverses applications (vaccination, chimie verte...). Le génome de *M. mycoides* subsp. *capri* (1079 Kb) a été synthétisé artificiellement, assemblé chez la levure par recombinaison puis transplanté dans une bactérie receveuse *M. capricolum* subsp. *capricolum*. Ceci a abouti au premier organisme "synthétique" *M. mycoides* JCVI-syn1.0 (Lartigue *et al.* 2007; Gibson *et al.* 2010). Afin de s'approcher au plus près des composants essentiels à la vie, à partir du génome de JCV-syn1.0, un nouveau clone synthétique a été construit, par délétion de gènes non essentiels, il est nommé *M. mycoides* JCV-syn3.0 et comporte 531 Kb et 473 gènes (Hutchison *et al.* 2016). Les développeurs de cette technologie cherchent maintenant à créer des mycoplasmes plateformes qui produiraient des molécules d'intérêt comme des antigènes et constitueraient des souches vaccinales.

1.3. Les mycoplasmes de ruminants et l'agalactie contagieuse

1.3.1. Impacts des mycoplasmoses chez les ruminants

Chez les ruminants (bovins, caprins, ovins), plusieurs espèces de mycoplasmes sont des pathogènes majeurs et engendrent des pertes économiques importantes en production laitière comme bouchère. Ces mycoplasmes sont à l'origine de divers signes cliniques comme des mammites, des arthrites, des pneumonies, des kérato-conjonctivites et parfois des anémies (Citti 2006). Trois mycoplasmoses des ruminants sont des maladies à déclaration obligatoire auprès de l'OIE (Office Internationale des Epizooties, organisation mondiale de la santé animale)

- La péripneumonie contagieuse bovine (PPCB) causée par *M. mycoides* subsp. *mycoides*
- La pleuropneumonie contagieuse caprine (PPCC) causée par *M. capricolum* subsp. *capripneumoniae*
- L'agalactie contagieuse (AC) des petits ruminants causée par plusieurs mycoplasmes (*M. capricolum* subsp. *capricolum*, *M. mycoides* subsp. *capri*, *M. agalactiae*, *M. putrefaciens*)

La PPCB et la PPCC sont présentes en Afrique ou en Asie. Dans les pays industrialisés, l'AC listée à l'OIE et la bronchopneumonie infectieuse enzootique (BPIE) à *M. bovis* constituent les mycoplasmoses au plus fort impact économique.

En France, un réseau d'épidémiologie-surveillance des mycoplasmoses de ruminants, VIGIMYC a été mis en place en 2003 (Chazel *et al.* 2010). Il joue un rôle de surveillance pour l'éventuelle introduction de la PPCB ou de la PPCC sur le territoire. Plus largement, il constitue un observatoire de la circulation de l'AC, de la BPIE et de l'ensemble des mycoplasmoses de ruminants en France.

1.3.2. L'agalactie contagieuse chez les petits ruminants

En 1923 J. Bridé et A. Donatien isolent pour la première fois *M. agalactiae* (Bridé et Donatien 1925), l'agent responsable de l'agalactie contagieuse chez les chèvres et les moutons. L'agalactie contagieuse peut également être causée par *M. capricolum* subsp. *capricolum*, *M. mycoides* subsp. *capri* et *M. putrefaciens* mais *M. agalactiae* reste l'agent majoritairement isolé chez les ovins. Ces bactéries sont responsables d'atteintes mammaire, articulaire, oculaire et parfois pulmonaire chez les jeunes animaux (Bergonier et Berthelot 2008).

L'agalactie contagieuse sévit principalement dans le pourtour méditerranéen et en Asie occidentale (Bergonier et Poumarat 1996). En France, l'agalactie contagieuse à *M. agalactiae* est aujourd'hui classée comme danger sanitaire de catégorie 2 à déclaration obligatoire (Arrêté ministériel n°117 du 7 mai 2017). Le traitement de l'agalactie contagieuse peut faire appel à des traitements antibiotiques (tétracyclines, macrolides et fluoroquinolones). Cependant, l'efficacité de ces traitements est illusoire à l'échelle d'un troupeau, elle permet tout au mieux une atténuation des symptômes mais aucunement un arrêt de la circulation de l'agent. Aucun vaccin ne dispose d'une AMM (Autorisation de mise sur le marché) en France mais des vaccins sont commercialisés à l'étranger. Ces vaccins ne permettent généralement qu'une amélioration des signes cliniques, avec une immunité partielle et de courte durée (Agnone *et al.* 2013).

1.3.3. *Mycoplasma agalactiae* : trois souches séquencées

Actuellement, les génomes de trois souches de *M. agalactiae* ont été séquencées (Tableau 3), et leurs analyses ont permis des avancées majeures pour la compréhension de la phylogénie et de la dynamique des génomes de mycoplasmes. La séquence complète et circularisée de la souche de référence PG2 55-5 a été obtenue en 2007 (Sirand-Pugnet *et al.* 2007b) et celle de la souche 5632 C1 en 2010 (Nouvel *et al.* 2010), deux souches isolées de chèvres en Espagne. Le génome d'une souche plus atypique, 14628, isolée chez un bouquetin en France a été séquencé en 2012 (Tardy *et al.* 2012), la séquence n'a pas été circularisée mais la quasi-totalité (99,3 %) du génome est assemblée en un seul *scaffold*.

Ces 3 génomes présentent une faible teneur en G+C (29,7 %) et une forte densité de gènes (88 %), correspondant à des caractéristiques typiques des mycoplasmes. Aux alentours de 750 séquences codantes (CDS, *Coding DNA Sequence*) sont répertoriées, dont la moitié est associée à une fonction. Un nombre non négligeable de CDS ont été annotées protéines hypothétiques (HP), certaines ont par la suite été détectées par protéomique mais beaucoup ne présentent pas d'homologie avec des gènes de fonctions connues. Les protéines présentes au niveau de la membrane plasmique des mycoplasmes jouent un rôle crucial du fait de l'absence de paroi cellulaire, avec notamment la lipoprotéine P40 responsable de l'attachement aux cellules de l'hôte (Fleury *et al.* 2002). Les lipoprotéines des loci *vpma* (*variable protein membrane og M. agalactiae*) permettent une hyper-variabilité de la surface cellulaire, grâce à des réarrangements génétiques entre les gènes *vpma* (Nouvel *et al.* 2009), une variabilité est également décrite pour les polysaccharides de surface (Gaurivaud *et al.* 2016). La souche 5632 est capable d'une remarquable variabilité au niveau de ces antigènes de surface car elle possède 23 gènes *vpma*, répartis-en 2 loci contre seulement 6 gènes chez la souche PG2. De nombreux éléments génétiques mobiles (EGM) sont présents chez cette souche, avec 3 copies d'ICE (*Integratif and Conjugatif Element*) de 27 Kb chacun (ICE₅₆₃₂-I, -II, -III) et 15 IS (*Insertion Sequence*). Chez la souche 14628, isolée d'un bouquetin, un prophage de 34 Kb est détecté. La présence de ces différents EGM fait de *M. agalactiae* un modèle d'étude des transferts de gènes horizontaux (HGT) chez les mycoplasmes.

Tableau 3 : Comparaison des génomes des 3 souches séquencées de *M. agalactiae*.

	PG2	5632	14628
Organisme hôte	Caprin	Caprin	Bouquetin
Origine géographique	Espagne	Espagne	France
Date d'isolement	1952	Avant 1991	2006
Taille du génome (nt)	877438	1006702	919593*
Pourcentage de G+C (%)	29,70	29,62	29,87
Densité de gène (%)	88,5	88,7	86,9
Nombre de CDS	713	815	719
Nombre de CDS annotées	399	518	443
Nombre de (C)HP	314	297	287
Nombre d'ICE	(1 vestige)	3 (+2 vestiges)	1 (+1 vestige)
Opéron ribosomique	2	2	2
Nombre ARN transfert	34	34	34
Nombre d'IS	1 (+2 pseudo)	15 (+3 pseudo)	0 (+ 2 pseudo)
Nombre de gène vpma	6 (1 locus)	23 (2 loci)	nd
Prophage	0	0	1
Numéro d'accession	CU179680.1	FP671138.1	AJPR00000000.1
Référence	(Sirand-Pugnet <i>et al.</i> 2007b)	(Nouvel <i>et al.</i> 2010)	(Tardy <i>et al.</i> 2012)

D'après Nouvel *et al.* 2010 ; Tardy *et al.* 2012a. * Génome non circularisé.

(C)HP : protéines hypothétiques (conservées) ; ICE : élément intégratif et conjugatif ; IS : séquence d'insertion ; nd : non déterminé.

2. La résistance aux antibiotiques chez les mycoplasmes

2.1. Les antibiotiques et le phénomène de résistance

2.1.1. Les différentes classes d'antibiotiques

Les antibiotiques (ATB) sont des substances naturelles ou synthétiques qui agissent spécifiquement et à faible dose afin d'induire la lyse (effet bactéricide) ou l'arrêt de la croissance (effet bactériostatique) des bactéries. Leur utilisation a permis une avancée majeure en médecine humaine et vétérinaire pour le traitement de maladies infectieuses, permettant une augmentation notable de l'espérance de vie. Il existe une grande variété d'antibiotiques, la plupart sont des dérivés de molécules naturellement produites par les microorganismes. Les ATB se fixent au niveau de leur cible et perturbent les processus cellulaires, ils peuvent être classés suivant leur mode d'action (Lewis 2013; Wilson 2013) (Figure 4) :

- ATB inhibant la synthèse de la paroi bactérienne en empêchant la formation du peptidoglycane, un composant essentiel de la paroi bactérienne (β -lactamines, glycopeptides).

- ATB désorganisant la membrane plasmique par une augmentant de la perméabilité membranaire (polymyxines).

- ATB inhibant la synthèse protéique en se fixant sur le ribosome (tétracyclines, aminoglycosides, macrolides, phénicolés).

- ATB désorganisant les acides nucléiques en se fixant aux topoisomérase de type II (fluoroquinolones), à l'ARN polymérase (rifamycines), ou à des enzymes impliquées dans la synthèse des bases purines (sulfamides et triméthoprimés).

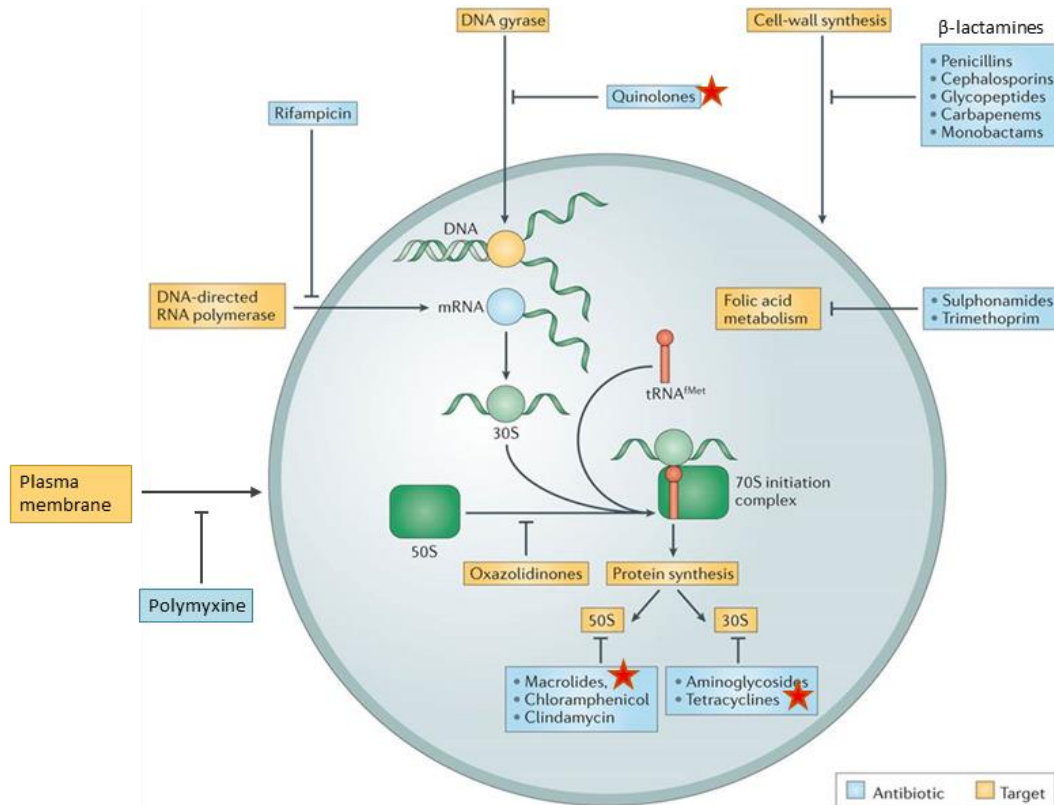


Figure 4 : Les différentes familles d'antibiotiques et leurs cibles cellulaires.

Les antibiotiques sont soulignés en bleu et leurs cibles respectives sont en jaune. Les étoiles rouges indiquent les antibiotiques utilisés pour traiter les mycoplasmes (adaptée de Lewis 2013).

2.1.2. La résistance, un problème de santé publique

Les bactéries ont développé des mécanismes de résistance leur permettant de contrer l'action des ATB. L'augmentation de l'antibiorésistance (nombre d'espèces bactériennes et de familles d'antibiotiques concernées) est un problème majeur de santé publique mondiale. Ces résistances limitent l'arsenal thérapeutique disponible et diminuent l'efficacité des traitements antibiotiques (Holmes *et al.* 2016). Dans une moindre mesure, cela pose aussi un problème économique avec des traitements efficaces qui sont plus coûteux (Levy et Marshall 2004). Cette augmentation de la résistance est due à une sur-utilisation et une mauvaise utilisation des ATB, en médecine humaine et vétérinaire combinés au pouvoir d'adaptation très efficace des bactéries.

L'Organisation Mondiale de la Santé (OMS) a mis en place un plan d'action pour lutter contre l'augmentation de la résistance aux ATB à l'échelle du globe, avec notamment une classification des ATB suivant leur degré d'importance

(<http://www.who.int/foodsafety/publications/antimicrobials-fifth/en/>). Les ATB hautement prioritaires depuis 2017 sont les céphalosporines, les glycopeptides, les macrolides et kétolides, les polymyxines et les fluoroquinolones. En médecine vétérinaire les céphalosporines et les fluoroquinolones sont considéré comme des ATB critiques depuis 2016 (Décret n°2016-317 du 16 mars 2016). L'OMS a également publié une liste de 12 bactéries pathogènes pour l'Homme pour lesquelles il est urgent de trouver de nouveaux traitements. *Acinetobacter baumannii*, *Pseudomonas aeruginosa* et les *Enterobacteriaceae* sont classés de haute priorité (<http://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>).

2.1.3. Les mécanismes moléculaires de la résistance

La résistance aux ATB est intrinsèque (naturelle) ou acquise. La résistance intrinsèque est présente chez tous les membres d'une espèce bactérienne. En revanche, la résistance acquise apparaît lorsque des souches d'une espèce bactérienne naturellement sensible à un ATB deviennent résistantes.

Les trois principaux mécanismes de résistance acquise sont : (1) l'altération de la cible de l'ATB, (2) l'altération de l'ATB lui-même ou (3) une diminution de la concentration intracellulaire d'ATB (Chopra et Roberts 2001; Alekshun et Levy 2007; Blair *et al.* 2014) (Figure 5).

(1) L'altération de la cible de l'ATB peut être due à des mutations dans le gène correspondant à la protéine ou à l'ARN cible. Elle peut aussi provenir de l'ajout de groupements chimiques, comme des méthylation pour les macrolides (Leclercq et Courvalin 1991). Une protéine bactérienne peut aussi se fixer sur la cible de l'ATB pour la protéger, c'est le cas de la protéine Tet(M) qui se fixe au ribosome pour empêcher la fixation des tétracyclines (Connell *et al.* 2003). Ces modifications entraînent un changement de structure 3D de la cible qui empêchent l'action des ATB.

(2) L'ATB peut être directement affecté par des enzymes le dégradant, comme les β -lactamases qui hydrolysent les β -lactamines. D'autres enzymes peuvent ajouter des groupements chimiques à l'ATB, comme des acétyltransférases et des phosphotransférases pour les aminoglycosides (Ramirez et Tolmasky 2010).

(3) La concentration intracellulaire d'un ATB est un paramètre important pour son activité. Cette concentration est modulée par un import diminué (défaut de pénétration) ou un export augmenté (défaut d'accumulation) (Poole 2005). Un défaut de pénétration est dû à des

mutations au niveau de canaux membranaires, appelés porines (ex : OmpF, OmpC). Lors d'un défaut d'accumulation, des pompes transmembranaires plus ou moins spécifiques détoxifient la bactérie, grâce à un efflux actif. Ces pompes sont réparties en 5 familles : les MFS (*Major Facilitator Superfamily*), les SMR (*Small Multidrug Resistance*), les MATE (*Multidrug And Toxic compound Extrusion*), les ABC transporteur (*ATP Binding Cassette*) et les RND (*Resistance Nodulation Division*). Ces pompes utilisent diverses sources d'énergie, des ions H⁺ (MFS, SMR, RND) Na⁺ (MATE) ou de l'ATP (ABC).

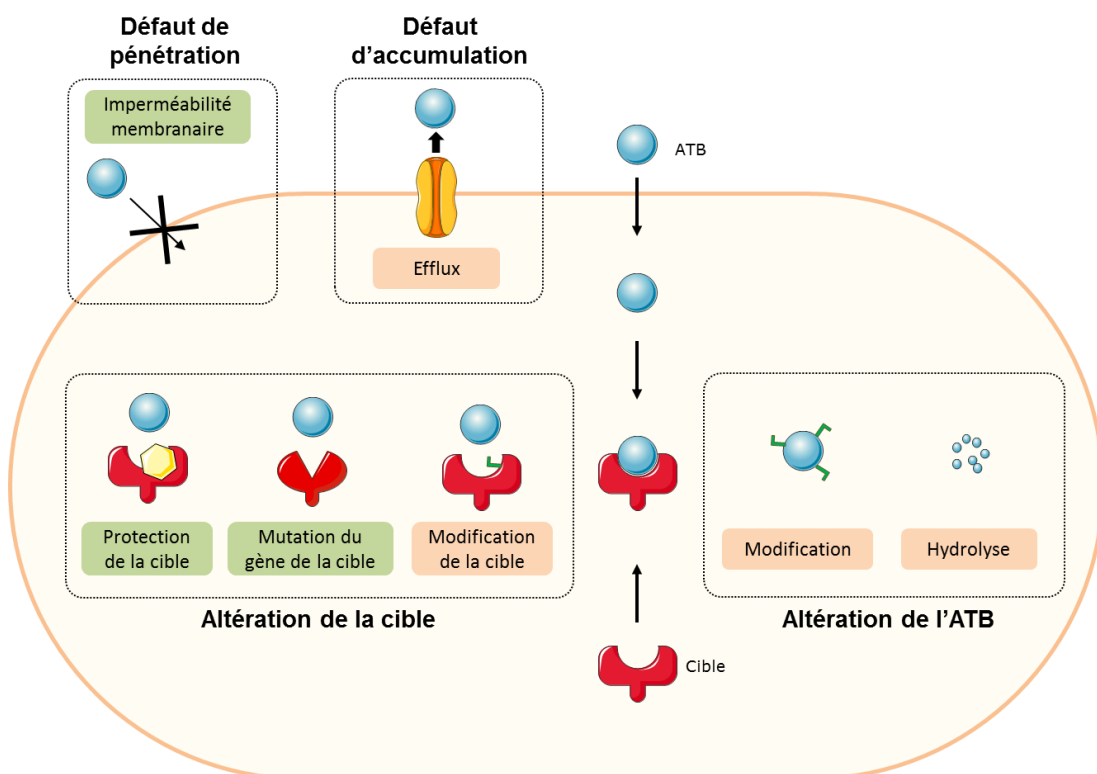


Figure 5 : Les principaux mécanismes de résistance aux antibiotiques (ATB).

Les mécanismes décrits chez les mycoplasmes sont indiqués en vert.

La majorité des phénomènes de résistance sont liés à l'acquisition par les bactéries d'EGM portant le déterminant de la résistance. Ces transferts génétiques horizontaux (HGT) peuvent avoir lieu entre bactéries d'une même espèce ou entre espèces différentes. Les bactéries présentes dans le sol et dans le microbiote intestinal constituent un réservoir de résistance (Broaders *et al.* 2013; Nesme et Simonet 2015). Les autres cas de résistance proviennent de mutations chromosomiques. Ces mutations se produisent de manière ponctuelle et les

modifications qu'elles entraînent sont stables et héréditaires. Elles ont un impact direct ou indirect sur la structure de protéines ou d'ARN cibles de l'ATB, ou sur les réseaux de régulation de gènes clés avec, par exemple, la surexpression de pompes à efflux (Blair *et al.* 2014).

2.1.4. Détermination du niveau de résistance

Le niveau de résistance d'une espèce bactérienne pour un ATB donné s'évalue par la mesure de la CMI (Concentration Minimale Inhibitrice). La CMI correspond à la plus faible concentration d'ATB inhibant toute croissance visible *in vitro*. Plusieurs méthodes permettent d'évaluer la CMI : la dilution d'ATB en milieu liquide ou solide, la diffusion à l'aide de disques ou de bandelettes (Etest) imprégnés d'ATB (Figure 6). Les méthodes de référence sont celles basées sur la dilution, elles nécessitent un inoculum calibré car sa concentration influence directement la mesure de la CMI. La densité optique (DO) ne pouvant être utilisée pour dénombrer les mycoplasmes, les cultures doivent être titrées au préalable en milieux liquide ou gélosé. Du fait de leur croissance lente, ces méthodes sont donc longues à mettre en place chez les mycoplasmes.

Les valeurs de CMI sont utilisées pour la surveillance de l'évolution des niveaux de résistantes. Il est nécessaire de standardiser les méthodes de mesure pour avoir une fiabilité inter- et intra-laboratoires (milieu de culture, temps d'incubation, densité de l'inoculum, souches contrôles...) (CLSI 2013). Les valeurs seuils de CMI sont basées sur des valeurs critiques déterminées à partir d'éléments cliniques, elles permettent de définir des catégories sensibles, intermédiaires et résistantes en contexte clinique. Ces valeurs seuils sont dépendantes de la souche bactérienne et de l'ATB utilisé. Des organismes de référence comme le CLSI (*Clinical & Laboratory Standards Institute*) ou l'EUCAST (*European Committee on Antimicrobial Susceptibility Testing*) déterminent ces valeurs. Des références sont disponibles pour les mycoplasmes humains (*M. pneumoniae*, *M. hominis*, *U. urealyticum*) (Waites *et al.* 2012; CLSI 2011) ainsi que des test commerciaux avec des dilutions en milieux liquide. Pour les mycoplasmes d'intérêt vétérinaire, il n'existe pas de référence officielle, les recommandations proposée par Hannan en 2000 sont très largement utilisées (Hannan 2000).

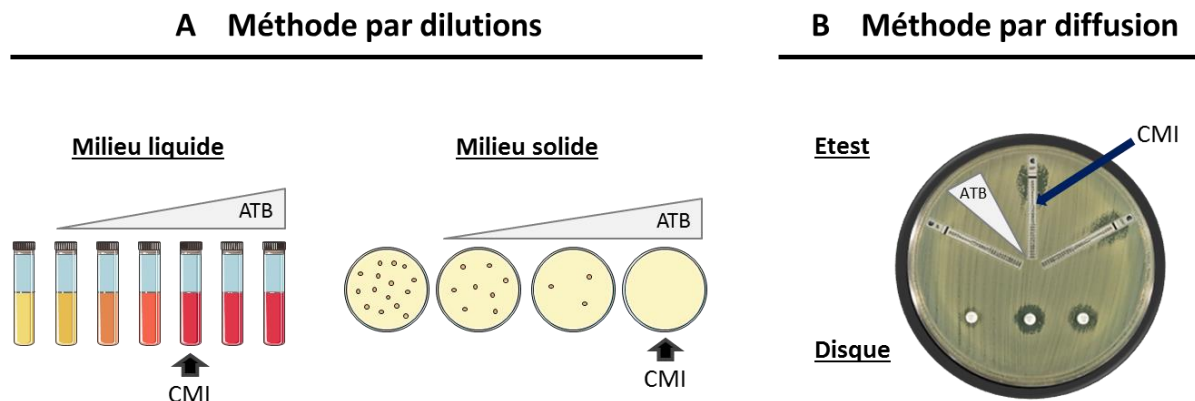


Figure 6 : Méthodes de mesure de la CMI.

(A) Méthodes par dilution avec des concentrations croissantes d'ATB qui sont ajoutés au milieu liquide ou solide. La CMI se lit à la première concentration d'ATB qui empêche la croissance, soit par l'indicateur coloré du milieu qui ne vire pas, soit par l'absence de colonies sur gélose. (B) Méthodes par diffusion, qui nécessitent l'ensemencement d'un tapis bactérien. Des bandelettes (Etest) ou des disques imprégnés d'ATB sont déposés sur la boîte. Les Etest contiennent des concentrations croissantes d'ATB, on peut donc lire directement la CMI. Avec les disques, la CMI est dépendante du diamètre d'inhibition (photo de la boîte de Pétri adaptée de Pottumarthy *et al.* 2003).

2.2. Mécanismes d'antibiorésistance chez les mycoplasmes

Les *Mollicutes* ont la particularité d'être intrinsèquement résistants à de nombreux ATB (Waites *et al.* 2014; Gautier-Bouchardon 2018). Les ATB disponibles pour le traitement des mycoplasmoses sont les tétracyclines, les MLSK (Macrolides, Lincosamides, Streptogramine, Kétolides), les aminoglycosides, les phénicolés et les fluoroquinolones (FQ). Comme chez toutes les bactéries, une augmentation de la résistance est observée chez les mycoplasmes. Ce constat est réalisé chez des pathogènes humains et animaux, avec les macrolides chez *M. pneumoniae* et chez *M. gallisepticum* (Pereyre *et al.* 2016; Gerchman *et al.* 2011), la plupart des ATB chez *M. hyopneumoniae* et *M. bovis* (Thongkamkoon *et al.* 2013; Gautier-Bouchardon *et al.* 2014; Heuvelink *et al.* 2016). Dans le cas de *M. agalactiae*, la résistance est modérée (Poumarat *et al.* 2016). Les niveaux de CMI sont en très légère augmentation pour tous les ATB chez les isolats de mouton, et uniquement pour les macrolides chez les isolats de chèvres.

2.2.1. Mécanismes de résistance intrinsèque chez les mycoplasmes

L'absence de paroi rend les mycoplasmes naturellement résistants aux ATB qui ciblent la paroi bactérienne comme les β -lactamines ou les glycopeptides. Les polymyxines et les sulfonamides/triméthoprimés sont également inefficaces dû à l'absence de leurs cibles, les lipopolysaccharides et la voie de biosynthèse de l'acide folique (Waites *et al.* 2014; Gautier-

Bouchardon 2018). De plus, les mycoplasmes sont intrinsèquement résistants aux rifampicines qui ciblent l'ARN polymérase. Ceci a été démontré initialement chez *Spiroplasma citri* comme étant dû à la position N526 de RpoB, la sous-unité β de l'ARN polymérase, qui empêche la fixation de l'ATB (Gaurivaud *et al.* 1996; Gautier-Bouchardon 2018).

2.2.2. Principal mécanisme de résistance par mutation des gènes codant les cibles

Le mécanisme de résistance majoritairement identifié chez les mycoplasmes est une altération de la cible de l'ATB par mutation des gènes correspondants, pour les tétracyclines, les macrolides et les FQ. Deux autres mécanismes ont été observés, une protection de la cible pour les tétracyclines et un défaut d'accumulation pour les FQ (Figure 5). Le niveau de résistance associé à chaque mutation de la cible est variable suivant la molécule d'ATB utilisée, l'espèce et la souche de mycoplasme. La position des mutations est systématiquement indiquée en suivant la numérotation d'*Escherichia coli*. Toutes les mutations décrites dans la littérature sont regroupées en Annexes avec leurs positions dans les gènes cibles pour les tétracyclines (Tableau A1), les macrolides (Tableau A2) et les FQ (Tableau A3).

Les tétracyclines se fixent à la sous-unité 30S du ribosome, ce qui empêche la fixation des ARNr, présents en un ou deux opérons chez les mycoplasmes. Plus précisément, les tétracyclines se fixent au niveau du site tet-1 (Figure 7), boucle formée par l'hélice 34 (nt 1055-1056 et 1196-1200) et l'hélice 31 (nt 964-967) de l'ARNr 16S (Gerrits *et al.* 2002). Chez les mycoplasmes résistants des mutations de l'ARNr 16S sont localisées au niveau de ce site tet-1. Les nucléotides 965 et 967 sont le plus souvent mutés chez *M. hominis*, *M. pneumoniae* et *M. bovis* (Dégrange *et al.* 2008; Amram *et al.* 2015).

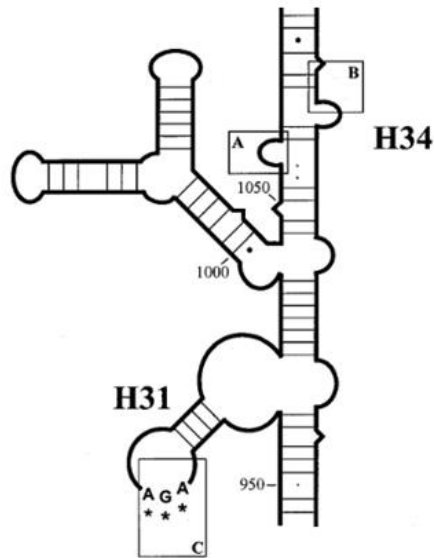


Figure 7 : La cible des tétracyclines : l'ARNr 16S

Structure secondaire d'une région de l'ARNr 16S (Gerrits *et al.* 2002). Le site de fixation des tétracyclines (tet-1) est formé par les nucléotides 1055-1056 et 1196-1200 de l'hélice 34 (H34) et par les nucléotides 964-967 de l'hélice 31 (H31).

Les macrolides et apparentés (MLSK) se fixent à la sous-unité 50S, cela induit une dissociation prématurée entre les ARNt et le ribosome, et aboutit à des polypeptides tronqués. Les macrolides se fixent à l'ARNr 23S, au niveau de la boucle centrale du domaine V et de l'épingle 35 du domaine II (Figure 8A). Ils interagissent aussi indirectement avec les protéines ribosomiques L4 et L22 situées le long du canal peptidique (Schlünzen *et al.* 2001). Chez les mycoplasmes résistants des mutations de l'ARNr 23S sont décrites, surtout retrouvées dans la boucle centrale du domaine V, aux positions 2058, 2059, 2062 et 2611. Des mutations sont également présentes sur l'épingle du domaine II de l'ARNr 23S. Les protéines ribosomiques L4 et L22 comportent des SNP ou des indel dans certains cas, mais ces mutations n'entraînent un effet que si des mutations du 23S sont déjà présentes (Khalil *et al.* 2017). Chez *M. hominis* une résistance intrinsèque à l'érythromycine est observée (Pereyre *et al.* 2002; Waites *et al.* 2014). Elle est due à 2 substitutions de l'ARNr 23S : A2057G et C2610U, qui influencent la structure de la boucle centrale du domaine V. La mutation A2057G est aussi observée chez *M. pulmonis*, *M. fermentans* et *M. hyopneumoniae* (Furneri *et al.* 2000; Pereyre *et al.* 2002). En comparant la séquence de l'ARNr 23S des mycoplasmes disponibles dans la base Molligen (Barré *et al.* 2004), à quelques exceptions près, tous les mycoplasmes des groupes phylogénétiques Spiroplasma et Pneumoniae ont un G2057 alors que tous ceux du groupe Hominis ont un A2057 et sont quant à eux certainement résistants (Figure 8B).

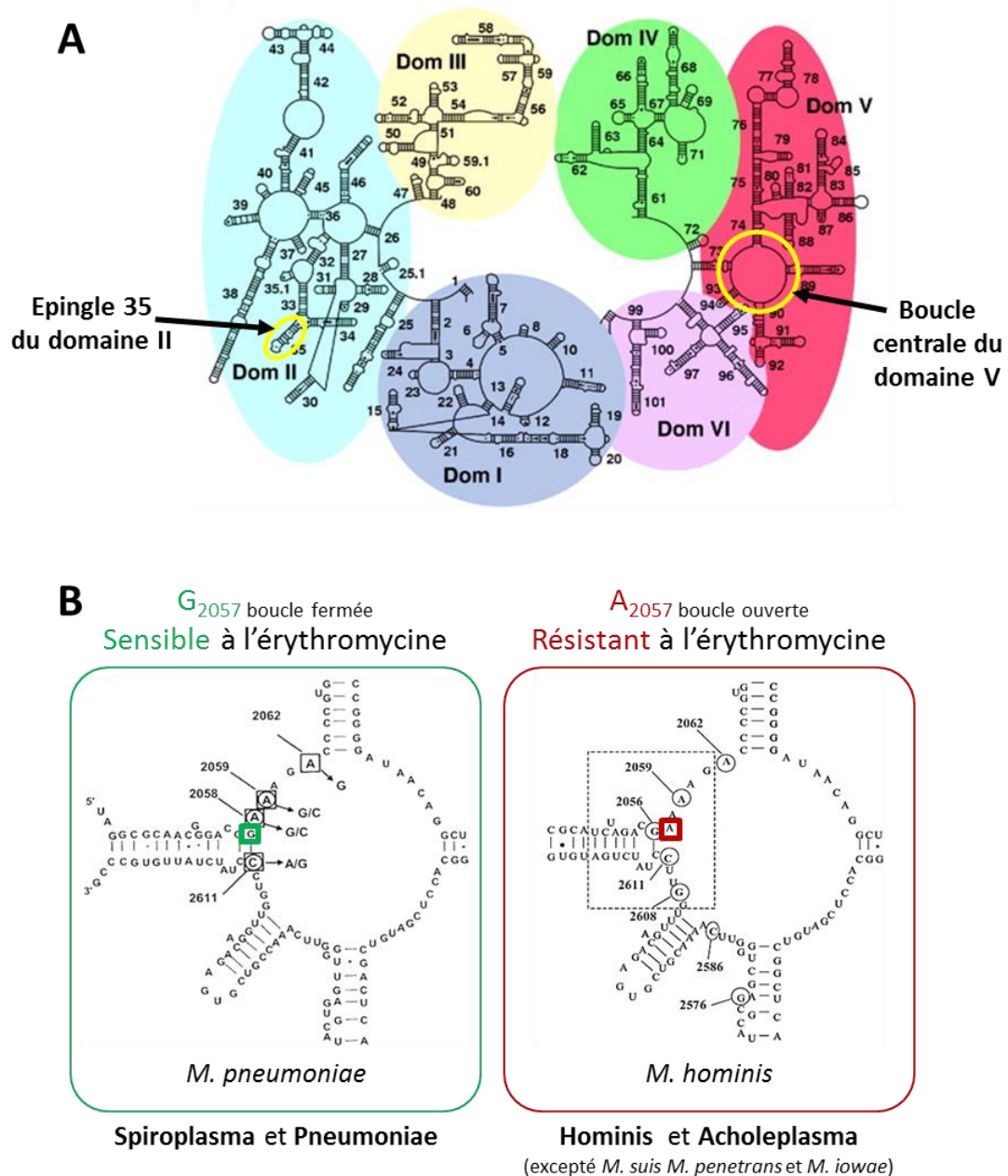


Figure 8 : La cible des macrolides : l'ARNr 23S

(A) Structure secondaire de l'ARNr 23S, avec les régions cibles des macrolides entouré en jaune, correspondant à l'épingle 35 du domaine II et la boucle centrale du domaine V, (adaptée de Ban *et al.* 2000). (B) Structure secondaire du domaine V chez les mycoplasmes sensible (en vert) ou résistant (en rouge) à l'érythromycine qui présentent un G ou un A en position 2057. La structure de *M. pneumoniae* est adaptée de Waites *et al.* 2014 et celle de *M. hominis* de Pereyre *et al.* 2006. Les séquences des ARNr 23S sont analysées avec MEGA7.

Les fluoroquinolones (FQ) inhibent l'activité des topoisomérases de type II (Redgrave *et al.* 2014; Correia *et al.* 2017). Ces enzymes sont impliquées dans la régulation du surenroulement de l'ADN, un processus essentiel pour la réplication et la transcription cellulaire. Les topoisomérases de type II sont des hétérotétramères, la DNA gyrase se compose de deux sous-unités GyrA et 2 GyrB et la topoisomérase IV de 2 ParC et 2 ParE (Figure 9A). La principale activité de la DNA gyrase est d'introduire du surenroulement négatif dans la molécule d'ADN et celle de la topoisomérase IV est de séparer les chromosomes homologues lors de la division cellulaire (procédé de décaténation). Les FQ se fixent au niveau des complexes de clivage ADN/enzyme ce qui empêche la religation de l'ADN, il y a donc formation d'un complexe ADN/enzyme/FQ et libération de fragments d'ADN double brin. En plus de l'inhibition enzymatique, les fragments d'ADN double brin entraînent une réponse SOS (Qin *et al.* 2015). Au niveau des mycoplasmes résistants, des mutations dans les gènes codant les topoisomérases de type II sont présentes. Elles sont surtout identifiées au niveau de *hotspot* dans les QRDR pour "*Quinolone Resistance-Determining Regions*" composés des acides aminés 67-106 pour GyrA, 64-103 pour ParC (Yoshida *et al.* 1990, 1991) (Figure 9B). D'autres mutations plus atypiques ont lieu dans les QRDR de GyrB (420 et 441) ou de ParE (426 et 447). Chez les mycoplasmes résistants, ces *hotspots* sont les positions GyrA 83, ParC 80 et 84. Les positions 84 et 87 de GyrA et 78, 79, 91 de ParC sont aussi souvent identifiées. Il a été montré qu'une accumulation de plusieurs mutations permet d'atteindre des niveaux élevés de résistance pour *M. bovis* ou *M. hominis* (Bébéar *et al.* 1998; Khalil *et al.* 2017).

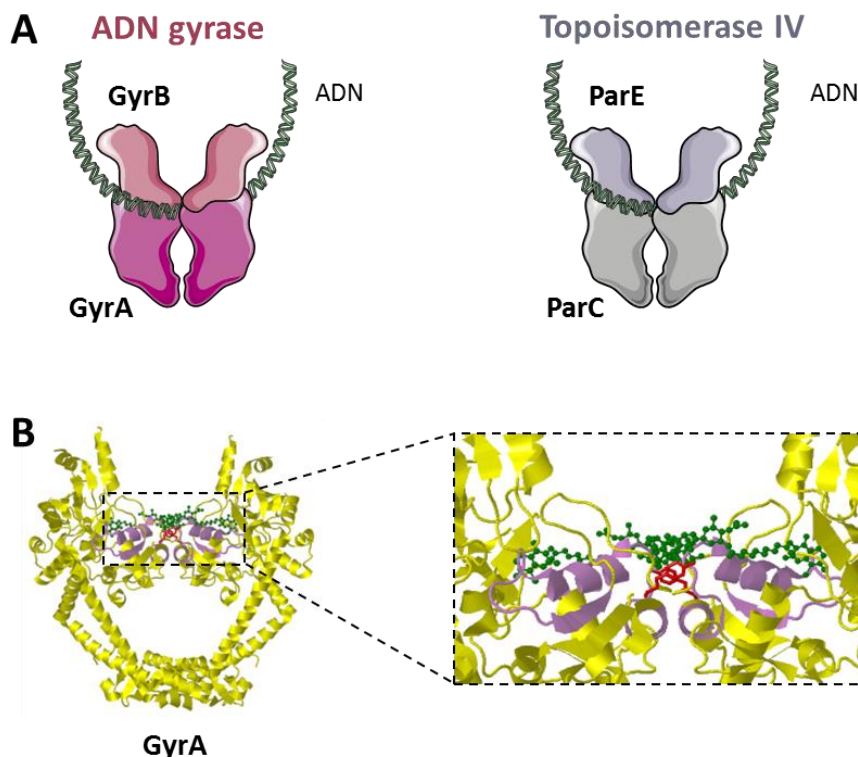


Figure 9 : Les cibles des fluoroquinolones : les topoisomérases de type II

(A) Représentation schématiques de l'ADN gyrase (en rose) composée des sous unités GyrA et GyrB et de la topoisomérase IV (en gris) composée des sous unités ParC et ParE. (B) Structure de la sous unité GyrA d'*E. coli* (PDB 4CKL), en jaune, complexée avec un antibiotique la simocyclinone D8, en vert. Le QRDR est indiqué en violet et le site actif de l'enzyme (Tyr₁₂₂) en rouge.

2.2.3. Autres exemples de mécanismes de résistance chez les mycoplasmes

Défaut d'accumulation de l'ATB par efflux :

Des pompes membranaires permettent d'exporter les molécules d'ATB par efflux actif. Chez les bactéries, des mutations au niveau de ces pompes n'entraînent en général qu'une légère augmentation de la résistance, favorisant ainsi l'apparition de résistance de plus haut niveau (Webber et Piddock 2003). Pour déceler l'implication de pompes dans le phénomène de résistance, l'assimilation d'ATB est mesurée en présence d'inhibiteurs (réserpine, CCCP, orthovanadate) qui ciblent différents types de pompes. Seulement deux exemples d'implication de système d'efflux dans la résistance aux FQ sont disponibles chez les mycoplasmes et correspondent à des pompes de type ABC transporteur. Ces défauts d'accumulation ont été uniquement mis en évidence *in vitro*, et leur implication *in vivo* reste à démontrer. Chez *M. hominis* en présence d'inhibiteur, l'assimilation de ciprofloxacine augmente, reflétant une action des pompes pour exporter l'ATB (Raheison *et al.* 2002). Il a été

montré que ceci est dû à la surexpression de deux pompes à efflux MD1 et MD2 (Raheison *et al.* 2005). Chez *M. mycoides* subsp. *capri* des isolats et des clones sélectionnés *in vitro* présentent une CMI à la ciprofloxacine et la norfloxacine plus faible en présence d'inhibiteur de pompe (Antunes *et al.* 2015).

Seul cas connu de résistance liée à un EGM le Tet(M) :

Un seul exemple de protection de la cible existe chez les mycoplasmes. Il s'agit de la protéine Tet(M) qui se fixe sur le ribosome et empêche la fixation des tétracyclines, sans affecter l'activité de traduction (Dégrange *et al.* 2008). Le gène *tet(M)* a été identifié chez des isolats de *M. hominis* et *U. urealyticum* (Roberts *et al.* 1985; Roberts et Kenny 1986), il est porté par un transposon *Tn916* provenant de *Streptococcus* (Burdett 1980).

Ce cas d'EGM impliqué dans résistance aux ATB est le seul démontré pour les mycoplasmes. Chez les mycoplasmes la résistance aux ATB est ainsi principalement associé à des mutations chromosomiques, à l'inverse de ce qui est habituellement observé chez les autres bactéries. Le grand nombre de données de séquençage acquises aujourd'hui identifie de plus en plus d'EGM chez les mycoplasmes, leur rôle dans la diffusion de la résistance aux ATB, ou dans d'autres processus adaptatifs reste à étudier.

3. Le Transfert de Gène Horizontal chez les mycoplasmes

3.1. Mécanisme et impact du HGT chez les bactéries

3.1.1. Définition du HGT

Le HGT est un processus permettant à une cellule (receveuse) d'acquérir du matériel génétique provenant d'une autre cellule (donneuse), sans en être le descendant direct. Il s'oppose au transfert vertical de gènes, faisant référence à l'hérédité. Des transferts horizontaux de gènes ont été identifiés au sein des 3 grands domaines du vivant, les procaryotes, les archées et les eucaryotes (Luo et Wasserfallen 2001; Andersson 2005; Thomas et Nielsen 2005). Des échanges ont même été identifiés entre ces différents domaines du vivant (Nelson *et al.* 1999; Rumpho *et al.* 2008).

Chez les procaryotes, la multiplication cellulaire se fait uniquement suivant un mode de division clonale. Le HGT permet un brassage génétique, reproduisant les effets d'une reproduction sexuée, avec des transferts de portions plus ou moins grandes du patrimoine génétique. L'acquisition de nouveaux gènes par la cellule receveuse peut lui permettre d'acquérir de nouveaux traits phénotypiques comme la résistance à un ATB, l'acquisition de facteurs de pathogénicité, de nouvelles propriétés métaboliques, ou une nouvelle spécificité d'hôtes (Ochman *et al.* 2000; Frost *et al.* 2005). Le HGT peut également avoir des conséquences négatives pour les génomes, avec l'altération ou la perte de certains gènes, mais seules les populations viables seront capables de se multiplier.

Les HGT ont été initialement découverts par les expériences de Griffith en 1928 (Griffith 1928), avec des injections à des souris d'un mélange de pneumocoques de la souche R, non virulente, et de la souche S, virulente mais tuée. Ces expériences démontrent un transfert de la virulence entre deux bactéries, appelé transformation. Ce n'est qu'en 1944 qu'Avery montre que ce facteur transformant est constitué d'ADN (Avery *et al.* 1944). L'impact des HGT a été mesuré en 1959 avec la découverte au Japon de souches d'entérobactéries multi-résistantes aux antibiotiques. L'équipe d'Ochiai a montré que des plasmides étaient à l'origine de cette dissémination de la résistance entre des *Shigella* spp. et des *Escherichia* spp. (Ochiai *et al.* 1959), et ils ont pu reproduire ces transferts *in vitro*. Actuellement, le HGT est considéré comme un des facteurs principaux de l'augmentation de la résistance aux ATB dans le monde (von Wintersdorff *et al.* 2016). Un vaste nombre de gènes de résistance sont portés par des plasmides, ils permettent de cibler différentes familles d'ATB comme les fluoroquinolones ou les β -lactamines (Robicsek *et al.* 2006; Brolund et Sandegren 2016).

Les HGT ont joué un rôle majeur dans l'évolution des espèces. Depuis les avancées technologiques permettant le séquençage de génomes à grande échelle, la communauté scientifique a pris conscience des nombreux échanges réalisés entre les espèces et de l'impact sur leur évolution et leur adaptation (Koonin 2009). Des événements de HGT peuvent être détectés *in silico*. Par exemple, le contenu en G+C d'une région transférée peut se trouver différent de celui du reste du génome, ce biais permet ainsi de les identifier. Des HGT peuvent aussi être détectés par la recherche du plus proche homologue. Lorsque le meilleur score est obtenu chez un organisme éloigné phylogénétiquement, cela suggère un événement de HGT plus ou moins récent (Zhaxybayeva 2009).

3.1.2. Les mécanismes de HGT

Il existe trois mécanismes principaux de HGT : la transformation, la transduction et la conjugaison (Figure 10). Ils permettent un transfert de matériel génétique, appelé EGM (Élément Génétique Mobile), d'une bactérie donneuse vers une bactérie receveuse. Une fois acquis par la bactérie receveuse, l'ADN est soit intégré au génome afin de se maintenir dans la cellule, soit circularisé pour se répliquer de manière autonome.

Transformation : Lors de la transformation bactérienne les bactéries dites compétentes acquièrent des fragments d'ADN nu extracellulaire (Chen et Dubnau 2004). Cet ADN étant non protégé, il peut rapidement être dégradé par des DNases présentes dans le milieu extracellulaire. Certaines bactéries sont naturellement compétentes. Cet état peut être permanent ou dépendant de l'environnement ou du cycle cellulaire. L'état de compétence d'une bactérie peut être induit *in vitro* par des traitements chimiques qui vont perméabiliser les membranes.

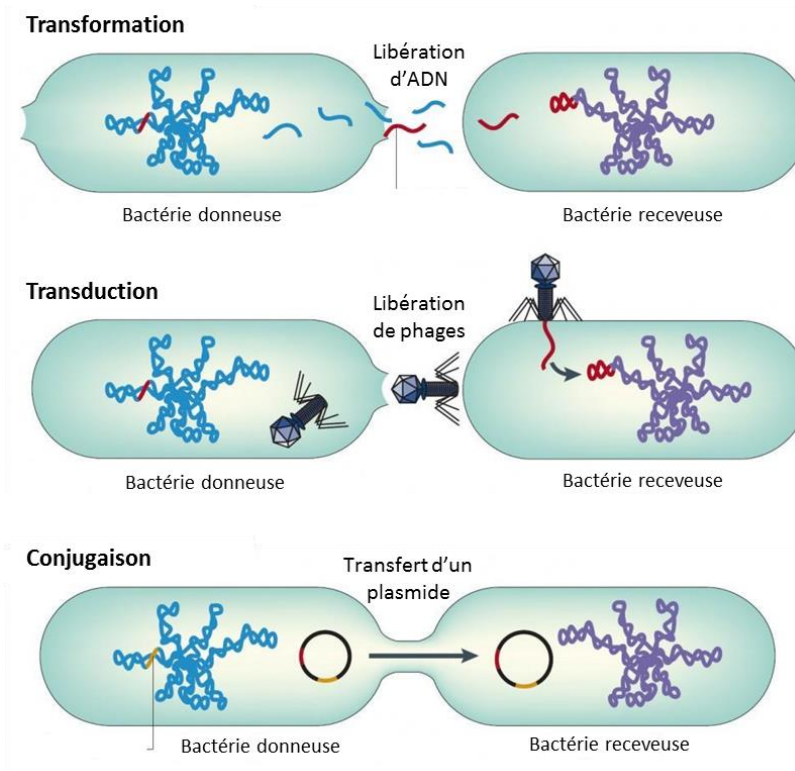


Figure 10 : Principaux mécanismes de HGT : transformation, transduction, conjugaison.

Le génome de la bactérie donneuse est bleu et celui de la receveuse en violet. Les fragments rouges représentent des gènes de résistance à un antibiotique et ceux en jaune un transposon (adaptée de Furuya et Lowy 2006).

Transduction : La transduction correspond à un transfert d'ADN par l'intermédiaire d'un virus appelé phage ou bactériophage (Miller 2001). Chez une bactérie donneuse infectée par un phage, lors de l'encapsidation il peut y avoir incorporation d'ADN de l'hôte dans la capsid. Soit cette encapsidation est aléatoire (transduction généralisée), soit elle concerne les régions d'ADN adjacente au virus qui s'excisent incorrectement (transduction spécialisée). Les phages libérés par lyse vont ensuite infecter d'autres bactéries, ces transferts sont donc dépendants de la gamme d'hôte des phages.

Conjugaison : La conjugaison consiste en un transfert d'ADN par l'intermédiaire d'un pore qui relie une bactérie donneuse à une bactérie receveuse. Contrairement à la transformation et à la transduction, cela nécessite donc un contact direct. Les systèmes de conjugaison sont composés de 3 modules : le relaxosome, le transférosome, et la protéine de couplage qui les connecte (de la Cruz *et al.* 2010).

- Le relaxosome est un complexe nucléoprotéique permettant l'initiation et la terminaison du transfert conjugatif. Il est constitué d'une relaxase, et de protéines accessoires. Ces protéines se fixent au niveau de l'origine de transfert (*oriT*) de l'EGM, ce qui facilite l'accès de la relaxase au site spécifique de coupure de l'ADN (Willetts et Wilkins 1984; Howard *et al.* 1995). La relaxase clive l'ADN et reste ensuite fixée à son extrémité 5'.

- Le transférosome est composé de protéines transmembranaires constituant un pore entre les deux cellules, un système de sécrétion de type IV (SSTIV) joue classiquement ce rôle. Chez certaines bactéries le transférosome code également pour la synthèse d'un filament extracellulaire, le pilus, qui sert à la reconnaissance de la bactérie receveuse (Fronzes *et al.* 2009).

- La protéine de couplage est associée à la membrane proche du SSTIV, et a une affinité pour le relaxosome. L'ADN simple ou double brin (plasmide, ICE, chromosome) est adressé au pore par la relaxase et transféré par un processus actif, nécessitant une hydrolyse d'ATP (Gomis-Rüth *et al.* 2002).

Le SSTIV le mieux décrit est celui d'*Agrobacterium tumefaciens*, bactérie capable d'injecter de l'ADN oncogène chez certaines plantes (Bourras *et al.* 2015). Les gènes codant ce système de sécrétion se trouvent sur un plasmide. Ce dernier se compose d'un opéron virB (virB1-virB11) et d'un opéron virD (virD1-virD4) (Alvarez-Martinez et Christie 2009) (Figure 11). La protéine VirD2 est une relaxase, VirD4 est une protéine de couplage associée à une activité ATPase, VirB4 et VirB11 sont également des ATPases. Les protéines VirB6 à VirB11 s'assemblent pour former le pore de conjugaison. Des homologues à ce système de conjugaison sont retrouvés chez d'autres espèces bactériennes (*Bacillus subtilis*, *Streptococcus agalactiae*) et permettent la formation d'un pore de conjugaison (Goessweiner-Mohr *et al.* 2013; Auchtung *et al.* 2016).

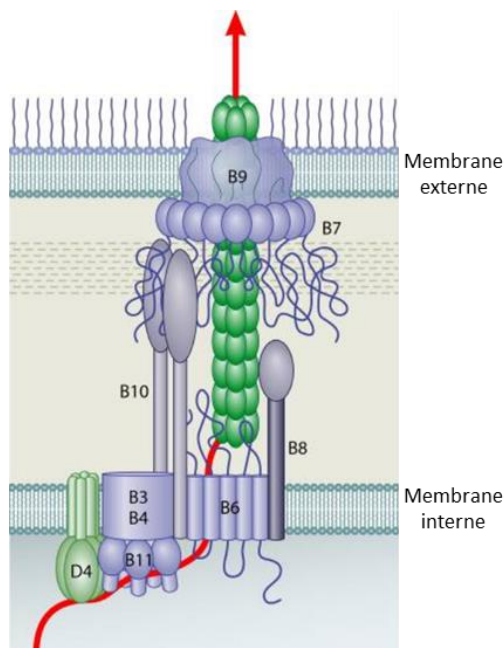


Figure 11 : Architecture d'un SSTIV et voie de translocation de l'ADN.

Représentation du système VirB-VirD chez *A. tumefaciens* avec les différentes protéines qui le composent et la flèche rouge qui montre le chemin suivi par l'ADN. (Alvarez-Martinez et Christie 2009).

D'autres mécanismes de HGT moins répandus sont également décrits chez certaines espèces (García-Aljaro *et al.* 2017). Il s'agit en l'occurrence d'un transfert de gènes via des vésicules membranaires, surtout étudié chez les bactéries à Gram négatif. Des vésicules lipidiques se forment au niveau de la membrane et incorporent des constituants cellulaires comme des protéines, des lipides, des ions mais aussi des acides nucléiques. Ce type de transfert est présent notamment chez *E. coli* O157 avec le transfert de gène codant la Shiga toxine (Kolling et Matthews 1999). Si des observations expérimentales suggèrent des transferts horizontaux pouvant impliquer des vésicules chez les acholeplasmies (Medvedeva *et al.* 2014), ces résultats restent à confirmer.

Le GTA pour "Gene Transfer Agent" est un mécanisme découvert initialement chez la bactérie marine *Rhodobacter capsulatus*, il ressemble à la transduction mais n'est pas directement associé à des phages (Lang *et al.* 2012). Des gènes codant pour une structure similaire à une capsid sont présents au niveau du chromosome bactérien mais ne sont pas soumis à une excision. L'ADN de la cellule est aléatoirement encapsidé, puis relargué lors de la lyse cellulaire. Les GTA dérivent probablement d'anciens prophages dégénérés insérés au niveau du génome.

Enfin, un mécanisme hybride entre transformation et conjugaison a été identifié chez *Thermus thermophilus* (souche HB27), appelé transjugaison (Blesa *et al.* 2017). L'ADN des cellules donneuses est transféré par conjugaison, et les cellules réceptrices récupèrent cet ADN grâce à leur compétence naturelle. Cette libération active d'ADN par conjugaison est dépendante de la protéine TdtA, qui s'associe en hexamère pour former un pore par lequel l'ADN est transféré.

3.1.3. Les éléments génétiques mobiles (EGM)

Les éléments génétiques mobiles comprennent les intégrons, les séquences d'insertions (IS), les transposons, les prophages, les plasmides et les ICE. Certains EGM sont capables de mobilité intra-cellulaire, au niveau du chromosome ou de plasmides (intégron, transposon), alors que d'autres éléments peuvent s'échanger entre bactéries et sont donc capables de mobilité inter-cellulaire ou HGT (prophage, plasmides, ICE). A l'exception des plasmides réplicatifs, les EGM ne sont pas capables de réplication autonome. Leur maintien au sein de la cellule est lié à la réplication du chromosome ou du plasmide réplicatif au sein duquel ils sont insérés. Les EGM peuvent également s'assembler entre eux pour former des éléments plus complexes. Ceci permet le transfert horizontal d'éléments non-autotransférables comme les transposons et les IS par le biais de plasmides ou d'ICE.

3.1.4. Traits adaptatifs ou phénotypes transférés par les EGM

La dissémination d'un EGM par HGT peut apporter un avantage sélectif à la bactérie receveuse via les gènes qu'il contient, comme par exemple, une résistance aux antibiotiques ou aux métaux lourds, des facteurs de virulence, la capacité d'utilisation de nouveaux métabolites ou la colonisation de nouvelles niches environnementales. Les plasmides portent notamment des gènes de résistance aux fluoroquinolones, β -lactamines ou macrolides (Leclercq et Courvalin 2002; Robicsek *et al.* 2006; Brolund et Sandegren 2016). Des gènes de résistance associés à des transposons sont aussi retrouvés, comme la résistance aux tétracyclines, au chloramphénicol ou à la kanamycine grâce aux transposons *Tn10*, *Tn9*, *Tn5* respectivement (Mahillon et Chandler 1998).

Un autre exemple de l'implication du HGT dans la résistance aux ATB est l'EGM de *Salmonella enterica* désigné SGI1, constitué de 43 Kb et contenant des gènes de résistance à 5 antibiotiques appartenant à des familles différentes : streptomycine, ampicilline, chloramphénicol, sulfaméthoxazole, tétracycline (Briggs et Fratamico 1999; Boyd *et al.* 2001).

Cet EGM est responsable d'épidémies de bactéries multi-résistantes chez l'Homme et chez l'animal de *S. enterica* Typhimurium DT104 (Threlfall 2000). Chez *Vibrio cholerae*, les ICE STX de 62 Kb permettent une diffusion de la résistance à la spectinomycine, aux sulfonamides et aux triméthoprimes (Waldor *et al.* 1996). Ces EGM sont aujourd'hui identifiés chez de nombreuses souches de *V. cholerae*, responsables d'épidémies (Hochhut *et al.* 2001).

Le transfert horizontal de la résistance aux antibiotiques est largement répandu et étudié, mais les EGM sont également porteurs d'autres traits phénotypiques, dont quelques exemples sont présentés ici. La bactérie symbiotique *Mesorhizobium loti* contient un îlot de symbiose de 500 Kb situé au niveau d'un ICE, et pouvant être transféré horizontalement (Sullivan et Ronson 1998). Ce groupe de gènes lui confère les propriétés de nodulation et de fixation de l'azote atmosphérique, qui sont essentielles pour établir une symbiose avec les plantes. Un autre phénotype transféré horizontalement est la voie de fermentation du sucrose chez *Salmonella enterica* (Hochhut *et al.* 1997). Un groupe de gènes impliqué dans le catabolisme du sucrose se trouvent sur l'ICE CTnscr94, de 100 pb. Au niveau du génome de *Pseudomonas putida* se trouve plusieurs copies d'un ICE de 105 Kb, appelé *clc* (Ravatn *et al.* 1998). Des gènes situés sur cet ICE permettent la dégradation du chlorocatechol, un composé chloré utilisé comme source de carbone. Le HGT a un impact majeur chez les bactéries, comme chez les mycoplasmes pour lesquels des EGM sont de plus en plus identifiés et des transferts *in silico* ont été détectés.

3.2. Evidence *in silico* de transferts chromosomiques chez les mycoplasmes

Le séquençage et la comparaison des génomes ont permis de détecter des échanges chromosomiques chez les mycoplasmes, alors que leur évolution était vue comme uniquement régressive (Sirand-Pugnet *et al.* 2007a; Citti *et al.* 2018). Ces échanges chromosomiques horizontaux ont été détectés *in silico* entre des espèces de mycoplasmes éloignées phylogénétiquement mais partageant la même niche écologique. Les transferts de gènes ont été détectés essentiellement par la recherche des BBH "*Best Blast Hit*", c'est-à-dire des gènes ayant la plus forte homologie chez d'autres organismes, et par reconstructions phylogénétiques à partir de certains gènes des mycoplasmes. Des échanges importants ont eu lieu entre *M. agalactiae* et des mycoplasmes du cluster Mycoides, tous deux pathogènes de ruminants. D'autres événements de transferts de moindre ampleur ont été détectés entre *M. hominis* et *U. parvum*, tous deux retrouvés chez l'Homme, et entre *M. synoviae* et *M. gallisepticum*, deux pathogènes de volaille (Figure 2).

3.2.1. Entre pathogènes de ruminants

Le génome du pathogène de petit ruminant *M. agalactiae* (souche PG2, clone 55-5) a été entièrement séquencé en 2007 (Sirand-Pugnet *et al.* 2007b). Une comparaison avec les autres génomes de mycoplasmes disponibles a montré que 134 gènes proviennent potentiellement d'échanges avec des mycoplasmes du cluster Mycoides éloigné phylogénétiquement, dont font partie *M. capricolum* subsp. *capricolum* et *M. mycoides* subsp. *capri*. Le cluster Mycoides rassemble des pathogènes de ruminants mais il est appartient au groupe phylogénétique des Spiroplasma alors que *M. agalactiae* fait partie du groupe Hominis (Figure 2). Ces échanges de gènes représentent 18 % du génome de *M. agalactiae*, ils sont regroupés au niveau de 11 régions réparties tout le long du génome, qui présentent une synténie conservée chez les mycoplasmes du cluster Mycoides (Figure 12). Parmi les séquences codantes échangées on retrouve des HP (50), des HP associées à la membrane (24) des transporteurs (18), des enzymes impliquées dans le métabolisme (19), des pseudogènes (21), une protéine de phage (1) et une transposase (1). Sur les 134 gènes associés à des échanges, un grand nombre code pour des protéines membranaires, ce qui suggère leur implication dans la colonisation des petits ruminants.

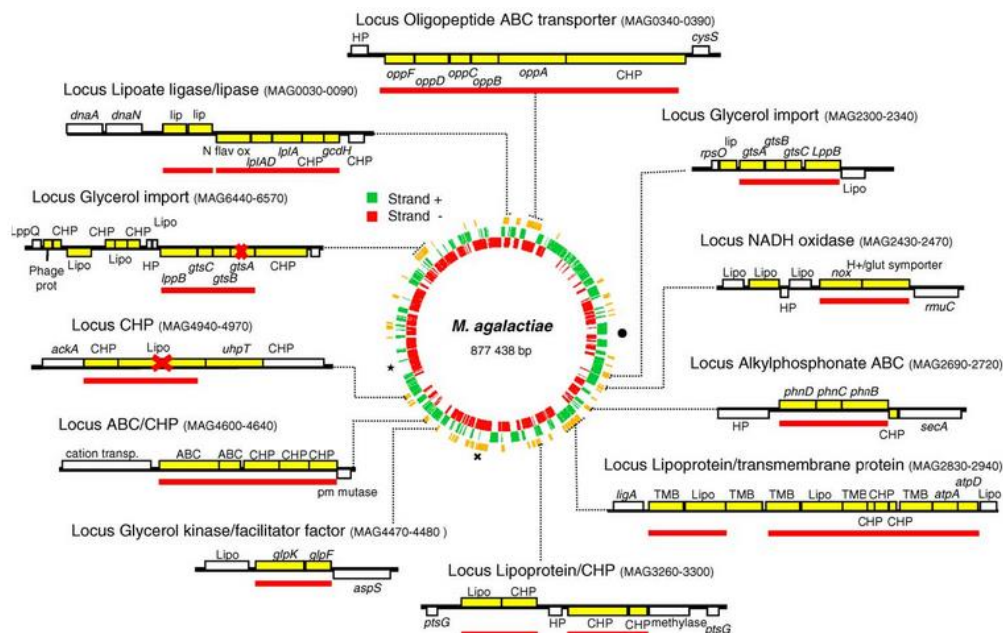


Figure 12 : Régions échangées entre *M. agalactiae* et le cluster Mycoides.

Les positions des 11 régions échangées sont indiquées en jaune sur le génome circulaire de la souche PG2 de *M. agalactiae*. La position des ADNr 16S-23S est indiquée par une étoile et celle de l'ADNr 5S par un point. Le vestige d'ICE de PG2 est lui indiqué par une croix noire. Les gènes qui présentent la même organisation chez les deux partenaires sont indiqués par un trait rouge horizontal. Les pseudogènes sont marqués d'une croix rouge. HP : protéines hypothétiques ; CHP : protéine hypothétique conservée ; Lipo : lipoprotéines ; TMB : protéine prédite comme étant transmembranaire. (Sirand-Pugnet *et al.* 2007b).

Un exemple de gènes transférés est le locus du transporteur ABC d'oligopeptides (*opp*), retrouvé en deux copies chez *M. agalactiae*. Une copie est semblable à celle du cluster Mycoïdes alors que l'autre copie est proche de *M. pulmonis*, un membre du groupe Hominis. Quatre séquences codantes sont suspectées de provenir d'un transfert horizontal entre *M. agalactiae* et des mycoplasmes du groupe Pneumoniae, et 10 CDS avec des bactéries ne faisant pas partie des *Mollicutes*.

3.2.2. Entre pathogènes humains

La souche PG21 de *M. hominis* a été séquencée en 2009 et comparée avec les génomes d'autres mycoplasmes humains, *M. genitalium* G37 et *U. parvum* serovar 3 (Pereyre *et al.* 2009). Lors de cette étude, un échange de gènes entre *M. hominis* et *U. parvum* a été mis en évidence, il correspond à 23 gènes, répartis sur 5 loci et représente 5 % du génome de *M. hominis*. Ces deux mycoplasmes sont retrouvés dans le tractus urogénital de l'Homme mais *M. hominis* fait partie du groupe phylogénétique Hominis et *U. parvum* du groupe Pneumoniae. Les CDS transférées correspondent à des systèmes de restriction/modification de type I (8) et de type III (3), à des HP conservées (7), à des pseudogènes d'une IS (2), à une ATPases (2) et à une intégrase/recombinase (1). Également une CDS codant pour une cytosine DNA méthyltransférase possède un homologue uniquement chez *M. mycoïdes* subsp. *mycoïdes*, et 5 CDS trouvent leurs plus proches homologues chez des bactéries autres que les *Mollicutes* (*Listeria*, *Streptococcus* et *Helicobacter*).

3.2.3. Entre pathogènes de volailles

Lors du séquençage de *M. synoviae* (souche 53) en 2005, l'analyse comparative avec les génomes déjà séquencés a mis en évidence des échanges chromosomiques de gènes entre *M. synoviae* et *M. gallisepticum* (Vasconcelos *et al.* 2005). Ces résultats ont été confirmés par une seconde étude en 2007 (Sirand-Pugnet *et al.* 2007b). Ces deux mycoplasmes sont des pathogènes des volailles mais appartiennent à des groupes phylogénétiques distincts : Hominis pour *M. synoviae* et Pneumoniae pour *M. gallisepticum* (Figure 2). Les échanges chromosomiques sont regroupés au niveau de 14 régions, la plus grande faisant 5,9 Kb et représentant 3 % du génome de *M. synoviae*. La plupart des gènes transférés n'ont pas d'homologie avec d'autres gènes de fonction connue et sont annotés comme protéines hypothétiques (HP). Parmi les autres gènes sont retrouvés : un ABC transporteur, une

peptidase signal de type I, un facteur d'élongation (EF-G) et des transposases. De plus des gènes liés à la pathogénicité ont été transférés entre ces deux espèces de mycoplasmes, il s'agit de sialidases, précédemment identifiées comme facteur de virulence chez *M. alligatoris* (Brown *et al.* 2004). Egalement, un locus comprenant plusieurs hémagglutinines pouvant jouer un rôle dans l'interaction de ces mycoplasmes avec leurs hôtes a été échangé (Noormohammadi *et al.* 1998). Six de ces gènes sont présents uniquement chez *M. synoviae* et *M. gallisepticum* (Sirand-Pugnet *et al.* 2007a), ce qui renforce l'hypothèse qu'ils aient un rôle dans l'adaptation à leurs hôtes, les volailles.

3.2.4. Entre pathogènes de plantes

Des échanges de gènes ont également été mis en évidence entre spiroplasmes, des *Mollicutes* infectant les plantes et leurs insectes vecteurs. Alors que la plupart des spiroplasmes du clade Mirum, sont associés aux insectes, *S. eriocheris* est un pathogène de crustacé, isolé chez le crabe (Wang *et al.* 2011). L'analyse de son génome suggère que 7 % de ses gènes ont été acquis par HGT avec des bactéries n'appartenant pas aux spiroplasmes (Lo *et al.* 2015). La comparaison avec le génome de *S. atrichopogonis*, une espèce très proche de *S. eriocheris* qui infecte les insectes a montré que ces échanges ont eu lieu au niveau de leur ancêtre commun. Chez *S. atrichopogonis* les gènes acquis par HGT sont devenue des pseudogènes, suggérant que leur acquisition n'a entraîné aucun bénéfice (Lo *et al.* 2015) (Figure 13). Ceci vient renforcer l'hypothèse que ces gènes échangés, qui sont principalement des transporteurs et des enzymes métaboliques ont joué un rôle dans le changement d'hôte de *S. eriocheris*.

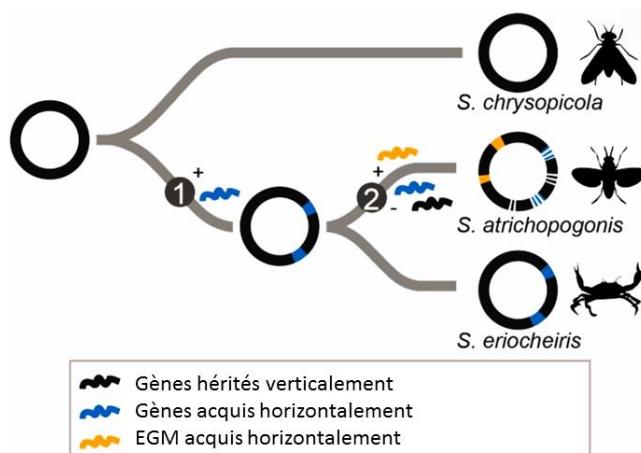


Figure 13 : HGT et évolution du clade Mirum des Spiroplasmes.

Les cercles représentent les chromosomes de chaque spiroplasma et de leur ancêtre. Les organismes hôtes sont représentés à droite. L'étape 1 correspond à l'acquisition de gènes et l'étape 2 à l'acquisition d'EGM et à la formation de nombreux pseudogènes (adaptée de Lo *et al.* 2015).

3.3. Les EGM présent chez les mycoplasmes : rôle majeur des ICE

L'augmentation du nombre de souches séquencées a mis en évidence plusieurs types EGM comme des *IS*, des prophages, des plasmides et des ICE, récapitulés dans les Tableaux 4, 5 et 6 (Marenda 2014). Aujourd'hui, les génomes des mycoplasmes sont donc appréhendés différemment, plutôt comme étant des mosaïques ayant subi différents événements de HGT. L'identification des EGM a contribué à la compréhension de l'évolution des mycoplasmes, mais en revanche les mécanismes de transfert de ces EGM restent pour le moment peu étudiés (Citti *et al.* 2018). A l'exception des *IS* que l'on retrouve chez quasiment tous les *Mollicutes*, les EGM sont surtout présents chez les espèces des groupes phylogénétiques Hominis, Spiroplasma et Acholeplasma, principalement chez les pathogènes de ruminants et de plantes. Seuls ceux retrouvés chez les espèces du genre *Mycoplasma* seront abordés ci-après. Le processus de transduction n'a jamais été démontré chez les mycoplasmes, il en est de même pour la transformation naturelle, mais il est possible de l'induire chez certains mycoplasmes par électroporation ou grâce au PEG (Renaudin *et al.* 2014). A ce jour, le transfert d'ICE par conjugaison a été démontré chez *M. agalactiae* (Dordet Frisoni *et al.* 2013; Baranowski *et al.* 2018).

3.3.1. Les éléments transposables (transposon et *IS*)

Les éléments transposables sont capables de se déplacer par transposition au sein du génome, leur insertion dans des gènes de la voie de biosynthèse des pigments du maïs a été initialement découverte en 1950 par McClintock (McClintock 1950). Les transposons sont des constituants majeurs des génomes eucaryotes (80 % du génome du maïs et 45 % du génome de l'Homme) et sont aussi très présents chez les bactéries (Mahillon et Chandler 1998; Wessler 2006). Les transposons de classe II sont présents chez les eucaryotes et les procaryotes alors que les transposons de classe I, aussi appelés rétrotransposons, nécessitent un intermédiaire ARN et ne sont retrouvés que chez les eucaryotes.

Les séquences d'insertion (*IS*) sont l'élément transposable le plus simple (de 0,7 à 2Kb), elles sont constituées d'un gène codant une transposase qui va reconnaître des séquences répétées inversées (IR, *inverted repeat*) situées en amont et en aval de l'*IS* (Figure 14). Lors de l'insertion d'une *IS* une séquence du génome hôte se retrouve dupliquée de part et d'autre, ce sont les séquences directement répétées (DR, *direct repeat*). Un transposon est formé d'une *IS* et d'un ou plusieurs gènes supplémentaires. Lorsque deux *IS* sont proches, la portion d'ADN qu'elles

encadrent peut se trouver mobilisée, ce sont les transposons composites (Figure 14). La transposition peut se faire par un mode copier-coller (réplication du transposon) ou par un mode couper-coller (excision du transposon). Suivant le site d'insertion du transposon, cela peut entraîner l'interruption ou la dérégulation d'un gène avec des insertions, délétions, duplications ou inversions, qui conduisent à d'importants réarrangements génomiques (Siguier *et al.* 2014).

Séquence d'insertion (IS)



Transposon



Transposon composite



Figure 14 : Organisation des transposons et IS.

Les gènes additionnels, en bleu, peuvent être des gènes de résistance aux antibiotiques. Les flèches blanches indiquent le sens d'orientation des IR (*inverted repeat*) et DR (*direct repeat*).

Des *IS* sont identifiées chez de nombreux mycoplasmes, appartenant aux 4 groupes phylogénétiques (Hominis, Spiroplasma, Pneumoniae, Acholeplasma). Elles appartiennent majoritairement aux familles *IS3*, *IS30*, *IS1634*, également des *IS* issues des familles *IS4*, *IS150*, *IS256* sont retrouvées. Ces séquences ont une taille comprise entre 1 et 15 Kb, et sont juxtaposées par des DR de 3 à 15 pb. Leur transposition se fait principalement par des transposases de type DDE (Mahillon et Chandler 1998). Chez certains mycoplasmes, le nombre d'*IS* est assez important, notamment chez *M. mycoides* subsp. *mycoides* PG1 où les *IS* représentent environ 13 % du génome (Westberg *et al.* 2004) ou chez *M. bovis* PG45 où elles représentent 6,4 % du génome (Lysnyansky *et al.* 2009). En raison de la forte densité de gènes chez les mycoplasmes, ces *IS* sont souvent responsables de l'interruption de séquences codantes. Le nombre d'*IS* est variable suivant la souche étudiée, par exemple pour *M. agalactiae* la souche PG2 contient uniquement un pseudogène d'*IS*, alors que la souche

5632 contient 15 IS appartenant à la famille des IS30 (Nouvel *et al.* 2010) (Tableau 3). Des IS homologues sont retrouvées chez différentes espèces, ce qui suggère des transferts horizontaux de ces éléments, comme entre *M. bovis* et *M. mycoides* subsp. *mycoides* (ISMbov2 et ISMbov3) (Thomas *et al.* 2005), entre *M. hyorhinis* et *M. hyopneumoniae* (IS1221) (Zheng et McIntosh 1995) et entre *M. fermentans* et *M. arthritidis* (IS1630) (Calcutt *et al.* 1999). La position de certaines IS suggère leur implication dans des réarrangements génomiques. Par exemple pour *M. bovis* une inversion de 540 Kb entre les souches PG45 et Hubei-1 se trouve encadrée par deux ISMbov3 chez Hubei-1 (Li *et al.* 2011).

