Determinants of growth rate in genome-reduced bacteria

Carolina Gallo López

DOCTORAL THESIS UPF / YEAR 2017

THESIS DIRECTORS
Dr. Luis Serrano Pubul & Dra. Maria Lluch Senar
Department EMBL/CRG Systems Biology Research Unit
Centre for Genomic Regulation (CRG)

A mis padres, mi hermana y mi familia

Acknowledgments

I would like to thank my supervisors Luis Serrano and Maria Lluch for giving me the opportunity to join the lab, for their guidance throughout this journey and for letting me work on both experimental and computational biology. Luis, I thank you for looking carefully at my experimental data and finding correlations. Maria, I thank you for teaching me to work with *Mycoplasma* and for the support when things did not go well.

Also, I want to acknowledge the members of my thesis committee: Ben Lehner, Juan Valcarcel and Jaume Pinyol. I am very grateful for their constructive feedback, critical comments and advices.

I thank all the member in the lab, speciatlly those who taught me programming in R and other computational analysis: Javier Delgado, Marie Trussart, Veronica Llorens, Samuel Miravet and Marc Weber. Also from the wet lab, thanks to Eva Yus with whom I did the screening of growth curves, together with Sira Martinez and Tony Ferrar. Tony, thank you also for the experiments with *Mollicutes* species. Thanks to Adria Sogues who taught me the methodology for ribosome profiling. Thanks to Carlos Piñero for his help in some experiments. Thank to Albert Fabregas and Claudia Capdevila for their help in molecular cloning. Moreover, I want to particularly thank Marie Trussart, Claire Lastrucci, Hanna Benistry and Besray Unal for their emotional support in the difficult times during my PhD, which were not few. And thanks to: Dan Shaw, best deskmate ever, Eva García, Raúl Burgos, Jae Song, Christina Kiel, Martin Shaefer, Violeta Beltrán and Reyes Perza. Thanks to all for all the moments outside the lab, for the beach volleyball games, for the dinners and the parties.

I am very grateful with all the people from the CRG core facilities. In the genomics facility: Jochen Jecht for his invaluable ideas and contributions and his team; in the proteomics facility: Eduard Sabido and Guadalupe Espadas for all

the explanations with the proteomics data; in the microscopy facility: Timmo Zimmermann and Raúl Gómez for his help with the optimization of the microcolony growth assay; in the proteins facility: Carlo Carolis and Katie Broadbent; in the bioinformatics facility: Luca Cozzuto, Antonio Hermoso and Juan Carlos Company and in the flow cytometry facility: Oscar Fornas and his team. From the UAB microscopy service, thanks to Alejandro Sánchez.

Big thanks to the Ministry of Economy, Industry and Competitiveness (MINECO), who granted me the predoctoral fellowship Severo Ochoa 2013 - 2017, reference SVP-2013-067670.

Last but not least, I want to thank my mother and my father, family and friends who from the long distance gave me all their support to study overseas and encouragement to continue and finishing my PhD.

Abstract

Understanding how the growth rate is regulated in bacteria is an ongoing challenge in biology and its controlled regulation would have a great impact in the biotechnological industry. Growth rates may be regulated by several genetic factors, but despite some of them are known, we are still unable to rationally increase bacterial growth rates. Most studies are done in fast-growing and highly complex bacteria with large and redundant genomes. *Mycoplasma pneumoniae*, from the *Mollicutes* class, is a simpler organism with one of the smallest genomes. Furthermore, *Mollicutes* species have wide range of growth rates and reduced genomes making them appealing for growth studies. In this thesis, we investigated the genetic determinants of growth rates in *M. pneumoniae* and in other *Mollicutes* species by different approaches. Our results corroborated some genetic factors reported to be associated to fast growth and found additional translational and metabolic determinants that have not been described before.

Keywords: growth rate, ribosomes, tRNA, transcription, metabolism, genome-reduced bacteria.

Resumen

Entender cómo se regula la tasa de crecimiento en bacterias es uno de los retos en curso en biología y su regulación controlada tendría un gran impacto en la industria biotecnológica. Las tasas de crecimiento pueden ser reguladas por varios factores genéticos, pero a pesar de que algunos de ellos son en parte conocidos, aún somos incapaces de incrementar las tasas de crecimiento racionalmente. La mayoría de estudios se han llevado a cabo en bacterias de crecimiento rápido y complejas con genomas grandes y redundantes. *Mycoplasma pneumoniae*, de la clase *Mollicutes*, es un organismo más simple con uno de los genomas más pequeño y con poca redundancia. Adicionalemente, las especies de *Mollicutes* tienen un amplio rango de tasas de crecimiento y genomas reducidos, lo cual las hace atractivas para estudios de crecimiento. En esta tesis, investigamos los determinantes genéticos de las tasas de crecimiento en *M. pneumoniae* y en otras especies de *Mollicutes* por medio de enfoques diferentes. Nuestros resultados corroboraron algunos de los ya reportados factores genéticos asociados a un crecimiento rápido y encontramos además determinantes traduccionales y metabólicos que no habían sido descritos anteriormente.

Palabras claves: tasa de crecimiento, ribosomas, tRNA, transcripción, metabolismo, bacterias de genoma reducido.

Preface

Many of the environmental factors that regulate growth rates were characterized in the 1950s, when the study of bacterial growth physiology began. At that time, the studies in *Salmonella typhimurium* concluded that in rich medium, cells exhibited higher growth rates and were larger than cells grown in poor medium. Later, several cellular components were also described in *Escherichia coli* as functions of growth rates and mathematical models were formulated with quantitative data available at the 1960's and later on. Despite of the important findings of the environmental and cellular factors that determine growth rates of those model bacteria, the causal relationship between genetic factors and growth rates were not fully understood.

With the emergence of systems biology, microbial growth physiology escape the biochemical reductionism to seek for deeper quantitative understanding on a whole-cell level. Nonetheless, the emergence of this field was accompanied with the development of new technologies capable of integrate measurements in a rapid, precise and detailed manner.

Altogether, these advances allowed physiologist to further investigate bacterial growth, although the awareness on the complexity of understanding these commonly used bacteria became evident. Consequently, in the search for lower complexity, scientists have focused in genome-reduced bacteria. In this thesis, we have studied the determinants of growth rates in the genome-reduced bacterium *Mycoplasma pneumoniae* and other *Mollicutes* species. Their genomes, as we said, have the advantage of low complexity, but they are difficult to handle. Most of them require biosafety level 2, lack of genetic tools and the methods to determine their growth rates are tedious, time-consuming and not simple.

In this thesis, first, we have developed new methodologies to measure growth rates in *M. pneumoniae* with high sensitivity and reproducibility. Second, we have ex-

panded our understanding on the factors regulating bacterial growth rates. Third, we have increased successfully the growth rate of *M. pneumoniae* by modifying its genome.

The determinants of faster growth found in *M. pneumoniae* will be implemented in the chassis that is being designed and developed in the laboratory to use as an attenuated live vaccine, which will be used in animals. Additionally, the knowledge acquired in this thesis will be incorporated in the first whole-cell mathematical model for *M. pneumoniae* that is currently been developed in our laboratory. The development of these kind of models is of great importance because of their predictive power of emergent cellular behaviors, which will contribute for the understanding of bacterial growth.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1. Bacterial growth physiology

The study of bacterial physiology started with several quantitative findings in the 1950-60s, for a historical review, see Cooper (1993); Nomura (1999); Schaechter (2015); Zaritsky and Woldringh (2015). These main findings determined how the cell size and chemical composition change with transitions between growth rates, from 20 to 150 min doubling time, afforded by a variety of media in balanced, steady state exponential growth of *Salmonella typhimurium* (Schaechter et al., 1958; Kjeldgaard and Schaechter, 1958). They described that there exists a large number of possible stable physiological states that cells could have and they were dependent on the medium conditions. These physiological states were, furthermore, characterized by a particular cell size and RNA, DNA and number of nuclei/cell composition, each of which could be described as exponential functions of the growth rates. They found that faster growing cells were bigger and slower growing cells were smaller. Moreover, these physiological states were independent of the growth temperature but rather dependent of the biosynthetic activities imposed by the media (Schaechter et al., 1958).

Following studies determined that the caused of cell size differences of the same species were the result of the different number of ribosomes imposed by the

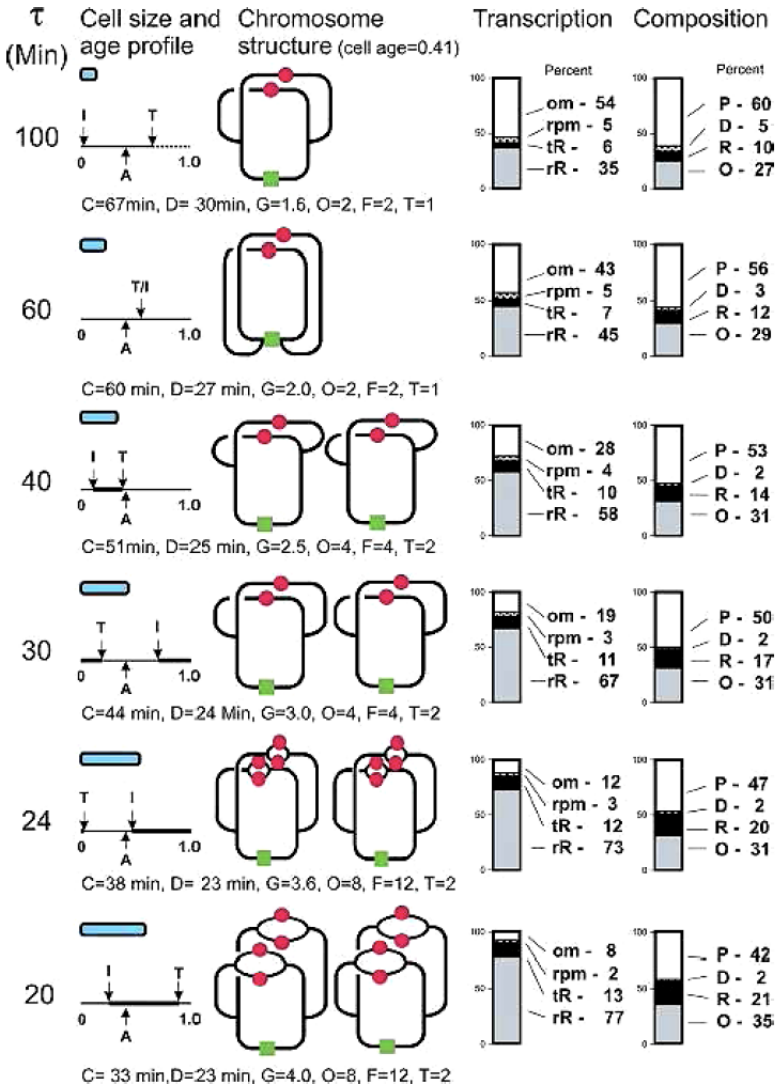


Figure 1.1: Relationships between growth rate, cell size, DNA replication, transcription, and macromolecular composition of *E. coli* B/r. (Left) Average cell size growing with a doubling time, τ, ranging from 100 to 20 min (growth rate, μ, from 0.6 to 3.0 doublings/h) is depicted by the shaded ovals. The cell age 0 (a newborn daughter cell) to 1 (a dividing mother cell) is presented for each μ. An average cell age (A), initiation (I) and termination (T) of DNA replication are indicated. The dashed portion of the age axis indicates that there is no DNA replication. The light line portions indicate periods with two forks per chromosome, and the heavy line portions periods with six forks. After termination, there are two chromosomes per cell, which are segregated to the daughter cells at age 1.0. **(Center)** Replicating chromosomes, for τ=24 or 20 min, the chromosome pattern indicates that replication has reinitiated and that each of these chromosomes has multiple (six) replication forks. The amount of DNA in genome equivalents (G), the numbers of origins (O), termini (T), and forks (F) are indicated. **(Right)** The synthesis rates of rRNA (rR), tRNA (tR), r- protein mRNA (rpm), and other mRNA (om), expressed in the bar graph as a percent of total transcription, and the macromolecular composition. Relative amounts of protein (P), DNA (D), RNA (R), and other components (O) as percent of the total cell mass are indicated. Figure taken from Dennis and Bremer (2008).

growth rates (Ecker and Schaechter, 1963). The concentration of ribosomes was found to be a linear function of the growth rate. Later, the concentration of additional cellular components as functions of growth rate became known in other bacterial species, and thus, the intrinsic factors determining different growth rates became known (Bremer and Dennis, 1996), for a comprehensive review see Dennis and Bremer (2008). Figure 1.1 shows the relationships between growth rate, cell size and the macromolecular composition found in *Escherichia coli* B/r, notably the RNA fraction of the total transcription increases with increasing growth rate and the ribosomal (r-) protein mRNA increases as a fraction of the total mRNA synthesis rate (Gausing, 1977). Similarly, the number of tRNA molecules increases with increasing growth rate, although, the number of tRNA molecules per ribosome was found to decrease at high growth rates (Dong et al., 1996). The higher tRNA values during slow growth was attributed to the instability of newly made rRNA at slow growth rates, which increases the ratio of tRNA to rRNA and a slight decrease in the ratio of tRNA to rRNA genes at fast growth rates due to increased gene dosages associated to initiations of chromosome replication (Voulgaris et al., 1999), since the rrn operons are located closer to the origin of replication than the tRNA genes. Later studies found similar correlations among 214 species of bacteria with different growth rates, an increase in copy number of rRNA and tRNA genes with increasing growth rates, as well as the proximity of these genes to the origin of replication were found, suggesting the result of growth optimization (Rocha, 2004; Couturier and Rocha, 2006; Vieira-Silva and Rocha, 2010).

Additional studies on DNA replication and cell division in *E. coli* B/r (Helmstetter and Cooper, 1968; Cooper and Helmstetter, 1968) found that the time of DNA replication is a constant and is independent of the growth rate and the richness of the medium. In contrast, the initiation of rounds of DNA replication depend on growth rate. Therefore, in slower growing *E. coli* B/r cells, there is a delay or a ”gap” between the end of a round of DNA synthesis and cell division thereby decreasing the growth rate while in faster growing cells, to which the DNA replication time is longer than their doubling time, several rounds of DNA replication are initiated. Thus, this hypothesis could explain how the DNA content could

increase exponentially with increasing growth rates and regulation on DNA initiation became an important factor for determining growth rates. These observations helped to formulate mathematical relationships for the model of the regulation of the chromosome cycle (Cooper and Helmstetter, 1968).

1.2. Bacterial cell division

Bacterial cells have evolved different mechanisms to ensure that their genetic material is passed on from a cell to its daughter cells (Thanbichler, 2010). Most bacteria divide by binary fission (Rivas-Marín et al., 2016). Typically, bacterial cell division is the progression of a series of events, such as DNA replication, chromosome segregation and cytokinesis. These processes are tightly coordinated and controlled and their disregulation are deleterious and often lethal.

DNA replication is usually initiated in a single chromosome region called origin of replication, called *oriC*. This initiation is performed by *trans*-acting initiator proteins (Mott and Berger, 2007), which not only bind the origin of replication, but also assemble other proteins to form the replisome or replication fork. Once the origin is duplicated, chromosome replication progresses bidirectionally until the replisome meets at the terminus locus, which is positioned directly opposite to the origin. As DNA duplication proceeds, the new synthesized chromosomes move apart to the nucleoid region of each daughter cell. After DNA segregation and replication termination, chromosomal copies are resolved by a specialized machinery and the cell divides in between the two newly formed nucleoids giving rise to two identical cells (Thanbichler, 2010).

1.2.1. DNA Replication

DNA replication consist of three consecutive phases: initiation, elongation and termination. It is presumably the most regulated process in cell division and in particular the initiation phase. Interestingly, this regulation varies across organisms. This process has been extensively studied and is the best understood in the Gram-negative model organism *Escherichia coli*. Jameson and Wilkinson

(2017) have very recently reviewed the similarities and differences of the regulatory mechanisms of DNA replication initiation in *E. coli* and the Gram-positive model organism *Bacillus subtilis*.

In the following section, the common principles and the differences discovered in *E. coli* and *B. subtilis* will be mentioned briefly. Moreover, although much less in known about the DNA replication in *Mollicutes* species (see section 1.3.1 for information about this group of bacteria), our current understanding on these species will be also included.

1.2.1.1. Initiation phase

As mentioned above, chomosomal replication is initiated at the origin of replication, *oriC*. Comparative studies have demonstrated that the origins differ in sequences, organizations and sizes (Wolanski et al., 2015; Jameson and Wilkinson, 2017). All bacterial replication origins contain multiple DnaA boxes (usually a 9-bp repeat), an AT-rich DNA unwinding element (DUE) (usually composed of three 13-mer repeat) and binding sites for accessory and regulatory proteins (Mott and Berger, 2007; Wolanski et al., 2015; Jameson and Wilkinson, 2017). Notably, the origins of *E. coli* and *B. subtilis* differ in two features, namely, the genomic context of the origin and the number of regions that forms *oriC*.

Genomic context of origins of replication The location and the genetic context are well preserved across bacteria. Most origins contain or are flanked by the *dnaA* gene (Briggs et al., 2012; Wolanski et al., 2015; Jameson and Wilkinson, 2017) and sometimes also the *dnaN* gene (Wolanski et al., 2015), and are surrounded by the *rnpA-rpmH-dnaA-dnaN-recF-gyrB-gyrA* gene cluster (Figure 1.2) (Briggs et al., 2012; Wolanski et al., 2015; Jameson and Wilkinson, 2017). Surprisingly, the origin, in *E. coli* and in some closely related Gram-negative species, has undergone significant rearrangements, although the gene organization around the *dnaA* gene has been conserved (Briggs et al., 2012) (Figure 1.2), *e.g.*, in *E. coli*, the origin has translocated 44 kb away from the *rnpA-rpmH-dnaA-dnaN-recF-gyrB-gyrA* cluster, thus, it is flanked by *gidA* and *mioC* genes (Jameson and Wilkinson, 2017).

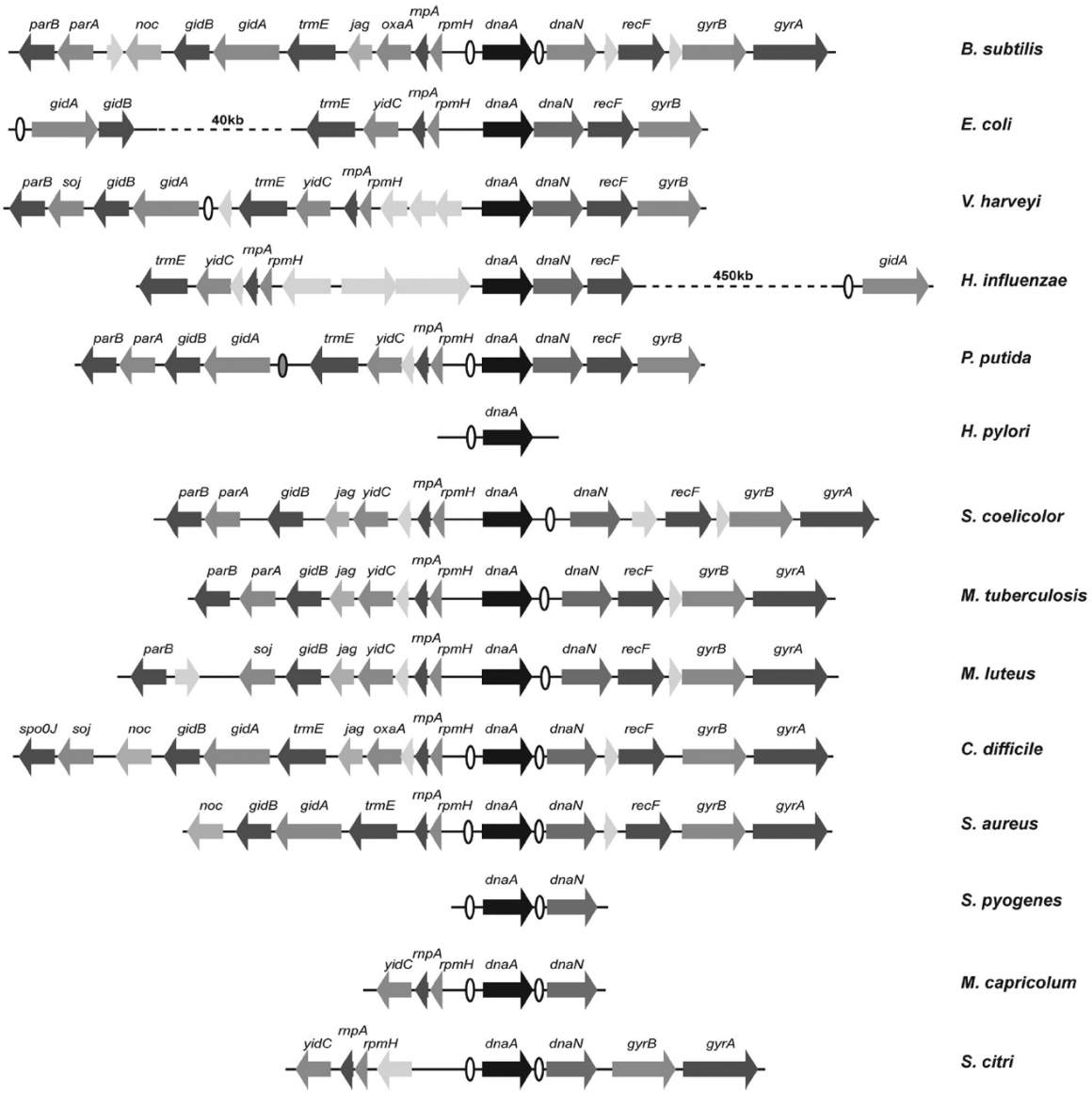


Figure 1.2: Conservation of gene synteny around the *dnaA/oriC* region across selected bacteria. The Gram-negative bacteria: *Escherichia coli*, *Vibrio harveyi*, *Haemophilus influenza*, *Pseudomonas putida* and *Helicobacter pylori*; the Gram-positive bacteria Actinobacteria (high G+C content): *Streptomyces coelicolor*, *Mycobacterium tuberculosis* and *Micrococcus luteus*; the Gram-positive bacteria Firmicutes (low G+C content): *Bacillus subtilis*, *Clostridium difficile*, *Staphylococcus aureus* and *Streptococcus pyogenes*; and the more-distantly related *Mollicutes* species: *Mycoplasma capricolum* and *Spiroplasma citri*. Open ellipses: DnaA-box clusters which form the origins of replication. Filled ellipse: the additional DnaA-box cluster which can function as an autonomously replicating sequence (*ars*) in *Pseudomonas putida*. Figure taken from Briggs et al. (2012).

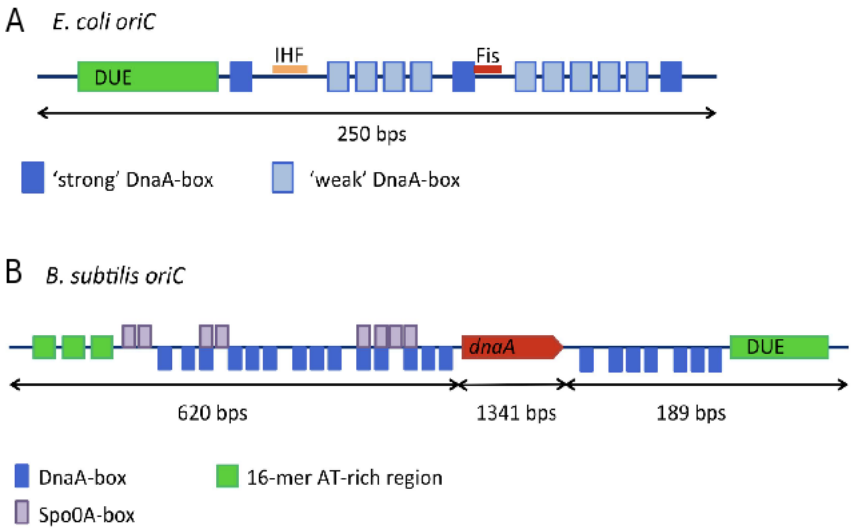


Figure 1.3: Scheme of the structure of origins of replication of *E. coli*, a continuous origin, and *B. subtilis*, a bipartite origin. (A) *E. coli* origin of replication: dark blue: strong DnaA-boxes, light blue: weak DnaA-boxes, green: the DNA unwinding element (DUE), orange and red: binding sites for accessory proteins integration host factor (IHF) and Fis, respectively. **(B)** *B. subtilis* origin of replication: blue: DnaA-boxes, red: the *dnaA* gene, green: the DNA unwinding element (DUE) and purple: Spo0A-boxes. Figure taken from Jameson and Wilkinson (2017).

Continuous and bipartite origins of replication The consensus organization of origins in closely related Bacilli species (*Clostridium difficile*, *Staphylococcus aureus*, *Streptococcus pyogenes* and *B. subtilis*) and more distantly related *Mollicutes* species (*Mycoplasma capricolum* and *Spiroplasma citri*) revealed a split or divided origin (Figures 1.2, 1.3.B), whose role is unknown but is required for functional replication (Briggs et al., 2012). Initially, this split origin was though to be found in few Gram-positive bacteria, however it was recently found in the Gram-negative bacteria *Helicobacter pylori* (Donczew et al., 2012). Consequently, a origin of replication may be continuous or bipartite depending on whether all the functional modules are included in one or two intergenic regions respectively (Wolanski et al., 2015; Jameson and Wilkinson, 2017). The two regions of the split origins contain a cluster of DnaA boxes, and only one of the regions contain the DUE region. These origins also vary in length: the continuous one ranges

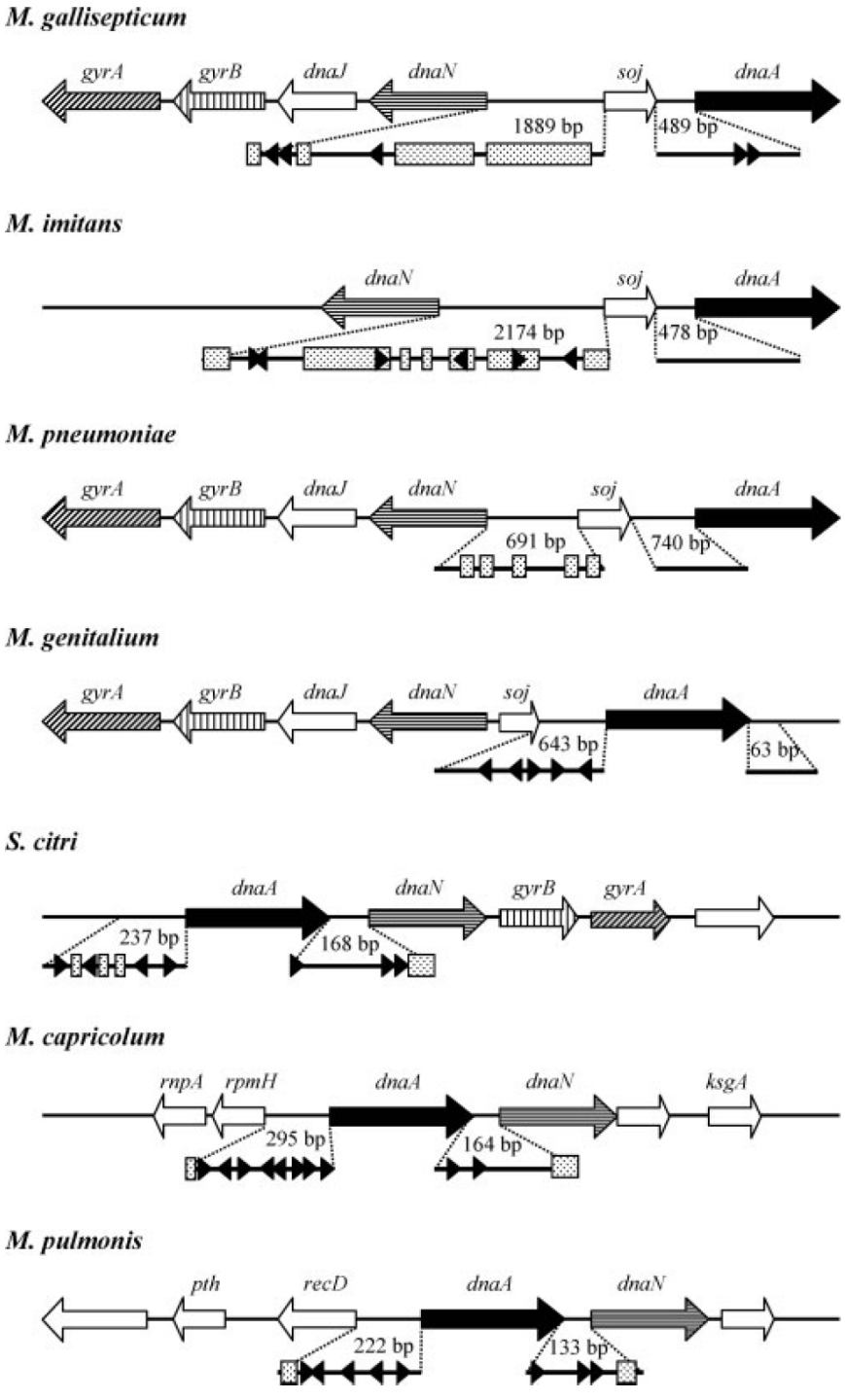


Figure 1.4: Conservation of gene synteny around the *oriC* region and the structure of putative *oriC* regions across *Mollicutes* species. Triangles: the locations of the DnaA box consensus sequences, shaded rectangles: the locations of AT-rich regions. Figure taken from Lee et al. (2008a).

from ∼250 (in *E. coli* (Figure 1.3.A)) to ∼950 bps (in *Streptomyces*), whereas the split one is longer and may reach up to ∼2000 bps, because it contains usually the *dnaA* gene. The replication origins of *Mollicutes* were described as having unusual properties like divergent organizations of their *oriCs e. g.*, differences in the number, orientation and sequence of the DnaA boxes and the localization of the AT-rich regions (Figure 1.4) Lartigue et al. (2003); Lee et al. (2008a). However, these conclusions should be reviewed because they were identified only by the minichromosome approach and no further characterization was performed (Wolanski et al., 2015). The Mycoplasmas belonging to the Pneumoniae phylogenetic group: *M. gallisepticum*, *M. imitans*, *M. pneumoniae* and *M. genitalium* present a conserved gene order (Figure 1.4). Nevertheless, the DNA sequence of the *oriC* regions in *M. gallisepticum* and *M. imitans* is less conserved Lee et al. (2008a). Although, several AT-rich cluster were found in *M. pneumoniae oriC* regions, none a single consensus DnaA box nonamers was found suggesting that it may have a less conserved sequence Lee et al. (2008a). In contrast, the gene order in the other species differs according to the phylogenetic group, *M. pulmonis* belonging to the Hominis phylogenetic group, and *S. citri* and *M. capricolum* belonging to the Spiroplasma phylogenetic group, and in particular *M. capricolum* which belongs to the mycoides cluster (Figure 1.4).

On the *oriC*, the DNA replication is initiated by the binding of the initiator protein, DnaA, to its DnaA boxes. Interestingly, DnaA belongs to the AAA+ ATPase familiy (ATPases Associated with diverse cellular Activities). It oligomerises and forms a nucleoprotein complex which induces DNA unwinding and facilitates binding of the DNA helicase, DnaB in *E. coli* or DnaC in *B. subtilis* (Table 1.1). The DNA helicase is loaded onto the unwound single-stranded DNA by the helicase loader, called DnaC in *E. coli* and DnaI in *B. subtilis* (Table 1.1). Exceptionally, it was found that DnaA of *M. capricolum* is four and ∼80 times more concentrated than that of *E. coli* and *B. subtilis*, respectively (Seto et al., 1997). This difference was explained by the assumption of the lower affinity of DnaA of *M. capricolum* to its non-conserved DnaA boxes or alternatively the smaller ratio of active molecules compared to those of other bacteria (Seto et al., 1997). Otherwise, it may also suggest alternative roles of DnaA in DNA replication in *M.*

capricolum (Seto et al., 1997). It was also shown that most of the DnaA proteins were localized in the membrane fraction, suggesting a membrane role in initiation of DNA replication (Skarstad and Boye, 1994; Seto et al., 1997).

Besides to the DnaA protein and the replication restart protein PriA, in general the low-G+C-content Firmicutes require two additional essential proteins, DnaD and DnaB (Table 1.1), to be bound to the origin and remodel DNA in order to load the helicase (Briggs et al., 2012; Jameson and Wilkinson, 2017). In several *Mollicutes*, no *dnaD* and *dnaB* genes have been found, instead, there is a single gene annotated as *dnaD-like*, which could retain both functions (Briggs et al., 2012). Fuerthermore, outside the Firmicutes phylum, no homologous genes have been found (Briggs et al., 2012).

Table 1.1: Protein factors involved in bacterial replication initiation in the model organisms *Escherichia coli* and *Bacillus subtilis* . Table modified from Briggs et al. (2012)

	Gram-negative	**Gram-positive**
Factors	***E. coli***	***B. subtilis***
Initiator	DnaA	DnaA
Helicase	DnaB[a]	DnaC[a]
Helicase loader	DnaC[a]	DnaI[a], DnaB and DnaD
Primase	DnaG	DnaG
Accessory/remodeling protein		DnaB[a,b], DnaD[b]
Accessory protein	DiaA	
Remodeling factor	IHF, Fis	
Histone-like protein	HU1, HU2	HBsu
Regulatory protein	Hda, SeqA	
Regulatory protein		YabA, Soj, SirA, Spo0A
Primosomal protein	PriA	PriA
Primosomal accessory protein	PriB, PriC, DnaT	

[a] For historical reasons, the *B. subtilis* replicative helicase is called DnaC and the helicase loader is DnaI. *B. subtilis* DnaB and DnaC are unrelated to the *E. coli* DnaB and DnaC.
[b] In several of the *Mollicutes*, there are not separate DnaB and DnaD proteins; there is a single DnaD-like protein which may combine functions from both.

There are several additional regulatory proteins that are not found in *E. coli*, in-

cluding YabA, Soj, SirA, and Spo0A. Likewise, there are other regulatory proteins that are found in *E. coli* but not in *B. subtilis*, including Hda, SeqA (Table 1.1) (Briggs et al., 2012). Katayama et al. (2010) and more recently Jameson and Wilkinson (2017) have reviewed these regulatory proteins with their functions for a comprehensive understanding.

1.2.1.2. Elongation phase

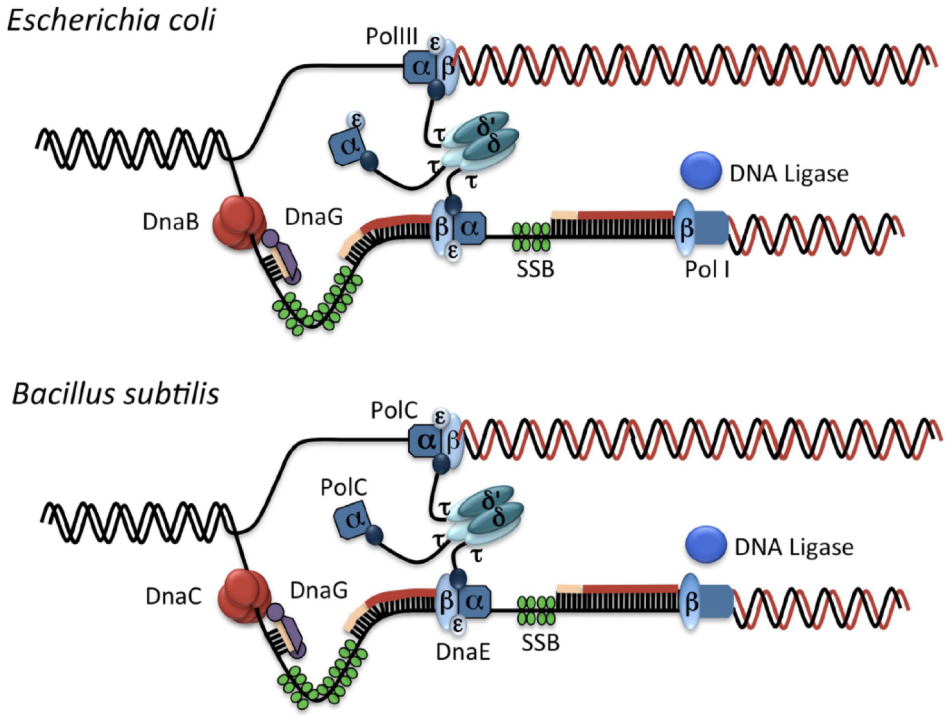


Figure 1.5: Scheme representing replisomes of *E. coli* and *B. subtilis*. The helicase, primase, DNA polymerase, the β-clamp and the clamp loader ($\tau_3\delta\delta$') at the replication fork are shown. SSB: single-stranded DNA binding protein. Figure taken from Jameson and Wilkinson (2017).

To prepare for the DNA synthesis in the elongation phase of DNA replication, the hexameric DNA helicase recruits the primase, DnaG, and the three processivity polymerases β-clamp, DnaN, which in turn recruits a pentameric clamp loader complex and DNA polymerases forming the replisome (Figure 1.5) (Jameson and Wilkinson, 2017). The pentameric clamp loader complex consists of a

$\tau_3\delta\delta$' structure, where the τ subunit, encoded by the *dnaX* gene, interacts with both the DNA helicase and the DNA polymerase coupling DNA unwinding and DNA extension. In *E. coli*, the replisome contains three DNA polymerase III (Pol III) complexes, which consists of a $\alpha\varepsilon\theta$ structure, where α , encoded by the *dnaE* gene, is the catalytic subunit, ε is the proofreading subunit and θ is a non-essential subunit presumably to stimulate the activity of ε (Figure 1.5). In *B. subtilis* instead, the replisome contains two different and essential DNA polymerases, PolC and DnaE (Dervyn et al., 2001). DnaE is closely related to the α subunit of Pol III in *E. coli* (Dervyn et al., 2001). PolC is essential for the synthesis of the leading strand, while for the lagging strand synthesis both polymerases are essential (Sanders et al., 2010).

1.2.1.3. Termination phase

DNA synthesis ends at the terminus region, *ter*, where the two bidirectional replication forks meet. Two different polar mechanisms, have evolved in *E. coli* and *B. subtilis*, where there is a permissive or non-permissive direction from which the *ter* site may be reached (Hill, 1992; Jameson and Wilkinson, 2017). In *E. coli*, the *ter* site is flanked on both sites by five non-palindromic 23-bp sites, *TerA-J* (Figure 1.6.A), to which the monomeric protein Tus (terminator utilisation substance) binds. The orientation of these *ter* sites determines whether a processive replication fork is halted or not. Hence, a replication fork is able to pass a *ter* site when advancing in the permissive direction, but is halted in the non-permissive direction (Berghuis et al., 2015). In *B. subtilis*, DNA is arrested by the binding of two homodimers of the replication termination protein (RTP) at A and B sites within the *ter* region (Figure 1.6.B) (Vivian et al., 2007).

1.2.2. Chromosome segregation and cytokinesis

Cell division genes were originally called *fts* (filamentous temperature sensitive) genes in *E. coli*. *fts* mutants grow without dividing, at the non-permissive temperature (usually 42°C), forming long filamentous cells that performed successfully DNA replication and chromosome segregation but not septa formation or cell envelope constriction (Margolin, 2000; Rowlett and Margolin, 2015). Later,

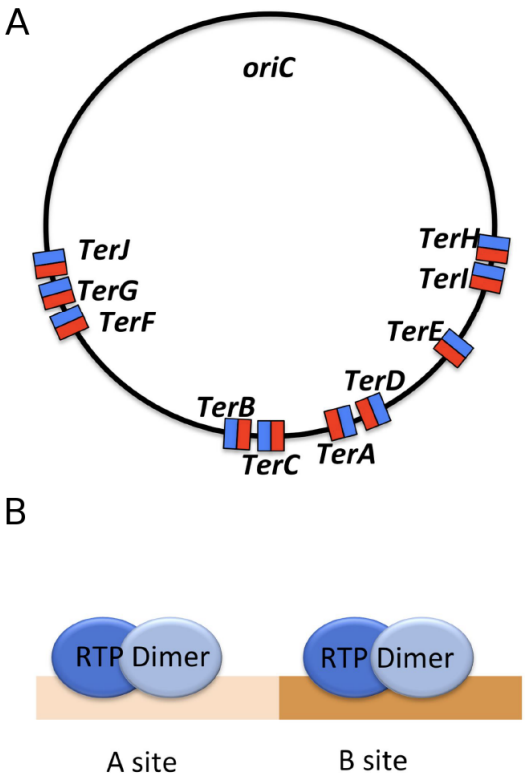


Figure 1.6: Scheme representing *ter* sites in *E. coli* and *B. subtilis*. (A) *ter* sites in *E. coli*: blue represents the permissive face, red represents the non-permissive face. **(B)** *ter* site in *B. subtilis*: RTP dimers binding at the A and B sites of the terminus region. Figure taken from Jameson and Wilkinson (2017).

these genes were found to be responsible of cytokinesis. Cytokinesis in bacteria is achieved by a macromolecular complex called the divisome. The formation of the divisome is initiated by the tubulin homologue, FtsZ (Adams and Errington, 2009). For a recent comprehensive review of bacterial chromosome organization and segregation see (Badrinarayanan et al., 2015).

1.2.2.1. FtsZ conservation

FtsZ, which is a GTPase (de Boer et al., 1992; RayChaudhuri and Park, 1992) polymerizes into a septum or ring-like structure near the division site. It is both an initiator of this Z-ring assembly and a scaffold for the divisome assembly (Adams and Errington, 2009). FtsZ is widely conserved in almost all bacteria, only few

groups lack an homologue of FtsZ, including *Chlamydiae*, *Planctomycetes*, and the *Mollicutes Phytoplasma sp.*, *Ureaplasma urealyticum*, and *Mycoplasma mobile*, it is also conserved in several archaes, chloroplasts of plants and some mitochondria (Margolin, 2000, 2005; Alarcón et al., 2007; Adams and Errington, 2009; Bernander and Ettema, 2010; Rivas-Marín et al., 2016).

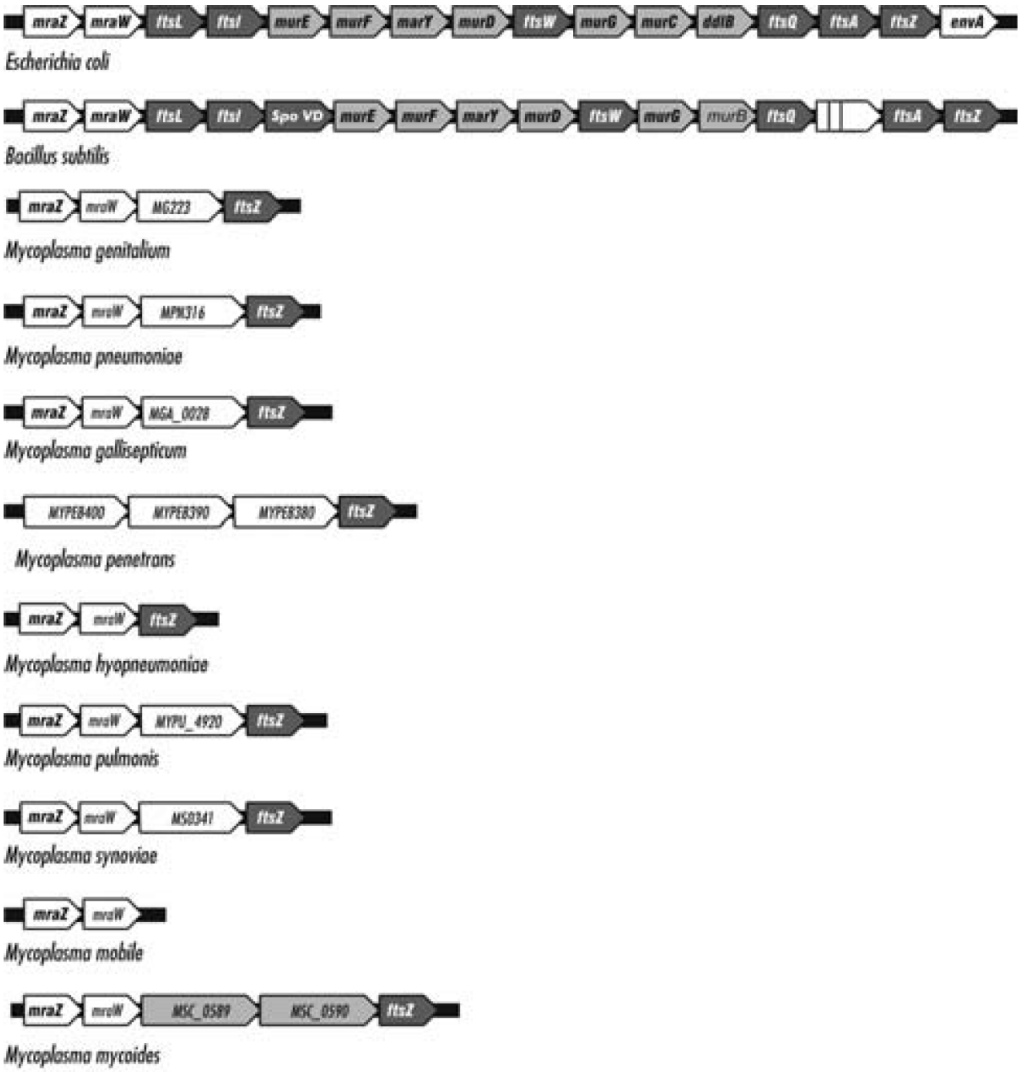


Figure 1.7: Conservation of gene synteny of the *dcw* cluster in *E. coli*, *B. subtilis* and *Mollicutes* species from the genus Mycoplasma. Figure taken from Alarcón et al. (2007).

1.2.2.2. The *mra* or *dcw* cluster

In bacteria, genes implicated in cell division and synthesis of the precursors of the cell wall are grouped in a single locus called either the murein region A (*mra*) or the division and cell wall (*dcw*) cluster (Mengin-Lecreulx et al., 1998) (Figure 1.7). The *dcw* cluster is composed of 16 and 17 genes in *E. coli* and *B. subtilis*, respectively, whereas in *Mollicutes* species, it is of three to four genes (Alarcón et al., 2007). In *E. coli*, the first two genes of the *dcw* cluster are *mraZ* and *mraW*. Downstream the cluster, the genes *murBCDEFG*, *mraY* and *ddlB* are responsible for peptidoglycan or murein precursor biosynthesis whereas the genes *ftsLIWQAZ* are responsible for cell division (Alarcón et al., 2007).

The gene products of *mraZ* (*yabB*) and *mraW* (*yabC*, *rsmH*), which are also highly conserved across bacteria, were initially assumed to be involved in cell wall biosynthesis and/or cell division because of their genomic position (Adams et al., 2005). However, their presence in Mycoplasma species (Figure 1.7), which lack a cell wall, suggested additional functions (Eraso et al., 2014). More recently, Eraso et al. (2014) described Mraz as a transcriptional regulator of the *dcw* cluster, negatively regulating its own expression, and genes outside thereof in *E. coli*. In contrast, MraZ was identified as an activator of only the *dcw* cluster in *M. gallisepticum* and similar to the study in *E. coli* Eraso et al. (2014), cells displayed a filamentous growth and tended to form aggregates as a consequence of incomplete cytokinesis (Fisunov et al., 2016). On the other hand, mraW, was identified as a S-adenosyl-dependent methyltransferase responsible for the N^4-methylation at position 1402 of the *E. coli* 16S rRNA, where the 16S rRNA interacts with the codon of the mRNA. (Carrión et al., 1999; Kimura and Suzuki, 2009). The N^4-methylation of $m^4Cm1402$ is involved in non-AUG initiation efficiency and UGA read-through, therefore $m^4Cm1402$ is important in increasing decoding fidelity by fine-tuning the shape and function of the P-site of the 30S subunit of the ribosome, where the codon-anticodon pairing occurs (Kimura and Suzuki, 2009).

The *ftsL*, *ftsI*, *ftsW* and *ftsQ* gene products are all transmembrane proteins and are involved in peptidoglycan synthesis during division (Weiss, 2004). Together

with the genes *murBCDEFG*, *mraY* and *ddlB* are highly conserved in most bacteria, with the exception of those that lack a cell wall, like *Mollicutes* species (Figure 1.7). In *Mollicutes*, the number of genes of *dcw* cluster is much lower, however the organization of genes in this group of bacteria is conserved (Figure 1.7) (Alarcón et al., 2007). The *dcw* cluster of the species of the clade Pneumoniae, such as *M. genitalium*, *M. pneumoniae* and *M. gallisepticum*, with the exception of *M. penetrans* has four genes: *mraZ*, *mraW*, a conserved hypothetical CDS, and *ftsZ*. In *M. penetrans*, *mraW* gene is located in another position of the genome (424743..425681) and a conserved *mraZ* gene was not found (Alarcón et al., 2007). Similarly, species of the clade Hominis *M. pulmonis* and *M. synoviae* with the exception of *M. hyopneumoniae* and *M. mobile* have four genes in the *dcw* cluster. In contrast, the *dcw* cluster in *M. hyopneumoniae* and *M. mobile* is composed of three CDSs: *mraZ*, *mraW* and *ftsZ* and two CDSs: *mraZ* and *mraW*, respectively (Figure 1.7) (Alarcón et al., 2007).

1.2.2.3. The divisome

In *E. coli*, besides the FtsZ, the septal ring is also constituted by the following proteins: FtsA, ZipA, ZapA, FtsEX, FtsK, FtsQ, FtsL, FtsB (formerly called YgbQ), FtsW, FtsI (which encodes Penicillin Binding Protein 3-PBP3), FtsN, AmiC and EnvC (Errington et al., 2003; Weiss, 2004). FtsA, ZipA and ZapA assembly early to the Z ring and modulate the assembly state of FtsZ, in particular FtsA and ZipA connect the septal ring to the cytoplasmic membrane. FtsK was shown to be a DNA translocase involved in decatenation of sister chromosomes by the topoisomerase IV and resolution of chromosome dimers by the XerCD recombinase. FtsI and FtsW are involve in the synthesis of peptidoglycan cell wall. AmiC and EnvC are murein hydrolases, that are shown to localize in the septal ring facilitating the separation of daughter cells by hydrolysis of peptidoglycan (Errington et al., 2003; Weiss, 2004). The rest of the proteins assemble later to the ring, have membrane and/or periplasmic localization and except for FtsEX (unknown function), are involved in the synthesis of septal peptidoglycan (Alarcón et al., 2007).

In *B. subtilis*, the homologous proteins involved in cell division, besides FtsZ,

are: FtsA, FtsW, DivIB (FtsQ homologue), DivIC (FtsL-like), and PBP-2B (PBP 3 homologue) (Alarcón et al., 2007). Their localization is completely interdependent for assembly at the division septum (Errington et al., 2003). EzrA and ZapA homologues are shown to be negative and positive regulators of FtsZ polymerization (Alarcón et al., 2007).

In *Mollicutes*, cell division has been poorly investigated and there are few studies on genes involved in this process (Alarcón et al., 2007). The genome reduction in *Mollicutes* species has lead to the loss of most cell division genes (Alarcón et al., 2007). Compared to the cell division machinery described in *E. coli* and *B. subtilis*, it seems that in Mycoplasmas, it involves a reduced number of proteins, nonetheless, the cell division process in this last group remains elusive (Alarcón et al., 2007).

In the study of Alarcón et al. (2007), they concluded that species of the clade Pneumoniae (*M. genitalium*, *M. gallisepticum*, *M. pneumoniae* and *U. urealyticum*) seem to have conserved more cell division genes than those of the clade Hominis (*M. hyopneumoniae*, *M. synoviae*, *M. pulmonis* and *M. mobile*). They found that the orthologue of the structural maintenance of chromosomes (SMC) protein, involved in chromosome segregation, is present in the ten genomes of *Mollicutes* studied (Alarcón et al., 2007). In contrast, only two species analyzed (*M. gallisepticum* and *U. urealyticum*) have FtsK, a protein involved in chromosome partitioning (Alarcón et al., 2007). Likewise, the gene encoding ParA, involved in chromosome partitioning, is present in *M. genitalium*, *M. pneumoniae*, *M. penetrans*, and *M. gallisepticum*; the genes encoding recombinases (xerD/xerC family), involved in resolution of the chromosomes, are present in *M. penetrans*, *M. pulmonis* and *U. urealyticum*. EzrA, a protein identified as a negative regulator of FtsZ polymerization, was present only in *M. gallisepticum*. Lastly, the genes encoding FtsH and FtsY are present in all species studied, although they are not directly involved in cell division but they may affect it (Alarcón et al., 2007). The absence of *ftsA*, *ftsQ*, etc. may indicate that FtsZ is sufficient for cell division and that the presence of other proteins in non-*Mollicutes* may play a role in coordination of the synthesis of the cell wall and septum by FtsZ-mediated cytokinesis

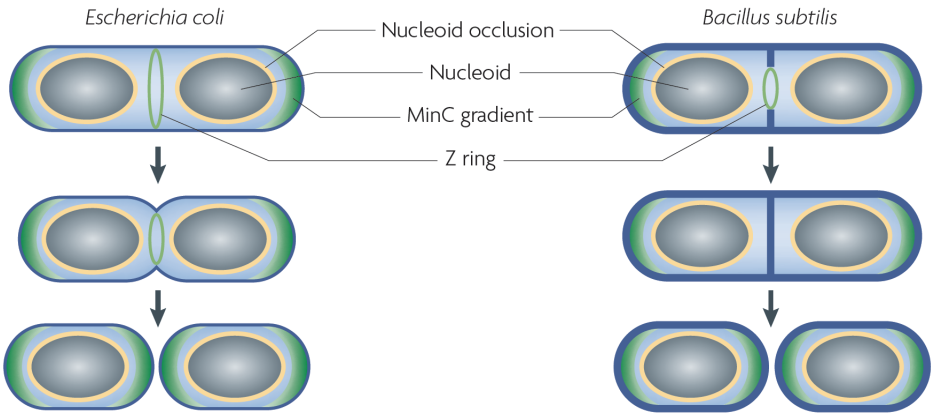


Figure 1.8: Cell division in the rod-shaped bacteria *E. coli* and *B. subtilis*. Two different modes of division. After chromosome replication and segregation into nucleoids the Z ring assembles at midcell. In *E. coli*, synthesis of the division septum is accompanied by constriction of the outer membrane. In *B. subtilis*, a cross wall of peptidoglycan initially divides the cell before it is degraded and remodeled to form the new, hemi-spherical cell poles. Figure taken from Adams and Errington (2009).

(Margolin, 2000). It may, therefore, suggest that FtsZ can provide the constructive force necessary in order to divide cells, nonetheless, *Mollicutes* species have unusual cytoskeletal-like proteins, implicated in cytadherence, that are also thought to be involved in cell division (Margolin, 2000).

1.2.2.4. Z ring regulation

In rod-shaped bacteria, such as *E. coli* and *B. subtilis*, cell division includes septation, and cellular constriction in *E. coli*, usually the Z ring assembly localizes at the midpoint of the cell and perpendicular to the long axis of the cell (Figure 1.8) (Adams and Errington, 2009). Although some bacteria, *e.g. Caulobacter crescentus*, tightly regulate both FtsZ synthesis and stability, in *E. coli* and *B. subtilis*, the concentration of FtsZ does not change significant either throughout the cell cycle or under different growth conditions (Adams and Errington, 2009). However, Z ring assembly is regulated by several proteins in *E. coli* (Figure 1.9) and *B. subtilis*. These regulatory proteins are described in detail by Rothfield et al. (2005); Lutkenhaus (2007) and will be briefly mentioned here. Firstly, nucleoid occlusion inhibits Z ring polymerization close to the nucleoid by the presence of two non

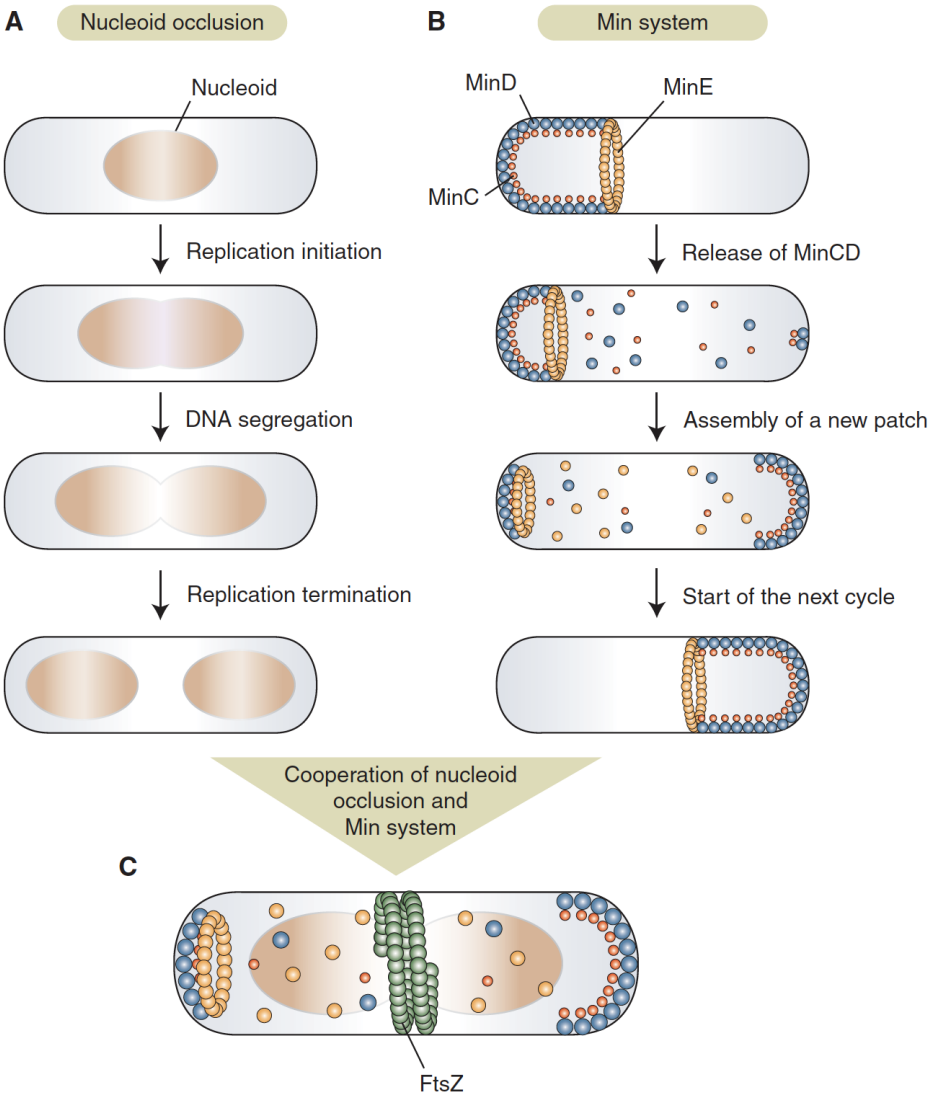


Figure 1.9: Positioning of the FtsZ ring by nucleoid occlusion and Min systems in *E. coli*. **(A)** Temporal and spatial regulation of cell division by nucleoid occlusion. Cell division initiates by positioning SlmA near midcell, thereby blocking FtsZ ring assembly. During chromosome replication and segregation, the two newborn daughter nucleoids move apart. Thus, the midcell region is cleared of SlmA, and FtsZ polymerization begins. **(B)** Inhibition of polar cell division events by the Min system. MinD binds to the cell division inhibitor MinC, on the cytoplasmic membrane, forming a polymeric layer that prevents FtsZ ring assembly in the cell pole. MinE is polimerized into a ring-shaped structure that releases MinCD. Free MinC and MinD reassemble at the opposite cell pole, thus restarting the cycle. **(C)** Cooperation of the nucleoid occlusion and Min systems. The combined action of SlmA and the Min system positions the FtsZ ring at midcell and ensures that divisome is formed only after the final phase of replication. Figure taken from Thanbichler (2010).

homologous DNA binding proteins, Noc (in *B. subtilis*) or SlmA (in *E. coli*) (Figure 1.9.A.) (Adams and Errington, 2009). Secondly, the Min system inhibits Z ring polymerization at the cell poles by creating a concentration gradient of MinC and MinD around the midpoint so that concentration is always highest at the cell poles. In *E. coli* the division site is delimited by the oscillation from pole to pole of MinC and MinD, which is determined by MinE (Figure 1.9.B.). In *B. subtilis*, most probably DivIVA and MinJ recruit MinC and MinD to the division site and retain them at the new cell poles (Adams and Errington, 2009), thus MinC and MinD do not oscillate (Thanbichler, 2010). Nonetheless, Rodrigues and Harry (2012) suggested that the Min system and the nucleoid occlusion are not required for the localization of the division site but rather these factors are only required for the efficient assembly of the Z ring and for ensuring the correct position of the division site in *B. subtilis*. Similarly, later studies in *E. coli* also concluded that there are additional systems that determine the localization of the Z ring at midcell (Bailey et al., 2014; Cambridge et al., 2014).

1.3. *Mollicutes* as model systems for cell division

1.3.1. The class *Mollicutes*

Mollicutes is a class of bacteria that is characterized by the absence of a cell wall (Table 1.2) (Dybvig and Voelker, 1996; Razin et al., 1998). The word mollicutes is derived from the Latin "*mollis*" meaning soft, and "*cutis*" meaning skin (Razin et al., 1998). Together with the classes *Bacilli* and *Clostridia*, *Mollicutes* belong to the phylum Firmicutes, so-called Tenericutes. The Firmicutes are Gram-positive bacteria characterized by having a relatively low G+C content in their genomes. They are also morphologically and physiologically diverse: they may be rod-shaped bacilli and spherical cocci, aerobic and anaerobic, spore-forming and non-spore-forming bacteria (Briggs et al., 2012).

Mollicutes are characterized by their small genome sizes ranging from 580 kbp (basepairs) (*Mycoplasma genitalium*) (Fraser et al., 1995) to 1358 kbp (*Mycoplasma penetrans*) (Sasaki et al., 2002). Thus, *M. genitalium* is the smallest self-replicating

Table 1.2: Properties distinguishing *Mollicutes* from other eubacteria. Table modified from Razin et al. (1998)

.

Property	***Mollicutes***	**Other eubacteria**
Cell wall	Absent	Present
Plasma membrane	Cholesterol present in most species	Cholesterol absent
Genome size	580 - 2220 kb	1050 - >10000 kb
G+C content of genome	23 - 40 mol%	25 - 75 mol%
No. of rRNA operons	1 or 2[a]	1 - 10 (*B. subtilis*)
5S rRNA length	104 113 nucleotides	>114
No. of tRNA genes	30 (*M. capricolum*), 33 (*M. pneumoniae*)	84 (*B. subtilis*), 86 (*E. coli*)
UGA codon usage	Tryptophan codon in *Mycoplasma*, *Ureaplasma*, *Spiroplasma*, *Mesoplasma*	Stop codon
RNA polymerase	Rifampin resistant	Rifampin sensitive

[a] Three rRNA operons in *Mesoplasma lactucae* (Bové, 1993)

microorganism, capable of independent growth and in axenic laboratory conditions (Fraser et al., 1995; Grosjean et al., 2014). Their small genomes are the result of a reductive evolution, from a Gram-positive ancestor (Maniloff, 1996), as a consequence of their adaptation to parasitic life styles Oshima et al. (2013). Consequently, these bacteria are obligate pathogens or commensals of a wide range of animals (mammals, birds, reptiles, fish, arthropods) and plants, with a preference for inhabiting mucous surfaces (Dybvig and Voelker, 1996; Razin et al., 1998; Rocha and Blanchard, 2002). Because of their reduced genomes, these bacteria have retained few biosynthetic capabilities, and hence they need to obtain most nutrients from their hosts, generally eukaryotic cells (Dybvig and Voelker, 1996). For example, *Mollicutes* lack pathways for the production of the cell wall, *de novo* biosynthesis of purines, a functional tricarboxylic acid cycle, a cytochrome-mediated electron transport chain system. (Dybvig and Voelker, 1996) and both *M. genitalium* and *Mycoplasma pneumoniae* lack all the genes involved in amino acid synthesis (Razin and Hayflick, 2010). Another feature of most *Mollicutes* and of the genus *Mycoplasma* is the requirement for sterols, which are necessary to provide some structural support to the osmotically fragile mycoplasmal plasma membrane (Waites and Talkington, 2004). All these factors make *in vitro* culture of these organisms very difficult and fastidious (Bové, 1993; Dybvig and Voelker,

1996; Carvalho et al., 2005). For a historical review and highlights of the class *Mollicutes*, see (Razin, 2010; Razin and Hayflick, 2010).

1.3.1.1. *Mollicutes* taxonomy

Mollicutes species have been commonly called mycoplasmas (from the Greek, "mykes" meaning fungus and "plasma" meaning formed, allusive to a fungus-like growth pattern) (Waites and Talkington, 2004; Razin and Hayflick, 2010), however, throughout the text, I will refer to mycoplasmas exclusively to the species from the genus *Mycoplasma*. The class *Mollicutes* comprises currently four orders (*Mycoplasmatales*, *Entomoplasmatales*, *Acholeplasmatales*, *Anaeroplasmatales*), five families and eight genera (Brown et al., 2007). The order *Mycoplasmatales* contains the single family *Mycoplasmataceae*, which in turn contains the genera *Mycoplasma* and *Ureaplasma*. The order *Entomoplasmatales* contains two families (*Entomoplasmataceae* and *Spiroplasmataceae*) and three genera (*Mesoplasma, Entomoplasma* and *Spiroplasma*). The order *Acholeplasmatales* contains the single family *Acholeplasmataceae*, which in turn contains the single genus *Acholeplasma*. The order *Anaeroplasmatales* contains the single family *Anaeroplasmataceae*, which in turn contains the two genera *Anaeroplasma* and *Asteroleplasma*. In addition, many *Mollicutes* have not yet been cultivated and most of them are phytoplasmas. Consequently, phytoplasmas are currently classified within the provisional genus "*Candidatus* Phytoplasma" (Brown et al., 2007).

1.3.1.2. *Mollicutes* evolution

Mollicutes species have evolved from a Gram-positive ancestor by degenerative or reductive evolution (Razin and Hayflick, 2010). According to the study of Maniloff (1996), the ancestral *Mollicutes* arose from the *Streptococcus* phylogenetic branch about 600 million years ago, presumably from a bacterium with a genome size around 2000 kb. Later, the *Mollicutes* phylogenetic tree split into two major branches, about 450 million years ago, probably from an organism with a genome size of 1700-2000 kb (Maniloff, 1996). Then, both branches evolved to produce *Mollicutes* sublines with genome sizes of 1200-1700 kb (Maniloff, 1996). One branch (the AAA branch) led to the Asteroleplasma, Anaeroplasma, and Ac-

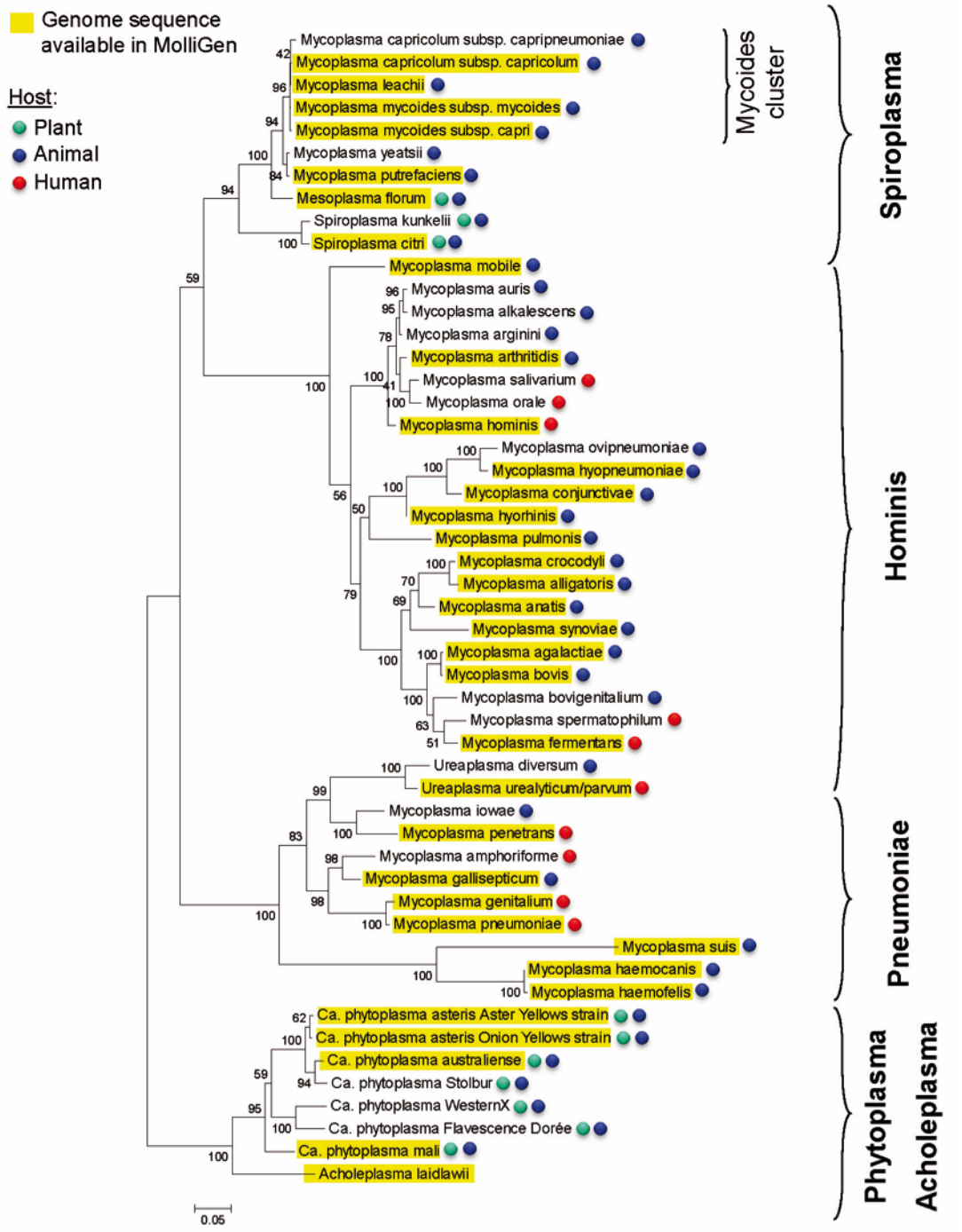


Figure 1.10: Phylogenetic tree inferred from 16S rDNA sequences of several *Mollicutes* from all phylogenetic groups and clusters. The tree was constructed by maximum likelihood method. Figure taken from the MolliGen database (Barré et al., 2004).

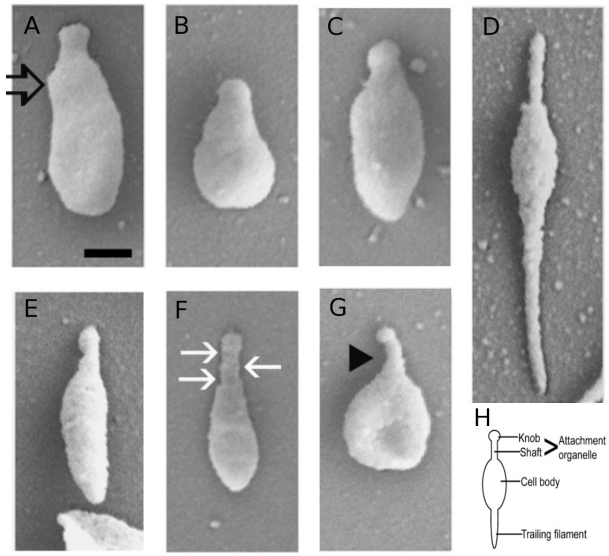


Figure 1.11: Scanning electron micrographs (SEM) of *Mycoplasma* cells grown on glass coverslips. Images are aligned so that the attachment organelle (AO) is at the top. **(A)** *M. gallisepticum*; **(B)** *M. imitans*; **(C)** *M. amphoriforme*; **(D)** *M. pneumoniae*; **(E)** *M. testudinis*; **(F)** *M. pirum*; **(G)** *M. genitalium*; **(H)** scheme of *Mycoplasma* cell. Scale bar, 250 nm. Figure modified from Hatchel and Balish (2008).

holeplasma branches, and the other (the SEM branch) led to the Spiroplasma, Entomoplasma, and Mycoplasma branches (Razin et al., 1998). The phytoplasmas subsequently arose from the Acholeplasma branch, and Ureaplasma arose from the Mycoplasma branch (Razin et al., 1998). The Entomoplasma branch also contains Mesoplasma and the M. mycoides sublines (Razin et al., 1998). Mycoplasma species with small (600-1100 kb) genomes subsequently arose independently on different sublines (Maniloff, 1996). Maniloff (1996) suggested that in both major branches, during their degenerate (reductive) evolution, genome reductions had occurred independently (Figure 1.10). The nature of the selective pressure that cause this degenerate evolution in *Mollicutes* is unknown (Waites and Talkington, 2004).