3.3.2. Les prophages

Les bactériophages peuvent suivre un cycle lytique ou lysogénique. Dans ce dernier cas, leur ADN s'intègre au chromosome ou au plasmide de leur bactérie hôte, et ils sont appelés prophages. Lors de dommages cellulaires, un processus d'induction permet au prophage d'être excisé et de rentrer dans un cycle lytique. Il y a libération de capsides qui vont infecter d'autres bactéries permettant un transfert de gènes par transduction (Brown-Jaque *et al.* 2015).

Des prophages ont été décrits chez les groupes phylogénétiques *Acholeplasma*, *Spiroplasma* et *Hominis*. Le prophage MAV1 chez *M. arthritidis* est celui qui a été le plus étudié (Voelker et Dybvig 1999) (Tableau 4). Pour certains ORF (cadre de lecture ouvert) une fonction putative a été attribuée : *repB* correspond à une protéine de réplication ; *marMP* à une ADN méthyltransferase C5 ; *int* à une intégrase et *exiS* à une excisase. Le prophage MAV1 induit une augmentation de la virulence lors d'infection de rats (Voelker *et al.* 1995), la lipoprotéine Vir étant le facteur de virulence candidat (Voelker et Dybvig 1999). D'autres études ont montré que cette même protéine Vir est impliquée dans la résistance à l'infection par d'autres phages (Clapper *et al.* 2004). Le rôle des prophages dans une fonction biologique reste pour le moment à élucider chez les autres espèces de mycoplasmes

Le séquençage de la souche 14628 de *M. agalactiae* a permis la détection d'un prophage de 34 Kb comportant 35 CDS (Tardy *et al.* 2012), dont la forme circulaire a été détectée par PCR. Une séquence homologue de ce prophage est retrouvée chez *M. conjunctivae* HCR/581T et *M. bovigenitalium* cl-51080L (Figure 15).

Tableau 4 : Description des prophages présents chez les *Mycoplasmes*.

Nom de l'espèce	Groupe phylo	Organisme hôte	Souche	Nom de l'EGM	Taille (Kb)	Nombre de CDS	Référence
<i>M. hominis</i>	Hominis	Humain	nd	MV 20P	nd	nd	(Jansson <i>et al.</i> 1982)
			LBD-4	MHoV-1	15,9	14	(Calcutt et Foecking 2015)
<i>M. hyorhinis</i>	Hominis	Porc	GDL-1	Hr-1	nd	nd	(Gourlay <i>et al.</i> 1983)
<i>M. bovirhinis</i>	Hominis	Bovin	nd	Br-1	nd	nd	(Gourlay <i>et al.</i> 1983)
<i>M. pulmonis</i>	Hominis	Rongeur	nd	P1	11,6	11	(Tu <i>et al.</i> 2001)
<i>M. arthritis</i>	Hominis	Rongeur	PG61	MAV1	15,6	15	(Voelker et Dybvig 1999)
<i>M. fermentans</i>	Hominis	Humain	PG18	ØMFV1	23,3	26	(Röske <i>et al.</i> 2004)
			II-29/1	ØMFV1	15,7	18	(Röske <i>et al.</i> 2004)
			K7	ØMFV1	18,8	19	(Röske <i>et al.</i> 2004)
<i>M. hyosynoviae</i>	Hominis	Porc	NPL-3	nd	nd	15	(Bumgardner <i>et al.</i> 2015)
			NPL-4	nd	nd	15	(Bumgardner <i>et al.</i> 2015)
			NPL-5	nd	nd	15	(Bumgardner <i>et al.</i> 2015)
<i>M. agalactiae</i>	Hominis	Petit ruminant	14628	nd	34,0	35	(Tardy <i>et al.</i> 2012)
<i>M. conjunctivae</i>	Hominis	Petit ruminant	HRC/581T	nd	nd	33	(Calderon-Copete <i>et al.</i> 2009)
<i>M. bovis genitalium</i>	Hominis	Ruminants	51080	nd	nd	32	(Manso-Silván <i>et al.</i> 2013)

nd : non déterminé

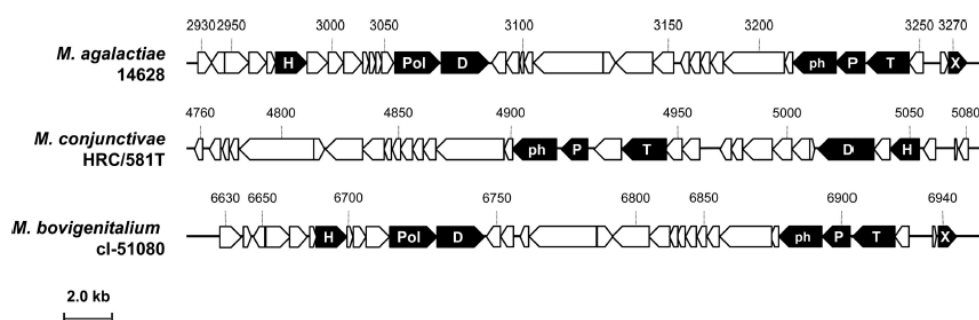


Figure 15 : Organisation génomique et comparaison du prophage de *M. agalactiae* 14628.

Les chiffres indiquent le numéro de CDS correspondant à chaque mycoplasme. Les CDS en noir sont ceux souvent retrouvés chez les phages. H : hélicase ; Pol : ADN polymérase ; D : ADN primase ; ph : protéine de la tête ; P : portail ; T : terminase ; X : Xer. (Tardy *et al.* 2012).

3.3.3. Les plasmides

Les plasmides sont composés d'ADN double brin. Ils se répliquent de manière autonome grâce à la machinerie de réplication de leur cellule hôte qui reconnaît leur origine de réplication (*oriC*). Ils ne sont pas indispensables à la bactérie mais lui confèrent souvent un avantage sélectif, dû aux gènes qu'ils contiennent. Les plasmides sont présents en nombre de copies variable dans une cellule (del Solar et Espinosa 2000). Certains plasmides sont incompatibles, ils ne peuvent pas coexister au sein d'une même cellule et font partie du même groupe d'incompatibilité (Couturier *et al.* 1988). Les plasmides se répliquent suivant deux mécanismes : thêta (θ) ou en cercle roulant (*rolling-circle*). La réplication en thêta démarre par la séparation des deux brins au niveau de l'*oriC*, puis s'ensuit une réplication dans les deux directions (Lilly et Camps 2015). La réplication en cercle roulant est initiée par une coupure simple brin puis l'ADN est synthétisé en tournant autour du brin non coupé (Ruiz-Masó *et al.* 2015). Le transfert horizontal des plasmides peut se faire par conjugaison et débute à l'origine de transfert (*oriT*). A partir de plasmides naturellement présents chez les bactéries, un grand nombre de vecteur ont été construits par génie génétique, ils permettent l'expression de protéines recombinantes, le clonage de gènes, la sélection de bactéries... Certains plasmides, dits conjugatifs ou autotransférables, possèdent tous les gènes nécessaires à leur transfert. D'autres n'en sont pas capables, ils sont dits mobilisables et peuvent être transférés en présence d'un plasmide conjugatif additionnel.

Breton *et al.* en 2012 (Breton *et al.* 2012) ont mené une analyse sur un grand nombre de souches de mycoplasmes pathogènes de ruminants des groupes Hominis et Spiroplasma (194 souches), qui a permis d'identifier de nouveaux plasmides (Figure 16). Des plasmides sont présents chez 5 mycoplasmes proches ou inclus dans le cluster Mycoides (*M. mycoides* subsp. *capri*, *M. capricolum* subsp. *capricolum*, *M. cottewii*, *M. leachii* et *M. yeastii*) (Tableau 5). Aucun plasmide n'a été détecté pour le groupe Hominis (*M. agalactiae* et *M. bovis*). Ces plasmides ont une taille restreinte, comprise entre 1 et 2 Kb. Un seul type de plasmide est présent dans chaque souche, à l'exception de *M. yeastii* GIH chez lequel deux types sont présents. Ces plasmides appartiennent tous à la famille pMV158, isolés pour la première fois chez *Streptococcus agalactiae* (Burdett 1980) à l'exception de pMyBK1. Ils partagent une même organisation, avec deux CDS et deux origines de réplifications (Figure 16). Rep correspond à une protéine de réplication et CopG à une protéine régulatrice du nombre de copie. Les deux origines de réplifications *dso* (*double strand origin*) et *sso* (*single strand origin*)

permettent une réplication en cercle roulant. L'histoire évolutive de ces plasmides n'est pas toujours en accord avec la phylogénie établie à partir de l'ARNr 16S, ce qui suggère qu'ils ont pu être transférés entre les différentes espèces de mycoplasmes de ruminants (Breton *et al.* 2012). Le plasmide pMyBK1 de *M. yeastii* GIH diffère des autres plasmides décrits, il est plus grand (3,4 Kb), ne fait pas partie de la même famille, et n'a pas la même organisation. Une des deux CDS de ce plasmide présente des similitudes avec des relaxases (MobV) impliquées dans la mobilisation de plasmides (Garcillán-Barcia *et al.* 2009). Des dérivés de ce plasmide peuvent se répliquer chez d'autres espèces de mycoplasmes, ce qui a conduit à la construction de vecteur navette compatible avec *E. coli* et certains mycoplasmes (Kent *et al.* 2012).

Chez les bactéries, les plasmides sont largement impliqués dans la résistance aux ATB, mais chez les mycoplasmes, aucun gène additionnel pouvant conférer un avantage sélectif n'a été identifié sur les plasmides. *S. citri*, un pathogène de plantes du groupe Spiroplasma, constitue une exception avec la présence de 6 plasmides, dont un de 35 Kb (pSci6) conférant la capacité de transmission aux insectes vecteurs (Saillard *et al.* 2008).

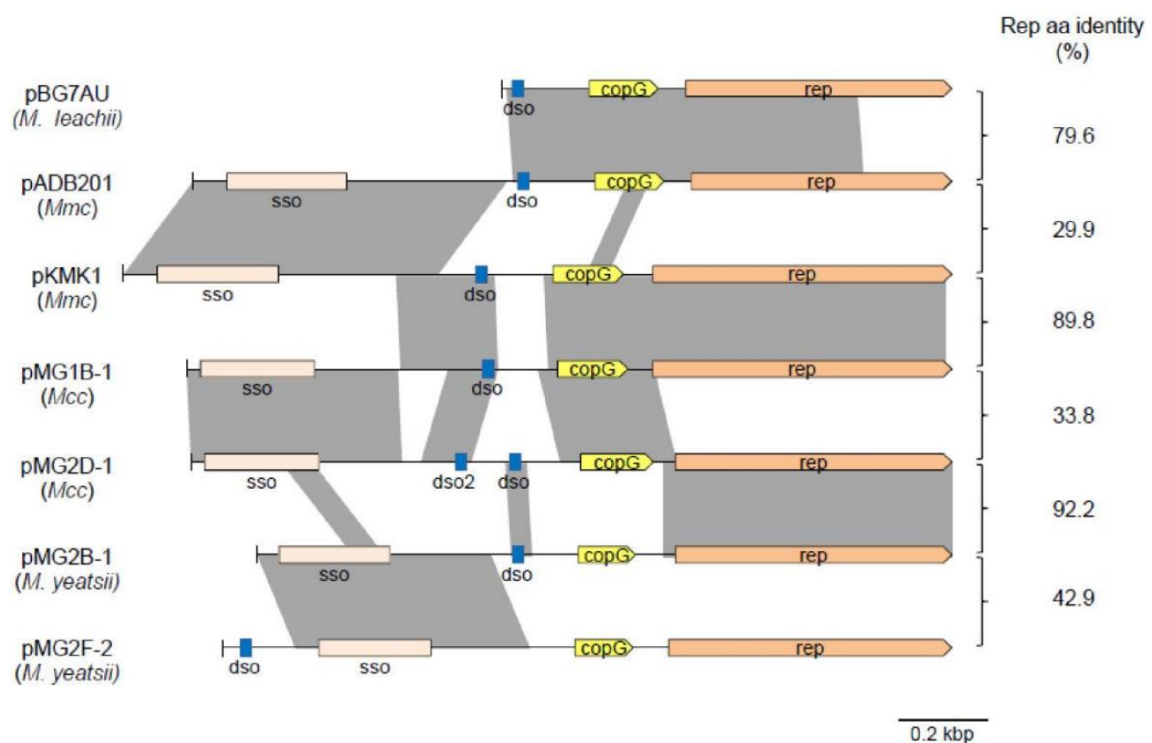


Figure 16 : Comparaison des plasmides de mycoplasmes de ruminants.

Les régions similaires détectées par ACT sont indiquées en gris. Le pourcentage d'acide aminé identique pour Rep est indiqué à droite. Mmc : *M. mycoides* subsp. *capri* ; Mcc : *M. capricolum* subsp. *capricolum* ; sso : *single strand origin* ; dso : *double strand origin* ; copG : régulation du nombre de copie ; rep : réplication (Breton *et al.* 2012).

Tableau 5 : Description des plasmides présents chez les *Mycoplasmes*.

Nom de l'espèce	Groupe phylo	Organisme hôte	Souche	Nom de l'EGM	Taille (Kb)	Nombre de CDS	Référence
<i>M. mycoides</i> subsp. <i>capri</i>	Spiroplasma	Petit ruminant	GC1176-2	pADB201	1,7	2	(Bergemann et Finch 1988)
			GM12	pKMK1	1,9	1*	(King et Dybvig 1992)
			95010	pMmc-95010	1,8	2	(Thiaucourt <i>et al.</i> 2011)
			4343	pMG1C-1	1,8	2	(Breton <i>et al.</i> 2012)
			14227	pMG1A-1	1,9	2	(Breton <i>et al.</i> 2012)
<i>M. capricolum</i> subsp. <i>capricolum</i>	Spiroplasma	Petit ruminant	14250	pMG2A-1	1,6	2	(Breton <i>et al.</i> 2012)
			11186	pMG2D-1	1,7	2	(Breton <i>et al.</i> 2012)
			14425	pMG1B-1	1,7	2	(Breton <i>et al.</i> 2012)
<i>M. cottewii</i>	Spiroplasma	Petit ruminant	VIS	pMG2C-1	1,6	2	(Breton <i>et al.</i> 2012)
			15104	pMG2E-1	1,0	2	(Breton <i>et al.</i> 2012)
<i>M. leachii</i>	Spiroplasma	Bovin	PG50	pBG7AU	1,0	2	(Djordjevic <i>et al.</i> 2001)
<i>M. yeastii</i>	Spiroplasma	Petit ruminant	GIH	pMyBK1	3,4	2	Kent and Calcutt 2008 direct submission
			GIH	pMG2B-1	1,6	2	(Breton <i>et al.</i> 2012)
			11181	pMG2F-1	1,7	2	(Breton <i>et al.</i> 2012)

* une seule CDS est annotée mais deux sont bien présente au niveau du plasmide pKMK1

3.3.4. Les Elément Intégratif et Conjugatif (ICE)

Les ICE sont des éléments intégratifs et conjugatifs, autotransférables. Ils combinent des éléments de transposon, car ils s'intègrent et s'excisent dans les génomes hôtes et également des propriétés de plasmides conjugatifs permettant leur transfert par conjugaison (Burrus *et al.* 2002; Wozniak et Waldor 2010; Bellanger *et al.* 2014). Chez la plupart des bactéries, les ICE intégrés au chromosome sont bordés par des sites répétés attR et attL, leur excision et leur circularisation se fait au niveau de ces sites répétés pour donner le site attP (ou attI) (Figure 17A). Leur intégration passe par une recombinaison du site attP au niveau de l'ICE et du site attB situé sur le chromosome. Le site d'intégration des ICE est généralement spécifique, un très grand nombre s'insère au niveau des gènes codant des ARN de transfert. La recombinaison des ICE au niveau du chromosome se fait généralement par des recombinases à tyrosine, d'autres recombinases de type DDE sont également retrouvées.

Les ICE sont constitués de plusieurs modules (Figure 17A) :

- Un module de recombinaison, codant pour l'intégrase (Int) et d'autres protéines permettant l'excision et l'intégration.

- Un module de conjugaison, où se trouve l'origine de transfert (*oriT*) et les protéines du pore conjugatif, souvent homologue au système de sécrétion de type 4 (SST4).

- Un module de régulation, permettant de contrôler le processus de recombinaison et de conjugaison en fonctions des conditions cellulaires.

Les ICE ont d'abord été considérés comme des éléments se répliquant uniquement grâce au chromosome, mais des études démontrent qu'une réplication extra-chromosomique est possible (Bellanger *et al.* 2014; Burrus 2017). Des structures génomiques dérivant des ICE et qui ont perdu certaines de leur fonction ont été mises en évidence, les IME et les CIME (Wozniak et Waldor 2010). Les IME (Éléments Mobilisables et Intégratifs) n'ont pas de système de conjugaison, ils nécessitent donc celui d'un autre ICE pour assurer leur transfert. Les CIME (Éléments Cis-Mobilisables) n'ont ni système de conjugaison ni intégrase, ils nécessitent l'intégration d'un ICE à leur côté pour être excisés et transférés ensemble. Depuis leurs découvertes, les ICE ont été identifiés comme les EGM les plus abondants chez les procaryotes, ce qui en fait des acteurs majeurs de l'évolution et de l'adaptation des bactéries (Guglielmini *et al.* 2011).

Le premier ICE décrit chez les mycoplasmes est celui de *M. fermentans* PG18, ICEF, présent en 4 copies de 23 Kb qui représentent 8 % du génome (Calcutt *et al.* 2002). Le deuxième exemple a été mis en évidence chez *M. agalactiae* 5632 avec l'ICEA de 27 Kb présent en 3 copies identiques (Marenda *et al.* 2006). Par la suite des ICE ont été recherchés chez d'autres mycoplasmes. Identifiés majoritairement au sein des groupes phylogénétiques Hominis et Spiroplasma, ils semblent très répandus chez les pathogènes de ruminants (*M. mycoides* subsp. *capri*, *M. capricolum* subsp. *capricolum*, *M. conjunctivae*, *M. bovis*, *M. putrefaciens*, *M. auris*, *M. yeatsii*, *M. bovigenitalium* et *M. alkalescens*) (Marenda 2014; Citti *et al.* 2018) (Tableau 6). Au sein du groupe Pneumoniae, seules deux CDS conservées, CDS5 et CDS17, ont été détectées chez *U. diversum* (ATCC 49782) (Marques *et al.* 2015) et aucun ICE ou élément évocateur d'ICE n'a pu être identifié jusqu'à présent chez les mycoplasmes du groupe Acholeplasma (Citti *et al.* 2018).

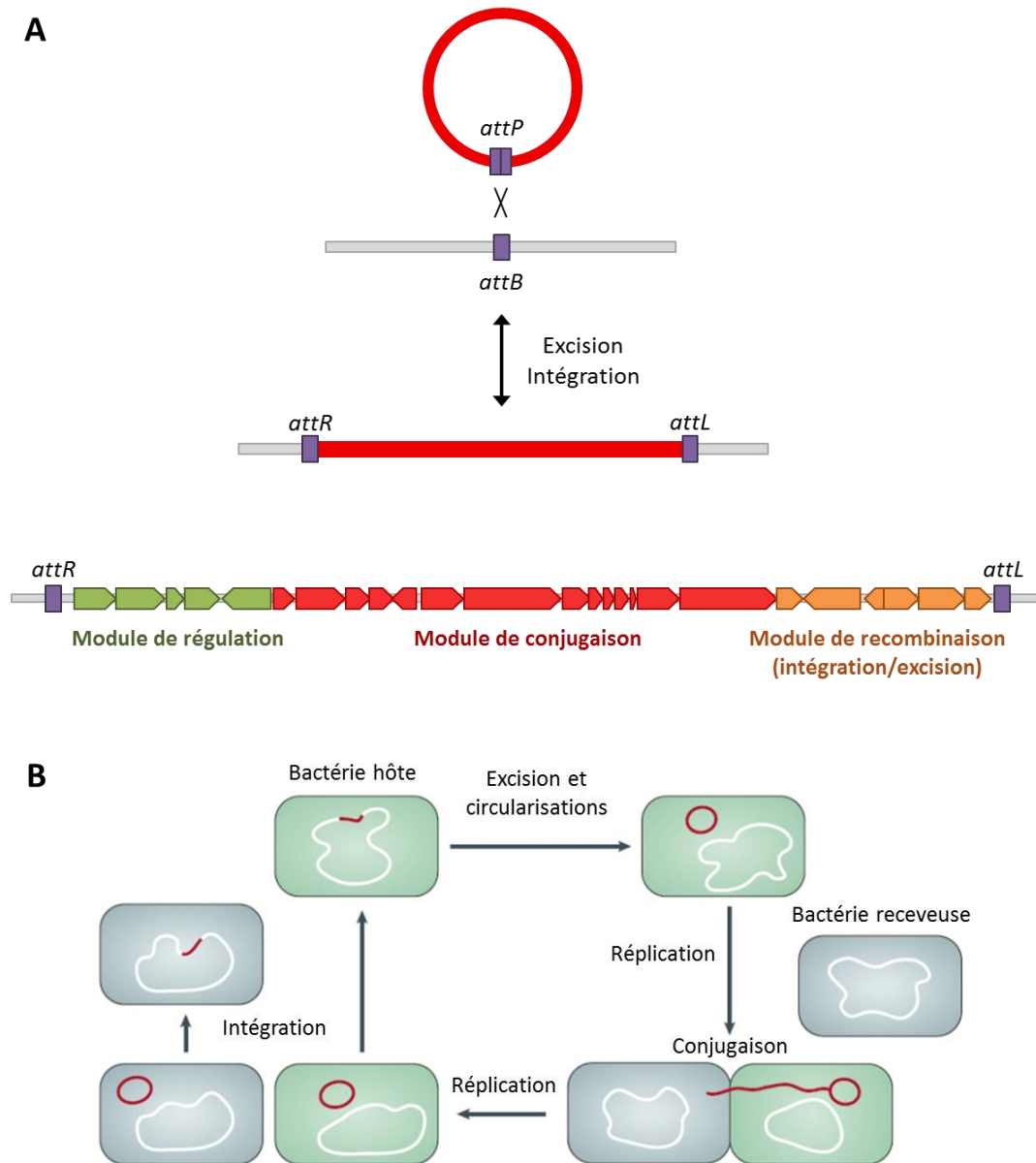


Figure 17 : Organisation et mode de dissémination des ICE.

(A) L'ICE est intégré au chromosome et encadré par les séquences attR et attL. L'excision produit une molécule circulaire, avec une recombinaison entre les sites attR et attL pour donner attP dans l'ICE et attB dans le chromosome. Les ICE sont composés de plusieurs modules indiqués chacun par une couleur. (B) Modèle de dissémination des ICE entre une bactérie donneuse (bactérie hôte) et une bactérie receveuse, suivant les étapes d'excision, de conjugaison et d'intégration (adaptée de Wozniak et Waldor 2010).

Les ICE de mycoplasmes, aussi appelés MICE pour "*Mycoplasma Integrative Conjugative Element*", ont une taille comprise entre 20 et 30 Kb et comportent 20 CDS en moyenne (Tardy *et al.* 2015; Citti *et al.* 2018) (Figure 18). Le nombre d'ICE varie au sein d'une même espèce, par exemple chez *M. fermentans* 7 ICE sont présents chez la souche M64 alors que seulement deux copies sont retrouvées chez la souche JER (Calcutt *et al.* 2002). Des ICE incomplets appelés des vestiges d'ICE (VICE) plus ou moins longs et souvent constitués de pseudogènes sont retrouvés chez certains mycoplasmes (Figure 18). Ces VICE ont pu conserver certaines fonctions et peuvent constituer des IME ou des CIME. Chez *M. agalactiae* la souche 5632 contient deux VICE (5,2 et 1,9 Kb) et la souche PG2 en comporte un seul, plus long (21,9 Kb) (Nouvel *et al.* 2010).

Chez les mycoplasmes de ruminants, deux catégories d'ICE se distinguent par leur organisation (Citti *et al.* 2018). Les ICE de type Spiroplasma, avec pour représentant l'ICEM de *M. mycoides* subsp. *capri* GM12 et ceux de type Hominis, avec pour représentant l'ICEA de *M. agalactiae* 5632 (Figure 18). La différence entre ces deux groupes se fait surtout au niveau de la CDS22, présente chez tous les ICE mais homologue à l'une ou l'autre type d'ICE. Cette classification des ICE en deux types ne suit pas celle des groupes phylogénétiques, surtout au niveau des vestiges d'ICE. Par exemple, le VICE de *M. agalactiae* PG2 est plus proche des ICE du cluster Mycoides que des autres ICE de *M. agalactiae* 5632, il en est de même avec le VICE de PG45 (Sirand-Pugnet *et al.* 2007b; Tardy *et al.* 2015). Ceci suggère que ces vestiges d'ICE proviennent d'événements de HGT entre différentes espèces et font échos aux transferts chromosomiques démontrés *in silico* entre *M. agalactiae* et le cluster Mycoides (Sirand-Pugnet *et al.* 2007b).

La présence d'une forme circulaire de l'ICE est nécessaire à son transfert par conjugaison. Des ICE circulaires ont été détectées par PCR pour tous les mycoplasmes contenant des ICE (Calcutt *et al.* 2002; Vasconcelos *et al.* 2005; Marena *et al.* 2006; Pinto *et al.* 2007; Tardy *et al.* 2015). Également une amplification des sites vides des ICE est cohérente avec leur circularisation et leur excision sur un mode "couper-coller". Comme montré chez *M. agalactiae*, le transfert des ICE se fait à faible fréquence et leur intégration est aléatoire dans le génome (Dordet Frisoni *et al.* 2013; Tardy *et al.* 2015). Du fait de la forte densité de gène chez les mycoplasmes, l'insertion des ICE au niveau de sites-non spécifiques entraîne souvent des interruptions de gènes.

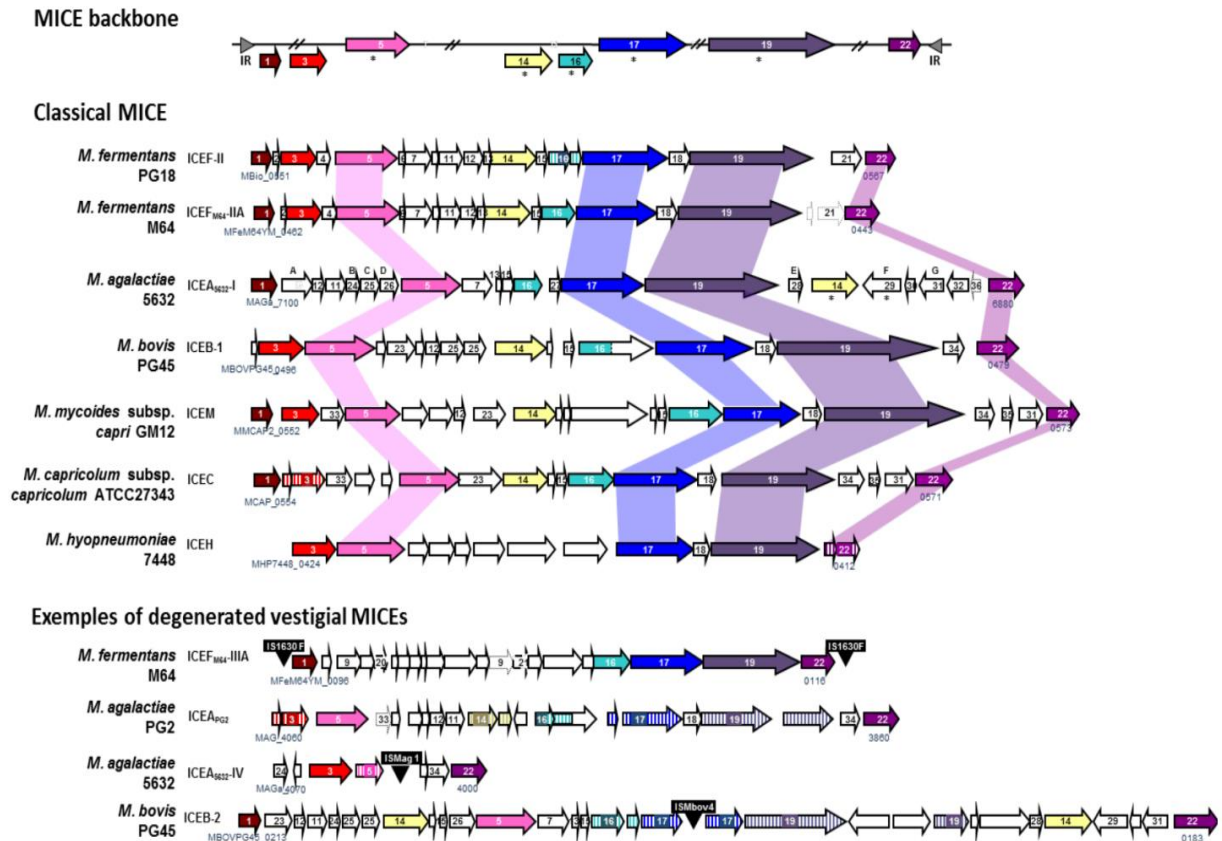


Figure 18 : Organisation génomique et comparaison des ICE chez les mycoplasmes.

Le nom de chaque ICE et de la souche dont il est isolé est indiqué à gauche. Un backbone des ICE se trouve sur la première ligne. Les CDS conservées pour tous les ICE sont sur la ligne, et celles qui sont parfois absentes ou tronquées sont en dessous de la ligne. Les triangles gris à chaque extrémité indiquent la position des IR. Les CDS contenant des domaines transmembranaires sont indiqués par une étoile. La couleur de chaque CDS est conservée pour les différents ICE. Les CDS hachurées représentent les pseudogènes. Pour les vestiges l'insertion d'IS est indiquée par les triangles noirs (Citti *et al.* 2018).

Les ICE des mycoplasmes contiennent une vingtaine de gènes qui présentent peu d'homologies avec les ICE des autres bactéries, excepté pour CDS5 et CDS17. CDS5 correspond à une protéine de couplage, homologue à TraG/VirD4 alors que CDS17 correspond à une ATPase impliquée dans le transport d'ADN, homologue à TraE/VirB4 (Calcutt *et al.* 2002; Marena *et al.* 2006; Tardy *et al.* 2015) (Figure 18). Une fonction putative est associée à certaines autres CDS. CDS22 correspond à une DDE recombinase impliquée dans l'intégration et l'excision des ICE (Dordet Frisoni *et al.* 2013; Guéillot *et al.* 2014). Une protéine SSB, interagissant avec l'ADN simple brin et le protégeant des nucléases est identifiée comme CDS12. CDS14 est une lipoprotéine, dont la présence est essentielle pour le transfert d'ICE chez *M. agalactiae*, elle est également identifiée chez *M. fermentans* (nommé p57) (Calcutt *et*

al. 2002; Dordet Frisoni *et al.* 2013). CDS3 contient des motifs appartenant aux relaxases, CDSG de protéine de partition et CDSH de méthyltransférase. Un backbone minimale de 4 CDS présentes chez tous les mycoplasmes a été défini, il s'agit des CDS5, 17, 19 et 22. D'autres CDS sont également très conservés ce sont les CDS1, 3, 14, 16 (Citti *et al.* 2018) (Figure 18).

Les ICE sont parfois associés à d'autres EGM, comme des *IS*, qui conduisent à des pseudogènes. C'est le cas pour *M. agalactiae* 5632 et *M. bovis* PG45, où des *IS* sont présentes au niveau de la CDS16 de l'ICE_{5632-II} et des CDS17 et 19 de l'ICEB-2. Chez les mycoplasmes du cluster Mycoïdes, des ICE et des plasmides sont présents chez les mêmes souches (Breton *et al.* 2012; Tardy *et al.* 2015), mais aucune interaction entre eux n'a pu être démontrée. Ils présentent cependant des homologies de séquences de 130 pb composées de motif répétés pouvant être impliqués dans des recombinaisons (Thiaucourt *et al.* 2011).

Les fonctions supplémentaires portées par les ICE de mycoplasmes n'ont pas été identifiées pour le moment, à l'exception d'un ICE chez *M. hominis* et *U. diversum* associé à un gène de résistance à la tétracycline *tet(M)*. Ce dernier a récemment été décrit comme homologue à un ICE de *Streptococcus* et différent de ceux retrouvés chez les autres mycoplasmes (Figure 19) (Calcutt et Foecking 2015). Cet ICE de 25 Kb dont la possible circularisation a été démontrée, se trouve chez la souche Spott de *M. hominis*. Il contient un transposon *Tn916* porteur d'un gène de résistance aux tétracyclines, *tet(M)*.

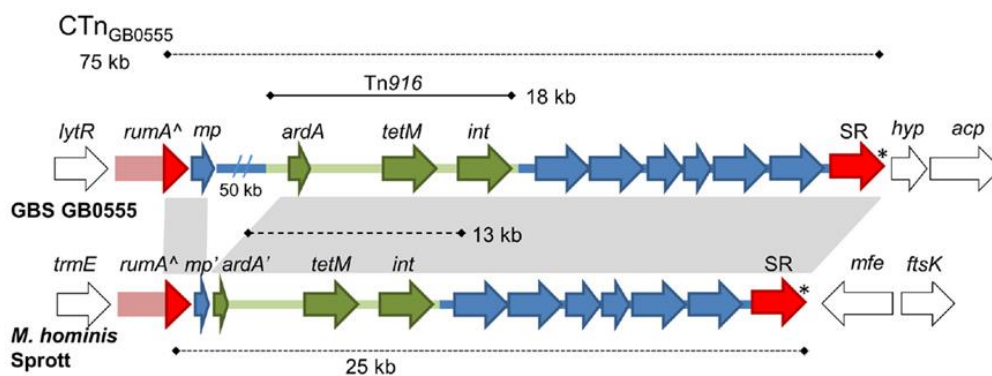


Figure 19 : Transposon conjugatif comportant *tetM* chez *M. hominis* et *S. agalactiae*.

Les flèches vertes correspondent à la région du *Tn916*, les flèches bleues et rouges indiquent la région du transposon conjugatif (ou ICE) chez *M. hominis* Spott et chez *S. agalactiae* GBS GB0555. Les portions grises montrent les régions homologues entre les deux, chez *M. hominis* une délétion est présente entre *mp* et *ardA* (adaptée de Calcutt et Foecking 2015).

Tableau 6 : Description des ICE présents chez les *Mycoplasmas*.

Nom de l'espèce	Groupe phylo	Organisme hôte	Souche	Nom de l'EGM	Taille (Kb)	Nombre de CDS	Référence
<i>M. mycoides</i> subsp. <i>capri</i>	Spiroplasma	Petit ruminant	GM12	ICEM	29,6	21	(Lartigue <i>et al.</i> 2009)
			95010	ICEMx2	30,0	21	(Thiaucourt <i>et al.</i> 2011)
<i>M. capricolum</i> subsp. <i>capricolum</i>	Spiroplasma	Petit ruminant	CK	ICEC	23,8	17	Glass 2005 direct submission
<i>M. conjunctivae</i>	Hominis	Petit ruminant	HRC/581	ICECJ x2	28,5	23	(Calderon-Copete <i>et al.</i> 2009)
<i>M. hyopneumoniae</i>	Hominis	Porc	7448	ICEH	22,3	14	(Vasconcelos <i>et al.</i> 2005)
			168	ICEH	23,4	19	(Liu <i>et al.</i> 2011)
			232	ICEH	22,5	19	(Minion <i>et al.</i> 2004)
<i>M. fermentans</i>	Hominis	Humain	PG18	ICEF x4	23,2	22	(Calcutt <i>et al.</i> 2002)
			JER	ICEF x2	24,4	23	(Rechnitzer <i>et al.</i> 2011)
			M64	ICEF x7	22,7-26,0	22-26	(Shu <i>et al.</i> 2011)
<i>M. agalactiae</i>	Hominis	Petit ruminant	5632	ICEA x3	27,2	23	(Marenda <i>et al.</i> 2006)
			PG2	vestige	nd	nd	(Sirand-Pugnet <i>et al.</i> 2007b)
<i>M. bovis</i>	Hominis	Bovin	PG45	ICEB	37,1	22	(Wise <i>et al.</i> 2011)
			Hubei	vestige	nd	nd	(Li <i>et al.</i> 2011)
			HB0801	vestige	nd	nd	(Qi <i>et al.</i> 2012)
<i>M. putrefacien</i>	Spiroplasma	Petit ruminant	9231	nd	nd	nd	(Dupuy <i>et al.</i> 2013)*
<i>M. auris</i>	Hominis	Petit ruminant	15026	nd	nd	nd	(Dordet-Frisoni <i>et al.</i> 2013)*
<i>M. yeatsii</i>	Spiroplasma	Petit ruminant	13926	nd	nd	nd	(Dordet-Frisoni <i>et al.</i> 2013)*
<i>M. bovigentialium</i>	Hominis	Bovin	51080	nd	nd	nd	(Manso-Silvan <i>et al.</i> 2013)*
<i>M. alkalescens</i>	Hominis	Bovin	14918	nd	nd	nd	(Manso-Silvan <i>et al.</i> 2013)*

nd : non determine ; * sequence draft

3.4. Demonstration *in vitro* de HGT par conjugaison chez *M. agalactiae*

3.4.1. Transfert des ICE de mycoplasmes

Le transfert horizontal des ICE n'a ete a ce jour explore que chez *M. agalactiae*. Le premier transfert a ete demonstre entre la souche 5632, qui possede 3 copies d'ICE et la souche PG2, depourvue d'ICE (Dordet Frisoni *et al.* 2013). Des experiences d'accouplement (appele *mating*) ont ete realisees avec la souche 5632 portant un marqueur de resistance insere dans l'ICE et la

souche PG2 portant un autre marqueur de résistance inséré dans le chromosome (Figure 20). Les phénomènes de transfert conjugatif observés se produisent à une fréquence très faible, de l'ordre de 1.10^{-8} transconjugants/CFU donneur, et l'intégration des ICE se fait de façon aléatoire dans le génome. L'acquisition d'un ICE confère des capacités de conjugaison à la cellule receveuse. En effet, la souche PG2 qui a acquis un ICE peut ensuite la transférer à d'autres souches (Figure 20).

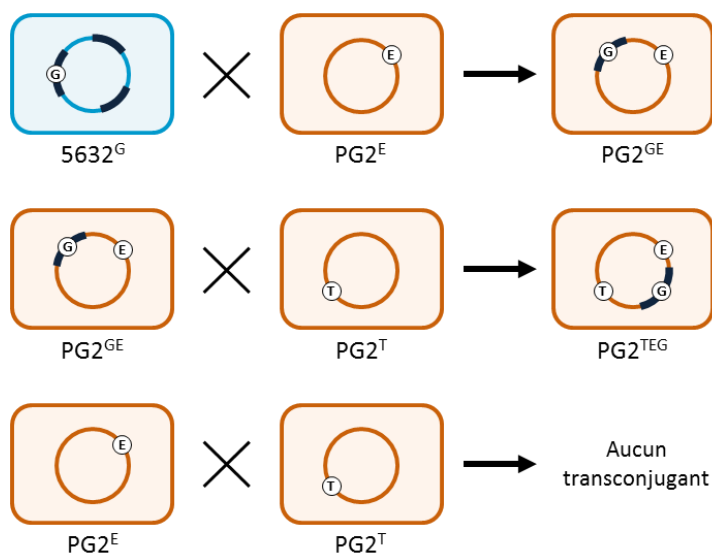


Figure 20: Transfert de l'ICE chez *M. agalactiae*.

Expériences de *mating* réalisées avec les souches 5632 et PG2. Les petits ronds indiquent les marqueurs de résistance à la gentamicine (G) à l'enrofloxacin (E) et à la tétracycline (T). Les ICE sont représentés en noir (adaptée de Dordet Frisoni *et al.* 2013).

Les ICE s'excisent et se circularisent par recombinaison au niveau de DR de 8 pb, formant ainsi une jonction de 6 pb. L'excision et la recombinaison des ICE se font grâce à la CDS22, une recombinase de type DDE, faisant partie de la famille des "*Mutator like*" (Guérillot *et al.* 2014). Ce type de transposase est principalement retrouvée associée à des transposons chez les autres bactéries. Une DDE transposase est également impliquée dans le transfert de certaines ICE chez *Streptococcus* mais dans ce cas leur intégration s'effectue au niveau de régions promotrices, et non aléatoirement comme chez les mycoplasmes (Brochet *et al.* 2009). Un locus de 7 protéines comportant des domaines transmembranaires (CDS5-CDS19) est porté par l'ICE, il contient les CDS5 et CDS17 (Figure 21). CDS5 est une protéine de couplage et CDS17 une ATPase permettant le transport d'ADN. Ce locus pourrait constituer le pore de conjugaison par lequel l'ADN est transféré.

Des expériences de mutagénèse par transposition ont permis d'identifier de nombreuses CDS essentielles au transfert de l'ICE (CDS1, A, C, 5, 7, 15, 16, 17, 19, 14, 30, G, 22) (Figure 21). La CDS14 est une lipoprotéine exprimée à la surface et requise pour le transfert. Son inactivation peut être complétée par un vecteur contenant le gène sauvage qu'il soit chez la donneuse ou chez la receveuse. Son rôle supposé est d'établir un contact entre les cellules afin d'initier le processus de transfert (Dordet Frisoni *et al.* 2013; Baranowski *et al.* 2018). Lorsque plusieurs copies d'ICE sont présentes, comme c'est le cas chez la souche 5632, un phénomène de trans-complémentation entre les différentes copies d'ICE a lieu pour la CDS14, mais pas pour d'autres CDS essentielles au transfert, comme les CDS5 et CDS22 (Baranowski *et al.* 2018).

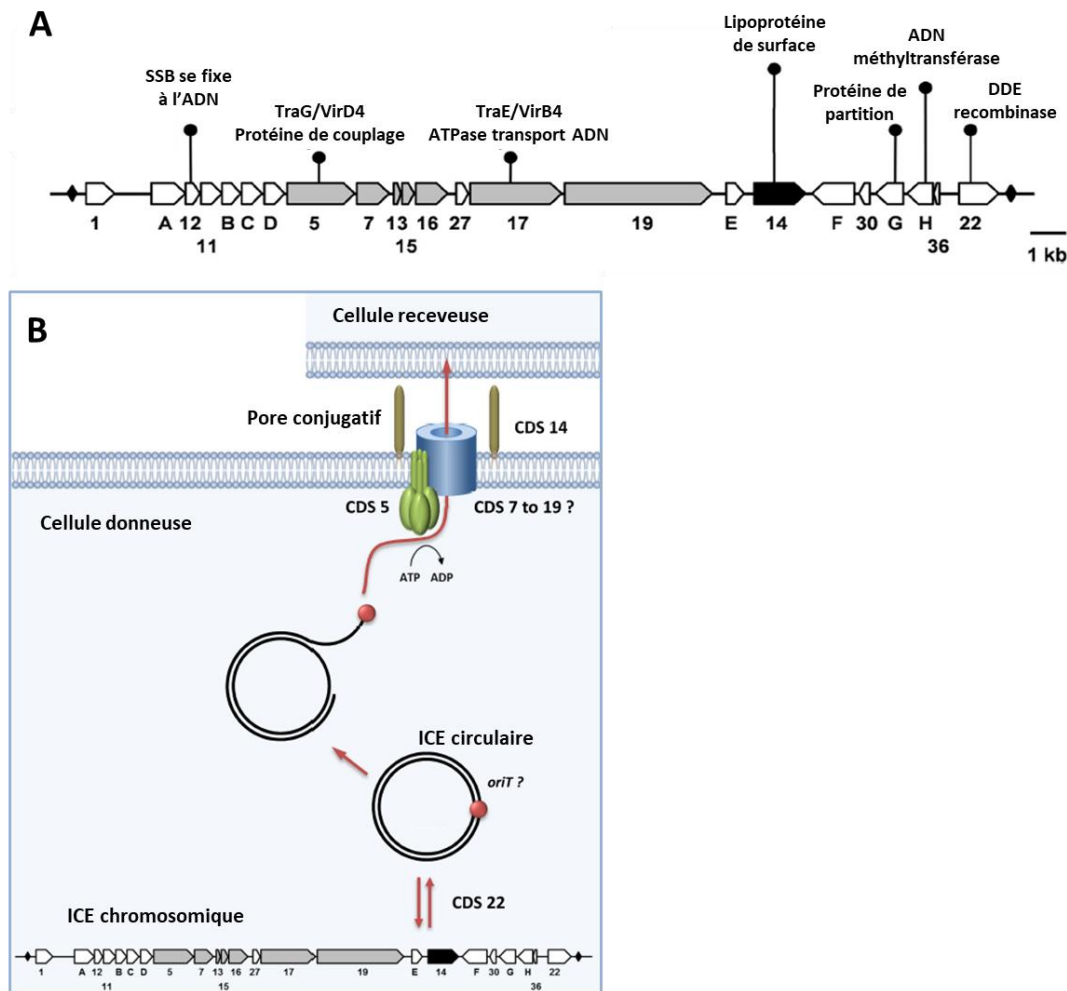


Figure 21 : Organisation et fonctions des CDS des ICE de la souche 5632.

(A) Le nombre ou la lettre correspondant à chaque CDS est indiqué dessous. Les CDS gris contiennent des domaines transmembranaires. La lipoprotéine, CDS14 est représentée en noir. (B) Schéma représentant le modèle actuel du transfert des ICE chez les mycoplasmes (adaptée de Baranowski *et al.* 2018).

3.4.2. Transfert chromosomique

Au cours des études sur le transfert d'ICE chez *M. agalactiae*, certains résultats ont suggéré un transfert d'ADN chromosomique, appelé MCT (*Mycoplasma Chromosomal Transfer*) mais dans le sens opposé au transfert d'ICE : de la souche PG2 (ICE négative) vers la souche 5632 (ICE positive) (Dordet-Frisoni *et al.* 2014). Des expériences de *mating* entre les souches PG2 et 5632, portant un marqueur de sélection à l'extérieur des ICE, ont permis d'obtenir des doubles résistants à une fréquence de 1.10^{-8} transconjugants/CFU donneur. Le MCT est résistant à l'ajout de DNase dans le milieu, ce qui indique qu'il se produit par conjugaison. Il a lieu en même temps que le transfert de l'ICE au sein de la même population et à la même fréquence, mais l'ADN transféré n'est pas directement associé aux ICE. Le MCT consiste au remplacement de larges régions chromosomiques (pouvant atteindre 80 Kb) par son équivalent chez la donneuse (PG2) par recombinaison homologe, comme observé lors d'un *crossing over*. L'analyse par séquençage d'un pool de transconjugants a révélé que n'importe quelle région du génome de 5632 pouvait être remplacée par sa contrepartie PG2 (Dordet-Frisoni *et al.* 2014) (Figure 22).

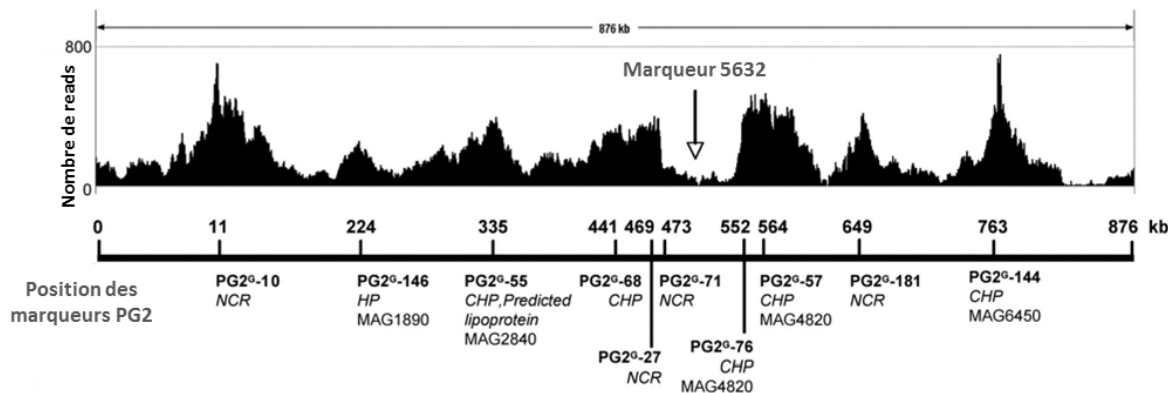


Figure 22 : Tout le génome de PG2 peut être transféré chez 5632.

A l'issue de *mating* entre la donneuse PG2 et la receveuse 5632, un pool de 200 transconjugants a été séquençé. Les *reads* correspondant à la souche PG2, donc aux fragments échangés, alignés sur le génome de PG2 sont visualisés avec IGV. Dix PG2 marquées par la gentamicine (G) et une 5632 marquée à la tétracycline (position relative indiquée par la flèche) ont été utilisées comme souches parentales (adaptée de Dordet-Frisoni *et al.* 2014).

Afin d'étudier les événements de recombinaison indépendamment des étapes de conjugaison et de transfert, des expériences de fusion membranaire à l'aide de PEG (polyéthylène glycol) ont été réalisées. Dans ces conditions, la souche 5632 reste toujours receveuse et la PG2

donneuse. L'ajout de PEG permet également d'obtenir des échanges d'ADN entre deux souches PG2, soit en l'absence d'ICE fonctionnel. Ceci suggère fortement que la recombinaison génomique n'est pas liée à la présence d'ICE. L'implication des ICE dans le MCT serait d'établir un contact cellulaire pour initier le transfert (Dordet-Frisoni *et al.* 2014).

Des mécanismes de transfert chromosomique ont été mis en évidence chez d'autres bactéries, comme par exemple le transfert *High frequency recombination* (Hfr ou *oriT*) identifié au départ chez *E. coli*. Dans le transfert Hfr, le plasmide conjugatif F peut s'intégrer au chromosome. Dans ce cas, une portion du plasmide, débutant à l'*oriT* accompagnée de régions chromosomiques est transférée (Lloyd et Buckman 1995). Ces régions chromosomiques suivent un gradient : plus elles sont proche de l'*oriT*, plus elles ont de chance d'être transférées. En général, le plasmide F n'est pas transmis dans sa totalité, la cellule receveuse n'acquiert donc pas de capacité de transfert. Le MCT de *M. agalactiae* reste atypique car il n'est pas physiquement associé à un EGM et toutes les régions sont transférables avec la même fréquence.

Chez *M. pulmonis*, une étude a montré le transfert d'un transposon par conjugaison (résistant à l'ajout de DNase) avec ou sans PEG (Teachman *et al.* 2002). Récemment, des transferts chromosomiques ont été mis en évidence chez un autre mycoplasme, *M. genitalium* (Torres-Puig *et al.* 2018) mais uniquement en condition de surexpression du facteur de transcription alternatif sigma (σ^{20}) et en présence de *recA*, impliqués dans la recombinaison. Aucun ICE n'a été détecté à ce jour chez *M. pulmonis* ou *M. genitalium* (Citti *et al.* 2018).

4. Objectifs et contexte de l'étude

Contexte

Comme pour beaucoup de bactéries, une augmentation de la résistance aux ATB est observée chez les mycoplasmes. Les bases moléculaires de cette résistance sont décrites par de nombreuses publications (voir les revues Waites *et al.* 2014; Lysnyansky et Ayling 2016; Gautier-Bouchardon 2018). Il s'agit essentiellement de mutations au niveau des gènes codant les cibles chez des souches de mycoplasmes résistantes isolées chez l'Homme et l'animal ou expérimentalement générées *in vitro*. Chez notre modèle d'étude, *M. agalactiae*, le nombre d'isolats résistants est relativement faible pour le moment (Poumarat *et al.* 2016) mais des niveaux élevés de résistance sont constatés chez les isolats actuels d'une espèce phylogénétiquement très proche, *M. bovis* (Gautier-Bouchardon *et al.* 2014).

Si les HGT constituent des mécanismes d'acquisition de la résistance dominants chez la plupart des autres bactéries, leurs rôles dans la résistance restent peu étudiés chez les mycoplasmes, à l'exception d'un transposon conjugatif porteur du gène *tet(M)* conférant la résistance aux tétracyclines (Roberts *et al.* 1985; Roberts et Kenny 1986; Dégrange *et al.* 2008). Des analyses *in silico* ont démontré le rôle potentiel des HGT dans l'histoire évolutive de ces bactéries minimales (Vasconcelos *et al.* 2005; Sirand-Pugnet *et al.* 2007b; Pereyre *et al.* 2009). Des transferts d'éléments conjugatifs (ICE) (Dordet Frisoni *et al.* 2013) mais aussi du transfert chromosomique (Dordet-Frisoni *et al.* 2014) ont récemment été mis en évidence chez *M. agalactiae*. La description et l'impact de ces mécanismes restent à explorer et *M. agalactiae* apparaît comme un modèle de choix pour l'étude du HGT chez les mycoplasmes. La caractérisation de l'implication du HGT dans le phénomène d'antibiorésistance chez les mycoplasmes pourrait permettre une meilleure détection, surveillance et prévention de la résistance.

Au-delà du problème de santé publique posé par ce phénomène auxquels les mycoplasmes ne font pas exception, l'étude des mécanismes d'acquisition de la résistance est aussi une voie de compréhension de mécanismes adaptatifs et évolutifs plus larges. Les mycoplasmes semblent avoir évolué vers une extrême spécialisation ou adaptation à leurs hôtes et certains sont des pathogènes majeurs et de véritables fléaux pour les filières touchées. La lutte contre les mycoplasmes bénéficierait d'une meilleure compréhension de leurs capacités d'adaptation, de l'acquisition de facteurs de virulence au franchissement de la barrière d'espèce hôte. Ainsi, nos

travaux s'inscrivent dans la problématique de l'antibiorésistance pour un enjeu de santé publique, celle de la lutte contre les maladies bactériennes avec des enjeux de santé humaine et animale, et celle plus fondamentale de la compréhension de processus biologiques d'adaptation et d'évolution de ces pathogènes minimaux.

Objectifs

Les objectifs de nos travaux peuvent se décliner autour des points suivants :

- Comprendre l'acquisition de la résistance aux FQ chez *M. agalactiae* par mutations spontanées *in vitro*.
- Observer *in vitro* et analyser des transferts horizontaux de la résistance aux FQ par MCT chez *M. agalactiae*.
- Comparer l'acquisition de la résistance aux FQ par mutations spontanées et par MCT.
- Evaluer les conséquences de ces deux mécanismes (mutation ou MCT) sur certains phénotypes comme le niveau de résistance aux FQ et le fitness.
- Etudier l'évolution expérimentale avec ou sans maintien d'une pression de sélection de différentes populations bactériennes.
- Avoir une meilleure compréhension de l'impact et des mécanismes du MCT chez *M. agalactiae*.

Démarche

Une sélection de mutants, spontanément résistants à l'enrofloxacin (Fluoroquinolone) a été effectuée, par des passages successifs en présence de concentrations croissantes d'ATB. En plus du séquençage des QRDR, la séquence complète de chaque mutant a été obtenue par Illumina, et une analyse de variants a été réalisée. Ce séquençage génomique total a fourni des informations supplémentaires au niveau de la dynamique et de l'évolution à court terme des génomes de mycoplasmes. Des transconjugants ont été sélectionnés par accouplements (*mating*) entre un clone sensible et un mutant spontané résistant aux FQ. Les gènes qui codent les cibles des FQ sont situés à 3 loci distants sur le génome de *M. agalactiae*, ce qui engendre une pression de sélection sur plusieurs positions du génome. Les transconjugants ont été séquencés entièrement par Illumina et une méthode a été mise au point pour définir les

fragments transférés. La stabilité génomique et phénotypique des mutants et des transconjugants ont été évaluées par passages en culture sans pression de sélection. Le niveau de résistance des mutants, des transconjugants et de leurs passages a été mesuré (CMI avec dilutions en milieu gélosé) et leur fitness a été déterminé par des tests de compétition. Pour évaluer l'étendue de ces observations et mieux comprendre l'impact et les mécanismes du MCT, des transconjugants obtenus dans différentes conditions de sélection ont été analysés et comparés. Plusieurs combinaisons de marqueurs ATB à différentes positions du génome ont été utilisées. Afin d'étudier les événements de recombinaison indépendamment des étapes de conjugaison et de transfert, des expériences de fusion artificielle des membranes par PEG ont aussi été réalisées. Les souches parentales de mutants, de transconjugants et d'hybrides PEG, à la base de nos expériences, ont été séquencées et comparées.

La présentation de nos travaux est divisée en trois chapitres au format d'articles scientifiques. La comparaison des mécanismes d'adaptation (mutations vs transfert chromosomique) est abordée au 1^{er} chapitre et une caractérisation plus fine du transfert chromosomique au 2^{ème} chapitre. Le 3^{ème} chapitre explique et détaille plus précisément les éléments méthodologiques utilisés pour l'analyse des données de génomique tout au long de ce travail de thèse et aborde les résultats supplémentaires non inclus dans les publications issues des 2 chapitres précédents. Chaque chapitre comporte sa propre discussion et une discussion globale visant à mettre l'ensemble des résultats en perspectives termine le manuscrit.

5. Références

- Agnone A, La Manna M, Sireci G, Puleio R, Usticano A, Ozdemir U, *et al.* A comparison of the efficacy of commercial and experimental vaccines for contagious agalactia in sheep. *Small Ruminant Research*. 1 mai 2013;112(1):230-4.
- Alekshun MN, Levy SB. Molecular Mechanisms of Antibacterial Multidrug Resistance. *Cell*. 23 mars 2007;128(6):1037-50.
- Alvarez-Martinez CE, Christie PJ. Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol Mol Biol Rev*. déc 2009;73(4):775-808.
- Amram E, Mikula I, Schnee C, Ayling RD, Nicholas R a. J, Rosales RS, *et al.* 16S rRNA Gene Mutations Associated with Decreased Susceptibility to Tetracycline in *Mycoplasma bovis*. *Antimicrob Agents Chemother*. 2 janv 2015;59(2):796-802.
- Andersson JO. Lateral gene transfer in eukaryotes. *Cell Mol Life Sci*. juin 2005;62(11):1182-97.
- Antunes A, Rainey FA, Wanner G, Taborda M, Pätzold J, Nobre MF, *et al.* A New Lineage of Halophilic, Wall-Less, Contractile Bacteria from a Brine-Filled Deep of the Red Sea. *J Bacteriol*. mai 2008;190(10):3580-7.
- Antunes NT, Assunção P, Poveda JB, Tavío MM. Mechanisms involved in quinolone resistance in *Mycoplasma mycoides* subsp. *capri*. *The Veterinary Journal*. juin 2015;204(3):327-32.
- Auchtung JM, Aleksanyan N, Bulku A, Berkmen MB. Biology of ICEBs1, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid*. 2016;86:14-25.
- Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of Pneumococcal types. *J Exp Med*. 1 févr 1944;79(2):137-58.
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*. 11 août 2000;289(5481):905-20.
- Baranowski E, Dordet-Frisoni E, Sagné E, Hygonenq M-C, Pretre G, Claverol S, *et al.* The Integrative Conjugative Element (ICE) of *Mycoplasma agalactiae*: Key Elements Involved in Horizontal Dissemination and Influence of Coresident ICEs. *MBio*. 3 juill 2018;9(4).
- Barré A, de Daruvar A, Blanchard A. MolliGen, a database dedicated to the comparative genomics of Mollicutes. *Nucleic Acids Res*. 1 janv 2004;32(Database issue):D307-310.
- Bébéar CM, Renaudin H, Charron A, Bové JM, Bébéar C, Renaudin J. Alterations in topoisomerase IV and DNA gyrase in quinolone-resistant mutants of *Mycoplasma hominis* obtained in vitro. *Antimicrobial agents and chemotherapy*. 1998;42(9):2304-2311.
- Bellanger X, Payot S, Leblond-Bourget N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiol Rev*. 1 juill 2014;38(4):720-60.

- Bergemann AD, Finch LR. Isolation and restriction endonuclease analysis of a mycoplasma plasmid. *Plasmid*. janv 1988;19(1):68-70.
- Bergonier D, Berthelot X. Mycoplasmoses des petits ruminants: Le syndrome de l'agalactie contagieuse. 2008;161(161):167-77.
- Bergonier D, Poumarat F. Agalactie contagieuse des petits ruminants: epidemiologie, diagnostic et contrôle. *Rev Sci Tech*. 1996;15:1431-75.
- Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJV. Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*. 1 déc 2014;13(1):42-51.
- Blanchard A, Bebear C. Mycoplasmas of Humans. In: *Molecular biology and pathogenicity of Mycoplasmas*. Kluwer Academic / Plenum Publishers. New York, NY: Razin Shmuel and Richard Herrmann; 2002. p. 45-71.
- Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. The transjugation machinery of *Thermus thermophilus*: Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet*. mars 2017;13(3):e1006669.
- Bourras S, Rouxel T, Meyer M. *Agrobacterium tumefaciens* Gene Transfer: How a Plant Pathogen Hacks the Nuclei of Plant and Nonplant Organisms. *Phytopathology*. oct 2015;105(10):1288-301.
- Boyd D, Peters GA, Cloeckert A, Boumedine KS, Chalus-Dancla E, Imberechts H, *et al*. Complete nucleotide sequence of a 43-kilobase genomic island associated with the multidrug resistance region of *Salmonella enterica* serovar Typhimurium DT104 and its identification in phage type DT120 and serovar Agona. *J Bacteriol*. oct 2001;183(19):5725-32.
- Breton M, Tardy F, Dordet-Frisoni E, Sagne E, Mick V, Renaudin J, *et al*. Distribution and diversity of mycoplasma plasmids: lessons from cryptic genetic elements. *BMC microbiology*. 2012;12(1):257.
- Bridé J, Donatien A. Le microbe de l'agalaxie du mouton et de la chèvre. *Annales de l'Institut Pasteur*. 1925;29:925-51.
- Briggs CE, Fratamico PM. Molecular characterization of an antibiotic resistance gene cluster of *Salmonella typhimurium* DT104. *Antimicrob Agents Chemother*. avr 1999;43(4):846-9.
- Broaders E, Gahan CGM, Marchesi JR. Mobile genetic elements of the human gastrointestinal tract. *Gut Microbes*. 1 juill 2013;4(4):271-80.
- Brochet M, Da Cunha V, Couvé E, Rusniok C, Trieu-Cuot P, Glaser P. Atypical association of DDE transposition with conjugation specifies a new family of mobile elements. *Mol Microbiol*. févr 2009;71(4):948-59.
- Brolund A, Sandegren L. Characterization of ESBL disseminating plasmids. *Infect Dis (Lond)*. 2016;48(1):18-25.
- Brown DR. Phylum XVI. Tenericutes. In: Krieg NR, Whitman WB, Bergey DH, éditeurs. *The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia,*

- Chlamydiae, and Planctomycetes. 2. ed. New York, NY: Springer; 2011. p. 567-724. (Bergey's manual of systematic bacteriology).
- Brown DR, Zacher LA, Farmerie WG. Spreading factors of *Mycoplasma alligatoris*, a flesh-eating mycoplasma. *J Bacteriol.* juin 2004;186(12):3922-7.
- Brown-Jaque M, Calero-Cáceres W, Muniesa M. Transfer of antibiotic-resistance genes via phage-related mobile elements. *Plasmid.* mai 2015;79:1-7.
- Bumgardner EA, Kittichotirat W, Bumgarner RE, Lawrence PK. Comparative genomic analysis of seven *Mycoplasma hyosynoviae* strains. *Microbiologyopen.* 18 févr 2015;
- Burdett V. Identification of tetracycline-resistant R-plasmids in *Streptococcus agalactiae* (group B). *Antimicrob Agents Chemother.* nov 1980;18(5):753-60.
- Burrus V. Mechanisms of stabilization of integrative and conjugative elements. *Curr Opin Microbiol.* août 2017;38:44-50.
- Burrus V, Pavlovic G, Decaris B, Guédon G. Conjugative transposons: the tip of the iceberg. *Mol Microbiol.* nov 2002;46(3):601-10.
- Calcutt MJ, Foeking MF. Complete Genome Sequence of *Mycoplasma hominis* Strain Sprott (ATCC 33131), Isolated from a Patient with Nongonococcal Urethritis. *Genome Announc.* 9 juill 2015;3(4):e00771-15.
- Calcutt MJ, Lavrrar JL, Wise KS. IS1630 of *Mycoplasma fermentans*, a novel IS30-type insertion element that targets and duplicates inverted repeats of variable length and sequence during insertion. *J Bacteriol.* déc 1999;181(24):7597-607.
- 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.* déc 2002;184(24):6929-41.
- Calderon-Copete SP, Wigger G, Wunderlin C, Schmidheini T, Frey J, Quail MA, *et al.* The *Mycoplasma conjunctivae* genome sequencing, annotation and analysis. *BMC Bioinformatics.* 16 juin 2009;10 Suppl 6:S7.
- Chazel M, Tardy F, Le Grand D, Calavas D, Poumarat F. Mycoplasmoses of ruminants in France: recent data from the national surveillance network. *BMC Veterinary Research.* 7 juin 2010;6(1):32.
- Chen I, Dubnau D. DNA uptake during bacterial transformation. *Nat Rev Microbiol.* mars 2004;2(3):241-9.
- Chopra I, Roberts M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews.* 1 juin 2001;65(2):232-60.
- Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. Toward automatic reconstruction of a highly resolved tree of life. *Science.* 3 mars 2006;311(5765):1283-7.

- Citti C. Les mycoplasmes : stratégie d'adaptation et de persistance de bactéries minimales. *Le Nouveau Praticien Vétérinaire élevages et santé*. 2006;(Août/Septembre/Octobre):203-9.
- Citti C, Blanchard A. Mycoplasmas and their host: emerging and re-emerging minimal pathogens. *Trends in Microbiology*. avr 2013;21(4):196-203.
- Citti C, Dordet-Frisoni E, Nouvel LX, Kuo CH, Baranowski E. Horizontal Gene Transfers in Mycoplasmas (Mollicutes). *Curr Issues Mol Biol*. 12 avr 2018;29:3-22.
- Citti C, Nouvel L-X, Baranowski E. Phase and antigenic variation in mycoplasmas. *Future Microbiology*. juill 2010;5(7):1073-85.
- Citti C, Watson-McKown R, Drosesse M, Wise KS. Gene Families Encoding Phase- and Size-Variable Surface Lipoproteins of *Mycoplasma hyorhinis*. *J Bacteriol*. mars 2000;182(5):1356-63.
- Clapper B, Tu A-HT, Elgavish A, Dybvig K. The vir gene of bacteriophage MAV1 confers resistance to phage infection on *Mycoplasma arthritidis*. *J Bacteriol*. sept 2004;186(17):5715-20.
- CLSI. Methods for Antimicrobial Susceptibility Testing for Human Mycoplasmas: Approved Guideline M43A. CLSI, Wayne, PA, USA. 2011.
- CLSI. Performance Standards for Antimicrobial Disk and Dilution susceptibility Tests for Bacteria Isolated from Animals; Approved Standard _ Fourth Edition. 2013.
- Connell SR, Trieber CA, Dinos GP, Einfeldt E, Taylor DE, Nierhaus KH. Mechanism of Tet(O)-mediated tetracycline resistance. *EMBO J*. 17 févr 2003;22(4):945-53.
- Correia S, Poeta P, Hébraud M, Capelo JL, Igrejas G. Mechanisms of quinolone action and resistance: where do we stand? *Journal of Medical Microbiology*. 2017;66(5):551-9.
- Couturier M, Bex F, Bergquist PL, Maas WK. Identification and classification of bacterial plasmids. *Microbiol Rev*. sept 1988;52(3):375-95.
- de la Cruz F, Frost LS, Meyer RJ, Zechner EL. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev*. janv 2010;34(1):18-40.
- Dégrange S, Renaudin H, Charron A, Bébéar C, Bébéar CM. Tetracycline resistance in *Ureaplasma* spp. and *Mycoplasma hominis*: prevalence in Bordeaux, France, from 1999 to 2002 and description of two tet(M)-positive isolates of *M. hominis* susceptible to tetracyclines. *Antimicrob Agents Chemother*. févr 2008;52(2):742-4.
- Djordjevic SR, Forbes WA, Forbes-Faulkner J, Kuhnert P, Hum S, Hornitzky MA, *et al*. Genetic diversity among *Mycoplasma* species bovine group 7: clonal isolates from an outbreak of polyarthritis, mastitis, and abortion in dairy cattle. *Electrophoresis*. oct 2001;22(16):3551-61.
- Dordet Frisoni E, Marena MS, Sagné E, Nouvel LX, Guépillot R, Glaser P, *et al*. ICEA of *Mycoplasma agalactiae* : a new family of self-transmissible integrative elements that confers conjugative properties to the recipient strain: *Mycoplasma* ICE on the move. *Molecular Microbiology*. sept 2013;89(6):1226-39.

- Dordet-Frisoni E, Baranowski E, Barré A, Blanchard A, Breton M, Couture C, *et al.* Draft Genome Sequences of *Mycoplasma auris* and *Mycoplasma yeatsii*, Two Species of the Ear Canal of Caprinae. *Genome Announc.* 13 juin 2013;1(3).
- Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX, Blanchard A, *et al.* Chromosomal Transfers in Mycoplasmas: When Minimal Genomes Go Mobile. *mBio.* 31 déc 2014;5(6):e01958-14.
- Dupuy V, Sirand-Pugnet P, Baranowski E, Barré A, Breton M, Couture C, *et al.* Complete Genome Sequence of *Mycoplasma putrefaciens* Strain 9231, One of the Agents of Contagious Agalactia in Goats. *Genome Announc.* 13 juin 2013;1(3).
- Edward DG, Freundt EA. The classification and nomenclature of organisms of the pleuropneumonia group. *J Gen Microbiol.* févr 1956;14(1):197-207.
- Edward DG, Freundt EA, Chanock RM, Fabricant J, Hayflick L, Lemcke RM, *et al.* Recommendations on nomenclature of the order Mycoplasmatales. *Science.* 31 mars 1967;155(3770):1694-6.
- Euzéby JP. List of Bacterial Names with Standing in Nomenclature: a Folder Available on the Internet. *International Journal of Systematic and Evolutionary Microbiology.* 1997;47(2):590-2.
- Fleury B, Bergonier D, Berthelot X, Peterhans E, Frey J, Vilei EM. Characterization of P40, a cytoadhesin of *Mycoplasma agalactiae*. *Infect Immun.* oct 2002;70(10):5612-21.
- Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, *et al.* The minimal gene complement of *Mycoplasma genitalium*. *Science.* 20 oct 1995;270(5235):397-403.
- Frey J. Mycoplasmas of animals. In: *Molecular biology and pathogenicity of Mycoplasmas.* Kluwer Academic / Plenum Publishers. New York, NY: Razin Shmuel and Richard Herrmann; 2002. p. 73-90.
- Fronzes R, Schäfer E, Wang L, Saibil HR, Orlova EV, Waksman G. Structure of a type IV secretion system core complex. *Science.* 9 janv 2009;323(5911):266-8.
- Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol.* sept 2005;3(9):722-32.
- Furneri PM, Rappazzo G, Musumarra MP, Tempera G, Roccasalva LS. Genetic basis of natural resistance to erythromycin in *Mycoplasma hominis*. *J Antimicrob Chemother.* 4 janv 2000;45(4):547-8.
- Furuya EY, Lowy FD. Antimicrobial-resistant bacteria in the community setting. *Nature Reviews Microbiology.* janv 2006;4(1):36-45.
- García-Aljaro C, Ballesté E, Muniesa M. Beyond the canonical strategies of horizontal gene transfer in prokaryotes. *Curr Opin Microbiol.* août 2017;38:95-105.
- Garcillán-Barcia MP, Francia MV, de la Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev.* mai 2009;33(3):657-87.

- Gaurivaud P, Baranowski E, Pau-Roblot C, Sagné E, Citti C, Tardy F. Mycoplasma agalactiae Secretion of β -(1 \rightarrow 6)-Glucan, a Rare Polysaccharide in Prokaryotes, Is Governed by High-Frequency Phase Variation. *Appl Environ Microbiol.* 1 juin 2016;82(11):3370-83.
- Gaurivaud P, Laigret F, Bove JM. Insusceptibility of members of the class Mollicutes to rifampin: studies of the *Spiroplasma citri* RNA polymerase beta-subunit gene. *Antimicrob Agents Chemother.* avr 1996;40(4):858-62.
- Gautier-Bouchardon AV. Antimicrobial Resistance in *Mycoplasma* spp. *Microbiol Spectr.* juill 2018;6(4).
- Gautier-Bouchardon AV, Ferré S, Le Grand D, Paoli A, Gay E, Poumarat F. Overall Decrease in the Susceptibility of *Mycoplasma bovis* to Antimicrobials over the Past 30 Years in France. *PLoS ONE.* 4 févr 2014;9(2):e87672.
- Gerchman I, Levisohn S, Mikula I, Manso-Silvan L, Lysnyansky I. Characterization of in vivo-acquired resistance to macrolides of *Mycoplasma gallisepticum* strains isolated from poultry. *Vet Res.* 2011;42(1):90.
- Gerrits MM, de Zoete MR, Arents NLA, Kuipers EJ, Kusters JG. 16S rRNA mutation-mediated tetracycline resistance in *Helicobacter pylori*. *Antimicrob Agents Chemother.* sept 2002;46(9):2996-3000.
- Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang R-Y, Algire MA, *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science.* 2 juill 2010;329(5987):52-6.
- Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, *et al.* Essential genes of a minimal bacterium. *PNAS.* 10 janv 2006;103(2):425-30.
- 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.* août 2000;68(8):4539-48.
- Goessweiner-Mohr N, Arends K, Keller W, Grohmann E. Conjugative type IV secretion systems in Gram-positive bacteria. *Plasmid.* nov 2013;70(3):289-302.
- Gomis-Rüth FX, de la Cruz F, Coll M. Structure and role of coupling proteins in conjugal DNA transfer. *Res Microbiol.* mai 2002;153(4):199-204.
- Gourlay RN, Wyld SG, Poulton ME. Some characteristics of mycoplasma virus Hr 1, isolated from and infecting *Mycoplasma hyorhinis*. Brief report. *Arch Virol.* 1983;77(1):81-5.
- Griffith F. The Significance of Pneumococcal Types. *J Hyg (Lond).* janv 1928;27(2):113-59.
- Guérillot R, Siguier P, Goubeyre E, Chandler M, Glaser P. The diversity of prokaryotic DDE transposases of the mutator superfamily, insertion specificity, and association with conjugation machineries. *Genome Biol Evol.* févr 2014;6(2):260-72.
- Guglielmini J, Quintais L, Garcillan-Barcia MP, de la Cruz F, Rocha EPC. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet.* août 2011;7(8):e1002222.

- Hannan PC. Guidelines and recommendations for antimicrobial minimum inhibitory concentration (MIC) testing against veterinary mycoplasma species. *Veterinary research*. 2000;31(4):373–395.
- Heuvelink A, Reugebrink C, Mars J. Antimicrobial susceptibility of *Mycoplasma bovis* isolates from veal calves and dairy cattle in the Netherlands. *Veterinary Microbiology*. 30 juin 2016;189:1-7.
- Hochhut B, Jahreis K, Lengeler JW, Schmid K. CTnscr94, a conjugative transposon found in enterobacteria. *J Bacteriol*. avr 1997;179(7):2097-102.
- Hochhut B, Lotfi Y, Mazel D, Faruque SM, Woodgate R, Waldor MK. Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. *Antimicrob Agents Chemother*. nov 2001;45(11):2991-3000.
- Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, *et al*. Understanding the mechanisms and drivers of antimicrobial resistance. *The Lancet*. 15 janv 2016;387(10014):176-87.
- Howard MT, Nelson WC, Matson SW. Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem*. 24 nov 1995;270(47):28381-6.
- Hutchison CA, Chuang R-Y, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, *et al*. Design and synthesis of a minimal bacterial genome. *Science*. 25 mars 2016;351(6280):aad6253.
- Jansson E, Backman A, Hakkarainen K, von Bonsdorff CH, Seniusová B, Miettinen A. Isolation and preliminary characterisation of mycoplasmavirus 20-P. *Med Biol*. avr 1982;60(2):116-20.
- Johansson KE, Pettersson B. Taxonomy of Mollicutes. In: *Molecular biology and pathogenicity of Mycoplasmas*. Kluwer Academic / Plenum Publishers. New York, NY: Razin Shmuel and Richard Herrmann; 2002. p. 1-29.
- Kent BN, Foecking MF, Calcutt MJ. Development of a novel plasmid as a shuttle vector for heterologous gene expression in *Mycoplasma yersiniae*. *J Microbiol Methods*. oct 2012;91(1):121-7.
- Khalil D, Becker CAM, Tardy F. Monitoring the Decrease in Susceptibility to Ribosomal RNAs Targeting Antimicrobials and Its Molecular Basis in Clinical *Mycoplasma bovis* Isolates over Time. *Microbial Drug Resistance*. 2 févr 2017;23(6):799-811.
- King KW, Dybvig K. Nucleotide sequence of *Mycoplasma mycoides* subspecies *Mycoides* plasmid pKMK1. *Plasmid*. juill 1992;28(1):86-91.
- Kolling GL, Matthews KR. Export of virulence genes and Shiga toxin by membrane vesicles of *Escherichia coli* O157:H7. *Appl Environ Microbiol*. mai 1999;65(5):1843-8.
- Koonin EV. Darwinian evolution in the light of genomics. *Nucleic Acids Res*. mars 2009;37(4):1011-34.
- Lang AS, Zhaxybayeva O, Beatty JT. Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev Microbiol*. 11 juin 2012;10(7):472-82.

- Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA, *et al.* Genome transplantation in bacteria: changing one species to another. *Science*. 3 août 2007;317(5838):632-8.
- Lartigue C, Vashee S, Algire MA, Chuang R-Y, Benders GA, Ma L, *et al.* Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science*. 25 sept 2009;325(5948):1693-6.
- Leclercq R, Courvalin P. Bacterial resistance to macrolide, lincosamide, and streptogramin antibiotics by target modification. *Antimicrob Agents Chemother*. juill 1991;35(7):1267-72.
- Leclercq R, Courvalin P. Resistance to Macrolides and Related Antibiotics in *Streptococcus pneumoniae*. *Antimicrob Agents Chemother*. sept 2002;46(9):2727-34.
- Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med*. déc 2004;10(12 Suppl):S122-129.
- Lewis K. Platforms for antibiotic discovery. *Nat Rev Drug Discov*. mai 2013;12(5):371-87.
- Li Y, Zheng H, Liu Y, Jiang Y, Xin J, Chen W, *et al.* The complete genome sequence of *Mycoplasma bovis* strain Hubei-1. *PLoS ONE*. 2011;6(6):e20999.
- Lilly J, Camps M. Mechanisms of Theta Plasmid Replication. *Microbiol Spectr*. févr 2015;3(1):PLAS-0029-2014.
- Liu W, Feng Z, Fang L, Zhou Z, Li Q, Li S, *et al.* Complete genome sequence of *Mycoplasma hyopneumoniae* strain 168. *J Bacteriol*. févr 2011;193(4):1016-7.
- Lloyd RG, Buckman C. Conjugational Recombination in *Escherichia Coli*: Genetic Analysis of Recombinant Formation in Hfr X F(-) Crosses. *Genetics*. mars 1995;139(3):1123-48.
- Lo W-S, Gasparich GE, Kuo C-H. Found and Lost: The Fates of Horizontally Acquired Genes in Arthropod-Symbiotic *Spiroplasma*. *Genome Biol Evol*. 8 août 2015;7(9):2458-72.
- Luo Y, Wasserfallen A. Gene transfer systems and their applications in Archaea. *Syst Appl Microbiol*. avr 2001;24(1):15-25.
- Lysnyansky I, Ayling RD. *Mycoplasma bovis*: Mechanisms of Resistance and Trends in Antimicrobial Susceptibility. *Front Microbiol*. 2016;7:595.
- Lysnyansky I, Calcutt MJ, Ben-Barak I, Ron Y, Levisohn S, Methé BA, *et al.* 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*. mai 2009;294(2):172-82.
- Lysnyansky I, Sachse K, Rosenbusch R, Levisohn S, Yogev D. The *vsp* Locus of *Mycoplasma bovis*: Gene Organization and Structural Features. *J Bacteriol*. sept 1999;181(18):5734-41.
- Mahillon J, Chandler M. Insertion Sequences. *Microbiol Mol Biol Rev*. 9 janv 1998;62(3):725-74.

- Manso-Silván L, Perrier X, Thiaucourt F. Phylogeny of the *Mycoplasma mycoides* cluster based on analysis of five conserved protein-coding sequences and possible implications for the taxonomy of the group. *Int J Syst Evol Microbiol.* oct 2007;57(Pt 10):2247-58.
- Manso-Silván L, Tardy F, Baranowski E, Barré A, Blanchard A, Breton M, *et al.* Draft Genome Sequences of *Mycoplasma alkalescens*, *Mycoplasma arginini*, and *Mycoplasma bovis*, Three Species with Equivocal Pathogenic Status for Cattle. *Genome Announc.* 13 juin 2013;1(3):e00348-13.
- 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. *Journal of Bacteriology.* 1 juin 2006;188(11):4137-41.
- Marenda MS. Genomic mosaics. In: *Mollicutes: Molecular biology and pathogenesis.* Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browning; 2014. p. 15-54.
- Marques LM, Guimarães AMS, Martins HB, Rezende IS, Barbosa MS, Campos GB, *et al.* Genome Sequence of *Ureaplasma diversum* Strain ATCC 49782. *Genome Announc.* 16 avr 2015;3(2).
- McClintock B. The Origin and Behavior of Mutable Loci in Maize. *Proc Natl Acad Sci U S A.* juin 1950;36(6):344-55.
- Medvedeva ES, Baranova NB, Mouzykantov AA, Grigoreva TY, Davydova MN, Chernova OA, *et al.* Extracellular vesicles of mycoplasmas and development of resistance to quinolones in bacteria. *Dokl Biochem Biophys.* janv 2014;454(1):34-7.
- Miller RV. Environmental bacteriophage-host interactions: factors contribution to natural transduction. *Antonie Van Leeuwenhoek.* juin 2001;79(2):141-7.
- 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.* nov 2004;186(21):7123-33.
- Neimark H, Johansson KE, Rikihisa Y, Tully JG. Proposal to transfer some members of the genera *Haemobartonella* and *Eperythrozoon* to the genus *Mycoplasma* with descriptions of « *Candidatus Mycoplasma haemofelis* », « *Candidatus Mycoplasma haemomuris* », « *Candidatus Mycoplasma haemosuis* » and « *Candidatus Mycoplasma wenyonii* ». *International Journal of Systematic and Evolutionary Microbiology.* 2001;51(3):891-9.
- Neimark H, Johansson K-E, Rikihisa Y, Tully JG. Revision of haemotrophic *Mycoplasma* species names. *International Journal of Systematic and Evolutionary Microbiology.* 2002;52(2):683-683.
- Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, *et al.* Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature.* 27 mai 1999;399(6734):323-9.
- Nesme J, Simonet P. The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. *Environmental Microbiology.* 1 avr 2015;17(4):913-30.

- Nocard E, Roux E. Le microbe de la Péripleumonie. *Annales de l'Institut Pasteur*. 1898;12:240-62.
- Noormohammadi AH, Markham PF, Duffy MF, Whithear KG, Browning GF. Multigene families encoding the major hemagglutinins in phylogenetically distinct mycoplasmas. *Infect Immun*. juill 1998;66(7):3470-5.
- Nouvel L-X, Marena M, Sirand-Pugnet P, Sagné E, Glew M, Mangenot S, *et al*. Occurrence, Plasticity, and Evolution of the vpma Gene Family, a Genetic System Devoted to High-Frequency Surface Variation in *Mycoplasma agalactiae*. *J Bacteriol*. juill 2009;191(13):4111-21.
- Nouvel LX, Sirand-Pugnet P, Marena MS, Sagné E, Barbe V, Mangenot S, *et al*. Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro-and micro-events that are shaping mycoplasma diversity. *BMC genomics*. 2010;11(1):86.
- Ochiai K, Yamanaka T, Kimura K, Sawada O. Inheritance of drug resistance (and its transfer) between *Shigella* strains and between *Shigella* and *E. coli* strains (in Japanese). *Hihon Iji Shimpou*. 1959;1861:34.
- Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. mai 2000;405(6784):299-304.
- Pereyre S, Gonzalez P, de Barbeyrac B, Darnige A, Renaudin H, Charron A, *et al*. Mutations in 23S rRNA Account for Intrinsic Resistance to Macrolides in *Mycoplasma hominis* and *Mycoplasma fermentans* and for Acquired Resistance to Macrolides in *M. hominis*. *Antimicrob Agents Chemother*. oct 2002;46(10):3142-50.
- Pereyre S, Goret J, Bébéar C. *Mycoplasma pneumoniae*: Current Knowledge on Macrolide Resistance and Treatment. *Front Microbiol*. 2016;7:974.
- Pereyre S, Renaudin H, Charron A, Bébéar C, Bébéar CM. Emergence of a 23S rRNA mutation in *Mycoplasma hominis* associated with a loss of the intrinsic resistance to erythromycin and azithromycin. *J Antimicrob Chemother*. avr 2006;57(4):753-6.
- Pereyre S, Sirand-Pugnet P, Beven L, Charron A, Renaudin H, Barré A, *et al*. Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital mycoplasmas. *PLoS Genet*. oct 2009;5(10):e1000677.
- Peterson SN, Fraser CM. The complexity of simplicity. *Genome Biol*. 2001;2(2):comment2002.1-comment2002.8.
- Pinto PM, Carvalho MO de, Alves-Junior L, Brocchi M, Schrank IS. Molecular analysis of an integrative conjugative element, ICEH, present in the chromosome of different strains of *Mycoplasma hyopneumoniae*. *Genetics and Molecular Biology*. 2007;30(1):256-63.
- Poole K. Efflux-mediated antimicrobial resistance. *J Antimicrob Chemother*. juill 2005;56(1):20-51.
- Pottumarthy S, Moland ES, Juretschko S, Swanzy SR, Thomson KS, Fritsche TR. NmCA Carbapenem-hydrolyzing Enzyme in *Enterobacter cloacae* in North America¹. *Emerg Infect Dis*. août 2003;9(8):999-1002.

- Poumarat F, Gautier-Bouchardon A, Bergonier D, Gay E, Tardy F. Diversity and variation of antimicrobial susceptibility patterns over time in *M. agalactiae* isolates collected from sheep and goats in France. *J Appl Microbiol*. 2 févr 2016;
- Qi J, Guo A, Cui P, Chen Y, Mustafa R, Ba X, *et al*. Comparative geno-plasticity analysis of *Mycoplasma bovis* HB0801 (Chinese isolate). *PLoS ONE*. 2012;7(5):e38239.
- Qin T-T, Kang H-Q, Ma P, Li P-P, Huang L-Y, Gu B. SOS response and its regulation on the fluoroquinolone resistance. *Ann Transl Med*. déc 2015;3(22):358.
- Raherison S, Gonzalez P, Renaudin H, Charron A, Bebear C, Bebear CM. Evidence of Active Efflux in Resistance to Ciprofloxacin and to Ethidium Bromide by *Mycoplasma hominis*. *Antimicrobial Agents and Chemotherapy*. 1 mars 2002;46(3):672-9.
- Raherison S, Gonzalez P, Renaudin H, Charron A, Bébéar C, Bébéar CM. Increased expression of two multidrug transporter-like genes is associated with ethidium bromide and ciprofloxacin resistance in *Mycoplasma hominis*. *Antimicrob Agents Chemother*. janv 2005;49(1):421-4.
- Ramirez MS, Tolmasky ME. Aminoglycoside modifying enzymes. *Drug Resist Updat*. déc 2010;13(6):151-71.
- Ravatt R, Studer S, Springael D, Zehnder AJ, van der Meer JR. Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. Strain B13. *J Bacteriol*. sept 1998;180(17):4360-9.
- Razin S, Herrmann R. *Molecular biology and pathogenicity of mycoplasmas*. New York, NY: Kluwer Academic/Plenum; 2002.
- Razin S, Yogev D, Naot Y. *Molecular Biology and Pathogenicity of Mycoplasmas*. *Microbiol Mol Biol Rev*. déc 1998;62(4):1094-156.
- Rechnitzer H, Brzuszkiewicz E, Strittmatter A, Liesegang H, Lysnyansky I, Daniel R, *et al*. Genomic features and insights into the biology of *Mycoplasma fermentans*. *Microbiology (Reading, Engl)*. mars 2011;157(Pt 3):760-73.
- Redgrave LS, Sutton SB, Webber MA, Piddock LJV. Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success. *Trends in Microbiology*. août 2014;22(8):438-45.
- Renaudin J, Breton M, Citti C. Molecular genetic tools for Mollicutes. In: *Mollicutes: molecular biology and pathogenesis*. Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browning; 2014. p. 55-76.
- Roberts MC, Kenny GE. Dissemination of the tetM tetracycline resistance determinant to *Ureaplasma urealyticum*. *Antimicrob Agents Chemother*. 2 janv 1986;29(2):350-2.
- Roberts MC, Koutsky LA, Holmes KK, LeBlanc DJ, Kenny GE. Tetracycline-resistant *Mycoplasma hominis* strains contain streptococcal tetM sequences. *Antimicrob Agents Chemother*. juill 1985;28(1):141-3.

- Robicsek A, Jacoby GA, Hooper DC. The worldwide emergence of plasmid-mediated quinolone resistance. *Lancet Infect Dis.* oct 2006;6(10):629-40.
- Rosengarten R, Wise KS. Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. *Science.* 19 janv 1990;247(4940):315-8.
- Röske K, Calcutt MJ, Wise KS. The *Mycoplasma fermentans* prophage phiMFV1: genome organization, mobility and variable expression of an encoded surface protein. *Mol Microbiol.* juin 2004;52(6):1703-20.
- Ruiz-Masó JA, Machón C, Bordanaba-Ruiseco L, Espinosa M, Coll M, Del Solar G. Plasmid Rolling-Circle Replication. *Microbiol Spectr.* févr 2015;3(1):PLAS-0035-2014.
- Rumpho ME, Worful JM, Lee J, Kannan K, Tyler MS, Bhattacharya D, *et al.* Horizontal gene transfer of the algal nuclear gene *psbO* to the photosynthetic sea slug *Elysia chlorotica*. *PNAS.* 18 nov 2008;105(46):17867-71.
- Saillard C, Carle P, Duret-Nurbel S, Henri R, Killiny N, Carrère S, *et al.* The abundant extrachromosomal DNA content of the *Spiroplasma citri* GII3-3X genome. *BMC Genomics.* 28 avr 2008;9:195.
- Schlünzen F, Zarivach R, Harms J, Bashan A, Tocilj A, Albrecht R, *et al.* Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature.* 25 oct 2001;413(6858):814-21.
- Shen X, Gumulak J, Yu H, French CT, Zou N, Dybvig K. Gene rearrangements in the *vsa* locus of *Mycoplasma pulmonis*. *J Bacteriol.* mai 2000;182(10):2900-8.
- Shu H-W, Liu T-T, Chan H-I, Liu Y-M, Wu K-M, Shu H-Y, *et al.* Genome sequence of the repetitive-sequence-rich *Mycoplasma fermentans* strain M64. *J Bacteriol.* août 2011;193(16):4302-3.
- Siguier P, Goubeyre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev.* sept 2014;38(5):865-91.
- Sirand-Pugnet P, Citti C, Barré A, Blanchard A. Evolution of mollicutes: down a bumpy road with twists and turns. *Research in Microbiology.* déc 2007a;158(10):754-66.
- Sirand-Pugnet P, Lartigue C, Marendia M, Jacob D, Barré A, Barbe V, *et al.* Being Pathogenic, Plastic, and Sexual while Living with a Nearly Minimal Bacterial Genome. *PLoS Genetics.* 2007b;3(5):e75.
- del Solar G, Espinosa M. Plasmid copy number control: an ever-growing story. *Mol Microbiol.* août 2000;37(3):492-500.
- Sullivan JT, Ronson CW. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a *phe*-tRNA gene. *Proc Natl Acad Sci U S A.* 28 avr 1998;95(9):5145-9.
- Tardy F, Baranowski E, Nouvel L-X, Mick V, Manso-Silvan L, Thiaucourt F, *et al.* Emergence of Atypical *Mycoplasma agalactiae* Strains Harboring a New Prophage and Associated with an Alpine Wild Ungulate Mortality Episode. *Applied and Environmental Microbiology.* 1 juill 2012;78(13):4659-68.

- Tardy F, Mick V, Dordet-Frisoni E, Marena MS, Sirand-Pugnet P, Blanchard A, *et al.* Integrative Conjugative Elements Are Widespread in Field Isolates of Mycoplasma Species Pathogenic for Ruminants. Elkins CA, éditeur. Applied and Environmental Microbiology. 1 mars 2015;81(5):1634-43.
- Teachman AM, French CT, Yu H, Simmons WL, Dybvig K. Gene Transfer in Mycoplasma pulmonis. J Bacteriol. févr 2002;184(4):947-51.
- Thiaucourt F, Manso-Silvan L, Salah W, Barbe V, Vacherie B, Jacob D, *et al.* Mycoplasma mycoides, from « mycoides Small Colony » to « capri ». A microevolutionary perspective. BMC Genomics. 16 févr 2011;12:114.
- 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. 15 avr 2005;245(2):249-55.
- Thomas CM, Nielsen KM. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nature Reviews Microbiology. sept 2005;3(9):711-21.
- Thongkamkoon P, Narongsak W, Kobayashi H, Pathanasophon P, Kishima M, Yamamoto K. In vitro susceptibility of Mycoplasma hyopneumoniae field isolates and occurrence of fluoroquinolone, macrolides and lincomycin resistance. J Vet Med Sci. 2013;75(8):1067-70.
- Threlfall EJ. Epidemic salmonella typhimurium DT 104--a truly international multiresistant clone. J Antimicrob Chemother. juill 2000;46(1):7-10.
- Torres-Puig S, Martínez-Torró C, Granero-Moya I, Querol E, Piñol J, Pich OQ. Activation of σ_{20} -dependent recombination and horizontal gene transfer in Mycoplasma genitalium. DNA Res. 1 août 2018;25(4):383-93.
- Tu AH, Voelker LL, Shen X, Dybvig K. Complete nucleotide sequence of the mycoplasma virus P1 genome. Plasmid. mars 2001;45(2):122-6.
- Vasconcelos ATR, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, *et al.* Swine and Poultry Pathogens: the Complete Genome Sequences of Two Strains of Mycoplasma hyopneumoniae and a Strain of Mycoplasma synoviae. J Bacteriol. août 2005;187(16):5568-77.
- Voelker LL, Dybvig K. Sequence analysis of the Mycoplasma arthritis bacteriophage MAV1 genome identifies the putative virulence factor. Gene. 11 juin 1999;233(1-2):101-7.
- Voelker LL, Weaver KE, Ehle LJ, Washburn LR. Association of lysogenic bacteriophage MAV1 with virulence of Mycoplasma arthritis. Infect Immun. oct 1995;63(10):4016-23.
- Waites KB, Duffy LB, Bébéar CM, Matlow A, Talkington DF, Kenny GE, *et al.* Standardized Methods and Quality Control Limits for Agar and Broth Microdilution Susceptibility Testing of Mycoplasma pneumoniae, Mycoplasma hominis, and Ureaplasma urealyticum. J Clin Microbiol. nov 2012;50(11):3542-7.

- Waites KB, Lysnyansky I, Bébéar C. Emerging antimicrobial resistance in mycoplasmas of humans and animals. In: Mollicutes: Molecular biology and pathogenesis. Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browing; 2014. p. 289-322.
- Waldor MK, Tschäpe H, Mekalanos JJ. A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. *J Bacteriol.* juill 1996;178(14):4157-65.
- Wang W, Gu W, Gasparich GE, Bi K, Ou J, Meng Q, *et al.* *Spiroplasma eriocheiris* sp. nov., associated with mortality in the Chinese mitten crab, *Eriocheir sinensis*. *Int J Syst Evol Microbiol.* avr 2011;61(Pt 4):703-8.
- Webber MA, Piddock LJV. The importance of efflux pumps in bacterial antibiotic resistance. *J Antimicrob Chemother.* 1 janv 2003;51(1):9-11.
- Weisburg WG, Tully JG, Rose DL, Petzel JP, Oyaizu H, Yang D, *et al.* A phylogenetic analysis of the mycoplasmas: basis for their classification. *J Bacteriol.* 12 janv 1989;171(12):6455-67.
- Wessler SR. Transposable elements and the evolution of eukaryotic genomes. *Proc Natl Acad Sci U S A.* 21 nov 2006;103(47):17600-1.
- Westberg J, Persson A, Holmberg A, Goesmann A, Lundeberg J, Johansson K-E, *et al.* The Genome Sequence of *Mycoplasma mycoides* subsp. *mycoides* SC Type Strain PG1T, the Causative Agent of Contagious Bovine Pleuropneumonia (CBPP). *Genome Res.* févr 2004;14(2):221-7.
- Willetts N, Wilkins B. Processing of plasmid DNA during bacterial conjugation. *Microbiol Rev.* mars 1984;48(1):24-41.
- Wilson DN. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology.* 16 déc 2013;12(1):35-48.
- von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, *et al.* Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol.* 2016;7:173.
- Wise KS, Calcutt MJ, Foecking MF, Roske K, Madupu R, Methe BA. Complete Genome Sequence of *Mycoplasma bovis* Type Strain PG45 (ATCC 25523). *Infection and Immunity.* 1 févr 2011;79(2):982-3.
- Woese CR, Maniloff J, Zablén LB. Phylogenetic analysis of the mycoplasmas. *Proc Natl Acad Sci USA.* janv 1980;77(1):494-8.
- Wozniak RAF, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Micro.* août 2010;8(8):552-63.
- Yamao F, Muto A, Kawauchi Y, Iwami M, Iwagami S, Azumi Y, *et al.* UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc Natl Acad Sci USA.* avr 1985;82(8):2306-9.
- Yoshida H, Bogaki M, Nakamura M, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy.* 1990;34(6):1271-1272.

Yoshida H, Bogaki M, Nakamura M, Yamanaka LM, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrB* gene of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 1991;35(8):1647–1650.

Zhaxybayeva O. Detection and quantitative assessment of horizontal gene transfer. *Methods Mol Biol*. 2009;532:195-213.

Zheng J, McIntosh MA. Characterization of IS1221 from *Mycoplasma hyorhinis*: expression of its putative transposase in *Escherichia coli* incorporates a ribosomal frameshift mechanism. *Mol Microbiol*. 1995;16(4):669-85.

CHAPITRE 1

CHAPITRE 1

Mycoplasmas under experimental antimicrobial selection: the unpredicted contribution of horizontal chromosomal transfer in accelerating short-term adaptation.

Les mycoplasmes sous sélection antimicrobienne expérimentale: contribution inédite du transfert chromosomique horizontal dans l'accélération de l'adaptation à court terme.

Marion Faucher^{1,2}, Laurent-Xavier Nouvel^{1*}, Emilie Dordet-Frisoni¹, Eveline Sagné¹, Eric Baranowski¹, Marie-Claude Hygonenq¹, Marc-Serge Marena³, Florence Tardy², Christine Citti^{1*}

¹IHAP, Université de Toulouse, INRA, ENVT, Toulouse, France.

²UMR Mycoplasmoses of ruminants, ANSES, VetAgro Sup, University of Lyon, Lyon, France.

³Asia-Pacific Centre for Animal Health, University of Melbourne, Melbourne, Victoria, Australia.

*Corresponding author

E-mail: c.citti@envt.fr (CC), x.nouvel@envt.fr (LXN)

Article soumis à Plos Genetics le 13 août 2018.

Résumé

Les mycoplasmes, large groupe de bactéries sans paroi, sont souvent décrits comme des bactéries minimales en raison de leur génome réduit et de leurs voies métaboliques limitées. Une évolution régressive par perte massive de gènes a longtemps été considérée comme l'unique moteur de leur évolution. Cette vision a récemment été bouleversée par la mise en évidence de transferts horizontaux de gène massifs et la découverte d'un mécanisme de conjugaison non-conventionnel qui permet des échanges de grands fragments chromosomiques. Dans cette étude, des expériences de conjugaison et d'évolution expérimentale combinées à du séquençage de nouvelle génération explorent la contribution et l'impact de ce phénomène sur l'évolution et l'adaptation des mycoplasmes. Le pathogène de ruminants *Mycoplasma agalactiae* a été utilisé comme organisme modèle et une fluoroquinolone, l'enrofloxacin (Enro) comme pression de sélection. Des lignées isogéniques avec différentes combinaisons de mutations spontanées dans les gènes cibles de l'Enro (*gyrA*, *gyrB*, *parC* et *parE*) associées à des niveaux croissant de résistance à l'Enro ont été générées. Chez *M. agalactiae*, les 4 gènes cibles sont situés à 3 loci distincts et distants, l'acquisition de ces mutations par une population sensible via un transfert chromosomique conjugatif a été testée. Nos résultats montrent que le temps nécessaire à l'acquisition de mutations qui confèrent un haut niveau de résistance à l'Enro peut être rapidement compressé, en une seule étape de conjugaison dans laquelle plusieurs allèles Enro^R sont transférés d'un mycoplasme résistant vers un sensible. En plus d'accélérer la dissémination de la résistance, les transferts chromosomiques observés remanient les génomes et engendrent des mosaïques. Ce mosaïcisme peut conférer aux sous populations résistantes des caractéristiques imprévisibles sans rapport direct avec la sélection antibiotique. Ainsi, nos résultats donnent une vision augmentée des processus évolutifs et adaptatifs qui façonnent les petits génomes de mycoplasmes au travers d'un mécanisme de conjugaison non-conventionnel.

Abstract

Horizontal Gene Transfer was long thought to be marginal in *Mycoplasma* a large group of wall-less bacteria often portrayed as minimal cells because of their reduced genomes (ca. 0.5 to 2.0 Mb) and their limited metabolic pathways. This view was recently challenged by the discovery of conjugative exchanges of large chromosomal fragments that equally affected all parts of the chromosome via an unconventional mechanism, so that the whole mycoplasma genome is potentially mobile. By combining next generation sequencing to classical mating and evolutionary experiments, the current study further explored the contribution and impact of this phenomenon on mycoplasma evolution and adaptation using the fluoroquinolone enrofloxacin (Enro), for selective pressure and the ruminant pathogen *M. agalactiae*, as a model organism. For this purpose, we generated isogenic lineages that displayed different combination of spontaneous mutations in Enro target genes (*gyrA*, *gyrB*, *parC* and *parE*) in association to gradual level of resistance to Enro. We then tested whether these mutations can be acquired by a susceptible population via conjugative chromosomal transfer knowing that in our model organism, the 4 target genes are scattered in three distinct and distant loci. Our data shows that under antibiotic selective pressure, the time scale of the mutational pathway leading to high-level of Enro resistance can be readily compressed into a single conjugative step, in which several Enro^R alleles were transferred from resistant to susceptible mycoplasma cells. In addition to acting as an accelerator for antimicrobial dissemination, mycoplasma chromosomal transfer reshuffled genomes beyond expectations and created a mosaic of resistant sub-populations with unpredicted and unrelated features. Our findings provide insights into the evolutionary processes that shaped the small mycoplasma genome and drive adaptability via an unconventional conjugative mechanism.

Author summary

Genome downsizing is often viewed as degenerative process of evolution. Such erosion has left current mycoplasmas with a minimal genome: for some species its size barely exceeds the amount of information needed for sustaining autonomous life. Despite such limitations, these simple bacteria showcase a baffling capacity for adaptation to complex environments such as that provided by the animal host. By using the enrofloxacin antibiotic as selective pressure, we performed a genome scale analysis of macro- and micro-events leading to antimicrobial resistance in mycoplasmas. Sexually competent cells were found to shortcut this process by using an unconventional mechanism of chromosomal transfer driving massive exchanges of DNA materials. Remarkably, this powerful mechanism was associated with a profound genomic reorganization that reshuffled genomes beyond expectations and created mosaicism. This finding emphasizes the extraordinary adaptability of mycoplasma and provides major insights into the evolutionary processes that shaped their small genome. While unconventional conjugative mechanisms are being documented in more complex bacteria, the reduced mycoplasma genome provides a simple model to study mosaicism and its role in bacterial evolution.

1. Introduction

The richness and diversity of bacterial taxa is fascinating, yet the understanding of the mechanisms underlying bacterial evolution and adaptation, at both population- and single-cell levels, remains unexplored in most species. To reduce this gap, tremendous efforts have been invested over the last decade in sequencing and decrypting bacterial genomes. One main outcome is that our traditional view of bacterial clonality and species boundaries is currently being challenged by the many facets of horizontal gene transfer (HGT), a key player of microbial diversification (Ochman et al. 2000; Wiedenbeck and Cohan 2011). In this phenomenon, the role of mobile genetic elements (MGE) is central (Burrus et al. 2002; Wozniak and Waldor 2010; Dordet Frisoni et al. 2013; Baranowski et al. 2018) and an increasing number of reports suggests that these might only represent the tip of the iceberg. Indeed, the conjugative transfer of large chromosomal fragments across genomes and their subsequent recombination might be more prominent and complex than first envisaged, with several emerging mechanisms that differ from the canonical Hfr- (or *oriT*-based) transfer initially described in *Escherichia coli* (Boritsch et al. 2016; Dordet-Frisoni et al. 2014; Gray et al. 2013; Blesa et al. 2017; Husain et al. 2017).

HGT was long thought to be marginal in *Mycoplasma* (class *Mollicutes*), a large group of wall-less bacteria often portrayed as minimal cells because of their reduced genomes (ca. 0.5 to 2.0 Mb) and limited metabolic pathways (Razin et al. 1998; Citti and Blanchard 2013). Despite this simplicity, several mycoplasma species are important pathogens of human and a wide range of animals (Razin et al. 1998; Rosengarten et al. 2001). Their current prevalence and impact on health highlights our failure in providing efficient preventive and therapeutic strategies, as well as the mycoplasma astonishing capacity to face the challenging environment of the host, to escape the immune response as well as to generate antimicrobial resistance (Citti et al. 2010; Citti and Blanchard 2013; Gautier-Bouchardon 2018). Mycoplasmas live in close contact with their host on which they rely for nutrients. Their adaptability is most likely the result of a co-evolution process: many loci generating high-frequency variation of the mycoplasma surface architecture have been selected in species infecting man and animals (Citti et al. 2005, 2010). These are controlled by localized, stochastic, error-prone or specific-recombination systems and encode for a broad range of molecules that are key in host-interactions, including sugar moieties (Citti et al. 2010; Gaurivaud et al. 2016). Thanks to the advances in genomics, *in silico* analyses have detected the occurrence of massive HGT across phylogenetically distant species that may account for unpredicted genome plasticity (Vasconcelos et al. 2005; Sirand-Pugnet et

al. 2007a). This hypothesis was further supported by experimental data showing the conjugative exchange of large chromosomal fragments among and within strains of *Mycoplasma agalactiae*, a ruminant pathogen and a model organism (Dordet-Frisoni et al. 2014). Congruent *in silico* and *in vitro* data further showed that these transfers equally affected all part of the chromosome via an unconventional mechanism, so that the whole mycoplasma genome is potentially mobile. While this has been formally demonstrated for *M. agalactiae* and *M. bovis*, growing evidences indicate that HGT is occurring in other species such as in *M. pulmonis* (Teachman et al. 2002), *M. genitalium* (Torres-Puig et al. 2018) and in other genera of the class *Mollicutes* such as in ureaplasmas or spiroplasmas (Pereyre et al. 2009; Citti et al. 2018; Xiao et al. 2011; Lo et al. 2015). HGT may have tremendous impact on the long and short-term evolution and adaptability of these organisms but has yet to be explored. Using an antibiotic selective pressure would offer a powerful way for addressing this issue *in vitro*; at the same time understanding the emergence of antibiotic resistance in pathogenic mycoplasmas is of primary importance for public health (Waites et al. 2014; Gautier-Bouchardon 2018).

The horizontal dissemination of MGE carrying antibiotic resistance genes, within and across bacterial species has been largely incriminated in the antibiotic crisis (Levy and Marshall 2004; von Wintersdorff et al. 2016). In mycoplasmas, the acquisition of resistance via HGT has long been ignored mainly because of (i) the lack of known conjugative plasmids as well as the paucity in other MGEs and in recombination systems and (ii) the scarcity of appropriate genetic tools combined to a fastidious growth under laboratory conditions, when compared to other more classical bacteria. The main genetic pathway described so far for the emergence of antibiotic resistance in these organisms is the occurrence, selection and fixation of chromosomal mutations in target genes (Waites et al. 2014; Gautier-Bouchardon 2018). Fluoroquinolones represents one of the most important classes of antibiotics that are effective on the wall-less mycoplasma cell. This antimicrobial inhibits the positive supercoils relaxing activity of the DNA gyrase and topoisomerase IV, as well as DNA decatenation following DNA replication and thus inhibiting DNA synthesis and leading to cell death (Redgrave et al. 2014; Correia et al. 2017). Mutations conferring quinolone resistance have been extensively reported in several taxa (Hopkins et al. 2005; Hooper and Jacoby 2015), including pathogenic mycoplasma species, and usually occur in 4 genes, *gyrA*, *gyrB*, *parC* and *parE*. These allow a structural change in the enzyme that limits antibiotic fixation, with the QRDR (Quinolone Resistance Determining Region) of GyrA and ParC being most often affected at key positions

(amino acids 83 for GyrA and 80, 84 for ParC according to *E. coli* numbering) (Kenny et al. 1999; Reinhardt et al. 2002; Khalil et al. 2016).

In this study, we experimentally explored the impact of conjugative chromosomal transfer on mycoplasma evolution and adaptation using a fluoroquinolone, the enrofloxacin (Enro), as selective pressure agent and *M. agalactiae*, as a model organism and an important pathogen of ruminants. For this purpose, we first generated spontaneous isogenic mutants displaying different combination of mutations in *gyrA*, *gyrB*, *parC* and *parE* together with various level of resistance to Enro. We then tested whether these mutations can be acquired by a susceptible population via conjugative chromosomal transfer knowing that in several mycoplasma species, including our model organism, the 4 target genes are located in distinct chromosomal loci; in *M. agalactiae* these are separated by at least 250 kb, with *parE* and *parC* being part of a same operon (Figs 1 and 2). Under antibiotic selective pressure, spontaneous mutants emerged stepwise following a similar pathway with HGT acting as an evolutionary accelerator that reshuffled genomes beyond expectations and created a mosaic of resistant sub-populations with unpredicted and unrelated features. Our findings provide insights into the evolutionary processes that shaped the small mycoplasma genome and drive adaptability via an unconventional conjugative mechanism.

2. Results

Selection of spontaneous mycoplasma mutants highly-resistant to enrofloxacin

Prior to testing the impact of Mycoplasma Chromosomal Transfer (MCT) in antimicrobial-resistance acquisition (see below), we analysed the evolutionary pathway leading to high resistance to enrofloxacin (Enro) in a set of spontaneous mutants derived from the strain PG2 55-5. For this purpose, six lineages namely MF26 and MF29 to MF33 (Table 1), were generated by rounds of single-colony bottleneck selection on solid medium containing stepwise concentrations of Enro (0.5 to 32 $\mu\text{g.mL}^{-1}$, with a two-fold step interval) (Fig 1A). At each step, single colonies were picked and analysed as described in Materials and Methods. Overall, we selected 22 individual isogenic clones with increased MIC (Minimal Inhibitory Concentration) values ranging from 1 to 64 $\mu\text{g.mL}^{-1}$, further designated as Enro^R clones, and sequenced their *parE*-, *parC*-, *gyrA*- and *gyrB*-QRDR (Quinolone Resistance Determining Region) (Yoshida et al. 1990, 1991) directly from the chromosome.

Sequence data revealed the occurrence of 1 to 3 single-point mutations in each of the 22 Enro^R clones (Table 1), with a total of 11 Single Nucleotide Polymorphisms (SNPs) detected in distinct positions, none being silent. All clones with increased MICs had at least one mutation within *parC* that was always a transition, C>T or G>A, and resulted in amino-acid changes, corresponding to codon 80 and/or 84 of the QRDR (further designated *parC*₈₀ and *parC*₈₄, respectively). In 68% of the mutants, an additional point mutation was present in *gyrA* that also resulted in codon change in its QRDR, most often at codon 83 and in some cases, at codon 87 or 81 (further designated *gyrA*₈₃, *gyrA*₈₇ and *gyrA*₈₁, respectively). Finally, additional mutations were occasionally found in *parE* and *gyrB*, affecting the canonical QRDR only in mutants MF29-1-3-6 and MF31-1, in *parE* codon 420 (*parE*₄₂₀) and *gyrB* codon 426 (*gyrB*₄₂₆), respectively (Table 1).

Within each lineage, the MIC value was shown to increase over the stepwise selection process together with the number of accumulated mutations. This is illustrated in Table 1 with MF26 which successively acquired a new mutation at each round of selection, in *parC*, *gyrA* and then *parE*, concomitantly to an increase of MIC from 1 to 32 µg.mL⁻¹. This pattern was observed for all lineages, with a few cases of one-step increase in MIC that were not linked to an additional mutation in the sequenced regions (Table 1, MF30-1>MF30-1-4 and MF29-1>MF29-1-3).

Overall, Enro^R mutations accumulate following a common pathway, emerging first in *parC*, then in *gyrA* (or for MF31 in *gyrB*) and last in *parE*. Thus the contribution of mutations in *parE* did not seem to be as critical towards resistance as those occurring in *parC* and *gyrA*, some mutants having high MIC value and no *parE* mutation (see for instance MF30-1-4-1).

Passaging of PG2 55-5 in broth medium only, without intermediate rounds of sub-cloning onto solid medium (Fig 1A), in presence of increasing concentration of Enro resulted in a selected PG2^{E10} population having a MIC of 64 µg.mL⁻¹ and 3 SNPs located in *parC*₈₀, *parC*₈₄ and *gyrA*₈₃, as in MF30-1-4-1, MF30-1-4-8 and MF29-1-3-1 (Table 1).

Table 1: Increase in the level of resistance to enrofloxacin in *M. agalactiae* mutants correlate with the accumulation of amino acid substitutions in QRDR of *parE*, *parC*, *gyrA* and *gyrB*.

Name ^(a)	MIC Enro ^(b)	Topoisomerase IV ^(d)				DNA gyrase ^(d)					Enro selection ^(b)		
		ParE				ParC		GyrA				GyrB	
		<u>420</u>	<u>422</u>	<u>426</u>	<u>439</u>	<u>80</u>	<u>84</u>	<u>81</u>	<u>83</u>	<u>87</u>	<u>426</u>	<u>469</u>	
		D (GAT)	A (GCT)	A (GCT)	P (CCG)	T (ACT)	D (GAT)	G (GGT)	S (TCA)	E (GAA)	N (AAC)	S (TCG)	
PG2 55-5 ^(c)	0.125												
MF32	1	-	-	-	-	I (ATT)	-	-	-	-	-	L (TTG)	0.5
MF32-1	8	-	-	-	-	I (ATT)	-	-	L (TTA)	-	-	L (TTG)	2
MF26 ●	1	-	-	-	-	I (ATT)	-	-	-	-	-	-	0.5
MF26-1 ●	16	-	-	-	-	I (ATT)	-	-	L (TTA)	-	-	-	4
MF26-1-2 ●	32	-	-	V (GTT)	-	I (ATT)	-	-	L (TTA)	-	-	-	32
MF33 ●	1	-	-	-	-	I (ATT)	-	-	-	-	-	-	0.5
MF33-1 ●	16	-	-	-	-	I (ATT)	-	-	-	K (AAA)	-	-	8
MF33-1-1 ●	32	-	-	-	S (TCG)	I (ATT)	-	-	-	K (AAA)	-	-	32
MF33-1-2 ●	64	-	-	-	S (TCG)	I (ATT)	-	-	-	K (AAA)	-	-	32
MF33-1-3 ●	64	-	S (TCT)	-	-	I (ATT)	-	-	-	K (AAA)	-	-	32
MF30 ●	2	-	-	-	-	-	N (AAT)	-	-	-	-	-	0.5
MF30-1 ●	4	-	-	-	-	-	N (AAT)	-	L (TTA)	-	-	-	2
MF30-1-4 ●	16	-	-	-	-	-	N (AAT)	-	L (TTA)	-	-	-	8
MF30-1-4-1 ●	64	-	-	-	-	I (ATT)	N (AAT)	-	L (TTA)	-	-	-	32
MF30-1-4-8 ●	64	-	-	-	-	I (ATT)	N (AAT)	-	L (TTA)	-	-	-	32
MF29	1	-	-	-	-	-	N (AAT)	-	-	-	-	-	0.5
MF29-1	4	-	-	-	-	-	N (AAT)	D (GAT)	-	-	-	-	2
MF29-1-3	16	-	-	-	-	-	N (AAT)	D (GAT)	-	-	-	-	8
MF29-1-3-1	64	-	-	-	-	I (ATT)	N (AAT)	-	L (TTA)	-	-	-	32
MF29-1-3-6	64	N (AAT)	-	-	-	-	N (AAT)	-	L (TTA)	-	-	-	32
MF31	1	-	-	-	-	-	N (AAT)	-	-	-	-	-	0.5
MF31-1	2	-	-	-	-	-	N (AAT)	-	-	-	K (AAA)	-	1
PG2 ^{E10} ●	64	-	-	-	-	I (ATT)	N (AAT)	-	L (TTA)	-	-	-	10

^(a) Black arrows on the left side indicate mutant's relationships within each isogenic lineage. Bullets designated the 14 clones further subjected to whole-genome sequencing.

^(b) Enrofloxacin concentration, in $\mu\text{g}\cdot\text{mL}^{-1}$.

^(c) *M. agalactiae* parental clone PG2 55-5 (NC_009497.1).

^(d) Mutations and consequent amino acid changes in coding regions encompassing QRDR as defined in *E. coli* numbering (underlined amino acids are in QRDR). Amino acid change and the corresponding codon are indicated, with mutated nucleotides in red. -: Identical to the parental PG2 55-5.

Mutations conferring increasing resistance to enrofloxacin preferentially accumulate in type II topoisomerase genes

Some discrepancies between the number of mutations and the MIC values (see above) raised the question of whether the level of resistance may be modulated by mutations occurring outside the QRDR regions. To address this issue, the genome of 13 clones belonging to 3 independent and representative lineages (MF26, MF33, and MF30) was fully sequenced by Illumina with a mean coverage of 3100X. SNPs and indels were identified by variant calling analyses using the PG2 55-5 parent clone as reference (Fig 1B). In several mutants (MF33, MF30, MF33-1, MF30-1-4, MF30-1-4-1, and MF30-1-4-8), additional SNPs were detected in the

3 target genes (*parE*₈₆, *parE*₁₁₂, *parC*₂₉₁, *parC*₅₇₄ *gyrB*₂₉ and *gyrB*₂₇₈), outside of the QRDR and of the regions previously sequenced (S2 Table). Surprisingly, none of these additional mutations was further selected upon passages under increasing antibiotic selection pressure. A few mutations (SNPs and indels) were also detected outside of the classical quinolone target genes (Fig 1B and S2 Table), with most occurring in homopolymeric tracts that are known as being prone to high frequency insertion-deletion (Citti et al. 2010). For instance, nt-707306 and/or nt-711627 both underwent a C deletion within a polyC of the so-called *spma* locus which encode phase variable membrane proteins (Nouvel et al. 2010). As well, the MF33-1-1 and -1-2 siblings both contained sub-populations displaying a large number of SNPs (14 and 11) within the highly variable *vpma* locus (Glew et al. 2000). Overall, sequenced genomes contained from 3 to 23 mutations, with a mean of 10.2 ± 5.8 mutations, with approximately half being fixed in the population (present in $\geq 95\%$ of the reads).

In parallel, the genome of the Enro^R PG2^{E10} was analysed and a total of 8 mutations were detected, 3 fixed and 5 unfixed (present in $< 95\%$ of the reads), when compared to the parental strain (Fig 1B and S2 Table). As expected, 3 SNPs were found in the Enro target genes, *parC*₈₀, *parC*₈₄, *gyrA*₈₃, and an additional one was detected in *parE*₁₁₂, outside of the QRDR initially sequenced by Sanger (see above). Interestingly, amongst the 3 studied lineages, this 4-SNPs' combination was only found in MF30-1-4-1 and MF30-1-4-8 which MIC of $64 \mu\text{g.mL}^{-1}$ is identical to that of PG2^{E10}. The other 4 mutations occurred outside the target genes and correspond to either highly variable loci (see above) or to mutations occurring only in minor subpopulations (unfixed). Based on competition fitness assays, these mutations did not appear to impose a cost on PG2^{E10} which has a fitness of 1.03 ± 0.10 when compared to PG2 55-5 parent (see Materials and Methods).

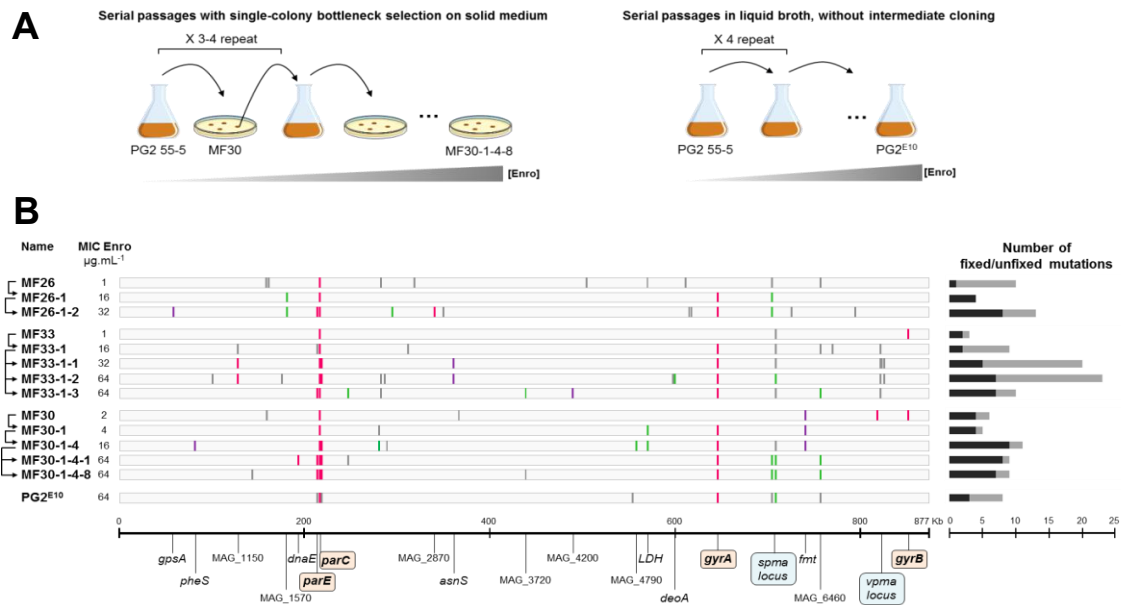


Fig 1. *In vitro* selection and genome analysis of spontaneous *M. agalactiae* Enro^R mutants. (A) Experimental selection of Enro^R *M. agalactiae* mutants by serial propagation of PG2 55-5 in broth with increasing concentration of enrofloxacin, with (left) or without (right) intermediate cloning on solid media. For bottleneck selection on solid media, a single colony was picked at each step to generate lineages of Enro^R mutants. For selection of PG2^{E10}, 1% of culture was transferred at each step in fresh media containing the antimicrobial. (B) Nature and localisation of mutations detected in genomes of Enro^R mutants selected in panel A. Genomes are represented by horizontal grey rectangles starting on left by the origin of replication. Mutations (SNPs and indels) are indicated by vertical lines in comparison to the reference PG2 55-5 genome (NC_009497.1). Black arrows on the left side indicate the parental relationships within each isogenic lineage, with the enrofloxacin MIC indicated in $\mu\text{g.mL}^{-1}$. Fixed mutations ($\geq 95\%$ of reads) are represented by coloured lines (synonymous SNP in violet, non-synonymous SNP in pink and indel in green) and grey lines represent unfixed mutations ($< 95\%$ of reads). Genes impacted by fixed mutations are indicated below, by their Genbank locus tag, or gene name when defined. Enrofloxacin target genes and variable loci are underlined in orange and blue, respectively. The total number of mutations for each clone (fixed mutations in black, unfixed in grey) is depicted by a diagram on the right side.

Horizontal chromosomal transfer of multiple, distant Enro^R-mutations accelerates resistance dissemination

Our hypothesis is that horizontal conjugative chromosomal transfer may act as a driving force of mycoplasma short-term evolution. We tested this hypothesis by addressing whether multiple, distant chromosomal Enro^R point-mutations can be simultaneously transferred from a resistant to a susceptible strain by HGT and selected in presence of the antimicrobial. For this purpose, two independent mating experiments (T5 and T6) were performed using (i) as donor the Enro^R PG2^{E10} mutant (see above) which displayed the highest MIC ($64 \mu\text{g.mL}^{-1}$) but the smallest number of fixed mutations (see Fig 1B and S2 Table) and (ii) as recipient, the 5632^{G3} clone previously derived from the Enro^S 5632 strain (MIC= $0.125 \mu\text{g.mL}^{-1}$) in which the gentamicin-resistance marker (Gm) is stably inserted as a proxy (Dordet-Frisoni et al. 2014) (see Materials and Methods) (Figs 2A and S1). Transconjugants were selected on solid media

containing 50 $\mu\text{g.mL}^{-1}$ of gentamicin in addition to Enro at concentrations ranging from 0.25 to 8 $\mu\text{g.mL}^{-1}$ (equal to 2 to 64 fold the MIC 5632^{G3}). Repeated attempts consistently yielded transconjugants colonies on solid medium containing 0.25 $\mu\text{g.mL}^{-1}$ of Enro (except for T5-5 obtained at 0.5 $\mu\text{g.mL}^{-1}$) with a low frequency ranging from $2.7.10^{-11}$ to $7.2.10^{-8}$ transconjugants per donor-CFU, depending on the 5632^{G3}:PG2^{E10} initial ratio (1:10 or 10:1, respectively, see Materials and Methods).

A total of 18 individual transconjugants were then picked and subjected to a series of PCR assays. These targeted the Gm marker and 11 distant loci that are distributed around the genome and discriminate 5632 from PG2 (S1 Fig). Of these, 8 were previously described (Dordet-Frisoni et al. 2014) and 3 were specifically designed in this study to distinguish PG2-*parC*, -*gyrA* and -*gyrB* from their 5632 counterparts (S1 Table). PCR data indicated that the transconjugant genotypes were a composite of PG2 and 5632 genomes (S1 Fig) except for T6-7 which was further shown by sequencing to have a chimeric *gyrA* (see below). They further designated 5632^{G3} as the recipient chromosome, with a majority of the PCR products being 5632-specific and the Gm marker constantly detected at the same position, as in the parent. This finding was in agreement with our previous findings showing that chromosomal transfer always occurred from PG2 (donor) to 5632 (recipient) (Dordet-Frisoni et al. 2014). Overall, 10 distinct PCR profiles were observed, with several transconjugants sharing identical profiles: T5-2 and T6-2; T6-4, -8, -9, -11, -12, -13; T5-4 and T6-1; T6-3 and T6-6 (S1 Fig).

Whole Genome Sequencing (WGS) by Illumina was performed with a subset of 13 transconjugants that were selected (i) to represent each of the 10 PCR profiles identified above and (ii) to include transconjugants with identical PCR profiles from independent (T5-4 and T6-1) or the same (T6-4, -8, -9) mating experiments. Sequence data confirmed that all displayed the 5632^{G3} chromosome as genetic background designating the corresponding strain as the recipient (Fig 2A). Further analyses demonstrated the systematic transfer of PG2^{E10} donor remote Enro target-genes containing (i) the mutated *parE/parC* operon (10/13 transconjugants) together with either the mutated-*gyrA* or the wild-type *gyrB* (wt) or (ii) the mutated *gyrA* alone (3/13) (Fig 2A). A close-up image of these regions is depicted in Fig. 3A and shows that two fixed mutations, corresponding to *parC*₈₀ and/or *gyrA*₈₃, were always associated to the transfer. Of note, the chimeric structure of T6-7 *gyrA* explains the PCR result obtained above (S1 Fig).

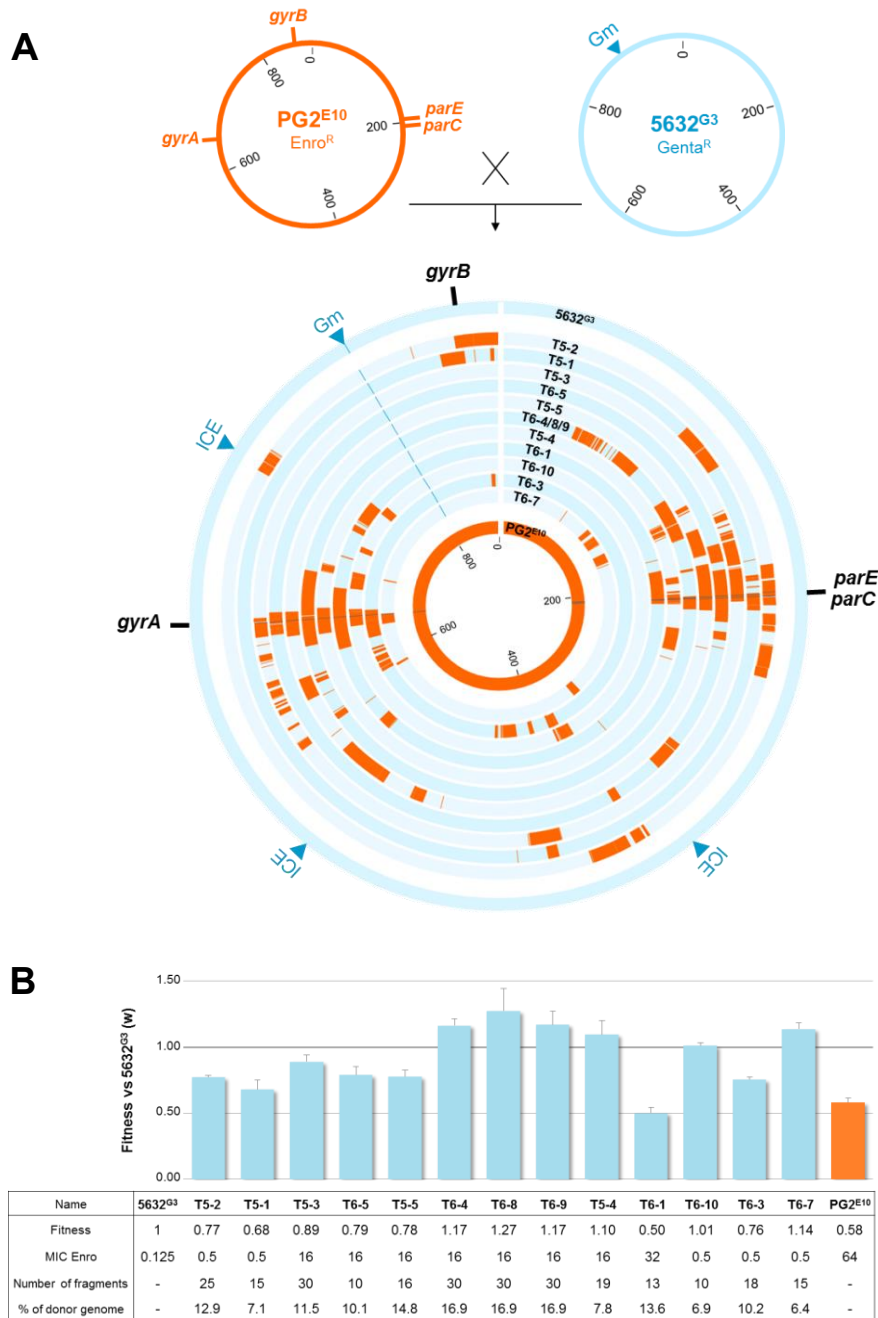


Fig 2. Mycoplasma transconjugant genomes and relative fitness.

(A) Mating experiments were conducted using the Enro^R PG2^{E10} as donor and the Enro^S, gentamicin-tagged 5632^{G3} as recipient. Transconjugants were selected on solid media containing enrofloxacin and gentamicin (see Materials and Methods), their mosaic genomes are represented as concentric circles flanked by parental genomes. Sequences specific of PG2^{E10} and 5632^{G3} are colour-coded in orange and blue, respectively. The gentamicin resistance marker (Gm) and genes targeted by Enro (*parE*, *parC*, *gyrA* and *gyrB*) are shown on the periphery, with vertical black lines indicating the position of mutations relative to PG2 55-5. ICE relative positions in 5632 genome are indicated by blue arrow. From inner to outer circle: PG2^{E10}, T6-7, T6-3, T6-10, T6-1, T5-4, T6-4/8/9, T5-5, T6-5, T5-3, T5-1, T5-2 and 5632^{G3}. (B) Transconjugants relative fitness (*w*) compared to the 5632^{G3} recipient clone is represented by histogram bars, and values are indicated in the table. For each transconjugants, the enrofloxacin MIC (in $\mu\text{g}\cdot\text{mL}^{-1}$), the number of PG2^{E10} fragments and the percentage of donor genome are indicated in the table. The percentage of donor genome was estimated by the total size of the PG2^{E10} transferred fragments divided by the size of the PG2^{E10} genome.

Of the 13 transconjugants analysed, 10 had received *parE* sequences from the donor. One, T5-2, has acquired two unfixed mutations, *parE*₁₁₂ and *parC*₈₄ that were pre-existing in PG2^{E10} subpopulations (92% and 60% of the reads, respectively). The remaining 9 transconjugants displayed one mutation in *parE*, not previously detected in the donor, that was either (i) an insertion of 3 nt resulting in adding an Ala residue at codon 390 or (ii) a non-synonymous SNP corresponding to codon 423 or 625. Whether the occurrence of these mutations in some transconjugants reflects the heterogeneity of the PG2^{E10} donor population, with subpopulations being selected here, or whether they have arisen independently after transfer is not known. Of note, T5-2 *parC* and T6-5 *gyrA* were more chimeric than in other transconjugants as if multiple recombination events have occurred to produce mosaic genes composed of 5632 and PG2 intermingled sequences (Fig 3A).

Interestingly, none of the transconjugants accumulated all 4 SNPs described for the PG2^{E10} in the Enro target genes. As well, none reached the 64 µg.mL⁻¹ MIC of the PG2^{E10} parental strain, the individual transconjugants' MIC values ranging from 0.5 to 32 µg.mL⁻¹ (Fig 3A). Of note, these values were higher than the concentration used for selection (0.25 µg.mL⁻¹). More specifically, transconjugants having concomitantly acquired the two distant PG2^{E10} loci containing the mutated *parE-parC* and the mutated *gyrA* had the highest MIC (16 µg.mL⁻¹ to 32 µg.mL⁻¹). In T5-1 and T5-2 that displayed the lowest MIC value (0.5 µg.mL⁻¹), the mutated *parE-parC* were co-transferred with the wild-type PG2^{E10}-*gyrB* instead of the mutated PG2^{E10}-*gyrA*. Since donor and recipient *gyrB* allelic sequences differ slightly, this event introduced amino acid changes in GyrB when compared to the parental 5632-background (S5 Table). PCR genotyping indicates that this same combination was also observed in T6-2, a transconjugants derived from the same partner but in an independent mating experiment (see S1 Fig). Finally, it is interesting to note that the co-transfer of the PG2^{E10} mutated-*gyrA* and its wt-*gyrB* was never observed. In agreement with data obtained with the spontaneous mutants, the transfer of mutated-donor *gyrA* only was not sufficient to confer the recipient strain with the Enro^R phenotype.

Overall, mutations conferring resistance to Enro with MIC values ranging from 16-32 µg.mL⁻¹ could be acquired by HGT within one mating experiment, while reaching the same levels of Enro^R MIC through spontaneous mutations would have required approximately 100 generations and multiple passages under selective pressure.

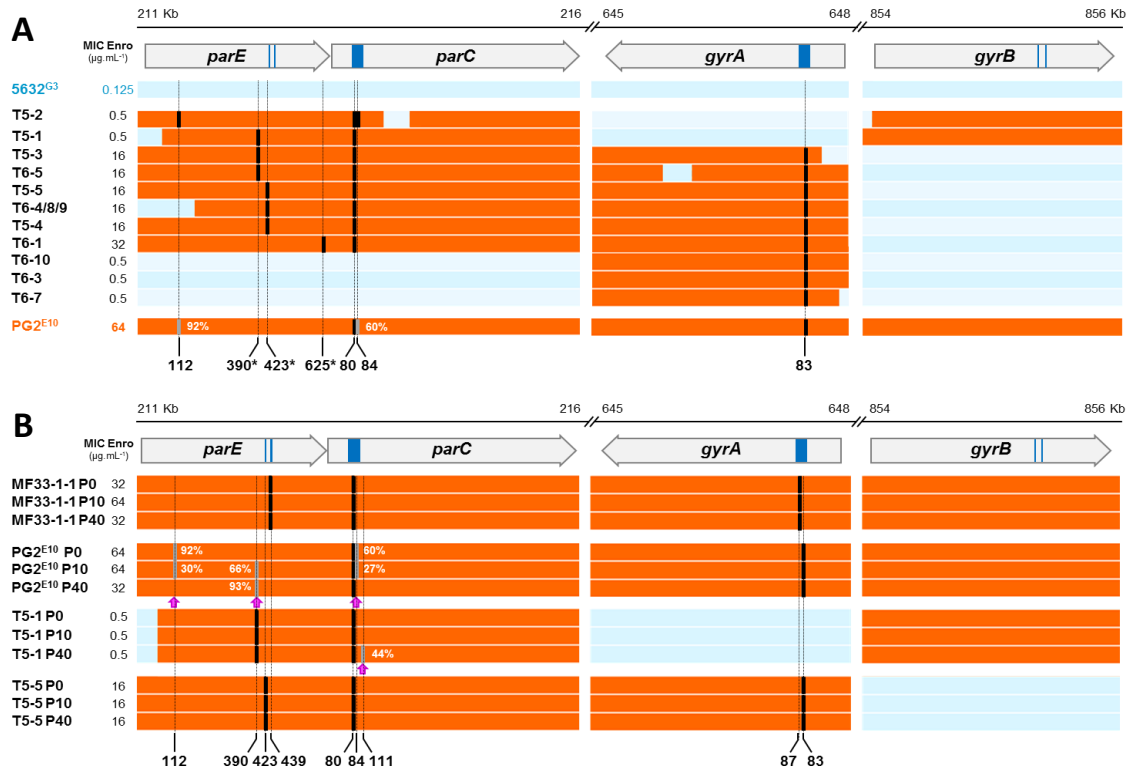


Fig 3. Features of genes targeted by the enrofloxacin in Enro^R mutants and transconjugants. (A) Schematic representing the 4 Enro target genes (*parE*, *parC*, *gyrA* and *gyrB*) of the 13 transconjugants analysed in Fig. 2. Each coding sequence is represented by an open arrow with the dark blue box correspond to the QRDR. The enrofloxacin MIC is indicated in µg.mL⁻¹. PG2 and 5632 specific sequences are represented in orange and blue, respectively. The positions of mutation are indicated in black (fixed mutations) or grey (unfixed mutations) and positions of amino acids are indicated based on *E. coli* numbering. Mutations not detected in the PG2^{E10} donor strain are indicated by a star. (B) Schematic representing the 4 Enro target genes (*parE*, *parC*, *gyrA* and *gyrB*) as above of 2 spontaneous mutants (MF33-1-1 and PG2^{E10}) and 2 transconjugants, (T5-1 and T5-5) after serial propagation in broth culture without selective pressure. Pink arrows indicated modifications observed during passages 10 and 40, corresponding to P10 and P40.

Antimicrobial selection drives the emergence of unrelated genotypes via unconventional HGT

The co-transfer of multiple loci and the possible occurrence of additional macro- and micro-heterogeneities were addressed in the 13 sequenced transconjugants. Reconstruction of the composite-genomes was performed as previously described with some minor modifications (Dordet-Frisoni et al. 2014) (see Materials and Methods). Briefly, reads generated by NGS (Next Generation Sequencing) were mapped onto the 5632 and PG2 reference genomes and reads perfectly matching to one of the other genome were retained. Analyses of the reconstructed genomes and more specifically of the PG2 inherited sequences confirmed that other fragments, unrelated to topoisomerase genes carrying Enro^R mutations, were also

exchanged (Fig 2A). This resulted in complex mosaic genomes, containing an average of 18 ± 7 PG2^{E10} fragments (Fig 2B) which size highly varied from 77 bp to 53429 bp. All transconjugants display distinct patterns of transferred fragments, except for 3, which had strictly identical genome sequences. These clones, namely T6-4, -8 and -9, were all selected from the same mating experiment and are most likely the result of the expansion of a single transconjugant as all were shown to be fitter than the parent or than other transconjugants produced during the same mating (i. e. T6-5 and T6-1) (Fig 2B). Competitive culture assays also indicated that there was no correlation between the number of fragments or the overall DNA amount that was exchanged and the fitness level (Fig 2B), with some combinations imposing a fitness cost while other conferring a fitness benefit.

Overall, the most frequently transferred regions were clustered within 20 kb around the selective Enro^R determinants, but distant loci were also exchanged in all transconjugants. This suggests that multiple events of genomic replacements by recombination occurred simultaneously. Although the PG2 and 5632 genomes are highly synthetic, some genes or regions are only present in one strain. Thus, in some cases, replacement of 5632 recipient genome by a PG2 fragment resulted in the loss or the gain of strain-specific genes. On average, 13 ± 11 5632-specific genes were lost for 6 ± 4 PG2-specific genes that were gained (S3 Table). One extreme case of replacement resulted in the deletion of a large region (ca. 22 genes, 27 Kb) which contained an integrated conjugative element (ICE) specific to 5632 and not present in PG2 (Marenda et al. 2006). This was observed in T5-2, T5-5, T6-1 and T6-7 transconjugants (Fig 2A) where the loss of the 22 genes was not due to the ICE excision but to recombination events occurring at homologous sites on each side of the ICE.

In addition, micro-complexity events were also observed, with transconjugants displaying short PG2-inherited fragments (180 ± 29 nt) that were defined by only one or two PG2-specific variations (SNPs or indels), and/or the occurrence within PG2-inherited fragments of short 5632 fragments defined by one or two 5632-specific variations.

Overall, chromosomal exchange by recombination of large or small fragments was shown to often occur within a coding sequence, resulting in chimeric PG2/5632 genes as illustrated above for *parC* and *gyrA*. On average 20 ± 7 genes were mosaic to various degrees, for each transconjugant.

In vitro multiplication without selection pressure reveals fine-mutational tuning

To evaluate the stability of the *Enro*^R spontaneous mutants and transconjugants in absence of selective pressure, two different spontaneous mutants, namely, MF33-1-1 (MIC= 32 $\mu\text{g.mL}^{-1}$) and PG2^{E10} (MIC= 64 $\mu\text{g.mL}^{-1}$), and two different transconjugants T5-1 (MIC= 0.5 $\mu\text{g.mL}^{-1}$) and T5-5 (MIC= 16 $\mu\text{g.mL}^{-1}$), were submitted to serial passages in broth medium 40 times (P0 to P40). WGS was performed with DNA extracted at P10 and P40 that correspond to approximately 150 and 600 generations, respectively (see Materials and Methods). Based on comparative analysis, the genomes of the two mutants and the two transconjugants were remarkably stable over this period of time, in agreement with the overall stability of their MIC over passages (Fig 3B and S4 Table).

Interestingly, the two unfixed mutations pre-existing in PG2^{E10} *parE* and *parC* gradually faded in favour of the wt (Fig 3B): one was detected in *parE*₁₁₂ in 92%, 30% and 0% of the reads and the other in *parC*₈₄ in 60%, 27% and 0% of the reads, at P0, P10 and P40 respectively. Concomitantly, an indel emerged in *parE*₃₉₀ at P10 (66%) and P40 (93%), that resulted in the insertion of an Ala residue. Interestingly, this same insertion occurred in 3 transconjugants: T5-1, T5-3 and T6-5, where it is fixed. As shown in Fig 3A, the absence of mutation in *parC*₈₄ in all but one transconjugant coincides with the presence of mutations in *parE*. Altogether, these data suggest that mutations in *parC* might have a slight fitness cost that tends to be compensated over passages by the introduction of mutations in *parE*, the functional partner of *parC*. Of note, other unfixed mutations were observed elsewhere in the genome during passages. In particular, MF33-1-1 displayed the highest number of unfixed mutations at P10 with 9 unfixed mutations, excluding those in *vpma* locus, most of which (6/9) being lost at P40 (S4 Table).

We then investigated the fitness of the two mutants and two transconjugants over passages in broth medium. Data presented in Fig 4 indicated that mutants MF33-1-1 and PG2^{E10} displayed a fitness similar to that of the PG2 55-5 ancestor (P0) that remained stable over passages (P10 and P40). In contrast, the fitness of transconjugants T5-1 and T5-5 at P0 was reduced by 30 to 20%, respectively and increased over successive passaging to reach 120 and 100% when compared to the recipient strain, 5632^{G3}. WGS showed that a few different, unfixed mutations accumulate over passages in both the mutants and the transconjugants, without any obvious link to fitness (see S4 Table).

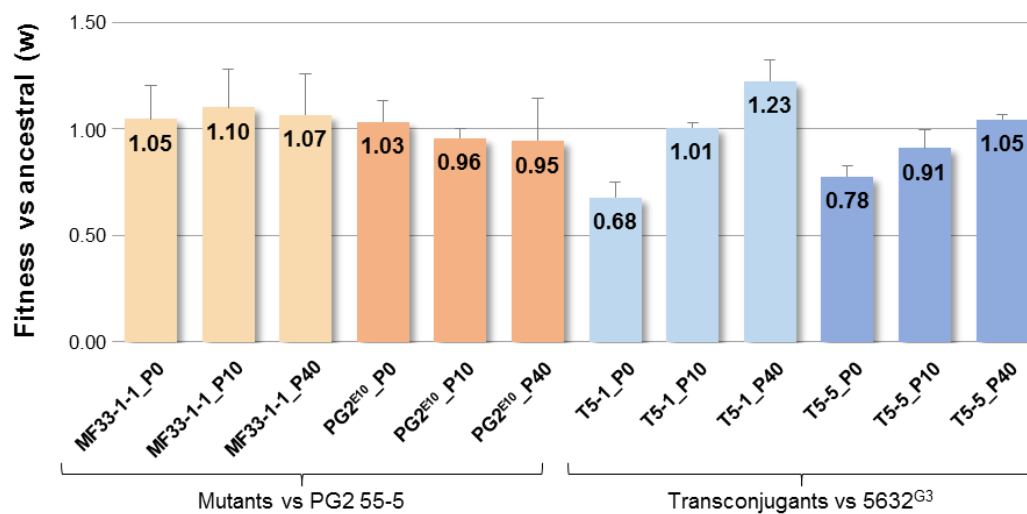


Fig 4. Fitness evolution of mutants and transconjugants over passages.

Relative fitness (w) compared to the ancestor clone is represented by histogram bars. Each fitness assay is performed by competitive assay with the ancestor clone. Fitness of spontaneous mutants PG2^{E10} and MF33-1-1 are evaluated versus PG2 55-5. Fitness of transconjugants T5-1 and T5-5 are evaluated versus 5632^{G3}. P0, P10 and P40 correspond to passages 0, 10 and 40 in the absence of enrofloxacin selection pressure.

3. Discussion

Over the past decade, HGT has increasingly attracted attention and is now recognized as a main driver of microbial innovation, with conjugation as one prominent mechanism (Ochman et al. 2000; Frost et al. 2005; Koonin 2016). Yet, knowledge about mycoplasma HGT processes, contribution and impacts is very limited with only a few publications dedicated to this topic (Teachman et al. 2002; Dordet-Frisoni et al. 2014; Citti et al. 2018; Torres-Puig et al. 2018). By combining next generation sequencing to classical mating and evolutionary experiments, the current study uncovered the role of an unconventional mechanism of gene transfer, in generating unexpected plasticity and adaptability of the small mycoplasma genome. Moreover, our data show that the sequential, mutational pathway towards the acquisition of high-level fluoroquinolone resistance can be readily compressed into a single step via the conjugative transfer of multiple chromosomal loci carrying Enro^R mutations, from highly resistant to susceptible mycoplasma cells. This event led in turn to the production mosaic genomes, which genetic diversities and fitness were beyond expectations.

Under evolutionary experimental conditions, high-level of Enro resistance in *M. agalactiae* can be reached via the emergence of spontaneous mutations when propagated with increasing concentrations of the antimicrobial, as shown for other *Mycoplasma* species (Reinhardt et al.

2002; Gruson et al. 2005; Khalil et al. 2016). Comparison of several, independent lineages indicate that these mutations accumulate following a similar trajectory: first in *parC* resulting in a 8 to 16-fold increase in resistance, followed by additional mutations in *gyrA* to reach a up to 128-fold increase. Higher resistance levels (up to 500 fold) were further achieved by combining either two mutations in *parC* with one in *gyrA* or, one mutation in each with one or more mutations in *parE*. WGS data further showed that only very few other mutations were selected and fixed outside of these genes, none that could account for the resistance phenotypes. Whether these play a role in counterbalancing a potential fitness cost during the selection process is not known but mutations in type II topoisomerase genes have no effect on PG2^{E10} fitness *in vitro* when compared to the ancestor strain ($w=1.03 \pm 0.10$). Overall, accumulation of fixed mutations in type II topoisomerase genes and high-levels of resistance were reached over several weeks of propagation, after approximately 200 or 100 generations depending on whether selection was performed with (bottleneck selection) or without intermediate cloning (Fig 1A), respectively. A limited number of reports addressed evolutionary trajectories of fluoroquinolones resistance in bacteria and each identified species-specific mutational trajectories with identical target-site mutations emerging in different order (Almahmoud et al. 2009; Sutera et al. 2014; Zhang et al. 2015). In *M. agalactiae*, the convergent outcome of parallel independent experiments strongly suggested intermolecular epistatic interactions between DNA topoisomerases in the mechanism of fluoroquinolone resistance.

The time scale of the mutational pathway leading to high-level of Enro resistance was reduced to a single conjugative step, in which several Enro^R alleles were transferred from resistant to susceptible mycoplasma cells. Such event required the physical contact and a form of sexual competence of the pair (Dordet Frisoni et al. 2013; Dordet-Frisoni et al. 2014), one partner being highly resistant while the other is susceptible. Independent mating experiments generated progenies with chimeric genomes made of the 5632 recipient chromosome in which sequences of the resistant PG2^{E10}-donor were transferred and recombined at homologous loci. Under the antimicrobial selective pressure, all transconjugants displayed the mutated donor *parE-parC* operon or/and *gyrA*, but resistance per se (from 4 up to 250-fold-increase in resistance) was reached only when both donor mutated loci were co-transferred. These data are in agreement with conclusions drawn from the analyses of co-evolved lineages (see above): alteration of both the topoisomerase IV and the DNA gyrase subunit A is critical for mycoplasmas' quinolones resistance.

Surprisingly, mutations not previously detected in neither of the parents even as a minor population, were observed in the transconjugants having acquired donor *parE-parC* sequences (corresponding to *parE*₃₉₀, *parE*₄₂₃ and *parE*₆₂₅) (Fig 3A). The mutation affecting *parE*₃₉₀ is also emerging in the donor strain after 40 serial passages in medium without Enro (Fig 3), raising the question of whether *parE* mutations (i) were pre-existing in the parent population at undetectable level and were preferentially selected after mating or (ii) occurred *de novo* during or just after mating. It is interesting to note that in the donor strain, while the mutation *parE*₃₉₀ emerged over passages, the mutations *parE*₁₁₂ and *parC*₈₄ were conversely being replaced by wild type (wt) sequences. Whether *parE*₃₉₀ is being beneficial to the transconjugants in the context of the experiment, either towards resistance or fitness, is not known.