1.3.1.3. Cell morphology and gliding motility

The cells are built essentially of three organelles: a plasma membrane, ribosomes and a circular double-stranded DNA molecule (Razin and Hayflick, 2010). The

lack of a cell wall explains the sensitivity to osmotic shock and detergents, resistance to penicillin, and formation of the peculiar fried-egg shaped colonies (Razin and Hayflick, 2010). The more predominant cell shape is a sphere (0.3 - 0.8 μm in diameter), however, *Mollicutes* cells display a variety of morphological shapes, including pear-shaped cells, flask-shaped cells with terminal tip structures (Figure 1.11), filaments of varying length (some branching), as well as helical filaments. Adopting this morphology in the absence of a rigid cell wall is due to the presence of a cytoskeleton in *Mollicutes* (Razin and Hayflick, 2010). The cytoskeleton-like structures were shown to serve in shaping cell morphology, in cell division and gliding motility (Balish, 2006; Balish and Krause, 2006; Hasselbring et al., 2006).

Many *Mollicutes*, displaying cell polarity, undergo gliding motility, in which the cells slide on solid surfaces in the direction of the attachment organelle (AO) (Miyata and Seto, 1999). Some of the average gliding speeds are 0.03, 0.11, 0.13, 0.3 - 0.4, 2.0 - 4.5 μm/s for *M. pirum* strain (str.) 70-159, *M. genitalium* str. G37, *M. gallisepticum* str. R_{low}, *M. pneumoniae* str. M129 and *M. mobile*, respectively (Miyata and Seto, 1999; Hatchel and Balish, 2008).

1.3.1.4. Metabolism and energy production

All the *Mollicutes* studied have truncated respiratory systems, they lack a complete tricarboxylic acid cycle and have no quinones and cytochromes (Razin and Hayflick, 2010). The principal pathway for ATP synthesis in *Mollicutes* is presumably substrate-level phosphorylation, which make them inefficient energy producers. Thus, it was shown that *Mollicutes*, excluding *Acholeplasma laidlawii* produce low ATP yields (Beaman and Pollack, 1983).

Based on their metabolism, *Mollicutes* are divided into fermentative and non-fermentative groups, and within each, the arginine or non-arginine-requiring species (Beaman and Pollack, 1983; Razin and Hayflick, 2010). Thus, most *Mollicutes* obtain energy through glycolysis and species that lack glycolytic enzymes are thought to obtain energy via the arginine hydrolase pathway (*e.g.* *Mycoplasma arthritidis* and *Mycoplasma hominis*) or through catabolism of urea (*e.g.* Ure-

aplasmas) (Dybvig and Voelker, 1996). The degradation of arginine by this pathway results in the equimolar generation of ATP, which is a very ineffective way to produce energy. The non-fermentative and non-arginine-requiring *Mollicutes* hydrolize urea and the intracellular pH increases generating a chemical potential difference across the membrane, which drives ATP formation through membrane-bound ATPase (Romano et al., 1980).

1.3.1.5. *Mollicutes* as minimal model organisms

For many years, scientists have devoted their lives for the search of knowledge and since the cell was discovered, in 1665 by Robert Hooke, it has intrigued us to know how cells work. Due to their complexity, some scientists have dedicated their work on simpler and minimal organisms, such us *Mollicutes* species, assuming that understanding organisms with smaller genomes would be less complex and easier (Glass et al., 2009). Because of the possession of the smallest and simplest genomes of any free-living and self-replicating organism, because of the minimum set of building organelles (plasma membrane, ribosomes and DNA) and the capacity to culture them axenically in a laboratory (Razin and Hayflick, 2010; Balish, 2014), *Mollicutes* species have become appealing to study the minimal cell concept. The interest in these bacteria, as model systems, led to the genome sequence of two of them, *Mycoplasma genitalium* (Fraser et al., 1995) and *Mycoplasma pneumoniae* (Himmelreich et al., 1996a), being among the first few bacteria whose genomes were sequenced (Glass et al., 2009). Studies on several bacteria, specially on *Mollicutes*, have helped to determine the minimal set of genes require for life (Itaya, 1995; Mushegian and Koonin, 1996; Hutchison et al., 1999; Glass et al., 2006; Posfai et al., 2006; Suzuki et al., 2015; Lluch-Senar et al., 2015b; Reuß et al., 2016). Thus, a study on *M. genitalium* has determined that the minimal set of genes essential for life consists of 387 protein-coding and 43 structural RNA genes (Glass et al., 2006), whereas a more recent study on *M. pneumoniae* found 342 essential ORF and 93 ORF categorized as fitness, meaning that their disruption may impaired cellular growth (Lluch-Senar et al., 2015b). These studies attempt to predict a still hypothetical, minimal synthetic *Mycoplasma laboratorium* (Reich, 2000). Additional studies, by the group of Craig Venter, have

made efforts in creating such a bacterium. Gibson et al. (2010) achieved to obtain a synthetic bacterium of 1079 kbp genome and 901 genes, called *Mycoplasma mycoides* JCVI-syn1.0. Ultimately, Hutchison et al. (2016) minimized two-fold the genome of synthetic bacterium JCVI-syn1.0, creating JCVI-syn3.0 with a genome size of 531 kbp and 473 genes.

The use of simple model organisms, with relatively small genomes, facilitates the characterization of core cellular processes and therefore improves our understanding on how an entire cell works (Glass et al., 2009; Suzuki et al., 2015). However, even these simple model organisms contain many genes with unknown functions (Suzuki et al., 2015) and very often multiple functions of a gene are still not consider.

1.3.2. *Mycoplasma pneumoniae*

Mycoplasma pneumoniae exclusively parasitizes humans. Its typical respiratory infection develops slowly and may present pharyngitis, sinus congestion, occasionally otitis media, and eventually include primary atypical pneumonia with fever and bibasilar pulmonary infiltrates, it is also implicated in asthma, and in extrapulmonary manifestations, including neurological, gastrointestinal and dermatological disorders (for a review, see (Razin et al., 1998; Waites and Talkington, 2004; Atkinson et al., 2008; Kishaba, 2016)). Among human mycoplasmas, *M. pneumoniae* is by far the best known and most studied (Atkinson et al., 2008).

M. pneumoniae infection in the respiratory tract may cause epithelial cells to lose their cilia, appear vacuolated, and show a reduction in oxygen consumption, glucose and amino acid uptake, and macromolecular synthesis, resulting in exfoliation and clinical manifestations such as the persistent, hacking cough (Atkinson et al., 2008). Although cytadherence is the initial step in disease development, it is unknown precisely how *M. pneumoniae* injures the respiratory epithelial cell after attachment. However, a variety of virulence factors are thought to be involved in cytotoxicity. For example, the synthesis of hydrogen peroxide and superoxide radicals by *M. pneumoniae* glycerol metabolism (Parrott et al., 2016) are

though to induce oxidative stress in the respiratory epithelium and it was shown that hydrogen peroxide confers hemolytic activity (Waites and Talkington, 2004). Although it was previously thought that hydrogen peroxide was in part the cause of the cytopathology by *M. pneumoniae*, it was recently discovered a new ADP-ribosylating and vacuolating cytotoxin, called Community Acquired Respiratory Distress Syndrome (CARDS) toxin, which also causes cytopathology both *in vitro* and *in vivo* (Kannan et al., 2005; Kannan and Baseman, 2006; Hardy et al., 2009; Kannan et al., 2010; Medina et al., 2012).

M. pneumoniae has a genome size of 816394 basepairs (bp) and possesses the highest G+C content among all *Mollicutes* species with 40 mol% (Himmelreich et al., 1996b). Since the sequence and annotation of its genome, it was found that it possesses all 10 reactions of glycolysis but that the tricarboxylic acid cycle and a complete electron transport chain containing cytochromes are absent (Pollack et al., 2002). Thus, *M. pneumoniae* generates ATP by fermentation of glucose to lactic and acetic acid, as well as, metabolizing glycerol and some other small carbohydrates (Atkinson et al., 2008).

The size of individual spindle-shaped cells of *M. pneumoniae* is 1 to 2 μm long and 0.1 to 0.2 μm wide (Figure 1.11.D.). With respect to a typical bacillus, which is 1 to 4 μm in length and 0.5 to 1.0 μm in width, the *M. pneumoniae* cell volume is less than 5% of that of a typical bacillus. Typical *M. pneumoniae* colonies are 100 μm in diameter, when cultivated on enriched medium such as SP-4 agar (Waites and Talkington, 2004). Due to their small sizes, their colonies need a stereomicroscope to visualize their morphology.

M. pneumoniae present an elongated shape with a tip structure, the attachment organelle (AO) (Figure 1.12), which is involved in adherence, gliding motility and cell division (Balish, 2006; Hasselbring et al., 2006). This AO is formed of a terminal button, paired plates, and a bowl (wheel) complex and is constituted of more than 15 different proteins (Miyata and Hamaguchi, 2016). Among the adhesins taking part of the AO (Figure 1.12), the P1, a protein concentrated in the attachment tip, is the major structure responsible for interaction with host cells. Other

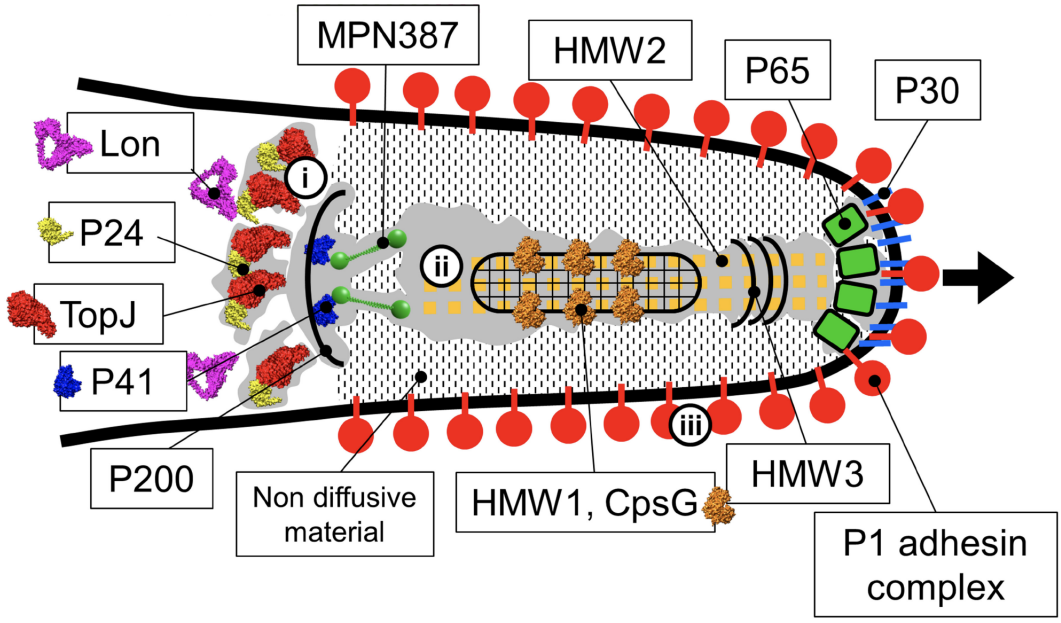


Figure 1.12: Scheme of the attachment organelle (AO) and its component proteins based on EM. The specific localization of proteins involved in formation of the AO is indicated. In the gliding mechanism, the force may be generated at the bowl complex (i), transmitted through the paired plates (ii), and modulated by the P1 adhesin complex (iii). The gliding direction is shown by a black arrow. Figure taken from Miyata and Hamaguchi (2016).

mediators in cytadherence in *M. pneumoniae* include proteins HMW1, HMW2, HMW3, HMW4, HMW5, P30, P90 and P65 that are thought to establish the polar structure (Waites and Talkington, 2004) and a complex with proteins P41 and P24 (which requires TopJ to be stable (Cloward and Krause, 2009)) at the base of the structure is also present (Atkinson et al., 2008). Lastly, an independently assembled complex of proteins B (P90), C (P40), and P1 complete the formation of the functional terminal AO to ultimately facilitate the extracellular parasitic association with host epithelial surfaces(Waites and Talkington, 2004). The AO is the leading end as *M. pneumoniae* cells move by gliding motility (Krause and Balish, 2001).

1.3.2.1. Cell division of *M. pneumoniae*

M. pneumoniae reproduces on agar and in broth culture by binary fission (Furness et al., 1968b,a; Furness, 1968) with previous duplication of its AO, which migrates

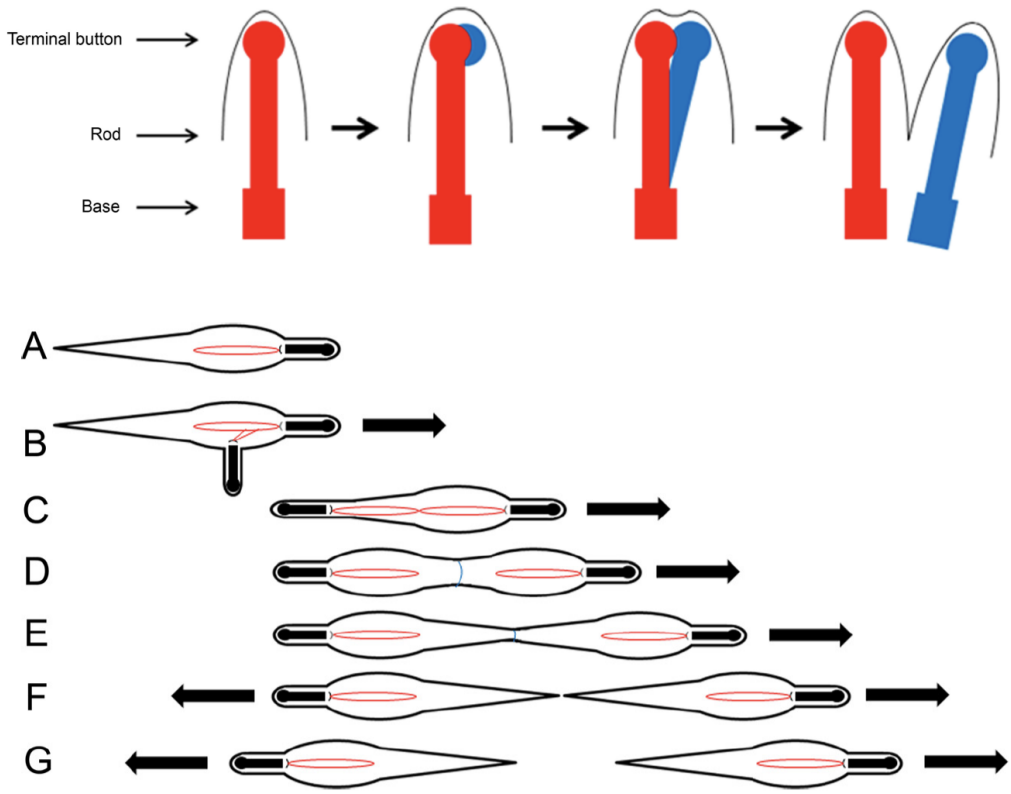


Figure 1.13: Model of duplication of the attachment organelle (AO) core and of cell division in *M. pneumoniae*. ***Upper:*** Model of template-driven duplication of the AO core in *M. pneumoniae*. Duplication of the core seems to begin at the distal terminal button and proceeds toward the cell interior in a TopJ- and P24-dependent manner. Red: original, motility-competent core. Blue: new, motility-incompetent core. ***Lower:*** Model of *M. pneumoniae* cell division. **(A)** In a predivisional cell, the AO contains an electron-dense core (solid black), to the base of which (black arc) the nucleoid (red) is attached. **(B)** Using the TopJ/P24-dependent mechanism, the core duplicates, resulting in a second AO with newly replicated DNA attached. **(C)** Movement of the cell body, driven by motility of only the old AO, pulls the new AO to the opposite pole of the cell. At this point, the new AO is not motility active. **(D)** Two nascent daughter cells become apparent as the old AO continues to stretch and pull the cell body. Constriction of FtsZ (blue) at midcell helps to the separation of the daughter cells. **(E)** The continued combination of pulling and constriction leads to the formation of a filament that connects the nascent daughter cells. **(F)** When the connecting filament ruptures, cytokinesis has occurred. The cell with the old AO continues its forward motion, and the other daughter cell remains inactive for some amount of time. **(G)** The new AO becomes activated for motility, and both daughter cells are now motile. This model is an extension of that of Hasselbring et al. (2006). Figures taken from Balish (2014).

to the opposite pole of the cell during DNA replication and before chromosome separation (Waites and Talkington, 2004). In broth, *M. pneumoniae* forms clusters of around eight cells (the usual colony-forming unit (CFU)), or elementary bodies as Furness et al. (1968b) called them, and on agar it forms colonies (Furness et al., 1968b,a). When batch cultured, in Eaton agent broth enriched with 20% uninactivated horse serum (Chanock et al., 1962), it has a growth curve typical of bacteria, with a lag period, a logarithmic phase, a stationary phase on reaching $4.1 \times 10^7 \pm 1.4 \times 10^7$ CFU per mL broth and a rapid death or logarithmic decline phase (Furness et al., 1968a). The mean generation time on agar (by CFUs) and in broth is 6.5 hours (Furness et al., 1968a). Interestingly, when the inoculum consisted of clusters of cells (CFU suspensions), it had a lag period of 12 ± 3.3 hours. In contrast, when the inoculum consisted of single cells, the lag phase lasted 28 ± 1 hours (Furness et al., 1968a). The increase in the lag phase duration was explained, by Furness et al. (1968a), by the lack of separation of cells after division and the formation of aggregates; thus, after having reached a constant cluster size, the clusters would separate into halves coinciding with the estimated generation time for CFU suspensions.

In many *Mollicutes*, FtsZ cytoskeletal protein and motility participate in cell division (Balish, 2014). As mentioned in the section 1.3.2, *M pneumoniae* cells exhibit an attachment organelle (AO) or terminal organelle, a polar extension of the cell involved in adherence, gliding motility and cell division (Balish, 2006; Hasselbring et al., 2006). An *ftsZ* knockout mutant of *M. genitalium* was successfully obtained indicating that *ftsZ* is not essential for cell division and that cytokinesis may be achieved by gliding motility (Lluch-Senar et al., 2010). Moreover, it was also suggested that non-adherent mutations might be lethal in the Δ*ftsZ* mutant (Lluch-Senar et al., 2010). Likewise, it was shown that nonadherent *M. pneumoniae* mutants are branched or filamentous, supporting the idea of the participation of the AO in cell division (Seto et al., 2001). These studies suggested synthetic lethality of loss of FtsZ and AO functions (Balish, 2014).

During the *M. pneumoniae* cell cycle, a new, adherence-active AO is buit next to the existing one at the initiation of DNA replication (Figure 1.13). The associ-

ation of DNA with the proximal end of the core of the AO was demonstrated by Scanning electron microscopy (SEM) in *M. pneumoniae* and other *Mycoplasma* species (Hatchel and Balish, 2008). Thus, before the new AO is motility active, the old AO pulls the cell body such that the new AO is moved apart to the opposite pole. The two nascent daughter cells become apparent as the old AO stretched and pulls forward the cell body. Together with the constriction of FtsZ at midcell, the daughter cells separate. Eventually, the new AO becomes active and the cell with the new AO starts to glide (Balish, 2014). However, in *M. pneumoniae* cell division, the AO duplication and cytokinesis are not tightly coordinated because multiple duplications of the AO were observed before cytokinesis of the daughter cells, indicating overlapping cycles (Hasselbring et al., 2006). Concerning the coordination of the AO duplication with the DNA replication is still unclear (Hasselbring et al., 2006).

Overall, in the model of *M. pneumoniae* cell division (Figure 1.13), cytokinesis is carried out by the combination of the presence of FtsZ and AO-mediated motility, which can only take place when the cell is adherent to a surface (Balish, 2014). Both components can individually perform inefficient cell division, but both are required for efficient cell division (Balish, 2014).

1.3.2.2. Genetic tools of *Mycoplasma pneumoniae*

Most of the established genetic tools available in the well-studied model organisms *E. coli* or *B. subtilis*, are unavailable for *Mollicutes* species.

Transposon mutagenesis Most studies in *Mollicutes* use suicide vectors or the transposons Tn*916* and Tn*4001* for study gene function (Lee et al., 2008a). In particular, transposons may be used for both random insertion of a gene of interest into the chromosome and for disruption of genes (Halbedel and Stülke, 2007). To increase the stability of mutants carrying transposon insertion, a mini-transposons of Tn*4001* was constructed in such a way that the transposase gene was outside the transposable elements to prevent the transposon from re-excision and jumping (Pour-El et al., 2002).

Replicative plasmids Some studies have successfully developed vectors for *Mollicutes* using homologous origins of replication (*oriC)* and selectable antibiotic resistance markers (Renaudin et al., 1995; Duret et al., 1999; Cordova et al., 2002; Lartigue et al., 2002, 2003; Janis et al., 2005; Lee et al., 2008a,b; Breton et al., 2012; Maglennon et al., 2013; Shahid et al., 2014). These *oriC* plasmids are able to replicate extrachromosomally but in many cases they integrate into the *oriC* region of the chromosome by homologous recombination after several *in vitro* passages. Remarkably, it was observed that plasmids containing shorter *oriC* regions are more stable and thus, they have been used to achieve targeted homologous recombination (Lee et al., 2008a). Despite of the efforts for developing artificial *oriC* plasmids in *M. pneumoniae*, they are still not stable in this badterium and therefore studies on *M. pneumoniae* rely on the random insertion of transposon mutagenesis (Balish, 2014).

Targeted gene knockout by homologous recombination The targeted gene knockout by homologous recombination has only been reported in a few *Mollicutes*, such as *M. genitalium*, *M. gallisepticum*, *M. pulmonis* and *Acholeplasma laidlawii* (Halbedel and Stülke, 2007). The inefficiency or the absence of homologous recombination events *in vitro* has been attributed to the presence of a non-functional RecU Holliday junction resolvase, MPN528a, in *M. pneumoniae* subtype 2 (*e.g.* FH) strains and to its absence in subtype 1 strains (*e.g.* M129), respectively (Sluijter et al., 2010). Supporting this results, doble crossover recombination events were observed *in vivo* in higher proportion in *M. pneumoniae* subtype 2 than in subtype 1 (Krishnakumar et al., 2010).

Site-specific recombination Being the first site-specific recombination system developed in *Mollicutes*, the Cre/lox recombination system was successfully implemented in *M. genitalium*, the closest relative of *M. pneumoniae* (Mariscal et al., 2016).

Inducible promoter system The first inducible promoter systems in *Mollicutes* was shown to tightly regulate gene expression by the tetracycline-inducible promoter P*xyl/tetO*$_2$ from *B. subtilis* (Breton et al., 2010). This system was eval-

uated in two *Mollicutes*, the plant pathogen *Spiroplasma citri* and the animal pathogen *Mycoplasma agalactiae* (Breton et al., 2010). This systems was tested successfully in the human pathogen *M. genitalium* with some modifications in the P*xyl/tetO*$_2$ promoter sequence to diminish transcriptional leakage (Mariscal et al., 2016).

1.3.2.3. Systems biology of *Mycoplasma pneumoniae*

The circular genome of *M. pneumoniae* M129 was completely sequenced in 1996 (Himmelreich et al., 1996a), it was among the fist genomes to be sequenced. It consists of 816,394 bp with 677 predicted open reading frames (ORFs) and 39 genes coding for RNAs (Himmelreich et al., 1996a). In contrast, *E. coli* genome consists of 4,639,221 bp and 4288 predicted protein-coding genes (Blattner et al., 1997), which accounts for almost six times larger than that of *M. pneumoniae*.

To date, there is a number of published omics-driven studies on *M. pneumoniae*, including genomics (Himmelreich et al., 1996a, 1997; Fukuda et al., 1999; Dandekar et al., 2000; Pachkov et al., 2007), epigenomics (Xia, 2003; Lluch-Senar et al., 2013), chromosome structure (Trussart et al., 2017), transcriptomics (Weiner III, 2003; Güell et al., 2009; Yus et al., 2012; Lloréns-Rico et al., 2015; Junier et al., 2016; Lloréns-Rico et al., 2016) , proteomics (Ueberle et al., 2002; Su et al., 2007; Kühner et al., 2009; Schmidl et al., 2010; van Noort et al., 2012; Borràs et al., 2013; Lluch-Senar et al., 2016) , metabolism (Desantis et al., 1989; Pollack et al., 1996, 1997, 2002; Halbedel et al., 2004; Halbedel and Stülke, 2005; Yus et al., 2009; Wodke et al., 2013) and the integration of those -omics data (Maier et al., 2011, 2013; Chen et al., 2015). All these studies have contributed to the increase of the comprehensive qualitative and quantitative understanding of the systems biology of *M. pneumonia*.

1.4. Objectives

The main objective of this thesis is to identify the genetic factors that regulate the growth rate in *M. pneumoniae* and in other *Mollicutes* species to ultimately increase the growth rate of *M. pneumoniae* to be used as a model organism for synthetic biology and more specifically to develop attenuated live vaccines applied in animals.

Despite of the existence of well-established methods to measure growth rates in many eubacteria, these methods are not applicable in the majority of *Mollicutes* species. Thus, in Chapter 2, in an attempt to find a high throughput and quicker method to estimate growth rates in *M. pneumoniae* and of adherent *Mollicutes* species, I developed and optimized two alternative methods and compared them with common methods currently used in these wall-less bacteria.

In Chapter 3, we aimed to rationally design fast-growing *M. pneumoniae* mutants by taking a hypothesis-driven approach, in which we reviewed the scientific literature on bacterial growth physiology and particularly looked at the genetic differences between fast- and slow-growing bacteria. Accordingly, I then generated mutants of *M. pneumoniae* by genetic modifications, performed their growth curves by using several methods to estimate growth rates and investigated more deeply the changes in expression of those genetic modifications.

In Chapter 4, we aimed to find fast-growing *M. pneumoniae* transformants from Serrano's lab collection by making a medium-throughput screening, in which we screened around 240 transformants (single transformants with gene disruptions, gene overexpressions or dominant negatives genes) for faster growers, and then, I validated those positives with a different method to measure growth rates. Furthermore, we grouped the genes of the faster-growing single mutants into three functional categories, namely, transcriptional, translational and metabolic genes, and then obtained the correspondent multiple mutants per functional category. Finally, I tested whether the multiple mutants showed an additive effect on growth.

In Chapter 5, we aimed to identify differently expressed genes of fast- and slow-growing *Mollicutes* species that correlate with fast growth by taking a data-driven approach, in which we compared transcriptomics and proteomics data of fast- and slow-growing *Mollicutes* species. We found the homologous genes among those species and grew them in the same culture media to prepare the samples for transcriptomics and proteomics procedures. I performed different normalization strategies in order to analyze the data and infer relevant conclusion.

Chapter 2

METHODS TO ESTIMATE BIOMASS AND MEASURE GROWTH RATES IN *MOLLICUTES* SPECIES

2.1. Introduction

A number of methods have been described to measure biomass and therefore to calculate growth rates. For instance, to estimate biomass the total mass; the mass of protein, DNA, ATP or other cell component; the amounts of substrate consumed or product formed; the metabolic rates; light scattering or turbidity; and viable cell counts are commonly measured (Maniloff, 1992). However, the determination of biomass in *Mollicutes* is particularly difficult due to their osmotic fragility, their tendency to form filaments, their small cell size and the low cell yields (10^7 to 10^9 cells per mL) (Stemke and Robertson, 1982; Stemler et al., 1987; Maniloff, 1992). Typically, *E. coli* cells grown exponentially in rich medium are rod-shaped and are roughly 3 μm long by 1 μm wide (Rowlett and Margolin, 2015), whereas *M. pneumoniae* barely reaches 1 to 2 μm long and 0.1 to 0.2 μm wide, which make difficult their observation by light microscopy because its size is very close to the resolution limit of light (200 nm). Ultimately, total mass determination in

Mollicutes is limited by lack of sensitivity, *e.g.* to achieve an accuracy $<2\%$ of dry mass, 50 mg are needed, however, the highest mollicute concentrations reached in culture are 0.5 mg of dry mass per mL (Maniloff, 1992).

Other methods assume that cell components and substrate consumption/product formation are proportionally constant to dry mass and growth yield, respectively. However, it is has been demonstrated that cell components, in a lesser extent the DNA, vary with the specific growth rate; thus, the faster the growth rate, the higher the cellular chemical composition (Schaechter et al., 1958; Snell, 1981; Wanner and Egli, 1990; Neidhardt, 1999; Dennis and Bremer, 2008; Klumpp et al., 2009; Thomas, 2015). Normally, protein determination is more widely used than genomic DNA content, because of its larger sensitivity, *i.e.* for accurate determination of DNA, around ten times more of cell mass is needed than that for protein (Maniloff, 1992), however there are a number of studies doing growth curves by DNA measurements (Burgos et al., 2012; Karas et al., 2014; Suzuki et al., 2015; Hutchison et al., 2016). ATP content in *Mollicutes* species is also very sensitive, rapid and correlates with viable cell counts (Saglio et al., 1979; Beaman and Pollack, 1981, 1983; Stemler et al., 1987; Calus et al., 2010).

Alternatively, for fermentative *Mollicutes* species, decrease in pH and glucose consumption were used as growth parameters in *M. gallisepticum*. The results showed that medium color change, tracked by adding the pH indicator phenol red (where medium color changes from red to yellow during growth) followed by measures of its optical density at 560 nm (OD_{560}), had a slightly higher correlation with viable counts than glucose consumption, although the reproducibility was poorer (Snell, 1981). Similarly, the medium color change was monitored, at absorbance at 550 nm (A_{550}) as an indicator of growth, during the early phases of *M. mycoides* subsp. *capri* batch cultures, when the medium does not show turbidity. The results showed an inverse linear relationship between the A_{550} and the log(CFU/mL) (Meur et al., 1988). Likewise, the medium color (pH) change has been used as an indicator of growth of *M. pneumoniae* (Yus et al., 2009) and *M. genitalium* (Karr et al., 2012) batch cultures in high-throughput set-ups by means of readings of the absorbance ratio at 430/560 nm ($A_{430/560}$) and A_{550}, respec-

tively, in 96-well plates in a Tecan plate reader . On the other hand, the rate of reduction of tetrazolium to form the red-colored formazan was used to estimate the amount of *M. pneumoniae* viable cells (Bredt, 1976) and correlated well with CFU/mL for 22 *Mycoplasma* and *Acholeplasma* species (Kirchhoff et al., 1992).

The small size of many *Mollicutes* species, the tendency to have a filamentous growth and to adhere to glass surfaces, the lack of a rigid cell wall and the low maximum titers prevent light microscopy as well as light scattering from measuring cell growth (Stemke and Robertson, 1982; Stemler et al., 1987) Despite these limitations, after comparing seven different methods for estimating *M. gallisepticum* growth, light scattering, measured at OD_{560}, was the quickest and with the highest reproducibility and correlation with viable counts (Snell, 1981). Even though turbidity may work with some species, for accurate biomass determinations, 100 µg of cell dry mass per mL are required, therefore viable counts are more often used (Maniloff, 1992). Viable cells can be counted by either colony-forming unit (CFU) or color-changing unit (CCU). CFU determination on agar underestimates the titers, specially in clumping cells of *Mollicutes* species (Stemler et al., 1987; Calus et al., 2010). CCU determination in broth proved to be more sensitive, gives higher estimates of cell counts and a higher correlation with DNA content (Stemke and Robertson, 1982), although some studies also found that CCU underestimates the actual cell concentration compared to ATP titres (Calus et al., 2010). Although, the drawback of viable cell methods is that they are cumbersome and required long incubation times, about 1 week for fast-growing or 2 for slow-growing *Mollicutes* species, like *M. pneumoniae*.

Moreover, *Mollicutes* species (*e.g. M. pneumoniae* and *M. genitalium*) are usually grown in statically incubated cultures, *i.e.* without shaking or stirring, allowing them to adhere to surfaces. This leads to the formation of concentration gradients of nutrients and toxic products, which impairs cell growth (Maniloff, 1992).

In an attempt to find a high throughput and quicker method to estimate growth of *M. pneumoniae* and of adherent *Mollicutes* species, I developed new alternative methods in broth cultures and compared them with common methods currently

used in these wall-less bacteria.

2.2. Materials and methods

2.2.1. Bacterial strains and culture conditions

One Shot®TOP10 Chemically Competent *E. coli* (Invitrogen, Cat. No. C4040) or NEB 5-alpha Competent *E. coli* (New England Biolabs, Cat. No. C2987H) was used for plasmid amplification. They were grown in Luria-Bertani (LB) broth or LB agar plates containing 100 μg/mL ampicillin at 37°C.

Wild type *M. pneumoniae* strain (str.) M129 (WT_{M129}) (ATTC 29342, subtype 1, broth passage no. 35) (Regula et al., 2000) was used. M129 *Δldh* mutant of *M. pneumoniae* str. M129 was gently provided by Jörg Stülke and obtained by haystack mutagenesis (Halbedel and Stülke, 2007). M129 *Δldh* is a pure mutant whose lactate dehydrogenase, encoded by the *ldh* (*mpn674*) gene, has been disrupted by the mini-transposon pMT85 (Zimmerman and Herrmann, 2005). *M. pneumoniae* cells were grown in modified Hayflick's medium (hereafter referred to as Hayflick) as previously described (Hayflick, 1965; Yus et al., 2009), the medium for M129 *Δldh* mutant was supplemented with 200 μg/mL gentamicin and were grown at 37°C in a humidified incubator gassed with 5% carbon dioxide (CO_2).

2.2.2. Transformation of *M. pneumoniae*

WT_{M129} cells were transformed with pMTn*Gm* plasmid, a negative control vector to use as a control of M129 *Δldh*, containing the *aac(6')-aph(2")* marker, which confers gentamicin resistance (Pich et al., 2006) or with pMTn*TetM438* plasmid (Pich et al., 2006) by electroporation (Hedreyda et al., 1993) with modifications. Briefly, a frozen stock of WT_{M129} was diluted 1:100 with modified Hayflick broth, the cells were grown in 75 cm^2 tissue culture flasks containing 30 mL medium and incubated at 37°C under 5% CO_2 to late exponential growth phase (for 72 h). Cells were washed twice, resuspended in precooled electroporation buffer (8 mM

HEPES, 272 mM sucrose, pH 7.4), scraped off and passed through a 25-gauge (G25) syringe needle ten times. Aliquots of 50 μL of cells in 0.1 cm cuvettes with 0.3 pmols of the corresponding plasmid were kept on ice for 15 min and later electroporated by using a Bio-Rad gene pulser (1250 V, 25 μF, 100 Ω). After 15 min on ice, 1 mL Hayflick broth was added and the cells were incubated for 2 h at 37°C in the presence of 5% CO_2. Aliquotes of 50uL were inoculated in 25 cm^2 tissue culture flasks containing 5 mL Hayflick broth supplemented with 200 μg/mL gentamicin. Broth cultures were harvested with a cell scraper, as soon as the medium turned orange, and stored at -70°C in fresh medium. Then, these stocks were used for inoculation of two more consecutive passages.

2.2.3. Growth curves conditions

To compare the different methods to measure growth in M129 *Δldh* and its negative control, M129 pMTn*Gm* transformants, growth curves were performed in batch cultures at 37°C under 5% CO_2. Each one of the methods aims at measuring number of cells in an indirect way, for instance by measuring cellular components: total cellular protein, total cellular ATP concentration and the medium color change of the culture by the production of acids.

Batch precultures were inoculated in 1:100 dilution from the stocks in 25 cm^2 tissue culture flasks containing 5 mL fresh Hayflick broth supplemented with gentamicin to late exponential phase (for 60 h). Medium was removed, cells were suspended in 1 mL fresh Hayflick broth without gentamicin, scraped off, passed through a G25 syringe needle ten times and used as the seed cultures to re-grow the growth curves explained in the coming sections.

2.2.3.1. Total cellular protein quantification in batch and semicontinuous cultures

Fresh prepared seed cultures were inoculated in 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL Hayflick broth without antibiotic. There were as many flasks as time points and duplicates. To extract total cellular protein at selected times, 200 μL of medium were kept for medium color (pH) change and

metabolite quantification (see section 2.2.3.3), the remaining was removed, cells were washed twice with phosphate-buffered saline (PBS; for 1 liter: 8 g NaCl, 0.2 g KCl, 0.24 g KH_2PO_4, 1.44 g $Na_2HPO_4 \cdot 2H_2O$, pH 7.4), scraped and pelleted by centrifugation at 14100 $\times$g for 10 min. PBS was removed and pelleted cells were stored at -70°C for further processing. After collecting samples twice per day for 7 days, frozen pelleted cells were suspended in 60 µL lysis buffer (4% Sodium Dodecyl Sulfate (SDS), 0.1M HEPES). Cell lysates were kept on ice and disrupted by using a Bioruptor®sonication system (Diagenode, B01010004) with an On/Off interval time of 30/30 sec at high frequency for 10 min. Finally, cell lysates were spun down, pipetted up and down to complete lysis and extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). Briefly, cell lysates taken from time points after 48 h of growth were diluted five-fold, before carrying out the assay, to be within the working range of the assay. The standards were prepared with Pierce™ Bovine Serum Albumin Standard (BSA) (Thermo Scientific, Cat. No. 23209) diluted at different concentrations with lysis buffer, following manufacturer's instructions. 25 µL of samples and BSA standards were added to a well of a 96-well plate, by duplicate. Then, 200 µL of BCA working reagent were added to each well and mixed in the Infinite™ M200 Tecan plate reader for 30 seconds. Samples were incubated at 37°C for 30 minutes, and after cooling down to room temperature, absorbance at 562 nm was measured using the Infinite™ M200 Tecan plate reader. Known concentrations from the BSA standards were used to make the standard curve and obtain protein concentration from each sample.

In contrast to batch cultures, in semicontinuous cultures the medium was changed every 12 h by fresh and pre-warm medium. The experiments were set up as described above, except that the fresh prepared seed cultures were not immediately inoculated but frozen and stored at -70°C and aliquots thereof were quantified by Pierce™ BCA Protein Assay Kit. 10 µg of these quantified frozen stocks were inoculated in 25 cm^2 tissue culture flasks containing 5 mL Hayflick broth without antibiotic.

To estimate the specific growth rate (h^{-1}) and doubling time (h), the exponen-

tial growth phase curve was transformed into the natural logarithm of the protein concentrations and it was fitted with a linear regression model, whose slope (corresponding to the specific growth rate) was determined with three adjacent points. The specific growth rate (μ) and the doubling time (DT) were estimated by using the equations 2.1 and 2.2, respectively.

$$\mu = \frac{lnP_2 - lnP_1}{t_2 - t_1} \tag{2.1}$$

where:

$t_2 - t_1$ = time interval
P_1 = protein concentration at t_1
P_2 = protein concentration at t_2

$$DT = \frac{ln2}{\mu} \tag{2.2}$$

Statistical data analysis of significant differences between the two mutants was performed by Student's unpaired t-Tests, with a two-tailed distribution and assuming unequal variances.

2.2.3.2. Total cellular ATP quantification

Fresh prepared seed cultures were inoculated in 1:70 dilution in 24-well cell culture plates containing 1 mL Hayflick broth without antibiotic. There were as many wells as time points and duplicates. Total cellular ATP at each time point was extracted and measured by using the ATP Bioluminescence Assay Kit HS II (Roche, Cat. No. 11 699 709 001). Briefly, the medium was removed completely from the well, cells were suspended in 500 μL of dilution buffer and 500 μL of cell lysis reagent, both reagents from the assay kit, and incubated for 5 min at room temperature to release ATP. Cells lysates were scraped with the tip and the suspension was pipetted up and down, recovered in a 1,5 mL tube and immediately frozen at -70°C for further processing. After collecting samples twice per day for 7 days, frozen ATP extracts were kept on ice and, to keep the samples within the working range of the assay, diluted 50-fold with dilution buffer from the kit, then 50 μL of

diluted samples were transferred into a 96-well tissue-culture treated BW isoplate (Perkin Elmer, Cat. No. 6005060). ATP standards and samples were injected automatically with 50 μL of luciferase reagent and measured by duplicate by using a Victor3 multilabel plate counter (Perkin Elmer, Model 1420). The specific growth rate and doubling time were estimated as in section 2.2.3.1 from the ATP concentrations.

2.2.3.3. Medium color (pH) change and metabolite quantification

As *M. pneumoniae* is a fermentative microorganism, it transforms glucose into acids (acetic and lactic acid). Thus, it lowers the pH of the medium and when a pH indicator, like phenol red, is used, medium pH change can be measured indirectly by the medium color change from red to yellow. 200 μL of medium, from the intracellular protein determination in batch cultures (section 2.2.3.1), were collected in 96-well plate and stored at -20°C for further processing. After collecting samples twice per day for 7 days, frozen medium samples were thawed and prepared for measuring medium color (pH) change, glucose consumption and acetic and lactic acid production. Medium color (pH) change was calculated by taking the ratio between the absorbance at 430 and 560 nm ($A_{430/560}$) taken with a settle time at 300 ms and 25 flashes, by the Infinite^{TM} M200 Tecan plate reader as previously described (Yus et al., 2009). Extracellular glucose, acetic and lactate acid concentrations in the medium were estimated by means of the commercial Glucose Colorimetric/Fluorometric Assay Kit (BioVision, Cat. No. K606-100), Acetic Acid kit (Megazyme, K-ACETRM), and Lactate Colorimetric/Fluorometric Assay Kit (BioVision, Cat. No. K607-100), respectively. The readings were carried out as described in the kit protocols in 96-well plate using the Infinite^{TM} M200 Tecan plate reader. Before performing the assay, to be within the working range of the assay, samples were diluted as follows: for glucose 1:500; for acetic acid 1:3 for time points between 30-60 h and 1:10 after 60 h; and for lactate, for M129 pMTn*Gm* 1:50 for time points before 30 h and 1:300 after 30h, for M129 *Δldh* 1:50. All samples and standards were measured by duplicate.

2.2.3.4. Colorimetric growth assay by serial dilutions

Fresh prepared seed cultures were used to make two five-fold and two ten-fold serial dilutions in Hayflick broth without antibiotic as previously described by Karr et al. (2012), with minor modifications. Dilutions were performed in duplicates and 200 µL of each diluted sample was plated in a 96-well plate excluding the use of peripheral wells to avoid the edge effect, thus reducing it to a 60-well plate. The 96-well plate was incubated in the Infinite^{TM} M200 Tecan plate reader at 37°C and the absorbance at 550 (A_{550}) was automatically taken, with a settle time at 300 ms and 25 flashes, every 20 min for 4 days. The specific growth rate (h^{-1}) and doubling time (h) were estimated as previously described (Karr et al., 2012) by using the equations 2.3 and 2.4, respectively.

$$\mu_{Karr} = \frac{ln(dilutionfactor)}{\Delta t} \tag{2.3}$$

where:

Δt = time required by more diluted cultures to reach the same A_{550}
$dilutionfactor$ = 5 or 10 if the serial dilutions were 5- or 10-fold, respectively

$$\tau_{Karr} = \frac{ln(2)}{\mu_{Karr}} \tag{2.4}$$

2.2.4. Automated high-throughput microcolony growth assay

Fresh prepared seed cultures of WT_{M129} and M129 pMTn*TetM438* cells were obtained as in section 2.2.3.1, but aside from passing the cells through a G25 syringe needle ten times to de-clump the cells, in order to separate the single cells from any remaining aggregates, the cell suspension was passed through a Millex-GV filter of 0.22 µm pore diameter (Merckmillipore, Ref. SLGV033RS). The filtered suspension was used to make four ten-fold serial dilutions in Hayflick broth. Dilutions were performed in duplicates. 96-well Imaging Plate Cover Glass (Mobitec, Prod. No. 5241-20) was previously coated or not with 200 µL of 3% gelatin (Sigam-Aldrich); 200µg/mL Concanavalin A Type IV (Sigam-Aldrich, Cat. No. C2010); or 0.1 mg/mL poly-L-Lysine (Sigam-Aldrich, Cat. No. P8920) during 24 h at 37°C, wells were washed twice with 200 µL of sterile Milli-Q water and dried

at 4°C for 24 h. 200 µL of each diluted sample were plated per well, excluding the peripheral wells, in the coated 96-well imaging plate and incubated at 37°C and under 5% CO_2 for different times (2, 6, 10, 16, 22 h) to allow cells to attach to the glass surface. After this time, medium was change to remove floating cells and the 96-well plate was incubated at 37°C in an ImageXpress®Micro Widefield High-Content Anaysis microscope (Figure 2.3.A.) at the Advanced Light Microscopy CRG facility. Time-lapse images of 9 fields per well were acquired automatically every 2 h for 5 days (Figure 2.3.B.) with a Leica Plan Apo 63X (1.35 numerical aperture) air objective using the MetaXpress®software (version 1.7 for Microsoft XP).

Image processing was done by using Fiji (ImageJ)-based scripts made by the Advanced Light Microscopy CRG facility and data analysis was done by using a homemade R studio-based script. Data may be analyzed by either a growth curve by percentage of area per field or by single microcolony area (Figure 2.3.C.). Although, only with the latter, the natural logarithm of the area was fitted with a linear regression model and the slope was calculated with adjacent points. Microcolony area growth curves presenting negative slopes, less than 10 adjacent points or an R-squared less than 0.9 were filtered out. The specific growth rate and doubling time were determined by using the equations 2.1 and 2.2, respectively.

2.3. Results

2.3.1. Growth curves of *M. pneumoniae* mutants to illustrate the problematic in following the growth rate

As mentioned in the introduction of this chapter, determining *M. pneumoniae* growth rate is not simple and many of the methods rely on the metabolic activity of the cell, which is not always related to the doubling time. A perfect example of this complexity is the M129 *Δldh* mutant, whose lactate dehydrogenase, encoded by the *ldh* (*mpn674*) gene, has been disrupted by a mini-transposon. *M. pneumoniae* can produce lactate from pyruvate, the final product of glycolysis, using Ldh when oxygen is in short supply, while it keeps a constant production

of acetic acid. Production of acetic acid results in the co-production of one ATP molecule and requires oxygen to balance redox, while production of lactate balances redox after the inbalance generated in glycolysis (Yus et al., 2009). In a normal growth curve of *M. pneumoniae,* the medium acidifies due to secretion of acetic and lactic acids. At the end of exponential growth phase, the ATP peaks and drops rapidly marking the beginning of the stationary growth phase. Figure 2.1 shows the growth of M129 *Δldh* and its negative control M129 pMTn*Gm* in batch cultures. We compared growth by the following parameters: intracellular protein (Figure 2.1.A.), ATP concentrations (Figure 2.1.B.), medium color (pH) change (Figure 2.1.C.), quantification of extracellular metabolites (glucose, acetate, lactate; Figure 2.1.D-F., respectively) and a colorimetric growth assay by serial dilutions (Figure 2.1.G-H.).

2.3.1.1. Total cellular protein quantification

The intracellular protein growth curves (Figure 2.1.A.) showed no significant differences between the growth of M129 pMTn*Gm* and M129 *Δldh* transformants, with doubling times during the first 30 h of growth of 10.10 ± 1.63 h and 11.37 ± 1.85 h (*P*-value = 0.54), respectively (Table 2.1). After 30 h of growth, the growth curves seem very similar.

2.3.1.2. Total cellular ATP quantification

In contrast to protein growth curves, after 30 h of growth (Figure 2.1.B.), M129*Δldh* begun to enter to the stationary phase and the concentration of ATP remained low and stable throughout the experiment, while M129 pMTn*Gm* presented an exponential increase and a rapid drop of ATP concentration, during the stationary phase, as soon as the metabolic activity slowed down. These results mistakenly suggest that M129 *Δldh* grows slower than its control M129 pMTn*Gm*. Nevertheless, during the first 30 h of intracellular ATP growth curves, the doubling times, which were 6.48 ± 0.31 h for M129 pMTn*Gm* and 7.01 ± 0.04 h for M129 *Δldh* (Table 2.1), did not showed significant differences between the growth of these mutants (*P*-value = 0.25).

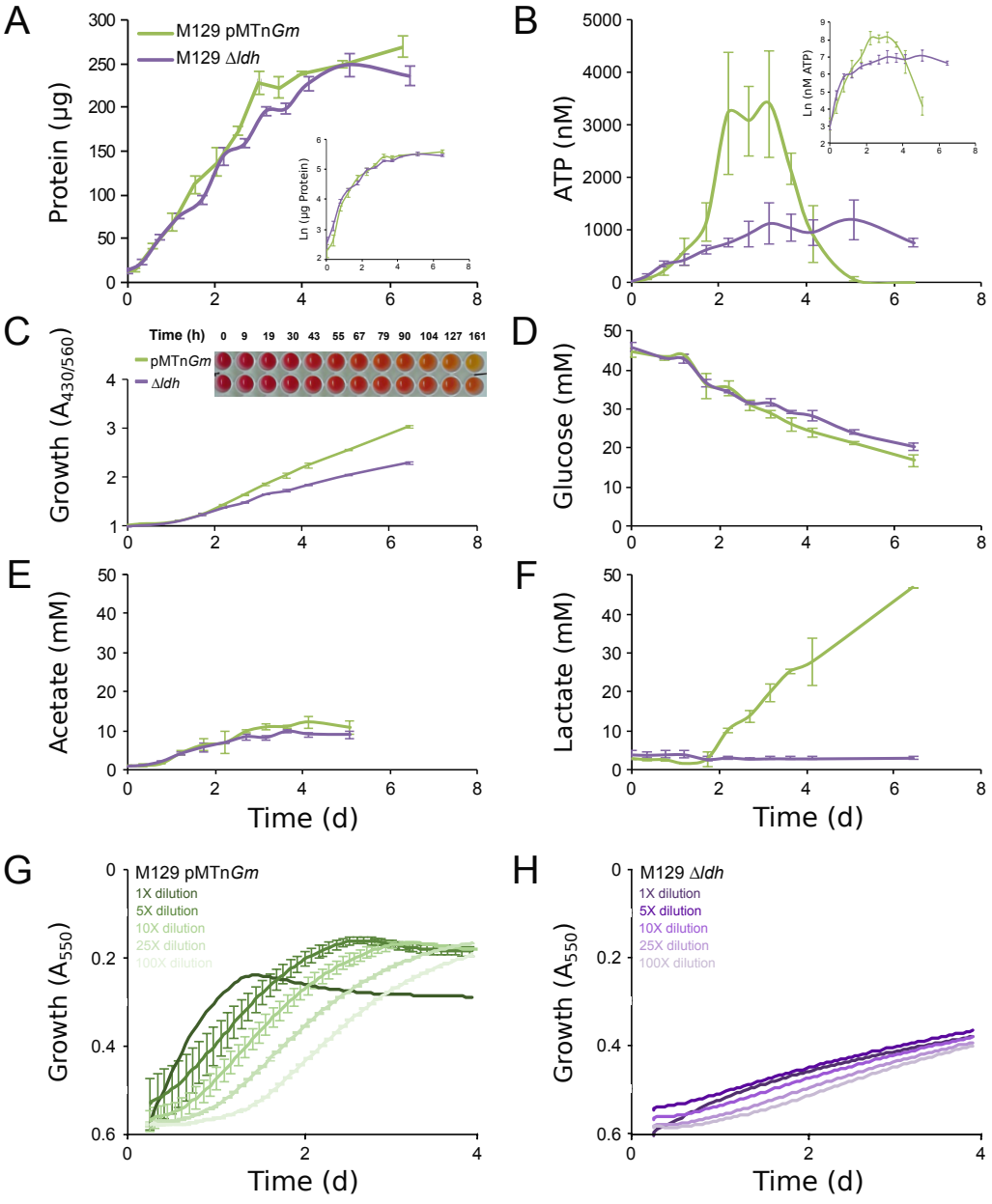


Figure 2.1: Growth curves (GC) (in days) of *M. pneumoniae* by different methods in batch cultures. M129 *Δldh* is indicated in violet and its negative control M129 pMTn*Gm* in green. The insets show the natural logarithm of the concentrations. $n = 4$ (two independent experiments with two technical replicates each) unless otherwise specified. Error bars = standard deviation (SD). **(A)** Total cellular protein as a measure of biomass. **(B)** Total cellular ATP concentration. **(C)** Medium color (pH) change, as a measure of growth, monitored at absorbance ratio at 430 and 560 nm ($A_{430/560}$). **(D, E, F)** Quantification of extracellular **(D)** glucose, **(E)** acetate, and **(F)** lactate concentration. **(G, H)** Medium color (pH) change of five cultures, with the inoculum diluted 1-, 5-, 10-, 25-, and 100-fold, $n = 2$ (two independent experiments) measured by absorbance at 550 of (A_{550}), of **(G)** M129 pMTn*Gm*, and **(H)** M129 *Δldh* mutants.

2.3.1.3. Medium color (pH) change and metabolite quantification

The medium color change measured as the ratio of absorbance at 430 and 560 nm ($A_{430/560}$) anticorrelates with pH, the lower the ratio, the higher the pH (Yus et al., 2009). During the first 30 h, the curves of medium color change did not show any difference in the two mutants (Figure 2.1.C). But, at the end of the curves, as expected, M129 *Δldh* acidified much less the medium than its control M129 pMTn*Gm*, mistakenly suggesting that M129 *Δldh* grows slower than its control M129 pMTn*Gm*. In fact, the reduced decrease of pH in M129 *Δldh* was due to the lack of lactate production (Figure 2.1.F), as a consequence of the *lactate dehydrogenase* gene disruption in this mutant. Thus, the minor mediun acidification in M129 *Δldh* was only the result of the production of acetic acid (Figure 2.1.E), which was similar to that of its control M129 pMTn*Gm*. Interestingly, the glucose consumption (Figure 2.1.D) was very similar at the beginning of the curve but from the third day onwards, it was slightly lower in M129 *Δldh* than in its control, although this difference was not significant (*P*-value = 0,12). Altogether, during the growth of M129 pMTn*Gm*, the fermentation of sugars resulted in the production of lactic and acetic acid which lowered much more the pH of the medium that that of M129 *Δldh* mutant, which only produced acetic acid.

2.3.1.4. Colorimetric growth assay by serial dilutions

The readings for this assay (A_{550}), in contrast to $A_{430/560}$, correlate with pH, the lower the A_{550}, the lower the pH. The curves of this assay showed a clear difference between M129 *Δldh* (Figure 2.1.H) and its control M129 pMTn*Gm* (Figure 2.1.G). In consistency with the results of the former method, M129 *Δldh* acidified the medium much less than its control M129 pMTn*Gm*. Thus, the doubling times calculated for M129 pMTn*Gm* and M129 *Δldh* were 6.61 $\pm$ 0.51 h and 2.29 $\pm$ 0.72 h, respectively (Table 2.1), being significantly different (*P*-value = 0.92×10^{-4}) and mistakenly suggesting that M129 *Δldh* grows much faster than its control M129 pMTn*Gm*.

Table 2.1: Estimated doubling times (DT) by different methods. The DT for protein and ATP concentration were calculated during the first 30 h of growth. For the color change dilutions assay, the Δt was taken when the diluted cultures reached an A_{550} = 0.4. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. For the microcolony growth assay, the shown DT corresponds to the coating with Concanavalin A and 10-fold dilution. Values indicate the mean $\pm$ the standard deviation (SD). - indicates not measured. Statistical significant differences compared to the corresponding left columns are indicated: NS: non-significant *P*-value, * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

	Type of cells	
Doubling time (h) by method	**M129 pMTn*Gm* (- ctrl.)**	**M129 *Δldh***
Batch cultures		
Protein concentration	10.10 ± 1.63	11.37 ± 1.85^{NS}
ATP concentration	6.48 ± 0.31	7.01 ± 0.04^{NS}
Medium color (pH) change dilutions	6.61 ± 0.51	$2.29 \pm 0.72^{***}$
Semicontinuous cultures		
Protein concentration 0-88 h	14.22 ± 0.08	-
Protein concentration 0-48 h	9.47 ± 0.24	$10.82 \pm 0.29^{***}$
(Corrected doubling time)	(9.47 ± 0.15)	$(10.76 \pm 0.17)^{***}$
	WT_{M129}	**M129 pMTn*TetM438* (- ctrl.)**
Batch cultures		
Microcolony growth assay	9.62 ± 1.91	$12.65 \pm 2.59^{***}$

2.3.2. Developing new methods to measure growth rates in adherent *Mollicutes* species

The methods described in the section 2.3.1 are not suitable for high-throughput screenings, except for the Colorimetric growth assay by serial dilutions, which is not suitable for metabolic mutants since they change the pH in a different way than the WT cells. Therefore, any output derived from medium color (pH) change experiments should be validated by other methods not relying on metabolic parameters. In the following, two new methods are described to measure growth rates in adherent *Mollicutes* species.

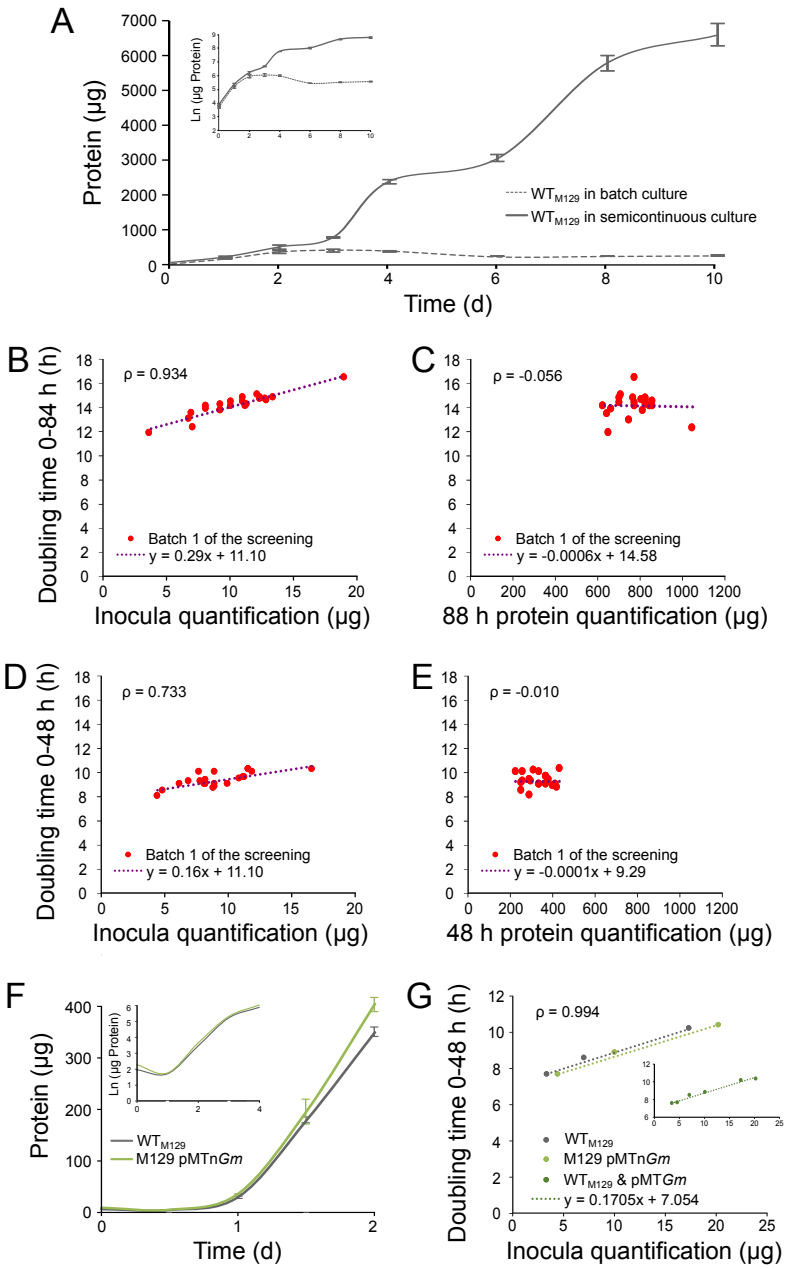


Figure 2.2: Developing a new method to measure growth curves (GC) based on intracellular protein concentrations in semicontinuous cultures. WT$_{M129}$ and M129 pMTn*Gm* are indicated in dark gray and green, respectively. The insets show the GC in the natural logarithm of protein amounts. $n = 4$ (two independent experiments with two technical replicates each). Error bars = SD. **(A)** Time courses (in days) of WT$_{M129}$ GC in batch and semicontinuous (changing medium every 12 h) cultures. **(B, C)** Doubling times of batch 1, of the screening from section 4.3.1, showed high correlation with inocula **(B)** but not with protein amount at 88 h **(C)**. **(D, E)** Doubling times of batch 1 showed still high correlation with inocula **(D)** and not with protein amount at 48 h **(E)**. **(F)** Time courses of WT$_{M129}$ and M129 pMTn*Gm* GC. **(G)** Correlation between doubling times and inocula for WT$_{M129}$ and M129 pMTn*Gm*.

2.3.2.1. Intracellular protein concentrations in semicontinuous cultures

As stated above, protein concentration measurements are much more sensitive than DNA, are quick and correlate very well with viable counts (Snell, 1981; Lin et al., 2008). However, the difficulty is still the low cell yield reached in batch cultures, which prevents detecting small differences in growth rates. One possibility to increase final biomass is to change the medium periodically renewing nutrients and discarding toxic products, such as acetic and lactic acids, which arrest cell growth. To see whether this was the case, WT_{M129} cells were grown for 10 days and the medium was changed every 12 h by fresh and pre-warm medium (thereafter called semicontinuous culture), in parallel for the control condition, WT_{M129} cells were grown in batch culture (without medium change). See section 2.2.3.1 for a detailed protocol to set up this experiment. Figure 2.2.A. shows that this is the case and a significant increase in cell yield in the semicontinuous compared to the batch culture.

In a semicontinuous culture of about 3.5 days (88 h) (Figure 2.2.A.), the cell yield is big enough (4-fold higher in semicontinuous than in batch culture), to probably differentiate close growth rates between different mutants. To test whether a semicontinuous culture could obtain large protein amounts to significantly differentiate doubling times, mutants from the batch 1 of the screening from the section 4.3.1, including WT_{M129} and M129 pMTn*Gm*, were grown in semicontinuous cultures and doubling times were calculated with two time points (0 and 88 h) with the equations 2.1 and 2.2. Interestingly, calculated doubling times (Table 2.1) highly correlated with inocula (Pearson correlation $\rho = 0.934$) (Figure 2.2.B.) but not with the protein amount at 88 h (Figure 2.2.C.). This correlation suggests that at 88 h the cells were not longer growing exponentially but rather they were in stationary growth phase.

To avoid that cells enter in stationary growth phase, the same experiment was repeated for a shorter time period (48 h). Because of this change, the correlation of the doubling time with inocula was lower (Pearson correlation $\rho = 0.733$) (Figure 2.2.D.) and still non-existent with the final protein amounts at 48 h (Fig-

ure 2.2.E.). To correct for this bias due to the variability in inocula, WT_{M129} and M129 pMTn*Gm* were grown in semicontinuous cultures for 48 h changing the inocula amount. Thus, the output of this experiment resulted in linear relationships between inocula amounts and doubling times (Figure 2.2.G.). The fitting equation 2.5 to the regression line of both WT_{M129} and M129 pMTn*Gm* was used to correct the doubling times by setting the inocula to 10 µg and increasing or decreasing the doubling time according to the slope of such line. Consequently, by correcting the doubling times, their standard deviation decreased. Effectively, this method showed that M129 *Δldh* grows significantly slower than its negative control transformants (*P*-value = 1.789×10^{-20}), their corrected doubling times were 10.76 ± 0.17 h and 9.47 ± 0.15 h, respectively (Table 2.1).

$$y = 0.1705x + 7.054 \tag{2.5}$$

Since this method is neither high-throughput nor automatic, it could not be used for screening but it could be used for validation of those mutants found to grow faster in the high-throughput screening in section 4.3.1.

2.3.2.2. Optimization of an automated high-throughput microcolony growth assay by time-lapse microscopy

The automated microcolony growth assay was developed for monitoring yeast growth rates of tens of thousands of microcolonies simultaneously (Levy et al., 2012; Ziv et al., 2013; Zackrisson et al., 2016). This technique could allow us to measure growth rates of single microcolonies from a population in a high-throughput and high-content manner. I attempted to adapt this technique, as described in section 2.2.4, to *M. pneumoniae* (Figure 2.3) and with the potential of using it for all adherent *Mollicutes* species. Several parameters were optimized (Figure 2.4.A.), *i.e. i)* surface coatings, *ii)* different fold dilutions of seed cultures, and *iii)* different times to allow the cells to attach to the glass surface.

With this method, at least ∼5000 microcolonies can be analyzed in a single run. Figure 2.4.B-D. and Table 2.1 show the calculated doubling times for WT_{M129}

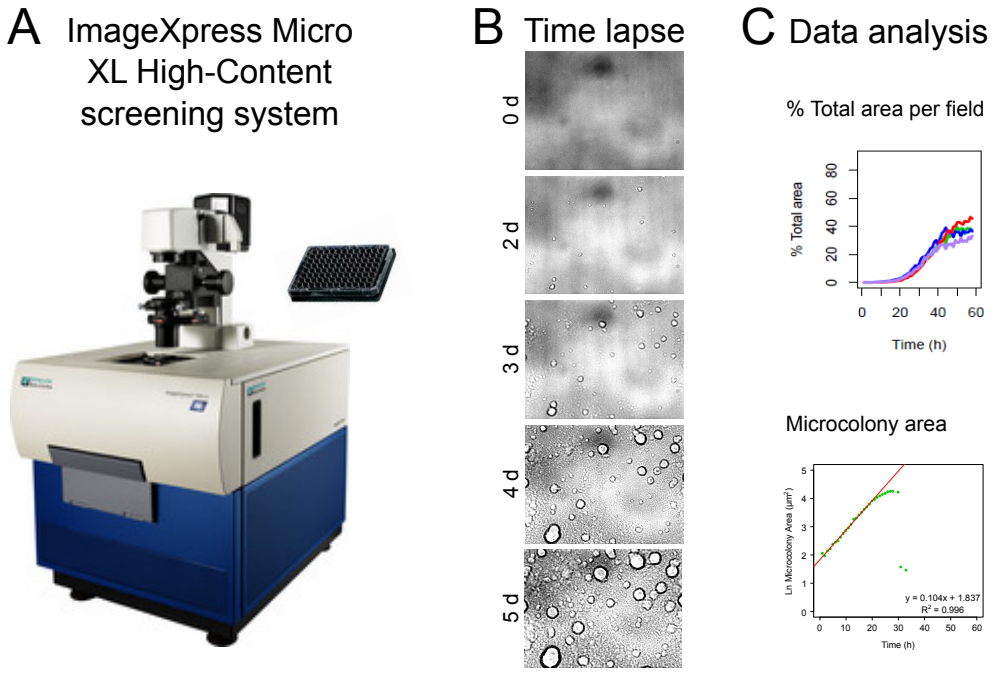


Figure 2.3: Developing a new method to measure growth rates based on an automated microcolony growth assay by high-throughput and high-content time-lapse microscopy. **(A)** Time-lapse images were acquired in an ImageXpress Micro XL High-Content screening microscope. **(B)** Time-lapse imaging series at 0, 2, 3, 4, and 5 days of incubation in the ImageXpress. **(C)** Data may be analyzed by either percentage of growth area per field or growth per microcolony area.

and the negative control transformants, M129 pMTn*TetM438*, with different well coatings and different dilutions. The results show, most of the time, that there is much more variability of M129 pMTn*TetM438* doubling times than that of WT_{M129}, which is expected since WT_{M129} is an isogenic line, whereas M129 pMTn*TetM438* is a population of cells carrying chromosomal insertions at random positions, which affects cell fitness by the gene disruption caused by the transposon insertion. Regarding the coating, cells did not grow in wells coated with 0.1 mg/mL poly-L-Lysine, but they grew in all other coatings and also without it. Interestingly, coating with concanavalin showed the lowest doubling times variance in all dilutions and in both WT_{M129} and M129 pMTn*TetM438*, whereas coating with 3% gelatin showed the highest doubling times variance. Unfortunately, there is not a clear tendency that would suggest which of the dilutions is the best one to use because the results are inconsistant. This method would be

A

Variables	Conditions	Growth
Surface coatings	No coating	+
	Poly-L-lysine	-
	Gelatin 3%	+
	Concanavalin A	+
Dilutions	1:10	++++
	1:100	+++
	1:1000	++
	1:10000	+
Cell attachment time	2 h	+
	6, 10 h	++
	16, 22 h	+++

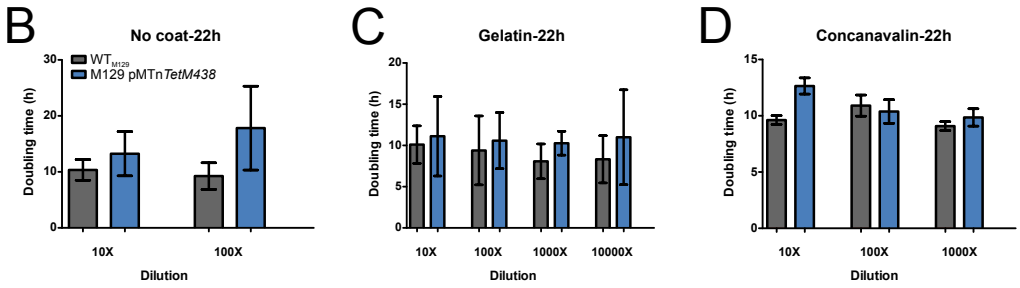


Figure 2.4: Optimization conditions of the automated microcolony growth assay. WT_{M129} and pMTn*TetM438* are indicated in dark gray and blue, respectively. $n = 2$ (two independent experiments). Error bars = SD. **(A)** Tested conditions to optimize the assay and growth outcomes (-, indicates no growth, and +, indicates the degree of growth). **(B, C, D)** The calculated doubling times for WT_{M129} and pMTn*TetM438* are plotted, after 22 h of cell attachment, in wells **(B)** with no coating; **(C)** coated with 3% gelatin; and **(D)** coated with 200 µg/mL Concanavalin A Type IV. Cells did not grow in wells coated with 0.1 mg/mL poly-L-Lysine.

of great advantage for the mycoplasmology field but it still needs more optimization to get more reproducible data with less variance. The main problem of this methodology is that it requires pure isolated transposon mutants, cells having the insertion at the same chromosomal position, to reduce growth variability, which would imply additional previous fastidious, intensive and time-consuming steps in the workflow and for a screening of several mutants would is not trivial.

2.4. Discussion

Overall, the comparison of growth curves of M129 *Δldh* and its negative control M129 pMTn*Gm*, with different methods, showed that certain growth param-

eters may mislead the conclusions on growth rates. pMTn*Gm* inserts randomly and in some cases affect fitness genes, resulting in an apparent lower growth rate (Lluch-Senar et al., 2015b). According to intracellular ATP concentrations and medium color (pH) change (Figure 2.1), it seemed that M129 *Δldh* grew much slower than its control, however with intracellular protein and extracellular glucose concentrations there was only a small difference in growth, making M129 *Δldh* slightly slower, but was not significant. Therefore, for M129 *Δldh* mutant, ATP concentrations did not correlate with protein concentrations. The misled difference observed in growth, with ATP concentrations and medium color (pH) change, was caused by the different metabolic activity that has M129 *Δldh*, suggesting that growth curves of mutants of metabolic genes or whose metabolism is altered should not be performed with intracellular ATP concentrations or medium color (pH) change but rather with protein concentrations or viable counts.

Similar results of ATP growth curves were obtained in studies in *Mollicutes* species, after an exponential ATP increase corresponding to the active growth phase, ATP decrease sharply during the stationary phase, while viable cell counts either CFU or CCU remained high (Saglio et al., 1979; Beaman and Pollack, 1981, 1983; Stemler et al., 1987), in agreement with intracellular protein concentrations (Figure 2.1). This dramatic ATP fell, which occurs as soon as the metabolic activity slows down, allows us to determine accurately the time when exponential phase ends (Saglio et al., 1979). Interestingly, the results suggest that M129 *Δldh* has a reduced energy charge, since the ATP concentrations increased proportionally to the intracellular protein concentration in the M129 pMTn*Gm* but not in M129 *Δldh*. The cause of this presumably low energy charge is unknown.

Both medium color (pH) change ($A_{430/560}$) and the colorimetric growth assay by serial dilutions (A_{550}) (Figure 2.1) are highly sensitive methods. The results of the colorimetric growth assay showed that the higher the inocula concentration, the poorer the reproducibility, probably due to variation in clumped cells and therefore inocula size. This low reproducibility was also observed previously by Snell (1981). To overcome this low reproducibility and obtain suspensions of singles cells, it could be necessary to sonicate from 2.5 to 10 min (Kim et al., 1966; Fur-

ness et al., 1968a,b) and/or serial filter the broth culture through 0.45 and 0.22 μm pore diameter filters (Karr et al., 2012). Notably, the 1-fold dilution growth curve of the colorimetric growth assay of the M129 pMTn*Gm* (Figure 2.1.G.) did not show the lag phase observed in the more diluted curves, instead it increased steeply. Similarly, the results of the serial dilutions growth curves showed that the higher the dilution factor of the inocula, the less steep the decrease in A_{550}, which correlates with a mild decrease in pH.

Altogether, the results of medium color (pH) change, extracellular glucose, acetate and lactate concentrations (Figure 2.1.C-F.) showed that during the first 2 days of growth, the production of acetic acid and hence the increase in medium acidification was buffered by HEPES, which is part of the reagents constituting Hayflick medium. After 2 days of growth, the accumulation of acetic acid and the beginning of lactic acid producion in M129 pMTn*Gm*, produced a rapid decrease in pH, as HEPES could not longer keep it constant. This time coincides with the beginning of the stationary phase in the ATP growth curve of M129 pMTn*Gm* (inset of Figure 2.1.B.) and probably also with the beginning of the late exponential phase in its protein growth curve (inset of Figure 2.1.A.). This make us conclude that in batch cultures, *M. pneumoniae* growth arrest seems to be caused by pH decrease (Shepard and Lunceford, 1965; Maier et al., 2011; Wodke et al., 2013) rather than nutrient depletion.