Surprisingly, none of the selected transconjugants reached the MIC of the donor, most likely because none displayed the exact combination of *parE-parC* and *gyrA* mutations found in the predominant PG2^{E10} population. One interesting observation is that all transconjugants carried the donor, mutated *gyrA* or the wt donor *gyrB*, none having inherited both genes from a single parent. While the two strains, 5632 and PG2, encode very similar GyrA and GyrB products these are not strictly identical, with 99.2 and 98.6% identity respectively (S5 Table). Since the DNA gyrase is composed of two GyrA and two GyrB subunits, all transconjugants expressed a modified version when compared to that of the recipient cell prior to mating. Whether this provides an advantage in the context of our experiment, or whether it reflects an epistatic phenomenon (Vogwill et al. 2016; Wong 2017) remains to be addressed.

The most unexpected outcome of this study was the extent of combinatorial variation obtained after mating. Mycoplasma Chromosomal Transfer (MCT) was initially shown to differ from classical Hfr- or *OriT*-mediated transfer (Hochhut et al. 2000) in that it affects nearly every position of the genome with equal efficiency (Dordet-Frisoni et al. 2014). These data were mainly obtained by combining classical mating and molecular experiments with NGS using pools of transconjugants and we then thought that MCT mainly concerned the transfer of one or two proximate loci during mating. Here, the simultaneous transfer of several remotely located, small and large fragments generated a diversity of intra-species genomic assortment that was beyond expectations (Fig 2A). Indeed, this phenomenon created within a single step a set of totally new genomes that were a combinatorial blend of the two parents. Thanks to the significant differences existing at the nucleotide level between the two parental strains (average 1 variation every 26 nt), it was possible to reconstruct the transconjugant

genomes with a high level of precision, revealing that besides the gain and loss of entire genes, MCT also generated chimeric genes. Overall, MCT affected from 6 to 17% of the genome regardless of whether these encoded house-keeping or accessory gene functions.

An average of 18 donor-fragments co-exists in the new transconjugant genomes, of which only 2 carried the selectable *Enro*^R mutations. This implies (i) that a large amount of unrelated fragments silently co-transferred along with the selectable marker, some of which may have conferred the cell with unpredictable phenotypes under different environments and (ii) that a large proportion of the transconjugants have not been selected and that most likely, the combinatorial possibilities of conjugative MCT are endless. Because MCT introduces variation instantly, one limitation of this phenomenon is the viability and adaptability rate of the resulting chimeric cells. Although both parents are of the same species, the overall success of the transconjugants depends on how well the donor and new chimeric genes interact with the remaining recipient's genes in a particular environment (Derbyshire and Gray 2014). For a few generations, the cell may have to cope with multimeric enzymes or products which sub-units are not of a perfect match depending of the turn-over of the recipient cell protein content. To a lower extent, this situation resembles genome transplantation used to engineer mycoplasmas and thus faced a number of similar issues (Labroussaa et al. 2016). Although several transconjugants turned out to be highly resistant to *Enro*, their initial selection could only be achieved in low concentration of *Enro* (0.25 $\mu\text{g.mL}^{-1}$). This raised the question of whether growth on selective media was impaired because of the low turnover of recipient wt *GyrA/ParC* or because of a synergistic effect of the *Enro* with the gentamicin used for selecting transconjugants.

Incorporating large amount of incoming donor DNA had a fitness cost for most transconjugants but not all. Surprisingly, this was counterbalanced after a few passages in media with even one transconjugant ending with a higher fitness than the recipient or the donor cell. In contrast, passaging had no effect on the fitness of *Enro*^R spontaneous mutants derived from the donor (Fig 4) suggesting that new genome configuration may require a certain period of time for fine tuning. In search for compensatory mutations, comparative analyses of WGS before and after passages were conducted that indicated the emergence of subpopulations with an overall low number of mutations, none of which could explain the improved fitness. Quantifying pathogen fitness in its entire life cycle is not trivial (Diard and Hardt 2017) and whether transconjugants selected in this study better perform than their parents in the animal host remains to be addressed.

Overall, our study unravelled the astonishing capacity of MCT to generate unlimited genome diversity. Long thought to be reluctant to HGT, this process was proposed to counteract the erosion of the mycoplasma genome (Naito and Pawlowska 2016), but our current data indicate that it also can serve the mycoplasma short-term adaptability to changing environment. The chromosomal mosaic formed upon mycoplasma mating resembles that recently described as resulting from Distributive Conjugative Transfer (DCT) in *Mycobacterium smegmatis* (Derbyshire and Gray 2014). The exact molecular mechanism driving DCT and MCT remains to be elucidated but recent reports (Boritsch et al. 2016; Sapriel et al. 2016; Blesa et al. 2017) suggest that a similar phenomenon also occur in other bacteria and that it might be more common than first thought. While this reinforces the central role played by HGT in promoting evolutionary adaptation, this is also challenging our view on what is a bacterial species and on our capacity to predicting the emergence of new phenotypes.

4. Materials and Methods

Bacterial strains and culture conditions

Mycoplasma agalactiae strains used in this study were PG2 clone 55-5 (Sirand-Pugnet et al. 2007b) and 5632 clone C1 (Nouvel et al. 2010), further referred as PG2 and 5632 for simplicity. The 5632 gentamicin-resistant clone (5632^{G3}), designated as 5632^G-3 in previous publication (Dordet-Frisoni et al. 2014), was previously obtained by stable insertion at nucleotide 919899 of the gentamicin-resistance gene, *aacA-aphD* (Dordet-Frisoni et al. 2014). All strains were propagated at 37°C in SP4 medium (Tully 1995) supplemented with 5 mM pyruvic acid (Sigma-Aldrich) and 45 µg.mL⁻¹ cefquinome (cobactan 4.5%, MSD Animal Health) and when necessary, enrofloxacin (Sigma-Aldrich) and/or gentamicin (Sigma-Aldrich) at specified concentrations.

The doubling time (or generation time) of *M. agalactiae* PG2 55-5 was obtained by count CFU at different times of the exponential growth phase and correspond to 3.3 ± 0.14 hours for one generation. The value of 7.2 generations in 24 hours was used as a reference for estimating the generation time of mutants and transconjugants, with only growth in broth medium being taken into account.

Selection of spontaneous enrofloxacin-resistant clones under selective antibiotic pressure

Mycoplasma cells from a 1 mL of mid-exponential culture were centrifuged for 15 min at 8000 g at room temperature and re-suspended in SP4 medium containing 0.5 $\mu\text{g.mL}^{-1}$ enrofloxacin. After 48h, 10 μl cultures of 10^8 to 10^9 CFU.mL⁻¹ were plated on SP4 agar plates with increasing enrofloxacin concentrations (0.5 to 2 $\mu\text{g.mL}^{-1}$). Colonies were only obtained on plates containing 0.5 $\mu\text{g.mL}^{-1}$ enrofloxacin, they were picked and propagated in SP4 broth medium with the same antimicrobial concentration. This represented the first step of selection that founded the isogenic lineages used in this study. One to 3 additional rounds of selection were similarly performed with increasing concentration of enrofloxacin (ranging from 0.5 to 32 $\mu\text{g.mL}^{-1}$), colonies growing on the highest concentration being picked and subjected to the next round. A total of 108 clones were obtained, among them 22 clones corresponding to 6 lineages were analysed. In parallel, PG2 55-5 culture (10^8 to 10^9 CFU.mL⁻¹) were propagated by serial passaging (dilution 1/50 or 1/100), in broth medium containing increasing enrofloxacin concentration (0.25, 0.5, 1 and 10 $\mu\text{g.mL}^{-1}$) to generate the resistant PG2^{E10} population.

Characterisation of spontaneous enrofloxacin-resistant clones

Mutations in the Quinolone Resistance Determining Region (QRDR) sequences of *gyrA*, *gyrB*, *parC* and *parE* were identified by direct sequencing with the BigDye Terminator chemistry (Heiner et al. 1998; Hudson et al. 2006) using primers listed in S1 Table and genomic DNA extracted with the chloroform method (Chen and Kuo 1993), and/or by whole genome sequencing (see below). Direct sequencing of genomic DNA was performed at the genomic platform of Get-Purpan (Toulouse, France). Of note, in the text and tables, the numbering of amino acid position of type II topoisomerases referred to the *Escherichia coli* K-12 strain, GyrA (AAC75291.1), GyrB (AAT48201.1), ParC (AAC76055.1) and ParE (AAA69198.1).

Selection of enrofloxacin-resistant transconjugants by mating

Mating experiments were performed as previously described (Dordet Frisoni et al. 2013). Briefly, the donor strain (PG2^{E10}) and the recipient strain (5632^{G3}) were grown individually in SP4 medium, during 24 h. The two cultures were mixed at a 5632^{G3}:PG2^{E10} cell ratio of 1:10 for one experiment (T5) and 10:1 for the second (T6) and then centrifuged for 5 min at 8000 g at room temperature. Cells were re-suspended in SP4 medium, incubated during 16 h at 37°C and

an aliquot of 300 μ l was plated in SP4 agar containing gentamicin (50 μ g.mL⁻¹) and different enrofloxacin concentrations (from 0.25 to 8 μ g.mL⁻¹). Colonies were only observed at 0.25 μ g.mL⁻¹ of enrofloxacin (except for one transconjugants T5-5, obtained at 0.5 μ g.mL⁻¹) they were picked before propagation in SP4 liquid medium containing the same concentration of antibiotics. Mating experiments using PG2 55-5 (Enro^S) and 5632^{G3} were used as negative control, to test the absence of enrofloxacin spontaneous resistant clones. The frequency of transconjugants was determined as the number of transconjugants, divided by the number of PG2^{E10} donor parental cells.

Characterisation of potential transconjugants by PCR

PG2- or 5632-specific PCR assays were used to determine the parental origin of 11 genomic loci across the transconjugant genomes (S1 Table). Of these, 8 were previously described (Dordet-Frisoni et al. 2014) and 3 that targeted *parC*, *gyrA* and *gyrB* were specifically designed for this study. PCR assays were conducted using genomic DNA extracted with the chloroform method (Chen and Kuo 1993) and primers listed in S1 Table. The presence and position of the 5632^{G3}-specific gentamicin resistance marker (Gm) was confirmed by a specific PCR using one primer inside the marker and the other in the flanking chromosomal sequence (S1 Table). All PCR amplifications were performed according to the recommendations of the Taq DNA polymerase suppliers (M0267S, New England Biolabs).

Determination of the Enrofloxacin Minimum Inhibitory Concentrations (MICs)

The enrofloxacin MICs were determined according to the recommendation of Hannan 2000 (Hannan 2000) using the agar dilution method as previously described (Khalil et al. 2016). Briefly, 1 μ L of each clone diluted to 10⁴-10⁵ CFU.mL⁻¹ was spotted on agar plates containing serial two-fold dilution of enrofloxacin (from 0.0625 to 64 μ g.mL⁻¹). MIC assays were performed in triplicates for each clone, and the median value was retained. The MIC was defined as the lowest concentration of enrofloxacin that prevented visible growth after 5 days at 37°C while, in parallel 30 to 300 CFU were observed on the antimicrobial-free control plate. Based on Hannan 2000 (Hannan 2000), we considered in this study isolates with MIC of \leq 0.5 μ g.mL⁻¹ as susceptible (Enro^S) while MIC \leq 1 μ g.mL⁻¹ and \geq 2 μ g.mL⁻¹ corresponded to intermediate and resistant isolates (Enro^R). Here, isolates with MIC \geq 16 μ g.mL⁻¹ were further referred as being highly resistant.

Fitness competition assays

A pairwise competition assay was performed to estimate the relative fitness of evolved strains versus their ancestor (i. e. transconjugants versus 5632^{G3} or spontaneous mutants versus PG2 55-5). For each pair, the evolved and the ancestor clones were mixed in SP4 medium at a 1:1 cell ratio (10^4 CFU.mL⁻¹). Serial dilutions of the starting (0h, T0) and final (18h, T18) co-cultures were plated on SP4 plates containing none or 4 µg.mL⁻¹ of enrofloxacin. After 5 days at 37°C, the number of CFU was determined for the ancestor and the evolved clones. Fitness of each clone relative to its ancestor was calculated according to the equation:

$$w = \text{Fitness}_{\text{evolved/ancestor}} = \ln(\text{evolved at T18}/\text{evolved at T0})/\ln(\text{ancestor at T18}/\text{ancestor at T0})$$

(Gagneux et al. 2006; Wisser and Lenski 2015). Concerning the fitness of Enro^S Genta^R transconjugants with MIC ≤ 0.5 µg.mL⁻¹, selection onto enrofloxacin solid media was obviously not feasible. These were then performed using 5632 as the ancestor and the gentamicin as selective antimicrobial (50 µg.mL⁻¹). Their fitness ratio was then corrected by multiplying by 1.17, a value equal to the fitness ratio of 5632 versus 5632G3. At least three replicates were performed for each assay, the mean value and the standard deviation (SD) were indicated. A value of 1 indicated a fitness of the evolved strain similar to the ancestor (5632^{G3} or PG2 55-5), a ratio lesser than 1 or greater than 1 indicated a fitness-cost or -benefit for the evolved strain, respectively.

Illumina Whole Genome Sequencing and bioinformatic analyses

Genomic DNA was extracted from mycoplasma cells using the phenol-chloroform method (Sambrook et al. 1989). Whole genome sequencing was performed at the GATC Biotech facility (Konstanz, Germany) using Illumina technology HiSeq (paired-end, 2x150 bp). An average of 2×10^7 reads by mutants or transconjugants was obtained, corresponding to an average of 3100X for coverage depth. One exception is the PG2^{E10} population that was sequenced by the Genome-Transcriptome facility of Bordeaux (France) using HiSeq (paired-end, 2x100 bp, 1.6×10^7 reads, coverage 1700X). All bioinformatics analyses were performed using the galaxy platform hosted by Genotoul, Toulouse, France (bioinfo.genotoul.fr). The reads of each clone (fastq file) were mapped on the reference genome *M. agalactiae* PG2 (NC_009497.1) or 5632 (NC_013948.1), using Burrows-Wheeler Aligner (BWA, MEM algorithms, Galaxy version 0.8.0) (Li and Durbin 2009). The quality of the alignments was controlled with Qualimap 2.2.1 (Okonechnikov et al. 2016). Calling variant analysis was performed using GATK3 (Galaxy version 3.5.0) for SNPs and indels detection (McKenna et al. 2010). Variations with a quality

lower than 10000 were excluded (Filter VCF file tool, Galaxy Version 1.0.0). Variations were considered as (i) fixed when present in $\geq 95\%$ of the reads or (ii) unfixed when present in $< 95\%$ of the reads, as a result of coexisting sub-populations (Sun et al. 2012). The percentage of each variation was calculated using the ratio AD/DP (AD: Allelic depths for the reference and alternative alleles; DP: Approximate read depth) provided by GATK. Alignments (bam file) and variations (vcf file) were visualized using the Integrative Genome Viewer (IGV 2.3.93) (Thorvaldsdóttir et al. 2013), Artemis 16.0.0 (Rutherford et al. 2000) and ACT 13.0.0 (Carver et al. 2005). Reconstruction of the composite genome of transconjugants (PG2/5632) was possible because of the frequent polymorphisms existing between PG2 and 5632, on average 1 variation every 26 nt calculated using Nucmer (Kurtz et al. 2004). This was performed by PG2 specific reads detection as follows: transconjugants reads were aligned on the 5632 genome, reads with mismatch were recovered (select lines tool, Galaxy version 1.0.1) and these reads were then aligned on the PG2 genome. Only reads with no mismatch and regions with a coverage higher than 15 reads were conserved. These mapped reads, corresponding to PG2 transferred regions, were manually curated using Artemis.

Data availability

All sequence data (fastq) have been deposited at the EMBL database, the European Nucleotide Archive (ENA) at <http://www.ebi.ac.uk/ena> under study accession numbers PRJEB28036 for spontaneous mutants and PRJEB27571 for transconjugants. Raw data accession numbers are summarized in Table S6.

Acknowledgements

We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigenae group for providing help and storage resources. We would like to thank Dr Pascal Sirand-Pugnet, University of Bordeaux, for providing assistance in sequencing PG2^{E10} and for helpful discussion. We also thank Agnès Tricot and Adélie Colin for their technical assistance in performing MICs assays.

5. References

- Almahmoud I, Kay E, Schneider D, Maurin M. Mutational paths towards increased fluoroquinolone resistance in *Legionella pneumophila*. *J Antimicrob Chemother.* 2009 Aug;64(2):284–93.
- Baranowski E, Dordet-Frisoni E, Sagné E, Hygonenq M-C, Pretre G, Claverol S, et al. The Integrative Conjugative Element (ICE) of *Mycoplasma agalactiae*: Key Elements Involved in Horizontal Dissemination and Influence of Coresident ICEs. *MBio.* 2018 Jul 3;9(4).
- Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. The transjugation machinery of *Thermus thermophilus*: Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet.* 2017 Mar;13(3):e1006669.
- Boritsch EC, Khanna V, Pawlik A, Honoré N, Navas VH, Ma L, et al. Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. *Proc Natl Acad Sci USA.* 2016 30;113(35):9876–81.
- Burrus V, Pavlovic G, Decaris B, Guédon G. Conjugative transposons: the tip of the iceberg. *Mol Microbiol.* 2002 Nov;46(3):601–10.
- Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics.* 2005 Aug 15;21(16):3422–3.
- Chen WP, Kuo TT. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res.* 1993 May 11;21(9):2260.
- Citti C, Blanchard A. Mycoplasmas and their host: emerging and re-emerging minimal pathogens. *Trends in Microbiology.* 2013 Apr;21(4):196–203.
- Citti C, Browing GF, Rosengarten R. Phenotypic diversity and cell invasion in host subversion by pathogenic mycoplasmas. In: *Mycoplasmas Molecular Biology Pathogenicity and Strategies for Control*. Horizon Bioscience. Wymondham: Blanchard Alain, Glenn Browing; 2005. p. 439–83.
- Citti C, Dordet-Frisoni E, Nouvel LX, Kuo CH, Baranowski E. Horizontal Gene Transfers in Mycoplasmas (Mollicutes). *Curr Issues Mol Biol.* 2018 Apr 12;29:3–22.
- Citti C, Nouvel L-X, Baranowski E. Phase and antigenic variation in mycoplasmas. *Future Microbiology.* 2010 Jul;5(7):1073–85.
- Correia S, Poeta P, Hébraud M, Capelo JL, Igrejas G. Mechanisms of quinolone action and resistance: where do we stand? *Journal of Medical Microbiology.* 2017;66(5):551–9.
- Derbyshire KM, Gray TA. Distributive Conjugal Transfer: New Insights into Horizontal Gene Transfer and Genetic Exchange in Mycobacteria. *Microbiol Spectr.* 2014;2(1).
- Diard M, Hardt W-D. Evolution of bacterial virulence. *FEMS Microbiol Rev.* 2017 01;41(5):679–97.
- Dordet Frisoni E, Marena MS, Sagné E, Nouvel LX, Guérillot R, Glaser P, et al. ICEA of *Mycoplasma agalactiae* : a new family of self-transmissible integrative elements that

- confers conjugative properties to the recipient strain: *Mycoplasma* ICE on the move. *Molecular Microbiology*. 2013 Sep;89(6):1226–39.
- Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX, Blanchard A, et al. Chromosomal Transfers in Mycoplasmas: When Minimal Genomes Go Mobile. *mBio*. 2014 Dec 31;5(6):e01958-14.
- Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol*. 2005 Sep;3(9):722–32.
- Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science*. 2006 Jun 30;312(5782):1944–6.
- Gaurivaud P, Baranowski E, Pau-Roblot C, Sagné E, Citti C, Tardy F. *Mycoplasma agalactiae* Secretion of β -(1→6)-Glucan, a Rare Polysaccharide in Prokaryotes, Is Governed by High-Frequency Phase Variation. *Appl Environ Microbiol*. 2016 Jun 1;82(11):3370–83.
- Gautier-Bouchardon AV. Antimicrobial Resistance in *Mycoplasma* spp. *Microbiol Spectr*. 2018 Jul;6(4).
- 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 Aug;68(8):4539–48.
- Gray TA, Krywy JA, Harold J, Palumbo MJ, Derbyshire KM. Distributive conjugal transfer in mycobacteria generates progeny with meiotic-like genome-wide mosaicism, allowing mapping of a mating identity locus. *PLoS Biol*. 2013 Jul;11(7):e1001602.
- Gruson D, Pereyre S, Renaudin H, Charron A, Bebear C, Bebear CM. In Vitro Development of Resistance to Six and Four Fluoroquinolones in *Mycoplasma pneumoniae* and *Mycoplasma hominis*, Respectively. *Antimicrobial Agents and Chemotherapy*. 2005 Mar 1;49(3):1190–3.
- Hannan PC. Guidelines and recommendations for antimicrobial minimum inhibitory concentration (MIC) testing against veterinary mycoplasma species. *Veterinary research*. 2000;31(4):373–395.
- Heiner CR, Hunkapiller KL, Chen S-M, Glass JI, Chen EY. Sequencing Multimegabase-Template DNA with BigDye Terminator Chemistry. *Genome Res*. 1998 May;8(5):557–61.
- Hochhut B, Marrero J, Waldor MK. Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in *Vibrio cholerae* O139. *J Bacteriol*. 2000 Apr;182(7):2043–7.
- Hooper DC, Jacoby GA. Mechanisms of drug resistance: quinolone resistance. *Ann N Y Acad Sci*. 2015 Sep;1354:12–31.
- Hopkins KL, Davies RH, Threlfall EJ. Mechanisms of quinolone resistance in *Escherichia coli* and *Salmonella*: recent developments. *Int J Antimicrob Agents*. 2005 May;25(5):358–73.
- Hudson P, Gorton TS, Papazisi L, Cecchini K, Frasca S, Geary SJ. Identification of a virulence-associated determinant, dihydrolipoamide dehydrogenase (Ipd), in *Mycoplasma*

- gallisepticum through in vivo screening of transposon mutants. *Infect Immun*. 2006 Feb;74(2):931–9.
- Husain F, Tang K, Veeranagouda Y, Boente R, Patrick S, Blakely G, et al. Novel large-scale chromosomal transfer in *Bacteroides fragilis* contributes to its pan-genome and rapid environmental adaptation. *Microb Genom*. 2017 Nov;3(11).
- Kenny GE, Young PA, Cartwright FD, Sjöström KE, Huang WM. Sparfloxacin Selects Gyrase Mutations in First-Step *Mycoplasma hominis* Mutants, whereas Ofloxacin Selects Topoisomerase IV Mutations. *Antimicrob Agents Chemother*. 1999 Oct;43(10):2493–6.
- Khalil D, Becker CAM, Tardy F. Alterations in the Quinolone Resistance-Determining Regions and Fluoroquinolone Resistance in Clinical Isolates and Laboratory-Derived Mutants of *Mycoplasma bovis*: Not All Genotypes May Be Equal. *Appl Environ Microbiol*. 2016 Feb;82(4):1060–8.
- Koonin EV. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Res*. 2016;5.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. *Genome Biol*. 2004;5(2):R12.
- Labroussaa F, Lebaudy A, Baby V, Gourgues G, Matteau D, Vashee S, et al. Impact of donor-recipient phylogenetic distance on bacterial genome transplantation. *Nucleic Acids Res*. 2016 30;44(17):8501–11.
- Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med*. 2004 Dec;10(12 Suppl):S122-129.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754–60.
- Lo W-S, Gasparich GE, Kuo C-H. Found and Lost: The Fates of Horizontally Acquired Genes in Arthropod-Symbiotic *Spiroplasma*. *Genome Biol Evol*. 2015 Aug 8;7(9):2458–72.
- 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. *Journal of Bacteriology*. 2006 Jun 1;188(11):4137–41.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Jan 9;20(9):1297–303.
- Naito M, Pawlowska TE. Defying Muller’s Ratchet: Ancient Heritable Endobacteria Escape Extinction through Retention of Recombination and Genome Plasticity. *mBio*. 2016 Jun 7;7(3):e02057-15.
- Nouvel LX, Sirand-Pugnet P, Marenda MS, Sagné E, Barbe V, Mangenot S, et al. Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro-and micro-events that are shaping mycoplasma diversity. *BMC genomics*. 2010;11(1):86.

- Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. 2000 May;405(6784):299–304.
- Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. 2016 Jan 15;32(2):292–4.
- Pereyre S, Sirand-Pugnet P, Beven L, Charron A, Renaudin H, Barré A, et al. Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital mycoplasmas. *PLoS Genet*. 2009 Oct;5(10):e1000677.
- Razin S, Yogev D, Naot Y. Molecular Biology and Pathogenicity of Mycoplasmas. *Microbiol Mol Biol Rev*. 1998 Dec;62(4):1094–156.
- Redgrave LS, Sutton SB, Webber MA, Piddock LJV. Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success. *Trends in Microbiology*. 2014 Aug;22(8):438–45.
- Reinhardt AK, Kempf I, Kobisch M, Gautier-Bouchardon AV. Fluoroquinolone resistance in *Mycoplasma gallisepticum*: DNA gyrase as primary target of enrofloxacin and impact of mutations in topoisomerases on resistance level. *J Antimicrob Chemother*. 2002 Oct;50(4):589–92.
- Rosengarten R, Citti C, Much P, Spargser 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;8:166–85.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics*. 2000 Oct;16(10):944–5.
- Sambrook J, Fritsch E, Maniatis T. *Molecular cloning: A laboratory manual: Vol. 2*. 2. ed. S.I.: Cold Spring Harbor; 1989.
- Sapriel G, Konjek J, Orgeur M, Bouri L, Frézal L, Roux A-L, et al. Genome-wide mosaicism within *Mycobacterium abscessus*: evolutionary and epidemiological implications. *BMC Genomics*. 2016 Feb 17;17:118.
- Sirand-Pugnet P, Citti C, Barré A, Blanchard A. Evolution of mollicutes: down a bumpy road with twists and turns. *Research in Microbiology*. 2007a Dec;158(10):754–66.
- Sirand-Pugnet P, Lartigue C, Marendra M, Jacob D, Barré A, Barbe V, et al. Being Pathogenic, Plastic, and Sexual while Living with a Nearly Minimal Bacterial Genome. *PLoS Genetics*. 2007b;3(5):e75.
- Skapski A, Hygonenq M-C, Sagné E, Guiral S, Citti C, Baranowski E. Genome-Scale Analysis of *Mycoplasma agalactiae* Loci Involved in Interaction with Host Cells. Neyrolles O, editor. *PLoS ONE*. 2011 Sep 23;6(9):e25291.
- Sun G, Luo T, Yang C, Dong X, Li J, Zhu Y, et al. Dynamic Population Changes in *Mycobacterium tuberculosis* During Acquisition and Fixation of Drug Resistance in Patients. *J Infect Dis*. 2012 Dec 1;206(11):1724–33.
- Sutera V, Levert M, Burmeister WP, Schneider D, Maurin M. Evolution toward high-level fluoroquinolone resistance in *Francisella* species. *J Antimicrob Chemother*. 2014 Jan;69(1):101–10.

- Teachman AM, French CT, Yu H, Simmons WL, Dybvig K. Gene Transfer in *Mycoplasma pulmonis*. *J Bacteriol*. 2002 Feb;184(4):947–51.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinformatics*. 2013 Mar;14(2):178–92.
- Torres-Puig S, Martínez-Torró C, Granero-Moya I, Querol E, Piñol J, Pich OQ. Activation of σ^{20} -dependent recombination and horizontal gene transfer in *Mycoplasma genitalium*. *DNA Res*. 2018 Apr 11;
- Tully JG. Culture medium formulation for primary isolation and maintenance of mollicutes. In: *Molecular and diagnostic procedures in mycoplasmaology: molecular characterization*. Academic Press. San Diego: J. G. Tully; 1995. p. 33–9.
- Vasconcelos ATR, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, 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 Aug;187(16):5568–77.
- Vogwill T, Kojadinovic M, MacLean RC. Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of *Pseudomonas*. *Proc Biol Sci*. 2016 May 11;283(1830).
- Waites KB, Lysnyansky I, Bébéar C. Emerging antimicrobial resistance in mycoplasmas of humans and animals. In: *Mollicutes: Molecular biology and pathogenesis*. Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browning; 2014. p. 289–322.
- Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev*. 2011 Sep;35(5):957–76.
- von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol*. 2016;7:173.
- Wiser MJ, Lenski RE. A Comparison of Methods to Measure Fitness in *Escherichia coli*. *PLoS ONE*. 2015;10(5):e0126210.
- Wong A. Epistasis and the Evolution of Antimicrobial Resistance. *Front Microbiol*. 2017;8:246.
- Wozniak RAF, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Micro*. 2010 Aug;8(8):552–63.
- Xiao L, Paralanov V, Glass JI, Duffy LB, Robertson JA, Cassell GH, et al. Extensive horizontal gene transfer in ureaplasmas from humans questions the utility of serotyping for diagnostic purposes. *J Clin Microbiol*. 2011 Aug;49(8):2818–26.
- Yoshida H, Bogaki M, Nakamura M, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 1990;34(6):1271–1272.
- Yoshida H, Bogaki M, Nakamura M, Yamanaka LM, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrB* gene of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*. 1991;35(8):1647–1650.

Zhang G, Wang C, Sui Z, Feng J. Insights into the evolutionary trajectories of fluoroquinolone resistance in *Streptococcus pneumoniae*. *Journal of Antimicrobial Chemotherapy*. 2015 Sep;70(9):2499–506.

6. Supporting information

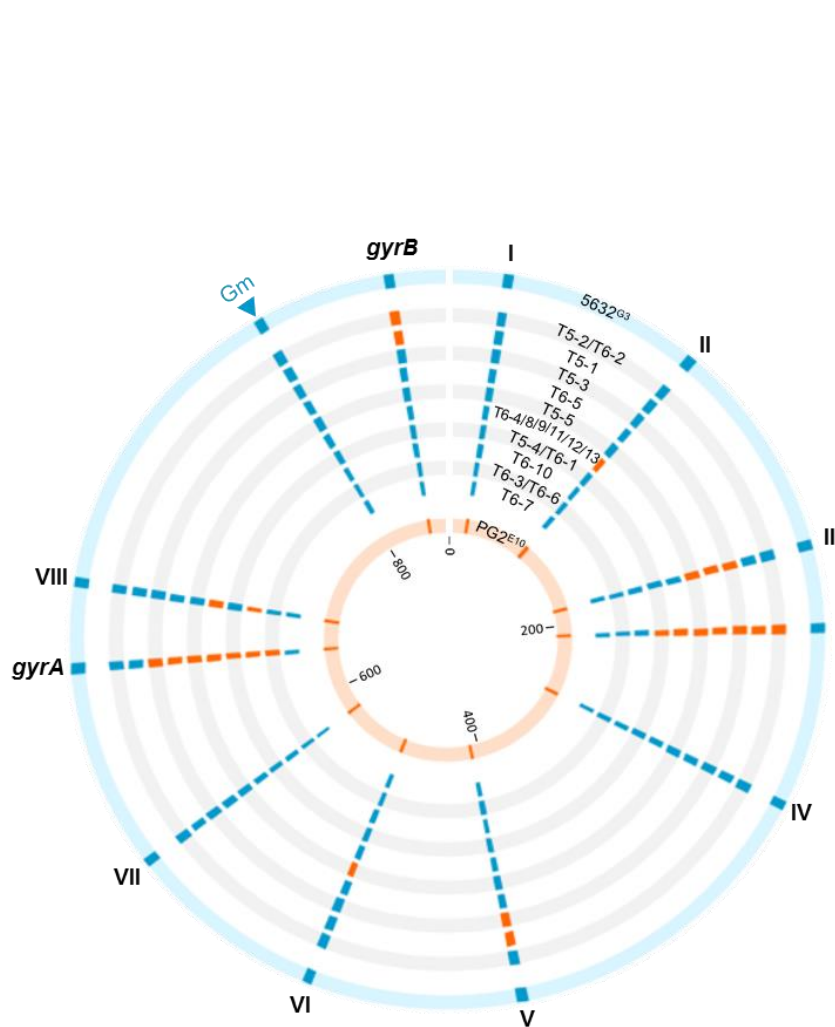


Figure S1: Loci targeted by specific-PG2 or -5632 PCR assays.

Loci targeted by specific-PG2 or -5632 PCR assays are depicted in orange and blue, respectively (see also Table S1 for primer list) and positioned along circles that represent the parental and transconjugant genomes. Genomes from inner to outer circle are: PG2^{E10}, T6-7, T6-3/T6-6, T6-10, T5-4/T6-1, T6-4/8/9/11/12/13, T5-5, T6-5, T5-3, T5-1, T5-2/T6-2 and 5632^{G3}.

Table S1 : Primers used for Sanger sequencing and PCR screening.**Primer for Sanger target sequencing of QRDR**

QRDR target	Primer name	Primer sequence (5'->3')	Primer position (PG2 genome)	Sequenced regions (bp)
ParE	QRDR_ParE_F	AGACAAAATTAAGCGTGCATCTGA	213099-213122	403
	QRDR_ParE_R	ATGCGCTCCATCAGTATCAGC	213481-213501	
ParC	QRDR_ParC_F	GCACTTCTGATGTGCGTG	214039-214057	364
	QRDR_ParC_R	TTTCAGAATCGTCAAATTAGGAGC	214378-214402	
GyrA	QRDR_GyrA_F	TTGTTTCTCGTCTTTGCCT	648140-648159	383
	QRDR_GyrA_R	TGGCTCCTTTTCAGTTGCATC	647777-647797	
GyrB	QRDR_GyrB2_F	GCTGTGCAAGCAAATATGGGT	855727-855747	336
	QRDR_GyrB_R	ATCGACGTCGGCATCAGTC	856044-856062	

Table S1 : Continued (2/2).

Primer for transconjugant characterisation by PCR

Locus target	Primer name	Primer sequence (5'->3')	Primer position PG2 genome	Primer position 5632 genome	Product size in PG2 (bp)	Product size in 5632 (bp)	Annealing temperature (°C)	Reference
I	MAG0220F	CGCCATTGAAGGAACCTGCTCC	-	21895-21916	0	947	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAG0220R	TATGCCTGCAGATCAGTCACCA	-	22820-22841				
II	MAGa0900F	ATGGCAATTGATTACTCCC	97294-97313	98968-98987	971	1516	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAGa0900R	ATGATGTTGGCCGTATTGG	98245-98264	100464-100483				
III	MAG1560F	GCCGGTGCCCGGTTAAT	182471-182488	-	1311	0	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAG1560R	TCGACACCGCATTTAGCAGCA	183761-183781	-				
IV	MAG2450_F	AGGACGCTGTAGAGTTTGCAAAGAA	283862-283886	-	1508	0	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAG2450_R	TACCACTAGCTGCCGACCA	285350-285369	-				
V	MAGa3840F	GAAGCACTAAAATCAGGCAA	410721-410740	440006-440025	1067	1223	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAGa3840R	GCGGCAGCAGAGTTATCAAG	409674-409693	438803-438822				
VI	MAGa4400F	CTCTGGCTACAGATTCATCA	491702-491721	508980-508999	607	490	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAGa4400R	GGTGAACGAGAACAATCGCC	492289-492308	509450-509469				
VII	pks1-F	AGAAGGCTTTAAGGCCGGTA	566932-566951	-	1085	609	55	(Dordet Frisoni <i>et al.</i> 2013)
	pks1R2	GCCTTTGGTTTCTGTTTTGC	567997-568016	619068-619087				
	ISMAG1_R	AAAGGCAATAGAAAGCACCGC	-	618479-618499				
VIII	MAG5890R	ACATTGCGATGTAGCGGAACAG	680401-680423	-	1318	494	55	(Dordet Frisoni <i>et al.</i> 2013)
	MAG5860F	ACCATTGAACTACATTTGCA	679106-679125	745715-745734				
	MAGa6510R	AAAACCGCAGACCCAAATGGT	-	746188-746208				
ParC spe	ParC_spPG2R	TTACTATTTCGAGTCTTGCCCTTA	215055-215077	-	1039	1039	55	This study
	QRDR_ParC_F	GCACTTCCTGATGTGCGTG	214039-214057	215927-215945				
	ParC_sp5632R	TTACTATTTCGAGTCTTCTCTG	-	216943-216965				
GyrA spe	GyrA_spPG2F	GATACTAAATACAGCAGTATGTTAAAA	648799-648825	-	1049	1049	58	This study
	QRDR_GyrA_R	TGGCTCCTTTTCAGTTGCATC	647777-647797	712116-712136				
	GyrA_sp5632F	GATACTAAATTCAGTAGTATGACAAGT	-	713138-713164				
GyrB spe	GyrB_spPG2R	TGTTGAATTTTATATAGCGGA	856152-856172	-	848	848	62	This study
	GyrB F	CAATACACATCAACCTTCCA	855325-855344	984590-984609				
	GyrB_sp5632R	TGCTGAATTTGTACAGTGGT	-	985417-985437				
Gm	5632G3	CCCGGAAGGTATATTTATGG	-	920198-920218	0	346	55	This study (Skapski <i>et al.</i> 2011)
	SG5	TTTTACACAATTATACGGACTTTATC	-	-				

Table S2 : Position, nature and impact of mutations in whole genome of Enro^R mutants.

Region	Nucleotide position ^(e)	Mutation ^(b)	Locus tag	Gene name ^(c)	Product ^(d)	Homo polymeric tract	Impact of SNP ^(f)	Name (MIC Enro) ^(e)																	
								MF26 (1)	MF26-1 (16)	MF26-1-2 (32)	MF33 (1)	MF33-1 (16)	MF33-1-1 (32)	MF33-1-2 (64)	MF33-1-3 (64)	MF30 (2)	MF30-1 (4)	MF30-1-4 (16)	MF30-1-4-1 (64)	MF30-1-4-8 (64)	PG2 ^{E10} (64)				
	60146	SNP TAC (Y) ->TAT (Y)	MAG_0500	<i>gpsA</i>	Glycerol-3-phosphate dehydrogenase		S	-	-	100	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	81704	SNP CTC (L) -> CTT (L)	MAG_0670	<i>pheS</i>	Phenylalanine-tRNA ligase subunit alpha		S	-	-	-	-	-	-	-	-	-	-	95	-	-	-	-	-	-	
	101548	SNP TGC (C) -> TTC (F)	MAG_0890	<i>hprk</i>	HPr kinase phosphorylase		NS	-	-	-	-	-	84	-	-	-	-	-	-	-	-	-	-	-	
	129186	SNP GTT (V) -> ATT (I)	MAG_1150	.	16S rRNA maturation RNase YbeY		NS	-	-	-	24	100	100	-	-	-	-	-	-	-	-	-	-	-	
	147623	Del A	MAG_1300	.	HP	6 A		-	-	-	-	-	-	-	-	-	-	-	-	-	-	28	-	-	
	159588	SNP A-> T	MAG_1380 /1390	<i>fold/pta</i>	Methenyltetrahydrofolate/Phosphate acetyltransferase			-	-	-	-	-	-	-	-	60	-	-	-	-	-	-	-	-	
	160396	Inser C	MAG_1390	<i>pta</i>	Phosphate acetyltransferase		Stop	60	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	161297	SNP CCA (P) -> CAA (Q)	MAG_1400	<i>ackA</i>	Acetate kinase		NS	36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	177174	Del T	MAG_1520	.	HP		Stop	-	-	-	-	83	-	-	-	-	-	-	-	-	-	-	-	-	
	183721	Del G	MAG_1570	.	HP	9 G		-	98	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	194231	SNP GAA (E) -> GAC (D)	MAG_1660	<i>dnaE</i>	DNA polymerase III subunit alpha		NS	-	-	-	-	-	-	-	-	-	-	-	-	100	-	-	-	-	
ParE	212252	SNP GCT (A) -> GTT (I)	MAG_1820	<i>parE</i> (codon 86)	DNA topoisomerase IV subunit B		NS	-	-	-	45	-	-	-	-	-	-	-	-	-	-	-	-	-	
	212329	SNP CAT (H) -> TAT (Y)	MAG_1820	<i>parE</i> (codon 112)	DNA topoisomerase IV subunit B		NS	-	-	-	-	-	-	-	-	-	-	-	-	100	100	-	92	-	
	213271*	SNP GCT (A) -> TCT (S)	MAG_1820	<i>parE</i> (codon 422)	DNA topoisomerase IV subunit B		NS	-	-	-	-	-	-	100	-	-	-	-	-	-	-	-	-	-	
	213284*	SNP GCT (A) -> GTT (V)	MAG_1820	<i>parE</i> (codon 426)	DNA topoisomerase IV subunit B		NS	-	-	100	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	213322*	SNP CCG (P) -> TCG (S)	MAG_1820	<i>parE</i> (codon 439)	DNA topoisomerase IV subunit B		NS	-	-	-	-	100	100	-	-	-	-	-	-	-	-	-	-	-	-

Table S2 : Continued (2/4).

ParC	214190*	SNP ACT (T) -> ATT (I)	MAG_1830	<i>parC</i> (codon 80)	DNA topoisomerase IV subunit A	NS	100	100	100	100	100	100	100	100	100	100	100	100	100
	214201*	SNP GAT (D) -> AAT (N)	MAG_1830	<i>parC</i> (codon 84)	DNA topoisomerase IV subunit A	NS	-	-	-	-	-	-	-	100	100	100	100	100	60
	214828	SNP CGT (R) -> TGT (C)	MAG_1830	<i>parC</i> (codon 291)	DNA topoisomerase IV subunit A	NS	-	-	-	-	-	-	-	-	100	-	-	-	-
	215656	SNP TCA (S) -> CCA (P)	MAG_1830	<i>parC</i> (codon 574)	DNA topoisomerase IV subunit A	NS	-	-	-	-	-	-	-	-	95	-	-	-	-
	248847	Inser G	MAG_2160 /2170	./.	HP/HP		-	-	-	-	-	-	100	-	-	-	24	-	-
	282992	Del C	MAG_2420 /2430	./.	Lipoprotein/Lipoprotein	9 C	-	-	-	-	-	-	-	94	97	-	-	-	-
	284389	Inser/del CT	MAG_2440 /2450	./.	PTS sugar transporter subunit IIB/Lipoprotein	10 CT	34	-	-	-	-	83	82	-	-	-	-	-	-
	287686	Inser AC	MAG_2470	.	Proton-glutamate symporter		-	-	-	-	-	40	-	-	-	-	-	-	-
	289813	Inser GC	MAG_2480 /2490	<i>rmuC</i> ./.	DNA recombination protein/HP		-	-	-	-	-	-	-	-	94	-	-	-	-
	297503	Del A	MAG_2540 /2550	./.	vpma-like lipoprotein/HP		-	-	100	-	-	-	-	-	-	-	-	-	-
	311997	Inser G	MAG_2680	.	HP		-	-	-	46	-	-	-	-	-	-	-	-	-
	320985	SNP C->A	MAG_2730 /2740	<i>secA</i> ./.	Preprotein translocase subunit/Alcohol dehydrogenase		37	-	-	-	-	-	-	-	-	-	-	-	-
	342374	SNP GAA (E) -> AAA (K)	MAG_2870	.	HP	NS	-	-	100	-	-	-	-	-	-	-	-	-	-
	352686	Del C	MAG_2950	.	Lipoprotein	9 C	-	-	49	-	-	-	-	-	-	-	-	-	-
	363277	SNP GGC (G) -> GGG (G)	MAG_3020	<i>asnS</i>	Asparagine-tRNA ligase	S	-	-	-	-	100	100	-	-	-	-	-	-	-
	367923	Inser C	MAG_3080	<i>efP</i>	Elongation factor P	Stop	-	-	-	-	-	-	-	28	-	-	-	-	-
	440268	Del A	MAG_3720	.	Asparaginyl-tRNA synthetase	7A Stop	-	-	-	-	-	-	-	-	-	-	-	26	-
	440496	Del A	MAG_3720	.	Asparaginyl-tRNA synthetase	7A Stop	-	-	-	-	-	-	97	-	-	-	-	-	-
	491621	SNP G->A	MAG_4200	.	Pseudogene		-	-	-	-	-	-	100	-	-	-	-	-	-
	506809	Inser G	MAG_4340	<i>adhT</i>	Zinc-dependent alcohol dehydrogenase	Stop	36	-	-	-	-	-	-	-	-	-	-	-	-
	559301	SNP GTT (V) -> GGT (G)	MAG_4780	.	ATP binding protein	NS	-	-	-	-	-	-	-	-	-	-	-	-	40
	559766	Inser C	MAG_4790	.	HP	Stop	-	-	-	-	-	-	-	-	100	-	-	-	-
	571958	Inser C	MAG_4900	<i>LDH</i>	L-lactate dehydrogenase	Stop	-	-	-	-	-	-	-	100	100	-	-	-	-

Table S2 : Continued (3/4).

	572337	SNP C->G	MAG_4900 /4910	<i>LDH/.</i>	L-lactate dehydrogenase/HP		35	-	-	-	-	-	-	-	-	-	-	-	-
	599476	Del T	MAG_5120	<i>deoA</i>	Pyrimidine-nucleoside phosphorylase		-	-	-	-	-	-	32	-	-	-	-	-	-
	599479	Inser C/SNP A->C (H->P)	MAG_5120	<i>deoA</i>	Pyrimidine-nucleoside phosphorylase		-	-	-	-	-	97	-	-	-	-	-	-	-
	613616	Inser G	MAG_r03	23S	23S ribosomal RNA		61	-	-	-	-	-	-	-	-	-	-	-	-
	615451	SNP C->A	MAG_r04	16S	16S ribosomal RNA		-	-	44	-	-	-	-	-	-	-	-	-	-
	620225	SNP T->A	MAG_r06	16S	16S ribosomal RNA		-	-	35	-	-	-	-	-	-	-	-	-	-
GyrA	647986*	SNP GAA (E) -> AAA (K)	MAG_5630	<i>gyrA</i> (codon 87)	DNA gyrase subunit A	NS	-	-	-	-	100	100	100	100	-	-	-	-	-
	647997*	SNP TCA (S) -> TTA (L)	MAG_5630	<i>gyrA</i> (codon 83)	DNA gyrase subunit A	NS	-	100	100	-	-	-	-	-	-	100	100	100	100
spma	707306	Del C	MAG_6060 /6070	<i>./.</i>	HP/HP	9 C	63	99	100	-	-	-	-	-	-	-	-	99	99
	711627	Del C	MAG_6090 /6100	<i>./.</i>	Lipoprotein/HP	10 C	-	-	-	93	94	94	95	91	-	-	94	96	95
	728393	SNP G->A	MAG_6120 /6130	<i>rpoB/.</i>	RNA polymerase subunit beta/Lipoprotein		-	-	36	-	-	-	-	-	-	-	-	-	-
	744149	SNP CCG (P) -> CCT (P)	MAG_6280	<i>fmt</i>	Methionyl-tRNA formyltransferase	S	-	-	-	-	-	-	-	-	100	100	100	-	-
	764266	Del G	MAG_6460	<i>.</i>	Lipoprotein	8 G	37	-	-	-	77	-	-	98	-	-	-	99	98
	771582	Inser GC	MAG_6540	<i>gtsB</i>	Glycerol transporter subunit B		-	-	-	-	44	-	-	-	-	-	-	-	-
	796940	SNP GAA (E) -> AAA (K)	MAG_6810	<i>.</i>	Ribonuclease J	NS	-	-	35	-	-	-	-	-	-	-	-	-	-
vpma	821978	SNP GAT (D) -> AAT (N)	MAG_7060	<i>vpmaW</i>	VSL W	NS	-	-	-	-	-	-	-	-	96	-	-	-	-
	823436	Del T	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y	14 T	-	-	-	-	87	89	85	83	-	-	-	-	-
	823627	SNP A->T	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		-	-	-	-	-	20	43	-	-	-	-	-	-
	823628	SNP T->G	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		-	-	-	-	-	20	44	-	-	-	-	-	-
	823632	SNP G->A	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		-	-	-	-	-	19	42	-	-	-	-	-	-
	823646	SNP A->G	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		-	-	-	-	-	26	37	-	-	-	-	-	-
	823652	SNP T->C	MAG_7070 /7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		-	-	-	-	-	23	33	-	-	-	-	-	-

Table S2 : Continued (4/4).

	825056	SNP A->T	MAG_7080 /7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U	-	-	-	-	-	17	-	-	-	-	-	-	-	-
	825057	SNP T->G	MAG_7080 /7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U	-	-	-	-	-	17	-	-	-	-	-	-	-	-
	825061	SNP G->A	MAG_7080 /7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U	-	-	-	-	-	17	-	-	-	-	-	-	-	-
vpma	827451	SNP G->A	MAG_7100 /7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase	-	-	-	-	-	24	37	-	-	-	-	-	-	-
	827457	SNP C->T	MAG_7100 /7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase	-	-	-	-	-	27	41	-	-	-	-	-	-	-
	827471	SNP T->C	MAG_7100 /7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase	-	-	-	-	-	35	49	-	-	-	-	-	-	-
	827475	SNP C->A	MAG_7100 /7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase	-	-	-	-	-	37	51	-	-	-	-	-	-	-
	827476	SNP A->T	MAG_7100 /7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase	-	-	-	-	-	37	51	-	-	-	-	-	-	-
GyrB	854621	SNP ACA (T) -> ATA (I)	MAG_7370	<i>gyrB</i> (codon 29)	Type IIA DNA topoisomerase subunit B	NS	-	-	-	100	-	-	-	-	-	-	-	-	-
	855389	SNP GGC (G) -> GAC (D)	MAG_7370	<i>gyrB</i> (codon 278)	Type IIA DNA topoisomerase subunit B	NS	-	-	-	-	-	-	-	100	-	-	-	-	-

(a) Nucleotide positions according to *M. agalactiae* PG2 55-5 genome (NC_009497.1). Fixed mutations in black, unfixed in grey. Enro target genes and variable loci underlined in orange and blue, respectively. *: mutations in target genes also identified by Sanger sequencing (see Table 1).

(b) For SNPs, the codon and the amino acid impacted by the mutation are indicated. "Del" corresponds to nucleotide deletion and "Insert" to nucleotide insertion.

(c) Gene names according to PG2 55-5 genome (NC_009497.1), dot: no gene name defined.

(d) HP: hypothetical protein ; VSL: Variable Surface Lipoprotein

(e) In brackets, Enrofloxacin MIC is indicated in $\mu\text{g.mL}^{-1}$. For fixed mutations, synonymous SNPs are in violet, non-synonymous SNPs in pink and indels in green; unfixed mutations are in grey. The percentage of each mutation is indicated (AD/DP ratio). -: Identical to the reference strain PG2 55-5.

(f) S: synonymous ; NS: non-synonymous ; stop : spot codon.

Table S3 : List of acquired and lost genes in transconjugants.

Locus tag	Product	T5-2	T5-1	T5-3	T6-5	T5-5	T6-4/8/9	T5-4	T6-1	T6-10	T6-3	T6-7
Acquired genes (PG2)												
Total number (mean 6 ±4)		7	4	12	11	10	9	3	7	2	3	2
MAG_0630	HP						■					
MAG_0840	HP						■					
MAG_1080	Pseudogene		■									■
MAG_1090	HP											■
MAG_1130	Pseudogene		■									
MAG_1330	Conserved HP					■	■					
MAG_1350	HP					■	■					
MAG_1550	Lipoprotein			■	■	■						
MAG_1560	Lipoprotein			■	■	■						
MAG_1590	HP			■	■	■						
MAG_1600	HP			■	■	■						
MAG_1610	HP			■	■	■						
MAG_1620	HP			■	■	■						
MAG_1710	HP	■		■	■	■			■			
MAG_1720	HP	■	■	■	■	■			■			
MAG_1770	HP	■		■	■	■		■	■			
MAG_1910	HP			■	■	■		■				
MAG_2180	HP			■	■	■		■				
MAG_3030	HP			■	■	■		■				
MAG_3250	HP	■		■	■	■		■				
MAG_3400	Conserved HP			■	■	■					■	
MAG_3410	Transposase			■	■	■					■	
MAG_4050	HP			■	■	■						
MAG_4060	HP			■	■	■						
MAG_4210	HP			■	■	■						
MAG_4240	Lipoprotein			■	■	■						
MAG_4960	HP			■	■	■	■					
MAG_5490	HP			■	■	■	■	■	■	■	■	
MAG_5670	HP			■	■	■	■	■	■	■	■	
MAG_5710	HP			■	■	■	■	■	■	■	■	
MAG_5880	HP			■	■	■	■	■	■	■	■	
MAG_6030	HP			■	■	■	■	■	■	■	■	

Table S3 : Continued (2/4).

Locus tag	Product	T5-2	T5-1	T5-3	T6-5	T5-5	T6-4/8/9	T5-4	T6-1	T6-10	T6-3	T6-7
Lost genes (5632)												
Total number (mean 13 ±11)		27	2	7	2	28	9	4	24	3	15	26
MAGa0830	HP						■					
MAGa0900	HP						■					
MAGa0930	HP						■					
MAGa1020	HP						■					
MAGa1570	Type III RM			■		■						
MAGa1590	IS			■		■						
MAGa1620	Conserved HP			■	■							
MAGa1850	HP		■		■	■	■					
MAGa2160	Transposase	■						■				
MAGa2180	Pseudogene	■						■				
MAGa2200	ATPase											
MAGa2210	Conserved HP	■						■				
MAGa2980	ICE											■
MAGa2990	ICE											■
MAGa3000	ICE											■
MAGa3010	ICE											■
MAGa3020	ICE											■
MAGa3030	ICE											■
MAGa3040	ICE											■
MAGa3050	ICE											■
MAGa3060	ICE											■
MAGa3070	ICE											■
MAGa3080	ICE											■
MAGa3090	ICE											■
MAGa3100	ICE											■
MAGa3110	ICE											■
MAGa3120	ICE											■
MAGa3130	ICE											■
MAGa3140	ICE											■
MAGa3150	ICE											■
MAGa3160	ICE											■
MAGa3170	ICE											■
MAGa3180	ICE											■
MAGa3190	ICE											■
MAGa3200	ICE											■
MAGa3210	ICE											■
MAGa3220	ICE											■
MAGa3460	HP											■

Table S3 : Continued (3/4).

Locus tag	Product	T5-2	T5-1	T5-3	T6-5	T5-5	T6-4/8/9	T5-4	T6-1	T6-10	T6-3	T6-7
MAGa3730	HP											
MAGa3740	HP											
MAGa3750	Pseudogene											
MAGa3950	Methyltransferase											
MAGa3960	Transposase											
MAGa3970	Deoxyribonuclease											
MAGa4000	HP											
MAGa4010	HP											
MAGa4020	HP											
MAGa4030	Transposase											
MAGa4040	HP											
MAGa4050	HP											
MAGa4060	HP											
MAGa4820	Transposase											
MAGa4850	ICE											
MAGa4860	ICE											
MAGa4870	ICE											
MAGa4880	ICE											
MAGa4890	ICE											
MAGa4900	ICE											
MAGa4910	ICE											
MAGa4920	ICE											
MAGa4930	ICE											
MAGa4940	ICE											
MAGa4950	ICE											
MAGa4960	ICE											
MAGa4970	ICE											
MAGa4980	ICE											
MAGa4990	ICE											
MAGa5000	ICE											
MAGa5010	ICE											
MAGa5020	ICE											
MAGa5030	ICE											
MAGa5040	ICE											
MAGa5050	ICE											
MAGa5060	ICE											
MAGa5490	HP											
MAGa5550	HP											
MAGa5640	IS											
MAGa6370	HP											
MAGa6510	HP											

Table S3 : Continued (4/4).

Locus tag	Product	T5-2	T5-1	T5-3	T6-5	T5-5	T6-4/8/9	T5-4	T6-1	T6-10	T6-3	T6-7
MAGa6520	HP											
MAGa6370	HP											
MAGa6880	ICE											
MAGa6890	ICE											
MAGa6900	ICE											
MAGa6910	ICE											
MAGa6920	ICE											
MAGa6930	ICE											
MAGa6940	ICE											
MAGa6950	ICE											
MAGa6960	ICE											
MAGa6970	ICE											
MAGa6980	ICE											
MAGa6990	ICE											
MAGa7000	ICE											
MAGa7010	ICE											
MAGa7020	ICE											
MAGa7030	ICE											
MAGa7040	ICE											
MAGa7050	ICE											
MAGa7060	ICE											
MAGa7070	ICE											
MAGa7080	ICE											
MAGa7090	ICE											
MAGa7100	ICE											
MAGa8410	HP											

In transconjugant genomes, replacement of 5632 sequences by PG2 fragments may result in PG2-specific gene gain (orange) and/or 5632 specific genes loss (blue). The total number of gain or loss gene is indicated for each transconjugants. Genes are indicated by their Genbank locus tag (HP: Hypothetical Protein, RM: Restriction Modification, ICE: Integrative and Conjugative Element, IS: Insertion Sequence).

Table S4 : Changes in mutation pattern of Enro^R mutants and transconjugants during multiplication without selection pressure.

Region	Nucleotide position ^(a)	Mutation ^(b)	Locus tag	Gene name ^(c)	Product ^(d)	Homo polymeric tract	Impact of SNP ^(f)	Passages ^(e)		
								P0	P10	P40
MF33-1-1										
	102994	SNP GAA (E) -> AAA (K)	MAG_0900	<i>lgt</i>	Prolipoprotein diacylglycerol transferase		NS	-	20	-
	110319	Inser GT	MAG_0970/0980	<i>rpmB/.</i>	Ribosomal protein 50S L28/Glutamine amidotransferase			-	20	-
	129186	SNP GTT (V) -> ATT (I)	MAG_1150	.	16S rRNA maturation RNase YbeY		NS	100	100	100
ParE	213322	SNP CCG (P) -> TCG (S)	MAG_1820	<i>parE</i> (codon 439)	DNA topoisomerase IV subunit B		NS	100	100	100
ParC	214190	SNP ACT (T) -> ATT (I)	MAG_1830	<i>parC</i> (codon 80)	DNA topoisomerase IV subunit A		NS	100	100	100
	284389	Del CT	MAG_2440/2450	<i>./.</i>	PTS sugar transporter subunit IIB/Lipoprotein	10 CT		-	23	-
	287762	Del A	MAG_2470	.	Proton-glutamate symporter			-	42	97
	287814	Inser G	MAG_2470	.	Proton-glutamate symporter			-	21	-
	352686	Del C	MAG_2950	.	Lipoprotein	9 C		-	48	96
	363277	SNP GGC (G) -> GGG (G)	MAG_3020	<i>asnS</i>	Asparagine-tRNA ligase		S	100	100	100
	567317	Del A	MAG_4870	.	Pseudogene			-	19	-
	588738	Inser GA	MAG_5030/5040	<i>./.</i>	P80 lipoprotein/HP			-	42	97
	634069	SNP C -> T	MAG_5470/7520	<i>rpsJ/.</i>	30S ribosomal protein S10/pseudo			-	-	61
GyrA	647986	SNP GAA (E) -> AAA (K)	MAG_5630	<i>gyrA</i> (codon 87)	DNA gyrase subunit A		NS	100	100	100
	652671	Inser C	MAG_5650	<i>hsdM</i>	Type I restriction-modification system subunit M			-	20	-
	711627	Del C	MAG_6090/6100	<i>./.</i>	Lipoproteine/HP	10 C		94	95	87
vpma	823436	Del T	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y	14 T		89	-	-
	823627	SNP A -> T	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y			20	-	-
	823628	SNP T -> G	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y			20	-	-
	823632	SNP G -> A	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y			19	-	-

Table S4 : Continued (2/3).

vpma	823646	SNP A -> G	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		26	25	52
	823652	SNP T -> C	MAG_7070/7080	<i>vpmaX/vpmaY</i>	VSL A/VSL Y		23	22	48
	825056	SNP A -> T	MAG_7080/7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U		17	38	78
	825057	SNP T -> G	MAG_7080/7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U		17	38	77
	825061	SNP G -> A	MAG_7080/7090	<i>vpmaY/vpmaU</i>	VSL Y/VSL U		17	37	75
	827451	SNP G -> A	MAG_7100/7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase		24	22	41
	827457	SNP C -> T	MAG_7100/7110	<i>vpmaZ/xer1</i>	Variable surface lipoprotein D/integrase-recombinase		27	24	46
	827471	SNP T -> C	MAG_7100/7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase		35	32	60
	827475	SNP C -> A	MAG_7100/7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase		37	33	63
	827476	SNP A -> T	MAG_7100/7110	<i>vpmaZ/xer1</i>	VSL D/integrase-recombinase		37	34	63
	852908	SNP GAT (D) -> ATA (I)	MAG_7360	.	HP		-	-	32
PG2^{E10}									
ParE	212329	SNP CAT (H) -> TAT (Y)	MAG_1820	<i>parE</i> (codon 112)	DNA topoisomerase IV subunit B	NS	92	30	-
	213174	Inser GCG	MAG_1820	<i>parE</i> (codon 390)	DNA topoisomerase IV subunit B	Inser Ala	-	66	93
ParC	214190	SNP ACT (T) -> ATT (I)	MAG_1830	<i>parC</i> (codon 80)	DNA topoisomerase IV subunit A	NS	100	100	100
	214201	SNP GAT (D) -> AAT (N)	MAG_1830	<i>parC</i> (codon 84)	DNA topoisomerase IV subunit A	NS	60	27	-
	223495	Del G	MAG_1890	.	HP	9G	-	-	64
	329667	SNP GAT (D) -> AAT (N)	MAG_2820	<i>ligA</i>	DNA ligase	NS	-	-	23
	443390	SNP GCT (A) -> GTT (V)	MAG_3750	<i>ktrA</i>	Potassium uptake protein	NS	-	-	92
	559301	SNP GTT (V) -> GGT (G)	MAG_4780	.	ATP binding protein	NS	40	-	-
	634069	SNP C -> T	MAG_5470/7520	<i>rpsJ</i> ./.	30S ribosomal protein S10/pseudo		-	-	75
GyrA	647997	SNP TCA (S) -> TTA (L)	MAG_5630	<i>gyrA</i> (codon 83)	DNA gyrase subunit A	NS	100	100	100
	703542	Del C	MAG_6030/6040	./.	HP/HP	10 C	-	71	91
spma	707306	Del C	MAG_6060/6070	./.	HP/HP	9 C	67	34	-
	711627	Del C	MAG_6090/6100	./.	HP/ Lipoprotein	10 C	97	94	94
	764266	Del G	MAG_6460	.	Lipoprotein	8 G	74	60	98

Table S4 : Continued (3/3).

Region	Nucleotide position ^(a)	Mutation ^(b)	Locus tag	Gene name ^(c)	Product ^(d)	Homo polymeric tract	Impact of SNP ^(f)	Passages ^(e)		
								P0	P10	P40
T5-1										
	187929	Del G	MAGa1620	.	HP	10 G		-	-	45
	214282	SNP GAT (D) -> AAT (N)	MAG_1830	<i>parC</i> (codon 111)	DNA topoisomerase IV subunit A		NS	-	-	44
	317265	SNP TCT (S) -> TTT (F)	MAGa2800	<i>glyS</i>	Glycine-tRNA synthetase		NS	-	18	-
	371751	SNP TTA (L) -> ATA (I)	MAGa3230	.	HP	6 T	NS	-	18	-
	391417	SNP A -> G	MAGa3370/3360	./.	HP/HP			-	-	35
T5-5										
	155280	Del C	MAGa1390/1400	./.	HP/HP	8 C		-	-	45
	245710	Del G	MAGa2120	<i>pknB</i>	Serine-threonine kinase			-	32	57
	320439	Inser T	MAGa2810	.	HP	7 T		-	18	40
	370917	SNP T -> C	MAGa3220/3230	CDS22/.	./HP			-	-	73
	430926	Del C	MAGa3730	.	HP	10 C		-	11	56
	489503	SNP CCC (P) -> CTC (L)	MAG_4180	<i>pnp</i>	Purine nucleoside phosphorylase		NS	-	-	45
	799514	SNP TCG (S) -> TTG (L)	MAGa6830	.	HP		NS	-	-	45
	891115	SNP AGG (R) -> AUG (M)	MAGa7530	.	HP		NS	-	14	35

(a) Nucleotide positions according to *M. agalactiae* PG2 55-5 genome or 5632 genome for 5632 fragments of transconjugants. Fixed mutations are indicated in black, unfixed in grey.

Enro target genes and variable locus are underlined in orange and blue, respectively.

(b) For SNPs, the codon and the amino acid impacted by the mutation are indicated. "Del" corresponds to nucleotide deletion and "Inser" to nucleotide insertion.

(c) Gene names according to PG2 55-5 (MAG) or 5632 (MAGa) genome, dot: no gene name defined.

(d) HP: hypothetical protein ; VSL: Variable Surface Lipoprotein

(e) P0, P10, P40 referred to passages 0, 10 and 40 respectively. For fixed mutations, synonymous SNPs are in violet, non-synonymous SNPs in pink and indels in green; unfixed mutations are in grey. The percentage of each mutation is indicated (AD/DP ratio). -: Identical to the reference strain PG2 55-5 or 5632.

(f) S: synonymous ; NS: non-synonymous ; stop : spot codon.

Table S5 : Variations in enrofloxacin target enzyme between PG2 and 5632 strains.

Alignment Needle PG2/5632	GyrA	GyrB	ParC	ParE
gene size (protein size)	2760 nt (919 aa)	1968 nt (655 aa)	2586 nt (861 aa)	1926 nt (641 aa)
% gene identity (SNP)	97.4% (72 SNPs)	95.2% (94 SNPs)	94.5% (143 SNPs)	95.8% (81 SNP)
% protein identity (amino acid)	99.2% (7 aa)	98.6% (9 aa)	97.2% (24 aa)	99.5% (3 aa)
% protein similarity (amino acid)	99.3% (6 aa)	100%	98.8% (10 aa)	99.7% (2 aa)

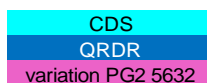
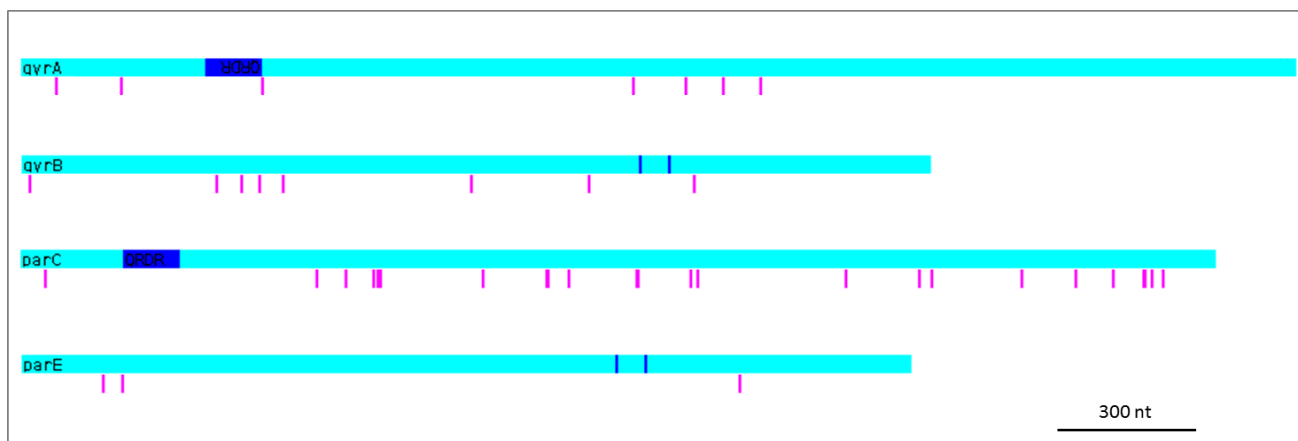


Table S6 : European Nucleotide Archive (EMBL-EBI) accession numbers.

Study	sample_alias	tax_id	scientific_name	sample_description	file_name	MD5	file_name	MD5	Sample_primary_accession	Sample_secondary_accession	Experiment_accession	Run_accession
PRJEB28036	MF26	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF26_1.fastq.gz	45eb00d4df70cf299393593618e500f9	MF26_2.fastq.gz	a8890d2282c792c96c04869163fd0ad0	ERS2632133	SAMEA4812231	ERX2735844	ERR2722073
PRJEB28036	MF26-1	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF26-1_1.fastq.gz	9573e1ca2bd4f36f03ad9ea6cb5f07ec	MF26-1_2.fastq.gz	2beb2ca6e0108ddf066c865313568b94	ERS2632134	SAMEA4812232	ERX2735845	ERR2722074
PRJEB28036	MF26-1-2	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF26-1-2_1.fastq.gz	cb69d3e5412dda2105a917880d2e79c0	MF26-1-2_2.fastq.gz	0d6352ae7a14f68c26ce3fae5c201314	ERS2632135	SAMEA4812233	ERX2735846	ERR2722075
PRJEB28036	MF30	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF30_1.fastq.gz	86c488a18f531687b20506a5004b7c94	MF30_2.fastq.gz	017b1572b7d3b33e2c571ba590f08853	ERS2632136	SAMEA4812234	ERX2735847	ERR2722076
PRJEB28036	MF30-1	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF30-1_1.fastq.gz	d2602cbacc48b621056fc6d9bb64bbd4	MF30-1_2.fastq.gz	0e4f44d7ad9dad579e7b3b61eafa164f	ERS2632137	SAMEA4812235	ERX2735848	ERR2722077
PRJEB28036	MF30-1-4	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF30-1-4_1.fastq.gz	b26a4c572ede4d00b5277ba696723bb7	MF30-1-4_2.fastq.gz	f67e4c4297c600d8fe87e39aef40cca2	ERS2632138	SAMEA4812236	ERX2735849	ERR2722078
PRJEB28036	MF30-1-4-1	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF30-1-4-1_1.fastq.gz	9d10a9e0dbf3b34d73d7e2c5c744cd8f	MF30-1-4-1_2.fastq.gz	eb2092434a4f27bc619e727bd6f51b5e	ERS2632139	SAMEA4812237	ERX2735850	ERR2722079
PRJEB28036	MF30-1-4-8	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF30-1-4-8_1.fastq.gz	63a74630fb7b7e4ad8ab025792f50c7b	MF30-1-4-8_2.fastq.gz	3e364ae9a8e2280277b72d93dce51d45	ERS2632140	SAMEA4812238	ERX2735851	ERR2722080
PRJEB28036	MF33	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF33_1.fastq.gz	7c58857b592f7ef71819f02d86bf2484	MF33_2.fastq.gz	892ce1caaa66ef8717aca8d258b6c5cc	ERS2632141	SAMEA4812239	ERX2735852	ERR2722081
PRJEB28036	MF33-1	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF33-1_1.fastq.gz	e499a8ae440e0f90260068c477f2c9ed	MF33-1_2.fastq.gz	7c69adaa3457555c291f3e70b48707f4	ERS2632142	SAMEA4812240	ERX2735853	ERR2722082
PRJEB28036	MF33-1-1	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF33-1-1_1.fastq.gz	b7c20094ca6a194b8b7fc2b564998987	MF33-1-1_2.fastq.gz	398cb1a56c6fdab7f80e6ac4ccd77a99	ERS2632143	SAMEA4812241	ERX2735854	ERR2722083

Table S6 : Continued (2/3).

PRJEB28036	MF33-1-2	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF33-1-2_1.fastq.gz	fd634ca5e47694f754854fe4ceeb18a3	MF33-1-2_2.fastq.gz	2d592477f579d9916eec678c25acb7fd	ERS2632144	SAMEA4812242	ERX2735855	ERR2722084
PRJEB28036	MF33-1-3	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	MF33-1-3_1.fastq.gz	9415de6cfc0ba716004da1e75de49674	MF33-1-3_2.fastq.gz	0d51562eeac1f33b4499caf990d5d4fb	ERS2632145	SAMEA4812243	ERX2735856	ERR2722085
PRJEB28036	MF33-1-1_P10	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant passages	MF33-1-1_P10_1.fastq.gz	6f6184e6c97a30cda f366ae44e03dd4b	MF33-1-1_P10_2.fastq.gz	6af8e46696bc41d65408ac3ee167b46f	ERS2632146	SAMEA4812244	ERX2735857	ERR2722086
PRJEB28036	MF33-1-1_P40	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant passages	MF33-1-1_P40_1.fastq.gz	c8b580bafec03bb b364e33980dd711f	MF33-1-1_P40_2.fastq.gz	aed1fa4c2790fe37fc8b33211cec3b8c	ERS2632147	SAMEA4812245	ERX2735858	ERR2722087
PRJEB28036	PG2E10	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant	PG2E10_1.fastq.gz	d22307c7e60cb401a b75a6699b634d9a	PG2E10_2.fastq.gz	2259cf4a9e0b92c4a0c08ca81a3445ad	ERS2632148	SAMEA4812246	ERX2735859	ERR2722088
PRJEB28036	PG2E10_P10	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant passages	PG2E10_P10_1.fastq.gz	c843e2cded9ac20e34733757bdd12911	PG2E10_P10_2.fastq.gz	3500c03523c0015cff205778bb658c9d	ERS2632149	SAMEA4812247	ERX2735860	ERR2722089
PRJEB28036	PG2E10_P40	2110	Mycoplasma agalactiae	PG2 55-5 enrofloxacin mutant passages	PG2E10_P40_1.fastq.gz	b8e3f27471a07dd64c215ad4f1793538	PG2E10_P40_2.fastq.gz	97b07a43a139a2ef21e428880d5f1491	ERS2632150	SAMEA4812248	ERX2735861	ERR2722090
PRJEB28036	PG2 55-5	2110	Mycoplasma agalactiae	PG2 55-5 parental strain	PG2 55-5_1.fastq.gz	7f59a746f7a57a3a76fbbb05f989b7e0	PG2 55-5_2.fastq.gz	21daa7f4470750328e3e6263410402a3	ERS2632151	SAMEA4812249	ERX2735862	ERR2722091
PRJEB27571	T5-1	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T5-1_1.fastq.gz	c0a6ed23989d4e787ab01bbd258a49cc	T5-1_2.fastq.gz	80a15cd2bd5f14f4cb1464369a5d42f6	ERS2631597	SAMEA4811695	ERX2735819	ERR2722048
PRJEB27571	T5-2	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T5-2_1.fastq.gz	5e10e3317f3c45ba2e751fddee76faf8	T5-2_2.fastq.gz	16a2de670599f05816f558b96b405cd8	ERS2631598	SAMEA4811696	ERX2735820	ERR2722049
PRJEB27571	T5-3	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T5-3_1.fastq.gz	9708fc177d50b0902ee1d5c48426fedd	T5-3_2.fastq.gz	b90bd6762df6aef7713c83987a679aa4	ERS2631599	SAMEA4811697	ERX2735821	ERR2722050
PRJEB27571	T5-4	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T5-4_1.fastq.gz	af92dc577d49c543d3ae7043d7603565	T5-4_2.fastq.gz	c55d91e68614b7d9101823e92feeb3a7	ERS2631600	SAMEA4811698	ERX2735822	ERR2722051
PRJEB27571	T5-5	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T5-5_1.fastq.gz	9139360774c0e24c8ae01989283a73e7	T5-5_2.fastq.gz	0667e7841352f0978a557d568030402b	ERS2631601	SAMEA4811699	ERX2735823	ERR2722052

Table S6 : Continued (3/3).

PRJEB27571	T6-1	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-1_1.fastq.gz	6b3a92adc9d17bd787653d20d4db80c6	T6-1_2.fastq.gz	4c9c3ca2a6f53a404574c43a012b49ee	ERS2631602	SAMEA4811700	ERX2735824	ERR2722053
PRJEB27571	T6-3	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-3_1.fastq.gz	417e70b49ac45622424be611748ad4be	T6-3_2.fastq.gz	8ab6658bfc10a85c80d535d53f19820b	ERS2631603	SAMEA4811701	ERX2735825	ERR2722054
PRJEB27571	T6-4	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-4_1.fastq.gz	76785aef29f98917e92b5173840f310	T6-4_2.fastq.gz	bb62e1b68900fa8d79b5385c5a5df1c9	ERS2631604	SAMEA4811702	ERX2735826	ERR2722055
PRJEB27571	T6-5	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-5_1.fastq.gz	784e03f58c3ef1b3cd92ffdfb9bfafb1	T6-5_2.fastq.gz	5c13bb61e0c7b2024e5b3110e8fda64c	ERS2631605	SAMEA4811703	ERX2735827	ERR2722056
PRJEB27571	T6-7	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-7_1.fastq.gz	595a24854581125cee27ab01d1888218	T6-7_2.fastq.gz	6bf3ecb139b10a2e269b795a6a6655d7	ERS2631606	SAMEA4811704	ERX2735828	ERR2722057
PRJEB27571	T6-8	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-8_1.fastq.gz	7aee484b5fa14a569eac2e1a885a4251	T6-8_2.fastq.gz	ba14d3bd98c2e515243ed6a5b07a427b	ERS2631607	SAMEA4811705	ERX2735829	ERR2722058
PRJEB27571	T6-9	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-9_1.fastq.gz	64f45f9ab3d1f8dd87f18e4b8db63497	T6-9_2.fastq.gz	832c33b6933f487b8b8536106bc4227a	ERS2631608	SAMEA4811706	ERX2735830	ERR2722059
PRJEB27571	T6-10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-10_1.fastq.gz	0bb337df9ef13685d09dfda92ba8cf85	T6-10_2.fastq.gz	e4f7b50d03f5eeb8b6326b9883e3a189	ERS2631609	SAMEA4811707	ERX2735831	ERR2722060
PRJEB27571	T5-1_P10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-1_P10_1.fastq.gz	6e66fcb0d024a505447070c13ad006f9	T5-1_P10_2.fastq.gz	e80810b175b3349e9a8978d51d85a2a9	ERS2631610	SAMEA4811708	ERX2735832	ERR2722061
PRJEB27571	T5-1_P40	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-1_P40_1.fastq.gz	002892dc890af4a76bcc2833c8ca65cb	T5-1_P40_2.fastq.gz	94e715777108b98874e7b79996bd73bc	ERS2631611	SAMEA4811709	ERX2735833	ERR2722062
PRJEB27571	T5-5_P10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-5_P10_1.fastq.gz	f9038a508e820dea ddcfe4a0d341b2b6	T5-5_P10_2.fastq.gz	1b34b953cc396c09c6b56030b9b01085	ERS2631612	SAMEA4811710	ERX2735834	ERR2722063
PRJEB27571	T5-5_P40	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-5_P40_1.fastq.gz	daf4951a8bb7a35b9899a6c813c5805	T5-5_P40_2.fastq.gz	b1339aeb580ef77d6089ba4430d9356f	ERS2631613	SAMEA4811711	ERX2735835	ERR2722064
PRJEB27571	5632G3	2110	Mycoplasma agalactiae	parental 5632 gentamicin resistant clone	5632G3_1.fastq.gz	d9581fd6da358dc2ab442992fa0bbd0d	5632G3_2.fastq.gz	647b559f130bad62d5f5ebb78b58ae4	ERS2631614	SAMEA4811712	ERX2735814	ERR2722037

Table S6 : Continued (3/3).

PRJEB27571	T6-1	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-1_1.fastq.gz	6b3a92adc9d17bd787653d20d4db80c6	T6-1_2.fastq.gz	4c9c3ca2a6f53a404574c43a012b49ee	ERS2631602	SAMEA4811700	ERX2735824	ERR2722053
PRJEB27571	T6-3	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-3_1.fastq.gz	417e70b49ac45622424be611748ad4be	T6-3_2.fastq.gz	8ab6658bfc10a85c80d535d53f19820b	ERS2631603	SAMEA4811701	ERX2735825	ERR2722054
PRJEB27571	T6-4	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-4_1.fastq.gz	76785aef29f98917e92b5173840f310	T6-4_2.fastq.gz	bb62e1b68900fa8d79b5385c5a5df1c9	ERS2631604	SAMEA4811702	ERX2735826	ERR2722055
PRJEB27571	T6-5	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-5_1.fastq.gz	784e03f58c3ef1b3cd92ffdfb9bfbf1	T6-5_2.fastq.gz	5c13bb61e0c7b2024e5b3110e8fda64c	ERS2631605	SAMEA4811703	ERX2735827	ERR2722056
PRJEB27571	T6-7	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-7_1.fastq.gz	595a24854581125cee27ab01d1888218	T6-7_2.fastq.gz	6bf3ecb139b10a2e269b795a6a6655d7	ERS2631606	SAMEA4811704	ERX2735828	ERR2722057
PRJEB27571	T6-8	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-8_1.fastq.gz	7aee484b5fa14a569eac2e1a885a4251	T6-8_2.fastq.gz	ba14d3bd98c2e515243ed6a5b07a427b	ERS2631607	SAMEA4811705	ERX2735829	ERR2722058
PRJEB27571	T6-9	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-9_1.fastq.gz	64f45f9ab3d1f8dd87f18e4b8db63497	T6-9_2.fastq.gz	832c33b6933f487b8b8536106bc4227a	ERS2631608	SAMEA4811706	ERX2735830	ERR2722059
PRJEB27571	T6-10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant	T6-10_1.fastq.gz	0bb337df9ef13685d09dfda92ba8cf85	T6-10_2.fastq.gz	e4f7b50d03f5eeb8b6326b9883e3a189	ERS2631609	SAMEA4811707	ERX2735831	ERR2722060
PRJEB27571	T5-1_P10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-1_P10_1.fastq.gz	6e66fcb0d024a505447070c13ad006f9	T5-1_P10_2.fastq.gz	e80810b175b3349e9a8978d51d85a2a9	ERS2631610	SAMEA4811708	ERX2735832	ERR2722061
PRJEB27571	T5-1_P40	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-1_P40_1.fastq.gz	002892dc890af4a76bcc2833c8ca65cb	T5-1_P40_2.fastq.gz	94e715777108b98874e7b79996bd73bc	ERS2631611	SAMEA4811709	ERX2735833	ERR2722062
PRJEB27571	T5-5_P10	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-5_P10_1.fastq.gz	f9038a508e820dea ddcfe4a0d341b2b6	T5-5_P10_2.fastq.gz	1b34b953cc396c09c6b56030b9b01085	ERS2631612	SAMEA4811710	ERX2735834	ERR2722063
PRJEB27571	T5-5_P40	2110	Mycoplasma agalactiae	5632G3xPG2E10 transconjugant passages	T5-5_P40_1.fastq.gz	daf4951a8bb7a35b9899a6c813c5805	T5-5_P40_2.fastq.gz	b1339aeb580ef77d6089ba4430d9356f	ERS2631613	SAMEA4811711	ERX2735835	ERR2722064
PRJEB27571	5632G3	2110	Mycoplasma agalactiae	parental 5632 gentamicin resistant clone	5632G3_1.fastq.gz	d9581fd6da358dc2ab442992fa0bbd0d	5632G3_2.fastq.gz	647b559f130bad62d5f5ebb878b58ae4	ERS2631614	SAMEA4811712	ERX2735814	ERR2722037

CHAPITRE 2

CHAPITRE 2

Horizontal chromosomal transfers and the extraordinary plasticity of the mycoplasma genome.

Le transfert chromosomique horizontal contribue à l'extraordinaire plasticité des génomes de mycoplasmes.

Emilie Dordet-Frisoni^{1#}, Marion Faucher^{1,2#}, Eveline Sagné¹, Eric Baranowski¹, Florence Tardy², Laurent Xavier Nouvel¹, and Christine Citti^{1*}.

¹ IHAP, Université de Toulouse, INRA, ENVT, Toulouse, France.

² UMR Mycoplasmoses of ruminants, ANSES, VetAgro Sup, University of Lyon, Lyon, France.

These authors contributed equally to this work.

*Corresponding author: c.citti@envt.fr

Article à soumettre à DNA Research.

Résumé

Le MCT (*Mycoplasma chromosomal transfer*) est un mécanisme de conjugaison, non conforme aux modèles classiques de transfert Hfr/*oriT*, qui permet le transfert horizontal de larges fragments d'ADN chez *Mycoplasma*, vaste groupe de bactéries dépourvues de paroi et au génome de taille réduite. Afin de mieux comprendre les mécanismes gouvernant ce processus, des transconjugants générés par *mating* ont été séquencés individuellement pour une analyse génomique approfondie. Les données obtenues montrent que le MCT est un processus performant de transfert d'ADN combinant des mécanismes de recombinaison homologue et non homologue. Les analyses des macro- et micro-événements provenant de l'incorporation d'ADN donneur dans le chromosome receveur révèlent un niveau de complexité inattendu. Le MCT ne se limite pas à l'échange horizontal d'un seul bloc d'ADN chromosomique, comme initialement suspecté, mais il implique le co-transfert de plusieurs fragments distants et non apparentés. En un seul événement, ce processus distributif peut conduire à l'échange de 17% du chromosome donneur, incluant des gènes accessoires comme des gènes de ménage. Les clones des transconjugants obtenus présentent des génomes hétérogènes hautement mosaïques. Ces résultats illustrent l'extraordinaire plasticité des génomes de mycoplasmes et mettent en lumière le rôle clef du MCT dans la diversification de génomes minimaux et l'adaptation aux changements de l'environnement.

Abstract

Mycoplasma Chromosomal Transfer (MCT) is an atypical conjugative mechanism that does not conform to classical Hfr/*oriT* based models and mediates the horizontal exchange of large DNA fragments in *Mycoplasma*, a large group of wall-less bacteria with nearly minimal genomes. To further understand the exact features underlying this process, individual mating-progenies were submitted to extensive genomic analysis at the single cell level. Data showed that MCT is an efficient process of DNA transfer combining both homologous and non-homologous recombination mechanisms. Analyses of macro- and micro-events resulting from donor-DNA incorporation into single-host chromosomes revealed an unexpected level of complexity. MCT was not limited to the horizontal exchange of a single block of chromosomal DNA, as initially thought, but involved the co-transfers of several remote, unrelated fragments. In a single event, this distributive process can drive the exchange of up to 17% of the donor-chromosome, equally including house-keeping and accessory genes. Individual mating-progenies were thus characterized by highly mosaic genomes, each being unique. These results emphasize the extraordinary plasticity of the mycoplasma genomes and point toward MCT as a key player in minimal genome diversification and adaptation to changing environments.

1. Introduction

By mixing genomes, sexual reproduction enables eukaryotic organisms to generate genetic diversity and eliminate deleterious mutations, resulting in better-fit progeny. In bacteria, one main driver of evolution and diversification is provided by horizontal gene transfer (HGT) (Soucy et al. 2015; García-Aljaro et al. 2017). This sexual, reproduction-like system is expected to produce less diversity compare to meiosis as it mainly concerns the spread of non-core, accessory genes which ones are not essential for survival and replication. Yet, HGT is key in allowing bacteria to adapt to specific niches or conditions via the exchange of genetic material which encode proteins conferring resistance to antibiotics, toxins, specific metabolic pathways, etc. (Ochman et al. 2000; Soucy et al. 2015)

During their evolution, bacteria have developed diverse strategies to perform HGT, with mechanisms that can be dependent or independent of mobile genetic elements (MGEs) (Frost et al. 2005; García-Aljaro et al. 2017). The three primary modes of HGTs are transformation, transduction and conjugation, with the latter being considered as the predominant mechanism allowing the rapid dissemination and acquisition of new, adaptive traits upon cell-to-cell contact (Frost et al. 2005; Halary et al. 2010). Conjugation traditionally required the donor bacterial cell to express (i) a 'conjugative pilus' or a multicomponent macromolecular membrane traversing structure that promotes cellular contact and maintains mating-pair formation and (ii) a conjugative machinery responsible for DNA transfer into the recipient, upon recognition of an origin of transfer (*oriT*) (Goessweiner-Mohr et al. 2014). This two-steps process is usually mediated by plasmids or integrative and conjugative elements (ICE) and mainly concerned the transfers of MGEs themselves. Yet, some conjugative mechanisms can also involve the transfer of chromosomal DNA regions as it was first described for the Hfr system of *Escherichia coli* (Wollman et al. 1956). There, Hfr- or *oriT*-chromosomal transfers are initiated from an *oriT* located in the genome, and are characterized by a gradient, with genes closer to the *oriT* being more frequently transferred and thus incorporated into the recipient chromosome (Smith 1991). Recently, we discovered a novel form of chromosomal conjugative transfer occurring in mycoplasmas that did not conform to this classic *oriT*-based model (Dordet-Frisoni et al. 2014; Faucher et al. 2018).

With over 200 species, Mycoplasmas represent of large group of bacteria that are characterized by the lack of a cell-wall and a minute-size genome (580 to 2,200 kb) (Razin and Hayflick 2010). Despite this simplicity, numerous mycoplasma species are important pathogens

of man, animals, insects and plants (Razin and Hayflick 2010). These bacteria derived from Gram positive ancestors and their evolution was long assimilated as regressive, by successive losses of genetic materials (Sirand-Pugnet et al. 2007). This paradigm was supported by the mycoplasma paucity in recombination systems, phages or conjugative elements, but has been challenged over the past years by comparative genomic studies : large DNA exchanges were evidenced in between mycoplasma species that are phylogenetically remote while sharing a same habitat (Vasconcelos et al. 2005; Sirand-Pugnet et al. 2007; Pereyre et al. 2009) and ICEs were detected in a large number of sequenced Mycoplasma genomes (Calcutt et al. 2002; Marenda et al. 2006; Dordet-Frisoni et al. 2013; Tardy et al. 2015) raising the prospect that these simple bacteria might be able to conjugate.

Indeed, our group highlighted two conjugative processes occurring among and within the *M. agalactiae* species (Ma), a ruminant pathogen and a model organism. The first one is the conventional horizontal dissemination of ICE, from ICE-positive to ICE-negative cells (Dordet-Frisoni et al. 2013). Mycoplasma ICEs (MICE) are a new family of large modular elements which encodes the machinery for its self-transmission and maintenance including conjugation, excision and integration into the chromosome of the recipient cell. MICE transfer occurs via a clean cut and paste mechanism and is not accompanied by the transfer of flanking chromosomal sequences (Dordet-Frisoni et al. 2013). The second conjugative mechanism discovered in Ma concerns the transfer of large blocks of chromosomal DNA from ICE-negative to ICE-positive cells that will be further designated MCT for Mycoplasma Chromosomal Transfer (Dordet-Frisoni et al. 2014). This atypical HGT mechanism is not directly linked to ICE- or other MGE-movement but takes advantage of the cellular contact and most likely of the conjugal pore expressed by the ICE. MCTs result in replacing large chromosomal regions of the recipient cell by the donor-counterparts via recombination (Dordet-Frisoni et al. 2014). Whole sequencing of mating-progeny populations revealed that, unlike classical *oriT*-based mechanisms, these DNA exchanges could equally affect any part of the donor chromosome (Dordet-Frisoni et al. 2014). Very recently, DNase-resistant HGT events have been reported in the human pathogen *M. genitalium* that depended on *recA* homologous recombination (Torres-Puig et al. 2018) but further investigations are needed to define whether this conjugative process is similar to that observed in *M. agalactiae*.

The emergence of new pathogens that have mosaic-like genomes (Kloesges et al. 2011; Lesic et al. 2012; Watt et al. 2018) resulted in the discovery of chromosomal conjugative transfers mechanism that also differ from the classical *oriT*- based model. For instance, in *Yersinia*

pseudotuberculosis, a chromosomal transfer more powerful and less constrain than the classical Hfr mechanism has been described (Lesic et al. 2012). This conjugative transfer requires the presence of an IS6-type element on a conjugative replicon and is able to convey the transfer of virtually any piece of chromosomal or plasmid DNA (Lesic et al. 2012). In *Thermus thermophilus*, an unconventional conjugation-like mechanism has been proposed to involved a sort of “push-pull” system termed transjugation (Blesa et al. 2017). There, the donor cell actively pushes out chromosomal or plasmid DNA and the recipient cell exploits its natural competence system to uptake the DNA. This mechanism is dependent of an ICE factor, but the nature and the diversity of the transferred chromosomal DNA remains to be explored. To our knowledge, the most investigated system of conjugative chromosomal transfer that does not conform to classical Hfr/*oriT* based model relates to that found in *Mycobacterium smegmatis* (Gray et al. 2013; Derbyshire and Gray 2014; Gray and Derbyshire 2018). This mechanism called distributive conjugal transfer (DCT) promotes the distributive transfer of multiple, non-contiguous genome segments from a donor to a recipient. As observed for MCT the exchange of donor chromosome occurs with equal efficiency regardless of their location. DCT is a chromosomally encoded process that leads to genetic mixing of two parental bacteria generating progenies with highly mosaics genomes (Gray et al. 2013).

The discovery of MTC was mainly based on the genomic analyses of mixed populations and more data from single cells were needed to progress on the understanding of this phenomenon. To address this issue, several individual mating-progeny derived from different mating-pairs or conditions were generated and sequenced. Single-genome analyses showed that MTC were much more complex than first expected and generated mosaic genomes with single progeny that could concomitantly inherit several remote, unrelated fragments, adding up to 17% of the total size of the donor genome. Data also indicated that incorporation of donor-DNA in the host chromosome relied on both homologous and non-homologous recombination mechanisms that were independent from the ICE conjugative machinery. Overall, MCT was shown to result in creating both macro- and micro-complexity, affecting house-keeping genes and accessory genes equally. We further discussed how the extraordinary plasticity of the mycoplasma minute-size genome may impact its evolution and promote adaptation to new environment.

2. Results

Overall experimental design and workflow

To analyze the features and extent of MCT, a set of individual transconjugants was first generated by experimental mating using a set of different parental pairs (Fig. 1A, Table 1). These were chosen to document and analyzed the transfer of various chromosomal loci and their associated selectable markers under different selective pressure. Mating pairs all derived from PG2 and 5632, two strains respectively known as MCT-donor and -recipient and for which the circularized genome sequence is available. More specifically, two 5632-recipient clones were chosen, 5632^G-3 and 5632^T-H3, that have the gentamicin (Gm) and the tetracycline resistance (Tet) genes stably inserted at positions nt-919899 and nt-529737 respectively (Dordet-Frisoni et al. 2014). As donors, we selected (i) three PG2-clones having the Gm stably inserted at different positions: nt-104815 for PG2^G-10, nt-473081 for PG2^G-71 and nt-648734 for PG2^G-181 (matings M1 to M3), (ii) the PG2^T-3 clone having the Tet at position nt-770055 (mating M4) and (iii) the PG2^{E10} mutant which enrofloxacin resistant phenotype, Enro^R, is due to three point mutations located in *parE* (nt-212329), *parC* (nt-214190) and *gyrA* (nt-697997) genes (matings M5 and M6). These two last matings were further referred as “high-selection pressure mating” because of the selection of the donor DNA relying on the transfer of two loci versus one in the other “classical mating” (M1-M4) (Fig. 1A). Overall, five independent mating experiments were conducted from which a total of 15 individual transconjugants were selected (Fig. 1A, Table S2), fully sequenced and analyzed.

In a second step, mycoplasma pairs used in M1 and M4 experiments (Fig. 1A) were artificially fused with PEG as previously described (Dordet-Frisoni et al. 2014) and 6 individual double-resistant clones were selected, fully sequenced and analyzed for comparative purposes.

Transconjugant and PEG-hybrid genomes were reconstructed as previously described. Briefly, reads mapping equally on both genomes were filtered out and only reads fully matching the donor reference genomes were retained for the identification of transferred-donor DNA fragments. Such a strategy was feasible because the 5632 and PG2 genomes are collinear and contained abundant single nucleotide polymorphisms (SNPs) allowing a clear distinction between parental DNA origins (SNPs averaging one per 11 to 26 nt depending on the software used, see material and methods). Of note, the largest chromosomal portion without SNP corresponds to the 909 bp encoded the two 23S rRNA genes.

Table 1 : Genetic modifications induced MCT in 5632 recipient cells

	CDSs Loss	CDSs Gain	Chimerics CDSs	Mutations	SNPs	Indels	% of non-synonymous mutations	SNPs in intergenic regions
T1-1	1	2	4	630	580	50	22	81
T1-2	35*	10	13	4092	3723	366	24	354
T2	13	29 [#]	7	2759	2469	289	27	267
T3	7	8	12	3795	3453	341	21	247
T4-1	5	5	5	1839	1702	136	25	111
T4-2	28*	5	9	2415	2230	185	25	230
T4-3	9	7	15	2007	1811	194	29	174
T5-1	2	4	20	2353	2190	163	21	158
T5-2	26*	7	23	3584	3349	235	22	251
T5-3	7	12	35	2625	2410	213	22	240
T5-4	4	3	17	2067	1916	151	20	214
T5-5	28*	10	15	3871	3511	356	25	356
T6-1	23*	7	11	3508	3280	227	23	333
T6-4	9	9	28	4028	3742	284	21	336
T6-5	2	11	16	2706	2500	205	23	266
H1-1	0	2	6	972	881	91	24	114
H1-2	4	3	13	1669	1517	150	19	151
H4-1	4	3	2	17	10	7	50	2
H4-2	4	4	7	873	783	90	32	72
H4-3	0	0	0	734	688	46	21	31
H4-4	5	4	14	2086	1936	150	22	143

* : loss of an ICE present in 5632 genome

#: gain of an degenerate vestigial form of ICE present in PG2 genome

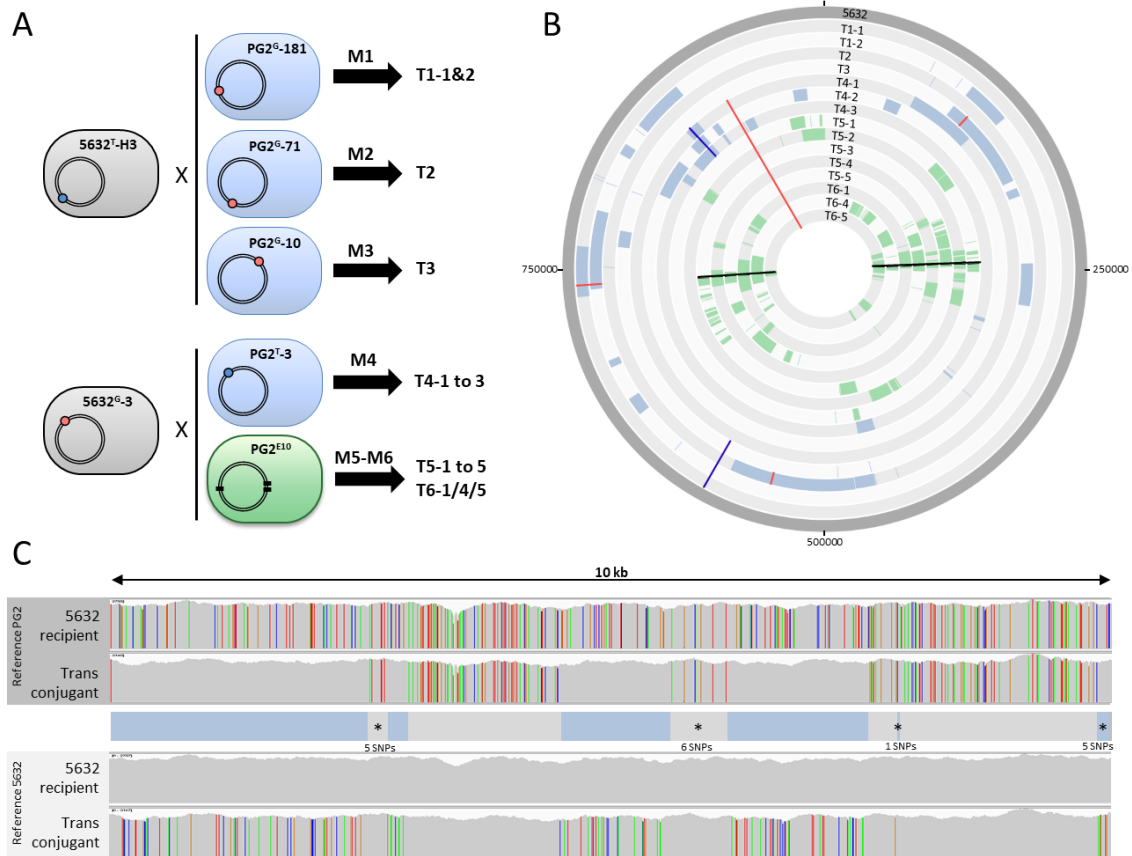


Figure 1. Mating experiments create in a single step a diversity of transconjugants with mosaic genomes

(A) Mating experiments (M1 to M6) conducted to produce mycoplasma transconjugants. The 5632 and PG2 parent strains carried the Tet (blue dot) or Gm (red dot) genetic-marker located in different loci depending on the mating. In PGE10 of mating M5 and M6, black thick lines correspond to the position of chromosomal mutations that confer enrofloxacin resistance (Enro^R); (B) DNA plotter representation depicting the mosaic nature of the 15 *M. agalactiae* transconjugants. Sequences specific of PG2, PG2E10 and 5632 are color-coded in blue, green and grey, respectively. The chromosome positions of Gm-marker, Tet-marker and Enro^R-mutations in transconjugants are indicated with red, blue and black lines respectively. From outer to inner circles: 5632, T1-1 and -2, T2, T3, T4-1 to -3, T5-1 to -5, T6-1, -4 and -5. (C) Example of micro-complexity regions observed in mycoplasma transconjugants. The 5632 recipient and the transconjugant T6-4 genomes were aligned to that of the PG2 (top) and 5632 (down) and the result obtained for a 10 kb segment is shown, as visualized with the IGV software. Each color bar represents a SNP with the 4 different colors representing the 4 different bases. The overall genotype of the T6-4 transconjugant is represented in between the alignments, with DNA segment corresponding to PG2-donor in blue and the 5632-recipient backbone in grey. Micro-complex regions are annotated by an asterisk. They correspond to small integrated PG2-fragments with ≤ 10 SNPs, or PG2 acquired segments interrupted by a small region of 5632 recipient cell of ≤ 10 SNPs. The number of SNPs in these micro-complex regions is indicated.

MCT creates genome mosaicism via the co-transfer of multiple, unrelated chromosomal fragments

Whole genome sequence analyses (WGS) of mating progenies first confirmed that each transconjugant displayed a 5632-backbone in which chromosomal regions were replaced by their PG2-donor counterparts (Fig. 1B). This exchange was evidenced by the concomitant loss of 5632 recipient-specific SNPs (Fig. 1C), the detection of PG2 donor-specific SNPs and the acquisition of the PG2-selectable marker. In very rare occasions, sequencing reads matched both the donor and the recipient SNPs almost equally: their mapping corresponded to regions containing repetitive elements such as the *vpma* multi-gene family (Nouvel et al. 2009) which precise organization and localization would require *de novo* assembly using long-read sequencing technology.

Comparative sequence analyses of the 15 individual transconjugants yield unexpected data. In a single mating event, 3 to 30 distinct chromosomal regions of the recipient-genome were replaced by donor sequences, with 85% of the exchanged fragments having no selectable marker (Fig. 1B). These PG2-chromosomal segments were co-inherited without obvious regional bias except around the donor-selective marker where a higher proportion of donor DNA fragments were detected (see below). Overall, these transfer events resulted in a variety of mosaic genomes as shown in Figure 1B, with no transconjugant displaying the same blend of the parental genomes. Of note, no PG2-donor region was found in all transconjugants.

The size of the inherited PG2-fragments varied from 49 bp to 64.7 kb, with a 6.7 kb average and a mean of 14 chromosomal fragments acquired per genome (Fig. 2, Table S4). Regarding the total amount of transferred-DNA, mating progenies contained from 29.4 to 147.9 kb of PG2-donor sequences, representing 3.3% to 14.8% of the transconjugant genomes. The number of exchanged CDSs ranged from 20 to 163, with both house-keeping (43%) and accessory genes (57%) being almost equally affected by MCT (Table S5). These findings supported our previous observation that all PG2 chromosomal regions are potentially transferable.

Despite the high synteny existing in between the 5632 and PG2 genomes, some transconjugants gained up to 29 CDS (3% of PG2 donor genome) or lost up to 35 CDS (3.8% of the 5632 recipient genome) (Table S5). In particular, in five transconjugants, a large proportion of gene losses was due to DNA recombination occurring on each side of one ICEA copy that resulted in its deletion (ICEA-III in T5-5 and ICEA-I in T1-2, T4-2, T5-2, and T6-1). ICEA is a 27-kb

integrative conjugative element that encodes 23 CDSs and occurs as three almost identical copies in 5632 while PG2 has none but a vestige. Sequence analyses showed that other CDSs present on both sides of the ICEA deletions were also exchanged (7 CDSs for T1-2, 98 CDSs for transconjugant T4-2, 6 CDSs for T5-2, 29 CDSs for T5-5 and 2 CDSs for T6-1) (Table S5) confirming that ICEA losses were not due to the ICE self-excision (Dordet-Frisoni et al. 2013), but to DNA replacement by with PG2-synthetic regions having no ICE. In another transconjugant, T2, the PG2-donor fragment that was acquired included the PG2-ICE vestige, allowing the gain of 26 new CDSs, the vast majority of which were annotated as pseudogenes (Table S5). Thus loss or gain of regions containing ICEA following MCT had a direct impact on the overall gene content and genome size considering the extremely small mycoplasma gene content (about 800 CDS in Ma genomes).

Gain or loss also affected CDSs that belong to the accessory genome and were not ICEA-encoded. A high proportion was annotated as hypothetical protein (71% of acquired CDSs and 50% of lost CDSs) and over 10% of the exchanged CDSs corresponded to membrane lipoproteins (Table S5). When compared to the PG2-donor, strain 5632 genome possess four additional genes that encode two type-II restriction modification systems (MAGa3950 and MAGa3970; MAGa4250 and MAGa4260) and one CDS coding for a type-III methylase (MAGa1570) (Table S5) (Nouvel et al. 2010). During MCT, replacements can result in these being deleted as illustrated by transconjugant T2 which lost all 4 CDS belonging to the two specific 5632 type-II RM systems (Table S5).

MCT is a powerful distributive chromosomal transfer process which induces gene replacement but also results in gain and/or loss of genetic material. Like meiosis it creates an extremely diverse blend of the two parental genomes.

MCT introduces micro-heterogeneity in mycoplasma mosaic genomes

Comparative genome analyzes at the nucleotide level revealed that the total number of SNPs introduced by MCT in the recipient 5632-chromosome varied from 630 to 4092 nt, with a mean of 2590 bases substitution and 226 indel polymorphisms (Table 1), 90% of these variations being located within CDS.

Most PG2-acquired regions corresponded to several kb DNA fragments in which a large number of SNPs could be detected (SNPs averaging one per 11 to 26 nt depending on the software

used, see Materials and Mvisualisationmethods). Our analyses also concomitantly uncovered micro-complex regions (Fig. 1C) that were composed (i) by a few donors SNPs (here, arbitrary by regions containing ≤ 10 of donor SNPs) or (ii) by inherited PG2-fragments interspaced by short interval (≤ 10 SNPs) of the 5632-recipient DNA. (Fig. 1C, Table S6). These micro-complex regions represented approximately 20% of the transferred material (Table S6) and were not *de novo* mutations as they were clearly of parental inheritance as illustrated in Figure 1C. Most likely, these are the result of a combination of “repair and recombination events” occurring between the recipient chromosome and, one or more of the transferred, donor-DNA fragments. Interestingly, about 25% of the overall SNPs introduced by the donor caused an amino acid substitution when compared to the original 5632 parental product. This proportion increases in smaller micro-complex regions (≤ 5 SNPs) with more than 40% of SNPs introducing non-synonymous mutations (Table S7).

Another source of micro-complexity is provided by chimeric-CDSs which appeared to be the result of one or more homologous recombination events occurring within the CDS themselves. Chimeric CDSs are thus a combination of both parental sequences and their number in transconjugants varied from 4 to 35 (Table S5 and S6). They corresponded to house-keeping as well accessory-genes, introducing numerous SNPs in gene sequences e.g. in *parC* (45 SNPs in T3, 97 SNPs in T5-2,), *ruvB* (9 SNPs in T4-3), *recA* (19 SNPs in T5-3 and 9 SNPs in T5-4), or in the P30 lipoprotein gene (40 SNPs in T5-1).

Overall, these data show that MCT induces two levels of complexity in transconjugants genomes: a macro- and a micro-complexity, both of which may have a significant impact on short and long term evolution.

The complexity of the transconjugants mosaic genomes is affected by the selection pressure

Comparison of transconjugant genomes derived from a same parental pair during classical mating experiment (T1-1 to -2, T4-1 to -3, T5-1 to -5 or T6-1/4/5), revealed that the number of exchanged chromosomal-fragments varied on average by 3-fold and the percentage of transferred PG2 genome by 2.2-fold among siblings (Fig. 2 and Table S4). The position and the nature of the selective-marker, Gm versus Tet, did not influence the extent of genome mosaicism: the mean value of the number of exchanged DNA fragments was comparable between progenies generated by 5632^T/PG2^G pair (mating M1, M2 and M3 depending on the

Gm position in the donor chromosome) or by the 5632^G/PG2^T pair (mating M4) with a mean of 7 fragments. Of note, the percentage of transferred PG2-DNA was slightly higher (10.4%) with the 5632^T/PG2^G pair than with the 5632^G/PG2^T (7.3%). This may be due to the respective position of the recipient and donor markers: the closer they get in the final transconjugant progenies the more the risk of losing the donor marker increased during incorporation of the donor fragment carrying the selective marker (Fig. 1).

When more than one donor chromosomal region was required for adaptation, here when using PG2 Enro^R mutant (PG2^{E10}) as donor (matings M5 and M6; high-selection pressure), we observed a higher degree of mosaicism of the progenies (Fig. 1 and 2). Indeed, T5 and T6 transconjugants had to acquire two diametrically opposed chromosomal mutated regions: *parE-C* (212329 - 214190) and *gyrA* (647997) to withstand the enrofloxacin selective pressure (Faucher et al. 2018). T5 and T6 Enro^R transconjugants in fact did inherit a mean of 19.8 ± 7.7 donor fragments compared to 7.1 ± 3.2 when the selection pressure applied on a single locus (mating M1 to 4) (Table S4). Interestingly, the total amount of acquired PG2-DNA was almost identical in both cases ($11.8\% \pm 3.4\%$ for T5-6 Gm/Enro^R transconjugants versus $9.1\% \pm 3.9\%$ for Gm/Tet). Compare to classical matings the acquired fragments in progenies selected under “high-selection pressure” were smaller (average of 11.7 kb versus 5.8 kb, respectively) (Fig. 2 and Table S2). The mosaicism surrounding the selective regions was higher in T5- and T6-Enro^R progeny where 40% of the exchanged fragments mapped within 50 kb from the donor Enro^R mutations compared to 20% for progeny obtained with the Gm or Tet inserted at one locus of the donor. Furthermore, the percentage of micro-variations related to the remnant 5632-DNA present between the PG2-donor regions was twice higher in T5 and T6 progenies than that observed for other “classical” matings. Progenies which selection relied on the acquisition of diametrically opposed chromosomal regions seem to have undergone a high rate of recombination events especially in regions surrounding chromosomal mutation markers. Despite these high recombinogenic properties, the amount of exogenous DNA integrated by the Enro-selected transconjugants is not superior, suggesting a maximum assimilation rate for exogenous DNA.

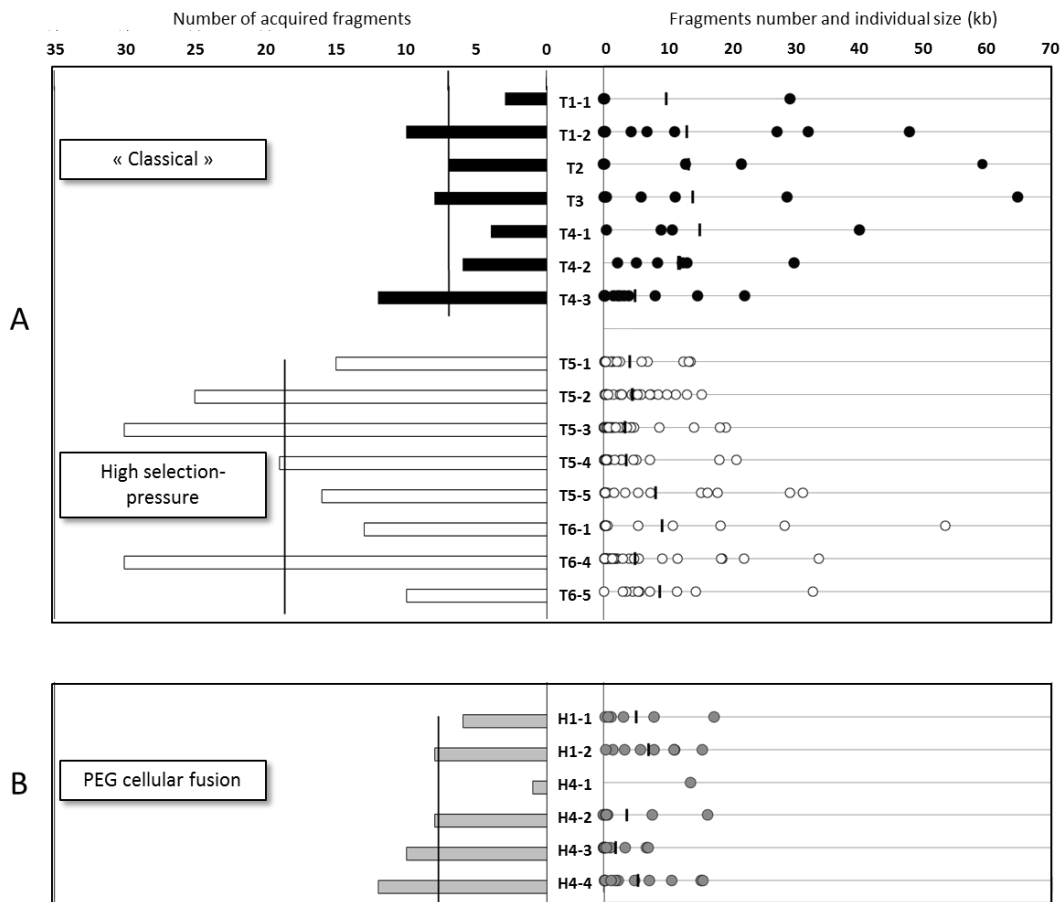


Figure 2. Number and size of PG2-donor fragments acquired by MCT in mating progeny

(A) Histogram and dot plot chart illustrating the number and size of PG2 acquired fragments in all 5632 mating progenies (T1-n to T6-n transconjugants). Data obtained with transconjugants produced by classical mating conditions (see Fig. 1, M1 to M4) are represented by black-histogram and -dots and those produced by “high selection-pressure” mating (see Fig. 1, M5 and M6) by white-histogram and -dots. The histogram represents the number of PG2-donor fragments acquired by each transconjugants. The average number of acquired DNA fragments of all transconjugants produced during (i) classical mating or (ii) high selection-pressure conditions is represented by a line. The dot plot chart illustrated the size and the number of each PG2-fragment integrated in 5632 transconjugants. Each dot represents a donor-fragment, which is positioned in the graphic depending of its size (kb). Black thick lines correspond to the mean size of all DNA fragments acquired in individual transconjugants. (B) Histogram and dot plot chart representing the number and size of incorporated fragments in hybrids progeny obtained after PEG-cellular fusion (H1-1&2 and H4-1 to -4, see Fig. 3). As panel A, the histogram represent the number of PG2-donor fragments and the dot charts the size and the number of each PG2- incorporated fragment in 5632 PEG-hybrids. The scale is the same as in panel A.

PEG artificial fusion of mycoplasma cells also generates genome mosaics but of lower complexity

We previously demonstrated that mycoplasma conjugation could be artificially bypassed by using PEG, a polymer which induced membrane fusion and resulted in chromosomal exchanges independently from ICE-encoded conjugative properties. To gain further insights into parameters affecting MCT and genome mosaicism, parental pairs previously used for the classical matings, M1 and M4, were respectively fused using PEG as previously described (Dordet-Frisoni et al. 2014) and six individual, double-resistant progenies were selected and analyzed as above (Fig. 2C and 3, Table S4). Their hybrid genomes, H1-1 and -2 and H4-1 to -4, are depicted in Figure 3 along with that obtained with the same pairs but in classical mating conditions (T1 and T4). As for classical conjugation, all PEG-hybrid progenies possessed a 5632 backbone with several incorporated PG2-DNA fragments and no influence of the nature and position of selection marker. The number of PG2-recombined fragments was similar to that observed during classical mating with a mean of 7.5 ± 3 exchanged DNA regions (Fig. 2, Table S4). As for the “high-selection pressure” the mean size of PG2 incorporated-DNA segments is smaller to that obtained in classical matings (11.7 kb) with an average of 6.1 kb (Fig. 2, Table S4). Regardless of the selective pressure, PEG-hybrid genomes showed less complexity compared to transconjugants, with the number of gained and/or lost CDS as well as number of acquired SNPs being lower (Table S6). In PEG-hybrid genomes, PG2 incorporated fragments were mostly concentrated around the selection marker (Fig. 3). Despite a greater simplicity of PEG-hybrid genomes, this result suggested that the recombination process observed for MCT does not require the presence of the conjugative pore machinery. Analysis of PEG-hybrid genomes reinforces the hypothesis that the ICE mainly contributes in promoting cell-to-cell contact, bridging the cell cytoplasmic content so that chromosomal exchange may take place.

Recombination regions reflects an highly permissive recombination system

When MCT was first discovered, we assumed that replacement of chromosomal recipient-loci by their donor counterparts occurred via homologous recombination. The Illumina sequencing of the 21 mosaic genomes provided us with a unique opportunity to revisit this issue. We then focused on homologous regions (HRs) flanking the inherited donor-fragments of the transconjugants and PEG-hybrid genomes.

HRs were defined as identical sequences in between donor and recipient DNA that were delimited at one end by the first recipient-SNP and, at the other, by the first donor SNP encountered on the chromosome (Fig. 4A). Alignment and motif search of HR surroundings did not reveal any specific sequence or bias in GC percent. HR-length varies from 2 to 394 nt, with a mean of 64 nt and no difference in HR-length was observed between the different experiments (Fig. 4B, Table S8).

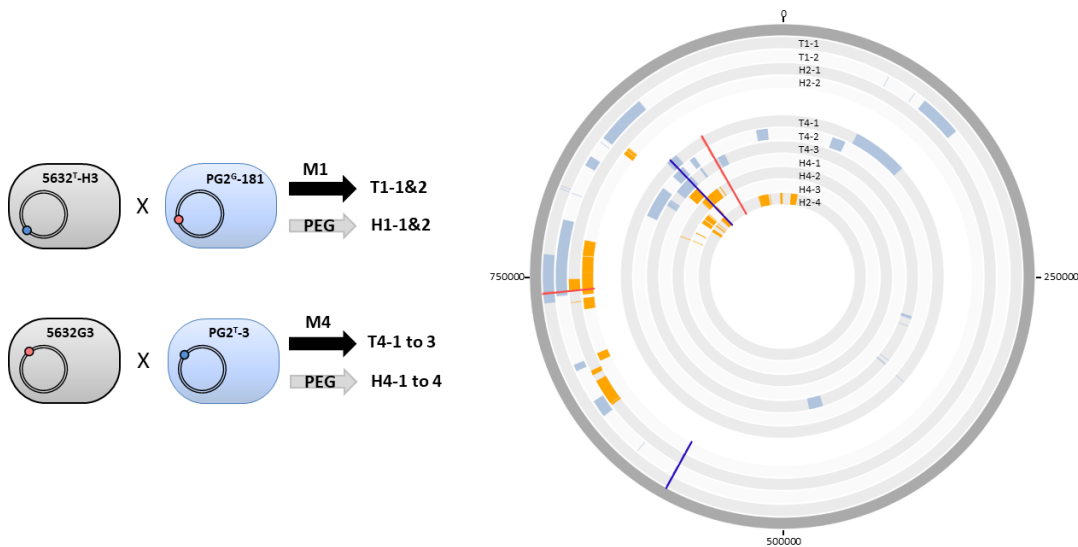


Figure 3. Artificial PEG fusion of mycoplasma cells also results in mosaic genomes

PEG cellular fusions were conducted with pairs used in mating M1 and M4 and resulted in PEG-hybrids (H1-1&2 and H4-1 to -4). 5632 and PG2 parents carried the Tet (blue dot) or Gm (red dot) genetic-marker at different loci. The DNA plotter depicts the mosaic nature of the 5 transconjugants (T1-1&2 and T4-1 to-3) and the 6 PEG-hybrids (H1-1&2 and H4-1 to -4). PG2-specific sequences are color-coded in blue for transconjugants and in orange for PEG-hybrids, and the 5632 genome backbone is depicted in grey. Chromosomal positions of Gm- and Tet-markers in progenies are indicated with red and blue line respectively. Outer circles: 5632, T1-1 and-2, H1-1&2 and inner circles T4-1 to3, H4-1 to 4, from outer to inner circles.

The minimal efficient processing segment (MEPS), previously defined by Shen *et al.* (Shen and Huang 1986) as the shortest length of sequence homology necessary for efficient recombination, can significantly vary among organisms (Shen and Huang 1986; Kung et al. 2013). In *E. coli*, efficient recombination has been observed with as little as 23 bp of sequence homology. For mycoplasmas, MEPS has never been determined, but in *M. agalactiae* transconjugant genomes 21.5% of HRs are micro-homologous region of less than 20 bp (Fig

4B), of which 27% were present on both side of the exchanged fragments. Yet the large proportion of micro-HRs raises the question of the existence of an illegitimate (non-homologous) recombination (IR) mechanism that joins DNA molecules at sites with only a few identical nucleotides.

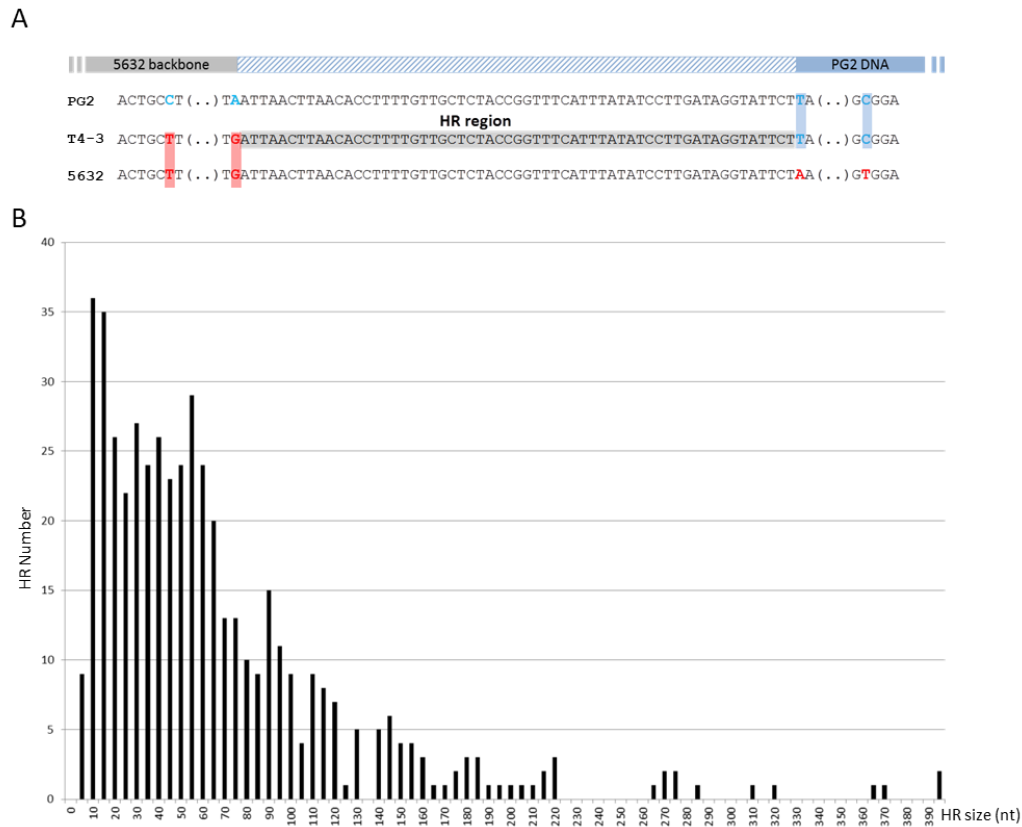


Figure 4 : Analyses of homologous region (HR) flanking the acquired DNA in transconjugants
 (A) Homologous regions (HR) were defined as identical sequences in between donor (grey) and recipient DNA (blue) that were delimited at one end by the first recipient-SNP (boxed in red) and, at the other, by the first donor SNP (boxed in blue) encountered on the chromosome. Only homologous regions with coverage of at least 15 X were taken into account and are represented here by a hatched blue rectangle. (B) Histogram illustrating the number of HR depending of their size (nt).

Features of MCT recipient versus donor cells

Chromosomal exchanges were documented between *M. agalactiae* strains PG2 and 5632 as well as in between the two related species *M. agalactiae* strain 5632 and *M. bovis* strain PG45 (Dordet-Frisoni et al. 2014). Each time, detection of tranconjugants required the presence of the strain 5632 and their fine genetic analyses designated the 5632 as the recipient strain. This

puzzling observation prompted us to test three other Ma strains that differ in their ICE content and history (year of isolation, geographical origin...) (Table S1). The transconjugant frequencies obtained with these strains ranged from 2.9×10^{-9} to 1×10^{-7} transconjugant/ total CFU and repeated attempts to generate transconjugant with the PG2-14628 pair always failed. Since an active ICE is occurring in this strain (Tardy et al. 2015), this indicated that other recipient cell factors are involved in MCT (Table S1).

All transconjugants were shown to display the 5632 backbone except for mating involving strain 4867. With this strain, all the progeny obtained with the PG2 as mating partner displayed a 4867 genomic backbone and in mating experiment involving 4867 and 5632, both parental backbones were detected in recipient cells (Table S1). The genomes of strain 4867 and of two transconjugants with a 4867 backbone, namely M161.11 from 5632vs4867 mating and M171.4 from PG2vs4867 mating, were sequenced by Illumina. WGS analyzes indicated that 4867 is a simplified version of the 5632 strain as both genome differ from 1084 SNPs (Table S9). Strain 4867 contains only one ICE copy instead of three in 5632, and lacks the IS30-like transposase present in multiple copies in the 5632 genome (Table S9). These differences enable the identification of the transconjugants genomic backbone but do not allow a precise mapping of the chromosomal DNA exchanges in M161.11 transconjugant. The PG2vs4867 transconjugant M171.4, however, contained six PG2-inherited DNA fragments, with two being large distant chromosomal DNA region of 23 and 42 kb (Fig. S1) while the remaining corresponded to micro-complex regions containing from 2 to 18 nt of the PG2-parental strain (Fig. S1). This transconjugant illustrates that MCT is able to create genetic mixing and genome mosaicism in 4867 recipient transconjugants.

Overall, this data showed that not only 5632 but also 4867, a strain similar to 5632, can act as recipient strain in MCT. The genomic simplicity of 4867 makes it a good model to further decipher the factors involve in MCT.

3. Discussion

This study formally established *Mycoplasma* Chromosomal Transfer (MCT) as a powerful conjugative process that creates a diversity of mosaic-genomes within a single-mating step. While the complexity of this phenomenon challenges our view on the biology of the simple

mycoplasma cell, this finding may have tremendous biological impacts in pathogenesis by providing a means for almost limitless, rapid adaptation to changing environments.

As shown by combining experimental mating experiments to WGS analyses of single-cell progeny, MTC allows a recipient genome to simultaneously acquire, multiple, large or small donor DNA fragments. In one mating-step, two clonal populations can generate a diversity of unique blends of hybrid progeny. MCT is an unusual conjugation process in that virtually any loci can be transferred in any combination as long as the resulting mosaic-genome is viable. A similar unconventional conjugative mechanism termed Distributive Conjugal Transfer (DCT) has been described in *Mycobacterium smegmatis* by Gray *et al.* (Gray *et al.* 2013). Both systems generate extensive genome-mosaicism and, to some extent, resemble eukaryote meiosis: in mycobacterial DCT or in MTC, 20-25% of donor DNA can be exchanged with their recipient orthologues and distributed around the chromosome into up to 30 large and small fragments (Derbyshire and Gray 2014) (This study). While MCT and DCT may appear identical from a genomic point of view, these systems belong to two distinct groups of bacteria that have little in common: mycoplasma and mycobacteria are phylogenetically remote, differs in their cell-envelop, have respectively small and large genomes with low and high GC content... As for conjugation, one main difference between DCT and MCT resides in their secretory systems. In mycobacteria, two Type VII secretion apparatus, ESX1- and -4 are essential for DCT and are, to our knowledge chromosomally encoded (Gray *et al.* 2016; Gray and Derbyshire 2018). In contrast, no Type VII secretion system has been detected in mycoplasmas, but several species possess a simplified version of Type IV Secretion System (T4SS) which is carried by an Integrative and Conjugative Elements (ICE) (Calcutt *et al.* 2002; Marena *et al.* 2006; Dordet Frisoni *et al.* 2013; Baranowski *et al.* 2018). Our group has previously demonstrated that such ICE element is self-transmissible and plays a key role in mycoplasma conjugation (Dordet-Frisoni *et al.* 2013). Hence, for MCT to occur, at least one of the two partners should carry an ICE (Dordet-Frisoni *et al.* 2013, 2014) and recent functional analysis confirmed the essentiality of ICE-encoded factors in conjugation (Baranowski *et al.* 2017). Mycoplasma conjugation is a complex phenomenon as two concomitant events take place during mating: the horizontal ICE self-transmission and chromosomal transfer, both occurring with a same frequency. Indeed, ICE can disseminate from ICE-positive to ICE-negative cells while chromosomal transfer is observed, opposite, from ICE-negative to ICE-positive cells, both DNA movements being physically independent (Dordet-Frisoni *et al.* 2014). This and data obtained with mating experiments involving different *M. agalactiae* strains, with or without ICE, suggest

that the presence of a T4SS is only required in the recipient cell. For mycobacterial DCT, ESX secretion apparatuses have been demonstrated to be essential both for donor and recipient cells (Gray and Derbyshire 2018), with no MGE or ICE being involved.

In absence of a cell-wall, mycoplasma cells can be artificial fused using PEG to bypass the first steps of conjugation (Dordet-Frisoni et al. 2014). This provides a wonderful means to address the influence of conjugative properties on DNA exchange and recombination. To our surprise, analyses of hybrid genomes generated by PEG cell-fusion also resulted in mosaic genomes all having the overall genetic background of the “ICE-positive strain”. Thus, the apparent polarity of the chromosomal DNA transfer (see above) might not be due to the ICE-conjugative machinery but rather to cytoplasmic factors that allow, or not, the survival and the incorporation of foreign DNA. Indeed, a difference in genes encoding DNA restriction-modification systems exists between the partners used in this study (Nouvel et al. 2010), with the ICE itself encoding a hypothetical DNA methyl-transferase (Marenda et al. 2006; Baranowski et al. 2017). Addressing experimentally the role of these factors in MCT is currently on the way but will require time due to the limited number of genetic tool available for mycoplasmas. Interestingly, hybrid-genomes obtained following PEG-fusion were less heterogeneous than those obtained during conjugation, with the exchanged fragments being more concentrated around the selective marker (Fig. 3). Identifying the factors responsible for this observation might not be trivial but will contribute to our understanding of the dynamic of the recombination events during MCT. In particular, one key question is whether multiple recombination events occur at the same time or sequentially, with their number depending of the length of the cell-to-cell contact. Preliminary data indicates that is unlikely due to the time of contact and co-incubation of the partners (data not shown).

When MCT was first discovered, our initial hypothesis was that DNA exchanges occurred via homologous recombination (Dordet-Frisoni et al. 2014). This was based on the analyses of a limited number of single transconjugant genomes and data further gathered in the current study indicate that this hypothesis may be not always valid. For homologous recombination to occur, RecA must recognize a sequence of sufficient similarity during strand exchange which one corresponds to the minimal efficient processing segment (MEPS) (Shen and Huang 1986). MEPS has never been determined for mycoplasmas but in *E. coli* its length depends on the recombination pathway: it is between 23-27 base pairs (bp) for the RecBC pathway and between 44-90 bp for RecF-dependent pathway (Shen and Huang 1986). Homologous Regions (HR) flanking donor-DNA in mycoplasma transconjugant genomes were relatively short, with

almost a third being less than 30 nt and ca. 7% \leq 5nt. Since mycoplasma genomes lack the MutSHL recombination system, these organisms might be more tolerant for sequence dissimilarity than *E. coli* or other classical bacteria (Sirand-Pugnet et al. 2007) and short HRs may reflect the high tolerance of the recombination machinery for dissimilarities as a consequence of an imperfect mismatch repair system. Mosaic-genomes generated during MCT or artificial cell fusion are thus most likely the results of both homologous and illegitimate recombinations, the latter being responsible for joining DNA molecules at sites of few or no homology (Brigulla and Wackernagel 2010). During natural transformation, several bacteria (de Vries and Wackernagel 2002; Prudhomme et al. 2002; Meier and Wackernagel 2003) used a homology-facilitated illegitimate recombination (HFIR) mechanism which enables the integration of foreign DNA (Brigulla and Wackernagel 2010). This mechanism involved an illegitimate recombination event in the foreign DNA adjacent to a homologous DNA (the anchor region) and occurred mostly within short stretches of sequence identity (3 to 12 nucleotides) with increased GC content. Mycoplasma micro-HR identified in our study have a mean GC% of 29.1 \pm 1.6%, a value no significantly different from the 29.7% calculated for the whole genome. This suggested that a system different than HFIR may occurred in mycoplasma. In *E. coli*, two short-homology-independent illegitimate recombination systems have been described that are either mediated by DNA gyrase (Shimizu et al. 1997) or topoisomerase I (Bierne et al. 1997). Interestingly, WGS analyses of Enro^R transconjugants revealed that their mosaic genomes are more complex than in other matings. One difference was the need for these transconjugants to acquire two diametrically opposed chromosomal mutated regions to withstand the Enrofloxacin selection pressure (Faucher et al. 2018). This constraint may have induced and/or promoted the selection of highly-recombinogenic transconjugants. Another explanation is the use of fluoroquinolone which one inhibits the bacterial DNA gyrase. Mutations in *gyrA*, and thus changes in the primary sequence of subunit A, and inhibition of the DNA gyrase by the quinolone have been shown to enhance illegitimate recombination and confer a hyper-recombination phenotype in *E. coli* (Ikeda et al. 1981; Ashizawa et al. 1999). In absence of data that would support one or this other, these hypotheses remained to be addressed.

MCT generates a unique blend of the parental genomes by combining macro genetic exchanges involving entire CDSs, operon and loci, with micro HGT-events that mixed genome at the nucleotide level. A previous comparative genomic analysis suggested that 134 CDSs of the MA genome had undergone HGT with members of the Mycoides cluster. These transferred

CDSs were distributed around the genome as clusters, an organization we thought reflected successive past HGT events. Experimental data gathered here and elsewhere supported our initial *in silico* findings and further demonstrated that MTC is a past and contemporary phenomenon with broad that have shaped species (Citti et al. 2018). Comparative sequence analyses also suggested that the *M. leachii* genome is a mosaic of the *M. capricolum* and *M. mycoides spp* genomes, all belonging of the Mycoides cluster (Tardy et al. 2009). Indeed, this cluster is composed of several species infecting ruminants, many of which are pathogenic and share the same habitat as *M. agalactiae*. Since, these are known to harbor ICE elements, they are likely to be able to conjugate.

An increasing number of mycoplasma genomic studies are detecting past HGT events (Citti et al. 2018; Lo et al. 2018), most often in species sharing the same host and in which ICE elements were identified. For instance, five gene clusters were predicted to have undergone HGT in between the phylogenetic distinct species *M. hominis* and *U. parvum*, two colonizers of the human urogenital tract (Pereyre et al. 2009). These data indicate that MCT is most likely not restricted to *M. agalactiae* or to ruminant mycoplasmas.

During MTC and depending on the genome sytheny, CDSs can be gained or lost, regardless of whether they are house-keeping or accessory genes. These events can result in the acquisition or the loss of entire biological pathways or of key adaptive properties and may have major impacts for mycoplasma survival and host-interactions. For instance, *M. agalactiae* has inherited an oligopeptide transport system from a Mycoides ancestor which is involved a range of biological events: biofilm formation, antimicrobial-compound production, and adaptation to specific environments such as milk from which the pathogen is often re-isolated (Sirand-Pugnet et al. 2007). In the current study and *in silico* analyzes (Sirand-Pugnet *et al.*, 2007), a large number of exchanges included genes encoding surface lipoprotein and membrane protein. In absence of a cell-wall, the mycoplasma membrane constitutes the primary interface with the host environment, mediating all types of direct and indirect interactions. Thus, change, acquisition or loss by MCT of this category of genes can dramatically affect the interaction with the host. For instance, mycoplasmas having acquired the MIP-MIB gene tandem via HGT may have an advantage in withstanding the humoral host response because of their products inducing the capture and cleavage of IgG (Arfi et al. 2016). MCT may also have a tremendous impact on diagnostic, typing, or epidemiological studies by blurring the border in between species and strains. Hence, extensive horizontal gene transfers in human

ureaplasmas (*Mycoplasmatacae* family) have questioned the utility of serotyping for diagnostic purposes (Xiao *et al.* 2011).

Simultaneously to the exchanges of large fragment, micro-complex regions are created, inducing minor nucleotides substitutions and sequence alteration. Yet, these fine modifications may be responsible for important phenotypic changes, such as the acquisition of high-level of antibiotic resistance following the transfer of a few point mutations (Faucher *et al.* 2018) (see above).

Overall, the mycoplasma minimal cells have retained a form of bacterial sex which outcomes may have a tremendous impact on their evolution and on their adaptability to complex hosts. The chromosomal transfer described here for mycoplasmas, by Gray *et al.* for *Mycobacteria* (Gray *et al.* 2013), and by an increasing number of reports for other bacterial genera most likely operates via different refine mechanisms but the end-process is identical, indicating that convergent evolution works towards creating always more diversity. At current, the occurrence and role of conjugative CTs in bacteria are probably underestimated because of methodological issues such as the difficulties in following these events within related species or strains or the high prevalence of MGE found in some genomes. Thanks to their minute-size genome, mycoplasma offers a simplified model (i) to decipher the intrinsic mechanisms of DNA transfer and integration responsible for this extraordinary plasticity and (ii) to address the biological impact of such mechanism using true pathogens.

Acknowledgments

This work was supported by INRA and ENVT. We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigenae group for providing help and storage resources. We would like to thank, and Dr Marc Serge Marendia, University of Melbourne for providing the PG2^{E10} strain, and Dr Pascal Sirand-Pugnet, University of Bordeaux for PG2^{E10} whole genome sequencing.

4. Materials and Methods

Mycoplasma strains and culture conditions

M. agalactiae (Ma) strains PG2, 5632 (Nouvel et al. 2010), 14628 (Tardy et al. 2012), 4055 and 4867 (Tardy et al. 2015) were used in this study (Table S1). Mycoplasmas were grown at 37°C in SP4 medium supplemented with 45 µg.mL⁻¹ cefquinome (cobactan 4.5%, MSD Animal Health). When needed, gentamicin (50 µg.mL⁻¹), tetracycline (2 µg.mL⁻¹) or enrofloxacin (0.25 µg.mL⁻¹) were added to the medium alone or in combination depending on selection needs. When necessary, mycoplasmas were subcloned by serial passages in broth and solid media. Mycoplasma cultures were stored at -80°C.

Mycoplasma genomic tagging by insertion of stable selective antibiotic makers

Antibiotic markers providing resistance to gentamicin (Gm) or tetracycline (Tet) were randomly introduced in the genome of *Ma* as previously described (Dordet Frisoni et al. 2013) by using a modified version of the transposon Tn4001. Briefly, mycoplasmas were transformed with the pMT85 plasmid which carries a mini-Tn composed of the gentamicin resistance *aacA-aphD* gene or the tetracycline resistance determinant *tetR* flanked by the two Tn4001 inverted repeats (IRs). Since the mini-Tn contains no transposase gene which one is located elsewhere on the pMT85, its insertion in the mycoplasma chromosome is stable. In addition to conferring antibiotic resistance to the cell (Baranowski et al. 2010), this provides a proxy to trace chromosomal sequences that have been transferred. Insertion of the mini-Tn was monitored by PCR, and its chromosomal location determined by direct sequencing of genomic DNA as previously described (Dordet-Frisoni et al. 2014) (Table S2). An enrofloxacin-resistant (Enro^R) strain of *M. agalactiae* PG2 (PG2^{E10}) with mutated Enro target genes (3 SNPs located in *parC* (214190), *parE* (212329) and *gyrA* (697997)) was obtained after four passages in medium containing increasing concentrations of enrofloxacin (0.25, 0.5, 1 and 10 µg.mL⁻¹) (Faucher et al. 2018) To enhance the chance of detecting recombination events between the different *Ma* genomes, mating with 4055, 4867 and 14628 strains versus PG2 and 5632 were conducted using pools of clones carrying the selective marker at different positions among the chromosome.

Mating experiments and genetic characterization of transconjugant progenies

Mating experiments were conducted as previously described (Dordet Frisoni et al. 2013) using individual (Fig 1 and Table S2) or pools (Table S1) of marked clones mixed in equal proportions (Table 3). Briefly, donor and recipient cells were cultured for 24 h in SP4 medium and mixed in a final volume of 2 ml at a 1:1 volume ratio (about 10^9 CFU.mL⁻¹). The mixture was centrifuged and the pellet resuspended in 1 ml SP4 before further incubation at 37°C for 16h. Cells were then seeded onto SP4-solid media supplemented with the appropriate antibiotics and incubated at 37°C. Individual, transconjugant colonies were randomly picked and grown into SP4 medium supplemented with gentamicin and tetracycline or gentamicin and enrofloxacin. Mycoplasma membrane cell fusion in the presence of PEG (polyethylene glycol 8000) was carried out as previously described (Teachman et al. 2002). The frequency of transconjugant was determined as the number of transconjugants divided by the total CFU. Genomic DNA was extracted from mycoplasma cells using the classical phenol-chloroform method. Before DNA extraction, transconjugants were cloned three times, with filtering through a 0.22 µm-pore filter. PCRs were carried out for the detection of antibiotic resistance markers and for genomic backbone determination using oligonucleotides described in Dordet-Frisoni *et al.* (Dordet-Frisoni et al. 2014) for 5632 vs PG2 transconjugants and other described in Table S3.

Illumina Whole Genome Sequencing and bioinformatic analyses

Illumina sequencing using HiSeq technology (paired-end, 2x150 bp, average of 3000X for coverage depth) was performed according to the manufacturer's instructions (Illumina, San Diego, CA) at the GATC Biotech facility (Konstanz, Germany). Bioinformatics analyses were conducted using the GenoToul Bioinformatics facility, Toulouse, France (<http://bioinfo.genotoul.fr/>). Quality of the sequencing reads was controlled with FastQC tool (Li and Durbin 2009) (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Each transconjugant or PEG-hybrid genomes was respectively aligned to the *M. agalactiae* genome PG2 55-5 (NC_009497.1), 5632 C1 (NC_013948.1) and 4867 using Burrows-Wheeler Aligner with MEM algorithm (Li and Durbin 2009). The quality of the alignments was controlled with Qualimap 2.2.1 (mean coverage and mean mapping quality) (Okonechnikov et al. 2016) and samtools flagstat (Li et al. 2009). The resulting alignments were converted to BAM, indexed and sorted using Samtools and visualized using the Integrative Genome Viewer (IGV 2.3.93)

(Thorvaldsdóttir et al. 2013), Artemis 16.0.0 (Rutherford et al. 2000) and ACT 13.0.0 (Carver et al. 2005).

Detection of parental inherited fragments in M1 to M6 mating transconjugants was possible because of the polymorphisms existing between PG2 and 5632. SNP and polymorphisms analyses between all parental strains were done with Breseq and Nucmer tools. Between 5632 and PG2 genome the polymorphism corresponds to 1 SNP every 11 nt, excluding ICE sequences, as calculated with Breseq v0.28.1 (Deatherage and Barrick 2014) or 1 SNP or indel every 26 nt using Nucmer software which only detect mutations in closely related nucleotide sequences and do not considered difference in gene content (Delcher et al. 2002). Alignment and motif search in region surrounding HR was done with MEME/MAST v5.0.1 (Bailey and Gribskov 1998) .

To define genomic region of different parental origin with no ambiguity, reads mapping equally on both genomes were filtered out and mismatches were not allowed for reads specifically mapping to one of the two parental genomes. Donor-specific reads, corresponding to inherited regions, were manually curated using Artemis and only reads with coverage higher than 15X and a feature size of at least 25 nt were taken into account. Circular plots were obtained with the Artemis DNA plotter tool.

De novo assembly of strain 4867 was done with using SPAdes v3.11.0 (Bankevich *et al.* 2012) and resulted in 8170 contigs. For comparative genomic and SNP analyses, only contigs superior to 1 kb were conserved (47 contigs).

Data availability/submission

All sequence data (fastq) have been deposited at the EMBL database, the European Nucleotide Archive (ENA) at <http://www.ebi.ac.uk/ena> with the accession number PRJEB27571.

5. References

- Arfi Y, Minder L, Di Primo C, Le Roy A, Ebel C, Coquet L, et al. MIB-MIP is a mycoplasma system that captures and cleaves immunoglobulin G. *Proc Natl Acad Sci USA*. 2016 Apr 25;
- Ashizawa Y, Yokochi T, Ogata Y, Shobuike Y, Kato J, Ikeda H. Mechanism of DNA gyrase-mediated illegitimate recombination: characterization of *Escherichia coli* *gyrA* mutations that confer hyper-recombination phenotype. *J Mol Biol*. 1999 Jun 11;289(3):447–58.
- Bailey TL, Gribskov M. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*. 1998;14(1):48–54.
- Baranowski E, Dordet-Frisoni E, Sagné E, Hygonenq M-C, Pretre G, Claverol S, et al. The Integrative Conjugative Element (ICE) of *Mycoplasma agalactiae*: Key Elements Involved in Horizontal Dissemination and Influence of Coresident ICEs. *MBio*. 2018 Jul 3;9(4).
- Baranowski E, Guiral S, Sagné E, Skapski A, Citti C. Critical role of dispensable genes in *Mycoplasma agalactiae* interaction with mammalian cells. *Infect Immun*. 2010 Apr;78(4):1542–51.
- Bierne H, Ehrlich SD, Michel B. Deletions at stalled replication forks occur by two different pathways. *EMBO J*. 1997 Jun 2;16(11):3332–40.
- Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. The transjugation machinery of *Thermus thermophilus*: Identification of TdtA, an ATPase involved in DNA donation. *PLOS Genetics*. 2017 Mar 10;13(3):e1006669.
- Brigulla M, Wackernagel W. Molecular aspects of gene transfer and foreign DNA acquisition in prokaryotes with regard to safety issues. *Appl Microbiol Biotechnol*. 2010 Apr;86(4):1027–41.
- 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 Dec;184(24):6929–41.
- Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics*. 2005 Aug 15;21(16):3422–3.
- Citti C, Dordet-Frisoni E, Nouvel LX, Kuo CH, Baranowski E. Horizontal Gene Transfers in *Mycoplasmas* (Mollicutes). *Curr Issues Mol Biol*. 2018 Apr 12;29:3–22.
- Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol*. 2014;1151:165–88.
- Delcher AL, Phillippy A, Carlton J, Salzberg SL. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res*. 2002 Jun 1;30(11):2478–83.
- Derbyshire KM, Gray TA. Distributive Conjugal Transfer: New Insights into Horizontal Gene Transfer and Genetic Exchange in *Mycobacteria*. *Microbiol Spectr*. 2014;2(1).
- Dordet Frisoni E, Marena MS, Sagné E, Nouvel LX, Guérillot R, Glaser P, et al. ICEA of *Mycoplasma agalactiae*: a new family of self-transmissible integrative elements that

- confers conjugative properties to the recipient strain. *Mol Microbiol.* 2013 Sep;89(6):1226–39.
- Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX, Blanchard A, et al. Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *MBio.* 2014;5(6):e01958.
- Faucher M, Nouvel LX, Dordet-Frisoni E, Sagne E, Baranowski E, Hygonenq M-C, et al. Mycoplasmas under experimental antimicrobial selection: the unpredicted contribution of horizontal chromosomal transfer in accelerating short-term adaptation. *PLoS Genetics.* 2018 submitted;
- Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol.* 2005 Sep;3(9):722–32.
- García-Aljaro C, Ballesté E, Muniesa M. Beyond the canonical strategies of horizontal gene transfer in prokaryotes. *Curr Opin Microbiol.* 2017 Aug;38:95–105.
- Goessweiner-Mohr N, Arends K, Keller W, Grohmann E. Conjugation in Gram-Positive Bacteria. *Microbiol Spectr.* 2014 Aug;2(4):PLAS-0004-2013.
- Gray TA, Clark RR, Boucher N, Lapierre P, Smith C, Derbyshire KM. Intercellular communication and conjugation are mediated by ESX secretion systems in mycobacteria. *Science.* 2016 Oct 21;354(6310):347–50.
- Gray TA, Derbyshire KM. Blending genomes: distributive conjugal transfer in mycobacteria, a sexier form of HGT. *Mol Microbiol.* 2018 Jun;108(6):601–13.
- Gray TA, Krywy JA, Harold J, Palumbo MJ, Derbyshire KM. Distributive conjugal transfer in mycobacteria generates progeny with meiotic-like genome-wide mosaicism, allowing mapping of a mating identity locus. *PLoS Biol.* 2013 Jul;11(7):e1001602.
- Halary S, Leigh JW, Cheaib B, Lopez P, Bapteste E. Network analyses structure genetic diversity in independent genetic worlds. *Proc Natl Acad Sci USA.* 2010 Jan 5;107(1):127–32.
- Ikeda H, Moriya K, Matsumoto T. In Vitro Study of Illegitimate Recombination: Involvement of DNA Gyrase. *Cold Spring Harbor Symposia on Quantitative Biology.* 1981 Jan 1;45(0):399–408.
- Kloesges T, Popa O, Martin W, Dagan T. Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Mol Biol Evol.* 2011 Feb;28(2):1057–74.
- Kung SH, Retchless AC, Kwan JY, Almeida RPP. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. *Appl Environ Microbiol.* 2013 Mar;79(5):1712–7.
- Lesic B, Zouine M, Ducos-Galand M, Huon C, Rosso M-L, Prévost M-C, et al. A Natural System of Chromosome Transfer in *Yersinia pseudotuberculosis*. *PLoS Genet.* 2012 Mar 8;8(3):e1002529.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009 Jul 15;25(14):1754–60.

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15;25(16):2078–9.
- Lo W-S, Gasparich GE, Kuo C-H. Convergent Evolution among Ruminant-Pathogenic *Mycoplasma* Involved Extensive Gene Content Changes. *Genome Biol Evol*. 2018 Aug 1;10(8):2130–9.
- 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. *Journal of Bacteriology*. 2006 Jun 1;188(11):4137–41.
- Meier P, Wackernagel W. Mechanisms of homology-facilitated illegitimate recombination for foreign DNA acquisition in transformable *Pseudomonas stutzeri*. *Mol Microbiol*. 2003 May;48(4):1107–18.
- Nouvel L-X, Marenda M, Sirand-Pugnet P, Sagné E, Glew M, Mangenot S, et al. Occurrence, plasticity, and evolution of the *vpma* gene family, a genetic system devoted to high-frequency surface variation in *Mycoplasma agalactiae*. *J Bacteriol*. 2009 Jul;191(13):4111–21.
- Nouvel LX, Sirand-Pugnet P, Marenda MS, Sagné E, Barbe V, Mangenot S, et al. Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro- and micro-events that are shaping mycoplasma diversity. *BMC Genomics*. 2010;11:86.
- Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. 2000 May 18;405(6784):299–304.
- Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. 2016 Jan 15;32(2):292–4.
- Pereyre S, Sirand-Pugnet P, Beven L, Charron A, Renaudin H, Barré A, et al. Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital mycoplasmas. *PLoS Genet*. 2009 Oct;5(10):e1000677.
- Prudhomme M, Libante V, Claverys J-P. Homologous recombination at the border: insertion-deletions and the trapping of foreign DNA in *Streptococcus pneumoniae*. *Proc Natl Acad Sci USA*. 2002 Feb 19;99(4):2100–5.
- Razin S, Hayflick L. Highlights of mycoplasma research--an historical perspective. *Biologicals*. 2010 Mar;38(2):183–90.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics*. 2000 Oct;16(10):944–5.
- Shen P, Huang HV. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics*. 1986 Mar;112(3):441–57.
- Shimizu H, Yamaguchi H, Ashizawa Y, Kohno Y, Asami M, Kato J, et al. Short-homology-independent illegitimate recombination in *Escherichia coli*: distinct mechanism from short-homology-dependent illegitimate recombination. *J Mol Biol*. 1997 Feb 21;266(2):297–305.