In order to overcome growth arrest after 2 days of growth in batch cultures, I attempted to eliminate the acid products and also renew nutrients by changing the medium every 12 h (Figure 2.2.A.). Cells in this semicontinuous culture turned out to increase 42-fold the initial biomass of M129 pMTn*Gm* (Figure 2.2.F.), while this increase in batch culture was only of 12-fold (Figure 2.1.A.). Nonetheless, that increase would have been much higher if the growth curves in semicontinuous cultures had not presented the observed lag phase compared to those in batch cultures. This lag phase is presumably caused by the use of frozen precultures as inoculum, which is done in order to quantify and normalize inocula, while for the batch culture experiments, the inocula came from fresh exponential phase precultures.

An alternative high-throughput method to the colorimetric growth assay was developed based on colony size measurements by an automated microscope capable of autofocusing and more importantly measuring 9 fields per well of a 96-well plate in less than 2 hours for a time course of 5 days (Figure 2.3). Colony size has been already used as a growth indicator in several studies on *Mollicutes* species (Smith, 1956; Shepard and Lunceford, 1965; Razin et al., 1977; Leland et al., 1982; Kihara et al., 1983; Dybvig et al., 1989; Pich et al., 2006; Hegde et al., 2015b; Hutchison et al., 2016). Colony diameter has been shown for several *Mollicutes* species to be inversely proportional to the number of colonies per area, thus, colony size is a valid parameter only when colonies are measured at the same density (Kihara et al., 1983). Then, this might explain the variability in doubling times found for the 10-fold serial dilutions (Figure 2.4.B-D.) and provide a corresponding tendency, observed in WT_{M129} cells, in which the doubling time decreases as the dilution factor increases. This, however, is difficult to see in M129 pMTn*TetM438*, due to its higher variance for being a population with randomly disrupted genes by transposon insertions, which alter cell fitness and hence single cell doubling times. Moreover, the highest variance of doubling times observed in wells coated with 3% gelatin, which is used to increase medium viscosity and hence reduce cell motility and maintain cell contact with the glass (Radestock and Bredt, 1977), might be due to the still remained cell motility. Therefore, in order to immobilize cells, even higher concentrations than 5% gelatin would be needed (Radestock and Bredt, 1977). Unlike gelatin, concanavalin A-coated wells showed the lowest variance of doubling times. Concanavalin A is a lectin that binds specifically to carbohydrate residues, such as mannose, glucose, or their glycosides on several mycoplasma cell membranes (Schiefer et al., 1975, 1978), agglutinating cells and immobilizing them. As a consequence, this probably improved the reproducibility of measuring microcolony areas resulting in very close determinations of doubling times.

Accurate and efficient methods to estimate growth rates in *Mollicutes* species has being, for long time, a topic of research and a bone of contention due to the inherent difficulty of working with these microorganisms. Despite the variability

of methods, there is not a perfect one and often it is necessary to use many of them to understand the growth physiology of the bacterium. Current technological advances may also improve and develop new methods to determine growth rates of *Mollicutes* species. Recently, single cell growth rates have been measured in microfluidic devices, with which a variety of additional growth parameters can be monitored, like, cell size (Wang et al., 2010; Campos et al., 2014; Taheri-Araghi et al., 2015; Yuan et al., 2017). This technology has obviously started in the model organisms *E. coli* and *B. subtilis* and although it is still not used among mycoplasmologist, a microdevice has been built to study the motility of *M. mobile* in microtracks (Hiratsuka et al., 2005).

Chapter 3

RATIONAL DESIGN OF FAST-GROWING MUTANTS OF *M. PNEUMONIAE*

3.1. Introduction

Recent data from 102 bacteria have showed correlations between the rRNA and tRNA gene content and growth rates. (Rocha, 2004; Vieira-Silva and Rocha, 2010). These correlations have also been found with ribosomal (r-) protein abundances (Rocha, 2004; Couturier and Rocha, 2006; Vieira-Silva and Rocha, 2010). As a consequence of genome reduction, *Mollicutes* species have only one or two rRNA operons in their genomes (except *Mesoplasma lactuacae*, which has three) (Amikam et al., 1984; Bové, 1993), whereas the model organisms *Escherichia coli* and *Bacillus subtilis* have seven and ten, respectively. Some studies have decrease the growth rates of *E. coli* and *B. subtilis* by gradually deleting 6 of its 7 and 9 of its 10 rRNA operons, respectively (Asai et al., 1999b,a; Yano et al., 2013).

Likewise, *Mollicutes* have fewer tRNA genes. For instance, *M. pneumoniae* has 34 tRNA species with different anticodons that are coded by 37 tRNA genes, whereas *B. subtilis* has 33 tRNAs encoded by 86 genes, and *E. coli* has 45 tRNAs encoded by 89 genes. Despite the small number of duplicated tRNA genes in

the *Mycoplasmas* (Bové, 1993), *M. pneumoniae* is the *Mollicutes* species with the largest number of tRNA genes (Sirand-Pugnet et al., 2007). The study by Rocha (2004) foud that fast-growing bacteria have more tRNA gene copy numbers (a median of 61 tRNAs) than slow growers (a median of 44 tRNAs), but fewer anticodons (34 and 39, respectively). Thus, fast growers have more abundant but less diverse tRNAs.

Due to the limited biosynthetic capabilities of *Mollicutes* species, these bacteria rely on uptake of external metabolites for survival and growth. Uptake of of several inorganic and organic substrates is mediated by ATP binding cassette (ABC) transporters (Wium et al., 2015; Staats et al., 2007). The peptide/opine/nickel uptake ABC transporter family, is involved in oligopeptide uptake. The oligopeptide transporter genes, from various *Mollicutes* species were compared and it was found that fast-growing species contained two copies of the *oppABCDF* operon (Staats et al., 2007), whereas slow-growing species, like *M. genitalium* had only one. The Opp transport permeases consists of five components, OppA, OppB, OppC, OppD and OppF. OppB and OppC create the transmembrane pore through which oligopeptides are imported. OppD and OppF link transport to the hydrolisis of ATP. OppA binds to the peptide and delivers it to the OppBCDF complex (Solomon et al., 2003).

In batch culture growth, *M. pneumoniae* undergoes a metabolic shift from acetic acid fermentation to lactic acid fermentation (Figure 2.1) (Wodke, 2012). This metabolic shift occurs in order to maintain a redox balance by the lactate dehydrogenase (Ldh), which catalyzes the conversion of piruvate into lactic acid with the oxidation of NADH to NAD^+. The oxidation of NADH by NADH oxidase (Nox), under aerobic conditions, could replaced the role of the Ldh in the regeneration of NAD^+ (Guo et al., 2012) to avoid medium acidification, which arrest cell cycle progression and growth. The overexpression of *noxE* gene in *Lactococcus lactis*, decreased the lactic acid yield from 21.15 ± 0.08 mM to 9.94 ± 0.07 mM, the intracellular NADH/NAD^+ ratios varied from 0.711 ± 0.005 to 0.383 ± 0.003 and the cell viability was increased (Guo et al., 2012).

In this Chapter, we aimed to rationally design fast-growing *M. pneumoniae* mutants by taking a hypothesis-driven approach, in which we reviewed the scientific literature on bacterial growth physiology and particularly looked at the genetic differences between fast- and slow-growing bacteria. Accordingly, I then generated mutants of *M. pneumoniae* by genetic modifications, performed their growth curves by using several methods to estimate growth rates and investigated more deeply the changes in expression of those genetic modifications.

3.2. Materials and methods

3.2.1. Bacterial strains and culture conditions

One Shot®TOP10 Chemically Competent *E. coli* (Invitrogen, Cat. No. C4040) or NEB 5-alpha Competent *E. coli* (New England Biolabs, Cat. No. C2987H) was used for plasmid amplification and cloning. ElectroMAX™ Stbl4™ Cells (Invitrogen, Cat. No. 11635-018) were used for pMT*Tc-rrn*$_{M129}$ cloning. They were grown in Luria-Bertani (LB) broth or LB agar plates containing 100 μg/mL ampicillin at 37°C, except Stbl4 cells, which were grown at 30°C.

Wild type *M. pneumoniae* strain (str.) M129 (WT_{M129}) (ATTC 29342, subtype 1, broth passage no. 35) (Regula et al., 2000), wild type *M. pneumoniae* str. FH (WT_{FH}) (ATTC 15531, subtype 2, broth passage no. 3) and wild type *M. genitalium* str. G37 (WT_{G37}) (ATCC 33530, broth passage no.) (Tully et al., 1983) were used. M129 *ΔglpD* mutant of *M. pneumoniae* str. M129 was gently provided by Jörg Stülke and obtained by haystack mutagenesis (Halbedel and Stülke, 2007). M129 *ΔglpD* is a pure mutant whose glycerol-3-phosphate dehydrogenase *glpD* (*mpn051*) gene has been disrupted by the mini-transposon pMT85 (Zimmerman and Herrmann, 2005). *M. pneumoniae* cells were grown in modified Hayflick's medium (hereafter referred to as Hayflick) as previously described (Hayflick, 1965; Yus et al., 2009). *M. genitalium* cells were grown in SP-4 medium (Tully et al., 1979). Static cell cultures were grown at 37°C in a humidified incubator gassed with 5% carbon dioxide (CO_2).

3.2.2. DNA manipulations

Genomic DNA of WT_{M129}, WT_{FH} and WT_{G37} was isolated by using the illustra bacteria genomicPrep Mini Spin Kit (GE Healthcare, 28-9042-59). Plasmidic DNA was obtained by using the QIAprep®Spin Miniprep kit (QIAGEN, Cat. No. 27106). DNA fragments were obtained by PCR amplification with the Phusion®High-Fidelity Polymerase (New England Biolabs, M0530). The purification of PCR products and digested fragments from agarose gels were performed by using the QIAquick®PCR Purification Kit (QIAGEN, Cat. No. 28106) and QIAquick®Gel Extraction Kit (QIAGEN, Cat. No. 28706), respectively. DNA was visualized by agarose gel electrophoresis with GelRed (Biotium, Cat. No. 41003) as stain. All oligonucleotides were ordered from Sigma-Aldrich purified by reverse phase. DNA Sequences were verified by Sanger sequencing with GATC Biotech. All primers used for molecular cloning and Sanger sequencing are listed in Tables 3.1 and 3.2.

3.2.2.1. Molecular cloning

All plasmids were constructed by using the Gibson assembly method (Gibson et al., 2009), unless otherwise specified, the master mix was provided by the Biomolecular Screening Protein Technologies CRG facility. All plasmids used in this study are listed in Table 3.3.

pMT*CmLox* Chloramphenicol resistance gene (cat, chloramphenicol acetyltransferase) was amplified by PCR from pMTn*Cat* using primers 5MCSlox66Cm and 3lox71Cm, which contain EcoRV and XbaI restriction sites at their 5' ends, respectively. Moreover, primer 5MCSlox66Cm contained several restriction sites and the 34-bp lox66 site, and primer 3lox71Cm contained the 34-bp lox71 site. The resulting 805 bp PCR fragment was digested with EcoRV and XbaI and cloned into EcoRV-XbaI-digested pMTn*Cat*, creating the pMT*CmLox* plasmid.

pMT*Tc-rrn$_{M129}$* Plasmid pMTn*TetM438* was amplified by PCR using primers Fw_Tn/Tc and Rv_Tn/Tc. The resulting 6.2 kb PCR-generated vector was digested with DnpI and purified. The rRNA operon (*mpnr01, mpnr02* and *mpnr03*)

Table 3.1: Sequence of primers used in this study for molecular cloning. Bold indicates overlapping regions between fragments and underlining indicates gene-specific sequences, unless otherwise noted.

Primer Name	Sequence (5'-3')
Cloning using restriction endonucleases	
pMT*CmLox*	
5MCSlox66Cm[a]	TAAGCTTGATATCG**CTGCAGCCCGGGGGATCCA***TACCGTTCGTATAATG TATGCTATACGAAGTTAT*AATTGTAGTATTTAGA
3lox71Cm[a]	AAAGCGGCCGC**TCTAGA***TACCGTTCGTATAGCATACATTATACGAAGTTAT*TTACGCCCCGCCCTG
Cloning by Gibson Assembly	
pMT*Tc-rrn$_{M129}$*	
F_Tn/Tc	**AGCTTGATATCGAATTCTAG**TATTTAG
R_Tn/Tc	**TATCGATACCGTCGACCTC**
F_rRNA:rRNA	**CGAGGTCGACGGTATCGATA**TAGAAAAATGCACTTTTTTCC
R_rRNA	**CTAGAATTCGATATCAAGCT**GCAAAGTAAAAACAGCTATAAAATG
pMT*CmLox-rrn$_{FH}$*	
pMTCmLox_F	**GCTTGATATCGCTGCAGC**
pMTCmLox_R	**AGAATAAAACTTTTTCAAAC**TGTTGATTTATCGATACCGTCGACC
rrnFH_F	**GTTTGAAAAAGTTTTATTCT**GAACCTATTAAATGTTAATAAAACTTATG TAAGTATCACGAACACTC
rrnFH_R	**GGGCTGCAGCGATATCAAGC**GCAAAGTAAAAACAGCTATAAAATG
pMT*Cm-OppA*	
TnTc_F[b]	***TAAAAGAGACTTAGGCCTAT****GCCTAGGTCTTTTTTATTGTTTTG*CGATAAGC TTGATATCGAATTC
TnCm_R	**ATACCGTCGACCTCGAGG**
pMpn456_F	**CCTCGAGGTCGACGGTAT**AATTTTTAAAACCTGAGTTTTAATTG
pMpn456_R	**CCGAGCAACAAGTATCTTTTTTTTAATTTCATGGATCGTTC**ATAAATG TTG
OppA_F	G**AACGATCCATGAAATTAAAAAAAAGATACTTGTTGCTCGG**
OppA_R[b]	***ATAGGCCTAAGTCTCTTTTATCTTCTTGAC***TTACTCATCAATGTAACCTGA ATCACG
pMT*Pac-OppBCDF*	
TnPuro2_F[b]	***GACTTAGGCCTATGCCTAGGTC****TTTTTTATTGTTTTG*TCGATAAGCTTGATA TCGAATTCCTG
TnPuro2_R	**TTGAACACTTTCGAATAC**TAATAAAAAAGAAAAAGGAACCTGGTTCT CAGTACCGTCGACCTCGAGGG
OppBCDF2_F	**GTATTCGAAAGTGTTCAA**GTATATTCTTAAACGTTTGGGTTTGGCAGT GTTAGCGATGTTCAAGTATATTCTTAAACGATTAG
OppBCDF2_R[b]	***GACCTAGGCATAGGCCTAAGTC****TCTTTTATCTTCTTGAC*TTAATTTATCGG CGTAGCC

[a] Bold indicates restriction sites at the 5' end of the primers. Italic indicates the lox sites.
[b] Italic indicates the 54-bp terminator of the rRNA operon of WT_{M129}.

Table 3.2: Sequence of primers used in this study to verify DNA sequences.

Primer Name	Sequence (5'-3')
Tnp_F	CGCATCTTCCCCAATCAAACAT
TnpS_F	TCAAACATCAGCCAATCG
TnS_R	TGTGTGGAATTGTGAGCG
TetR_R	GAAATCAGTAGAATTGCCCCA
16Srrn1S_F	TGGAATTTCATGTGGAGC
16Srrn2S_F	TTCTCGGGTCTTGTACAC
23Srrn1S_F	AAACATCTCAGTAGCCAC
23Srrn2S_F	GATAGCGGATGACTTGTG
23Srrn3S_F	GATGGAGTGACAAAGAAG
CmR_R	ACTGACTGAAATGCCTCAAA
OppA1_F	TTGAGCTGGGTTGATAATGC
OppA2_F	AATGCAAACATCACTGAGAC
OppBC_F	GAGTCTTTGTTTGGAGTACA
OppBC2_F	GAAATGAGGTGGTGGAAGTG
OppC2_F	AGCACAAAATCTACCACCAC
OppCD_F	TGAATTGCAGTAGTAGTAGC
OppCD2_F	AGAGTACATCCAAGCAGCTA
OppD_F	AGAGGGATTGATCTTAAGGT
OppDF_F	AAACCCAACAATGACGATAG
OppDF2_F	CAAGCCCCTAAAGTTGAACC
OppF1_F	TGGGGTTATTGGTGAGAGTG
OppF2_F	AGTTGAAACATTGCTTGGTG
OppF21_F	GCGAATTCCTCTTTAACACC
OppF22_F	AAAAGTGAGGCAAACAAAGC
PuroR_R	GGTGCTTCACTGTTTTCTTG

and its 147 bp-promoter region upstream were amplified from genomic DNA of WT_{M129} by PCR using primers Fw_rRNA:rRNA and Rv_rRNA. The resulting 5.1 kb PCR-generated insert and the 6.2 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*Tc-rrn*$_{M129}$.

pMT*CmLox-rrn*$_{FH}$ Plasmid pMT*CmLox* was amplified by PCR using primers pMTCmLox_F and pMTCmLox_R. The resulting 5.0 kb PCR-generated vector was digested with DnpI and purified. The rRNA operon (*mpne0108-rrs*, *mpne0109-rrl* and *mpne0110-rrf*) and the 233 bp-promoter region upstream were amplified

Table 3.3: Plasmids used in this study.

Plasmid Name	Description	Reference
pMTn*TetM438*	pMTn*4001* containing the tetracycline (*tetM*) resistance coding region under the 22 bp-promoter located upstream of the *mg438* translational start codon.	Pich et al. (2006)
pMTn*Gm*	pMTn*4001* containing the *aac(6')-aph(2")* marker, which confers gentamicin resistance.	Pich et al. (2006)
pMTn*Cat*	It is a derivative of pMTn*TetM438*, in which the EcoRI-BamHI *tetM438* selectable marker was replaced by an EcoRI-BamHI chloramphenicol (*cat*) resistance cassette.	Burgos and Totten (2014)
p*GmRCre*	It is a suicide vector, derivative of pBSK that contains the *aac(6')-aph(2")* gene, which confers gentamicin resistance, and *Cre* recombinase gene. Both genes are under the 22 bp-promoter located upstream of the *mg438* translational start codon.	Mariscal et al. (2016)
miniTn*4001-Puro-1*	pMTn*4001* containing the puromycin (*pac*) resistance gene under the 168 bp-Spiralin promoter.	Paetzold et al. (2013)
pMT*Tc-ncrrn357*	pMTn*TetM438* containing the ncRNA357 gene under the 22 bp-promoter located upstream of the *mg438* translational start codon.	Lluch-Senar et al. (2015b)
pMT*Pac-438gRNAeNT2*	miniTn*4001-Puro-1* containing a protospacer against *venus* gene (eNT2) followed by a scaffold and a terminator sequence derived from *Staphylococcus pyogenes*, and under the 22 bp-promoter located upstream of the *mg438* translational start codon.	(Piñero and Broto, in preparation)
pMT*Tc-rrn$_{M129}$*	pMTn*TetM438* containing the rRNA operon and its 147 bp-promoter region of WT$_{M129}$.	This study
pMT*CmLox*	It is a derivative of pMTn*Cat*, in which the chloramphenicol (*cat*) resistance cassette is flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively.	This study
pMT*CmLox-rrn$_{FH}$*	pMT*CmLox* containing the rRNA operon of WT$_{FH}$ and the 233 bp-promoter region of WT$_{M129}$.	This study
pMT*Cm-OppA*	pMTn*Cat* containing the *oppA* (*mg321*) gene from WT$_{G37}$ under the control of the 170 bp-promoter region of its orthologous gene from WT$_{M129}$ (*mpn456*) and the downstream 54-bp terminator of the rRNA operon of WT$_{M129}$.	This study
pMT*Pac-OppBCDF*	miniTn*4001-Puro-1* containing the *oppBCDF* (*mg077, mg078, mg079* and *mg080*) operon from WT$_{G37}$ under the control of the 87 bp-promoter region of its orthologous operon from WT$_{M129}$ (*mpn215, mpn216, mpn217* and *mpn218*) and downstream the 54-bp terminator of the rRNA operon of WT$_{M129}$.	This study
pMT*Pac-37tRNAs*	miniTn*4001-Puro-1* containing the 37 tRNA genes from WT$_{M129}$ under the control of their own promoters and conserving the operon organization.	This study
pMT*Pac-noxE*	miniTn*4001-Puro-1* containing the *noxE* (*llmg0408*) gene from *Lactococcus lactis* MG1363 under the control of the 179 bp-pmp200 promoter and 45 bp of the *pmp200* coding region.	This study
pMT*CmLox-gpsA*	pMT*CmLox* containing the *gpsA* (*mype2300*) gene from *Mycoplasma penetrans* HF2 under the control of the 105 bp-*glpd* (*mpn051*) promoter and downstream the 54-bp terminator of the rRNA operon of WT$_{M129}$.	This study

from genomic DNA of WT_{FH} by PCR using primers rrnFH_F and rrnFH_R. The resulting 5.1 kb PCR-generated insert and the 5.0 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*CmLox-rrn$_{FH}$*.

pMT*Cm-OppA* Plasmid pMTn*Cat* was amplified by PCR using primers TnTc_F and TnCm_R. The resulting 5.0 kb PCR-generated vector was digested with DnpI and purified. The 170 bp-promoter region of *oppA* (*mpn456*) gene was amplified from genomic DNA of WT_{M129} by PCR using primers pMpn456_F and pMpn456_R. The *oppA* (*mg321*) gene was amplified from genomic DNA of WT_{G37} by PCR using primers OppA_F and OppA_R. Moreover, primers OppA-R and TnTc-F contained the 54-bp terminator of the rRNA operon of WT_{M129}. The resulting 220 bp and 2.8 kb PCR-generated fragments, and the 5.0 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*Cm-OppA*.

pMT*Pac-OppBCDF* Plasmid miniTn*4001-Puro-1* was amplified by PCR using primers TnPuro2_F and TnPuro2_R. The resulting 5.4 kb PCR-generated vector was digested with DnpI and purified. The *oppBCDF* (*mg077, mg078, mg079* and *mg080*) operon was amplified from genomic DNA of WT_{G37} by PCR using primers OppBCDF2_F and OppBCDF2_R. Moreover, primers OppBCDF2_R and TnPuro2_F contained the 54-bp terminator of the rRNA operon of WT_{M129} and primers OppBCDF2_F and TnPuro2_R contained the 87 bp-promoter region of the orthologous operon from WT_{M129} (*mpn215, mpn216, mpn217* and *mpn218*). The resulting 6.2 kb PCR-generated insert, and the 5.4 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*Pac-OppBCDF*.

pMT*Pac-37tRNAs* The 37 tRNA genes from WT_{M129} with their own promoters were designed in such a way to conserve the operon organization. The 4.8 kb fragment containing the 37 tRNA genes was synthesized by GenScript with the XhoI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with XhoI and SalI and cloned into XhoI-SalI-digested miniTn*4001-Puro-1*, generating the pMT*Pac-37tRNAs* plasmid.

pMT*Pac-noxE* *noxE* (*llmg0408*) gene from *Lactoccocus lactis* str. MG1363 was placed downstream of the 179 bp-pmp200 promoter and 45 bp of the *pmp200* coding region (Zimmerman and Herrmann, 2005). This 1.6 kb fragment was synthesized by GenScript with the XhoI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with XhoI and SalI and cloned into XhoI-SalI-digested miniTn*4001-Puro-1*, generating the pMT*Pac-noxE* plasmid.

pMT*CmLox-gpsA* *gpsA* (*mype2300*) gene from *Mycoplasma penetrans* str. HF2 was placed downstream of the 105 bp-*glpd* (*mpn051*) promoter and upstream of the 54-bp terminator of the rRNA operon of WT_{M129}. This 1.2 kb fragment was synthesized by GenScript with the XhoI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with XhoI and SalI and cloned into XhoI-SalI-digested pMT*CmLox*, generating the pMT*CmLox-gpsA* plasmid.

3.2.3. Transformation of *M. pneumoniae*

WT_{M129}, WT_{FH} and WT_{G37} cells were transformed with some of the plasmids from Table 3.3 by electroporation (Hedreyda et al., 1993) as previously described in section 2.2.2, unless otherwise specified. WT_{G37} was grown in SP-4 medium instead. The transformants were selected and cultured in broth medium supplemented with the correspondent antibiotic, namely, 200 μg/mL gentamicin, 2 μg/mL tetracycline, 3.3 μg/mL puromycin, or 20 μg/mL chloramphenicol.

3.2.3.1. Increasing the rRNA operon copy number of *M. pneumoniae*

In order to increase the rRNA copy number to more than two copies of WT_{M129}, which only has one copy, four cycles of consecutive transformations were carried out with plasmids pMT*CmLox-*rrn_{FH} and p*GmRCre* or for the control transformants, with plasmids pMT*CmLox* and p*GmRCre* (Table 3.3). Briefly, a complete cycle consisted in a first trasnformation of WT_{M129} with plasmid pMT*CmLox-*rrn_{FH} and for the negative control, with pMT*CmLox*. As a result, M129 pMT*CmLox-*rrn_{FH} transformants contained two rRNA copies (*2rrn*), one endogenous ($rRNA_{M129}$) and one exogenous ($rRNA_{FH}$) randomly inserted in the cell chromosome, whereas the M129 pMT*CmLox* control transformants contained only the endogenous $rRNA_{M129}$

copy (*1rrn*). After two consecutive passages, both M129 pMT*CmLox-rrn*$_{FH}$ and M129 pMT*CmLox* resulting transformants were subjected to a second transformation with plasmid p*GmRCre*. Unlike a mini-transposon plasmid, p*GmRCre* remains in cells as long as the selection pressure is maintained, and during this time, the Cre recombinase is expressed and excises any gene flanked by lox sites, the chloramphenicol (*cat*) resistance gene in this case. Thus, the selection pressure was maintained for 10 days with 200 µg/mL of gentamicin at 37°C in the presence of 5% CO_2. Then, the medium was removed, cells were suspended in 1 mL Hayflick broth, scraped off and inoculated in a 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL Hayflick broth without any antibiotic. After two consecutive passages, the cycle was repeated three more times to obtain M129 pMT*CmLox-rrn*$_{FH}$ - p*GmRCre* transformants containing from two to five rRNA copies and the corresponding controls, M129 pMT*CmLox* - p*GmRCre* containing only the endogenous rRNA$_{M129}$ copy.

3.2.4. Growth curves conditions

Growth curves were performed as previously described in section 2.2.3 with some modification for the medium color (pH) change method described below.

3.2.4.1. Medium color (pH) change

Fresh prepared seed cultures, grown to late exponential phase (for 60 h) in 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL fresh Hayflick broth supplemented with the corresponding antibiotic, were frozen and stored at -70°C and a 100 µL aliquot thereof was twice pelleted at 14100 ×g for 10 min and washed with PBS (for 1 liter: 8 g NaCl, 0.2 g KCl, 0.24 g KH_2PO_4, 1.44 g $Na_2HPO_4{\cdot}2H_2O$, pH 7.4). After two cycles of centrifugation and washing, PBS was removed and pelleted cells were suspended in 100 µL lysis buffer (10 mM Tris·HCl, 6 mM $MgCl_2$, 1 mM EDTA, 100 mM NaCl, 0.1% Tx-100, pH 8, and 1× Protease Inhibitor Cocktail (Roche)). Cell lysates were kept on ice and spun down, pipetted up and down to complete lysis and extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). 1 µg of these quantified frozen stocks was inoculated per well and in duplicates

in a 96-well plate, excluding the peripheral wells to avoid the edge effect, containing 200 μL Hayflick broth without antibiotic per well. The 96-well plate was incubated at 37°C in the Infinite^{TM} M200 Tecan plate reader. The absorbance at 430 and 560 were automatically taken, with a settle time at 300 ms and 25 flashes, every 20 min for 5 days. Medium color (pH) change was measured by calculating the ratio between the absorbance at 430 and 560 nm ($A_{430/560}$), as previously described (Yus et al., 2009).

3.2.5. Visualization of cell size by Scanning Electron Microscopy (SEM)

Batch precultures were inoculated in 1:100 dilution from the stocks in 25 cm^2 tissue culture flasks containing 5 mL fresh Hayflick broth supplemented with 2 μg/mL tetracycline, except for WT_{M129}, to mid exponential phase (for 50 h). Exponential growth phase cultures were suspended with 700 μL fresh Hayflick broth without antibiotic, scraped off, passed through a G25 syringe needle ten times and through a Millex-GV filter of 0.22 μm pore diameter (Merckmillipore, Ref. SLGV033RS). 150 μL of the filtered suspension was used to regrown cells for 12 h on sterile glass coverslips, previously placed in 24-well plates containing 800 μL Hayflick broth without antibiotic. After 12 h of incubation at 37°C, cells were washed twice with fresh PBSand fixed with 2.5% glutaraldehyde, 0.1M PBS. Then, the samples were dehydrated, critical point dried and sputter coated with 20 nm of gold by the Electron Microscopy facility at the Universitat Autònoma de Barcelona. Samples were observed by using a FESEM Zeiss Merlin microscope.

3.2.6. Transcriptomics

Frozen cells were inoculated in 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL Hayflick broth with 2 μg/mL tetracycline, except for WT_{M129}. There were as many flasks as time points and duplicates. To extract total cellular RNA at selected times, cells were washed twice with PBS, immediately lysed in 700 μL QIAzol lysis reagent (QIAGEN, Cat. No. 79306) and scraped off. Cell lysates were stored at -70°C for further processing. Total RNA, including small RNAs,

was extracted from all samples by using the miRNeasy Mini Kit (QIAGEN, Cat. No. 217004) following the manufacturer's instructions and including the DNase treatment. RNA concentration and integrity were measured by using a Nanodrop (Thermo Scientific) and a Bioanalyzer (Agilent). Library preparation and RNA sequencing (RNA-seq) were carried out by the Genomics CRG facility. RNA libraries were prepared by using the TruSeq Stranded mRNA Library Prep Kit (Illumina). A paired-end sequencing run with a read length of 50 bases was performed in a Hi-Seq 2000 platform (Illumina).

Raw reads were mapped to the *M. pneumoniae* reference genome (NC_000912, NCBI) using MAQ (default parameters, and 1 mismatch allowed) (Li et al., 2008). Reads mapping only to a unique position of the genome were considered. Counts per gene were extracted from the pileups using our genome annotation and normalized with gene length. Expression levels were calculated using a homemade R studio-based script to determine the CPKM (Counts Per Kilobase per Million Mapped Counts) values (Lluch-Senar et al., 2015a), a measure that is similar to RPKM (Reads Per Kilobase per Million Mapped Reads) for single-end reads. The CPKM values were transformed to log2 values.

3.2.6.1. Quantification of relative expression of rRNA copies by customized RNA-seq

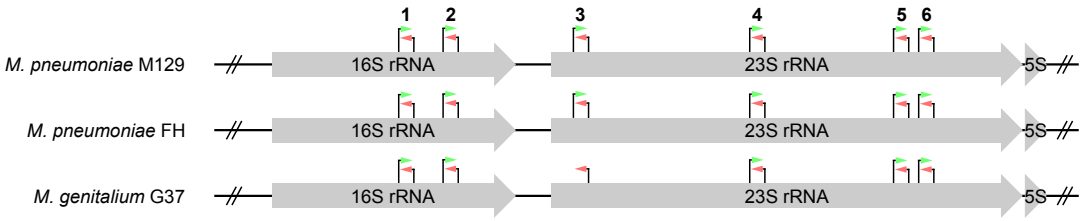


Figure 3.1: Schematic representation of rRNA regions to quantify the relative expression of rRNA copies. Primers are shown as colored arrows: Red indicate the primers used for a customized cDNA synthesis and green those used for a customized PCR. Regions are numbered as follows: (1) region amplified by primers FH837, (2) region amplified by primers FH1128, (3) region amplified by primers FH1920, (4) region amplified by primers FH3038, (5) region amplified by primers FH3851, and (6) region amplified by primers FH4029.

Table 3.4: Sequence of primers used in this study for customized cDNA synthesis and PCR amplification of rRNA regions. Bold indicates random bases and underlining indicates gene-specific sequences.

Primer Name	Sequence (5'-3')
Customized cDNA synthesis	
FH4029_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**CCACCGGATCACTATGACCT
FH3038_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**TCAATCGGTTTGGGAGATTC
FH1920_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**CACACGGAATCACATTCGTT
FH3851_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**CCAGTCAAACTGCCCACCTA
FH1128_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**TTGACGTCATCCCTTCCTTC
FH837_R	TTCAGACGTGTGCTCTTCCGATCT**NNNNNN**CCGTTTGAGTTTCATTCTTGC
1920_R12N	TTCAGACGTGTGCTCTTCCGATCT**NNNNNNNNNNNN**CACACGGAATCACATTCGTT
Specific gene primers for PCR	
FH4029_F	CCCTACACGACGCTCTTCCGATCTCGGTCGCCTCCTAAAAGGTA
FH4029_F2	CCCTACACGACGCTCTTCCGATCTACGGTCGCCTCCTAAAAGGTA
FH3038_F	CCCTACACGACGCTCTTCCGATCTCTTGTGGTAGACGAGCGTTG
FH3038_F2	CCCTACACGACGCTCTTCCGATCTACTTGTGGTAGACGAGCGTTG
FH1920_F	CCCTACACGACGCTCTTCCGATCTGAGCAATCCGGTAGTTGGAA
FH1920_F2	CCCTACACGACGCTCTTCCGATCTTGAGCAATCCGGTAGTTGGAA
FH3851_F	CCCTACACGACGCTCTTCCGATCTCCCGTGAAGCTTTACTGTAGC
FH3851_F2	CCCTACACGACGCTCTTCCGATCTTCCCGTGAAGCTTTACTGTAGC
FH1128_F	CCCTACACGACGCTCTTCCGATCTCAGCTCGTGTCGTGAGATGT
FH1128_F2	CCCTACACGACGCTCTTCCGATCTGCAGCTCGTGTCGTGAGATGT
FH837_F	CCCTACACGACGCTCTTCCGATCTTGTGGGGAGCAAATAGGATT
FH837_F2	CCCTACACGACGCTCTTCCGATCTATGTGGGGAGCAAATAGGATT
1920_F+1	CCCTACACGACGCTCTTCCGATCTHGAGCAATCCGGTAGTTGGAA
1920_F+2	CCCTACACGACGCTCTTCCGATCTHYGAGCAATCCGGTAGTTGGAA
1920_F+3	CCCTACACGACGCTCTTCCGATCTHYYGAGCAATCCGGTAGTTGGAA
1920_F+4	CCCTACACGACGCTCTTCCGATCT**NN**YTGAGCAATCCGGTAGTTGGAA
1920_F+5	CCCTACACGACGCTCTTCCGATCT**NNN**YTGAGCAATCCGGTAGTTGGAA
1920_F+6	CCCTACACGACGCTCTTCCGATCT**NNNN**YTGAGCAATCCGGTAGTTGGAA
1920_F+7	CCCTACACGACGCTCTTCCGATCT**NNNNN**YTGAGCAATCCGGTAGTTGGAA

To quantify the relative expression of endogenous and exogenous rRNA copies, we quantified the expression of rRNA by a customized cDNA synthesis, PCR amplification and sequencing of regions of rRNA having species-specific sequences (Figure 3.1). We firstly did a preliminary test of the methodology by quantifying an exogenous rRNA copy of WT_{M129} ($rRNA_{M129}$) in *M. pneumoniae* FH and *M. genitalium* G37 transformants. Briefly, WT_{FH} and WT_{G37} were transformed with pMT*Tc-rrn*$_{M129}$ or its negative control pMTn*TetM438*, as the vector containing only the antibiotic resistance gene, and pMT*Tc-ncrrn357*, as the vector express-

ing the ncRNA357 (Table 3.3). After two consecutive passages, cells were grown for 48 h, harvested and lised, and its RNA was extracted as mentioned above in section 3.2.6. Extracted RNA was used to synthesize cDNA and amplified by PCR the 6 regions shown in Figure 3.1, which was carried out by the Genomics CRG facility. All primers used for the customized cDNA synthesis and PCR amplification are listed in Table 3.4. Library preparation were validated by using a Bioanalyzer (Agilent). A paired-end sequencing run was performed in a Mi-Seq (Illumina) by the Genomics CRG facility.

PCR duplicates were removed from raw paired-end data by identifying those containing identical read 2 sequences and keeping only that with the best read quality using a custom-made C++ routine. Then, the reads containing the species-specif sequences for each rRNA region were quantified by means of bash scripting.

After the preliminary test was done, to get statistically larger identification of rRNA molecules, we increased the number of random nucleotides from six in FH1920_R to twelve in 1920_R12N (Table 3.4) and we quantified the relative expression of exogenous and endogenous rRNA copies of M129 transformants containing two, three, four and five rRNA copies, which were obtained as described above in section 3.2.3.1. Only the region (3) (Figure 3.1), which contains the insertion and deletion (InDel) that differentiates the endogenous $rRNA_{M129}$ from the exogenous $rRNA_{FH}$, was cDNA synthesized with primer 1920_R12N and PCR amplified with the mixture of specific gene primers 1920_F+1 to F+7 and the universal PCR and index primer provided in the NEBNext Singleplex or Multiplex oligos for Illumina (New England Biolabs).