- Sirand-Pugnet P, Lartigue C, Marena M, Jacob D, Barré A, Barbe V, et al. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS Genet*. 2007 May 18;3(5):e75.
- Smith GR. Conjugal recombination in *E. coli*: myths and mechanisms. *Cell*. 1991 Jan 11;64(1):19–27.
- Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: building the web of life. *Nat Rev Genet*. 2015 Aug;16(8):472–82.
- Tardy F, Baranowski E, Nouvel L-X, Mick V, Manso-Silván L, Thiaucourt F, et al. Emergence of atypical *Mycoplasma agalactiae* strains harboring a new prophage and associated with an alpine wild ungulate mortality episode. *Appl Environ Microbiol*. 2012 Jul;78(13):4659–68.
- Tardy F, Maigre L, Poumarat F, Citti C. Identification and distribution of genetic markers in three closely related taxa of the *Mycoplasma mycoides* cluster: refining the relative position and boundaries of the *Mycoplasma* sp. bovine group 7 taxon (*Mycoplasma leachii*). *Microbiology (Reading, Engl)*. 2009 Nov;155(Pt 11):3775–87.
- Tardy F, Mick V, Dordet-Frisoni E, Marena MS, Sirand-Pugnet P, Blanchard A, et al. Integrative conjugative elements are widespread in field isolates of *Mycoplasma* species pathogenic for ruminants. *Appl Environ Microbiol*. 2015 Mar;81(5):1634–43.
- Teachman AM, French CT, Yu H, Simmons WL, Dybvig K. Gene transfer in *Mycoplasma pulmonis*. *J Bacteriol*. 2002 Feb;184(4):947–51.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinformatics*. 2013 Mar;14(2):178–92.
- Torres-Puig S, Martínez-Torró C, Granero-Moya I, Querol E, Piñol J, Pich OQ. Activation of σ_{20} -dependent recombination and horizontal gene transfer in *Mycoplasma genitalium*. *DNA Res*. 2018 Aug;25(4):383–93.
- Vasconcelos ATR, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, 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 Aug;187(16):5568–77.
- de Vries J, Wackernagel W. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc Natl Acad Sci USA*. 2002 Feb 19;99(4):2094–9.
- Watt AE, Browning GF, Leone AR, Bushell RN, Stent A, Cutler RS, et al. A Novel *Glaesserella* sp. Isolated from Pigs with Severe Respiratory Infections Has a Mosaic Genome with Virulence Factors Putatively Acquired by Horizontal Transfer. *Appl Environ Microbiol*. 2018 Jun 1;84(11).
- Wollman EL, Jacob F, Hayes W. Conjugation and genetic recombination in *Escherichia coli* K-12. *Cold Spring Harb Symp Quant Biol*. 1956;21:141–62.

6. Supplementary material

M171.4

start	stop	fragments length (bp)
5085	28196	23111
30423	30975	552
177789	177848	59
457981	458066	85
678985	679122	137
711789	753801	42012

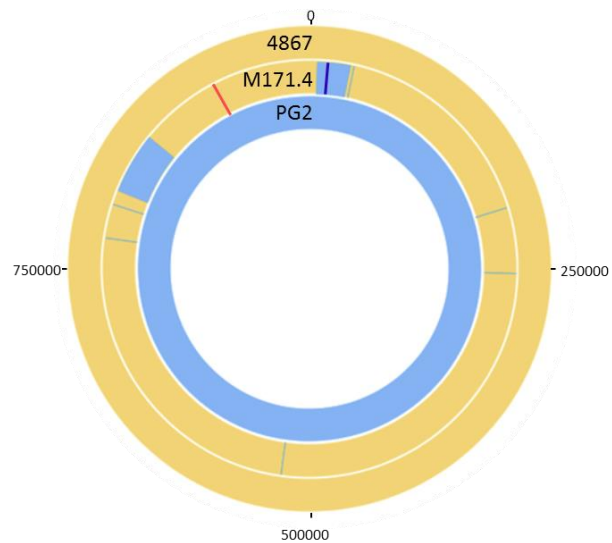


Figure S1 : Mosaic genome of the M171.4 transconjugant coming from PG2 vs 4867 mating.

Position and length of PG2-donor fragments in the M171.4 transconjugant possessing a 4867 genomic backbone. The DNA plotter depicting the mosaic nature of the M171.4 transconjugant. Chromosomal positions of Gm- (nt-922013) and Tet-markers (nt-20014) are indicated with red and blue line respectively. PG2-specific sequences are color-coded in blue, the 4867 and PG2 genome backbones are depicted in yellow and blue respectively. From outer to inner circles 4867, M171.4 and PG2.

Table S1 : *M. agalactiae* strain features, mating efficiencies and identification of resulting transconjugants

Strains origins					
Strain	Host	Associated symptoms /sample	Date of isolation	Geographical origin	ICE content
PG2	Caprine	Unknown	1952	Spain	no
5632	Caprine	Arthritidis (joint)	1991	Spain	ICE (x3)
14628	Ibex	Pneumonia (lung)	2006	France	ICE
4055	Caprine	Mastitis (milk)	1987	France	no
4867	Caprine	Mastitis (milk)	1990	Spain	ICE

Matings					
	Mating couple		ICE	Mating frequency	Progeny Backbone
	ICE				
5632^G	+	PG2^T	-	$4.8 \pm 1.1 \times 10^{-8}$	100% 5632
5632^G	+	4867^T	+	$2.9 \pm 1 \times 10^{-9}$	70% 5632 19% 4867 11% Ni
PG2^G	-	4867^T	+	$1.2 \pm 1.3 \times 10^{-9}$	98% 4867 2% Ni
5632^G	+	14628^T	+	$1.1 \pm 0.4 \times 10^{-9}$	100% 5632
PG2^G	-	14628^T	+	0	ND
5632^G	+	4055^T	-	$1 \pm 0.03 \times 10^{-7}$	100% 5632
PG2^G	-	4055^T	-	0	ND

identification on at least 12 different transconjugants

Ni: non-identifiable

ND: not determined

Table S2 : List of mating partners and chromosomal position of their respective selective markers

Transconjugants	Partner 1	Tag position	Partner 2	Tag position	Experimental conditions
T1-1	PG2 ^{G181}	648734	5632 ^{TH3}	529737	Mating
T1-2	PG2 ^{G181}	648734	5632 ^{TH3}	529737	Mating
H1-1	PG2 ^{G181}	648734	5632 ^{TH3}	529737	PEG-induced cell fusion
H1-2	PG2 ^{G181}	648734	5632 ^{TH3}	529737	PEG-induced cell fusion
T2	PG2 ^{G71}	473081	5632 ^{TH3}	529737	Mating
T3	PG2 ^{G10}	104815	5632 ^{TH3}	529737	Mating
T4-1	PG2 ^{T3}	770055	5632 ^{G3}	919899	Mating
T4-2	PG2 ^{T3}	770055	5632 ^{G3}	919899	Mating
T4-3	PG2 ^{T3}	770055	5632 ^{G3}	919899	Mating
H4-1	PG2 ^{T3}	770055	5632 ^{G3}	919899	PEG-induced cell fusion
H4-2	PG2 ^{T3}	770055	5632 ^{G3}	919899	PEG-induced cell fusion
H4-3	PG2 ^{T3}	770055	5632 ^{G3}	919899	PEG-induced cell fusion
H4-4	PG2 ^{T3}	770055	5632 ^{G3}	919899	PEG-induced cell fusion
T5-1	PG2 ^{E10}	212329 * 214190 647997	5632 ^{G3}	919899	Mating
T5-2	PG2 ^{E10}	212329 * 214190 647998	5632 ^{G3}	919899	Mating
T5-3	PG2 ^{E10}	212329 * 214190 647999	5632 ^{G3}	919899	Mating
T5-4	PG2 ^{E10}	212329 * 214190 648000	5632 ^{G3}	919899	Mating
T5-5	PG2 ^{E10}	212329 * 214190 648001	5632 ^{G3}	919899	Mating
T6-1	PG2 ^{E10}	212329 * 214190 648002	5632 ^{G3}	919899	Mating
T6-4	PG2 ^{E10}	212329 * 214190 648003	5632 ^{G3}	919899	Mating
T6-5	PG2 ^{E10}	212329 * 214190 648004	5632 ^{G3}	919899	Mating

* position of chromosomal mutations inducing enrofloxacin resistance , all other antibiotic tags corespond to resistance gene randomly introduce by transposition

Table S3 : Oligonucleotides used in this study

Primer name	Oligonucleotide sequence (5' -> 3')	products size (bp)				
		4055	4867	14628	5632	PG2
<i>M. agalactiae</i> strains		4055	4867	14628	5632	PG2
MAG0220F	CGCCATTGAAGGAACTTGCTCC	946	946	0	946	0
MAG0220R	TATGCCTGCAGATCAGTCACCA					
MAGa0900F	ATGGCAATTGATTTACTCCC	1500	1500	1500	1500	900
MAGa0900R	ATGATGTTGGCCGTATTTGG					
MAG1560F	GCCGGTGCCCGGGTTAAT	0	0	0	0	1311
MAG1560R	TCGACACCGCATTTAGCAGCA					
MAG2450F	AGGACGCTGTAGAGTTTGCAAAGAA	800	0	0	0	1500
MAG2450R	TACCACTAGCTGCCGCACCA					
MAGa 3840F	GAAGCACTAAAATCAGGCAA	1223	1223	1067	1223	1067
MAGa 3840R	GCGGCAGCAGAGTTATCAAG					
MAGa4400F	CTCTGGCTACAGATTCATCA	490	490	0	490	607
MAGa4400R	GGTGAACGAGAACAATCGCC					
MAG5860F	ACCATTGAACTACATTTGCA					
MAGa6510R	AAAACCGCAGACCCAAATGGT	0	494	3000	494	1318
MAG5890R	ACATTGCGATGTAGCGGGAACAG					
II-F1	TTTATGGCGGTGCAACTGGT	/	0	/	539	/
Left-1	TAATGGCCAAGAGTTCAAAAGCAA					
M8.4_P1	TTGCCACCCTTATTCATCATC	/	0	/	800	/
MAGb1260_F	GTAAATGGAGGTGGCAACTTATTCCA ATACTTGCTCT					

Table S4 : Extent and size of PG2-donor chromosomal DNA fragments acquired during mating and PEG-cellular fusion

		Number of acquired fragments	Fragments mean size (bp)	Total size of DNA transferred (bp)	% of PG2 genome acquired
Classical mating					
M1	T1-1	3	9790	29371	3.35%
	T1-2	10	12995	129949	14.81%
M2	T2	7	13443	94103	10.72%
M3	T3	8	13957	111657	12.73%
	T4-1	4	15081	60324	6.88%
M4	T4-2	6	11853	71115	8.10%
	T4-3	12	4959	59506	6.78%
	mean	7.14 ± 3.18	11725 ± 3424	79432 ± 34504	9.05% ± 3.9
High selection-pressure mating					
	T5-1	15	4146	62191	7.09%
	T5-2	25	4526	113155	12.90%
M5	T5-3	30	3352	100567	11.46%
	T5-4	19	3594	68286	7.78%
	T5-5	16	8126	130011	14.82%
	T6-1	13	9166	119164	13.58%
M6	T6-4	30	4929	147872	16.85%
	T6-5	10	8832	88318	10.07%
	mean	19.75 ± 7.7	5834 ± 2447	103696 ± 29750	11.82% ± 3.4
PEG cellular fusion					
PEG-M1	H1-1	6	5119	30714	3.50%
	H1-2	8	7085	56678	6.46%
	H4-1	1	13586	13586	1.55%
PEG-M4	H4-2	8	3658	25606	2.92%
	H4-3	10	1951	19507	2.22%
	H4-4	12	5427	59693	6.80%
	mean	7.5 ± 3.78	6138 ± 4040	34297 ± 19400	3.91% ± 2.2

Table S5 : PG2-CDSs acquired by transconjugants and PEG-hybrids.

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4	
<i>dnaA</i>	0010	0010									■												■	
<i>dnaN</i>	0020	0020									■													■
<i>lip</i>	0050	0030									■												■	
<i>lip</i>	0060	0040									■												■	
	0070	0050																					■	
	0080	0060																					■	
<i>lplA</i>	0090	0070																					■	
	0100	0080																					■	
<i>gcdH</i>	0110	0090																					■	
	0120	0100																					■	
	0130	0110																					■	
	0140	0120																					■	
	0150	0130																					■	
	0160	0140																					■	
	0170	0150																					■	
	0180	0160																					■	
<i>dgk</i>	0190	0170																					■	
<i>dgk</i>	0200	0180																					■	
	0210	0190																					■	
	NO	0200																					■	
	0220	0210																					■	
	NO	0220																					■	
	0230	0230																					■	
	0240	NO																					■	
	0250	0240																					■	
	0260	0250																					■	
	0270	0260																					■	
	0280	0270																					■	
	0290	0280																					■	
	0300	0290																					■	
	0310	0300																					■	
	0320	0310																					■	
	NO	0320																					■	
	0330	0330																					■	
<i>oppF</i>	0340	0340																					■	
<i>oppD</i>	0350	0350																					■	
<i>oppC</i>	0360	0360																					■	
<i>oppB</i>	0370	0370																					■	
<i>oppA</i>	0380	0380																					■	
	0390	0390																					■	
<i>cysS</i>	0400	0400																					■	

Table S5 : Continued (2/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	0410	0410						■															
	0420	0420						■															
<i>rpmG</i>	0430	0430						■															
	0440	NO																					
<i>nusG</i>	0450	0440						■															
	0460	0450						■															
<i>pyrH</i>	0470	0460						■															
<i>rfi</i>	0480	0470						■															
	0510	0480						■															
	0520	0490						■															
<i>gpsA</i>	0530	0500						■															
<i>ksgA</i>	0540	0510						■															
	0550	0520																					
<i>trmE</i>	0560	0530																					
	0570	0540																					
<i>gap</i>	0580	0550																					
<i>serS</i>	0590	0560																					
<i>epsG</i>	0600	0570						■															
<i>rpmH</i>	0610	7500		■				■															
<i>rnpA</i>	0620	0580		■				■															
	0630	0590						■															
<i>lplA</i>	0640	0600						■															
<i>lip</i>	0650	0610						■															
	0660	0620						■															
	NO	0630																					
<i>asnA</i>	0670	0640						■															
<i>polC</i>	0680	0650						■															
	0710	0660						■															
<i>pheS</i>	0720	0670						■															
<i>ung</i>	0730	0680						■															
<i>pheT</i>	0740	0690						■															
<i>glyA</i>	0750	0700						■															
<i>pip</i>	0760	0710						■															
<i>nifS</i>	0770	0720		■				■															
<i>nifU</i>	0780	0730		■				■															
<i>mucB</i>	0790	0740						■															
<i>nadD</i>	0800	0750						■															
<i>rluB</i>	0810	0760						■															
	0820	0770						■															
	0830	NO																					
<i>fba</i>	0840	0780					■	■															

Table S5 : Continued (3/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	0850	0790																					
<i>rplK</i>	0860	0800																					
<i>rplA</i>	0870	0810																					
	0880	0820																					
	0890	0830																					
	NO	0840																					
	0900	NO																					
<i>cspR</i>	0910	0850																					
	0920	0860																					
	0930	NO																					
	0940	0870																					
	0950	0880																					
<i>hprK</i>	0960	0890																					
<i>lgt</i>	0970	0900																					
<i>trxB</i>	0980	0910																					
	0990	0920																					
<i>pdhA</i>	1000	0930																					
<i>pdhB</i>	1010	0940																					
	1020	NO																					
<i>pdhC</i>	1030	0950																					
<i>pdhD</i>	1040	0960																					
<i>rpmB</i>	1050	0970																					
	1060	0980																					
	1070	0990																					
	1080	NO																					
	1090	1000																					
<i>oppB</i>	1100	1010																					
<i>oppC</i>	1110	1020																					
<i>oppD</i>	1120	1030																					
<i>oppF</i>	1130	1040																					
	1140	1050																					
	1150	1060																					
	1160	1070																					
	NO	1080																					
	NO	1090																					
	1170	1100																					
	1180	1110																					
	1190	1120																					
	NO	1130																					
	1200	1140																					
	1210	1150																					

Table S5 : Continued (4/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>cdd</i>	1220	1160				■				■													
<i>era</i>	1230	1170				■				■													
<i>pepP</i>	1240	1180				■				■													
<i>proS</i>	1250	1190				■				■													
	1260	1200				■				■													
	1270	1210				■				■													
	1280	1220				■				■													
	1290	1230				■				■													
<i>pepF</i>	1300	1240				■				■							■						
<i>potA</i>	1310	1250				■				■													
<i>potB</i>	1320	1260				■				■													
<i>potC</i>	1330	1270				■				■													
	1340	1280				■				■													
<i>rpmF</i>	1350	1290				■				■													
	1360	1300				■				■													
<i>thiI</i>	1370	1310				■				■													
	1380	1320				■				■													
	NO	1330				■				■													
	1390	1340				■				■													
	NO	1350				■				■													
	1400	1360				■				■													
<i>valS</i>	1410	1370				■				■													
<i>folD</i>	1420	1380				■				■													
<i>pta</i>	1430	1390				■				■													
<i>ackA</i>	1440	1400				■				■													
<i>coaD</i>	1450	1410				■				■													
<i>engB</i>	1460	1420				■				■													
	1470	1430				■				■													
<i>pyk</i>	1480	1440				■				■													
	1490	1450				■				■													
<i>dnaK</i>	1500	1460				■				■													
<i>gidA</i>	1510	1470				■				■													
	1520	1480				■				■													
<i>ldhD</i>	1530	1490				■				■													
<i>lip</i>	1540	1500				■				■													
	1550	1510				■				■													
	1560	1520				■				■													
	1570	NO				■				■													
	1580	1530				■				■													
	1590	NO				■				■													
<i>tig</i>	1600	1540				■				■													

Table S5 : Continued (5/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	NO	1550																					
	NO	1560																					
	1610	1570																					
	1620	NO																					
	1630	1580																					
	NO	1590																					
	NO	1600																					
	NO	1610																					
	NO	1620																					
	1640	1630																					
	1650	1640																					
<i>polA</i>	1660	1650																					
<i>dnaE</i>	1670	1660																					
	1680	1670																					
	1690	1680																					
	1700	1690																					
	1710	1700																					
	NO	1710																					
	NO	1720																					
	1720	1730																					
<i>gidA</i>	1730	1740																					
	1740	1750																					
	1750	1760																					
	NO	1770																					
	1760	1780																					
	1770	1790																					
<i>def</i>	1780	1800																					
	1790	1810																					
<i>parE</i>	1800	1820																					
<i>parC</i>	1810	1830																					
	1820	1840																					
<i>rnhB</i>	1830	1850																					
	1850	NO																					
<i>gidB</i>	1870	1860																					
<i>prs</i>	1880	1870																					
	1890	1880																					
	1900	1890																					
	1910	1900																					
	NO	1910																					
<i>bsn</i>	1920	1920																					
<i>tkt</i>	1930	1930																					

Table S5 : Continued (6/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>ppa</i>	1940	1940				█					█						█						
<i>rpsT</i>	1950	1950				█					█						█						
	1960	1960				█					█						█						
	1970	1970				█					█												
	1980	1980				█					█												
<i>hit</i>	1990	1990				█					█												
	2000	2000				█					█												
<i>rpmG</i>	2010	2010				█																	
<i>tpx</i>	2020	2020				█											█						
	2030	2030				█											█						
<i>ftsY</i>	2040	2040				█							█										
	2050	2050				█							█										
<i>metS</i>	2060	2060				█							█										
	2070	2070				█							█										
<i>vacB</i>	2080	2080				█						█					█						
	2090	2090				█						█											
<i>gmk</i>	2100	2100				█							█										
	2110	2110				█							█										
<i>pknB</i>	2120	2120				█							█										
<i>engC</i>	2130	2130				█							█										
<i>rpe</i>	2140	2140				█							█				█						
	2150	2150				█							█										
	2160	NO				█							█										
	2170	2160				█							█										
	2180	NO				█							█										
	2190	2170				█							█					█					
	NO	2180				█							█										
	2200	NO				█							█										
	2210	NO				█							█										
<i>pyrG</i>	2300	2190				█						█											
<i>ileS</i>	2310	2200				█						█											
<i>lsp</i>	2320	2210				█						█											
	2330	2220				█						█											
<i>ruvA</i>	2390	2230																					
<i>ruvB</i>	2400	2240							█														
<i>secD</i>	2410	2250							█				█										
<i>truB</i>	2420	2260							█				█										
	2430	2270							█				█										
<i>ribF</i>	2440	2280																					
<i>rpsO</i>	2450	2290																					
<i>lip</i>	2460	2300																					

Table S5 : Continued (7/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>gtsA</i>	2470	2310																					
<i>gtsB</i>	2480	2320																					
<i>gtsC</i>	2490	2330																					
	2500	2340							■														
	2510	2350																					
	2520	2360																					
<i>rplI</i>	2530	2370																					
<i>dnaB</i>	2540	2380																					
	2550	2390																					
	2570	2400																					
	2580	2410																					
	2590	2420																					
	2600	2430																					
	2610	2440																					
	2620	2450																					
<i>nox</i>	2630	2460																					
	2640	2470																					
<i>rmuC</i>	2650	2480																					
	2660	2490																					
	NO	2500																					
	2670	2510																					
	2680	2520																					
	2690	2530																					
	NO	2540																					
	2700	2550																					
	NO	2560																					
	2710	2570																					
	NO	2580																					
<i>rpsB</i>	2720	2590																					
<i>tsf</i>	2730	2600																					
	2740	2610																					
	2750	2620																					
<i>nox</i>	2760	2630																					
	2770	2640																					
<i>rpoD</i>	2780	2650																					
<i>dnaG</i>	2790	2660																					
<i>glyS</i>	2800	2670																					
	2810	2680							■														
<i>phnD</i>	2820	2690																					
	2830	2700																					
<i>phnB</i>	2840	2710							■														

Table S5 : Continued (8/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>secA</i>	2850	2720							■								■						
	2860	2730					■		■								■						
	2870	2740															■						
<i>topA</i>	2880	2750															■						
<i>rpsF</i>	2890	2760															■						
<i>ssb</i>	2900	2770															■						
	2910	2780															■						
	2920	2790															■						
<i>deoB</i>	2930	2800															■						
	2940	2810															■						
<i>ligA</i>	2950	2820															■						
	2960	2830															■						
	2970	2840															■						
ICEA-II	2980	NO																					
	2990	NO																					
	3000	NO																					
	3010	NO																					
	3020	NO																					
	3030	NO																					
	3040	NO																					
	3050	NO																					
	3060	NO																					
	3070	NO																					
	3080	NO																					
	3090	NO																					
	3100	NO																					
	3110	NO																					
	3120	NO																					
	3130	NO																					
	3140	NO																					
	3150	NO																					
	3160	NO																					
	3170	NO																					
	3180	NO																					
	3190	NO																					
3200	NO																						
3210	NO																						
3220	NO																						
	3230	2850																					
	3240	2860														■							
	3250	2870																					

Table S5 : Continued (9/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	3260	2880																					
	3270	2890																					
	3280	2900																					
	3290	2910																					
	3300	2920																					
<i>atpA</i>	3310	2930																					
<i>atpD</i>	3320	2940																					
	3330	2950																					
	3340	NO																					
	3350	2960																					
	3360	2970																					
	3370	2980																					
<i>pcrA</i>	3380	2990																					
	3390	3000																					
	3400	3010																					
<i>asnS</i>	3410	3020																					
	NO	3030																					
	3420	3040																					
	3430	3050																					
	3440	3060																					
	3450	3070																					
	3460	NO																					
<i>efp</i>	3470	3080																					
	3480	3090																					
	3500	3100																					
	3510	3110																					
	3520	3120																					
<i>msrA</i>	3530	3130																					
	3540	3140																					
<i>mutM</i>	3550	3150																					
<i>pgiB</i>	3560	3160																					
<i>pgiB</i>	3570	3170																					
<i>rpsA</i>	3580	3180																					
<i>eno</i>	3590	3190																					
<i>tufA</i>	3600	3200																					
	3610	3210																					
	3620	3220																					
<i>gcp</i>	3630	3230																					
	3640	3240																					
<i>ptsG</i>	NO	3250																					
	3760	3260																					

Table S5 : Continued (10/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	3770	3270									■												
	NO	3280																					
	NO	3290																					
	NO	3300																					
	NO	3310																					
<i>ptsG</i>	3650	3320							■														
	3660	3330							■														
	NO	3340																					
	3670	3350							■														
	3680	3360							■														
	3690	3370							■														
	3700	NO							■														
	3710	3380							■														
	3720	3390							■														
	3730	NO																					
	NO	3400																					
	NO	3410																					
	3740	NO																					
	3750	NO																					
	3780	3420			■				■														
<i>trpS</i>	3790	3430			■	■			■														
<i>thrS</i>	3800	3440			■	■			■														
	3810	3450			■	■			■														
	3820	3460			■	■			■														
<i>p30</i>	3830	3470			■	■			■	■													
	3840	3480			■	■			■														
	3850	3490			■	■			■														
<i>atpB</i>	3860	3500			■	■			■														
<i>atpE</i>	3870	3510			■	■			■														
<i>atpF</i>	3880	3520			■	■			■														
<i>atpH</i>	3890	3530			■	■			■														
<i>atpA</i>	3900	3540			■	■			■														
<i>atpG</i>	3910	3550			■	■			■														
<i>atpD</i>	3920	3560			■	■			■														
<i>atpC</i>	3930	3570			■	■			■														
	3940	3580			■	■			■														
	3950	NO			■	■			■														
	3960	NO			■	■			■														
	3970	NO			■	■			■														
	3980	3590			■	■			■														
	3990	3600			■	■			■														

Table S5 : Continued (11/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	4000	NO																					
	4010	NO																					
	4020	NO																					
	4030	NO																					
	4040	NO																					
	4050	NO																					
	4060	NO																					
	4070	3610																					
	4080	3620																					
<i>lon</i>	4090	3630																					
<i>leuS</i>	4100	3640																					
	4110	3650																					
<i>upp</i>	4120	3660																					
<i>ssb</i>	4130	3670																					
<i>pepO</i>	4140	3680																					
	4150	3690																					
<i>gcdH</i>	4160	3700																					
	4170	NO																					
<i>ftsZ</i>	4180	3710																					
	4190	3720																					
<i>mraW</i>	4200	3730																					
<i>mraZ</i>	4210	3740																					
<i>ktrA</i>	4220	3750																					
<i>ktrB</i>	4230	3760																					
	4240	3770																					
	4250	NO																					
	4260	NO																					
<i>uvrB</i>	4270	3780																					
<i>uvrA</i>	4280	3790																					
	4290	3800																					
	4300	3810																					
	4310	3820																					
	4330	3830																					
<i>hpt</i>	4340	3840																					
<i>greA</i>	4350	3850																					
	NO	3860																					
	NO	3870																					
	NO	3880																					
	NO	3890																					
	NO	3900																					
	NO	3910																					

Table S5 : Continued (12/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	NO	3920			█																		
	NO	3930			█																		
	NO	3940			█																		
	NO	3950			█																		
	NO	3960			█																		
	NO	3970			█																		
	NO	3980			█																		
	NO	3990			█																		
	NO	4000			█																		
	NO	4010			█																		
	NO	4020			█																		
	NO	4030			█																		
	NO	4040			█																		
	NO	4050			█									█									
	NO	4060			█								█										
	NO	4070			█																		
	NO	4080			█																		
	NO	4090			█																		
	NO	4100			█																		
	NO	4110			█																		
	4360	4120			█																		
	4370	4130			█																		
	4380	4140			█																		
	4390	4150			█																		
<i>alaS</i>	4400	4160			█																		
<i>trmU</i>	4410	4170			█																		
<i>deoD</i>	4420	4180			█									█									
	4430	4190			█									█									
	4440	4200			█									█									
	NO	4210			█									█									
	4450	4220			█									█									
	4460	4230			█									█									
	NO	4240			█									█									
	4470	4250			█									█									
	4480	4260			█									█									
	4490	4270			█									█									
	4500	NO																					
<i>adhT</i>	4510	4280																					
	4520	4290																					
	NO	4300																					
	4540	NO																					

Table S5 : Continued (13/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	4530	4310																					
	NO	4320																					
	4550	4330																					
<i>adhT</i>	4560	4340																					
	4570	4350																					
	4580	4360																					
	4590	4370																					
	4600	4380																					
<i>rnc</i>	4610	4390																					
<i>plsX</i>	4620	4400																					
	4630	4410																					
	4640	4420																					
	4650	4430																					
	4660	4440																					
	4670	4450																					
	4680	4460																					
<i>glpK</i>	4690	4470																					
<i>glpF</i>	4700	4480																					
<i>aspS</i>	4710	4490																					
<i>hisS</i>	4720	4500																					
	4730	4510																					
	4740	4520																					
	4750	4530																					
	4760	4540																					
<i>rplM</i>	4770	4550																					
<i>rpsI</i>	4780	4560																					
	4800	4570																					
<i>gtaB</i>	4810	4580																					
	4820	NO																					
	4840	4590																					
	4850	NO																					
	4860	NO																					
	4870	NO																					
	4880	NO																					
<i>ICEA-III</i>	4890	NO																					
	4900	NO																					
	4910	NO																					
	4920	NO																					
	4930	NO																					
	4940	NO																					

Table S5 : Continued (14/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4	
ICEA-III	4950	NO												■										
	4960	NO												■										
	4970	NO												■										
	4980	NO												■										
	4990	NO												■										
	5000	NO												■										
	5010	NO												■										
	5020	NO												■										
	5030	NO												■										
	5040	NO												■										
	5050	NO												■										
	5060	NO																						
5070	4600													■	■									
5080	4610													■	■									
5090	4620													■	■									
5100	4630													■	■									
5110	4640													■	■									
5120	4650													■	■									
5130	4660													■	■									
nusB	5140	4670												■	■									
acpD	5150	4680												■	■									
ptsH	5160	4690												■	■									
oppB	5170	4700												■	■									
5180	4710													■	■									
5190	4720													■	■									
5200	4730													■	■									
5210	4740													■	■									
infC	5220	4750												■	■									
rpml	5230	4760												■	■									
rpIT	5240	4770												■	■									
5250	4780													■	■									
5260	4790													■	■									
licA	5270	4800												■	■									
lepA	5280	4810												■	■									
5290	4820													■	■									
5300	4830													■	■									
5310	4840		■											■	■									
NO	4850		■											■	■									
5320	4860		■											■	■									
5330	NO		■											■	■									

Table S5 : Continued (15/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	5340	4870		█														█					
<i>argS</i>	5350	4880										█			█			█					
	5360	4890		█								█						█					
<i>ldh</i>	5370	4900		█								█				█		█					
	5380	4910		█								█						█					
<i>pta</i>	5390	4920										█						█					
<i>ackA</i>	5400	4930		█														█					
	5410	4940		█														█					
	5420	4950		█														█					
	NO	4960																					
<i>uhpT</i>	5430	4970																█					
	5440	4980																█					
<i>nfo</i>	5450	4990																█					
	5460	5000										█						█					
<i>trxA</i>	5470	5010										█						█					
<i>uvrC</i>	5480	5020																█					
	5490	NO																					
	5500	5030																					
	5510	5040																					
	5520	5050															█		█				
	5530	5060										█		█				█					
	5540	5070										█		█				█					
	5550	NO																					
	5560	5080										█		█									
	5570	5090																					
	5580	5100		█								█		█									
<i>deoD</i>	5590	5110		█								█		█							█		
<i>deoA</i>	5600	5120		█								█		█							█		
<i>deoC</i>	5610	5130		█																	█		
<i>tpiA</i>	5620	5140		█																	█		
	5630	5150										█		█							█		
	5640	NO										█		█							█		
	5650	5160										█		█							█		
	5660	5170										█		█									
	5670	5180										█		█									
<i>rplQ</i>	5680	5190										█		█									
<i>rpoA</i>	5690	5200										█		█									
<i>rpsK</i>	5700	5210										█		█									
<i>rpsM</i>	5710	5220										█		█									
<i>infA</i>	5720	5230												█									

Table S5 : Continued (16/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>map</i>	5730	5240										■											
<i>adk</i>	5740	5250										■											
<i>secY</i>	5750	5260										■											
	5800	NO																					
	5810	NO																					
	5820	NO																					
	5830	NO																					
	5840	NO																					
	5850	NO																					
	5860	NO																					
	5870	NO																					
	5880	NO																					
	5890	NO																					
	5900	5270										■			■								
	5910	5280																					
<i>rplO</i>	5920	5290										■			■								
<i>rpsE</i>	5930	5300										■			■								
<i>rplR</i>	5940	5310										■			■								
<i>rplF</i>	5950	5320										■			■								
<i>rpsH</i>	5960	5330										■			■								
<i>rpsN</i>	5970	5340										■			■								
<i>rplE</i>	5980	5350										■			■								
<i>rplX</i>	5990	5360										■			■								
<i>rplN</i>	6000	5370										■			■								
<i>rpsQ</i>	6010	5380										■			■								
<i>rpmC</i>	6020	7510										■			■								
<i>rplP</i>	6030	5390										■			■								
<i>rpsC</i>	6040	5400										■			■								
<i>rplV</i>	6050	5410										■			■								
<i>rpsS</i>	6060	5420										■			■								
<i>rplB</i>	6070	5430										■			■								
<i>rplW</i>	6080	5440										■			■								
<i>rplD</i>	6090	5450										■			■								
<i>rplC</i>	6100	5460						■				■			■								
<i>rpsJ</i>	6110	5470						■				■			■								
<i>hup</i>	6120	7520						■				■			■								
	6130	5480						■				■			■								
	NO	5490						■				■			■								
<i>recA</i>	6140	5500						■				■	■		■						■		
	6150	5510						■				■	■		■						■		
	6160	5520						■				■	■	■	■						■		

Table S5 : Continued (17/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>rsuA</i>	6170	5530						■					■	■	■	■			■				
<i>rpmA</i>	6180	5540						■					■	■	■	■							
<i>rplU</i>	6190	5550						■				■	■	■	■	■							
	6200	5560						■				■	■	■	■	■	■						
<i>mscL</i>	6210	5570						■				■	■	■	■	■							
<i>grpE</i>	6220	5580						■				■	■	■	■	■			■				
<i>hrcA</i>	6230	5590	■					■				■	■	■	■	■							
<i>rpmE</i>	6240	7490	■					■				■	■	■	■	■							
<i>rpsD</i>	6250	5610						■				■	■	■	■	■	■						
	6260	5620						■				■	■	■	■	■							
<i>gyrA</i>	6270	5630		■				■				■	■	■	■	■							
<i>hsdS</i>	6280	5640		■				■				■	■	■	■	■			■				
<i>hsdM</i>	6290	5650						■				■	■	■	■	■							
	6300	5660						■				■	■	■	■	■							
	NO	5670						■				■	■	■	■	■							
	6310	5680						■				■	■	■	■	■						■	
	6320	5690						■				■	■	■	■	■							
	6330	5700						■				■	■	■	■	■							
	NO	5710			■			■				■	■	■	■	■							
<i>hsdS</i>	6340	5720						■				■	■	■	■	■							
<i>hsdM</i>	6350	5730						■				■	■	■	■	■							
	6360	5740						■				■	■	■	■	■							
	6370	NO						■				■	■	■	■	■							
	6380	5750						■				■	■	■	■	■							
	6390	5760						■				■	■	■	■	■							
	6400	5770						■				■	■	■	■	■							
<i>gltX</i>	6410	5780						■				■	■	■	■	■							
	6420	5790	■					■				■	■	■	■	■							
<i>metK</i>	6430	5800						■				■	■	■	■	■							
	6440	5810						■				■	■	■	■	■							
<i>ffh</i>	6450	5820						■				■	■	■	■	■							
	6460	5830						■				■	■	■	■	■							
	6470	5840						■				■	■	■	■	■							
	6480	5850						■				■	■	■	■	■							
<i>pgk</i>	6490	5860						■				■	■	■	■	■							
	6510	NO						■				■	■	■	■	■							
	6520	NO						■				■	■	■	■	■							
	6530	5870						■				■	■	■	■	■							
	NO	5880						■				■	■	■	■	■							
	6540	5890						■				■	■	■	■	■							
	6550	5900						■				■	■	■	■	■							

Table S5 : Continued (18/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
<i>fusA</i>	6560	5910		■		■		■								■			■				
	6570	5920		■		■		■					■			■				■			
<i>rpsG</i>	6580	5930		■				■					■										
	6590	5940		■				■					■										
<i>rpsL</i>	6600	5950		■				■					■										
	6610	5960		■				■					■										
	6620	5970		■				■					■										
	NO	5980		■				■															
	6650	5990						■															
	6660	6000						■															
	6670	6010						■															
	6680	NO						■															
	6690	NO						■															
	6700	NO						■															
	6710	NO						■															
	NO	6020						■															
	NO	6030						■								■							
	NO	6040						■								■							
	NO	6050						■															
	6720	6060						■															
	6730	NO						■															
	6740	NO						■															
	6760	NO						■															
	6770	NO						■															
	6780	NO						■															
	6790	NO						■															
	6750	NO						■															
	NO	6070						■															
	6800	NO						■															
	6810	6080						■															
	6820	6090	■	■		■		■													■	■	
	6830	6100	■	■		■		■															
<i>rpoC</i>	6840	6110						■														■	
<i>rpoB</i>	6850	6120						■								■							
	6860	6130						■								■							
	6870	6140		■				■								■							
<i>ICEA-I</i>	6880	NO		■				■															
	6890	NO		■				■															
	6900	NO		■				■															
	6910	NO		■				■															
	6920	NO		■				■															

Table S5 : Continued (19/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4	
ICEA-I	6930	NO	█	█				█			█				█									
	6940	NO	█	█				█			█				█									
	6950	NO	█	█				█			█				█									
	6960	NO	█	█				█			█				█									
	6970	NO	█	█				█			█				█									
	6980	NO	█	█				█			█				█									
	6990	NO	█	█				█			█				█									
	7000	NO	█	█				█			█				█									
	7010	NO	█	█				█			█				█									
	7020	NO	█	█				█			█				█									
	7030	NO	█	█				█			█				█									
	7040	NO	█	█				█			█				█									
	7050	NO	█	█				█			█				█									
	7060	NO	█	█				█			█				█									
	7070	NO	█	█				█			█				█									
	7080	NO	█	█				█			█				█									
	7090	NO	█	█				█			█				█									
7100	NO	█	█				█			█				█										
7110	6150																							
7120	6160																							
7130	6170																							
rplL	7140	6180													█									
rplJ	7150	6190													█									
7160	6200																							
lysS	7170	6210		█											█									
7180	6220								█						█								█	
7190	6230								█						█							█	█	
clp	7200	6240																				█		
rpsP	7210	6250																				█		
trmD	7220	6260																				█		
rplS	7230	6270																				█		
fmt	7240	6280																				█		
7250	6290			█																		█		
ptsI	7260	6300																				█		
7270	6310			█																		█		
7280	6320			█																		█		
7290	6330																					█		
7300	6340																					█		
7310	6350																					█		
7320	6360																					█		
sgaB	7330	6370																				█	█	

Table S5 : Continued (20/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4	
	7340	6380	█	█					█										█			█		
	7350	6390	█	█					█											█			█	
	7360	6400	█	█					█											█				
	7370	6410	█	█					█											█				
	7380	6420	█	█					█											█				
	7390	6430	█	█					█											█				
	7400	6440	█	█					█											█				
	7410	6450	█	█					█											█				
	7420	NO	█	█					█											█				
	7430	NO	█	█					█											█				
	7440	NO	█	█					█											█				
	7450	NO	█	█					█											█				
	NO	6460	█	█					█											█				█
	7460	6470	█	█					█											█				█
	7470	6480	█	█					█											█				█
	NO	6490	█	█					█											█				█
	NO	6500	█	█					█											█				█
	7480	6510	█	█					█											█				█
	7490	6520	█	█					█											█				█
<i>gtsC</i>	7500	6530	█	█					█											█			█	█
<i>gtsB</i>	7510	6540	█	█					█											█				█
<i>gtsA</i>	7520	6550	█	█					█											█				█
<i>gtsA</i>	NO	6560	█	█					█											█				█
	7530	6570	█	█					█											█				█
	7540	6580	█	█					█											█				█
	7550	6590	█	█					█											█				█
<i>nadE</i>	7560	6600	█	█					█											█				█
	7570	6610	█	█					█											█				█
<i>lip</i>	7580	6620	█	█					█											█				█
	7590	6630	█	█					█											█				█
<i>dnal</i>	7600	6640	█	█					█											█				█
	7610	6650	█	█					█											█				█
	7620	6660	█	█					█											█				█
	7630	6670	█	█					█											█				█
	7650	6680	█	█					█											█				█
	7660	6690	█	█					█											█				█
	7670	6700	█	█					█											█				█
	NO	6710	█	█					█											█				█
	7680	6720	█	█					█											█				█
<i>rnhB</i>	7690	6730	█	█					█											█				█
	7700	6740	█	█					█											█				█

Table S5 : Continued (21/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	7710	6750																					
<i>chrA</i>	7720	6760																					
<i>chrA</i>	7730	6770																					
	7740	6780																					
	7750	6790																					
	7760	6800																					
	7770	6810																					
	7780	6820																					
	7790	6830																					
<i>dnaH</i>	7800	6840																					
<i>tmk</i>	7810	6850																					
<i>recR</i>	7820	6860																					
	7830	NO																					
<i>dnaX</i>	7840	6870																					
	7850	6880																					
<i>pgm</i>	7860	6890																					
	7870	6900																					
	7880	6910																					
	7900	6920																					
<i>pgsA</i>	7910	6930																					
<i>cdsA</i>	7920	6940																					
	7930	NO																					
	7940	6950																					
<i>apt</i>	7950	6960																					
<i>infB</i>	7960	6970																					
	7970	6980																					
<i>nusA</i>	7980	6990																					
	7990	7000																					
<i>tdk</i>	8000	7010																					
<i>pepA</i>	8010	7020																					
	8020	7030																					
	8030	7040																					
	8040	NO																					
	8050	NO																					
<i>vpmaV</i>	NO	7050																					
<i>vpmaW</i>	NO	7060																					
<i>vpmaX</i>	8060	7070																					
	8070	NO																					
	8090	NO																					
<i>vpmaY</i>	8170	7080																					
<i>vpmaU</i>	8100	7090																					

Table S5 : Continued (22/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
	8110	NO																					
	8120	NO																					
	8130	NO																					
	8140	NO																					
	8150	NO																					
	8160	NO																					
	8180	NO																					
	8190	NO																					
	8200	NO																					
	8210	NO																					
<i>vpmaZ</i>	8080	7100																					
<i>xer1</i>	8220	7110																					
	8230	NO																					
	8250	7120																					
	8260	7130																					
	8270	7140																					
	8280	7150																					
<i>gatB</i>	8290	7160																					
<i>gatA</i>	8300	7170																					
	8310	7180																					
<i>rluC</i>	8320	7190																					
<i>spcB</i>	8330	7200																					
<i>scpA</i>	8340	7210																					
<i>plsC</i>	8350	7220																					
<i>acpS</i>	8360	7230																					
<i>engA</i>	8370	7240																					
<i>cmk</i>	8380	7250																					
	8390	7260																					
	8400	7270																					
	8410	NO																					
<i>dnaJ</i>	8420	7280																					
<i>rbfA</i>	8440	7290																					
<i>tyrS</i>	8460	7300																					
	8470	7310																					
	8480	7320																					
	8500	7330																					
	8510	7340																					
<i>prfA</i>	8520	7350																					
	8530	7360																					
<i>gyrB</i>	8540	7370																					
<i>trxA</i>	8550	7380																					

Table S5 : Continued (23/23).

Name	MAGa	MAG	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4	
<i>smpB</i>	8560	7390						■		■	■													■
	8570	7400						■		■	■													
	8580	7410						■		■	■													
<i>mgtA</i>	8600	7420						■		■	■													
	8610	7430						■		■	■													
	8620	7440						■		■	■													
<i>ftsH</i>	8630	7450						■		■	■													
	8640	7460								■	■													■
<i>pth</i>	8650	7470								■	■													■
<i>recD</i>	8660	7480									■													■
<i>rpmJ</i>	NO	7530																						■

Name: gene name ; MAGa: locus tag 5632 ; MAG: locus tag PG2 ; NO: no ortholog.

Black square: 5632 gene; White square: PG2 gene; Grey square: chimeric PG2/5632 gene.

Table S6 : Features of micro-complex regions detected *M. agalactiae* transconjugants and hybrids

	T1-1	T1-2	T2	T3	T4-1	T4-2	T4-3	T5-1	T5-2	T5-3	T5-4	T5-5	T6-1	T6-4	T6-5	H1-1	H1-2	H4-1	H4-2	H4-3	H4-4
Fragments number	3	10	7	8	4	6	12	15	25	30	19	16	13	30	10	6	8	1	8	10	12
Fragments defined by ≤10 donor SNPs	1	3	4	4	1	0	4	4	3	8	5	4	2	7	1	1	1	0	3	5	2
Donor fragments separated by ≤10 recipient SNPs	0	0	1	1	1	1	2	3	11	7	7	5	3	13	2	1	2	0	2	4	2
Total of micro-variations	1	3	5	5	2	1	6	7	14	15	12	9	5	20	3	2	3	0	5	9	4
% micro-variations/ fragments number	14.3	14.3	33.3	29.4	22.2	7.7	24.0	22.6	27.5	24.6	30.8	27.3	18.5	32.8	14.3	15.4	17.6	0.0	29.4	42.9	16.0
Average of % micro-variations	20.7 ± 9.1							24.8 ± 6.2							20.2 ± 14.0						

Table S7 : List of SNPs detected in micro-complex region ≤ 5 nt and their impact in amino acid sequences

	Micro-variations number	Mnemonic	sequence in 5632 nt (AA)	sequence in PG2 nt (AA)	non silencious mutations
T1-1	3	MAG6100	GGT(G) GAT(D) GTG(V)	GGG(G) GAA (E) GTT(V)	X
T1-2	7		AAC(ATTTAAC)AGGAGGCAAT	AAC()AGGAGGCAAT	
T2	1		CTT (L)	CTG(L)	
	3		GAT(D) CTA(L) GAT(I)	GGT(G) TTA(L) AAT(N)	X
T3	3	MAG2120	GCC(A) GCTG(V) GCT(A)	GCT(A) GTA(V) CCT(P)	X
	5	MAG3430	GCA(A) ; ACA(T) GCA(A) ATAACA (IT)	GAA(E) ACT(T) GCT(A) ATTGCA (IA)	X
	4	MAG3630	TTG(L) AGT(S) TAC(Y) AAT(N)	TAG(L) AGA(R) TAT(Y) GAT(D)	X
	2		(C)9 TAAAC	(C)10 TTA AAC	
T4-1	1	MAG0650	GCA(A)	ACA(T)	X
T4-2	1	MAG7370	GAC(D)	AAC(N)	X
T4-3	1	MAG2340	CAG(A)	CAA(T)	X
	2	MAG2680	ACG(V) TCG(E)	CGG(G) TGG(W)	X
	5	MAG6210	CTA TGG TTA CGG TGT AGG (IGICGGT)	TTA TGG CTA CGG GTT GGC (IGICGGT)	
	5	MAG6900	GTA(V) GAC(D) GAT(D) AAA(K) GTT(V)	GTC(V) GGC(G) GAC(D) AAG (K) GTC(V)	X
	2	MAG6210	GCCAAT (AN)	GCAAAC (AN)	
T5-1	5	MAG1120	ATG(M) GGA(G) GAA(E) TTA(L) AGT(S)	ATT(I) GGC(G) GAG(E) CTA(L) AGC(S)	X
	3	MAG1660	AAG(K) CCA(P) GGG(G)	AAA(K) CCT(P) GGA(G)	
	2	MAG1780	ACT(T) TCT(S)	ATT(I) ACT(T)	X
	3	MAG1120	AAC(N) TTA(L) AAC(N)	AGC(S) CTA(L) AAT(N)	X
T5-2	4	MAG 3260	GGC(G) AGT(S) AGT(S) AGT(S)	GGA(G)- AAT(N)- AAT(N)- AGC(S)	X
	3	MAG 6240	TCG(S) TTG(L) GCC(A)	TCA(S) TTA(L) TCC(S)	X

Table S7 : Continued (2/4).

	Micro-variations number	Mnemonic	sequence in 5632 nt (AA)	sequence in PG2 nt (AA)	non silencing mutations
T5-2	3	MAG 6250	AAT(N) TAT(Y) ATT(I)	AAC(N) TAC(Y) CTT(L)	X
	3	MAG1980	CTA(L) GTC(V) TAC(Y)	TTA(L) GTT(V) TAT(Y)	
	1	MAG3130	AAG(K)	AAA(K)	
	1	MAG3260	GTA(V)	TTA(L)	X
	1	MAG3260	ACG(T)	ACT(T)	
	2	MAG3270	AAA(K) ACA(T)	AAG(K) ATA(I)	X
	4	MAG 6210	CTT(L) ACC(T) GTA(D) CTG(L)	CTC(L) ACA(T) GAC(D) TTG(L)	
	2	MAG6230	GTT(V) -ACC(T)	GTC(V) ACG(T)	
	1	MAG6240	CCA(P)	CCT(P)	
3	MAG7440	AGT(S) GAA (E)	AGC(S) GCG(A)	X	
T5-3	1	MAG1670	TTC(F)	TAC(Y)-	X
	5	MAG 1840	ACT(T) TGT(C) GGA(G) AGA(R) ATT(I)	ACA(T) TGC(C) GGT(G) AAA(K) ATC(I)	X
	3	MAG 3570	AAT(N) AGT(S) GCT(A)	AAC(N) AAT(N) GTT(V)	X
	3	MAG 4920	ATC(I) TCG(S) GAT(D)	ATT(I) TCT(S)- GAC(D)	
	5	MAG 5150	ATC(I) GCA(A) TCC(S) CTA(L) CAT(H)	AAC(N) GTA(V) TCA(S) TTA(L) CGT(R)	X
	1	MAG1670	TTC(F)	TTT(F)	
	3	NCR	GTC(V) GTG(V)	GTA(V) GTA(V)	
	5	MAG 5080	GAA(E) TAT(Y) GAT(D) AGT(S) AAT(N)	GAG(E) TAC(Y) GAC(D) AGC(S) AAC(N)	
2	MAG5260				
T5-4	1	MAG1370	GAC(D)	GAT(D)	
	2	MAG 1380	AAA(K) GGC(G)	AAT(N)GGT(G)	X
	1	MAG1430	CTA(L)	TTA(L)	
	3	MAG 2250	TCA(S) AAA(K) TTG(L)	TCT(S) AAG(K) TTA(L)	

Table S7 : Continued (3/4).

	Micro-variations number	Mnemonic	sequence in 5632 nt (AA)	sequence in PG2 nt (AA)	non silencing mutations
T5-4	1	MAG1430	CTA(L)	TTA(L)	
	4	MAG 1380	CCC(P) GCA(A) ACC(T) TGC(C)	CCT(P) GCT(A) ACT(T) TGT(C)	
	1	MAG5920	ACT(T)	ACA(T)	
	3	MAG 5960	TCA(S) AAC(N) CCT(P)	TCT(S) AAT(N) CCA(P)	
T5-5	3	MAG 1360	TTA(L) CTA(L) TAT(Y)	TTG(L) TTA(L) TAC(Y)	
	5	MAG 1580	GGC(G) GCA(A) ATC(I) ATC(I) CGC(R)	GGT(G) GCC(A) ATT(I) ATA(I) CGT(R)	
	3	MAG 4830	AGG(R) AAA(K) GAC(D)	AGA(R) AAG(K) GAT(D)	
	3	MAG 1360	AGC(S) GAA(E) TTA(L)	AGT(S) GAG(E) CTA(L)	
	3	MAG 4230	CCT(P) GTT(V)	TCA(S)) ATT(I)	X
	2	MAG 4240	GTT(V) AGT(S)	CTT(L) AGC(S)	X
	1	MAG4480	AAA(K)	AAC(N)	X
T6-1	4	MAG 1540	GCT(A) GGC(G) GAG(E) GGT(G)	GCC(A) GGT(G) GAA(E) GGC(G)	
	3	MAG 4600	AAC(N) GCA(A) AAT(N)	AAT(N) GCT(A) AAC(N)	
	1	MAG4480	TCC(S)	TCT(S)	
	1	MAG6210	ACC(T)	ACA(T)	
T6-4	1	MAG0600	GAG(E)	GAA (E)	
	5	MAG 0650	GTT(V) CTA(L) GAT(D) AGG(R) TCA(S)	GTG(V) CTG(L) GAC(D) CGG(R) TCT(S)	
	1	MAG5050	TTG(L)	TTA(L)	
	4	MAG0500	AAA(K) GTG(V) CTA(L) ATA(I)	AAG(K) GTA(V) CAA(Q) ATC(I)	X
	1	MAG0500	ACT(T)	ACA(T)	
	5	MAG 0510/20	GGT(G) GAG(E) CCT(P) TTA(L) ATT(I)	GGC(G) GAA(E) CTT(L) TCA(S) ATC(I)	X
	1	MAG0570	CTA(L)	TTA(L)	
	5	MAG 0590	GCT(A) GTA(V) GAT(D) TTT(F) ATT(I)	GCC(A) GTG(V) GAC(D) TTC(F) ATA(I)	

Table S7 : Continued (4/4).

	Micro-variations number	Mnemonic	sequence in 5632 nt (AA)	sequence in PG2 nt (AA)	non silencious mutations
T6-4	4	MAG 0600	AAT(N) AAG(K) TTA(L) GGT(G)	AAC(N) AAA(K) CTA(L) GGC(G)	
	2	MAG0650	GAT(D) ATT(I)	GAC(D) GTT(V)	X
	4	MAG 0650	GAT(D) CTG(L) AAC(N) CCT(P)	GAC(D) CTT(L) AAT(N) CCC(P)	
	1	MAG0650	CCA(P)	TCA(S)	X
	1	MAG0650	CAT(H)	CAC(H)	
	4	MAG 1470	AGT(S) AAA(K) TTG(L)	ACT(T) AGG(R) TTA(L)	X
	1	MAG1880	TCA(S)	TCT(S)	
T6-5	1	MAG5610	GAC(D)	AAC(N)	X
H1-1	2	MAG4880	GGc(G) CAA(Q)	CAG(Q) GGT(G)	
H1-2	4	MAG6380	ACT(L) TCG(K) TTA(L) ACC(T)	GCT(L) TTG(K) CTA(L) ATC(I)	X
H4-1					
H4-2	1				
	4	MAG6760	GGT(G) CAA(V) AAA(K) GGT(G)	GGC(G) TAA(I) AAG(K) GGC(G)	X
	2		GCC	ACT	
	5	MAG6760	1/5 non silent (I) ->(T)		X
H4-3	3	MAG6680	GCC(A)	ACC(T)	X
	1	MAG6370	GCA(R)	GCG(R)	
	2	MAG6380	ACT(T) GTT(V)	ACC(T) GTC(V)	
	1	MAG6110	ATC(I)	ATA(I)	
	2	MAG6380	TCA(S)	CCA(P)	X
H4-4	1	MAG0250	GTT(V)	GTG(V)	
	3	MAG6370	AAG(K) CAT(H) GAG (E)	AAA(K) CAC(H) GAA (E)	

NCR: non-codin region

Table S8 : Sizes of the Homologous Region (HR) surrounding the PG2-inherited DNA fragments

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
T1-1	642190..671308	29118	53	45
	710826..710875	49	5	6
	712514..712718	204	55	40
T1-2	68369..68496	127	55	272
	87422..87723	301	57	46
	94998..122105	27107	319	88
	565195..576353	11158	22	211
	597640..601994	4354	10	39
	645869..693639	47770	58	2
	710826..710881	55	5	5
	729662..736494	6832	81	43
	747140..747425	285	5	10
	749929..781889	31960	16	10
T2	72993..73064	71	72	1
	391065..391135	70	33	15
	405624..405777	153	36	49
	406283..419098	12815	97	49
	419274..478430	59156	82	32
	478433..499977	21544	13	56
	659240..659534	294	11	88
	T3	93477..158150	64673	46
163425..169371		5946	25	4
215665..244325		28660	7	48
244426..244887		461	105	79
407088..407295		207	63	63
429886..430228		342	78	117
672936..684212		11276	10	115
711547..711639		92	147	7
T4-1	67546..76591	9045	58	80
	76595..116611	40016	30	52
	319305..319769	464	58	31
	766673..777472	10799	70	112
T4-2	48711..61077	12366	28	26
	717227..747048	29821	109	58
	761528..774557	13029	5	64
	783745..788959	5214	25	6
	852781..855009	2228	30	13
	855013..863470	8457	42	85

Table S8 : Continued (2/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
T4-3	260981..264175	3194	175	65
	272064..272251	187	98	92
	313107..313449	342	89	392
	317271..318857	1586	136	31
	395947..410712	14765	109	161
	735501..735732	231	72	58
	735739..739732	3993	52	52
	739931..742146	2215	29	13
	752375..774488	22113	47	61
	785922..788502	2580	127	48
	806563..806723	160	55	40
	808138..816278	8140	13	64
	T5-1	112619..126202	13584	87
126730..126922		193	47	31
127204..139695		12492	44	23
194107..194287		181	106	62
202460..203985		1526	14	40
204096..205478		1383	48	50
208524..208716		193	181	6
209121..209456		336	71	54
212153..219028		6876	37	45
219089..221717		2629	5	17
405780..411811		6032	47	27
428422..429348		927	5	75
845464..858782		13319	119	29
864308..864689		382	20	35
872773..874910		2138	98	50
T5-2	864483..2456	15412	34	28
	197548..205010	7463	110	71
	205899..214493	8595	23	2
	214771..219159	4389	8	23
	223527..229242	5715	52	25
	229707..232554	2848	220	69
	232582..233313	732	32	36
	239748..240137	390	11	32
	240232..251588	11357	4	35
	251797..256693	4897	143	8
	357939..359527	1589	51	74
	360699..366567	5869	150	86
	369033..376334	7302	176	14

Table S8 : Continued (3/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'	
T5-2	376336..389363	13028	35	18	
	389365..389788	424	67	200	
	389790..390080	291	38	16	
	390093..390410	318	28	6	
	730534..735880	5347	28	50	
	735965..738576	2612	10	14	
	738614..741446	2833	86	146	
	741448..741659	212	53	28	
	742574..742980	407	95	217	
	743090..743469	380	44	23	
	832037..832796	760	12	46	
	854487..864471	9985	26	83	
	T5-3	162565..163125	561	35	50
		164142..168460	4319	50	20
170646..171679		1034	38	113	
174679..176046		1368	18	56	
176333..179471		3139	20	92	
180660..194851		14192	141	3	
194853..195000		148	49	98	
196797..197221		425	112	25	
197449..216542		19094	17	113	
216551..217009		459	111	20	
402127..402273		147	44	10	
402498..420677		18180	20	38	
421269..421417		149	65	60	
568575..573366		4792	23	47	
573514..573673		160	36	36	
582090..583899		1810	65	140	
592247..594983		2737	89	98	
595138..595394		257	38	50	
596978..598178		1201	41	82	
598756..600277		1522	107	8	
602679..602920		242	103	42	
603139..607486		4348	35	117	
609130..610308		1179	46	128	
610313..612825		2513	119	nd in 16S rrn	
622105..622600		496	11	11	
624195..624900		706	74	59	
625273..629013		3741	17	11	
634893..635782	890	121	185		

Table S8 : Continued (4/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
T5-3	639399..648184	8786	110	56
	648748..650719	1972	14	62
	153842..154491	650	24	156
	154860..157677	2818	51	152
	157679..157859	181	125	55
	157951..158731	781	36	210
	159115..159317	203	158	41
	159406..162262	2857	61	75
	162418..163239	822	59	46
	164140..164271	132	50	79
	196383..201570	5188	92	16
T5-4	207461..209267	1807	14	75
	209828..217126	7299	125	62
	217663..218211	549	8	15
	236781..254899	18119	8	94
	261870..262042	173	110	23
	636013..656766	20754	55	41
	685748..686041	294	93	14
	686043..690793	4751	8	68
	690810..691256	447	53	18
	691755..692215	461	44	17
T5-5	149525..152914	3390	86	28
	153985..154322	338	72	50
	154328..154619	292	7	8
	167225..185040	17816	309	25
	185132..185372	241	3	15
	191447..222596	31150	32	15
	247840..248268	429	10	33
	477340..477978	639	65	131
	489499..496909	7411	38	70
	496916..498565	1650	6	76
	498572..498765	194	19	23
	518506..523954	5449	67	29
	523956..553108	29153	35	2
	564697..564967	271	201	45
	592736..608001	15266	98	367
637720..654041	16322	14	50	
	179380..179564	185	92	62
T6-1	189305..217670	28366	83	8
	340329..340615	287	47	14

Table S8 : Continued (5/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
T6-1	539780..540078	299	35	60
	540134..550984	10851	41	59
	568872..569193	322	17	34
	620460..673888	53429	365	62
	701392..701753	362	38	41
	701755..702008	254	26	27
	725432..730838	5407	101	65
	733819..734519	701	283	136
	735499..735904	406	73	23
	735906..754200	18295	53	155
T6-4	56245..60268	4024	10	44
	60363..60793	431	86	20
	60795..61519	725	63	63
	61705..67262	5558	17	53
	67264..69337	2074	6	85
	69600..71805	2206	38	92
	71948..72156	209	143	65
	73296..74545	1250	74	191
	75183..76906	1724	45	176
	76909..77524	616	58	101
	77541..77741	201	23	93
	77743..77911	169	28	5
	77913..78154	242	10	53
	80209..81450	1242	32	167
	85651..85955	305	56	25
	89106..89629	524	141	173
	91208..109804	18597	65	74
	139449..157859	18411	28	55
	158166..169753	11588	94	13
	169784..170459	676	31	136
	192405..194059	1655	53	212
	212497..221735	9239	26	17
	221737..221950	214	11	87
	571119..575909	4791	50	53
	577096..577916	821	17	5
	591778..591968	191	137	53
	601346..604442	3097	41	47
	606056..607486	1431	80	117
628732..650719	21988	270	62	
650777..684449	33673	66	99	

Table S8 : Continued (6/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
T6-5	180229..191725	11497	48	32
	196660..229384	32725	82	10
	235416..241004	5589	26	47
	245965..246041	77	11	30
	311450..316074	4625	79	17
	316654..331060	14407	26	59
	360109..365569	5461	17	25
	639714..643343	3630	110	262
	643345..646363	3019	102	83
	646656..653943	7288	194	218
H1-1	248848..249190	342	146	5
	567873..569135	1262	4	15
	569144..586446	17302	16	40
	590974..594102	3128	31	109
	640879..641595	716	44	23
	648561..656525	7964	18	40
H1-2	597469..603312	5843	19	267
	635033..642973	7940	37	5
	645546..660978	15432	155	26
	661135..672298	11163	37	30
	672558..683636	11078	184	53
	749838..753167	3329	64	70
	753413..754906	1493	41	180
	755517..755917	400	145	141
H4-1	756585..770171	13586	85	19
H4-2	369689..369733	44	115	153
	710826..710880	54	157	153
	762357..763035	678	5	5
	763038..770652	7614	4	6
	771680..788005	16325	12	42
	792443..792866	423	34	98
	792987..793455	468	5	61
	793461..793741	280	117	64
H4-3	655680..655829	149	15	394
	710826..710871	45	5	6
	723203..723364	161	37	13
	723371..724478	1107	7	274

Table S8 : Continued (7/7).

	Transferred fragment Position in PG2	Fragments size (nt)	Size of HR in 5'	Size of HR in 3'
H4-3	739801..746569	6768	54	53
	746809..753839	7030	58	34
	754016..754193	177	28	154
	754264..754487	223	65	109
	756683..757151	468	76	109
	769024..772403	3379	58	22
H4-4	14614..25285	10671	24	13
	25430..25621	191	106	1
	737201..742114	4913	31	85
	748530..750872	2342	4	22
	753808..753995	187	34	90
	762736..763035	299	88	55
	763039..770262	7223	12	6
	837008..838839	1831	37	79
	839000..854241	15241	91	31
	855954..857172	1218	126	21
	873516..11656	15577	39	64

Table S9 : Comparative genome analysis of *M. agalactiae* 5632 and 4867 strains.

Evidence	Position	Mutation	Annotation	Gene	Description
MC JC	23,536	Δ1,531 bp		<i>MAGa0240</i> –[<i>MAGa0250</i>]	<i>MAGa0240</i> , [<i>MAGa0250</i>]
RA	74,586	C→G	Q254E (CAA→GAA)	<i>lip</i> →	Esterase/lipase
RA	138,128	A→G	K131K (AAA→AAG)	<i>MAGa1270</i> →	Hypothetical protein
RA	182,855	Δ1 bp	coding (345/1659 nt)	<i>MAGa1580</i> ←	Type III restriction modification system:Methylase
RA	182,856	Δ1 bp	coding (344/1659 nt)	<i>MAGa1580</i> ←	Type III restriction modification system:Methylase
RA	224,017	Δ1 bp	coding (76/963 nt)	<i>MAGa1890</i> ←	Hypothetical protein, predicted lipoprotein
RA	225,375:1	+G	pseudogene (26/927 nt)	<i>MAGa1900</i> ←	Pseudogene of hypothetical protein
MC JC	248,573	Δ1,531 bp		[<i>MAGa2150</i>]–[<i>MAGa2170</i>]	[<i>MAGa2150</i>], <i>MAGa2160</i> , [<i>MAGa2170</i>]
RA	314,082	C→T	G392D (GGC→GAC)	<i>rpoD</i> ←	DNA directed RNA polymerase sigma factor(Sigma A)
JC	338,486	+ATATATAT	intergenic (+86/-110)	<i>ligA</i> → / → <i>MAGa2960</i>	DNA ligase/Putative transmembrane protein
MC JC	343,469	Δ27,512 bp		<i>MAGa2980</i> – <i>MAGa3220</i>	24 genes
RA	380,686	T→C	L106S (TTA→TCA)	<i>MAGa3280</i> →	Conserved hypothetical protein
RA	387,221:1	+C	coding (253/285 nt)	<i>MAGa3330</i> →	Hypothetical protein, predicted lipoprotein
RA	432,171:1	+G	pseudogene (43/1158 nt)	<i>MAGa3760</i> →	Pseudogene of conserved hypothetical protein, DUF285 family (C-terminal part)
MC JC	453,253	Δ1,531 bp		<i>MAGa3960</i> –[<i>MAGa3970</i>]	<i>IS 30 like transposase family</i>
RA	464,106	T→G	I10I (ATA→ATC)	<i>MAGa4040</i> ←	Conserved hypothetical protein, truncated inC terminal
RA	467,845	A→G	S181P (TCA→CCA)	<i>MAGa4080</i> ←	Conserved hypothetical protein
RA	507,574:1	+G	intergenic (+147/-31)	<i>deoD</i> → / → <i>MAGa4430</i>	Purine nucleoside phosphorylase (Inosinephosphorylase) (PNP)/Hypothetical protein
RA	507,583	2 bp→TC	intergenic (+156/-21)	<i>deoD</i> → / → <i>MAGa4430</i>	Purine nucleoside phosphorylase (Inosinephosphorylase) (PNP)/Hypothetical protein
JC	519,657	+AGAGAGAGA GAGAGAGAG	intergenic (-91/+226)	<i>MAGa4500</i> ← / ← <i>adhT</i>	Hypothetical protein, predicated lipoprotein/Alcohol dehydrogenase
RA	521,984	Δ1 bp	coding (326/951 nt)	<i>MAGa4520</i> ←	Hypothetical protein, predicted lipoprotein
RA	523,643:1	+G	intergenic (-20/+474)	<i>MAGa4530</i> ← / ← <i>MAGa4540</i>	Hypothetical protein, DUF285 family/Conserved hypothetical protein, predictedlipoprotein, DUF285 family
JC	525,385	Δ8 bp	intergenic (-75/+320)	<i>MAGa4540</i> ← / ← <i>MAGa4550</i>	Conserved hypothetical protein, predictedlipoprotein, DUF285 family/Hypothetical protein, predicted lipoprotein
MC JC	560,518	Δ27,512 bp		<i>MAGa4850</i> – <i>MAGa5060</i>	22 genes

Table S9 : Continued (2/3).

Evidence	Position	Mutation	Annotation	Gene	Description
RA	588,703	T→A	L169M (TTG→ATG)	<i>MAGa5070</i> →	ABC transporter, permease protein
RA	588,815	G→A	R206Q (CGA→CAA)	<i>MAGa5070</i> →	ABC transporter, permease protein
RA	591,019	G→A	S58N (AGT→AAT)	<i>MAGa5090</i> →	Conserved hypothetical protein
RA	616,235	Δ1 bp	coding (253/1056 nt)	<i>MAGa5320</i> →	unnamed protein product
MC JC	617,021	Δ1,531 bp		[<i>MAGa5320</i>]-[<i>MAGa5340</i>]	[<i>MAGa5320</i>], <i>MAGa5330</i> , [<i>MAGa5340</i>]
RA	683,476	T→A	V12V (GTA→GTT)	<i>vpmaD2</i> ←	Variable surface lipoprotein D2 (VpmaD2)
RA	683,497	T→C	K5K (AAA→AAG)	<i>vpmaD2</i> ←	Variable surface lipoprotein D2 (VpmaD2)
RA	683,544	G→A	intergenic (-33/-249)	<i>vpmaD2</i> ← / → <i>xer1</i>	Variable surface lipoprotein D2 (VpmaD2)/Integrase recombinase
RA	696,281	2 bp→CG	coding (484-485/903 nt)	<i>rplD</i> ←	50S ribosomal protein L4
RA	730,547	Δ1 bp	intergenic (+300/-34)	<i>MAGa6360</i> → / → <i>MAGa6370</i>	Hypothetical protein, predicted lipoprotein/unnamed protein product
RA	767,905	Δ1 bp	coding (895/2559 nt)	<i>MAGa6670</i> ←	Hypothetical protein, predicted lipoprotein
RA	773,201	Δ1 bp	intergenic (-61/+81)	<i>MAGa6720</i> ← / ← <i>MAGa6730</i>	Hypothetical product, predicted lipoprotein/Hypothetical protein, predicted lipoprotein
JC	774,999	Δ3 bp	intergenic (-56/+82)	<i>MAGa6730</i> ← / ← <i>MAGa6740</i>	Hypothetical protein, predicted lipoprotein/Hypothetical protein
RA	779,817:1	+C	intergenic (-51/+276)	<i>MAGa6760</i> ← / ← <i>MAGa6770</i>	Hypothetical protein/Hypothetical protein, predicted lipoprotein
JC	788,609	Δ3 bp	intergenic (-46/+198)	<i>MAGa6790</i> ← / ← <i>MAGa6800</i>	Hypothetical protein/Hypothetical protein
RA	789,920	C→T	G84G (GGG→GGA)	<i>MAGa6810</i> ←	Aminopeptidase
RA	794,160	Δ1 bp	intergenic (-44/+349)	<i>MAGa6820</i> ← / ← <i>MAGa6830</i>	Hypothetical protein, predicted lipoprotein/Conserved hypothetical protein
RA	805,470	G→A	P546S (CCT→TCT)	<i>rpoC</i> ←	DNA directed RNA polymerase beta' chain
RA	810,268	A→G	C156R (TGT→CGT)	<i>rpoB</i> ←	DNA directed RNA polymerase beta chain
RA	865,591	G→A	A191V (GCA→GTA)	<i>MAGa7340</i> ←	Transport protein SGAT
RA	868,51	A→G	intergenic (+30/+143)	<i>MAGa7360</i> → / ← <i>MAGa7370</i>	Conserved hypothetical protein/Conserved hypothetical protein
MC JC	926,014	Δ1,531 bp		<i>MAGa7930</i> -[<i>MAGa7940</i>]	<i>MAGa7930</i> , [<i>MAGa7940</i>]
RA	942,901	A→T	L9L (CTT→CTA)	<i>vpmaE</i> ←	Variable surface lipoprotein E (VpmaE)
RA	942,904	T→C	L8L (TTA→TTG)	<i>vpmaE</i> ←	Variable surface lipoprotein E (VpmaE)
RA	942,952	A→G	intergenic (-25/-790)	<i>vpmaE</i> ← / → <i>vpmaB</i>	Variable surface lipoprotein E (VpmaE)/Variable surface lipoprotein B (VpmaB)

Table S9 : Continued (3/3).

Evidence	Position	Mutation	Annotation	Gene	Description
RA	942,958	T→C	intergenic (-31/-784)	<i>vpmaE</i> ← / → <i>vpmaB</i>	Variable surface lipoprotein E (VpmaE)/Variable surface lipoprotein B (VpmaB)
RA	942,96	A→G	intergenic (-33/-782)	<i>vpmaE</i> ← / → <i>vpmaB</i>	Variable surface lipoprotein E (VpmaE)/Variable surface lipoprotein B (VpmaB)
RA	943,721	T→C	intergenic (-794/-21)	<i>vpmaE</i> ← / → <i>vpmaB</i>	Variable surface lipoprotein E (VpmaE)/Variable surface lipoprotein B (VpmaB)
RA	943,774	2 bp→AT	coding (33-34/711 nt)	<i>vpmaB</i> →	Variable surface lipoprotein B (VpmaB)
RA	943,777	G→A	L12L (CTG→CTA)	<i>vpmaB</i> →	Variable surface lipoprotein B (VpmaB)
RA	943,783	A→T	S14S (TCA→TCT)	<i>vpmaB</i> →	Variable surface lipoprotein B (VpmaB)
RA	949,117	Δ1 bp	coding (476/792 nt)	<i>vpmaI</i> →	Variable surface lipoprotein I (VpmaI)
RA	949,625	G→A	L8L (TTG→TTA) ‡	<i>vpmaF1</i> →	Variable surface lipoprotein F1 (VpmaF1)
RA	951,499	G→T	R178M (AGG→ATG)	<i>vpmaJ</i> →	Variable surface lipoprotein J (VpmaJ)

Unassigned missing coverage evidence

start	end	size	←reads	reads→	gene
674235–675679	675679	88 [0]	[43] 45	<i>MAGa5800</i>	<i>MAGa5800</i>
675807–675930	677647–675933	45 [43]	[43] 44	<i>vpmaA</i> –[<i>vpmaB</i>]	<i>vpmaA</i> , [<i>vpmaB</i>]
683633–686101	686114	46 [43]	[0] 89	<i>xer1</i> – <i>MAGa5890</i>	<i>xer1</i> , <i>MAGa5890</i>
819110–836334	840541–836336	46 [43]	[43] 44	[<i>MAGa6940</i>]– <i>MAGa7100</i>	17 genes [<i>MAGa6940</i>], <i>MAGa6950</i> , <i>MAGa6960</i> , <i>MAGa6970</i> , <i>MAGa6980</i> , <i>MAGa6990</i> , <i>MAGa7000</i> , <i>MAGa7010</i> , <i>MAGa7020</i> , <i>MAGa7030</i> , <i>MAGa7040</i> , <i>MAGa7050</i> , <i>MAGa7060</i> , <i>MAGa7070</i> , <i>MAGa7080</i> , <i>MAGa7090</i> , <i>MAGa7100</i>

CHAPITRE 3

CHAPITRE 3

Apport du séquençage NGS dans l'analyse de génomes de mycoplasmes.