3.2.7. Proteomics

Sample preparation - FASP digestion (with SDS)

Total mycoplasma protein extracts were obtained as in section 2.2.3.1, frozen pelleted cells were instead suspended in 100 µL lysis buffer (4% SDS, 0.1M HEPES). Extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). Total protein extracts of two biological replicates

were analyzed by mass spectrometry (MS) by the CRG/UPF Proteomics Unit as following. 30 μg of samples were reduced with 10 mM dithiothreitol (30 nmols, 30 min, 56°C), alkylated in the dark with 55 mM iodoacetamide (60 nmols, 30 min, 25ºC) and digested with 1 μg LysC (Wako, Cat. No. 129-02541) overnight at 37ºC and then with 1 μg of trypsin (Promega, Cat. No. V5113) for eight hours at 37ºC following Wiśniewski et al. (2009) filter-aided sample preparation (FASP) procedure. After digestion, the peptide mix was acidified with formic acid and desalted with a MicroSpin C18 column (The Nest Group, Inc) prior to LC-MS/MS analysis.

LC-MS/MS analysis

The peptide mixes were analyzed using a LTQ-Orbitrap Velos Pro mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an EasyLC (Thermo Fisher Scientific (Proxeon), Odense, Denmark). Peptides were loaded onto the 2-cm Nano Trap column with an inner diameter of 100 μm packed with C18 particles of 5 μm particle size (Thermo Fisher Scientific) and were separated by reversed-phase chromatography using a 25-cm column with an inner diameter of 75 μm, packed with 1.9 μm C18 particles (Nikkyo Technos Co., Ltd. Japan). Chromatographic gradients started at 93% buffer A (0.1% formic acid in water) and 7% buffer B (0.1% formic acid in acetonitrile) with a flow rate of 250 nL/min for 5 minutes and gradually increased 65% buffer A and 35% buffer B in 120 min. After each analysis, the column was washed for 15 min with 10% buffer A and 90% buffer B.

The mass spectrometer was operated in data-dependent acquisition (DDA) mode and full MS scans with 1 micro scans at resolution of 60,000 were used over a mass range of m/z 350-2,000 with detection in the Orbitrap. Auto gain control (AGC) was set to 1E6, dynamic exclusion (60 seconds) and charge state filtering disqualifying singly charged peptides was activated. In each cycle of DDA analysis, following each survey scan the top twenty most intense ions with multiple charged ions above a threshold ion count of 5,000 were selected for fragmentation at normalized collision energy of 35%. Fragment ion spectra produced via collision-induced dissociation (CID) were acquired in the Ion Trap, AGC was set

to 5E4, isolation window of 2.0 m/z, activation time of 0.1 ms and maximum injection time of 100 ms was used. All data were acquired with Xcalibur software (version 2.2).

Data Analysis

Proteome Discoverer software suite (version 2.0, Thermo Fisher Scientific) and the Mascot search engine (version 2.5, Matrix Science, Perkins et al. (1999)) were used for peptide identification. Samples were searched against the *Mycoplasma pneumoniae* str. M129 HomoConTrans19 database, comprising all putative *M. pneumoniae* proteins longer than 19 amino acids (after in silico translation of *M. pneumoniae* genome in the six putative frames) with a list of common contaminants and all the corresponding decoy entries (87,059 entries). Trypsin was chosen as enzyme and a maximum of three miscleavages were allowed. Carbamidomethylation on cysteines was set as a fixed modification, whereas oxidation of methionine and protein acetylation at the N-terminus were used as variable modifications. Searches were performed using a peptide tolerance of 7 ppm, a product ion tolerance of 0.5 Da. Resulting data files were filtered for false discovery rates (FDR) $< 5\%$. Protein top 3 area have been calculated with unique peptides per protein. The areas were transformed to log2 values.

3.2.8. Western immunobloting

Total mycoplasma protein extracts were obtained as in section 2.2.3.1, frozen pelleted cells were instead suspended in 100 µL lysis buffer (4% Sodium Dodecyl Sulfate (SDS), 0.1M HEPES). Extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). The protein extracts were mixed with SDS loading buffer 5X (0.5 M DTT, 35% glycerol, 0.25% bromophenol blue, 10% SDS, and 0.25 M Tris·HCl pH 6.8), were boiled at 98°C for 10 min. Next, 30 µg of protein preparations were loaded in a precast polyacrylamide NuPAGE™ 12% Bis-Tris protein gels, 1.0 mm, 10-well (Invitrogen, Cat. No. NP0341BOX) and run with NuPAGE™ MOPS SDS Running Buffer (20X) (Life technologies, Cat. No. P0001) for 1.5 h at 150V. After electrophoresis, separated proteins were transferred using the iBlot® 7-Minute Blotting System to a

nitrocellulose transfer membrane (Invitrogen, Cat. No. IB301001). After protein transfer, the membrane was blocked with BSA in TBS-Tween solution (5% BSA, 0.05% TWEEN 20 and for 1 liter: 10 mL 1M Tris·HCl pH 7.5 and 20 mL 5M NaCl) for 1 h at room temperature, followed by three washes for 5 min with TBS-Tween solution (0.05% TWEEN 20 and for 1 liter: 10 mL 1M Tris·HCl pH 7.5 and 20 mL 5M NaCl)). Then, the membrane was incubated for 1 h at room temperature with polyclonal antibodies, produced in rabbits, to detect the ribosomal proteins at a 1:2.000 dilution (Maier et al., 2011). After primary antibody incubation, the membrane was washed three times for 5 min with TBS-Tween solution and incubated for 1 h at room temperature with a anti-rabbit IgG-Peroxidase (Sigma-Aldrich, Cat. No. A045) at a 1:5.000 dilution. After secondary antibody incubation, the membrane was washed three times for 5 min with TBS-Tween solution and developed using Pierce^{TM} ECL Western Blotting Substrate (Thermo Scientific, Cat. No. 32106) according to the manufacturer's instructions. Western blot bands were quantified by densitometry using the software ImageJ.

3.2.9. Ribosomal profiling

Batch cultures were inoculated in 1:300 dilution from the stocks in 300 cm^2 tissue culture flasks containing 150 mL fresh Hayflick broth to late exponential phase (60 h). After this time of growth, chloramphenicol was added to the medium in a final concentration of 0.1 mg/mL and flasks were incubated for 1 min at room temperature. Medium was removed and cells were scraped off and passed through a G25 syringe needle ten times. 200 µL was taken for genomic DNA extraction, as described in section 3.2.2, and DNA quantification. The remaining cell suspension was centrifuged at 8000 $\times$g for 5 min at room temperature. Medium was removed and pelleted cells were resuspended in 1 mL of lysis buffer (50 mM HEPES pH7, 100 mM NaCl, 10 mM $MgCl_2$, chloramphenicol 1 mM, 0.4% Triton, 0.1 % NP40, 1:50 diluted DNAse I, 5 mM $CaCl_2$, 125 mM superase, 50 mM octyl-glucoside and 1$\times$ Protease Inhibitor Cocktail (Roche)), quick frozen in liquid nitrogen and stored at -80^{o}C. Frozen lysates were thawed and lysed by passing them twice through a prechilled cell disruptor vessel (Parr Instrument Company, Cat. No. 4639) applying 60 bars of pressure. In RNase free conditions, lysates

were centrifuged at 16100 ×g for 15 min at 4°C and the supernatant was recovered. Lysates and blank (lysis buffer) were diluted 1:100 with water and A_{260} was measured by NanoDrop to determine the nucleic acid concentration. The concentration of clarified lysates was calculated and lysates were adjusted to same DNA quantity. Linear 5-40% sucrose gradients were prepared in a gradient station. Equivalent amount (according to DNA content) of lysates were layered onto the gradients and these latter were equilibrated with lysis buffer. Total cell lysates were separated by ultracentrifugation at 217,000 ×g (35,000 r.p.m.) for 2.5 h at 4 °C in an SW40 rotor with accelerate max and decelerate no break parameters. The resulting ribosomal profiles of the gradient were fractionated from top to bottom with a density gradient fractionator and measured at 254 nm.

3.3. Results and discussion

In this chapter, we tested several hypothesis in order to decrease *M. pneumoniae* doubling time, such as (1) increasing the rRNA operon copy number from two to five, (2) duplicating the tRNA genes, (3) duplicating the *oppABCDF* genes, (4) expressing NoxE of *Lactococcus lactis*, and (5) expressing GpsA of *M. penetrans*. The results of each hypothesis are shown in the following sections.

3.3.1. Increasing the rRNA operon copy number of *M. pneumoniae* decreases its doubling time

3.3.1.1. *M. pneumoniae* with two rRNA operons

In a prelimary approach, to increase the rRNA operon copy number to two copies, WT_{M129} was transformed with plasmid pMT*Tc-rrn*$_{M129}$ or its negative controls: pMTn*TetM438* and pMT*Tc-ncrrn357*. A variety of cellular parameters, described below, of these resulting transformants were measured, such as (1) doubling times, (2) cell size, (3) transcriptomes, (4) proteomes, (5) ribosomal-protein (r-protein) abundances, and (6) ribosome abundances.

Growth curves

Growth curves were performed as in section 2.2.3 (Figure 3.2). The intracellular

Table 3.5: Estimated doubling times (DT) of *M. pneumoniae* containing two $rrn_{MpnM129}$ copies and its controls by different methods. The DT for protein and ATP concentration were calculated during the first 30 h of growth. For the color change dilutions assay, the Δt was taken when the diluted cultures reached an A_{550} = 0.4. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. For the microcolony growth assay, the shown DT corresponds to the coating with Concanavalin A and 10-fold dilution. Values indicate the mean ± SD. - indicates not measured. Statistical significant differences compared to the corresponding left/most left columns are indicated: NS: non-significant *P*-value, * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

	Type of cells		
Doubling time (h) by method	**WT_{M129}**	**M129 pMTn*TetM438* (- ctrl.-*1rrn*)**	**M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*)**
Batch cultures			
Protein concentration	-	21.4 ± 2.23	$11.8 \pm 0.61^{*}$
ATP concentration	-	6.42 ± 0.004	6.04 ± 0.53^{NS}
Medium color (pH) change dilutions	6.47 ± 0.14	7.97 ± 0.60^{NS}	$7.04 \pm 0.42^{NS/NS}$
Microcolony growth assay	9.62 ± 1.91	$12.65 \pm 2.59^{***}$	$10.66 \pm 1.66^{NS/NS}$
Semicontinuous cultures			
Protein concentration 0-48 h	8.99 ± 0.09	$9.57 \pm 0.07^{***}$	$8.97 \pm 0.08^{***/NS}$
(Corrected DT)	(9.30 ± 0.05)	$(9.35 \pm 0.07)^{*}$	$(8.96 \pm 0.04)^{***/***}$

protein growth curves (Figure 3.2.A.) showed a significant difference between the growth of M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) and its control M129 pMTn*TetM438* (-ctrl.-*1rrn*) transformants, with doubling times between 18 and 30 h of growth of 11.8 ± 0.61 h and 21.4 ± 2.23 h (*P*-value = 0.028), respectively (Table 3.5). Similarly, the final protein concentration reached by the negative control transformants of around 250 μg was reached 24 h earlier by *2rrn* transformants.

In contrast to protein growth curves, neither before nor after 30 h of growth with the intracellular ATP growth curves (Figure 3.2.B.), a significant difference was found between the doubling times of *2rrn* (6.04 ± 0.53 h) and its negative control transformants (6.42 ± 0.004 h) (*P*-value = 0.49) (Table 3.5), although the former was slightly faster.

The medium color (pH) change, measured as the ratio of absorbance at 430 and 560 nm ($A_{430/560}$), showed a significant difference from the 24 h of growth indicating a faster medium acidification of the medium, hence a faster growth in

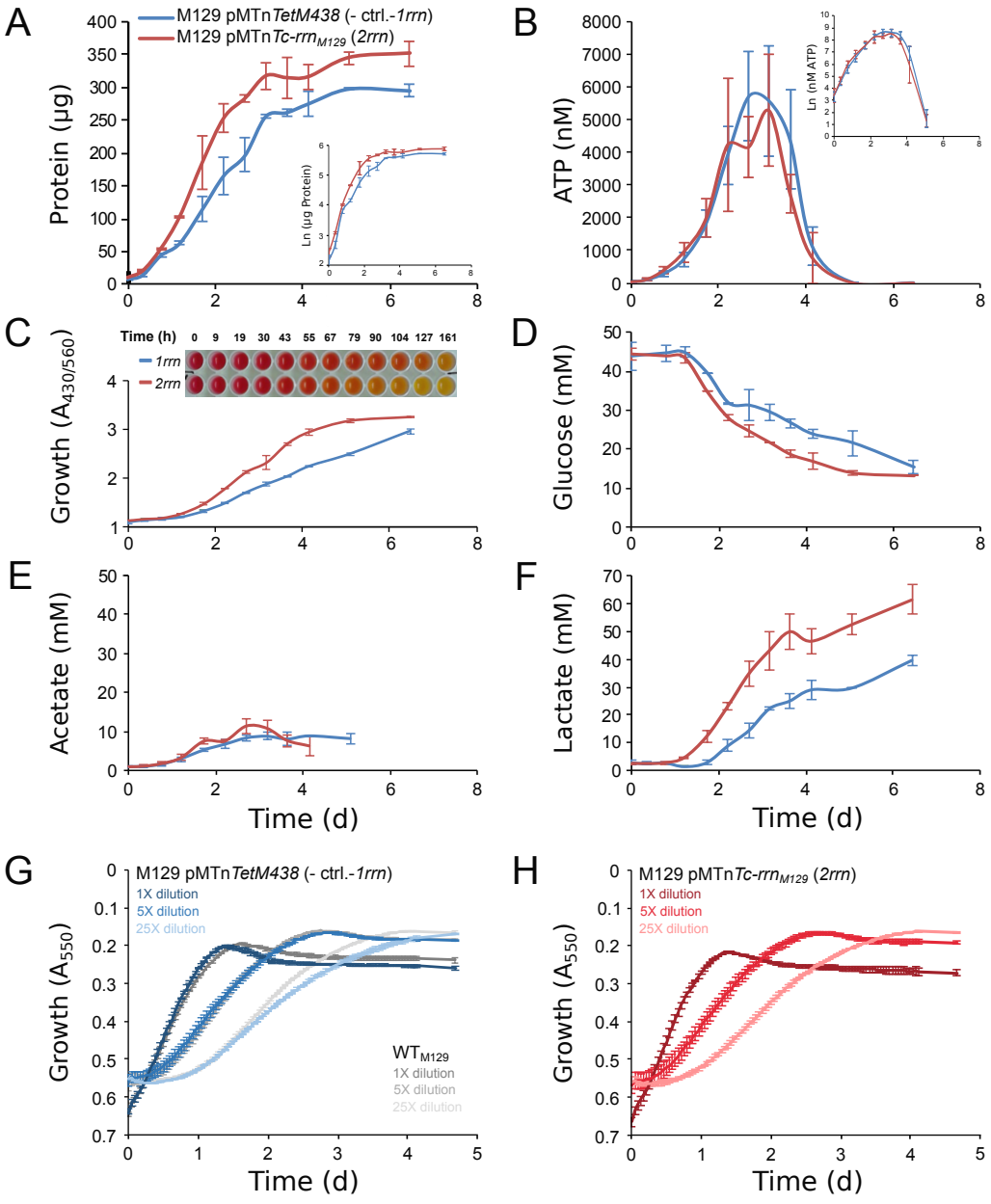


Figure 3.2: Growth curves (GC) (in days) of *M. pneumoniae* containing two $rrn_{MpnM129}$ copies and its control by different methods in batch cultures. WT_{M129} is indicated in grey, M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) in red and its control M129 pMTn*TetM438* (- ctrl.-*1rrn*) in blue. The insets show the GC in the natural logarithm of the concentrations. $n = 4$ (two independent experiments with two technical replicates each) unless otherwise specified. Error bars = SD. **(A)** Total cellular protein as a measure of biomass. **(B)** Total cellular ATP concentration. **(C)** Medium color (pH) change measured by $A_{430/560}$. **(D, E, F)** Quantification of extracellular **(D)** glucose, **(E)** acetate, and **(F)** lactate concentration. **(G, H)** Medium color (pH) change of three cultures, diluted 1-, 5-, and 25-fold, $n = 2$ (two independent experiments) measured by A_{550}, of **(G)** WT_{M129}, M129 pMTn*TetM438*, and **(H)** M129 pMT*Tc-rrn*$_{M129}$.

the *2rrn* transformants (Figure 3.2.C). The faster decrease of pH in *2rrn* transformants was the effect of higher lactate concentration in the medium than that of the negative control transformants (Figure 3.2.F) as a consequence of a significant higher consumption of glucose by the former from the first day onwards (*P*-value = 0.015) (Figure 3.2.D). In contrast, the acetic acid concentration in the medium did not show any difference (Figure 3.2.E).

The curves of the colorimetric growth assay by serial dilutions, measured at absorbance of 550 nm (A_{550}), suggested that M129 pMTn*TetM438* transformants grow slightly slower than WT_{M129} but it was not significant (*P*-value = 0.075) (Figure 3.2.G) (Table 3.5). Moreover, in consistency with the results of the former method, the curves suggested that *2rrn* (Figure 3.2.H) grow slightly faster than M129 pMTn*TetM438* transformants, although with this method, the difference was not significant (*P*-value = 0.22).

Furthermore, the methods developed in this study and described in section 2.3.2 were used in this section. Similarly to the growth curves of protein in batch cultures, the protein curves in semicontinuous cultures showed that *2rrn* grow significantly faster than its negative control transformants (*P*-value = 1.8×10^{-11}), see Table 3.5 for their raw and corrected doubling times. Interestingly, this method could even showed that *2rrn* grow significantly faster than the WT_{M129} with the corrected doubling times (*P*-value = 3.0×10^{-14}). Alike with previous methods, this method could also showed that the negative control transformants grow significantly slower than WT_{M129} (*P*-value = 0.013).

The calculated doubling times by using the second method developed in this study, the microcolony growth assay, are shown in Figure 3.3. The data from wells coated with Concanavalin A, showed the lower dispersion, so this data from the 10-fold dilution were used for the analysis. This method showed successfully that the negative control transformants grow significantly slower than WT_{M129} (*P*-value = 0.3×10^{-3}) but failed to show a significant difference between *2rrn* and its negative control transformants (*P*-value = 0.056) (Table 3.5).

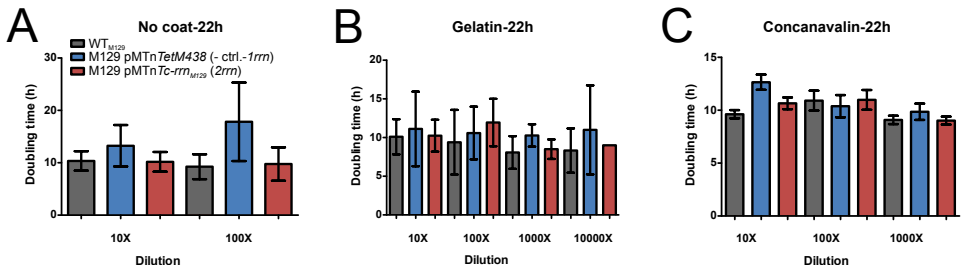


Figure 3.3: Doubling times by using the automated microcolony growth assay. WT_{M129}, M129 pMTn*TetM438* (- ctrl.-*1rrn*) and M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) are indicated in dark gray, blue and red, respectively. n = 2 (two independent experiments). Error bars = SD. **(A, B, C)** The calculated doubling times for WT_{M129}, M129 pMTn*TetM438* and M129 pMT*Tc-rrn*$_{M129}$ are plotted after 22 h of cell attachment, in wells **(A)** with no coating; **(B)** coated with 3% gelatin; and **(C)** coated with 200 µg/mL Concanavalin A Type IV.

In summary, the methods that were able to detect significant differences, holding true the hypothesis that duplicating the rRNA operon of *M. pneumoniae* increases its growth rate were the medium color (pH) change at $A_{430/560}$ and the intracellular protein concentrations in both batch and semicontinuous cultures.

An attempt to estimate cell size

The increase in growth rate found in M129 pMT*Tc-rrn*$_{M129}$ transformants ,containing two rRNA operon copies, measured by intracellular protein concentration might also be explained by an increase in protein concentration per cell, leading to bigger cells, instead of an increase in cell division rate. To estimate a possible difference in cell size between M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) transformants containing two rRNA copies and their controls containing one copy, scanning electron microscopy (SEM) of WT_{M129}; M129 pMTn*TetM438* negative control transformants (- ctrl.-*1rrn*); M129 pMT*Tc-ncrrn357* negative control transformants (- ctrl.-*1rrn*); and M129 pMT*Tc-rrn*$_{M129}$ transformants (*2rrn*), was performed. Estimation of cell size was attempted but due to the variability found in cell morphology and the filamentous growth, it was not possible to do it (Figure 3.4). Furthermore, the difference found with the corrected doubling times was of a decrease of 0.39 h in semicontinuous culture (Table 3.5), which represents a 4.2% of decrease with respect to the doubling time of the negative control M129 pMTn*TetM438* transformants. Therefore, it is very unlikely to find such a difference in cell size

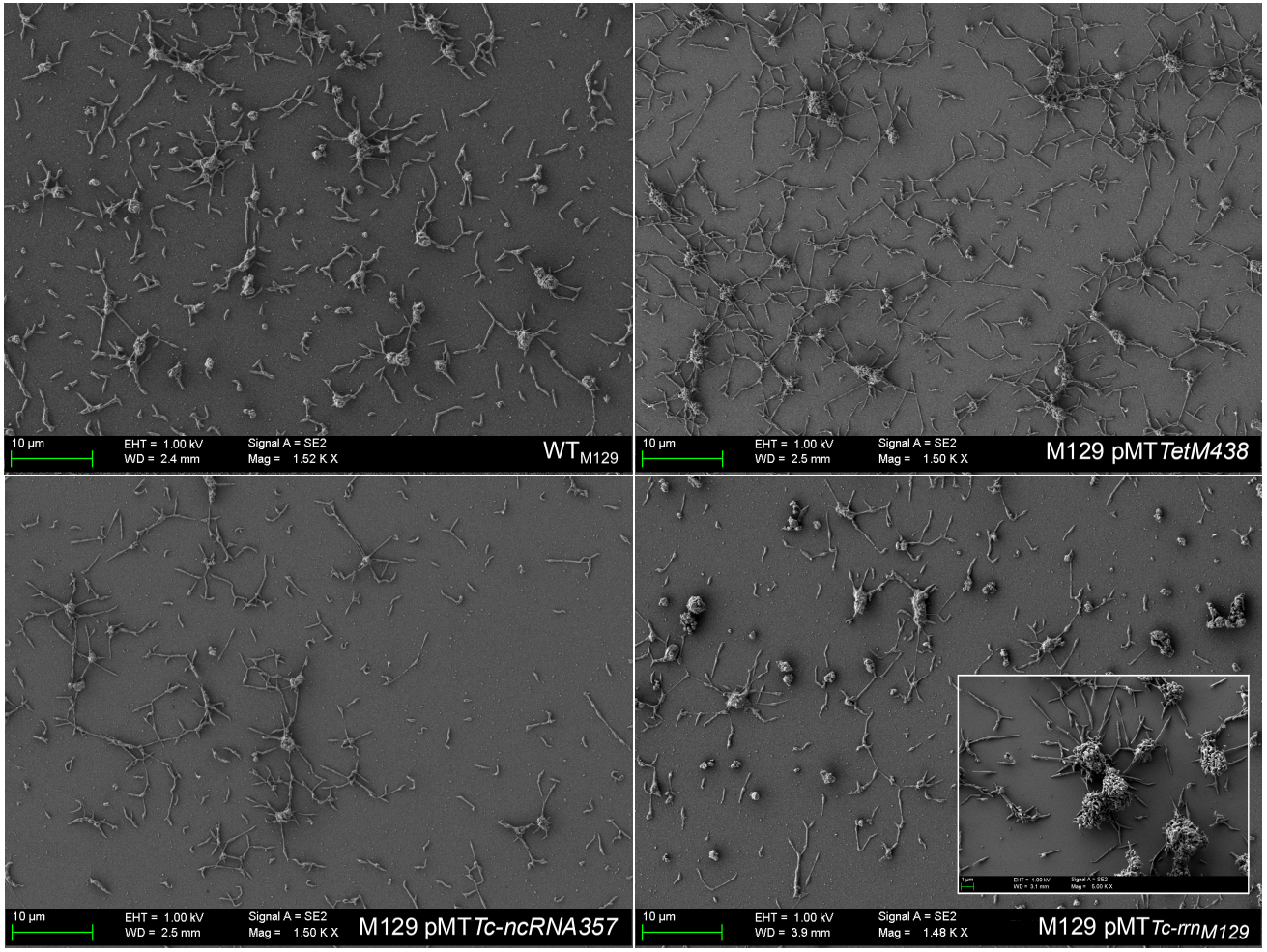


Figure 3.4: SEM of *M. pneumoniae* single cell and clumped transformants grown on glass coverslips. Images of WT_{M129}, M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) and its controls M129 pMTn*TetM438* (- ctrl.-*1rrn*), and M129 pMT*Tc-ncrrn357* (vector expressing the ncRNA357-*1rrn*), after 12 h of incubation at 37°C in Hayflick broth medium. Inset shows cell clumps, this sample was neither passed through a G25 syringe needle nor through a 0.22 µm filter.

estimations.

To estimate and compare properly cell size of transformants carrying two rRNA copies and those with one rRNA copy, it is necessary to isolate single transformants and select those to which the transposon insertion does not affect cell morphology. Typically, after a transformation with a transposon plasmid, the obtained transformants are cells with different mutated genes, some of these mutations will surely induce cell morphology changes, which will make more difficult to statistically determine any difference in cell sizes. Nonetheless, estimation of cell size is very complicated in *M. pneumoniae* due to morphological and size changes during its growth phases as previously observed (Domermuth et al., 1964; Zucker-Franklin et al., 1966; Bredt, 1968; Kammer et al., 1970; Biberfield and Biberfield, 1970). By phase contrast microscopy, Bredt (1968) observed that from 7 to 22 h of growth, *M. pneumoniae* filaments branched off to a network-like structure and after 24 h grape-like structures started to grow out from the colonies into the broth. Similarly, by electron microscopy, Zucker-Franklin et al. (1966) described *M. pneumoniae* as very pleomorphic during its exponential growth phase: round, oval, and tear-drop shapes were observed, and moreover, very variable diameters from 100 μm to 1000 μm were measured. Furthermore, by SEM, Biberfield and Biberfield (1970); Kammer et al. (1970) observed spherical organisms from a culture incubated for 2 h to 2 days; irregular, filamentous organisms with several protrusions and constrictions from a culture incubated from 2 to 6 days; and larger asymmetric spherical organisms from 6 to 10 days.

rRNA abundances

To test whether indeed the rRNA expression is increased in M129 pMT*Tc-rrn*$_{M129}$ transformants, containing two rRNA operon copies, RNA-seq was performed at 6, 24, 48 and 96 h of growth in batch cultures. Unfortunately, we were unable to determine a significant increase in the rRNA expression due to several reason. First, the rRNA is the most abundant RNA and most of the reads map already to the rRNA sequence. Second, it is quite difficult to compare between the samples because rRNA reads are usually used to normalize between samples and thus compare mRNA expressions. Third, the exogenous rRNA copy is identical to the

endogenous one, therefore, we were not even able to determine whether the exogenous copy is expressed. For these reasons, we added an exogenous rRNA copy of a closely related *M. pneumoniae* strain and quantified the relative expression of rRNA by a customized RNA-seq protocol as shown in section 3.3.1.2.

Cellular protein, r-proteins and ribosomal abundances

As mentioned above, it has been shown that faster growth rates are the result of an increase in protein synthesis determined by an increase in the number of ribosomes. To test, whether this is the case in M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) transformants, mass spectrometry (MS) of total protein extracts was performed. The volcano plot, shown in Figure 3.5.A., indicated a significant fold increase of most of the r-proteins in *2rrn* transformants. Similarly, some r-proteins of the large and small ribosomal subunits were found to be upregulated by western immunoblotting (Figure 3.5.B.).

Moreover, we performed ribosomal profiling to quantify the number of ribosomes in *2rrn* transformants (Figure 3.5.C.). The results of this assay showed that there was an increase in the number of monosomes (70S) in *2rrn* transformants. Taken together, these results suggest that there is indeed an increase in growth rate in *2rrn* transformants. The increase in growth rate was the result of the upregulation of r-proteins, due to the presumably upregulation or rRNA expression, which led to the increase of number of ribosomes per cell and, therefore, the increase in protein synthesis. This increase in protein synthesis is the factor that determines the increase in growth rate as previously determined in other bacteria.

3.3.1.2. *M. pneumoniae* with two to five rRNA operons

In a secondary approach, to increase the rRNA operon copy number to more than two copies, WT_{M129} was transformed with plasmid pMT*CmLox-rrn*$_{FH}$ and p*GmRCre* or with its respective control plasmid as described in 3.2.3.1. From these resulting transformants, their doubling times and quantification of rRNA relative expression were measured and are described below.

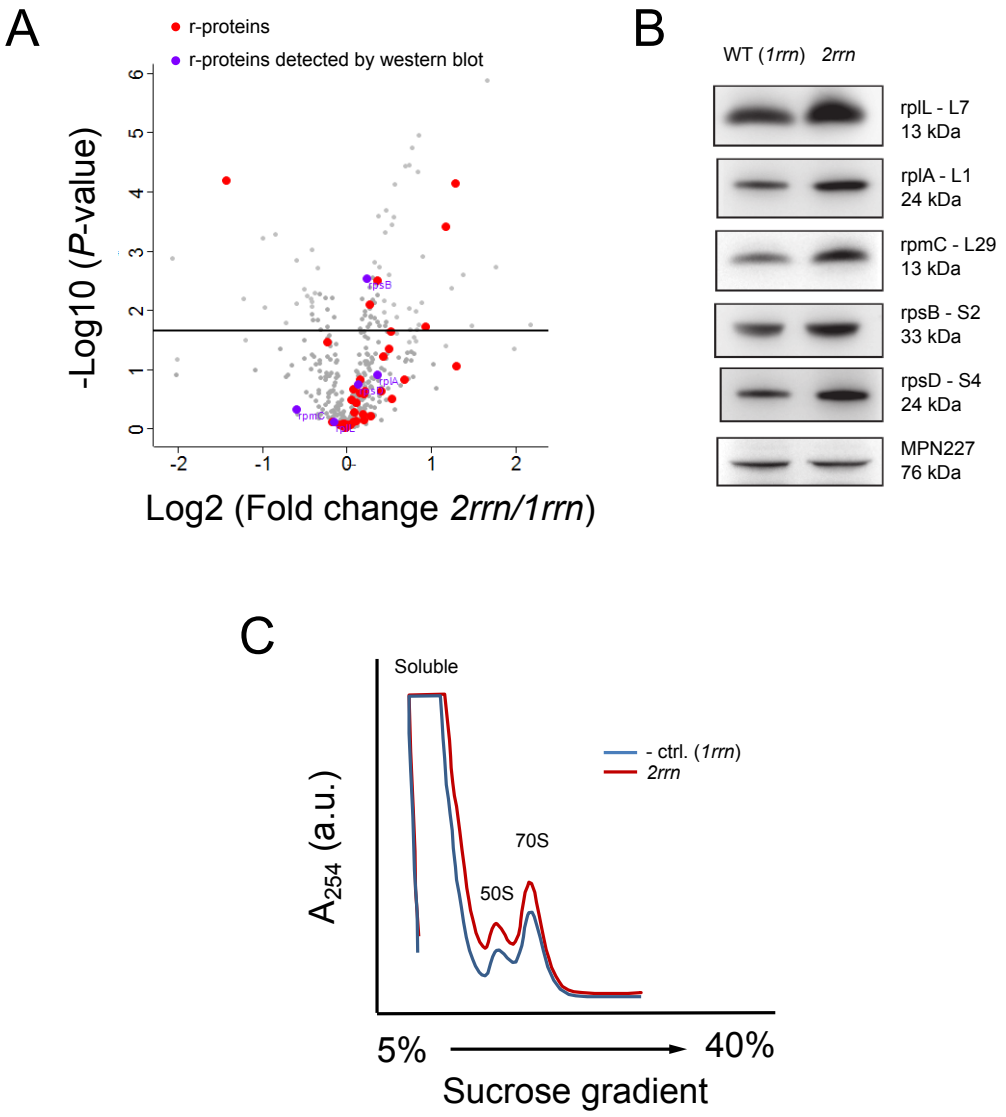


Figure 3.5: Ribosomal (r-) proteins and ribosomes abundances of *M. pneumoniae* transformants. Samples of WT_{M129} and M129 pMT*Tc-rrn*$_{M129}$ (*2rrn*) transformants are shown. **(A)** Volcano plot of mass spectrometry data showing protein abundances fold changes. **(B)** Detection of r-proteins by western immunobloting. MPN227 was used as loading control. **(C)** Ribosomal profiles indicate the distribution and the abundance of cytosolic (soluble), large ribosomal subunit (50S) and monosomes (70S) in the gradient.