Résumé

Grâce aux avancées technologiques en matière de séquençage NGS (*Next Generation Sequencing*), notamment l'augmentation du nombre de *reads* et la diminution des coûts, le séquençage de génomes complets est de plus en plus utilisé en microbiologie pour aborder un certain nombre de questions biologiques. En parallèle, de nombreux outils d'analyse de ces séquences ont été développés. Certains sont mieux adaptés que d'autres pour répondre au mieux aux questions posées, le choix de ces outils est ainsi une étape cruciale. Notre modèle d'étude *Mycoplasma agalactiae* a la capacité de générer des génomes mosaïques complexes via un transfert d'ADN chromosomique atypique, le MCT (*Mycoplasma Chromosomal Transfer*). Chacun de ces génomes est unique et constitué d'un mélange des génomes parentaux. Certains comportent jusqu'à 30 fragments acquis par transfert, dont la taille varie de quelques nucléotides à plusieurs dizaines de Kb. Afin de mieux comprendre le mécanisme du MCT, la reconstruction précise des génomes mosaïques est primordiale, mais constitue un défi. Ce travail de thèse s'appuie sur l'analyse d'un grand nombre de génomes mosaïques, et a nécessité la mise en place d'une méthode optimisée et automatisée de tri des *reads* de séquençage Illumina. Le séquençage Illumina a également été utilisé pour caractériser les variations chez des populations isogéniques obtenues en présence ou non de sélection antibiotique. Cette technique de séquençage donne accès à une profondeur de couverture importante qui peut présenter des variations. L'analyse de ces différences de couverture a notamment permis de révéler des informations importantes sur le gain, la perte ou la duplication de loci tel que les ICE ou le locus *ps3*.

1. Introduction

Le séquençage de génomes complets est de plus en plus utilisé comme base à la compréhension de nombreux mécanismes biologiques, par exemple ceux responsables de la résistance aux antibiotiques ou de l'évolution des populations bactériennes (Frese et al. 2013; Sharma et al. 2014; Ramanathan et al. 2017).

Nos études ont montré que les mycoplasmes ont la capacité de s'adapter par mutations spontanées ou par transfert horizontal de gènes (HGT), ce qui leur permet de faire face à une pression de sélection antibiotique (Chapitres 1 et 2). Les mutations sélectionnées à l'échelle de la population ont été détectées par des analyses de variants chez des mutants de *Mycoplasma agalactiae* spontanément résistants à l'enrofloxacin (Chapitre 1). Des transconjugants aux génomes mosaïques ont été obtenus *in vitro* chez *M. agalactiae* via un mécanisme atypique de HGT alliant conjugaison et recombinaison homologue (Dordet-Frisoni et al. 2014) (Chapitre 1 et 2). Ils correspondent à un réassemblage des génomes parentaux, avec le remplacement de certaines régions de la souche receveuse 5632 par celles de la souche donneuse PG2. Les fragments transférés sont nombreux et de taille très variable, la détermination précise des régions appartenant à chaque parent, et donc la "reconstruction" des génomes mosaïques, constitue un défi. Une méthode automatisée basée sur des alignements et des tris de *reads* Illumina a été mise en place. Elle s'appuie sur les différences de séquence entre les deux génomes parentaux. La comparaison des mécanismes d'adaptation (mutations vs transferts chromosomiques) a été abordée au chapitre 1 et une caractérisation plus fine du transfert chromosomique est décrite au chapitre 2. Ce 3^{ème} chapitre vient en complément des deux précédents, en expliquant et détaillant les éléments méthodologiques utilisés pour l'analyse des données issues du séquençage.

Les clones parentaux utilisés lors de nos expériences *in vitro* (PG2 55-5 pour les mutants spontanés, PG2^{E10} et 5632^{G3} pour les transconjugants) ont été (re)séquencés par Illumina, comme contrôles. La comparaison des résultats de séquençage avec les séquences initialement déterminées par Sanger (Sirand-Pugnet et al. 2007, Nouvel et al. 2010), a permis d'identifier des mutations ou des erreurs de séquençage. Les *reads* issus des clones parentaux ont aussi été utilisés pour une comparaison fine des variations existantes entre les souches PG2 et 5632, et ont servi de jeux d'entraînements dans la mise en place de nos méthodes d'analyses. L'alignement des *reads* sur un génome de référence, analyse qui précède l'analyse de variants, a révélé des informations importantes comme la duplication d'une région ou la

variation d'éléments génétiques mobiles. Ces résultats complémentaires non décrits dans les précédents chapitres sont également présentés ici.

Au total, au cours de cette thèse, 52 génomes de *M. agalactiae* ont été entièrement séquencés par Illumina (HiSeq, *paired-end* 2x150 nt). Pour chaque génome, 2.10^7 reads ont été obtenus en moyenne, avec une profondeur de couverture supérieure à 2700 X.

2. Choix des méthodes de séquençage et d'analyse

2.1. Le séquençage NGS

Le principe du séquençage NGS

Le séquençage de dernière génération (NGS, *Next Generation Sequencing*) utilise différentes technologies développées depuis 2005. Aussi appelé séquençage à haut débit (*High Throughput Sequencing*), il permet de générer un très grand nombre de données de séquences (jusqu'à plus de 1000 Gb). Les diverses techniques se distinguent par la méthode d'amplification, le type de réaction de séquençage, la taille des lectures ou *reads* obtenus, et présentent des biais et taux d'erreurs variables (Tableau 1). Ces techniques permettent le séquençage de tout type d'acides nucléiques (ADN : génomique, ARN : transcriptomique) provenant de tout organisme (Homme, animaux, plantes, végétaux, bactéries, virus...). Le séquençage peut être ciblé sur une région d'intérêt ou s'intéresser au génome entier (WGS, *Whole-Genome Sequencing*). En plus de la séquence nucléique, certaines techniques comme le PacBio permettent aussi de connaître les états de méthylation de l'ADN (méthylome).

Tableau 1 : Comparaison des technologies de NGS.

	Sanger	Ion torrent Life technologies	454 Roche	HiSeq/Miseq Illumina	PacBio Pacific Bioscience
Méthode d'amplification	PCR classique	PCR en émulsion	PCR en émulsion	Bridge amplification	Lecture d'une seule molécule en temps réel
Méthode de séquençage	Séquençage par synthèse	Détection de proton	Pyro-séquençage	Séquençage par synthèse	Séquençage par synthèse
Taille des reads	400-900 pb	200 pb	700 pb	50-250 pb (paired-end)	5-30 Kb
Taux d'erreur	0,01 %	2 %	0,1 %	0,1 %	13 %
Nombre de reads par run	1	5.10 ⁶	1.10 ⁶	3.10 ⁹	5.10 ⁴
Durée d'un run	30min-3h	2-4h	24h	2-10 jours	30min-2h
Avantage	Reads longs et haute-fidélité	Rapidité	Longs reads	Nombre important de reads	Reads très longs et rapidité
Inconvénient	Une seule lecture	Erreurs homopolymères	Erreurs homopolymères	Reads courts et run long	Taux d'erreur très élevé
Nombre de Gb	2-80	2	1	1000	1
Cout par Gb (\$)	2400000	500	10000	30	600

D'après (Pareek *et al.* 2011; Liu *et al.* 2012; Escobar-Zepeda *et al.* 2015).

Le séquençage NGS se compose de 3 grandes étapes :

- Préparation de la librairie : l'acide nucléique est fragmenté aléatoirement, par sonication ou digestion enzymatique, puis filtré selon sa taille. Des adaptateurs sont fixés à chaque extrémité pour une amplification de l'ensemble des fragments par PCR. Un marquage spécifique de différents génomes ou ADN, appelé multiplexage, est possible à cette étape. Il permet d'amplifier et de séquencer différents ADN ensemble au cours d'une même réaction de séquence tout en les différenciant par la suite lors de l'analyse.

- Réaction de séquençage : plusieurs réactions de séquence, appelé *run* se font simultanément avec l'obtention de *reads*, ou lectures, qui correspondent à la séquence de chaque fragment.

- Analyse des données brutes : Les données obtenues sont des fichiers texte au format fastq, ils rassemblent un très grand nombre de données en peu d'espace mémoire. Dans ce type de fichier chaque *read* est associé à 4 lignes (un identifiant de séquence, la séquence nucléotidique, la qualité de chaque nucléotide et une ligne contenant un +, avec parfois le nom de la séquence).

Deux paramètres sont habituellement calculés pour évaluer l'étendue des séquences obtenues : la profondeur et l'étendue de la couverture (Figure 1) (Sims *et al.* 2014). La profondeur de la couverture, notée X , est définie par le nombre de *reads* présents pour chaque base séquencée. La profondeur moyenne obtenue pour une région donnée correspond

à :

$$\frac{L \times N}{G}$$

Où L = longueur des *reads* ; N = nombre de *reads* pour cette région et G = taille de la région.

L'étendue de la couverture se définit comme la zone couverte par un nombre défini de *reads*, elle est exprimée en pourcentage de la région d'intérêt.

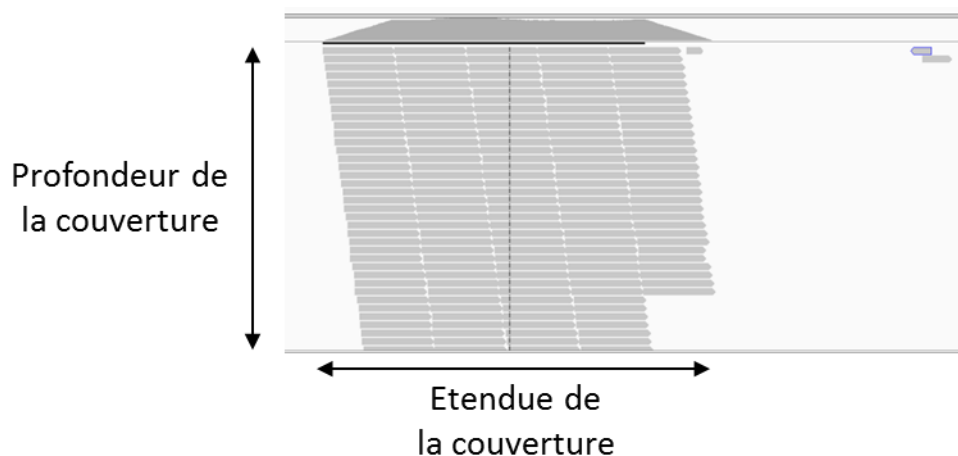


Figure 1 : Profondeur et étendue de la couverture d'un alignement

Visualisation d'un alignement de *reads* alignés sur une séquence de référence. Le nombre de *reads* empilés définit la profondeur de la couverture pour une région donnée. La région couverte par un nombre minimal de *reads* correspond à l'étendue de la couverture.

La technique Illumina

Parmi les différentes techniques disponibles (Tableau 1), la technique Illumina est celle dont l'usage est le plus répandu à ce jour. L'Illumina permet d'obtenir un très grand nombre de *reads*, la profondeur obtenue peut ainsi donner accès aux polymorphismes présents dans des sous-populations de l'échantillon séquencé. Pour répondre à des questions d'adaptation et étudier la survenue et l'évolution de mutations au sein de populations bactérienne, l'Illumina constitue une technique de choix, c'est pour ces raisons que notre travail y a principalement fait appel.

En Illumina (Figure 2) la PCR est dite "en *bridge*" (en pont) : chaque brin d'ADN dénaturé est fixé à une amorce immobilisée sur un support et les brins complémentaires sont synthétisés par PCR. Chaque brin fixé forme un pont en s'hybridant localement avec l'amorce complémentaire de l'autre extrémité. L'opération est répétée un grand nombre de fois pour former des *clusters*, ou regroupements, d'un même fragment d'ADN. Le séquençage se fait ensuite par synthèse : le brin d'ADN est lu par une polymérase qui ajoute les nucléosides triphosphates (NTPs) correspondants marqués par 4 fluorochromes différents. Bien que ces technologies NGS donnent accès à un très grand nombre de *reads*, il existe des biais de PCR ou d'amplification et des biais de séquençage dus au taux d'erreurs élevé de la polymérase.

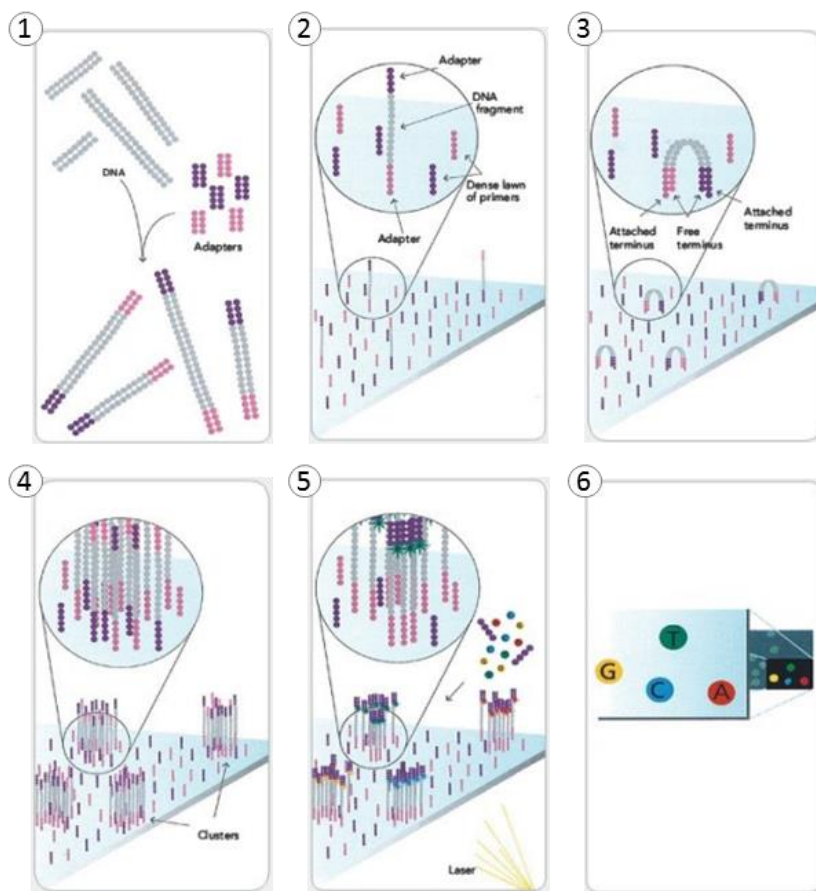


Figure 2 : Etapes d'un run de séquençage par la technique Illumina.

- 1) Fragmentation de l'ADN puis fixation d'adaptateurs aux 2 extrémités.
 - 2) Fixation des fragments d'ADN simple brin à un support grâce aux adaptateurs.
 - 3) Fixation de la seconde extrémité du fragment, ce qui entraîne la formation d'un pont (*bridge*), puis amplification des fragments par PCR.
 - 4) Des *clusters* regroupant le même fragment simple brin sont obtenus.
 - 5) Séquençage par synthèse avec l'ajout d'une polymérase et des 4 NTPs marqués par 4 fluorochromes différents.
 - 6) A chaque cycle, les NTPs incorporés sont détectés grâce à leur fluorochrome. Ceci permet la lecture de la séquence complète du fragment.
- (adaptée de Mardis 2008).

2.2. Analyse des données de séquençage

Avant l'analyse des données, le contrôle de la qualité des *reads* est la première étape nécessaire. Pour l'analyse en elle-même, deux approches différentes peuvent être employées : l'alignement ou l'assemblage. La première se base sur une séquence de référence et la deuxième est une approche sans à priori ou *de novo*, ces deux approches peuvent parfois être combinées.

Un grand nombre d'outils de bio-informatique sont disponibles pour réaliser ces différentes analyses à partir de données de NGS, ceux décrits par la suite sont présentés dans le Tableau 2. Ces outils s'utilisent principalement en ligne de commande sous GNU/Linux. Toutefois, certains ont été regroupés au sein de plateformes web pour un usage et une accessibilité simplifiés aux utilisateurs moins initiés aux langages informatiques. Ces plateformes, comme la plateforme de bio-informatique de Toulouse Midi-Pyrénées (Galaxy, GenoToul), offrent de hautes performances par la connexion à des serveurs de calculs et l'automatisation ou l'enchaînement de programmes par la création de flux de travail ou *workflow*.

Tableau 2 : Logiciels utilisés pour les analyses de séquençages NGS.

Analyse	Logiciel	Référence
Contrôle qualité	FastQC	-
	Qualimap 2.2.1	(Okonechnikov <i>et al.</i> 2016)
Assemblage	Pipeline AL5	(Tritt <i>et al.</i> 2012)
Alignement	BWA	(Li et Durbin 2009)
	Bowtie 2	(Langmead <i>et al.</i> 2009)
Analyse de variant	GATK	(McKenna <i>et al.</i> 2010)
	Breseq	(Deatherage et Barrick 2014)
	Nucmer	(Kurtz <i>et al.</i> 2004)
Visualisation	IGV_2.3	(Thorvaldsdóttir <i>et al.</i> 2013)
	Artemis Release 16.0.0	(Rutherford <i>et al.</i> 2000)
	ACT Release 13.0.0	(Carver <i>et al.</i> 2005)

Contrôle de la qualité des données

Le séquençage NGS génère un très grand nombre de données mais avec un taux d'erreur important (Tableau 1). Il est nécessaire de contrôler la qualité des données brutes avant de commencer les analyses. Un score de qualité associé à chaque nucléotide est présent dans les données au format fastq. Ce score, noté Q ou appelé score de Phred, est relié à la probabilité d'erreur P pour chaque nucléotide par l'équation :

$$Q = -10 \times \log(P)$$

Par exemple, pour un nucléotide donné si Q=10, la probabilité qu'il soit faux est de 1 sur 10 ; si Q=30, la probabilité est de 1 sur 1000.

Des logiciels comme FastQC génèrent des graphiques et des tables d'évaluation de la qualité à partir des données brutes au format fastq. FastQC rassemble, entre autres, les valeurs suivantes : la qualité de chaque nucléotide, la qualité de chaque *reads*, la proportion de chaque base d'ADN, le pourcentage en G+C, la distribution de la longueur des *reads*, etc.

Approche par assemblage *de novo*

A l'inverse de l'alignement sur une séquence de référence, l'assemblage *de novo* est une analyse sans à priori. Elle peut être utilisée pour analyser la séquence d'organismes pour lesquels il n'existe pas de référence ou pour lesquels il y a eu d'importants réarrangements.

Les différents *reads* qui contiennent des séquences chevauchantes sont assemblés entre eux en séquences plus longues appelées *contigs*. Ces *contigs* sont ensuite assemblés entre eux pour former des *scaffolds*. L'obtention d'un seul *contig* ou *scaffold* qui correspond à l'assemblage maximal d'une même molécule d'ADN (génom complet) n'est pas toujours possible. Pour minimiser le nombre de *scaffolds* obtenus, il est préférable d'utiliser des techniques qui génèrent des lectures longues ou *long reads*. Toutefois, ces techniques comme le PacBio donnent des profondeurs de couverture limitées et un taux d'erreur important (Eid *et al.* 2009). Pour une meilleure fidélité de séquence, il peut être utile de les combiner à des techniques générant une profondeur de couverture élevée avec des *reads* courts, comme l'Illumina. Au-delà des problèmes bio-informatiques de compatibilité des données générées et des outils utilisables, cela nécessite de réaliser tout en double (bibliothèques, réactions de séquençage et analyses) et présente donc un coût plus élevé. Une réorganisation de l'ordre et de l'orientation des *contigs* par comparaison à priori avec un génome proche peut constituer une alternative.

Le processus d'assemblage nécessite plusieurs étapes avec un nettoyage des séquences, un assemblage en *contigs* et en *scaffolds*, un contrôle de la qualité. Le pipeline A15 (*Andrew And Aaron's Awesome Assembly*) utilisé lors de nos assemblages de séquences Illumina a l'avantage de regrouper plusieurs logiciels et d'enchaîner automatiquement les étapes les unes après les autres (Tritt *et al.* 2012). Il a été mis au point pour analyser des génomes d'archées à partir de *reads* Illumina *paired-end*, et se compose de 3 étapes majeures : nettoyage des séquences (par SGA), assemblage des *contigs* (par IDBA), assemblage en *scaffold* (par SSPACE).

Comme nous disposons de références assemblées proches des génomes étudiés, nos travaux ont essentiellement fait appel à de l'Illumina. Le génome de la souche de référence PG2 a également été séquencé par PacBio afin de bénéficier des avantages des *reads* longs obtenus avec cette technique.

Approche par alignement sur une référence

Alignement des *reads*

L'approche par alignement des *reads* nécessite de choisir une séquence de référence, idéalement la plus proche possible du génome séquencé. Les logiciels recherchent pour chaque *read* le meilleur alignement sur la séquence de référence (position et identité) et ils leur attribuent des coordonnées de position et un score d'alignement.

Deux logiciels d'alignement de *reads* courts sur un génome de référence sont présentés ci-dessous : BWA et Bowtie 2. Ils utilisent les données brutes au format fastq et produisent des fichiers au format SAM (*Sequence Alignment Map*), composés de 2 parties : un en-tête et un tableau. L'en-tête contient des informations sur le génome de référence et le programme utilisé. Dans le tableau, chaque ligne correspond à un *read*, avec 11 à 12 colonnes qui indiquent le nom du *read*, sa position sur la référence, la qualité de l'alignement, la séquence nucléotidique, le nombre de SNP (*Single Nucleotide Polymorphism*) et d'indel (*insertion-deletion*) vis à vis de la référence, etc. Les fichiers SAM sont de taille très importante et peuvent être compressés au format BAM (*Binary Alignment Map*). Les BAM regroupent les mêmes informations mais sous forme de fichiers texte et non de tableaux, pour être déchiffrés par la plupart des logiciels ils nécessitent d'être associés à un fichier index (.bai).

BWA (*Burrows-Wheeler Alignment*) est un logiciel d'alignement de *reads* courts sur une référence peu divergente (Li et Durbin 2009). Selon la taille et la nature des *reads*, 3 choix d'algorithmes sont possibles : BWA-backtrack (*reads* de plus de 100 pb), BWA-SW (*reads* de 70

pb à 1 Mb contenant de nombreux indel) et BWA-MEM (*reads* de 70 pb à 1 Mb). Pour nos travaux, nous nous sommes appuyés sur ce dernier qui est le plus performant et le plus fréquemment utilisé.

Nous avons également utilisé Bowtie 2 car il est intégré avec l'analyse de variants effectuée par Breseq. Bowtie 2 s'utilise lors d'un alignement sur un génome de référence avec des *reads* de 50 pb à 1 Kb (Langmead *et al.* 2009). Deux types d'alignements sont possibles : "*end-to-end alignment*" où la totalité du *read* est alignée ; "*local alignment*" où un alignement seulement partiel du *read* est autorisé.

La qualité des alignements a ensuite été vérifiée à l'aide du logiciel Qualimap (Okonechnikov *et al.* 2016). A partir des fichiers SAM ou BAM, il génère des graphiques et des tables qui indiquent le nombre total de *reads*, le nombre de *reads* alignés, leur taille moyenne, la couverture moyenne et la qualité moyenne et le nombre de SNP et d'indel.

Visualisation de l'alignement

Les alignements de *reads* peuvent être visualisés à l'aide de logiciels qui donnent un accès graphique à l'étendue de la couverture d'un génome, à la profondeur de couverture de chaque position et aux différentes variations par rapport au génome de référence.

- Le logiciel IGV (*Integrative Genome Viewer*) peut lire de nombreux formats comme les BAM (Thorvaldsdóttir *et al.* 2013). Sur la visualisation générée, chaque SNP est indiqué en couleur par rapport au génome de référence (Figure 3A).

- Artemis permet de visualiser les fichiers BAM mais également d'annoter un génome (Figure 3B) (Rutherford *et al.* 2000).

- ACT (*Artemis Comparison Tool*) est utilisé pour comparer plusieurs génomes 2 à 2 à partir des séquences génomiques et des résultats de blast réalisés entre ces génomes (Carver *et al.* 2005).

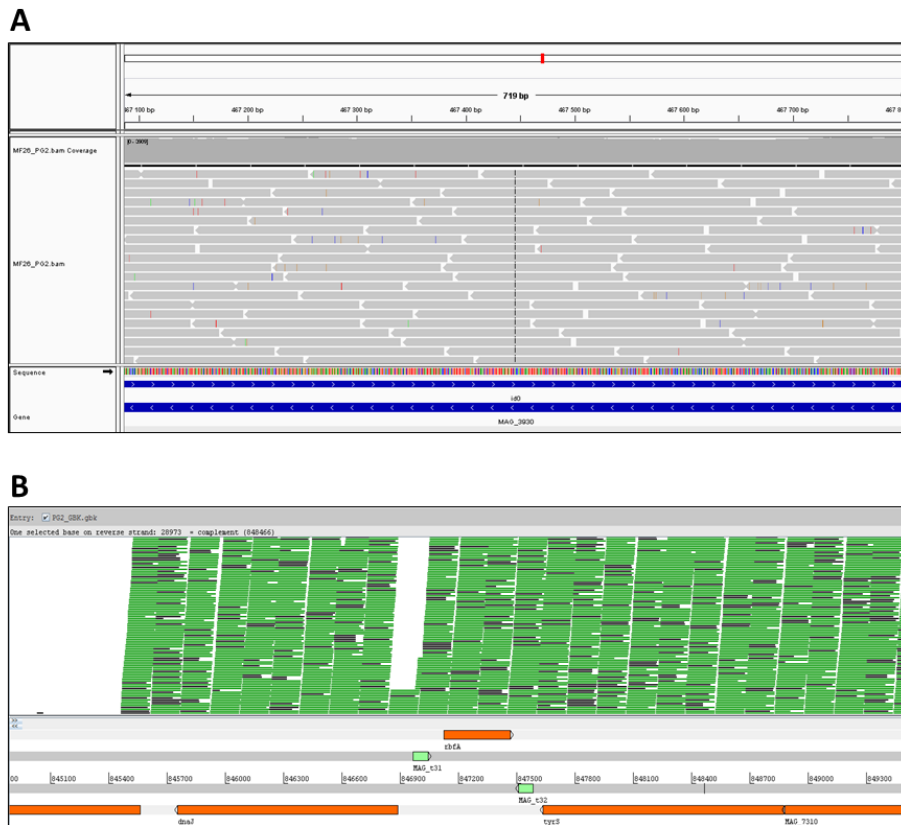


Figure 3 : Visualisation d'un alignement de reads sur une référence avec IGV ou Artemis.

(A) Alignement sur un génome de référence avec IGV, chaque *read* est représenté par une flèche grise, avec les variations indiquées en couleurs. Le trait bleu horizontal représente le génome de référence. (B) Alignement sur un génome de référence avec Artemis, chaque *read* est représenté par un trait vert. Les flèches orange et vert clair indiquent les CDS du génome de référence.

Recherche de variants vis-à-vis d'une référence

A partir d'un alignement (fichier au format SAM ou BAM), des logiciels détectent les SNP et les indel par rapport à la référence. Ils trient les variations présentes dans la population pour écarter les mésappariements dus à des erreurs de séquençage. Nous présenterons ici trois logiciels : GATK-3, Breseq et Nucmer. Pour choisir le mieux adapté à chaque utilisation, une comparaison de ces 3 logiciels sur les mêmes jeux de données a été réalisée (Paragraphe 3.1).

GATK-3 (*Genome Analysis Toolkit*) identifie chaque variation (SNP et indel) par rapport au génome de référence (McKenna *et al.* 2010). Pour cela, il procède à un réalignement local des *reads*, à une recalibration de la qualité des nucléotides puis à une analyse de variants. Il génère un fichier au format VCF (*Variant Calling Format*) lisible avec un tableur type Excel ou visualisable sur un génome de référence avec IGV ou Artemis. Un VCF est composé d'un en-tête et d'un tableau listant chaque variation, sa position, la modification qu'elle entraîne, son

score de qualité, etc. GATK-3 calcule, entre autres, des paramètres tels que AD (*Allelic Depths*, nombre de *reads* porteurs d'un allèle donné) et DP (*Approximate read Depth*, nombre total de *reads*). Le ratio AD/DP permet d'évaluer la proportion de chaque variation au sein de la population. Par exemple, $AD_1=2500$, $DP=5000$, indique que l'allèle 1 est retrouvé dans 50 % de la population séquencée ($2500/5000 \times 100 = 50$). Plus la profondeur de couverture est importante, plus la présence et la fréquence de sous-populations minoritaires peuvent être déterminées.

Breseq est un autre programme pour l'analyse de variants à partir de *reads* Illumina (Deatherage et Barrick 2014). Ce programme a été développé pour analyser les séquences des souches d'*Escherichia coli* issues du projet d'évolution expérimentale LTEE (*Long Term Evolution Experiment*) débuté il y a 30 ans (Pennisi 2013). Il peut être utilisé pour des génomes microbiens dont la séquence est similaire à la référence. L'analyse débute par un alignement des *reads* sur le génome de référence par Bowtie 2, suivie de l'analyse et de l'annotation de chaque variation. Breseq détecte également les séquences de jonctions absentes de la référence (*new sequence junctions*) souvent associées à l'insertion d'éléments génétiques mobiles. Utilisé en mode "polymorphisme", les variations des sous-populations sont analysées. En mode "consensus", l'analyse se limite à la population majoritaire. Il est recommandé de limiter le nombre de *reads* utilisés à 80-120X, compromis entre rapidité d'analyse et discrimination des variants. Dans notre cas avec une couverture supérieure à 2700X, cela diminue fortement l'information disponible au niveau du polymorphisme des sous-populations par rapport à GATK.

Nucmer (*Nucleotide MUMmer*) est une extension du package MUMmer (Delcher *et al.* 2002; Kurtz *et al.* 2004). MUMmer permet un alignement de deux séquences au format fasta par la recherche des motifs MUMs (*Maximal Unique Matches*), motifs parfaitement identiques entre les deux séquences. Le programme comble ensuite les vides entre les MUMs par des alignements comprenant le moins de variations possibles. Pour finir, Nucmer liste toutes les variations entre les deux génomes et génère un tableau. Son utilisation nécessite des génomes au format fasta qui peuvent être obtenus à la suite d'un assemblage *de novo* ou d'une séquence consensus générée à partir d'un alignement sur une séquence de référence. Par rapport aux deux logiciels précédents, le recours à ce format (fasta) induit la perte de l'information sur les sous-populations de variants non majoritaires.

3. Résultats et discussion

3.1. Détection de variants

Deux génomes circularisés de *M. agalactiae* sont disponibles dans les banques de données : PG2 clone 55-5 (Sirand-Pugnet *et al.* 2007) (NC_009497.1) et 5632 clone C1 (Nouvel *et al.* 2010) (NC_013948.1). Ces clones ou leurs dérivés, à la base de toutes nos expériences, ont été (re)séquencés par Illumina. Des recherches de variants entre PG2 et 5632 ont été réalisées à l'aide de 3 logiciels (GATK3, Breseq, Nucmer). Le nombre de variations entre les deux souches diffère suivant le logiciel et le génome de référence utilisé (Tableau 3). Une comparaison fine des variations identifiées par ces 3 logiciels montre que les divergences résultent essentiellement d'une difficulté à identifier correctement les indel. Par comparaison aux deux autres logiciels, Nucmer est celui qui détecte le plus de variations, en moyenne une variation tous les 26 nt. Il détecte des variations par excès (faux positifs) mais aucun faux négatif n'a pu être identifié. Bien que moins spécifique, il est donc le plus sensible et a été retenu pour le profil type des variations entre PG2 et 5632 utilisé comme référence dans l'analyse des génomes des transconjugants.

Tableau 3 : Variations entre PG2 et 5632 détectées à partir de 3 logiciels

Logiciel		Référence PG2		Référence 5632	
Alignement	Recherche de variants	Nombre de variations (% SNP)	1 variation tous les ... nt	Nombre de variations (% SNP)	1 variation tous les ... nt
BWA-MEM	GATK3	28856 (95,6 %)	30	29982 (95,1 %)	34
Bowtie2	Breseq	27465 (92,2 %)	32	27215 (92,2 %)	37
MUMmer	Nucmer	37052 (94,1 %)	24	36434 (94,3 %)	28

A l'inverse de Nucmer, GATK3 se base sur les données brutes Illumina (*reads*) et peut donner accès aux sous-populations de variants qui coexistent. Il a ainsi été utilisé pour détecter les variants (SNP et indel) sur les génomes des expériences d'évolution expérimentale avec ou sans pression de sélection. La première étape consiste en un l'alignement des données brutes (fastq) sur un génome de référence avec BWA (algorithme MEM). La qualité des données brutes et des alignements a été vérifiée avec FastQC et Qualimap, respectivement et une

recherche de variants a été effectuée avec GATK3, les variations de qualité inférieure à 10000 ont été écartées. Un schéma des étapes de l'analyse de variants est présenté dans la Figure 4.

La proportion de chaque variation au sein de la population totale est obtenue par le ratio AD/DP (nombre de *reads* porteurs d'un allèle donné / nombre total de *reads*). Des seuils sont classiquement fixés (Sun *et al.* 2012), les variations détectées dans 95 % des *reads* ou plus ont été considérées comme fixées dans l'ensemble de la population, celles détectées dans moins de 95 % des *reads* comme non fixées.

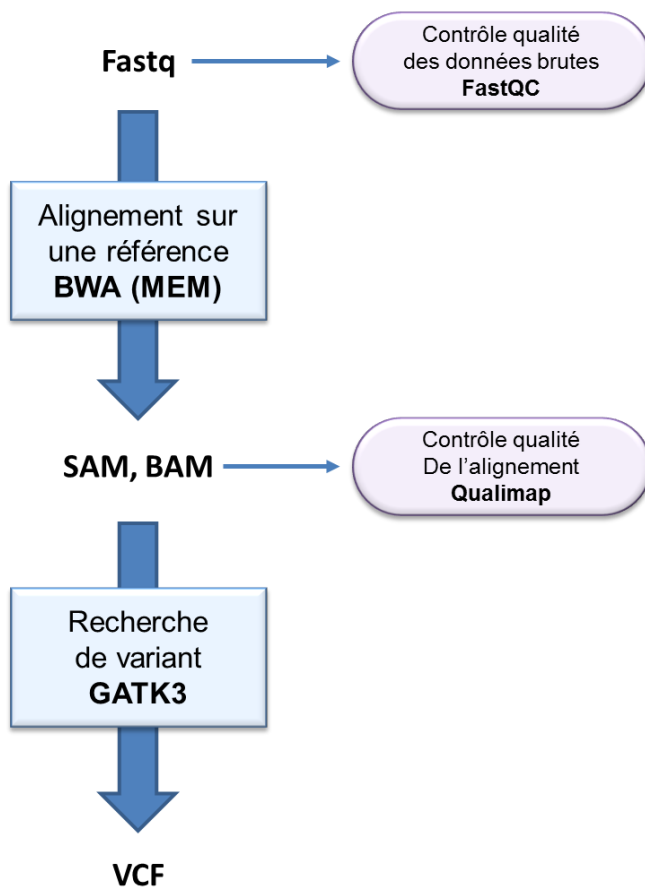


Figure 4 : Etapes réalisées lors d'une analyse de variants vis-à-vis d'une référence.

A partir des données brutes (fastq) un alignement sur le génome de référence est réalisé par BWA, avec l'algorithme MEM. L'alignement génère un fichier au format SAM qui est ensuite transformé en BAM moins volumineux. A partir de l'alignement (BAM) une recherche de variants est effectuée avec GATK3 et aboutit à un fichier VCF. La qualité des données brutes est vérifiée avec FastQC et la qualité de l'alignement avec Qualimap.

Le clone 5632^{G3} dérive de la souche 5632 par l'insertion d'une cassette de résistance à la gentamicine (*aacA-aphD*) par transposition (Dordet-Frisoni *et al.* 2014). Ce clone et une culture de celui-ci après 10 passages en culture liquide (P10) ont été séquencés par Illumina. Une analyse de variants par rapport à 5632 révèle 14 mutations chez 5632^{G3}, dont 6 fixées et 8 non fixées (Tableau 4). Après 10 passages en culture liquide sans pression de sélection, les 14 mutations initiales sont toujours présentes et 10 mutations non fixées apparaissent. La majorité des mutations non fixées se trouvent dans les loci *vpma*, régions connues comme variables car soumises à de nombreux réarrangements génétiques qui contribuent à l'hypervariabilité des protéines membranaires observée chez *M. agalactiae* (Citti *et al.* 2010).

Une sous culture du clone PG2 55-5 séquencée par Illumina a été comparée au génome de référence PG2 55-5 obtenu en 2007 par Sanger (Sirand-Pugnet *et al.* 2007). Cette sous culture comporte 5 mutations au total, 2 fixées et 3 non fixées. Les 2 mutations fixées sont des délétions au niveau d'homopolymères intergéniques. Les 3 mutations non fixées sont situées dans un locus appelé ps3. A l'issue de 40 passages (P40) sans pression de sélection, 3 nouvelles mutations fixées sont détectées ailleurs dans le génome (Tableau 4).

A partir de la sous culture du clone PG2 55-5 reséquencé, des mutants spontanément résistants à l'enrofloxacin ont été sélectionnés. Leurs séquences ont été alignées sur PG2 55-5 pour déterminer les mutations potentiellement responsables de la résistance. Ces clones résistants présentent en moyenne 10 mutations par rapport à PG2 55-5. Les résultats de l'ensemble des mutants séquencés sont présentés dans le chapitre 1, à titre d'illustration, ceux du mutant MF26-1 sont présentés ici (Tableau 4). Il comporte 9 mutations dont les 5 précédemment identifiées au niveau de PG2 55-5. Parmi les 4 mutations apparues lors du processus de sélection, 2 mutations décrites comme responsables du phénotype de résistance (Redgrave *et al.* 2014) sont situées dans les gènes codant les cibles de l'enrofloxacin (*parC* et *gyrA*).

Tableau 4 : Analyse de variant chez différents clones.

	Position nucléotide	Mutation	Genbank locus tag	Nom du gène	Produit	Homo polymère Impact du SNP	Pourcentage de reads mutés
5632^{G3}							P0 P10
	286425	SNP A -> T	MAGa2550/2560 ./.		Protéine de liaison à l'hémolysine/ARNt Gly		- 31
	370727	Del A	MAGa3220 .		CDS22	6 A	100 100
	387221	Inser C	MAGa3340 .		HP	8 C	98 96
	525384	Del AG	MAGa4540/4550 ./.		HP/HP	15 AG	- 35
	526453	Inser AG	MAGa4550/4560 ./adhT		HP/Alcool déshydrogénase	12 AG	83 73
	588703	SNP TTG (L) -> ATG (M)	MAGa5070 .		Transporteur de type ABC	NS	100 100
	588815	SNP CGA (R) -> CAA (Q)	MAGa5070 .		Transporteur de type ABC	NS	100 100
	591019	SNP AGT (S) -> AAT (N)	MAGa5090 .		HP	NS	100 100
vpma	677740	Del TAT	MAGa5820/5830 <i>vpmaB/vpmaC</i>		LSV B/LSV C		- 88
	678694	SNP TTG (L) -> TTC (F)	MAGa5830 <i>vpmaC</i>		LSV C	NS	- 44
	720155	SNP AAA (K) -> AAG (K)	MAGa6310 <i>hsdS</i>		Sous unité du système de RM de type I	S	50 50
	774998	Del C	MAGa6730/6740 ./.		HP/HP	13 C	- 49
	788608	Del C	MAGa6790/6800 ./.		HP/HP	11 C	- 33
	794159	Del C	MAGa6820/6830 ./.		HP/HP	10 C	93 86
	939298	Del A	MAGa8050/8060 <i>vpmaF2/vpmaX</i>		LSV F2/LSV X		42 36
	940889	SNP G -> A	MAGa8070/8080 <i>vpmaW/vpmaC</i>		LSV W/LSV C		22 32
	941932	Inser T	MAGa8080/8090 <i>vpmaC/vpmaE</i>		LSV C/LSV E		62 57
	941938	SNP G -> A	MAGa8080/8090 <i>vpmaC/vpmaE</i>		LSV C/LSV E		68 67
vpma	943721	SNP T -> C	MAGa8090/8100 <i>vpmaE/vpmaB</i>		LSV E/LSV B		79 75
	943774	SNP TCG (S) -> TCA (S)	MAGa8100 <i>vpmaB</i>		LSV B	S	- 60
	943775	SNP CTG (L) -> TTG (L)	MAGa8100 <i>vpmaB</i>		LSV B	S	- 58
	943777	SNP CTG (L) -> CTA (L)	MAGa8100 <i>vpmaB</i>		LSV B	S	- 58
	943783	SNP TCA (S) -> TCT (S)	MAGa8100 <i>vpmaB</i>		LSV B	S	- 54
	989455	SNP GCA (A) -> ACA (T)	MAGa8600 <i>mgta</i>		ATPase	NS	100 100
PG2 55-5							P0 P40
	151276	Del C	MAG_1330/1340 ./.		HP/HP	8 C	98 99
	183721	Del G	MAG_1570 .		HP	9 G	- 96
	331447	Del AT	MAG_2820/2830 <i>ligA</i> ./.		DNA ligase/Protéine transmembranaire	9 AT	98 98
	443390	SNP GCA (A) -> GAA (V)	MAG_3750 <i>ktrA</i>		Transporteur potassium	NS	- 99
	634214	SNP G -> A	MAG_5470/7520 <i>rps</i> ./.		Protéine ribosomique S10/Pseudogène		- 100
	761525	SNP TTG (L) -> TTA (L)	MAG_6430 .		HP	S	20 -
ps3	763036	SNP A -> G	MAG_6450/6460 ./.		HP/Lipoprotéine		50 49
	763685	Del 11nt	MAG_6460 .		Lipoprotéine		48 48
Mutant MF26-1							P0
PG2 55-5	151276	Del C	MAG_1330/1340 ./.		HP/HP	8 C	90
	183721	Del G	MAG_1570 .		HP	9 G	98
	214190	SNP ACT (T) -> ATT (I)	MAG_1830 <i>parC</i>		Topoisomerase IV sous unité A	NS	100
PG2 55-5	331447	Del AT	MAG_2820/2830 <i>ligA</i> ./.		ADN ligase	9 AT	98
	647997	SNP TCA (S) -> TTA (L)	MAG_5630 <i>gyrA</i>		ADN gyrase sous unité A	NS	100
	707306	Del C	MAG_6060/6070 ./.		HP/HP	9 C	99
PG2 55-5 (ps3)	761525	SNP TTG (L) -> TTA (L)	MAG_6430 .		HP	S	23
	763036	SNP A -> G	MAG_6450/6460 ./.		HP/Lipoprotéine		49
	763685	Del 11nt	MAG_6460 .		Lipoprotéine		47

Variations détectées lors d'une analyse de variants de 5632^{G3} (référence 5632) et de PG2 55-5 et un mutants MF26-1 (référence PG2 55-5). Pour chaque variation, sa position, la modification qu'elle entraîne, le gène qu'elle affecte et le pourcentage de reads muté (ratio AD/DP) sont indiqués. Variations fixées en bleu foncé, non fixées en gris. Les locis remarquables sont indiqués en bleu clair à gauche. Les alignements sont réalisés avec BWA-MEM et l'analyse de variants avec GATK-3. (Del= Délétion; Inser= Insertion; HP= Protéine hypothétiques; LSV= Lipoprotéine de surface variable; S= Synonyme; NS= Non-synonyme).

3.2. Reconstruction des génomes mosaïques des transconjugants

Les transconjugants possèdent des génomes mosaïques PG2/5632. Les génomes de PG2 et 5632 présentent un pourcentage d'identité élevé (82,92 % selon *NCBI Genome Symmetrical identity*) et une synténie conservée à l'exception de quelques gènes présents uniquement chez l'une des 2 souches. En revanche, ces 2 génomes présentent des variations suffisantes pour identifier les fragments issus de PG2 (ou PG2^{E10}) de ceux issus de 5632 (ou 5632^{G3}) (Figures 5B, 5D). Un alignement des *reads* d'un transconjugant sur 5632 (Figure 5A) permet d'identifier les fragments correspondants au génome de 5632 (*reads* alignés sans mésappariement représentés en gris) et les fragments correspondants au génome de PG2 (*reads* alignés avec mésappariements représentés en couleurs). L'alignement des *reads* de ce même transconjugant sur la référence PG2 donne le résultat inverse (Figure 5C).

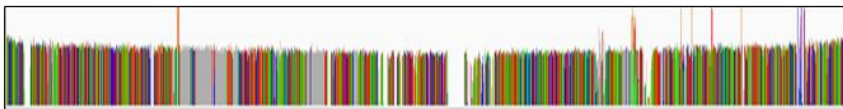
A Transconjugant aligné sur 5632



B PG2^{E10} aligné sur 5632



C Transconjugant aligné sur PG2



D 5632^{G3} aligné sur PG2



Figure 5 : Visualisation IGV d'alignements des *reads* sur le génome des 2 souches parentales.

Alignements (BWA-MEM) des *reads* d'un transconjugant sur le génome de PG2 (A) ou de 5632 (C), les *reads* de PG2^{E10} sont alignés sur 5632 (B) et ceux de 5632^{G3} sur PG2 (D). Les zones grises indiquent des *reads* identiques à la séquence de référence. Chaque trait vertical de couleur désigne une mutation par rapport au génome de référence. Les 2 génomes 5632 et PG2 présentent un grand nombre de variations entre eux (B et D). Les génomes des transconjugants sont hybrides (PG2/5632), en alignant les *reads* sur 5632 les régions colorées correspondent aux séquences issues de PG2 (A) et en les alignant sur PG2 les régions colorées correspondent aux séquences issues de 5632 (C).

Ces alignements ne permettent cependant qu'une définition approximative des régions de PG2 ayant transférées vers 5632. La définition précise des fragments demande de vérifier si entre deux variations observées, d'autres variations entre PG2 et 5632 existent ou s'il s'agit d'une région commune à PG2 et 5632 (Figure 6). Pour ce faire, le profil des variations entre PG2/5632 obtenu avec Nucmer (Paragraphe 3.1) peut être utilisé. Ce type d'analyse nécessite de "scanner" manuellement tous les génomes de transconjugants pour comparer précisément chaque variation.

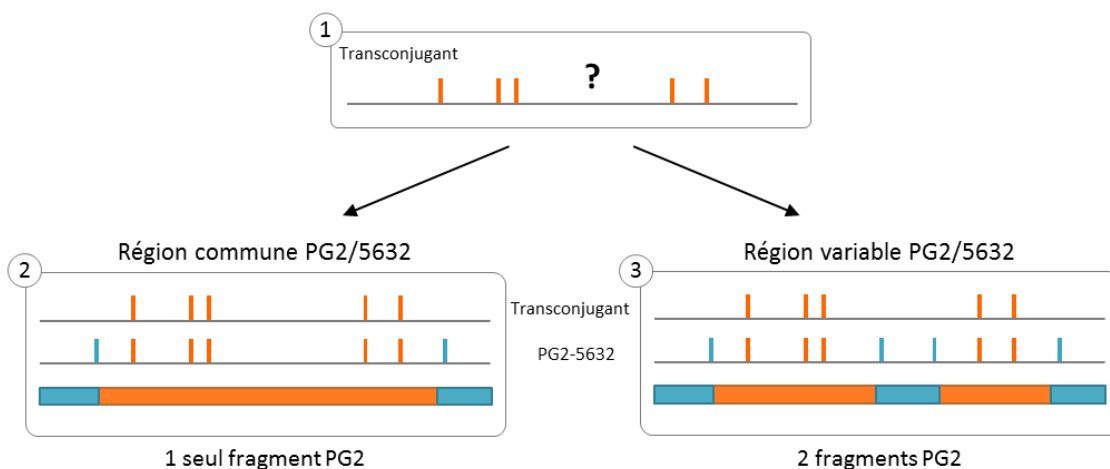


Figure 6 : Définition des fragments des transconjugants.

Les variations représentées par les traits verticaux orange correspondent à PG2 et celles en bleue à 5632. Lorsqu'on aligne les *reads* d'un transconjugant sur 5632, on obtient des variations correspondant à PG2 (1). Mais entre deux variations PG2 soit la région est commune (2), soit des variations existent entre PG2 et 5632 (3). Dans le premier cas un seul fragment PG2 est défini (2), alors que pour le second, deux fragments de PG2 sont définis (3).

Une méthode plus automatisée a été mise au point pour déterminer les fragments de PG2 transférés chez 5632 et les résultats détaillés sont présentés dans les chapitres 1 et 2, nous reprendrons ici plus en détail la démarche (Figure 7). Elle repose sur des étapes d'alignement et de tri des *reads* en deux catégories : ceux qui ont des *mismatches* (mésappariements) et ceux qui n'ont aucun *mismatch* (totalement appariés) par rapport au génome de référence PG2 ou 5632. La première étape consiste à aligner les *reads* du transconjugant sur 5632, puis à récupérer ceux qui ont des *mismatches*. Ceci permet d'éliminer les *reads* spécifiques de 5632 et les *reads* de séquences communes à PG2 et 5632 pour ne conserver que ceux spécifiques de PG2. Ces *reads* sont ensuite alignés sur PG2, un "bruit de fond" est observé avec des *reads* contenant des *mismatches* répartis tout le long du génome avec une faible couverture (environ

300X). Les *reads* qui n'ont aucun *mismatch* sont alors conservés, ils correspondent aux régions PG2 du génome du transconjugant.

Les alignements sont réalisés sur Galaxy avec BWA-MEM. Le tri des *reads* se fait à l'aide de l'outil "*Select lines*". Le pattern "NM:i:0" permet d'identifier tous les *reads* qui s'alignent parfaitement sur la référence (aucun *mismatch*). On peut ensuite sélectionner les *reads* qui possèdent ce pattern (No-mismatch) ou non (Mismatch). L'analyse des génomes des transconjugants montre qu'en moyenne 71 % des *reads* correspondent à des régions de 5632, 8 % à des régions de PG2 et 21 % à des régions communes. Pour finir, à l'aide du logiciel Artemis les fragments correspondants à PG2 sont définis lorsqu'au minimum 15 *reads* sont présents, les parties restantes du génome correspondent à 5632. Cela permet d'annoter le génome des transconjugants sur la référence PG2, les positions nucléotidiques des fragments sont ainsi définies par des coordonnées sur le génome de PG2. Pour chaque zone ambiguë et pour les fragments très courts une vérification manuelle est effectuée au niveau du Bam. Le même procédé appliqué à partir de *reads* issus de 5632^{G3} identifie quelques régions spécifiques de PG2 (5 kb au total), ces régions correspondent donc à des "faux positifs" générés par cette méthode de tri des *reads*. Ces faux positifs, détectés aussi sur les génomes de transconjugants sont manuellement supprimés.

Entre deux variations qui définissent un fragment mosaïque chez un transconjugant (une spécifique à PG2 et une à 5632), la séquence entre PG2 et 5632 est identique. Les zones de recombinaison des fragments PG2 ne peuvent donc pas être exactement définies au nucléotide près. Ainsi, les bords du fragment ne peuvent être déterminés précisément, seuls sont définissables son bord minimal (variation de type PG2) et son bord maximal (variation de type 5632). Lors de nos analyses, nous avons choisi de définir nos fragments PG2 par leur taille maximale, c'est à dire la région comprise entre deux variations correspondant à 5632 (Figure 6).

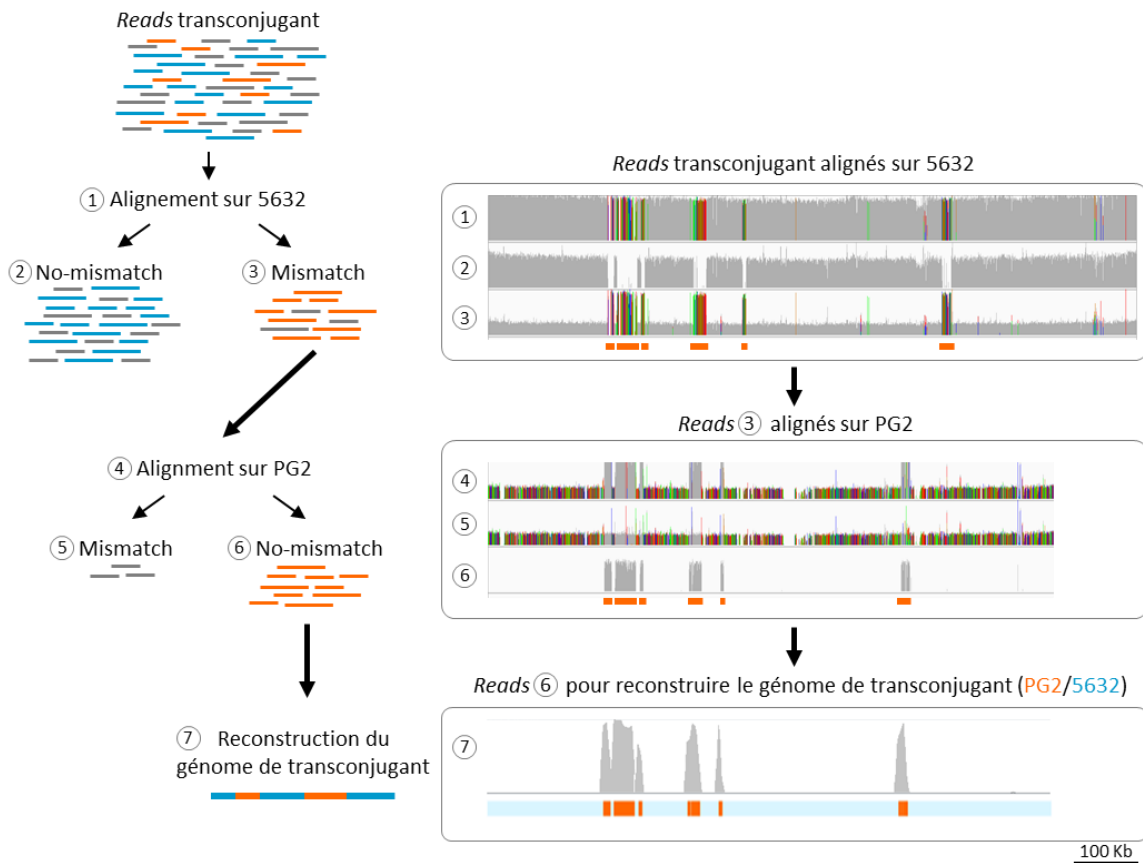


Figure 7 : Méthode de tri des *reads* spécifiques de PG2 pour reconstruire les génomes des transconjugants.

A gauche est présenté le schéma de la méthode et à droite les résultats des alignements correspondants (réalisé avec BWA-MEM et visualisé avec IGV). A gauche : les petits traits horizontaux représentent des *reads*, en gris ceux communs PG2/5632, en bleu ceux spécifique de 5632 et en orange ceux spécifique de PG2. A droite : les zones grises indiquent des *reads* identiques à la référence (no-mismatch), les traits verticaux en couleur indiquent les variations vis-à-vis de la référence (mismatch). Les traits orange horizontaux identifient les fragments de type PG2. La dernière étape (7) se fait avec Artemis qui permet de créer les fragments génomiques de type PG2 à partir des *reads* triés.

3.3. Analyse des variations de couverture

Des variations de couverture sont observées pour certaines régions des génomes séquencés. Artéfacts de séquençage, d'assemblage ou différences entre la référence et le clone séquencé, elles peuvent révéler des éléments intéressants sur la structure de nos génomes (Figure 8).

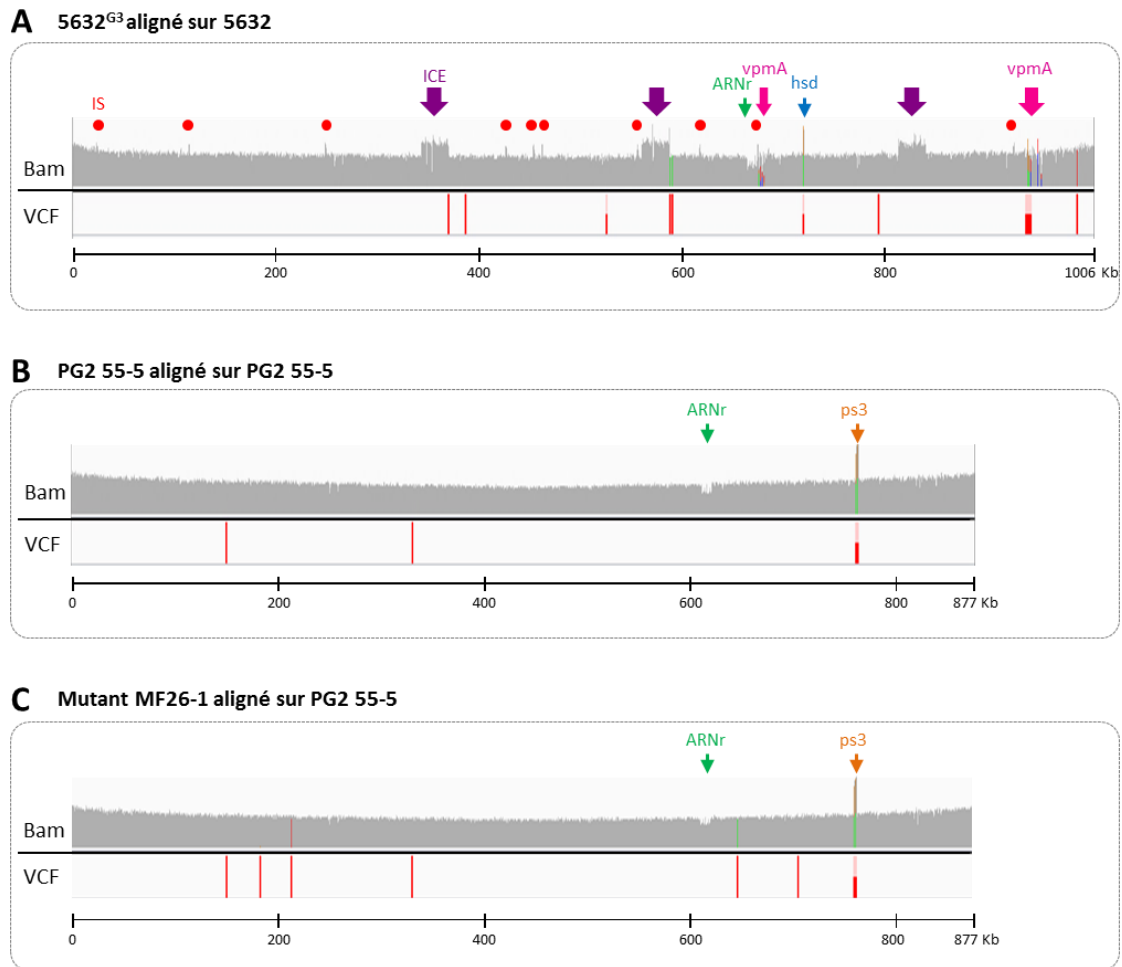


Figure 8 : Alignement et analyse de variant chez 5632^{G3}, PG2 55-5 et un mutant résistant.

Visualisation des alignements des *reads* issus de 5632^{G3} sur 5632 (A) des *reads* de PG2 55-5 en Illumina sur PG2 55-5 (B) et du mutant MF26-1 sur PG2 55-5 (C). Pour les alignements (Bam), les zones grises indiquent les *reads* identiques à la référence et les traits verticaux colorés désignent les variations vis-à-vis de la référence. Les flèches et les points montrent les régions pour lesquelles la couverture est différente de la moyenne du génome (IS: Séquence d'insertion; ICE: Élément intégratif et conjugatif; *vpmA*: loci des gènes de protéines de surface hypervariables de *M. agalactiae*). Pour l'analyse de variant (VCF) les barres verticales rouges indiquent les positions des mutations (voir la liste Tableau 4). Les alignements sont réalisés avec BWA-MEM (Bam), l'analyse de variants avec GATK3 (VCF) et la visualisation avec IGV.

Éléments génétiques mobiles chez 5632

Le génome de 5632 contient 3 copies d'ICE (*Integrative and Conjugative Element*) quasi-identiques et des IS (*Insertion Sequence*) (Marenda *et al.* 2006; Nouvel *et al.* 2010). L'alignement des *reads* de 5632^{G3} sur le génome de 5632 montre des augmentations de la profondeur de couverture, localement, au niveau de ces éléments génétiques mobiles (Figure 8A). La couverture est, par exemple, d'environ 3700X au niveau des ICE pour une profondeur moyenne de 2700X sur l'ensemble du génome. Ces augmentations de couverture

peuvent résulter de l'insertion de nouvelles copies d'éléments génétiques mobiles ailleurs dans le génome ou d'une surexpression de leurs formes extra-chromosomiques circulaires. Une détection des copies d'ICE et de leur forme circulaire a été réalisée par *Southern blot*, voir ci-dessous (Figure 9B).

Variation du nombre de copie d'ICE chez les transconjugants

Dix transconjugants constitués d'un génome hybride PG2^{E10}/5632^{G3} ont été séquencés. Le clone 5632^{G3} contient 3 copies d'ICE identiques de 27 Kb (ICE-I ; ICE-II et ICE-III) absentes chez PG2^{E10}. L'alignement des *reads* des deux parents et des transconjugants sur 5632 montre une profondeur de couverture variable au niveau des ICE par rapport au reste du génome (Figure 9A). PG2^{E10} ne contient pas d'ICE, il n'y a donc aucun *read* correspondant. Chez 5632^{G3} la couverture des ICE est supérieure au reste du génome. Pour les transconjugants la couverture des ICE est parfois plus faible que la moyenne (T5-2 ; T5-5 ; T6-1), parfois plus élevée (T6-4 ; T6-8 ; T6-9), ou bien équivalente au reste du génome (T5-1 ; T5-3 ; T5-4 ; T6-5). Les diminutions de la couverture au niveau des ICE suggèrent une délétion d'une copie d'ICE alors que les cas d'augmentation suggèrent la présence d'une copie supplémentaire ou une abondance de la forme circulaire extra-chromosomique.

Un *Southern blot* a été réalisé afin de mieux comprendre ces différences de profondeur, en utilisant une sonde spécifique des ICE (Figure 9B). Après digestion enzymatique cette sonde permet de distinguer les 3 copies d'ICE au niveau du chromosome et également de détecter la forme circulaire (Marenda *et al.* 2006). Les 3 transconjugants qui montrent une diminution de la couverture des ICE ont chacun une copie manquante (ICE-I pour T5-2 et T6-1 ; ICE-III pour T5-5). Après la reconstruction de leurs génomes mosaïques, il apparaît que les régions contenant ces ICE ont été remplacées par recombinaison homologue par un fragment de PG2^{E10} (Figure 10). La forme circulaire est détectée par *Southern blot* pour les transconjugants dont le séquençage montre une couverture élevée au niveau des ICE (Figure 9B). C'est donc l'abondance de forme circulaire qui induit une surreprésentation des *reads* correspondants aux ICE pour ces transconjugants (T6-4 ; T6-8 ; T6-9). Dans une moindre mesure, la même chose est observée pour 5632^{G3}.

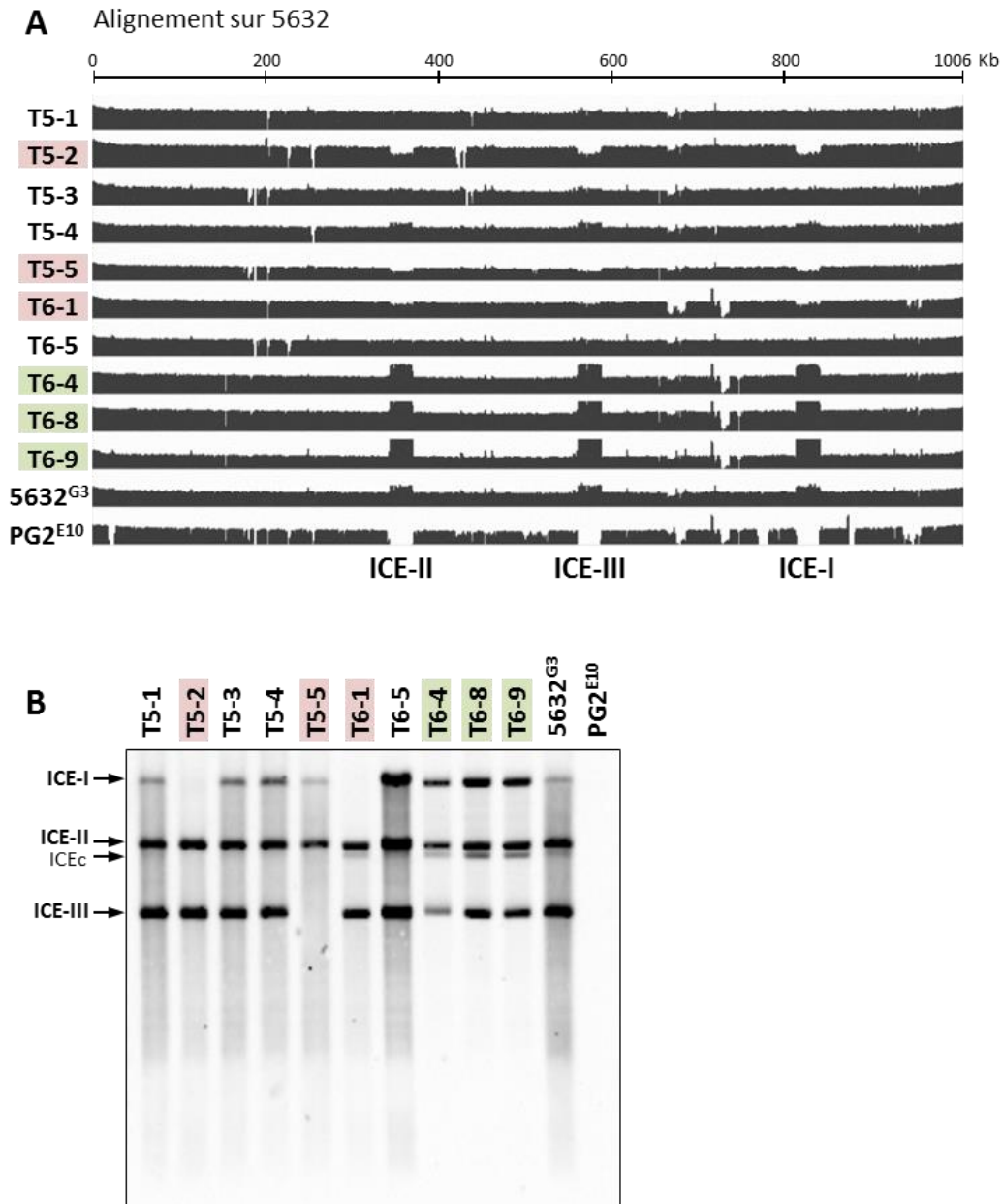


Figure 9 : Variation du niveau d'expression des ICE des transconjugants.

(A) Visualisation avec IGV des *reads* issus de 10 transconjugants et des 2 parents (5632^{G3} et PG2^{E10}) alignés sur le génome de 5632 (BWA-MEM), les positions des 3 copies d'ICE sont indiquées en bas. (B) *Southern blot* permettant de visualiser les 3 copies d'ICE et la forme circulaire chez les transconjugants et 5632^{G3}. ADN génomiques digérés par EcoRV qui contient un site dans la CDS22. Séparation des fragments de restrictions par migration en gel d'agarose à 1 %. Révélation avec une sonde ADN s'hybridant avec la séquence de la CDS22, marquée à la digoxigénine (Marenda *et al.* 2006). Les tailles des fragments sont de 13,2 Kb pour l'ICE-I; 8,9 Kb pour l'ICE-II; 2,1 Kb pour l'ICE-III et 4,5 Kb pour l'ICEc (circulaire). Les transconjugants qui ont perdu une copie d'ICE sont surlignés en rouge, ceux pour qui la forme circulaire est surexprimée en vert.

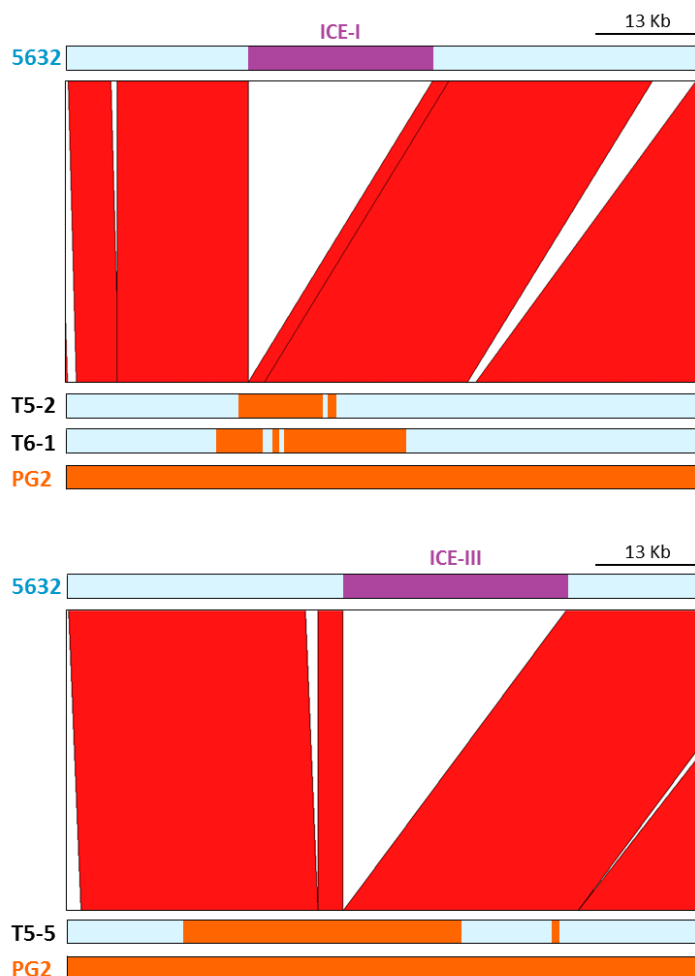


Figure 10 : Délétion des ICE due à la recombinaison d'un fragment PG2.

Comparaison du génome de 5632 (en bleu) et de PG2 (en orange) avec ACT. Les zones rouges indiquent les régions communes aux 2 génomes. Les ICE (ICE-I et ICE-III) (en violet) ne sont présentes que chez 5632. Chez les 3 transconjugants (T5-2; T6-1; T5-5) la recombinaison avec un fragment PG2 a eu lieu en amont et en aval de l'ICE, ce qui aboutit à sa délétion dans le génome du transconjugant.

Opérons ribosomiques et loci *vpma*

Chez *M. agalactiae*, les opérons ribosomiques sont en double copie. Pour tous les génomes séquencés par Illumina, la profondeur de couverture est diminuée au niveau de ces gènes codant les ARNr 23S et 16S (Figure 8). Plus qu'une diminution globale de couverture à ces loci, un niveau inconstant de couverture est observé, donnant un profil avec plusieurs "petites vagues" comme le montre un zoom sur ces opérons de 5632^{G3} (Figure 11). La technique Illumina est sensible aux variations du taux en G+C des séquences d'ADN (Oyola *et al.* 2012) en lien avec des biais de PCR lors de la préparation des bibliothèques (Benita *et al.* 2003). Les mycoplasmes ont des génomes globalement riches en A+T, mais l'opérons ribosomique 23S-16S contient 45 % de G+C, contre 29 % en moyenne sur tout le génome pour *M. agalactiae*.

L'inconstance de couverture observée ici est donc à mettre en relation avec cette forte variation de contenu en G+C.

Chez 5632, des variations de couverture sont également présentes au niveau des deux loci *vpma* (Figures 8A; 11). Ces loci sont constitués de gènes codant pour des protéines membranaires variables avec de nombreuses séquences communes. Des réarrangements génomiques ont lieu entre les différents gènes *vpma* de ces loci et entraînent de nouvelles configurations des protéines membranaires (Nouvel *et al.* 2009). Ces réarrangements permettent d'expliquer les variations de couverture (1000X-3200X) et également les mésappariements détectés (Tableau 4).

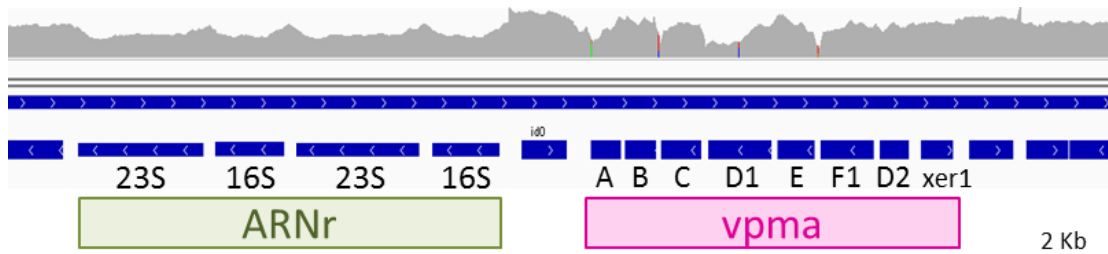


Figure 11 : Variation de couverture de l'opéron ribosomique et d'un locus *vpma*.

Visualisation d'un alignement des *reads* de 5632^{G3} sur 5632 (nt 663200-687900) avec un zoom au niveau de l'opéron ribosomique 23S-16S (en vert) et d'un des 2 locus *vpma* (en rose). Les zones grises indiquent les *reads* identiques à la séquence de référence et les traits verticaux colorés désignent les variations vis-à-vis de la référence. Les alignements sont réalisés avec BWA-MEM (Bam), l'analyse de variants avec GATK3 (VCF) et visualisés avec IGV.

Positionnement de l'insertion d'un marqueur de sélection

Le clone 5632^{G3} a été généré par l'insertion d'un gène de résistance à la gentamicine (*aacA-aphD*) chez 5632 par transposition (Dordet-Frisoni *et al.* 2014). Le transposon utilisé est modifié et ne contient pas de transposase afin de rendre son insertion stable dans le génome. Le clone 5632^{G3} contient donc une large insertion de 3,6 Kb absente chez 5632 au niveau de la CDS MAGa7850. Au point correspondant à cette insertion, l'alignement des *reads* de 5632^{G3} sur le génome de 5632 met en évidence une diminution progressive de la profondeur de la couverture puis un pic de couverture donnant un profil à l'allure d'une "cheminée" (Figure 12A). Ce pic constitue le *direct repeat* (DR) de 8 nt (nt 919899-919906) dupliqué à l'insertion du transposon chez 5632 au niveau de la MAGa7850. Une amplification par PCR à l'aide d'une amorce spécifique de la région insérée (transposon) et d'une amorce située proche du site d'insertion sur le chromosome confirme la position d'insertion du marqueur

(Figure 12B). L'assemblage *de novo* des *reads* de 5632^{G3} par A15 permet d'obtenir 431 *scaffolds*, dont un *scaffold* de 57 Kb contenant le fragment inséré de 3,6 kb absent de la référence 5632 (Figure 12C). Le site d'insertion du marqueur de sélection, retrouvé en amont et en aval du fragment inséré, entre les nt 919899 et 919906, peut-être visualisé de manière précise à l'aide de cet assemblage.

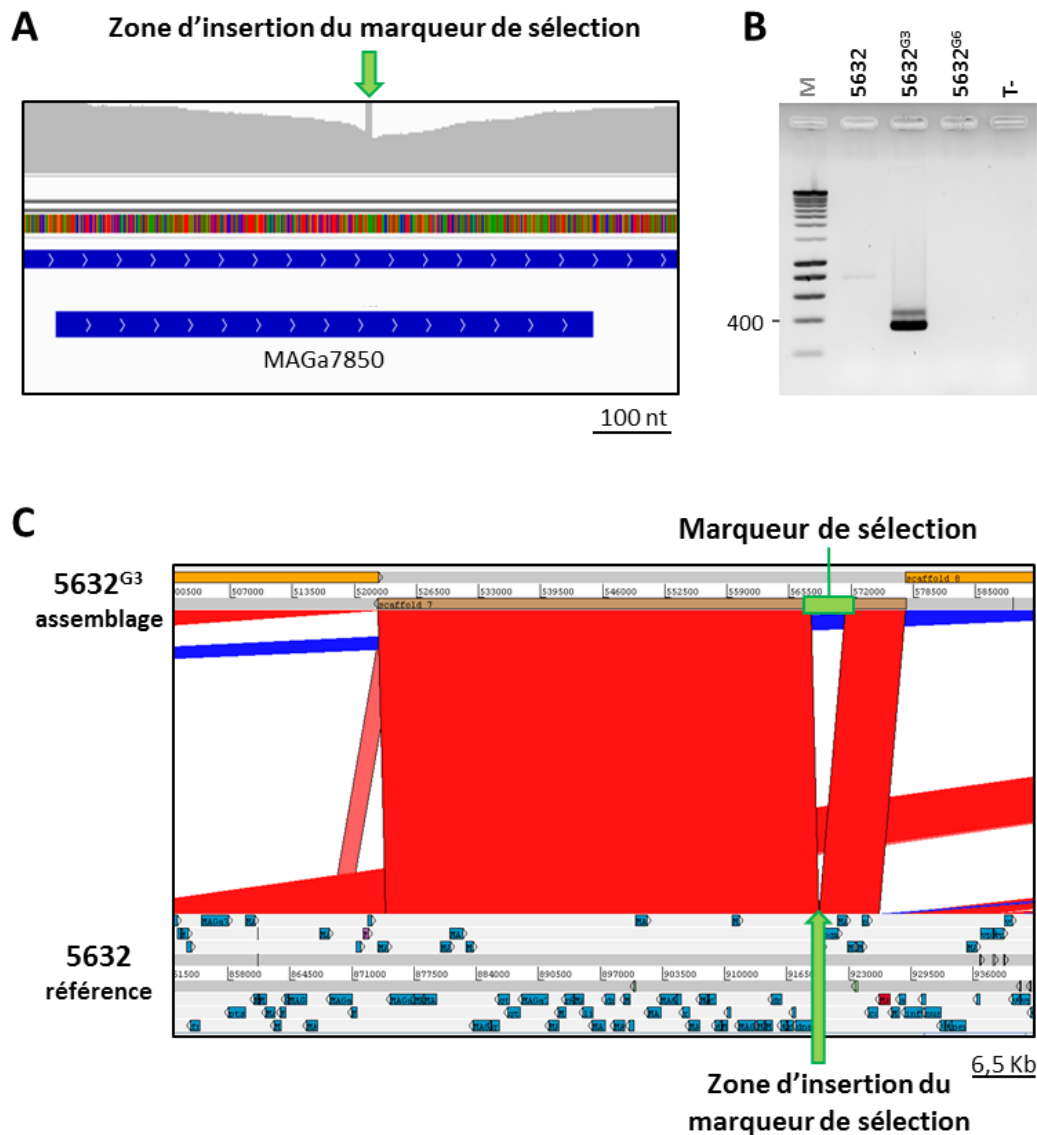


Figure 12 : Détection de l'insertion d'un marqueur de sélection chez 5632^{G3}.

(A) Visualisation par IGV du point d'insertion d'un marqueur de sélection par un alignement de 5632^{G3} sur 5632. La flèche verte désigne le *direct repeat* (DR) généré à l'insertion du transposon (nt 919899-919906). (B) Amplification par PCR du marqueur de sélection chez 5632^{G3} avec les amorces G3 et SG5. Amplicon obtenu : 346 pb. Témoins négatifs : souche 5632 non marquée et clone 5632^{G6} marqué par une insertion à un endroit différent. M : marqueur de taille en pb. Gel d'agarose à 1,8 %. (C) Visualisation avec ACT de l'assemblage *de novo* de 5632^{G3} (en haut) comparé au génome de référence de 5632 (en bas). Les zones rouges indiquent les régions communes aux deux génomes. La région de 3,6 kb représentée en vert sur le *scaffold* de 57 Kb (marron) correspond au transposon porteur du gène de résistance à la gentamicine.

Détection d'une duplication du locus ps3

L'alignement des *reads* du génome de PG2 55-5 séquencé par Illumina sur le génome de PG2 55-5 préalablement séquencé par la méthode Sanger montre une profondeur de couverture moyenne de 7500X (Figure 8B). Toutefois, une forte augmentation du nombre de *reads* (couverture de 17000X) est observée au niveau d'une région de 2,7 Kb (nt 761502-764250) appelée locus ps3 (Figure 13A). Ce locus ps3 est constitué de 4 gènes (Figure 13A) : MAG_6440 présente des homologues avec une protéine de prophage putative ps3, MAG_6460 code une lipoprotéine, MAG_6430 et MAG_6450 codent des protéines hypothétiques. Trois variations sont observées au niveau de ce locus (Figure 13A, Tableau 4) : une substitution synonyme, un SNP intergénique et une délétion de 11 nucléotides, respectivement aux positions 761525, 763036 et 763685. Ces variations ne sont détectées que dans 50 % des *reads* au maximum. Ces résultats sont cohérents avec une duplication du locus et la présence de variations sur une seule des deux copies.

Deux PCR, mises au point pour confirmer cette hypothèse en amplifiant la forme dupliquée et/ou non dupliquée permettent bien l'amplification de deux copies répétées en tandem (Figure 13). Les *reads* courts Illumina (150 nt en moyenne) ne permettent pas d'obtenir un assemblage détectant la duplication sur un seul *contig*. En revanche, l'utilisation de *reads* plus longs (2 kb en moyenne) issus d'un séquençage de PG2 55-5 par PacBio, permet de visualiser la duplication en tandem (Figure 13D). Sur le *contig* obtenu, une séquence de 13 nt (ccattttat) située à l'intérieur des MAG_6430 et MAG_6460 est retrouvée au début et à la fin de chaque copie (Figure 13B). La duplication a eu lieu au niveau de cette séquence et aboutit à des copies tronquées de ces deux gènes. Elle est stable même après 40 passages en culture et est observée sur tous les mutants dérivant de PG2 55-5 (17 clones au total). Cette duplication n'avait pas été détectée lors du séquençage Sanger.

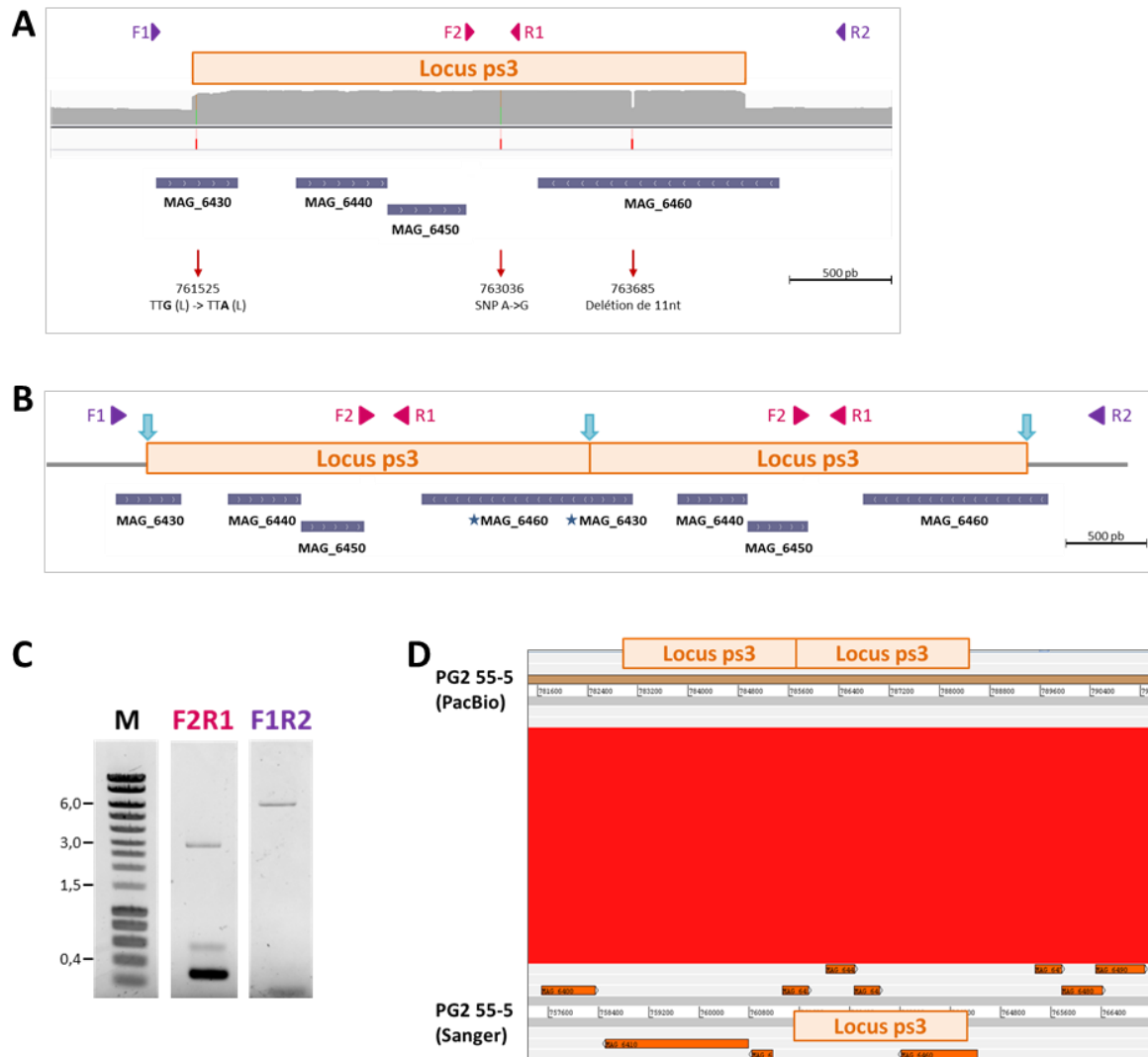


Figure 13 : Duplication du locus ps3 chez PG2 55-5.

(A) L'alignement des *reads* de PG2 55-5 (visualisé avec IGV) montre une couverture plus élevée au niveau du locus ps3 (nt 760801-764980). Les 4 MAG constituant ce locus sont représentées en bleu foncé et les 3 positions de mutations sont indiquées par des traits et des flèches rouges. (B) Représentation du locus ps3 dupliqué et des 4 MAG associées. Les flèches bleu clair indiquent les régions répétées de 13 nt (ccattttatttt). Les flèches violettes et rose désignent les positions des 4 amorces (F1, R1, F2, R2). Les étoiles bleues indiquent les MAG tronquées à cause de la duplication du locus. (C) Amplification par PCR du locus ps3 dupliqué à l'aide des amorces F2R1 et F1R2 chez PG2 55-5. F2R1 : produits de 309 pb et 3045 pb, F1R2 : 6232 pb. M : marqueur de taille en Kb. Gel d'agarose à 0,8 %. (D) Comparaison ACT entre l'assemblage *de novo* de PG2 55-5 obtenu par séquençage PacBio (en haut) et le génome de référence de PG2 55-5 obtenu par séquençage Sanger (en bas).

4. Conclusions

Depuis ces dernières années, des avancées importantes dans le domaine du NGS ont eu lieu. Elles permettent d'obtenir un très grand nombre de données plus rapidement et à moindre coût. Ces progrès aboutissent aujourd'hui à des séquenceurs de très petite taille comme le Minlon (Oxford Nanopore) (Lu *et al.* 2016) qui génère rapidement des *reads* longs d'une dizaine de Kb sur un appareil de la taille d'une clef USB. Le recours au séquençage de génomes complets est donc de plus en plus utilisé pour répondre à des questions biologiques. Plusieurs techniques de NGS et un très grand nombre de logiciels d'analyse des données sont disponibles. Chacun comporte des avantages et des inconvénients qui doivent être pris en compte en fonction de l'analyse que l'on souhaite réaliser. Bien souvent un compromis doit être trouvé entre la profondeur et l'étendue de la couverture que l'on souhaite. Nos travaux démontrent l'importante complémentarité des différents outils d'analyses de séquences pour détecter des variants ou des événements de réorganisation (profondeurs des *reads*, qualité de lecture, *reads* courts et longs...).

Lors de cette thèse, le séquençage de clones de *M. agalactiae* PG2 55-5 et de 5632^{G3} par Illumina confirme les séquences des références PG2 et 5632 obtenues en Sanger mais permet également de détecter une duplication en tandem du locus *ps3*. Les mycoplasmes sont considérés comme des bactéries ayant un taux de mutation élevé du fait de l'absence de systèmes de réparation de l'ADN connus (Delaney *et al.* 2012). Avec 3 ou 10 mutations au bout de 10 ou 40 passages en culture sans pression de sélection, le séquençage de clones de *M. agalactiae* montre toutefois une relative stabilité. Ceci a également été montré chez *M. bovis*, une espèce très proche, avec seulement 46 SNP et une large délétion (14,2 Kb) observés après 180 passages (Rasheed *et al.* 2017). Le génome de 5632 semble plus variable que celui de PG2 puisqu'il présente 10 nouvelles mutations fixées au bout de 10 passages contre seulement 3 pour PG2 au bout de 40 passages. Ces mutations sont essentiellement (6/10) retrouvées au niveau des loci *vpma* de 5632 qui codent des lipoprotéines hypervariables. La souche 5632 possède en effet un répertoire *Vpma* étendu : 23 gènes *vpma* sont répartis en 2 loci chez 5632 contre 6 gènes à un seul locus chez PG2. Le génome de 5632 est aussi dans son ensemble beaucoup plus plastique par rapport à celui de PG2 (Nouvel *et al.* 2009) avec d'autres répertoires de lipoprotéines étendus (*Spma*, *Drp*...) et des éléments génétiques mobiles (*IS* et *ICE*). Tout ceci pourrait contribuer à une labilité plus importante du génome de 5632.

M. agalactiae PG2 55-5 contient un locus ps3 dupliqué en tandem. Trois gènes sur les quatre composant ce locus et d'autres situés à proximité ont été identifiés comme liés à des transferts de gènes horizontaux avec des mycoplasmes du cluster Mycoides (Sirand-Pugnet *et al.* 2007). *M. agalactiae* et les membres du cluster Mycoides (*M. mycoides* subsp. *capri*, *M. mycoides* subsp. *mycoides*, *M. capricolum* subsp. *capricolum*, *M. capricolum* subsp. *capripneumoniae*, *M. leachii*) sont éloignés phylogénétiquement mais partagent les mêmes hôtes ruminants. Un des gènes du locus ps3 (MAG_6440) est également présent chez 5632, il présente une forte homologie (87 % d'homologie nucléotidique) avec CDSE retrouvée au niveau des ICE. Ce gène a été annoté comme une protéine putative de prophage Ps3 car il contient un motif DUF 4065 associé à des protéines phagiques mais dont le rôle exact reste à ce jour inconnu (DUF, *Domain of Unknown Fonction*). Des homologues de cette protéine Ps3 sont aussi identifiés chez d'autres mycoplasmes et d'autres bactéries comme des streptocoques. Chez les mycoplasmes, les homologues de Ps3 ne sont retrouvés que chez des mycoplasmes de ruminants comme *M. bovis* et ceux du cluster Mycoides, à l'intérieur ou à l'extérieur des ICE. Le locus ps3 pourrait être le vestige d'un élément phagique. Si des phages ont été identifiés *in silico* chez les mycoplasmes et notamment chez *M. agalactiae* (Tardy *et al.* 2012), l'étude de la biologie des phages chez les mycoplasmes est un champ encore peu exploré dont les applications fondamentales et thérapeutiques sont évidentes.

A la suite des expériences d'accouplement (*mating*) entre les souches PG2 et 5632, 10 PCR spécifiques de chaque parent ont été utilisées pour caractériser les génomes chimériques des transconjugants. Afin de les analyser plus précisément, ils ont été séquencés par Illumina mais leurs génomes mosaïques compliquent les analyses avec jusqu'à 30 fragments transférés. Une méthode a été mise au point pour définir les fragments de la donneuse (PG2) et de la receveuse (5632) et donc reconstruire la séquence des transconjugants. Elle est applicable lorsque la séquence complète des deux parents est disponible, et qu'un nombre suffisant de variations existe entre les deux parents. L'analyse de nombreux génomes de transconjugants, obtenus dans différentes conditions, permet de mieux comprendre le mécanisme de transfert chromosomique mis en jeu chez les mycoplasmes (Chapitres 1 et 2). Au-delà, l'ensemble de ces travaux a également permis de mettre à la disposition de l'équipe d'accueil une méthode standardisée d'analyse des génomes hybrides et une banque de transconjugants finement définis dont les caractéristiques pourront alimenter de futurs projets de recherche.

5. Matériels et méthodes

Cultures bactériennes

Les clones utilisés pour nos expériences sont PG2 clone 55-5 (Sirand-Pugnet *et al.* 2007), 5632 clones C1 et G3 (Nouvel *et al.* 2010; Dordet-Frisoni *et al.* 2014). Les mycoplasmes ont été cultivés à 37 °C en milieu SP4 (Tully 1995), supplémenté avec 5mM d'acide pyruvique (Sigma-Aldrich) et 45 $\mu\text{g}\cdot\text{mL}^{-1}$ de cefquinome (cobactan 4,5 %, MSD Animal Health). Selon les nécessités de sélection, des antibiotiques comme l'enrofloxacin, la gentamicine ou la tétracycline (Sigma-Aldrich) ont été ajoutés au milieu. Les expériences de mating sont décrites dans les sections matériels et méthodes des chapitres 1 et 2. La sélection des mutants (MF26-1, PG2^{E10}) est décrite dans le chapitre 1.

PCR

Toutes les amplifications PCR ont été réalisées suivant les recommandations du fournisseur de la Taq DNA polymérase (M0267S, *New England Biolabs*). Une PCR, spécifique de position, a été mise au point pour amplifier le marqueur de résistance à la gentamicine chez le clone 5632^{G3} à l'aide des amorces SG5 et G3 à une température d'hybridation de 55°C (Tableau 5). Les clones 5632 et 5632^{G6} ont été utilisés comme témoins négatifs. Les amplicons ont été visualisés par migration sur un gel d'agarose à 1,8 % et révélation au bromure d'éthidium. Deux autres PCR ont été utilisées pour amplifier le locus ps3 dupliqué ou non, à partir d'ADN génomique de PG2 55-5 à l'aide des amorces F2/R1 et F1/R2 (Tableau 5), la température d'hybridation des PCR est de 51°C, et les produits ont migré sur un gel d'agarose à 0,8 %.

Southern blot

Des expériences de *Southern blot* ont permis la visualisation des 3 copies d'ICE et de leur forme circulaire chez les transconjugants et 5632^{G3}. Les ADN génomiques ont été digérés par EcoRV qui contient un seul site dans la CDS22, située à une extrémité des ICE. Les fragments de restriction ont migré sur un gel d'agarose à 1 %, transférés sur membrane de nitrocellulose et hybridés avec une sonde ADN spécifique de la séquence de la CDS22 (22-F) (Tableau 5), marquée à la digoxigénine (Marenda *et al.* 2006). Les tailles des fragments sont de 13,2 Kb pour l'ICE-I; 8,9 Kb pour l'ICE-II; 2,1 Kb pour l'ICE-III et 4,5 Kb pour l'ICEc (circulaire).

Tableau 5: Liste des oligonucléotides

Nom	Séquence (5'→3')	Produit (pb)	Référence
PCR insertion gentamicine 5632 ^{G3}			
G3	CCCGGAAGGTATATTTATGG	346	Skapski <i>et al.</i> 2011
SG5	TTTACACAATTATACGGACTTTATC		
PCR duplication locus ps3			
F2	GTGCTGACAAAAATTGACC	309 ; 3045	
R1	TTTCTAGCACACTCACAAGA		
F1	TCATCACTCTGCATTGAATCT	6332	
R2	AAGTAGTTTAATCACCGCTCA		
<i>Southern blot</i> ICE			
22-F	TGAGACCAGCAAGCTGAAGA	-	Marenda <i>et al.</i> 2006

Séquençage et analyses bioinformatiques

L'extraction des ADN génomiques a été réalisée à l'aide de phénol-chloroforme (Sambrook *et al.* 1989) à partir de cultures de mycoplasmes en milieu SP4. Le séquençage complet de 52 génomes de *M. agalactiae* a été effectué par Illumina, à Eurofin GATC Biotech (Constance, Allemagne) (HiSeq, *paired-end*, 2x150 nt : 2.10⁷ reads, 2700X), à l'exception de PG2^{E10} séquencée par la plateforme Genome-Transcriptome de Bordeaux (France) (HiSeq, *paired-end*, 2x100 nt : 1.6x10⁷ reads, 1700X).

Les analyses bio-informatiques ont été réalisées à l'aide de la plateforme Galaxy (Genotoul, Toulouse, France). La qualité des séquences a été vérifiée avec le logiciel FASTQC. Les *reads* de chaque clone (fastq) ont été alignés sur les génomes de référence PG2 (NC_009497.1) et 5632 (NC_013948.1) avec BWA-MEM (Li et Durbin 2009). La qualité de chaque alignement a été contrôlée avec Qualimap 2.2.1 (Okonechnikov *et al.* 2016). Les analyses de variants ont été effectuées avec GATK3 (McKenna *et al.* 2010) et les variations de qualité inférieure à 10000 ont été exclues. Ces analyses de variants ont aussi été réalisées pour comparaison avec Breseq et Nucmer (Deatherage et Barrick 2014; Kurtz *et al.* 2004). Les alignements (fichier bam) et les variations (fichier vcf) ont été visualisés à l'aide d'IGV 2.3.93 (Thorvaldsdóttir *et al.* 2013), Artemis 16.0.0 (Rutherford *et al.* 2000) et ACT 13.0.0 (Carver *et al.* 2005). Les assemblages *de novo* de PG2 55-5 et 5632^{G3} ont été réalisés, respectivement, à partir de *reads* Pacbio à l'aide de SPAdes v3.11.0 (Bankevich *et al.* 2012) ou Illumina à l'aide du pipeline A15 (Tritt *et al.* 2012).

Liste des outils utilisés sur la plateforme Galaxy (<http://bioinfo.genotoul.fr/>) :

- FastQC: Read QC reports using FastQC (Galaxy Version 0.52)
- Filter a VCF file (Galaxy Version 1.0.0)
- GATK 3 (Galaxy Version 3.5-0)
- Map with BWA-MEM (Galaxy Version 0.8.0)
- SAM-to-BAM convert SAM to BAM (Galaxy Version 2.1)
- SAM to FASTQ (Galaxy Version 0.1) Fastq extraction
- Select lines that match an expression (Galaxy Version 1.0.1)

6. Références

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* mai 2012;19(5):455-77.
- Benita Y, Oosting RS, Lok MC, Wise MJ, Humphery-Smith I. Regionalized GC content of template DNA as a predictor of PCR success. *Nucleic Acids Res.* 15 août 2003;31(16):e99.
- Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics.* 15 août 2005;21(16):3422-3.
- Citti C, Nouvel L-X, Baranowski E. Phase and antigenic variation in mycoplasmas. *Future Microbiology.* juill 2010;5(7):1073-85.
- Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol.* 2014;1151:165-88.
- Delaney NF, Balenger S, Bonneaud C, Marx CJ, Hill GE, Ferguson-Noel N, et al. Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLoS Genet.* févr 2012;8(2):e1002511.
- Delcher AL, Phillippy A, Carlton J, Salzberg SL. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 1 juin 2002;30(11):2478-83.
- Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX, Blanchard A, et al. Chromosomal Transfers in Mycoplasmas: When Minimal Genomes Go Mobile. *mBio.* 31 déc 2014;5(6):e01958-14.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science.* 2 janv 2009;323(5910):133-8.

- Escobar-Zepeda A, Vera-Ponce de León A, Sanchez-Flores A. The Road to Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics. *Front Genet.* 2015;6:348.
- Frese KS, Katus HA, Meder B. Next-Generation Sequencing: From Understanding Biology to Personalized Medicine. *Biology (Basel).* 1 mars 2013;2(1):378-98.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. *Genome Biol.* 2004;5(2):R12.
- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 15 juill 2009;25(14):1754-60.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of next-generation sequencing systems. *J Biomed Biotechnol.* 2012;2012:251364.
- Lu H, Giordano F, Ning Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics Proteomics Bioinformatics.* oct 2016;14(5):265-79.
- Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends in Genetics.* 1 mars 2008;24(3):133-41.
- 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. *Journal of Bacteriology.* 1 juin 2006;188(11):4137-41.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 9 janv 2010;20(9):1297-303.
- Nouvel L-X, Marenda M, Sirand-Pugnet P, Sagné E, Glew M, Mangenot S, et al. Occurrence, Plasticity, and Evolution of the *vpma* Gene Family, a Genetic System Devoted to High-Frequency Surface Variation in *Mycoplasma agalactiae*. *J Bacteriol.* juill 2009;191(13):4111-21.
- Nouvel LX, Sirand-Pugnet P, Marenda MS, Sagné E, Barbe V, Mangenot S, et al. Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro-and micro-events that are shaping mycoplasma diversity. *BMC genomics.* 2010;11(1):86.
- Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics.* 15 janv 2016;32(2):292-4.
- Oyola SO, Otto TD, Gu Y, Maslen G, Manske M, Campino S, et al. Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes. *BMC Genomics.* 3 janv 2012;13:1.
- Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet.* nov 2011;52(4):413-35.

- Pennisi E. The Man Who Bottled Evolution. *Science*. 15 nov 2013;342(6160):790-3.
- Ramanathan B, Jindal HM, Le CF, Gudimella R, Anwar A, Razali R, et al. Next generation sequencing reveals the antibiotic resistant variants in the genome of *Pseudomonas aeruginosa*. *PLoS ONE*. 2017;12(8):e0182524.
- Rasheed MA, Qi J, Zhu X, Chenfei H, Menghwar H, Khan FA, et al. Comparative Genomics of *Mycoplasma bovis* Strains Reveals That Decreased Virulence with Increasing Passages Might Correlate with Potential Virulence-Related Factors. *Front Cell Infect Microbiol*. 2017;7:177.
- Redgrave LS, Sutton SB, Webber MA, Piddock LJV. Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success. *Trends in Microbiology*. août 2014;22(8):438-45.
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics*. oct 2000;16(10):944-5.
- Sambrook J, Fritsch E, Maniatis T. *Molecular cloning: A laboratory manual: Vol. 2*. 2. ed. S.I.: Cold Spring Harbor; 1989.
- Sharma P, Gupta SK, Rolain J-M. Whole genome sequencing of bacteria in cystic fibrosis as a model for bacterial genome adaptation and evolution. *Expert Rev Anti Infect Ther*. mars 2014;12(3):343-55.
- Sims D, Sudbery I, Illott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*. févr 2014;15(2):121-32.
- Sirand-Pugnet P, Lartigue C, Marends M, Jacob D, Barré A, Barbe V, et al. Being Pathogenic, Plastic, and Sexual while Living with a Nearly Minimal Bacterial Genome. *PLoS Genetics*. 2007;3(5):e75.
- Skapski A, Hygonenq M-C, Sagné E, Guiral S, Citti C, Baranowski E. Genome-Scale Analysis of *Mycoplasma agalactiae* Loci Involved in Interaction with Host Cells. Neyrolles O, éditeur. *PLoS ONE*. 23 sept 2011;6(9):e25291.
- Sun G, Luo T, Yang C, Dong X, Li J, Zhu Y, et al. Dynamic Population Changes in *Mycobacterium tuberculosis* During Acquisition and Fixation of Drug Resistance in Patients. *J Infect Dis*. 1 déc 2012;206(11):1724-33.
- Tardy F, Baranowski E, Nouvel L-X, Mick V, Manso-Silvan L, Thiaucourt F, et al. Emergence of Atypical *Mycoplasma agalactiae* Strains Harboring a New Prophage and Associated with an Alpine Wild Ungulate Mortality Episode. *Applied and Environmental Microbiology*. 1 juill 2012;78(13):4659-68.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinformatics*. mars 2013;14(2):178-92.
- Tritt A, Eisen JA, Facciotti MT, Darling AE. An integrated pipeline for de novo assembly of microbial genomes. *PLoS ONE*. 2012;7(9):e42304.

Tully JG. Culture medium formulation for primary isolation and maintenance of mollicutes. In: Molecular and diagnostic procedures in mycoplasmaology: molecular characterization. Academic Press. San Diego: J. G. Tully; 1995. p. 33-9.

**DISCUSSION
ET PERSPECTIVES**

DISCUSSION ET PERSPECTIVES

Depuis plus d'un siècle, les antibiotiques permettent un recul de la mortalité associée aux maladies infectieuses. Leur utilisation est aujourd'hui remise en question par l'augmentation alarmante de la résistance bactérienne devenue un problème majeur de santé publique au niveau mondial. En adéquation avec le concept "Une seule santé", l'OMS, l'OIE et la FAO se sont unies dans la lutte contre la résistance aux antimicrobiens en 2015 (http://www.who.int/foodsafety/areas_work/antimicrobial-resistance/tripartite/en/). Cette lutte bénéficie de la mobilisation de la communauté scientifique pour une meilleure compréhension que ce soit au niveau moléculaire ou sociologique.

Bien que causées par des bactéries atypiques sans paroi, les mycoplasmoses, maladies qui touchent l'Homme et l'ensemble des animaux, ne font pas exception au problème d'antibiorésistance (Waites *et al.* 2014). Les HGT constituent des mécanismes majeurs d'acquisition de la résistance chez la plupart des bactéries (Barlow 2009; von Wintersdorff *et al.* 2016), ils ont longtemps été considérés comme absents chez les mycoplasmes lesquels contiennent peu de plasmides (Marenda 2014). La description d'un transposon conjugatif (ICE) porteur d'un gène de résistance à la tétracycline chez *Mycoplasma hominis* (Roberts *et al.* 1985) et la détection d'échanges de gènes *in silico* chez de nombreux mycoplasmes (Sirand-Pugnet *et al.* 2007a; Pereyre *et al.* 2009; Vasconcelos *et al.* 2005; Lo *et al.* 2015; Chang *et al.* 2014) justifient un questionnement nouveau sur les HGT et l'antibiorésistance chez les mycoplasmes. La détection de HGT massifs et la maîtrise d'une reproduction expérimentale de transferts chromosomiques *in vitro* en présence d'ICE chez *M. agalactiae* (Dordet Frisoni *et al.* 2013) font de cette espèce un modèle d'étude du HGT chez les mycoplasmes. Ces travaux de thèse se sont attachés à étudier, chez *M. agalactiae*, un mécanisme conjugatif original appelé MCT (*Mycoplasma Chromosomal Transfer*) et à évaluer son impact sur l'acquisition de la résistance aux antibiotiques.

Le séquençage global de pools de transconjugants a précédemment montré au laboratoire que tout fragment du génome de la souche donneuse pouvait être transféré par MCT (Dordet-Frisoni *et al.* 2014) mais des informations sur plusieurs génomes individuels de transconjugants constituaient un manque. Le séquençage génomique complet de transconjugants, associé à la définition des fragments échangés par une méthode mise au point lors de cette thèse, ont révélé le caractère distributif du MCT et la diversité génétique insoupçonnée qu'il produit. En effet, des fragments aléatoires autres que ceux associés aux marqueurs de sélection sont

transférés et les transconjugants présentent des génomes mosaïques uniques. Les transconjugants comportent de 3 à 30 fragments transférés (en moyenne 14) dont la taille est très diverse (de 49 à 64673 pb). L'échange de ces fragments peut entraîner le gain ou la perte de larges régions mais aussi produire des événements de micro-complexité pouvant avoir un impact sur des régulations génétiques fines. Ces réarrangements génomiques massifs générés par MCT pourraient avoir des conséquences importantes sur de nombreux processus adaptatifs comme la résistance aux antibiotiques.

Notre étude montre que la résistance aux fluoroquinolones acquise par l'accumulation de mutations spontanées dans les gènes cibles (*gyrA*, *gyrB*, *parC* et *parE*) peut être transférée horizontalement par MCT chez *M. agalactiae*. Le nombre de générations nécessaires à l'obtention de transconjugants résistants est largement inférieur au nombre de générations nécessaires à la sélection de mutants spontanés équivalents. Le MCT accélère donc l'acquisition de la résistance par le transfert de plusieurs mutations chromosomiques distantes en une seule étape. Bien que la réalité de ces observations *in vitro* reste à démontrer *in vivo*, ce transfert de résistance pourrait avoir de lourdes conséquences en termes de diffusion de la résistance dans les cas d'infections naturelles. Des études de détection de mosaïcisme *in silico* sur les séquences de souches résistantes circulantes pourraient permettre de révéler des événements de MCT associés à la résistance *in vivo*. Ces études pourraient être réalisées sur des isolats de *M. bovis*, mycoplasme très proche de *M. agalactiae* et également capable de MCT (Dordet Frisoni *et al.* 2013), pour lequel des niveaux élevés de résistance sont décrits en France (Gautier-Bouchardon *et al.* 2014).

Le mosaïcisme généré par MCT peut aussi perturber le diagnostic des mycoplasmoses. Il pourrait expliquer les difficultés de typage sérologique rencontrées par exemple chez les mycoplasmes du cluster Mycoïdes ou ceux du genre *Ureaplasma* (Le Grand *et al.* 2004; Tardy *et al.* 2009; Xiao *et al.* 2011; Paralanov *et al.* 2012). De même, l'épidémiologie moléculaire basée sur des techniques de typage ADN comme la MLST (*MultiLocus Sequence Typing*) pourrait se trouver biaisée. Des événements de MCT expliqueraient par exemple les incongruences de topologie d'arbres phylogénétiques basés sur des gènes individuels ou concaténés comme mis en évidence pour le gène *lepA* mosaïque entre *M. mycoïdes* et *M. capricolum* (Manso-Silvan *et al.* 2007) ou pour le gène *fusA* entre *M. bovis* et *M. agalactiae* (Manso-Silvan *et al.* 2012).

Le MCT peut aussi avoir un impact sur d'autres processus adaptatifs comme le *fitness* qui se trouve modifie chez certains transconjugants. Les combinaisons genetiques obtenues peuvent

ainsi, dans certains cas, entraîner un gain de *fitness* qui améliore la croissance des transconjugants par rapport à la souche receveuse. Ces transconjugants pourraient donc rapidement envahir leur environnement. Dans d'autres cas, les nouvelles combinaisons apparaissent coûteuses en termes de *fitness*. Toutefois, ces transconjugants qui présentent un *fitness* réduit retrouvent un *fitness* équivalent à la souche receveuse après 10 à 40 passages en culture en l'absence de pression de sélection antibiotique. Un temps d'ajustement semble donc parfois nécessaire pour s'adapter à certaines compositions génétiques. Il est aussi probable que certaines combinaisons ne soient pas viables, et ne sont donc pas observées. (Baltrus 2013; San Millan et MacLean 2017).

Le MCT nécessite une reprogrammation du génome, comme observé dans les expériences de biologie synthétique avec la transplantation de génome entre deux espèces de mycoplasmes (Lartigue *et al.* 2007). La machinerie cellulaire de la souche receveuse doit être capable de lire les informations du génome de la souche donneuse. Il a été montré que l'efficacité de transplantation diminue avec l'augmentation de la distance phylogénétique entre les deux espèces partenaires. Par exemple, *M. mycoides* subsp. *capri* ou même *M. leachii* peuvent être transplantés avec succès chez *M. capricolum* subsp. *capricolum*, mais pas *S. citri* (Labroussaa *et al.* 2016). Jusqu'à présent, du MCT inter-espèces a été mis en évidence entre *M. agalactiae* et *M. bovis* (Dordet-Frisoni *et al.* 2014) mais n'a pas pu être obtenu *in vitro* entre *M. agalactiae* et d'autres espèces comme les mycoplasmes du cluster Mycoides qui certes, partagent le même hôte ruminant, mais sont phylogénétiquement très distantes. À l'image des observations faites pour la transplantation génomique, l'effet de la distance phylogénétique, associé à celui de la différence d'hôte, mériterait d'être testé par d'autres tentatives de *matings* inter-espèces, des expériences sont actuellement menées sur ce sujet au laboratoire.

Des transferts chromosomiques intra-espèces chez un mycoplasme humain, *M. genitalium*, ont récemment été démontrés *in vitro* (Torres-Puig *et al.* 2018). Tout comme le MCT, le mécanisme de ces transferts s'opère en présence de DNase, toutefois aucune ICE n'a été détectée chez *M. genitalium* et le caractère distributif n'a pas été documenté. Ces transferts sont conditionnés à la surexpression d'un facteur sigma alternatif (σ^{20}) qui favoriserait les recombinaisons génomiques. L'évaluation des niveaux d'expression, la surexpression ou l'inhibition de tels facteurs, chez *M. agalactiae* en conditions de *mating*, restent à étudier.

Les HGT sont très souvent abordés par le transfert d'éléments génétiques mobiles mais des études récentes mettent en évidence de plus en plus de mécanismes de transferts chromosomiques au sein de genre bactériens divers comme *Mycoplasma*, *Mycobacterium*,

Bacteroides, *Thermus*, *Streptomyces* et *Yersinia* (Dordet-Frisoni *et al.* 2014; Derbyshire et Gray 2014; Husain *et al.* 2017; Blesa *et al.* 2017; Thoma et Muth 2012; Lesic *et al.* 2012). Ces données suggèrent que le transfert chromosomique pourrait être un mécanisme répandu chez les bactéries. Il pourrait avoir un impact majeur sur leur évolution et leur adaptation, en participant notamment à la diversité et à la plasticité génétique des populations bactériennes. La possible universalité de tels remaniements génomiques pose la question des forces évolutives qui maintiennent la frontière entre espèces.

Parmi ces études, un mécanisme de transfert conjugatif et distributif, appelé DCT (*Distributive Conjugal Transfer*), similaire au MCT de *M. agalactiae*, est décrit, a minima *in silico*, chez des espèces de plusieurs genres mycobactériens : *Mycolicibacterium*, *Mycobacteroides* et *Mycobacterium* (Derbyshire et Gray 2014; Mortimer et Pepperell 2014; Boritsch *et al.* 2016; Sapriel *et al.* 2016). Le DCT a été observé *in vitro* chez *Mycobacterium smegmatis*¹ (Wang *et al.* 2003) et chez *M. canettii* (Boritsch *et al.* 2016). Bien caractérisé chez *M. smegmatis* (Gray *et al.* 2013), le DCT entraîne le remplacement de fragments génomiques de la receveuse par ceux de la donneuse par recombinaison homologue et génère des génomes hautement mosaïques comme le MCT. En revanche, il est conditionné à la présence d'un locus Esx qui code un système de sécrétion de type VII (Derbyshire et Gray 2014) alors que le MCT requiert, chez le receveur, la présence d'une ICE qui porte les gènes d'un système de sécrétion de type IV. Si les systèmes de sécrétion peuvent avoir une implication forte dans les transferts de gènes comme par exemple en association à de la prédation chez *Vibrio cholerae* (Borgeaud *et al.* 2015), leur rôle dans le DCT comme le MCT semble se concentrer autour du pore de conjugaison mais sa structure et sa mise en place restent à préciser.

La très grande majorité des transferts horizontaux décrits fait appel, directement ou indirectement, à des éléments génétiques mobiles (EGM). L'analyse des EGM de 384 génomes bactériens indique qu'ils sont présents en plus grande quantité chez les bactéries intracellulaires facultatives que chez les bactéries strictement extracellulaires et intracellulaires obligatoires (Newton et Bordenstein 2011). Cependant, chez les MRE (*Mycoplasma-Related Endobacteria*), bactéries intracellulaires obligatoires apparentées au *Mollicutes* qui résident dans les cellules eucaryotes de champignon, les EGM sont abondants et représentent environ 5 % du génome (Naito *et al.* 2015; Torres-Cortés *et al.* 2015). La présence d'EGM dans les génomes permet non seulement de mobiliser de l'ADN mais également de

¹ Espèce récemment renommée *Mycolicibacterium smegmatis*: VALIDATION LIST no. 181. *Int. J. Syst. Evol. Microbiol.*, 68 (2018), 1411-1417.

créer des régions répétées, utilisées lors de recombinaisons (Thomas et Nielsen 2005). Ainsi, la présence d'EGM serait cruciale pour ces MRE endosymbiontes en leur permettant d'éviter, grâce au HGT, l'extinction selon l'hypothèse du cliquet de Muller (schéma évolutif d'accumulation de mutations délétères pouvant conduire les petites populations bactériennes à l'extinction) (Takeuchi *et al.* 2014; Naito et Pawlowska 2016a). Le processus de HGT serait pour les MRE un déterminant vital du brassage génétique et contribueraient à l'hétérogénéité des populations et à leur longévité évolutive (Naito et Pawlowska 2016b). Cette hypothèse pourrait être extrapolée à l'ensemble des mycoplasmes capable de MCT.

Des études indiquent que l'évolution des bactéries serait plus souvent gouvernée par les HGT que par les mutations (Ochman *et al.* 2000; Wiedenbeck et Cohan 2011). L'analyse de 116 génomes de procaryotes décèle 14 % des ORF provenant d'événement de HGT (Nakamura *et al.* 2004). Ces gènes transférés font partie de trois catégories de fonctions bactériennes : l'enveloppe cellulaire, la régulation de l'ADN et le pouvoir pathogène ou la virulence (production de toxine et résistance aux ATB). De nombreux gènes sont acquis par transfert dans les génomes indépendamment d'une pression de sélection directe, ils peuvent ensuite dériver, être éliminés ou conservés suivant leur apport adaptatif (Wiedenbeck et Cohan 2011). Le pouvoir adaptatif d'un gène donné dépend du contexte génétique de la bactérie et de l'environnement dans lequel elle se trouve. Un gène transféré peut avoir un rôle pour une bactérie environnementale qui n'est pas le même que pour une bactérie pathogène (Martínez 2018). Le HGT a notamment un rôle majeur dans l'adaptation des bactéries pathogènes et donc sur les pathologies qu'elles entraînent, avec le transfert, entre autres, de facteurs de virulence ou de gènes de résistance aux antibiotiques (Martínez 2018; Partridge *et al.* 2018).

Un exemple remarquable de l'implication de HGT dans l'évolution d'une espèce pathogène est l'agent de la peste, *Yersinia pestis* qui a évolué très rapidement à partir de son ancêtre *Y. pseudotuberculosis*. Cette évolution s'est faite par l'acquisition de gènes de virulence portés par des plasmides, associée à une réduction génomique suite à de nombreux réarrangements médiés par des *IS*. Ces pertes de gènes ont entraîné la perte de toxines vis-à-vis d'insectes et ainsi autorisé une nouvelle voie de transmission de cette bactérie par les puces (Wren 2003).

De même chez les *Mollicutes*, une étude récente (Lo *et al.* 2018) a montré que les mycoplasmes du cluster *Mycoides*, adaptés aux ruminants tels que nous les connaissons aujourd'hui, sont la résultante d'une succession de gains et pertes de gènes chez un mesoplasme/entomoplasme ancestral, commensal d'insectes. Ce changement d'hôte a probablement été rendu possible par les transferts de gènes (18 % du génome) mis en

évidence entre *M. agalactiae* et les mycoplasmes du cluster Mycoides qui sont phylogénétiquement distants (Sirand-Pugnet *et al.* 2007b). Des évènements de MCT pourraient être impliqués dans ces transferts de gènes massifs répartis à plusieurs endroits du génome. Ces échanges questionnent tout de même sur la niche écologique, lieu de rencontre ou de co-infections entre mesoplasmes/entomoplasmes et mycoplasmes de ruminants. Des pistes sont envisageables comme des insectes piqueurs de ruminants, des acariens des fourrages ou les ruminants eux-mêmes... Des mycoplasmes du groupe Mycoides peuvent, en effet, être isolés d'acariens présents dans des oreilles de chèvres, ces dernières constituent une niche où siègent des mélanges d'espèces mycoplasmiques (Tardy *et al.* 2011).

Finalement, ces transferts de gènes ont contribué à des incongruences entre la taxonomie et la phylogénie : des espèces du genre *Mycoplasma* qui infectent les ruminants sont phylogénétiquement plus proches des *Entomoplasma* et *Spiroplasma* qui infectent les insectes et les plantes ! S'il est toujours difficile de réconcilier phylogénie et taxonomie, une vague de phylogénie déferle actuellement sur la taxonomie, avec l'avènement de la génomique. Certains auteurs comme Gupta et ses collaborateurs ont ainsi récemment essayé une réconciliation (Gupta *et al.* 2018). Comme celle-ci passe par la création d'un ordre, de deux familles et de cinq genres supplémentaires (*Mycoplasmoides*, *Malomycoplasma*, *Metamycoplasma*, *Mesomycoplasma*, *Mycoplasmaopsis*, etc), cette réconciliation soulève aussi un vent de discorde entre spécialistes où bactériologistes et même mycoplasmologistes risquent de se perdre. Le HGT a donc aussi des conséquences sur les querelles scientifiques. Que l'histoire et les HGT donnent raison aux uns ou aux autres, ce qui est certain c'est qu'il faudra toujours s'adapter et que les bactéries sont pour cela un modèle d'innovations.

Les études de métagénomique conduisent aujourd'hui nos réflexions vers une vision plus globale et vers une prise en compte du concept de l'holobionte. Par exemple, des études décrivent le microbiote intestinal comme un réservoir immense de gènes de résistance portés par des bactéries commensales bénéfiques, voire inconnues et non cultivables (Sommer *et al.* 2009; Forslund *et al.* 2014). Cette source immense d'information bactérienne doit être prise en compte dans le problème mondial de l'antibiorésistance. La métagénomique montre que les bactéries doivent être considérées comme appartenant à un continuum d'écosystème où se tissent de nombreux échanges et interactions (Theis *et al.* 2016) et le transfert chromosomique pourrait être un important contributeur de l'écologie microbienne.

Références :

- Baltrus DA. Exploring the costs of horizontal gene transfer. *Trends in Ecology & Evolution*. 1 août 2013;28(8):489-95.
- Barlow M. What antimicrobial resistance has taught us about horizontal gene transfer. *Methods Mol Biol*. 2009;532:397-411.
- Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. The transjugation machinery of *Thermus thermophilus*: Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet*. mars 2017;13(3):e1006669.
- Borgeaud S, Metzger LC, Scignari T, Blokesch M. The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science*. 2 janv 2015;347(6217):63-7.
- Boritsch EC, Khanna V, Pawlik A, Honoré N, Navas VH, Ma L, *et al*. Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. *Proc Natl Acad Sci USA*. 30 2016;113(35):9876-81.
- Chang T-H, Lo W-S, Ku C, Chen L-L, Kuo C-H. Molecular evolution of the substrate utilization strategies and putative virulence factors in mosquito-associated *Spiroplasma* species. *Genome Biol Evol*. mars 2014;6(3):500-9.
- Derbyshire KM, Gray TA. Distributive Conjugal Transfer: New Insights into Horizontal Gene Transfer and Genetic Exchange in Mycobacteria. *Microbiol Spectr*. 2014;2(1).
- Dordet Frisoni E, Marena MS, Sagné E, Nouvel LX, Guérillot R, Glaser P, *et al*. ICEA of *Mycoplasma agalactiae* : a new family of self-transmissible integrative elements that confers conjugative properties to the recipient strain: *Mycoplasma* ICE on the move. *Molecular Microbiology*. sept 2013;89(6):1226-39.
- Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX, Blanchard A, *et al*. Chromosomal Transfers in Mycoplasmas: When Minimal Genomes Go Mobile. *mBio*. 31 déc 2014;5(6):e01958-14.
- Forslund K, Sunagawa S, Coelho LP, Bork P. Metagenomic insights into the human gut resistome and the forces that shape it. *Bioessays*. mars 2014;36(3):316-29.
- Gautier-Bouchardon AV, Ferré S, Le Grand D, Paoli A, Gay E, Poumarat F. Overall Decrease in the Susceptibility of *Mycoplasma bovis* to Antimicrobials over the Past 30 Years in France. *PLoS ONE*. 4 févr 2014;9(2):e87672.
- Gray TA, Krywy JA, Harold J, Palumbo MJ, Derbyshire KM. Distributive conjugal transfer in mycobacteria generates progeny with meiotic-like genome-wide mosaicism, allowing mapping of a mating identity locus. *PLoS Biol*. juill 2013;11(7):e1001602.
- Gupta RS, Sawhani S, Adeolu M, Alnajjar S, Oren A. Phylogenetic framework for the phylum Tenericutes based on genome sequence data: proposal for the creation of a new order Mycoplasmoidales ord. nov., containing two new families Mycoplasmoidaceae fam. nov. and Metamycoplasmataceae fam. nov. harbouring Eperythrozoon, Ureaplasma and five novel genera. *Antonie Van Leeuwenhoek*. sept 2018;111(9):1583-630.

- Husain F, Tang K, Veeranagouda Y, Boente R, Patrick S, Blakely G, *et al.* Novel large-scale chromosomal transfer in *Bacteroides fragilis* contributes to its pan-genome and rapid environmental adaptation. *Microb Genom.* nov 2017;3(11).
- Labroussaa F, Lebaudy A, Baby V, Gourgues G, Matteau D, Vashee S, *et al.* Impact of donor-recipient phylogenetic distance on bacterial genome transplantation. *Nucleic Acids Res.* 30 2016;44(17):8501-11.
- Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA, *et al.* Genome transplantation in bacteria: changing one species to another. *Science.* 3 août 2007;317(5838):632-8.
- Le Grand D, Saras E, Blond D, Solsona M, Poumarat F. Assessment of PCR for routine identification of species of the *Mycoplasma mycoides* cluster in ruminants. *Vet Res.* déc 2004;35(6):635-49.
- Lesic B, Zouine M, Ducos-Galand M, Huon C, Rosso M-L, Prévost M-C, *et al.* A natural system of chromosome transfer in *Yersinia pseudotuberculosis*. *PLoS Genet.* 2012;8(3):e1002529.
- Lo W-S, Gasparich GE, Kuo C-H. Found and Lost: The Fates of Horizontally Acquired Genes in Arthropod-Symbiotic *Spiroplasma*. *Genome Biol Evol.* 8 août 2015;7(9):2458-72.
- Lo W-S, Gasparich GE, Kuo C-H. Convergent evolution among ruminant-pathogenic *Mycoplasma* involved extensive gene content changes. *Genome Biol Evol.* 8 août 2018;
- Manso-Silván L, Dupuy V, Lysnyansky I, Ozdemir U, Thiaucourt F. Phylogeny and molecular typing of *Magalactiae* and *Mbovis* by multilocus sequencing. *Veterinary Microbiology.* 28 déc 2012;161(1-2):104-12.
- Manso-Silván L, Perrier X, Thiaucourt F. Phylogeny of the *Mycoplasma mycoides* cluster based on analysis of five conserved protein-coding sequences and possible implications for the taxonomy of the group. *Int J Syst Evol Microbiol.* oct 2007;57(Pt 10):2247-58.
- Marenda MS. Genomic mosaics. In: *Mollicutes: Molecular biology and pathogenesis.* Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browning; 2014. p. 15-54.
- Martínez JL. Ecology and Evolution of Chromosomal Gene Transfer between Environmental Microorganisms and Pathogens. *Microbiol Spectr.* janv 2018;6(1).
- Mortimer TD, Pepperell CS. Genomic signatures of distributive conjugal transfer among mycobacteria. *Genome Biol Evol.* 30 août 2014;6(9):2489-500.
- Naito M, Morton JB, Pawlowska TE. Minimal genomes of mycoplasma-related endobacteria are plastic and contain host-derived genes for sustained life within Glomeromycota. *Proc Natl Acad Sci USA.* 23 juin 2015;112(25):7791-6.
- Naito M, Pawlowska TE. Defying Muller's Ratchet: Ancient Heritable Endobacteria Escape Extinction through Retention of Recombination and Genome Plasticity. *mBio.* 7 juin 2016a;7(3):e02057-15.
- Naito M, Pawlowska TE. The role of mobile genetic elements in evolutionary longevity of heritable endobacteria. *Mob Genet Elements.* févr 2016b;6(1):e1136375.

- Nakamura Y, Itoh T, Matsuda H, Gojobori T. Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nature Genetics*. juill 2004;36(7):760-6.
- Newton ILG, Bordenstein SR. Correlations between bacterial ecology and mobile DNA. *Curr Microbiol*. janv 2011;62(1):198-208.
- Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. mai 2000;405(6784):299-304.
- Paralanov V, Lu J, Duffy LB, Crabb DM, Shrivastava S, Methé BA, *et al*. Comparative genome analysis of 19 *Ureaplasma urealyticum* and *Ureaplasma parvum* strains. *BMC Microbiol*. 30 mai 2012;12:88.
- Partridge SR, Kwong SM, Firth N, Jensen SO. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clinical Microbiology Reviews*. 1 oct 2018;31(4):e00088-17.
- Pereyre S, Sirand-Pugnet P, Beven L, Charron A, Renaudin H, Barré A, *et al*. Life on arginine for *Mycoplasma hominis*: clues from its minimal genome and comparison with other human urogenital mycoplasmas. *PLoS Genet*. oct 2009;5(10):e1000677.
- Roberts MC, Koutsky LA, Holmes KK, LeBlanc DJ, Kenny GE. Tetracycline-resistant *Mycoplasma hominis* strains contain streptococcal tetM sequences. *Antimicrob Agents Chemother*. juill 1985;28(1):141-3.
- San Millan A, MacLean RC. Fitness Costs of Plasmids: a Limit to Plasmid Transmission. *Microbiol Spectr*. 2017;5(5).
- Sapriel G, Konjek J, Orgeur M, Bourli L, Frézal L, Roux A-L, *et al*. Genome-wide mosaicism within *Mycobacterium abscessus*: evolutionary and epidemiological implications. *BMC Genomics*. 17 févr 2016;17:118.
- Sirand-Pugnet P, Citti C, Barré A, Blanchard A. Evolution of mollicutes: down a bumpy road with twists and turns. *Research in Microbiology*. déc 2007a;158(10):754-66.
- Sirand-Pugnet P, Lartigue C, Marena M, Jacob D, Barré A, Barbe V, *et al*. Being Pathogenic, Plastic, and Sexual while Living with a Nearly Minimal Bacterial Genome. *PLoS Genetics*. 2007b;3(5):e75.
- Sommer MOA, Dantas G, Church GM. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science*. 28 août 2009;325(5944):1128-31.
- Takeuchi N, Kaneko K, Koonin EV. Horizontal gene transfer can rescue prokaryotes from Muller's ratchet: benefit of DNA from dead cells and population subdivision. *G3 (Bethesda)*. 19 févr 2014;4(2):325-39.
- Tardy F, Maigre L, Poumarat F, Citti C. Identification and distribution of genetic markers in three closely related taxa of the *Mycoplasma mycoides* cluster: refining the relative position and boundaries of the *Mycoplasma sp. bovine* group 7 taxon (*Mycoplasma leachii*). *Microbiology (Reading, Engl)*. nov 2009;155(Pt 11):3775-87.
- Tardy F, Maigre L, Tricot A, Poumarat F, Nguyen L, Le Grand D. Comparison of isolates of *Mycoplasma mycoides* subspecies capri from asymptomatic and septicemic goats. *J Comp Pathol*. janv 2011;144(1):70-7.

- Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TCG, *et al.* Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems*. avr 2016;1(2).
- Thoma L, Muth G. Conjugal DNA transfer in *Streptomyces* by TraB: is one protein enough? *FEMS Microbiol Lett*. déc 2012;337(2):81-8.
- Thomas CM, Nielsen KM. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*. sept 2005;3(9):711-21.
- Torres-Cortés G, Ghignone S, Bonfante P, Schüßler A. Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: Transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proc Natl Acad Sci USA*. 23 juin 2015;112(25):7785-90.
- Torres-Puig S, Martínez-Torró C, Granero-Moya I, Querol E, Piñol J, Pich OQ. Activation of σ^{20} -dependent recombination and horizontal gene transfer in *Mycoplasma genitalium*. *DNA Res*. 11 avr 2018;
- Vasconcelos ATR, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, *et al.* Swine and Poultry Pathogens: the Complete Genome Sequences of Two Strains of *Mycoplasma hyopneumoniae* and a Strain of *Mycoplasma synoviae*. *J Bacteriol*. août 2005;187(16):5568-77.
- Waites KB, Lysnyansky I, Bébéar C. Emerging antimicrobial resistance in mycoplasmas of humans and animals. In: *Mollicutes: Molecular biology and pathogenesis*. Caister Academic Pr. Norfolk: Christine Citti and Glenn F. Browning; 2014. p. 289-322.
- Wang J, Parsons LM, Derbyshire KM. Unconventional conjugal DNA transfer in mycobacteria. *Nat Genet*. mai 2003;34(1):80-4.
- Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev*. sept 2011;35(5):957-76.
- von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, *et al.* Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol*. 2016;7:173.
- Wren BW. The *Yersinia* — a model genus to study the rapid evolution of bacterial pathogens. *Nature Reviews Microbiology*. oct 2003;1(1):55-64.
- Xiao L, Paralanov V, Glass JI, Duffy LB, Robertson JA, Cassell GH, *et al.* Extensive horizontal gene transfer in ureaplasmas from humans questions the utility of serotyping for diagnostic purposes. *J Clin Microbiol*. août 2011;49(8):2818-26.

Annexes

Tableaux annexes

Les tableaux ci-dessous présentent les positions des mutations chez les mycoplasmes résistants aux tétracyclines (Tableau A1), aux macrolides (Tableau A2) et aux fluoroquinolones (Tableau A3). Pour les ARNr 16S et 23S les positions sont en nucléotides, pour les protéines elles sont en acides aminées codé par une (L4, L22) ou 3 lettres (GyrA, GyrB, ParC, ParE). Toutes les positions sont numérotées par référence à *E. coli*. Un tiret signifie qu'aucune mutation n'a été détectée, et "nd" que cette région n'a pas été séquencée. Lorsque des isolats cliniques sont étudié "*in vivo*" est indiqué, pour des mutants sélectionnées en laboratoire "*in vitro*" est indiqué.

Tableau A1: Positions des mutations chez les mutants résistants aux tétracyclines

Espèce	Groupe phylo	Hôte	ARN 16S	tet(M)	<i>in vivo</i> / <i>in vitro</i>	Référence
<i>M. pneumoniae</i>	Pneumoniae	Homme	968, 1193	nd	<i>in vitro</i>	(Dégrange <i>et al.</i> 2008b)
<i>U. urealyticum</i>	Pneumoniae	Homme	nd	tet(M)	<i>in vivo</i>	(Roberts et Kenny 1986)
<i>M. hominis</i>	Hominis	Homme	346, 965, 966, 967, 1054	nd	<i>in vitro</i>	(Dégrange <i>et al.</i> 2008b)
			-	tet(M)	<i>in vivo</i>	(Dégrange <i>et al.</i> 2008a)
<i>M. bovis</i>	Hominis	Ruminants	965, 966, 967, 1058, 1195, 1196	-	<i>in vitro</i>	(Sulyok <i>et al.</i> 2016)
			965, 967, 1012, 1153, 1268	-	<i>in vivo</i>	
			965, 967, 1058	-	<i>in vivo</i>	(Amram <i>et al.</i> 2015)
			965, 967	nd	<i>in vivo</i>	(Khalil <i>et al.</i> 2017)
<i>M. agalactiae</i>	Hominis	Ruminants	320, 376, 458, 461, 691, 723, 965, 966, 967, 1184, 1272, 1304	nd	<i>in vivo</i>	(Prats-van der Ham <i>et al.</i> 2018)
<i>M. mycoides</i> subsp. <i>capri</i>	Spiroplasma	Ruminants	212, 735, 1290	nd	<i>in vivo</i>	(Prats-van der Ham <i>et al.</i> 2018)

Tableau A2 : Positions des mutations chez les mutants résistants aux macrolides.

Espèce	Groupe phylo	Hôte	ARN 23S (domaine V)	ARN 23S (domaine II)	L4	L22	<i>in vivo</i> / <i>in vitro</i>	Référence
<i>M. genitalium</i>	Pneumoniae	Homme	A2058G, A2059G	nd	nd	nd	<i>in vitro</i>	(Tagg <i>et al.</i> 2013)
<i>M. pneumoniae</i>	Pneumoniae	Homme	A2062G, C2611A	-	H70R	ΔIPRA 111-114, A114T, P112R	<i>in vitro</i>	(Pereyre <i>et al.</i> 2004)
<i>U. parvum</i>	Pneumoniae	Homme	G2056U, G2057U/A, G2058G	-	W65R, T70K, G71V	K91N	<i>in vitro</i>	(Pereyre <i>et al.</i> 2007)
<i>M. gallisepticum</i>	Pneumoniae	Aviaire	G2057A, A2058G, A2059G, A2503U	nd	nd	nd	<i>in vitro</i>	(Wu <i>et al.</i> 2005)
			A2058G, A2059G	-	-	-	<i>in vivo</i>	(Gerchman <i>et al.</i> 2011)
<i>M. hominis</i>	Hominis	Homme	A2059G, C2611U	-	-	-	<i>in vivo</i>	(Pereyre <i>et al.</i> 2002)
			C2056A, A2059G, A2062C, G2576A, C2586U, G2608C, C2611U	G792A	H184L	R79K	<i>in vitro</i>	(Pereyre <i>et al.</i> 2006)
<i>M. bovis</i>	Hominis	Ruminants	A2058G	-	-	-	<i>in vivo</i>	(Kong <i>et al.</i> 2016)
			A2058G, A2059C/G	G748A, C752U	G185V/L/A/R/W, T186P	Q90H	<i>in vivo</i>	(Lerner <i>et al.</i> 2014)
			A2059G	A534U, G748A, G954A, U1248C, C1371U	-	-	<i>in vivo</i>	(Sulyok <i>et al.</i> 2016)
			A2059G/U, A2063U, C2067U	A534U, G748A/C, 753GUG, C1371U	-	-	<i>in vitro</i>	
			A2058G	G748A	-	-	<i>in vivo</i>	(Khalil <i>et al.</i> 2017)
<i>M. agalactiae</i>	Hominis	Ruminants	A2058G	-	-	S89L, Q90K/H	<i>in vivo</i>	(Prats-van der Ham <i>et al.</i> 2017)
			A2059G, C2611U	-	-	S89L, Q90K	<i>in vitro</i>	
<i>M. capricolum</i> subsp. <i>capricolum</i>	Spiroplasma	Ruminants	A2059G, A2062C	-	-	A89D	<i>in vivo</i>	(Prats-van der Ham <i>et al.</i> 2018)
			A2058G	-	-	A89D	<i>in vitro</i>	

Tableau A3 : Positions des mutations chez les mutants résistants aux fluoroquinolones.

Espèce	Groupe phylo	Hôte	GyrA	GyrB	ParC	ParE	<i>in vivo</i> / <i>in vitro</i>	Référence
<i>M. genitalium</i>	Pneumoniae	Homme	Met83Ile	nd	Gly78Cys, Asp79Asn, Ser80Ile/Arg, Asp84Asn/Tyr/His/Gly, Val100Ile	nd	<i>in vivo</i>	(Tagg <i>et al.</i> 2013)
<i>M. pneumoniae</i>	Pneumoniae	Homme	Asp83Asn	Asp426Asn, Arg447Lys, Glu466Gly	Gly78Cys, Ala80Val, Asp84Asn	Pro439Ser	<i>in vitro</i>	(Gruson <i>et al.</i> 2005)
<i>U. urealyticum</i>	Pneumoniae	Homme	Gln83Arg, Asp95Glu	-	Thr122Ala, Thr133Ala	-	<i>in vivo</i>	(Bebear <i>et al.</i> 2000)
<i>U. parvum</i>	Pneumoniae	Homme	-	-	Ser80Leu	-	<i>in vivo</i>	(Meygret <i>et al.</i> 2018)
			Gln83Arg, Asp95Glu	-	Ser80Leu, Glu84Lys, Ala123Thr, Ala134Thr	-	<i>in vivo</i>	(Bébéar <i>et al.</i> 2003)
			Gln87Lys	-	Ser80Leu	Val410Thr	<i>in vivo</i>	(Meygret <i>et al.</i> 2018)
<i>M. gallisepticum</i>	Pneumoniae	Aviaire	Gly81Ala, Ser83Arg/Ile/Asn, Ala84Pro, Glu87Gly/Lys	Asp426Asn, Asn464Asp, Glu465Gly/Lys	Ala64Ser, Ser80Leu/Trp, Ser81Pro, Glu84Gly/Lys	Asp420Asn, Ser463Leu, Cys467Phe	<i>in vitro</i>	(Reinhardt <i>et al.</i> 2002)
			Ser83Ile/Asn, Glu87Lys	-	Ser80Leu	-	<i>in vivo</i>	(Lysnyansky <i>et al.</i> 2012)
<i>M. hominis</i>	Hominis	Homme	Asp82Asn, Ser83Leu, Glu87Lys, Ala93Thr	Val450Phe, Ala453Phe	Arg73His, Ser80Ile	Asp420Asn, Arg441Lys, Glu460Lys	<i>in vitro</i>	(Gruson <i>et al.</i> 2005)
			Ser83Leu, Glu87Lys	-	Ser81Pro, Glu84Gly	Asp420Asn	<i>in vivo</i>	(Bebear <i>et al.</i> 2000)
			Ser83Leu, Ser84Trp	-	Ser81Pro, Glu84Lys	Asp420Asn	<i>in vitro</i>	
			Ser83Leu, Glu87Lys	nd	Ser80Ile, Glu84Lys	nd	<i>in vitro</i>	(Kenny <i>et al.</i> 1999)
			Ser83Leu, Ser84Trp	-	Asp69Tyr, Arg73His, Ser80Ile, Glu84Lys	Asp426Asn	<i>in vitro</i>	(Bébéar <i>et al.</i> 1998)
			Ser83Trp	-	Ser80Ile, Lys134Arg	Ala462Val	<i>in vivo</i>	(Bébéar <i>et al.</i> 2003)
			Ser83Leu	-	Ser80Ile, Glu84Gly, Lys134Arg	Asp420Asn	<i>in vivo</i>	(Meygret <i>et al.</i> 2018)
<i>M. bovis</i>	Hominis	Ruminants	Ser83Leu/Phe	-	Ser80Ile, Ser81Pro	-	<i>in vivo</i>	(Sato <i>et al.</i> 2013)
			Gly81Cys, Ser83Leu	-	Gly78Leu, Ser80Ile, Ser81Tyr, Asp84Tyr	-	<i>in vitro</i>	
			Ser83Phe	-	Asp84Asn	-	<i>in vivo</i>	(Lysnyansky <i>et al.</i> 2009b)

Tableau A3 : Suite (2/2).

Espèce	Groupe phylo	Hôte	GyrA	GyrB	ParC	ParE	<i>in vivo</i> / <i>in vitro</i>	Référence
<i>M. bovis</i>	Hominis	Ruminants	Gly81Asp/Asn, Ser83Phe/Tyr, Glu87Lys/Gly/Val	-	Ser80Ile, Asp84Asn/Tyr/Lys, Thr98Arg	-	<i>in vitro</i>	(Khalil <i>et al.</i> 2016)
			Asp82Asn, Ser83Phe/Tyr, Glu87Lys/Gly	His37Asn, Ile423Asn	Gly78Cys, Ser80Ile, Asp84Asn/Tyr/His	-	<i>in vitro</i>	(Sulyok <i>et al.</i> 2016)
			Ser83Phe	Val320Ala	Ser80Ile	-	<i>in vivo</i>	
<i>M. agalactiae</i>	Hominis	Ruminants	nd *	Asn424Lys	Gly78Cys, Asp79Asn, Thr80Ile, Asp84Asn/Tyr	Gly429Ser, Glu459Lys	<i>in vitro</i>	(Tatay-Dualde <i>et al.</i> 2017a)
				Glu387Lys	Thr80Ile, Asp84Asn/Tyr	Asn185Asp	<i>in vivo</i>	
<i>M. hyopneumoniae</i>	Hominis	Porcin	Ala83Val	-	Ser80Tyr	-	<i>in vivo</i>	(Vicca <i>et al.</i> 2007)
<i>M. synoviae</i>	Hominis	Aviaire	Asn87Lys	Ser401Tyr, Ser402Asn	Asp79Asn, Thr80Ile, Ser81Pro, Asp84His/Tyr	Asp420Asn	<i>in vivo</i>	(Lysnyansky <i>et al.</i> 2013)
<i>M. capricolum</i> subsp. <i>capricolum</i>	Spiroplasma	Ruminants	Ser83Ile, Cys84Thr, His154Arg	Asp426Gly, Arg447Lys, Asn465Lys	Asp79His, Ser80Ile, Asp84Asn, Glu150Lys	-	<i>in vitro</i>	(Tatay-Dualde <i>et al.</i> 2017b)
			Glu87Lys	Pro343Leu	Ser80Thr, Asp84Gly		<i>in vivo</i>	
<i>M. mycoides</i> subsp. <i>capri</i>	Spiroplasma	Ruminants	Arg47Cys, Ser83Ile, Ala84Val, Glu87Lys	Pro445Ser	Asp84Asn/Gly/Tyr	Asn458Lys/Ile, Glu459Lys	<i>in vitro</i>	(Antunes <i>et al.</i> 2015)

* mutation non détectées, l'oligonucléotide utilisé pour le séquençage est situé au niveau du QRDR.

Liste des communications

Communications orales

M. Faucher, L-X. Nouvel, E. Dordet-Frisoni, M-S. Marena, F. Tardy, C. Citti
In mycoplasmas: HGT may accelerate the emergence of antimicrobial resistance.
 22nd Congress of the International Organization for Mycoplasmology (IOM) 9-12 Juillet 2018,
 Portsmouth, New Hampshire, USA.

E. Dordet-Frisoni, M. Faucher, E. Sagne, L-X. Nouvel, C. Citti
Horizontal distributive transfers and the extraordinary plasticity of the mycoplasma genome.
 22nd Congress of the International Organization for Mycoplasmology (IOM) 9-12 Juillet 2018,
 Portsmouth, New Hampshire, USA.

Posters

M. Faucher, L-X. Nouvel, E. Dordet-Frisoni, M-S. Marena, F. Tardy, C. Citti.
Chromosomal DNA transfer: an accelerator for the emergence of antibiotic resistance in mycoplasma
 Journées d'animation scientifique du département Santé Animale de l'INRA, 9-12 Mai 2016,
 Poitiers, France.

M. Faucher, L-X. Nouvel, E. Dordet-Frisoni, M-S. Marena, F. Tardy, C. Citti.
Emergence de l'antibio-résistance chez les mycoplasmes : accumulation séquentielle de mutations et accélération par échange chromosomique.
 Journée des doctorants et des post-doctorants de l'ANSES, 6 Octobre 2016, Paris, France.

M. Faucher, L-X. Nouvel, E. Dordet-Frisoni, M-S. Marena, F. Tardy, C. Citti.
Horizontal chromosomal transfer: an accelerator for the emergence of antimicrobial resistance in mycoplasmas?
 7th Congress of European microbiologists (FEMS), 9-13 Juillet 2017, Valence, Espagne.

M. Faucher, L-X. Nouvel, E. Dordet-Frisoni, M-S. Marena, F. Tardy, C. Citti
In mycoplasmas: horizontal chromosomal transfer accelerates the emergence of antimicrobial resistance.
 22nd Congress of the International Organization for Mycoplasmology (IOM) 9-12, Juillet 2018,
 Portsmouth, New Hampshire USA.

AUTHOR : Marion Faucher

TITLE : Le transfert horizontal de gènes chez les mycoplasmes : de l'acquisition de l'antibiorésistance à la dynamique des génomes

SUMMARY

Mycoplasmas are wall-less bacteria often portrayed as minimal cells because of their reduced genomes. Several species are pathogenic and have a significant economic impact on livestock production, especially for ruminants. Mycoplasmas are also concerned with the worldwide increase in antibiotic resistance. In contrast to the majority of bacteria, these simple bacteria are deprived of conjugative plasmids that are frequently implicated in the horizontal dissemination of resistance genes: in mycoplasmas antibiotic resistance mainly relies on chromosomal mutations in target genes. In Mycoplasmas, the horizontal gene transfer (HGT) has long been underestimated. Recently, two conjugative mechanisms of HGT were described in *Mycoplasma agalactiae*: the transfer of an integrative and conjugative element (ICE), and the unconventional transfer of chromosomal DNA further designed by "MCT" for Mycoplasma Chromosomal Transfer. Our current study focused on exploring MCT mechanisms and on estimating its impact on antibiotic resistance dissemination. Comparative genomic analyses were performed from the sequencing (i) of spontaneous resistant mutants and (ii) of transconjugants selected by mating experiments and selected based on their resistance. Data revealed that MCT generated the simultaneous transfer of multiple, unrelated donor-fragments following a distributive process. In one conjugative step involving two strains, MCT generated a variety of highly mosaic genomes. This phenomenon was also shown to accelerate the dissemination of antibiotic resistance, by allowing in one step the acquisition of multiple and dispersed mutations associated with resistance. Due to the limitless ability of this phenomenon in reshuffling genomes, MCT may offer a valuable contribution in other adaptive processes such as virulence or host specificity. Finally, the distributive nature and the extent of MCT explain the origin of genes transfers detected *in silico* in several mycoplasma species. MCT is certainly a major player in the evolution of these minimal bacteria and a key factor of their persistence and virulence.

KEY-WORDS : *Mycoplasma agalactiae*, horizontal gene transfer, antibiotic resistance, mutations, chromosomal transfer, adaptation.

AUTEUR : Marion Faucher

TITRE : Le transfert horizontal de gènes chez les mycoplasmes : de l'acquisition de l'antibiorésistance à la dynamique des génomes

DIRECTEURS DE THESE : Christine CITTI, Florence TARDY, Xavier NOUVEL

LIEU ET DATE DE SOUTENANCE : Ecole National Vétérinaire de Toulouse, le 8 Novembre 2018

RESUME

Les mycoplasmes sont des bactéries atypiques, dépourvues de paroi et souvent considérées comme des cellules minimales du fait de la taille réduite de leur génome. De nombreuses espèces sont pathogènes et ont un impact économique important dans les filières d'élevage, notamment pour les ruminants. Les mycoplasmes n'échappent pas au phénomène mondial de la résistance aux antibiotiques. Contrairement à la plupart des autres bactéries, les mycoplasmes ne contiennent pas de plasmides conjugatifs souvent incriminés dans la dissémination horizontale de gènes de résistance, la base moléculaire principalement décrite étant la mutation chromosomique des gènes cibles. De façon générale, le transfert horizontal de gènes (HGT) chez les mycoplasmes a longtemps été sous-estimé. Récemment, deux mécanismes de HGT ont été décrits chez *Mycoplasma agalactiae* : le transfert d'élément conjugatif et intégratif (ICE), et le transfert non-conventionnel de régions chromosomiques par conjugaison appelé MCT (*Mycoplasma Chromosomal Transfer*). Nos travaux se sont attachés à explorer ce dernier mécanisme et à évaluer son impact sur l'acquisition de la résistance aux antibiotiques. Une analyse de génomique comparative a été conduite à partir du séquençage de nombreux mutants spontanément résistants et de transconjugants générés par des expériences de *mating* et sélectionnés pour leur résistance. Nos résultats montrent que le MCT conduit de façon distributive au transfert simultané de nombreux fragments. En une seule étape de conjugaison impliquant deux souches, ce phénomène génère une population variée de génomes hautement mosaïques. Il accélère la dissémination de l'antibiorésistance, permettant l'acquisition de plusieurs mutations distantes associées à la résistance en un seul événement. De par les multiples possibilités de réassemblage génomique qu'il produit, le MCT pourrait avoir des conséquences importantes sur d'autres processus adaptatifs comme la virulence ou la spécificité d'hôte. Enfin, les modalités distributives et l'ampleur du MCT expliquent l'origine des transferts de gènes précédemment détectés *in silico* entre de nombreux mycoplasmes. Ce phénomène pourrait donc avoir eu des répercussions importantes sur l'évolution de ces bactéries minimales et être un facteur clé de leur persistance et virulence actuelles.

MOTS-CLES : *Mycoplasma agalactiae*, transfert horizontal de gènes, antibiorésistance, mutations, transfert chromosomique, adaptation.

DISCIPLINE ADMINISTRATIVE : Pathologie, Toxicologie, Génétique et Nutrition.

INTITULE ET ADRESSE DU LABORATOIRE :

Interaction Hôtes-Agents Pathogènes (UMR 1225), INRA, ENVT
Equipe Pathogénèse des infections à mycoplasmes
23 Chemin des Capelles BP 87614
31076 Toulouse Cedex 3