Growth curves

Growth curves of these transformants were performed by the methods that were

Table 3.6: Estimated doubling times (DT) of FH and M129 transformants containing from two to five rRNA copies and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each (n = 4). Statistical significant differences compared to the corresponding control/WT_{M129} and controls to WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT		
WT_{M129} (*1rrn*)	11.04 ± 0.48	(11.58 ± 0.33)
WT_{FH} (*1rrn*)	17.82 ± 1.41	(18.05 ± 1.39)
FH transformants		
FH pMTn*TetM438* (*1rrn*)	$14.59 \pm 1.43^{***}$	$(15.12 \pm 1.43)^{***}$
FH pMT*Tc-ncrrn357* (*1rrn*)	$21.06 \pm 1.79^{***}$	$(21.46 \pm 1.79)^{***}$
FH pMT*Tc-rrn$_{M129}$* (*2rrn*)	$12.80 \pm 0.24^{***}$	$(13.09 \pm 0.18)^{**}$
M129 transformants		
1st cycle of *rrn$_{FH}$-Cre*		
M129 pMT*CmLox*1 (*1rrn*)	$10.34 \pm 0.63^{**}$	$(10.29 \pm 0.31)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*1 (*2rrn*)	$10.17 \pm 0.15^{NS/***}$	$(10.10 \pm 0.14)^{*/***}$
M129 pMT*CmLox*1 - p*GmRCre*1 (*1rrn*)	$12.17 \pm 0.51^{***}$	$(12.28 \pm 0.46)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*1 - p*GmRCre*1 (*2rrn*)	$10.71 \pm 0.18^{***/*}$	$(10.67 \pm 0.15)^{***/***}$
2nd cycle of *rrn$_{FH}$-Cre*		
M129 pMT*CmLox*2 - p*GmRCre*1 (*1rrn*)	10.98 ± 0.11^{NS}	$(10.93 \pm 0.10)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*2 - p*GmRCre*1 (*3rrn*)	$10.56 \pm 0.24^{***/**}$	$(10.39 \pm 0.17)^{***/***}$
M129 pMT*CmLox*2 - p*GmRCre*2 (*1rrn*)	$10.23 \pm 0.19^{***}$	$(10.26 \pm 0.17)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*2 - p*GmRCre*2 (*3rrn*)	$9.91 \pm 0.18^{***/***}$	$(10.04 \pm 0.15)^{***/***}$
3rd cycle of *rrn$_{FH}$-Cre*		
M129 pMT*CmLox*3 - p*GmRCre*2 (*1rrn*)	$9.74 \pm 0.22^{***}$	$(9.80 \pm 0.19)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*3 - p*GmRCre*2 (*4rrn*)	$9.90 \pm 0.19^{*/***}$	$(9.87 \pm 0.17)^{NS/***}$
M129 pMT*CmLox*3 - p*GmRCre*3 (*1rrn*)	$9.82 \pm 0.16^{***}$	$(9.84 \pm 0.12)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*3 - p*GmRCre*3 (*4rrn*)	$10.88 \pm 0.23^{***/NS}$	$(10.68 \pm 0.22)^{***/***}$
4th cycle of *rrn$_{FH}$-Cre*		
M129 pMT*CmLox*4 - p*GmRCre*3 (*1rrn*)	$10.29 \pm 0.09^{***}$	$(10.48 \pm 0.06)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*4 - p*GmRCre*3 (*5rrn*)	$10.66 \pm 0.14^{***/**}$	$(10.81 \pm 0.11)^{***/***}$
M129 pMT*CmLox*4 - p*GmRCre*4 (*1rrn*)	$9.73 \pm 0.23^{***}$	$(10.10 \pm 0.13)^{***}$
M129 pMT*CmLox-rrn$_{FH}$*4 - p*GmRCre*4 (*5rrn*)	$10.43 \pm 0.15^{***/***}$	$(10.29 \pm 0.12)^{***/***}$

able to detect significant differences in section 3.3.1.1, namely, medium color (pH) change (see methods in section 3.2.4.1) in batch cultures (Figure 3.6) and intracellular protein at two time points in semicontinuous cultures (see methods in section 2.3.2.1), from which doubling times were estimated (Table 3.6 and Figure

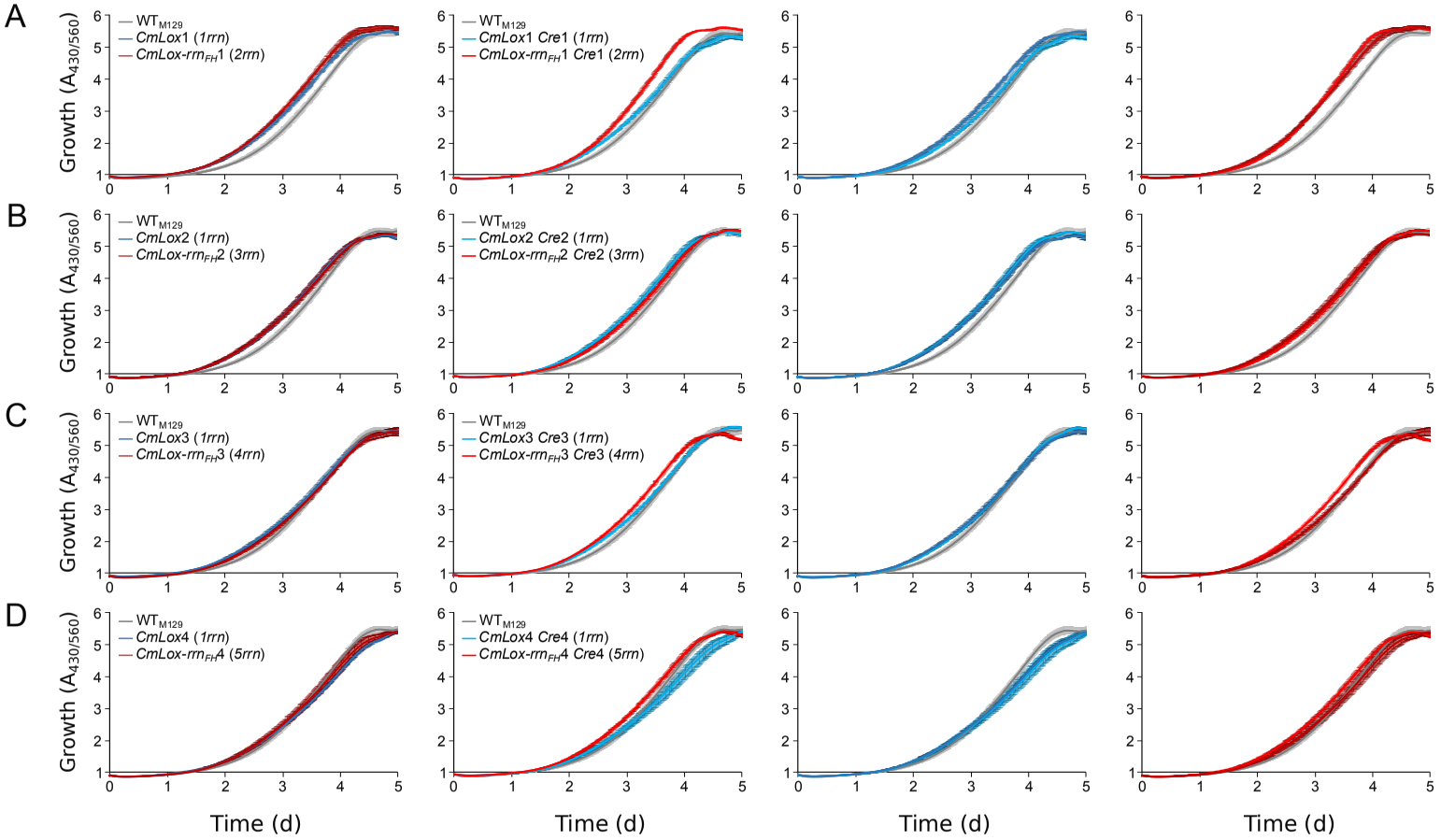


Figure 3.6: Growth curves (GC) of *M. pneumoniae* containing from two to five rRNA copies and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. WT_{M129} is indicated in grey, M129 transformants containing from two to five rRNA copies in red and their controls in blue. $n = 2$ (two independent experiments). Error bars = SD. **(A)** First, **(B)** second, **(C)** third and **(D)** fourth cycle of pMT*CmLox-rrn$_{FH}$* and p*GmRCre* transformation or with its control plasmid pMT*CmLox* and p*GmRCre*. Columns 3 and 4 rearrange the GC from columns 1 and 2 for better visualization.

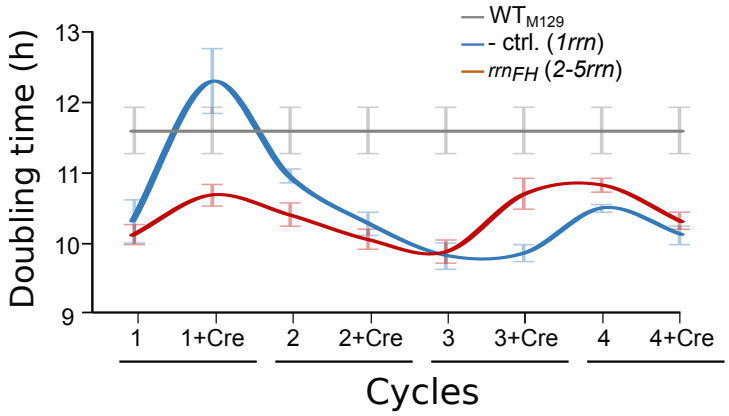


Figure 3.7: Doubling times of *M. pneumoniae* containing from two to five rRNA copies and their controls. Data from Table 3.6 $n = 4$ (two independent experiments with two technical replicates each). Error bars = SD.

3.7).

The medium color (pH) change, measured at $A_{430/560}$, showed a significant difference with both M129 pMT*CmLox-rrn*$_{FH}$1 (*2rrn*) and M129 pMT*CmLox-rrn*$_{FH}$ - p*GmRCre*1 (*2rrn*) with respect to their controls, from the fist cycle of transformations, indicating a faster medium acidification of the medium, hence a faster growth in the *2rrn* transformants, as already mentioned above (Figure 3.6). In the second, third and four cycles of transformations, there is not a observed significant difference between *3rrn*, *4rrn* and *5rrn* transformants and their controls, respectively.

From the methods developed in this study, only the intracellular protein concentrations in semicontinuous cultures was used. The great advantage of this method is that it determines differences with high statistical significance, it is also rapid and reproducible. The results of corrected doubling times showed that *2rrn* and *3rrn* transformants grow significantly faster than their controls (Table 3.6 and Figure 3.7). In contrast, *4rrn* and *5rrn* transformants grow significantly slower than their controls.

Surprisingly, in general with both methods, negative control transformants from

the four cycles of transformation, except in M129 pMT*CmLox*1 - p*GmRCre*1 transformants by intracellular protein, were found to grow significantly faster than WT_{M129} (Figure 3.6.Column 3, Figure 3.7 and Table 3.6) suggesting that this particular vector with the chloramphenicol resistance gene increases somehow the growth rate of *M. pneumoniae*. Consequently, *2rrn*, *3rrn*, *4rrn* and *5rrn* transformants turn out to grow significantly faster than WT_{M129} measured by intracellular protein.

In summary, the hypothesis that increasing the rRNA operon copy number of *M. pneumoniae* would increase its growth rate, holds true up to 3 rRNA copies, suggesting that is another factor that is limiting high growth rate in transformants with 4 and 5 rRNA copies.

Quantification of relative expression of rRNA copies by customized RNA-seq

To quantify relative expressions of endogenous and exogenous rRNA copies, 6 regions of the rRNA operon were quantified with a customized RNA-seq protocol (Figure 3.1). A preliminary test was performed by quantifying the exogenous rRNA copy of WT_{M129} ($rRNA_{M129}$) in *M. pneumoniae* FH and *M. genitalium* G37 transformants, which had been transformed with plasmid pMT*Tc-rrn*$_{M129}$ or its controls: pMTn*TetM438* and pMT*Tc-ncrrn357* (Table 3.7). Doubling times of *M. pneumoniae* FH transformants were estimated by intracellular protein at two time points in semicontinuous culture (see methods in section 2.3.2.1) (Table 3.6). Similarly to *M. pneumoniae* M129 transformants with 2 rRNA operon copies, the doubling time of *M. pneumoniae* FH transformants with 2 rRNA operon copies also decreased significantly with respect to the controls. M129 pMT*Tc-ncrrn357* transformants were thought to do not change *M. pneumoniae* growth rate (Lluch-Senar et al., 2015b), however with the corrected doubling times, we found that it actually increases its doubling time making it a bad control of a vector expressing a non-coding RNA.

According to the rRNA operon sequences of *M. pneumoniae* FH ($rRNA_{FH}$) and *M. pneumoniae* M129 ($rRNA_{M129}$), there are one insertion and deletion (InDel) and

Table 3.7: Preliminary test for the quantification of relative expression of endogenous and exogenous rRNA copies by customized RNA-seq. The numbers of specific reads of FH and G37 transformants are shown per region sequenced. The number of rRNA copies is indicated in brackets in each transformant. Values indicate the mean ± SD of two independent experiments (n = 2) of number of specific reads and its percentage in brackets.

Type of cells (No. of total rRNA copies)	**No. of specific reads per rRNA region (%)**	
	Endogenous	**Exogenous**
FH transformants	**$rRNA_{FH}$**	**$rRNA_{M129}$**
Region (3)		
FH pMTn*TetM438* (*1rrn*)	249 ± 4	0
	(100.00 ± 0.00)	(0)
FH pMT*Tc-ncrrn357* (*1rrn*)	209 ± 2	0
	(100.00 ± 0.00)	(0)
FH pMT*Tc-*rrn_{M129} (*2rrn*)	129 ± 11	36 ± 3
	(78.23 ± 2.59)	(21.77 ± 2.59)
G37 transformants	**$rRNA_{G37}$**	**$rRNA_{M129}$**
Region (1)		
G37 pMTn*TetM438* (*1rrn*)	1,306 ± 25	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-ncrrn357* (*1rrn*)	1,311 ± 16	2 ± 1
	(99.89 ± 0.04)	(0.11 ± 0.04)
G37 pMT*Tc-*rrn_{M129} (*2rrn*)	601 ± 72	616 ± 50
	(49.29 ± 0.98)	(50.71 ± 0.98)
Region (2)		
G37 pMTn*TetM438* (*1rrn*)	768 ± 125	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-ncrrn357* (*1rrn*)	584 ± 255	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-*rrn_{M129} (*2rrn*)	370 ± 30	400 ± 36
	(48.04 ± 0.26)	(51.96 ± 0.26)
Region (5)		
G37 pMTn*TetM438* (*1rrn*)	1,194 ± 87	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-ncrrn357* (*1rrn*)	1,279 ± 66	1 ± 1
	(99.96 ± 0.04)	(0.04 ± 0.04)
G37 pMT*Tc-*rrn_{M129} (*2rrn*)	600 ± 6	815 ± 21
	(42.41 ± 0.84)	(57.59 ± 0.84)
Region (6)		
G37 pMTn*TetM438* (*1rrn*)	449 ± 13	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-ncrrn357* (*1rrn*)	469 ± 103	0
	(100.00 ± 0.00)	(0)
G37 pMT*Tc-*rrn_{M129} (*2rrn*)	218 ± 20	134 ± 10
	(61.98 ± 0.49)	(38.02 ± 0.49)

two Single Nucleotide Polymorphisms (SNPs) that allow to distinguish the two rRNA operons from the two strains. The InDel is located in region (3), SNP1 in (4) and SNP2 in (6). After the custom-RNA-seq, the two SNPs turned out not to exist. Thus, region (3), in the 23S rRNA, was the unique region of the rRNA operon that possesses an InDel enabling to distinguish $rRNA_{FH}$ from $rRNA_{M129}$. Therefore, this region allowed to quantify the expression of the endogenous ($rRNA_{FH}$) and exogenous ($rRNA_{M129}$) copies in *M. pneumoniae* FH transformants. As expected, no reads of the exogenous $rRNA_{M129}$ were observed in the control transformants FH pMTn*TetM438* and FH pMT*Tc-ncrrn357*, whereas in FH pMT*Tc-rrn*$_{M129}$, $21.77 \pm 2.59\%$ of reads were detected to come from exogenous $rRNA_{M129}$ and the remaining $78.23 \pm 2.59\%$ came from the endogenous $rRNA_{FH}$.

In contrast to *M. pneumoniae* strains, there are much more sequence differences between rRNA operons from *M. pneumoniae* FH ($rRNA_{FH}$) and *M. genitalium* G37 ($rRNA_{G37}$). Thus, five regions from Figure 3.1 were processed by the customized RNA-seq. Region (3) could not be used because the forward primers used in PCR (FH1920_F and F2) did not hybridize in the endogenous $rRNA_{G37}$. After the custom-RNA-seq, the SNP from region (4) turned out not to exist, hence only regions (1), (2), (5), and (6) enabled to distinguish endogenous $rRNA_{G37}$ from exogenous $rRNA_{M129}$ in *M. genitalium* G37 transformants. As expected, no reads or very few of the exogenous $rRNA_{M129}$ were observed in the control transformants G37 pMTn*TetM438* and G37 pMT*Tc-ncrrn357*, whereas in G37 pMT*Tc-rrn*$_{M129}$, averaging the regions $49.57 \pm 7.16\%$ of reads were detected to come from exogenous $rRNA_{M129}$ and the remaining $50.43 \pm 7.16\%$ came from the endogenous $rRNA_{G37}$.

With the preliminary test, we could detect a very low number of specific reads due to the low number of random nucleotides in the customized cDNA primers. Therefore, to get a larger number of reads identifying rRNA molecules, we increased the number of random nucleotides from six in FH1920_R to twelve in 1920_R12N (Table 3.4) to amplify region 3, and we quantified the relative expression of exogenous and endogenous rRNA copies in M129 transformants containing two, three, four and five rRNA copies (Table 3.8). As controls, we quantified

Table 3.8: Quantification of relative expression of endogenous and exogenous rRNA copies by customized RNA-seq. The numbers of specific reads from region (3) of WT_{M129}, WT_{FH} and M129 transformants (containing from one to five rRNA copies which is indicated in brackets) are shown. Values indicate the mean ± SD of two independent experiments (n = 2) of number of specific reads and its percentage in brackets.

Type of cells (No. of total rRNA copies)	**No. of specific reads in region (3) (%)**	
	$rRNA_{M129}$	**$rRNA_{FH}$**
WT		
WT_{M129} (*1rrn*)	998,283 ± 404	126 ± 6
	(99.99 ± 0.00)	(0.01 ± 0.00)
WT_{FH} (*1rrn*)	672 ± 25,973	1,104,919 ± 118,898
	(0.06 ± 0.04)	(99.94 ± 0.04)
M129 transformants		
1st cycle of *rrn_{FH}-Cre*		
M129 pMT*CmLox*1 (*1rrn*)	1,062,517 ± 107,557	158 ± 92
	(99.99 ± 0.01)	(0.01 ± 0.01)
M129 pMT*CmLox*1 - p*GmRCre*1 (*1rrn*)	1,031,654 ± 36,348	147 ± 19
	(99.99 ± 0.00)	(0.01 ± 0.00)
M129 pMT*CmLox-rrn_{FH}*1 (*2rrn*)	623,416 ± 38,099	406,413 ± 29,534
	(60.55 ± 0.28)	(39.45 ± 0.28)
M129 pMT*CmLox-rrn_{FH}*1 - p*GmRCre*1 (*2rrn*)	540,003 ± 12,098	392,063 ± 12,111
	(57.94 ± 0.21)	(42.06 ± 0.21)
2nd cycle of *rrn_{FH}-Cre*		
M129 pMT*CmLox*2 - p*GmRCre*1 (*1rrn*)	1,003,508 ± 43,355	147 ± 52
	(99.99 ± 0.00)	(0.01 ± 0.00)
M129 pMT*CmLox*2 - p*GmRCre*2 (*1rrn*)	1,132,186 ± 23,053	329 ± 112
	(99.97 ± 0.01)	(0.03 ± 0.01)
M129 pMT*CmLox-rrn_{FH}*2 - p*GmRCre*1 (*3rrn*)	420,568 ± 2,236	620,687 ± 10,093
	(40.39 ± 0.26)	(59.61 ± 0.26)
M129 pMT*CmLox-rrn_{FH}*2 - p*GmRCre*2 (*3rrn*)	375,963 ± 39,690	545,975 ± 59,506
	(40.79 ± 0.08)	(59.21 ± 0.08)
3rd cycle of *rrn_{FH}-Cre*		
M129 pMT*CmLox*3 - p*GmRCre*2 (*1rrn*)	655,363 ± 24,166	52 ± 22
	(99.99 ± 0.00)	(0.01 ± 0.00)
M129 pMT*CmLox*3 - p*GmRCre*3 (*1rrn*)	653,597 ± 58,380	96 ± 68
	(99.99 ± 0.01)	(0.01 ± 0.01)
M129 pMT*CmLox-rrn_{FH}*3 - p*GmRCre*2 (*4rrn*)	223,066 ± 6,171	477,524 ± 12,945
	(31.84 ± 0.01)	(68.16 ± 0.01)
M129 pMT*CmLox-rrn_{FH}*3 - p*GmRCre*3 (*4rrn*)	237,315 ± 10,082	504,956 ± 20,974
	(31.97 ± 0.02)	(68.03 ± 0.02)
4th cycle of *rrn_{FH}-Cre*		
M129 pMT*CmLox*4 - p*GmRCre*3 (*1rrn*)	649,152 ± 55,218	99 ± 39
	(99.99 ± 0.00)	(0.01 ± 0.00)
M129 pMT*CmLox*4 - p*GmRCre*4 (*1rrn*)	659,560 ± 27,449	236 ± 142
	(99.97 ± 0.02)	(0.03 ± 0.02)
M129 pMT*CmLox-rrn_{FH}*4 - p*GmRCre*3 (*5rrn*)	166,417 ± 1,343	468,117 ± 3,139
	(26.23 ± 0.03)	(73.77 ± 0.03)
M129 pMT*CmLox-rrn_{FH}*4 - p*GmRCre*4 (*5rrn*)	183,523 ± 9,205	519,442 ± 24,764
	(26.10 ± 0.05)	(73.90 ± 0.05)

WT_{M129} and WT_{FH}, few reads of $rRNA_{FH}$ and $rRNA_{M129}$, respectively, were detected. These reads are most probably the product of PCR or sequencing errors and their values are negligible. Similarly to WT controls, a negligible number of reads of $rRNA_{FH}$ were observed in the control transformants M129 pMT*CmLox* and M129 pMT*CmLox* - p*GmRCre* in all cycles, whereas in M129 pMT*CmLox-*rrn_{FH} and M129 pMT*CmLox-*rrn_{FH} - p*GmRCre* in all cycles the number of reads detected from exogenous $rRNA_{FH}$ increased as the number of rRNA copies increased in M129 transformants. As a result, on average 40.8 $\pm$ 1.3%, 59.4 $\pm$ 0.3%, 68.1 $\pm$ 0.1%, and 73.8 $\pm$ 0.1% of the reads came from exogenous $rRNA_{FH}$ in M129 transformants containing two, three, four and five rRNA copies, respectively. Indeed, this indicates that the relative expression of the exogenous $rRNA_{FH}$ increases when increasing the $rRNA_{FH}$ copy number, although this increase is not proportional.

3.3.2. Duplicating the tRNA genes of *M. pneumoniae*

To test whether duplicating the tRNA genes of *M. pneumoniae* would increase its growth rate, WT_{M129} was transformed with plasmid pMT*Pac-37tRNAs* or its controls: miniTn*4001-Puro-1*, containing the only the antibiotic resistance gene; and pMT*Pac-438gRNAeNT2*, a vector expressing the gRNAeNT2. From these resulting transformants, their doubling times were measured and are described below.

Growth curves

Growth curves of these transformants were performed by medium color (pH) change (see methods in section 3.2.4.1) (Figure 3.8) and colorimetric growth assay by serial dilutions (see methods in section 2.2.3.4) in batch cultures and intracellular protein at two time points in semicontinuous cultures (see methods in section 2.3.2.1), doubling times were estimated from these two latter methods (Table 3.9).

Both the medium color (pH) change, measured at $A_{430/560}$ (Figure 3.8.A) and the curves of the colorimetric growth assay by serial dilutions, measured at A_{550} (Table 3.9) suggested that negative control M129 miniTn*4001-Puro-1* (*37tRNA*)

Table 3.9: Estimated doubling times (DT) of duplicated tRNA genes and control transformants by different methods. M129 miniTn*4001-Puro-1*, M129 pMT*Pac-438gRNAeNT2* and M129 pMT*Pac-37tRNAs* are indicated as - ctrl.*37tRNA*, gRNA-*37tRNA* and *74tRNA*, respectively. For the color change dilutions assay, the Δt was taken when the diluted cultures reached an A_{550} = 0.4. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean $\pm$ SD. - indicates not measured. Statistical significant differences compared to the corresponding left/most left columns are indicated: NS: non-significant *P*-value, * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

Doubling time (h) by method	**Type of cells**			
	WT_{M129}	**- ctrl.*37tRNA***	**gRNA-*37tRNA***	***74tRNA***
Batch cultures				
Medium color (pH) change dilutions	6.47 ± 0.14	$7.31 \pm 0.15^{*}$	-	$7.36 \pm 0.55^{-/NS/NS}$
Semicontinuous cultures				
Protein concentration 0-48 h	8.97 ± 0.22	$9.36 \pm 0.23^{***}$	$9.15 \pm 0.25^{*/*}$	$9.19 \pm 0.20^{NS/*/**}$
(Corrected DT)	(9.15 ± 0.20)	$(9.24 \pm 0.16)^{NS}$	$(8.55 \pm 0.23)^{***/***}$	$(9.05 \pm 0.08)^{***/***/NS}$

transformants grow significantly slower than WT_{M129} (*P*-value = 0.02). Unfortunately, the curves did not show any significant difference between M129 pMT*Pac-37tRNAs* (*74tRNA*) transformants, containing 74 tRNA genes, and its negative control transformants (*P*-value = 0.91), although the former grow slightly slower. Unexpectedly, M129 pMT*Pac-438gRNAeNT2* or, for short, gRNA (*37tRNA*) control transformants grow significantly faster than negative control transformants and WT_{M129}, indicating that it is not a good control (Figure 3.8.A).

From the methods developed in this study, only the intracellular protein concentrations in semicontinuous cultures was used, because it determines differences with high statistical significance, it is also rapid and reproducible. In contrast with the former method, this method could indeed detect significant differences with the corrected doubling times between the *74tRNA* and its negative control-*37tRNA* transformants (*P*-value = 0.0003), growing faster than the latter. In contrast, *74tRNA* transformants were found to grow slower than its control gRNA-*37tRNA* transformants (*P*-value = 4.2×10^{-9}) (Table 3.9). In fact, as also observed with the medium color (pH) change (Figure 3.8.A), the control gRNA-*37tRNA* transformants do not seem to be a good control, because they actually significantly decrease the doubling time of the negative control -*37tRNA* transformants (*P*-value = 8.1×10^{-11}) and even that of WT_{M129} (*P*-value = 1.0×10^{-8}). Unlike the previous method, the corrected doubling times of this method could not show a

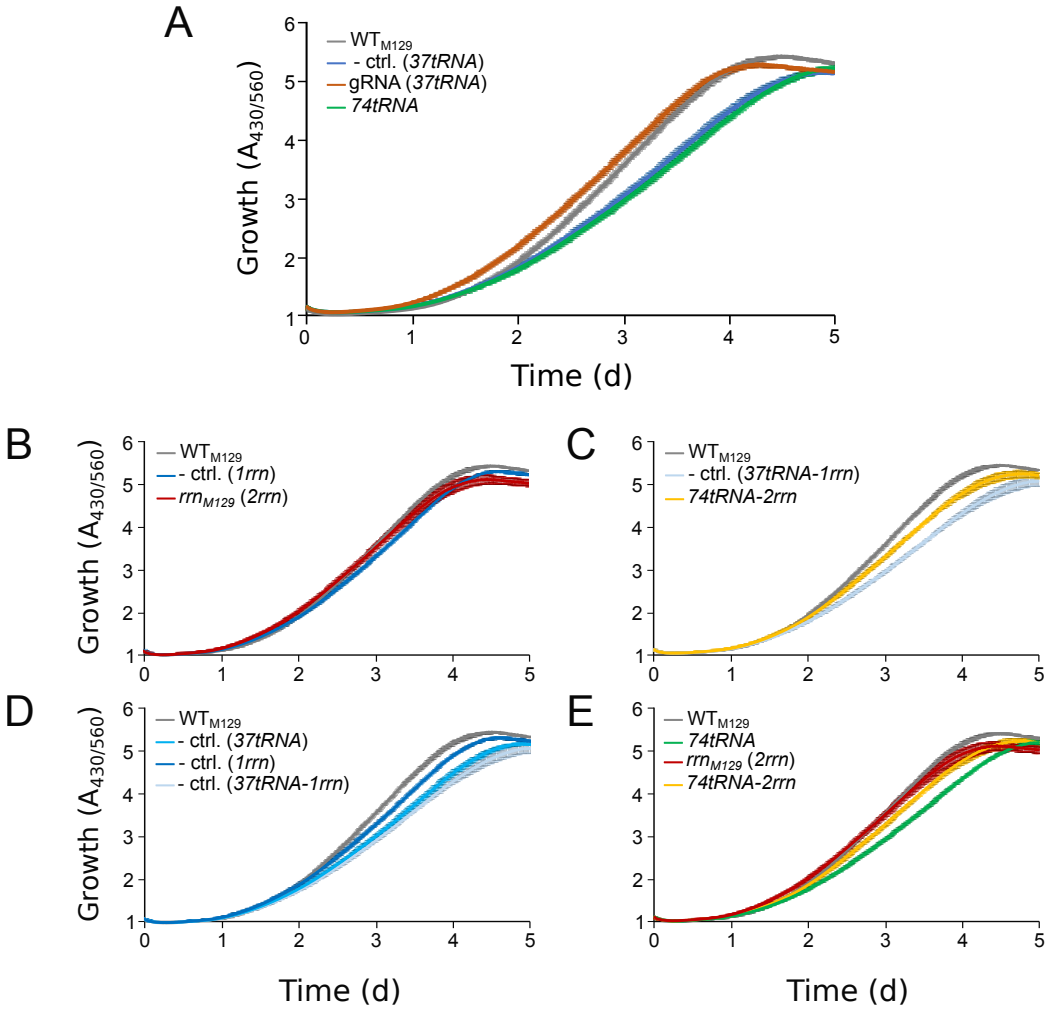


Figure 3.8: Growth curves (GC) of *M. pneumoniae* containing duplicated tRNA genes, rRNA operon and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. n = 2 (two independent experiments). Error bars = SD. **(A)** M129 miniTn*4001-Puro-1*, M129 pMT*Pac-438gRNAeNT2* and M129 pMT*Pac-37tRNAs* are indicated as - ctrl. (*37tRNA*), gRNA (*37tRNA*) and *74tRNA*, respectively. **(B)** M129 pMTn*TetM438* and M129 pMT*Tc-rrn*$_{M129}$ are indicated as - ctrl. (*1rrn*) and rrn$_{M129}$ (*2rrn*), respectively. **(C)** M129 miniTn*4001-Puro-1* - pMTn*TetM438* and M129 pMT*Pac-37tRNAs* - pMT*Tc-rrn*$_{M129}$ are indicated as - ctrl. (*37tRNA-1rrn*) and *74tRNA-2rrn*, respectively. **(D, E)** Rearrangement of **(A-C)** for better visualization of controls in **(D)** and samples in **(E)**.

significant difference between the negative control -*37tRNA* (M129 miniTn*4001-Puro-1*) transformants and WT_{M129} (P-value = 0.20), although the former grow slightly slower than the latter.

Table 3.10: Estimated doubling times (DT) of *M. pneumoniae* containing duplicated tRNA genes, rRNA operon and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each ($n = 4$). Statistical significant differences compared to the corresponding control/WT_{M129} or WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT_{M129}	8.97 ± 0.22	(9.15 ± 0.20)
M129 miniTn*4001-Puro-1* (- ctrl.-*37tRNA*)	$9.36 \pm 0.23^{***}$	$(9.24 \pm 0.16)^{NS}$
M129 pMT*Pac-37tRNAs* (*74tRNA*)	$9.19 \pm 0.20^{*/**}$	$(9.05 \pm 0.08)^{***/NS}$
M129 pMTn*TetM438* (- ctrl.-*1rrn*)	$7.99 \pm 0.30^{***}$	$(8.41 \pm 0.11)^{***}$
M129 pMT*Tc-rrn$_{M129}$* (*2rrn*)	$8.23 \pm 0.25^{*/***}$	$(8.20 \pm 0.15)^{***/***}$
M129 miniTn*4001-Puro-1* - pMTn*TetM438* (*37tRNA-1rrn*)	9.37 ± 0.17	$(9.36 \pm 0.11)^{**}$
M129 pMT*Pac-37tRNAs* - pMT*Tc-rrn$_{M129}$* (*74tRNA-2rrn*)	$9.09 \pm 0.09^{***/NS}$	$(8.88 \pm 0.07)^{***/***}$

Consequently, only the intracellular protein concentrations in semicontinuous cultures method was able to detect a significant difference, holding true the hypothesis that duplicating the tRNA genes of *M. pneumoniae* increases its growth rate. In order to test whether there was an additive increase in growth rate by duplicating both tRNA genes and the rRNA operon of *M. pneumoniae*, WT_{M129} was transformed with plasmids pMT*Pac-37tRNAs* and pMT*Tc-rrn$_{M129}$* or their negative control vectors miniTn*4001-Puro-1* and pMTn*TetM438*. From these resulting transformants, growth curves were performed by medium color (pH) change (Figure 3.8.C-E.) in batch culture and intracellular protein at two time points in semicontinuous culture, from which doubling times were estimated (Table 3.10).

With the medium color (pH) change, measured at $A_{430/560}$, M129 pMT*Pac-37tRNAs* - pMT*Tc-rrn$_{M129}$* double transformants, *74tRNA-2rrn* for short, seem to grow slightly slower than M129 pMT*Tc-rrn$_{M129}$* (*2rrn*) (Figure 3.8.E), although the difference is not significant.

In contrast with the former method, with the corrected doubling times calculated with the intracellular protein concentrations in semicontinuous cultures, *74tRNA-2rrn* double transformants were found to decrease the doubling time of the hy-

pothetical double transformants. This decrease in doubling time was, on average, additive, *74tRNA* single transformants grow 0.19 h faster than its negative control vector-*37tRNA*, and *2rrn* single transformants grow 0.21 h faster than its negative control vector-*1rrn*, summing 0.40 h (Table 3.10). Finally *74tRNA-2rrn* double transformants were found to grow 0.48 h faster than its negative control vector-*37tRNA-1rrn*. As a result, *74tRNA-2rrn* double transformants were found to grow slightly faster than the hypothetical double transformants, obtained by the addition of decrease in doubling times of *74tRNA* and *2rrn* single transformants with respect to their controls and consequently faster than WT_{M129}.

In summary, we showed that both hypothesis independently and grouped in double transformants are true. Thus, increasing both tRNA genes and rRNA operon copy number of *M. pneumoniae* increases its growth rate.

3.3.3. Duplicating the *oppABCDF* genes of *M. pneumoniae*

To test whether duplicating the *oppABCDF* genes of *M. pneumoniae* would increase its growth rate, WT_{M129} was transformed with plasmid pMT*Cm-OppA* and/or pMT*Pac-OppBCDF* or with their negative control vectors pMTn*Cat* and/or miniTn*4001-Puro-1* to obtain single and double transformants. From these resulting transformants, their doubling times were measured and are described below.

Growth curves

Growth curves of these transformants were performed by medium color (pH) change in batch cultures (see methods in section 3.2.4.1) (Figure 3.9) and intracellular protein at two time points in semicontinuous cultures (see methods in section 2.3.2.1), from which doubling times were estimated (Table 3.11).

The medium color (pH) change, measured at $A_{430/560}$ (Figure 3.9) suggested that neither single (OppA and OppBCDF) nor double transformants (OppA-BCDF) with duplicated *opp* genes grow significantly faster than their negative control transformants.

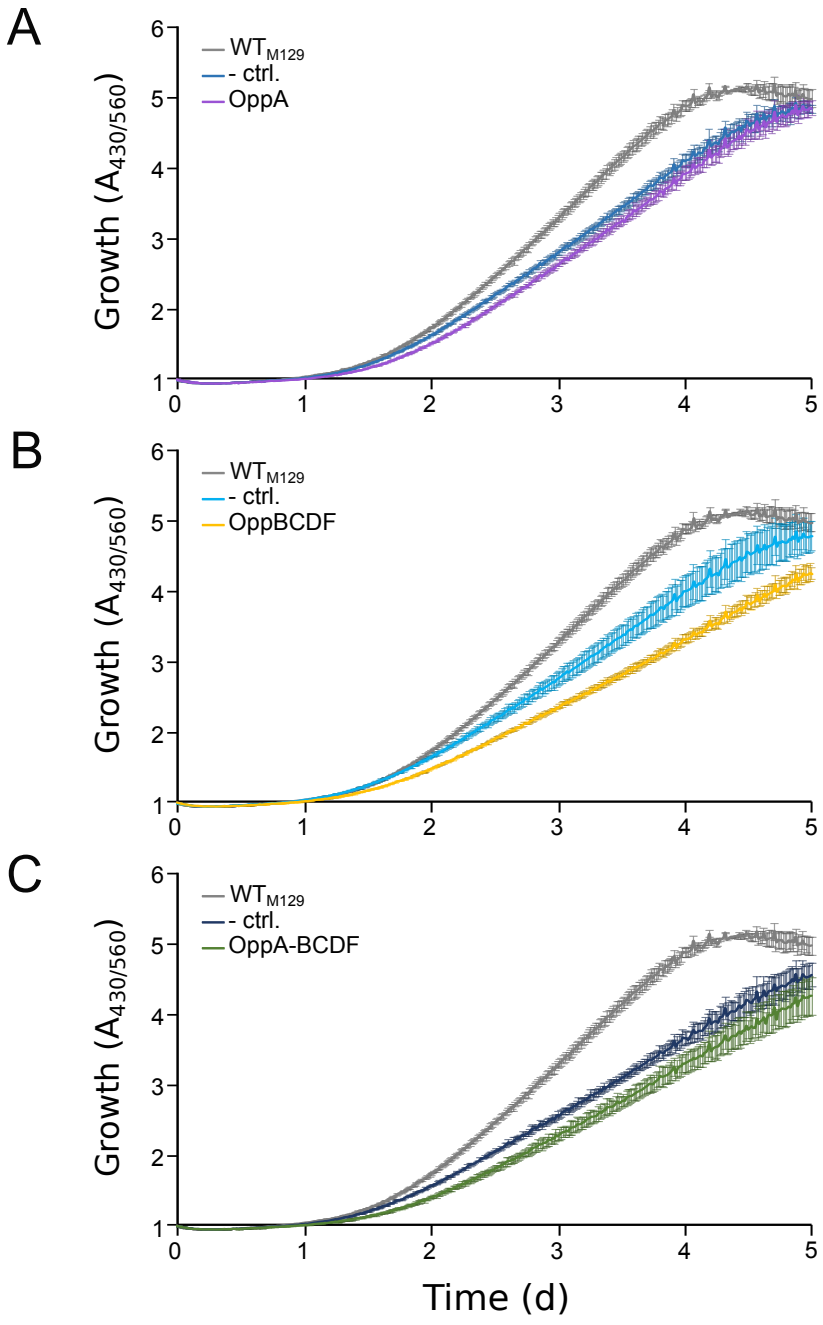


Figure 3.9: Growth curves (GC) of *M. pneumoniae* containing two sets of *opp* genes and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. $n = 2$ (two independent experiments). Error bars = SD. **(A)** M129 pMTn*Cat* and M129 pMT*Cm-OppA* are indicated as - ctrl. and OppA, respectively. **(B)** M129 miniTn*4001-Puro-1* and M129 pMT*Pac-OppBCDF* are indicated as - ctrl. and OppBCDF, respectively. **(C)** M129 pMTn*Cat* - miniTn*4001-Puro-1* and M129 pMT*Cm-OppA* - pMT*Pac-OppBCDF* are indicated as - ctrl. and OppA-BCDF, respectively.

Table 3.11: Estimated doubling times (DT) of *M. pneumoniae* containing two sets of *opp* genes and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each (n = 4). Statistical significant differences compared to the corresponding control/WT_{M129} or WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT_{M129}	8.97 ± 0.22	(9.15 ± 0.21)
M129 pMTn*Cat* (- ctrl.)	$8.13 \pm 0.23^{***}$	$(8.61 \pm 0.20)^{***}$
M129 pMT*Cm-OppA* (OppA)	$9.47 \pm 0.26^{***/***}$	$(9.06 \pm 0.26)^{***/NS}$
M129 miniTn*4001-Puro-1* (- ctrl.)	$9.36 \pm 0.23^{***}$	$(9.24 \pm 0.16)^{NS}$
M129 pMT*Pac-OppBCDF* (OppBCDF)	$9.23 \pm 0.19^{NS/**}$	$(9.31 \pm 0.18)^{NS/*}$
M129 pMTn*Cat* - miniTn*4001-Puro-1* (- ctrl.)	$8.67 \pm 0.22^{***}$	$(8.72 \pm 0.19)^{***}$
M129 pMT*Cm-OppA* - pMT*Pac-OppBCDF* (OppA-BCDF)	$8.73 \pm 0.13^{NS/***}$	$(8.89 \pm 0.08)^{**/***}$

Similarly, with the intracellular protein concentrations in semicontinuous cultures, the OppA single and OppA-BCDF double transformants with duplicated *opp* genes grow significantly slower than their negative control transformants (Table 3.11). The results of corrected doubling times showed that OppA transformants grow significantly slower than its negative control transformants (P-value = 6.3 $\times$ 10^{-6}). In contrast, the difference found between OppBCDF transformants and its negative control transformants was not significant (P-value = 0.23), although OppBCDF grow slightly slower. The resulting OppA-BCDF double transformants grow significantly slower than its negative control transformants (P-value = 0.003). However, OppA-BCDF double transformants were found to decrease the doubling time of the hypothetical double transformants. OppA single transformants grow 0.45 h slower than its negative control transformants, and OppBCDF single transformants grow 0.07 h slower than its negative control transformants, summing 0.52 h (Table 3.11). Finally OppA-BCDF double transformants were found to grow only 0.17 h slower than its negative control. As a result, OppA-BCDF double transformants are shown to decrease the doubling time of the hypothetical double transformants, obtained by the addition of the increase in doubling times of OppA and OppBCDF single transformants with respect to their controls.

As previously mentioned in section 3.3.2, the corrected doubling times of this

method could not show a significant difference between the negative control M129 miniTn*4001-Puro-1* transformants and WT_{M129} (*P*-value = 0.20).

Surprisingly and as noted above in section 3.3.1.2, M129 pMTn*Cat* single and M129 pMTn*Cat* - miniTn*4001-Puro-1* double transformants were found to grow significantly faster than WT_{M129} by intracellular protein method (Table 3.11) but not by medium color (pH) change (Figure 3.9.A, C.), in contrast to the results observed in in section 3.3.1.2. Consequently, OppA-BCDF double transformants turn out to grow significantly faster than WT_{M129} measured by intracellular protein.

In summary, medium color (pH) change in batch cultures does not hold true the hypothesis that by duplicating the *opp* genes of *M. pneumoniae* increases its growth rate. However, with the corrected doubling times calculated by intracellular protein in semicontinuous cultures, we showed that the independent expression of OppA and OppBCDF by themselves decreases the growth rate of *M. pneumoniae*, but their combined expression slightly increases the growth rate of the hypothetical OppA-BCDF double transformants.

3.3.4. Expressing NoxE of *Lactococcus lactis* in *M. pneumoniae*

To test whether expressing the NoxE of *L. lactis* would increase growth rate of *M. pneumoniae*, WT_{M129} was transformed with plasmid pMT*Pac-noxE* or its negative control vector miniTn*4001-Puro-1*. From these resulting transformants, their doubling times were measured and are described below.

Growth curves

Growth curves of these transformants were performed by medium color (pH) change (see methods in section 3.2.4.1) (Figure 3.10) and colorimetric growth assay by serial dilutions (see methods in section 2.2.3.4) in batch cultures and intracellular protein at two time points in semicontinuous cultures (see methods in section 2.3.2.1), doubling times were estimated from these two latter methods (Table 3.12).

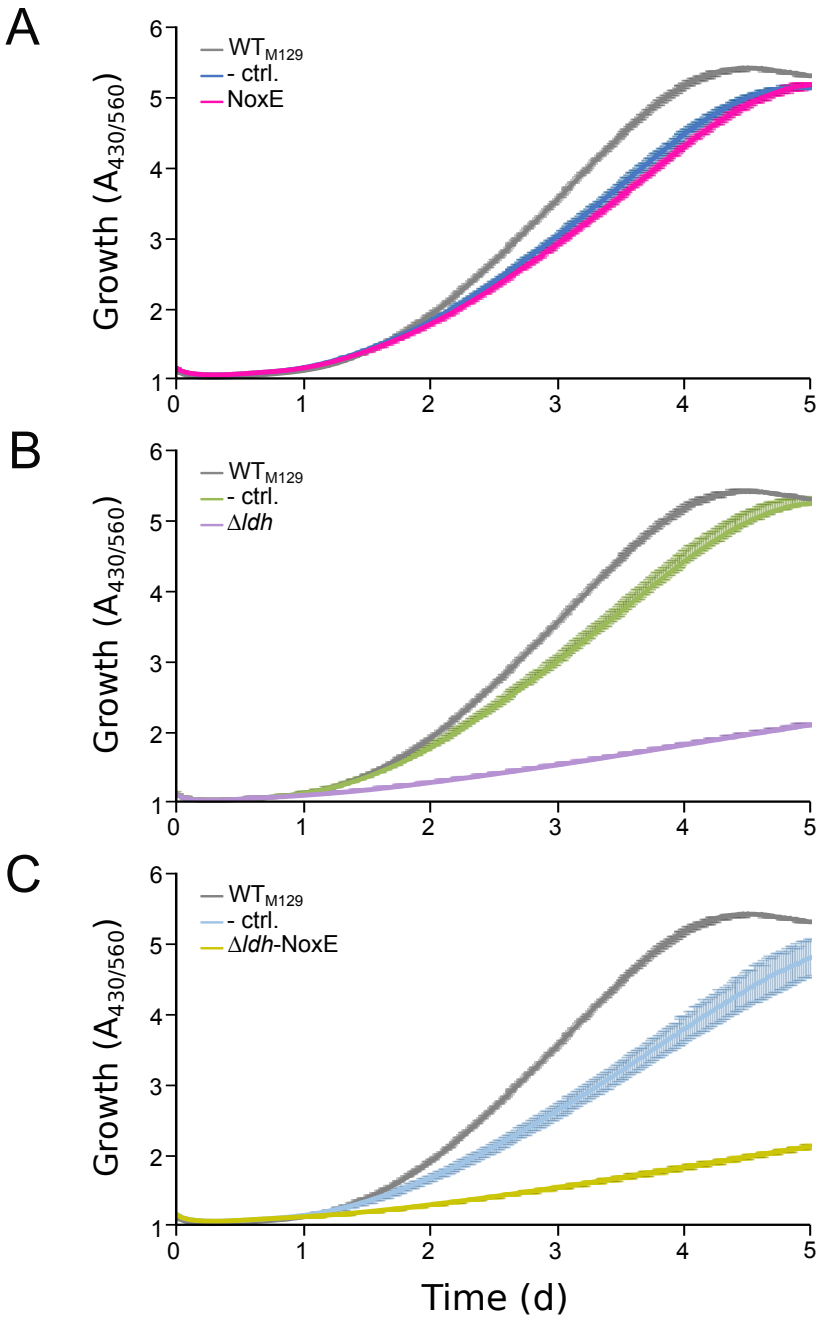


Figure 3.10: Growth curves (GC) of *M. pneumoniae* expressing NoxE, with the gene disruption of *glpD* and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. $n = 2$ (two independent experiments). Error bars = SD. **(A)** M129 miniTn*4001-Puro-1* and M129 pMT*Pac-noxE* are indicated as - ctrl. and NoxE, respectively. **(B)** M129 pMTn*Gm* and M129 *Δldh* are indicated as - ctrl. and *Δldh*, respectively. **(C)** M129 pMTn*Gm* - miniTn*4001-Puro-1* and M129 *Δldh* - pMT*Pac-noxE* are indicated as - ctrl. and *Δldh*-NoxE, respectively.

Table 3.12: Estimated doubling times (DT) of *M. pneumoniae* expressing NoxE and its controls by different methods. For the color change dilutions assay, the Δt was taken when the diluted cultures reached an A_{550} = 0.4. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean $\pm$ SD. Statistical significant differences compared to the corresponding left/most left columns are indicated: NS: non-significant *P*-value, * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

	Type of cells		
Doubling time (h) by method	**WT_{M129}**	**M129 miniTn*4001-Puro-1* (- ctrl.)**	**M129 MTn*Puro-noxE* (NoxE)**
Batch cultures			
Medium color (pH) change dilutions	6.47 ± 0.14	$7.31 \pm 0.15^{*}$	$7.33 \pm 0.52^{NS/NS}$
Semicontinuous cultures			
Protein concentration 0-48 h	8.97 ± 0.22	$9.36 \pm 0.23^{***}$	$10.04 \pm 0.36^{***/***}$
(Corrected DT)	(9.15 ± 0.20)	$(9.24 \pm 0.16)^{NS}$	$(9.79 \pm 0.29)^{***/***}$

Both the medium color (pH) change, measured at $A_{430/560}$ (Figure 3.10.A) and the curves of the colorimetric growth assay by serial dilutions, measured at A_{550} (Table 3.12), did not show any significant difference between M129 MTn*Puro-noxE* (NoxE) transformants and its negative control transformants (*P*-value = 0.96), although the former grow slightly slower. As mentioned above in section 3.3.2, these methods indicated that negative control M129 miniTn*4001-Puro-1* transformants grow significantly slower than WT_{M129} (*P*-value = 0.02).

From the methods developed in this study, only the intracellular protein concentrations in semicontinuous cultures was used, because it determines differences with high statistical significance, it is also rapid and reproducible. In contrast with the former method, this method could indeed detect significant differences with the corrected doubling times between the NoxE and its negative control transformants (*P*-value = 2.3×10^{-7}), as well as with WT_{M129} (*P*-value = 5.3×10^{-8}), growing slower than these latter two (Table 3.12). As previously mentioned and unlike the previous method, the corrected doubling times of this method could not show a significant difference between the negative control (M129 miniTn*4001-Puro-1*) transformants and WT_{M129} (*P*-value = 0.20), although the former grow slightly slower than the latter.

Consequently, these two methods do not hold the hypothesis showing that the

Table 3.13: Estimated doubling times (DT) of *M. pneumoniae* expressing NoxE, with the gene disruption of *glpD* and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each (n = 4). Statistical significant differences compared to the corresponding control or WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT_{M129}	8.97 ± 0.22	(9.15 ± 0.20)
M129 miniTn*4001-Puro-1* (- ctrl.)	$9.36 \pm 0.23^{***}$	$(9.24 \pm 0.16)^{NS}$
M129 pMT*Pac-noxE* (NoxE)	$10.04 \pm 0.36^{***}$	$(9.79 \pm 0.29)^{***}$
M129 pMTn*Gm* (- ctrl.)	$9.47 \pm 0.24^{***}$	$(9.47 \pm 0.15)^{***}$
M129 *Δldh*	$10.82 \pm 0.29^{***}$	$(10.76 \pm 0.17)^{***}$
M129 pMTn*Gm* - miniTn*4001-Puro-1* (- ctrl.)	$9.87 \pm 0.15^{***}$	$(9.54 \pm 0.12)^{***}$
M129 *Δldh* - pMT*Pac-noxE* (*Δldh*-NoxE)	$11.21 \pm 0.18^{***}$	$(10.97 \pm 0.18)^{***}$

growth rate of *M. pneumoniae* decreases when expressing NoxE. A secondary hypothesis was to test whether expressing NoxE in a genomic context lacking the lactate dehydrogenase, encoded by the *ldh* (*mpn674*) gene, would increase final biomass of *M. pneumoniae*. Lactate dehydrogenase catalyzes the reversible reaction of converting lactate from pyruvate, the final product of glycolysis, with the concomitant conversion of NADH to NAD^+, when oxygen is absent or in short supply. NADH oxidase, encoded by the *noxE* gene from *L. lactis*, converts NADH and oxygen to NAD^+ and water. The function replacement of Ldh by NoxE would balance redox as the Ldh does but it would probably delay the acidification thereby delaying growth arrest and increasing final biomass. Therefore, to test this secondary hypothesis, M129 *Δldh* and M129 pMTn*Gm* were transformed with plasmid pMT*Pac-noxE* and its negative control vector miniTn*4001-Puro-1*, respectively. From these resulting transformants, growth curves were performed by medium color (pH) change (Figure 3.10.C.) in batch culture and intracellular protein at two time points in semicontinuous culture, from which doubling times were estimated (Table 3.13).

With the medium color (pH) change, measured at $A_{430/560}$, *Δldh*-NoxE double transformants do not seem to grow differently from M129 *Δldh* (Figure 3.10.B-

C).

In contrast with the former method, with the corrected doubling times calculated with the intracellular protein concentrations in semicontinuous cultures, *Δldh*-NoxE double transformants were found to decrease the doubling time of the hypothetical double transformants. NoxE single transformants grow 0.55 h slower than its negative control transformants, and M129 *Δldh* single transformants grow 1.29 h slower than its negative control transformants, summing 1.84 h (Table 3.13). Finally *Δldh*-NoxE double transformants were found to grow only 1.43 h slower than its negative control transformants. As a result, *Δldh*-NoxE double transformants are shown to decrease the doubling time of the hypothetical double transformants, obtained by the addition of the increase in doubling times of NoxE and M129 *Δldh* single transformants with respect to their controls.

In summary, we showed that the expression of NoxE by itself decreases the growth rate of *M. pneumoniae*, however its expression in the *Δldh* genomic context slightly increases the growth rate of the hypothetical *Δldh*-NoxE double transformants.

3.3.5. Expressing GpsA of *M. penetrans* in *M. pneumoniae*

With the same purpose of maintaining a redox balance, we expressed GpsA, a NAD-dependent glycerol-3-phosphate dehydrogenase, which carries out the same reaction as the GlpD, a glycerol-3-phosphate dehydrogenase, but uses NAD^+ or $NADP^+$ as hydrogen acceptor instead of using oxygen. To test whether expressing the GpsA of *M. penetrans* would increase growth rate of *M. pneumoniae*, WT_{M129} was transformed with plasmid pMT*CmLox-gpsA* or its negative control vector pMT*CmLox*. From these resulting transformants, their doubling times were measured and are described below.

Growth curves

Growth curves of these transformants were performed by medium color (pH) change in batch cultures (see methods in section 3.2.4.1) (Figure 3.11) and intracellular protein at two time points in semicontinuous cultures (see methods in

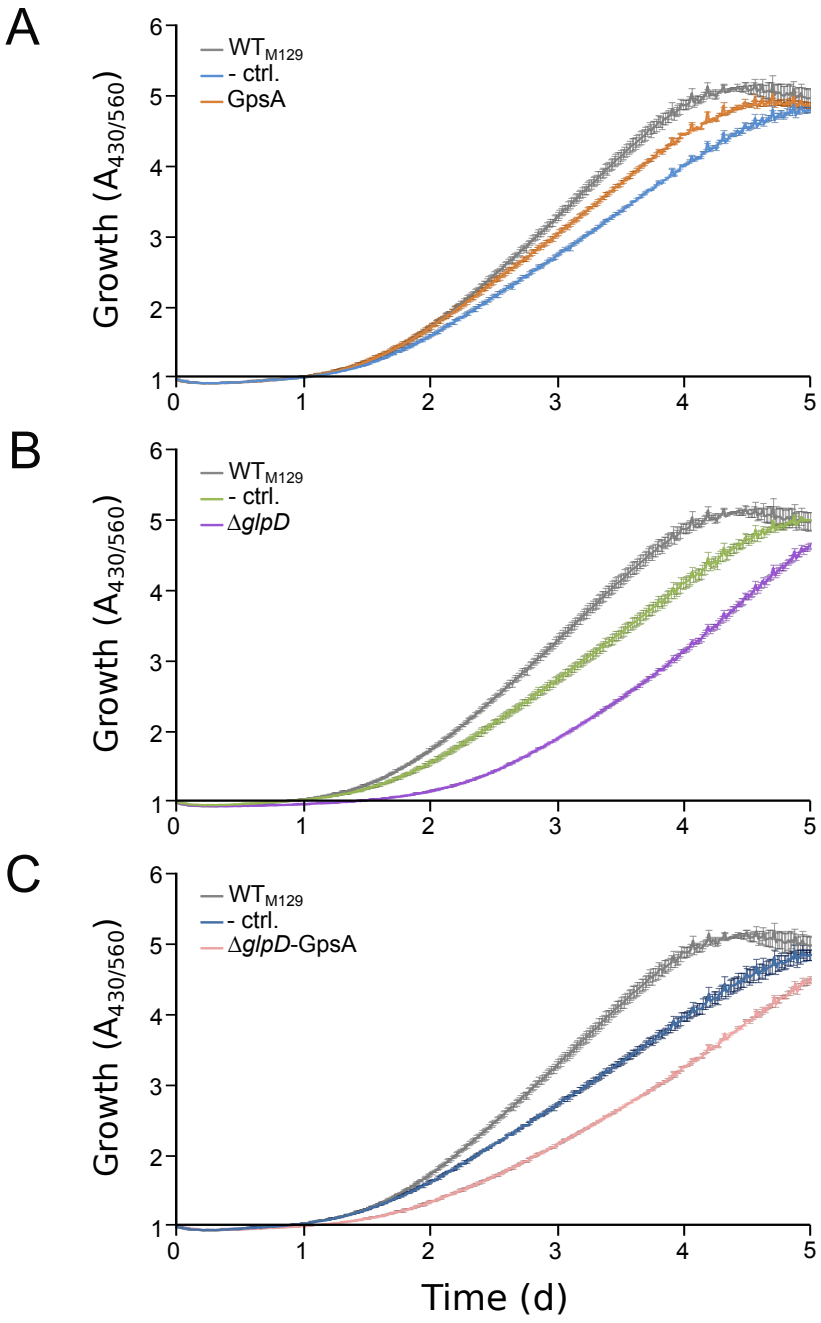


Figure 3.11: Growth curves (GC) of *M. pneumoniae* expressing GpsA, with the gene disruption of *glpD* and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. $n = 2$ (two independent experiments). Error bars = SD. **(A)** M129 pMT*CmLox* and M129 pMT*CmLox-gpsA* are indicated as - ctrl. and GpsA, respectively. **(B)** M129 pMTn*Gm* and M129 *ΔglpD* are indicated as - ctrl. and *ΔglpD*, respectively. **(C)** M129 pMTn*Gm* - pMT*CmLox* and M129 *ΔglpD* - pMT*CmLox-gpsA* are indicated as - ctrl. and *ΔglpD*-GpsA, respectively.

Table 3.14: Estimated doubling times (DT) of *M. pneumoniae* expressing GpsA, with the gene disruption of *glpD* and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each (n = 4). Statistical significant differences compared to the corresponding control or WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT_{M129}	9.30 ± 0.15	(9.07 ± 0.11)
M129 pMT*CmLox* (- ctrl.)	$9.56 \pm 0.34^{**}$	$(9.36 \pm 0.27)^{***}$
M129 pMT*CmLox-gpsA* (GpsA)	9.54 ± 0.19^{NS}	$(9.26 \pm 0.11)^{NS}$
M129 pMTn*Gm* (- ctrl.)	$9.58 \pm 0.10^{***}$	$(9.22 \pm 0.03)^{***}$
M129 *ΔglpD*	$9.85 \pm 0.16^{***}$	$(9.48 \pm 0.07)^{***}$
M129 pMTn*Gm* - pMT*CmLox* (- ctrl.)	$9.78 \pm 0.17^{***}$	$(9.38 \pm 0.13)^{***}$
M129 *ΔglpD* - pMT*CmLox-gpsA* (*ΔglpD*-GpsA)	$10.76 \pm 0.34^{***}$	$(10.15 \pm 0.33)^{***}$

section 2.3.2.1), from which the doubling times were estimated (Table 3.14).

The medium color (pH) change, measured at $A_{430/560}$ (Figure 3.11.A) showed that M129 pMT*CmLox-gpsA* (GpsA) transformants grow significantly faster than its negative control M129 pMT*CmLox* transformants.

In contrast with the former method, with the intracellular protein concentrations in semicontinuous cultures, the corrected doubling times showed that GpsA transformants grow slightly faster than its negative control transformants but the difference was not significant (P-value = 0.30) (Table 3.14).

Consequently, these two methods seem to hold true the hypothesis showing that the growth rate of *M. pneumoniae* increases when expressing GpsA, however the increase was not significant by the intracelular protein method. A secondary hypothesis was to test whether replacing the function of Glpd by GpsA would increase final biomass. This function replacement would balance redox as the Ldh does but it would probably delay the acidification thereby delaying growth arrest. Moreover, this would avoid the hydrogen peroxide production by the GlpD, which is one of the virulent factors of *M. pneumoniae* in its host cells. Since the ultimate

objective is to use *M. pneumoniae* as a chassis for animal vaccine production, it is of our interest to reduce the pathogenic factors. Therefore, to test this secondary hypothesis, M129 *ΔglpD* and M129 pMTn*Gm* were transformed with plasmids pMT*CmLox-gpsA* and its negative control vector pMT*CmLox*, respectively. From these resulting transformants, growth curves were performed by medium color (pH) change (Figure 3.11.C.) in batch culture and intracellular protein at two time points in semicontinuous culture, from which doubling times were estimated (Table 3.14).

With the medium color (pH) change, measured at $A_{430/560}$, *ΔglpD*-GpsA double transformants do not seem to grow differently from M129 *ΔglpD* transformants, they grow significantly slower than its negative control transformants (Figure 3.11.B-C).

In contrast with the former method, with the corrected doubling times calculated with the intracellular protein concentrations in semicontinuous culture, M129 *ΔglpD* transformants grow significantly slower than its negative control M129 pMTn*Gm* transformants (*P*-value = 1.9×10^{-7}). Similarly, *ΔglpD*-GpsA double transformants grow significantly slower than its negative control M129 pMTn*Gm* - pMT*CmLox* transformants (*P*-value = 3.8×10^{-8}) and they were found to further increase the doubling time of the hypothetical double transformants. GpsA single transformants grow 0.10 h faster than its negative control transformants, in contrast M129 *ΔglpD* single transformants were found to grow 0.26 h slower than its negative control transformants, summing 0.16 h for the hypothetical double transformants (Table 3.14). Finally *ΔglpD*-GpsA double transformants were found to grow 0.51 h slower than its negative control vector. As a result, *ΔglpD*-GpsA double transformants are shown to increase the doubling time of the hypothetical double transformants.

In contrast to as noted above in sections 3.3.1.2 and 3.3.3, M129 pMT*CmLox* single and M129 pMTn*Gm* - pMT*CmLox* double transformants were found to grow significantly slower than WT_{M129} by medium color (pH) change (Figure 3.11.A, C.) and intracellular protein method (Table 3.14).

In summary, it is not completely clear whether the expression of GpsA by itself increases the growth rate of *M. pneumoniae*, because the two methods were contradictory. However, the expression of GpsA decreases the growth rate of *M. pneumoniae* in the $\Delta glpD$ genomic context.

Chapter 4

MEDIUM-THROUGHPUT SCREENING OF FAST-GROWING MUTANTS OF *M. PNEUMONIAE*

4.1. Introduction

In this Chapter, we aimed to find fast-growing *M. pneumoniae* transformants from Serrano's lab collection by making a medium-throughput screening, in which we screened around 240 transformants (single transformants with gene disruptions, gene overexpressions or dominant negatives genes) for faster growers, and then, I validated those positives with a different method to measure growth rates. Furthermore, we grouped the genes of the faster-growing single mutants into three functional categories, namely, transcriptional, translational and metabolic genes, and then obtained the correspondent multiple mutants per functional category. Finally, I tested whether the multiple mutants showed an additive effect on growth.

4.2. Materials and methods

4.2.1. Bacterial strains and culture conditions

One Shot®TOP10 Chemically Competent *E. coli* (Invitrogen, Cat. No. C4040) or NEB 5-alpha Competent *E. coli* (New England Biolabs, Cat. No. C2987H) was used for plasmid amplification and cloning. They were grown in Luria-Bertani (LB) broth or LB agar plates containing 100 μg/mL ampicillin at 37°C.

Wild type *M. pneumoniae* strain (str.) M129 (WT_{M129}) (ATTC 29342, subtype 1, broth passage no. 35) (Regula et al., 2000) was used. The cell lines used in this study belong to Dr. Luis Serrano's lab collection and are detailed in Appendix A, Table A.1 A.2 (Yus *et al.*, in preparation). Briefly, proteins were flag-tagged (DYKDDDKG) in their N- or C-terminus, and expression was confirmed by western blotting with M2 monoclonal anti-flag (Sigma). In some cases, when the tag was foreseen to interfere with the protein function, they were expressed without. In general, promoter from the *tuf* gene (*mpn665*) was used for overexpression, unless otherwise indicated (in the cases in which the protein was toxic, the endogenous promoter was used instead). Dominant negative mutants or deletions were done in some cases. In few cases, TAP-tagged clones from Anne-Claude Gavin's collection (EMBL) were used (Kühner et al., 2009). Transposon insertion mutants were gently provided by Jörg Stülke and obtained by haystack mutagenesis (Halbedel and Stülke, 2007) with the mini-transposon plasmid pMT85 (Zimmerman and Herrmann, 2005), which contains the gentamicin resistance *aac(6')-aph(2")* gene. *M. pneumoniae* cells were grown in modified Hayflick's medium (hereafter referred to as Hayflick) as previously described (Hayflick, 1965; Yus et al., 2009). Static cell cultures were grown at 37°C in a humidified incubator gassed with 5% carbon dioxide (CO_2).

4.2.2. DNA manipulations

Genomic DNA of WT_{G37} was isolated by using the illustra bacteria genomicPrep Mini Spin Kit (GE Healthcare, 28-9042-59). Plasmidic DNA was obtained by us-

ing the QIAprep®Spin Miniprep kit (QIAGEN, Cat. No. 27106). DNA fragments were obtained by PCR amplification with the Phusion®High-Fidelity Polymerase (New England Biolabs, M0530). The purification of PCR products and digested fragments from agarose gels were performed by using the QIAquick®PCR Purification Kit (QIAGEN, Cat. No. 28106) and QIAquick®Gel Extraction Kit (QIAGEN, Cat. No. 28706), respectively. DNA was visualized by agarose gel electrophoresis with GelRed (Biotium, Cat. No. 41003) as stain. All oligonucleotides were ordered from Sigma-Aldrich purified by reverse phase. DNA Sequences were verified by Sanger sequencing with GATC Biotech. All primers used for molecular cloning and Sanger sequencing are listed in Tables 4.1 and 4.2.

4.2.2.1. Molecular cloning

All plasmids were constructed by using the Gibson assembly method (Gibson et al., 2009), the master mix was provided by the Biomolecular Screening Protein Technologies CRG facility. All plasmids used in this study are listed in Table 4.3.

pMT*TcLox* Plasmid miniTn*4001-Puro-1* was amplified by PCR using primers F_lox71XbaI.bb and R_lox66.bb. The resulting 4.3 kb PCR-generated vector was digested with DnpI and purified. The tetracycline (*tetM*) resistance coding region, under the 22 bp-promoter located upstream of the *mg438* translational start codon, was amplified from plasmid pMTn*TetM438* by PCR using primers F_lox66.Tc and R_lox71.Tc, which contained the 34-bp *lox66* and *lox71* sites, respectively. The resulting 2.0 kb PCR-generated insert and the 4.3 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*TcLox*.

pMT*PacLox* Plasmid miniTn*4001-Puro-1* was amplified by PCR using primers F_lox71XbaI.bb and R_lox66.bb. The resulting 4.3 kb PCR-generated vector was digested with DnpI and purified. The puromycin (*pac*) resistance gene, under the 168 bp-Spiralin promoter, was amplified from plasmid miniTn*4001-Puro-1* by PCR using primers F_lox66.Puro and R_lox71.Puro, which contained the 34-bp *lox66* and *lox71* sites, respectively. The resulting 1.1 kb PCR-generated insert and

Table 4.1: Sequence of primers used in this study for molecular cloning. Lower case indicates restriction sites, bold indicates overlapping regions between fragments and underlining indicates gene-specific sequences, unless otherwise noted.

Primer Name	Sequence (5'-3')
Cloning by Gibson Assembly	
F_lox71XbaI.bb[a]	*ATAACTTCGTATAATGTATGCTATACGAACGGTA*tctagaAGCGGCCGCTTT
R_lox66.bb[a]	*ATAACTTCGTATAGCATACATTATACGAACGGTA*agatctcccgggctgcagGAATTC GATATCAAGCTTATCGAT
pMT*TcLox*	
F_lox66.Tc[a]	**agatct*TACCGTTCGTATAATGTATGCTATACGAAGTTAT***ggatccTAGTATTTA GAATTAATAAAGTATGAAAATT
R_lox71.Tc[a]	**tctaga*TACCGTTCGTATAGCATACATTATACGAAGTTAT***AGAACTAGTggatcc
pMT*PacLox*	
F_lox66.Puro[a]	**agatct*TACCGTTCGTATAATGTATGCTATACGAAGTTAT***ggatccCAAGATCC AATGACCATGG
R_lox71.Puro[a]	**tctaga*TACCGTTCGTATAGCATACATTATACGAAGTTAT***AGAACTAGTGGA TCGGATCCACCAACTTTAATTTAGATC
pMT*GmLox*	
F_lox66.438.Gm[a]	***CCGTTCGTATAATGTATGCTATACGAAGTTAT***GggatccTAGTATTTAGAAT TAATAAAGTTAGTATTTAGAATTAATAAAGTATGAATATAGTTGAAAATG AAATATGTATAAG
R_lox71.Gm[a]	**tctaga*TACCGTTCGTATAGCATACATTATACGAAGTTAT***AGAACTAGTGGA TCggatccCTATTTAATACTAATGTCTTTTATAATAGCTTTTC
pUC57IR*GmLox-RpoBCD*	
F_pMTGmLox	**CTAGGTCTTTTTTATTGTTTTGgtcgac**TACCGTTCGTATAATGTATGCT ATAC
R_pMTGmLox[b]	**TGACCATGATTACGCCAAGC***GATAAAGTCCGTATAATTGTGTAAAA*G
F_IR[b]	**TTGTGAGCGGATAACAATTTCACACGAAT***GATAAAGTCCGTATAATTGT GTAAAA*gcgatcgcCTA
R_IR[b]	**AATGGCTGAAAATCAAAGCCATCATGATAGGCGATCGC***TTTTACACA ATTATACGGACTTTATC*att
pMT*CmLox-RpoA*	
F_p_rpmj	**AAGCTTGATATCGCTGCA**TTAGTTGAAGTGGAATTTTCCCCTTATGAT TTAAAGTTAGGTCGAA
R_p_rpmj	**ACTTTAAAAATTTTTCCAT**AGTATTTTAATTATTTTCTACCTATTATTCG ACCTAACTTTAAA
F_rpoA	**ATGGAAAAATTTTTAAAGT**ACG
R_rpoA	**GTATGGATCCCCCGGGC**TTAAGAACGGAATTTTAAACC

[a] Italic indicates the lox sites.
[b] Italic indicates the 26-bp IS256 inverted repeat.

the 4.3 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*PacLox*.

Table 4.2: Sequence of primers used in this study to verify DNA sequences.

Primer Name	Sequence (5'-3')
Tnp_F	CGCATCTTCCCCAATCAAACAT
TnpS_F	TCAAACATCAGCCAATCG
TnS_R	TGTGTGGAATTGTGAGCG
TetR_R	GAAATCAGTAGAATTGCCCCA
CmR_R	ACTGACTGAAATGCCTCAAA
PuroR_R	GGTGCTTCACTGTTTTCTTG
GmS_R	CCCAAGGCTCTGTATAATG

pMT*GmLox* Plasmid miniTn*4001-Puro-1* was amplified by PCR using primers F_lox71XbaI.bb and R_lox66.bb. The resulting 4.3 kb PCR-generated vector was digested with DnpI and purified. The *aac(6')-aph(2")* gene, which confers gentamicin resistance and is under the 22 bp-promoter located upstream of the *mg438* translational start codon, was amplified from plasmid pMTn*Gm* by PCR using primers F_lox66.438.Gm and R_lox71.Gm, which contained the 34-bp *lox66* and *lox71* sites, respectively. The resulting 1.6 kb PCR-generated insert and the 4.3 kb-purified vector were added in the Gibson assembly master mix to create plasmid pMT*GmLox*.

pUC57IR*GmLox-RpoBCD* *rpoBC* (*mg341* and *mg340*) operon from WT_{G37} was placed downstream of the 150 bp-promoter region of its orthologous operon from WT_{M129} (*mpn516* and *mpn515*). Downstream *rpoBC* operon, the *rpoD* (*mg249*) gene from WT_{G37} was placed under the control of the 100 bp-promoter region of its orthologous gene from WT_{M129} (*mpn352*) and downstream the 54-bp terminator of the rRNA operon of WT_{M129}. Two, three and three opal (TGA) stop codons, found in *rpoB*, *rpoC* and *rpoD* genes, respectively, were modified for thryptophan (TGG) codon. This 10 kb fragment was synthesized by GenScript with the AsiSI and SalI restriction sites at its 5' and 3' ends, respectively, and cloned in pUC57. Since this fragment was impossible to subclone in pMT*GmLox* once cloned in pUC57, the resulting plasmid pUC57-RpoBCD was linearised with the SalI enzyme and purified to be modified as following. The *aac(6')-aph(2")* gene, which confers gentamicin resistance and is under the 22 bp-promoter lo-

Table 4.3: Plasmids used in this study.

Plasmid Name	Description	Reference
pMTn*TetM438*	pMTn*4001* containing the tetracycline (*tetM*) resistance coding region under the 22 bp-promoter located upstream of the *mg438* translational start codon.	Pich et al. (2006)
pMTn*Gm*	pMTn*4001* containing the *aac(6')-aph(2")* marker, which confers gentamicin resistance.	Pich et al. (2006)
pMTn*Cat*	It is a derivative of pMTn*TetM438*, in which the EcoRI-BamHI *tetM438* selectable marker was replaced by an EcoRI-BamHI chloramphenicol (*cat*) resistance cassette.	Burgos and Totten (2014)
miniTn*4001-Puro-1*	pMTn*4001* containing the puromycin (*pac*) resistance gene under the 168 bp-Spiralin promoter.	Paetzold et al. (2013)
pMT*CmLox*	It is a derivative of pMTn*Cat*, in which the chloramphenicol (*cat*) resistance cassette is flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively.	This study, section 3.2.2.1
pMT*TcLox*	miniTn*4001-Puro-1* backbone containing the tetracycline (*tetM*) resistance coding region, under the 22 bp-promoter located upstream of the *mg438* translational start codon, and flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively.	This study
pMT*PacLox*	miniTn*4001-Puro-1* backbone containing the puromycin (*pac*) resistance gene, under the 168 bp-Spiralin promoter, and flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively.	This study
pMT*GmLox*	miniTn*4001-Puro-1* backbone containing the *aac(6')-aph(2")* gene, which confers gentamicin resistance and is under the 22 bp-promoter located upstream of the *mg438* translational start codon, and flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively.	This study
pUC57IR*GmLox-RpoBCD*	pUC57 containing the *aac(6')-aph(2")* gene flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively. *rpoBC* (*mg341* and *mg340*) operon and *rpoD* (*mg249*) gene from WT_{G37} are placed under the control of around 100 bp-own promoter region of their orthologous genes from WT_{M129} (*mpn516, mpn515* and *mpn352*). *aac(6')-aph(2")* gene with *lox* sites and the *rpoBCD* genes are flanked by the 26-bp IS256 inverted repeats.	This study
pMT*CmLox-RpoA*	pMT*CmLox* containing the *rpoA* (*mg177*) gene from WT_{G37} under the control of the 73 bp-promoter region of its orthologous operon from WT_{M129}, which begins with the *rpmj* (*mpn188*) gene.	This study
pMT*PacLox-5transla*	pMT*PacLox* containing the *rpsJ, rplC, rpmC, trpS* and *nusB* (*mg150, mg151, mg159, mg126* and *mg027*) genes from WT_{G37} under the control of approximately 100 bp-own promoter region of their orthologous genes from WT_{M129} (*mpn164, mpn165, mpn173, mpn265* and *mpn030*).	This study
pMT*TcLox-4metab*	pMT*TcLox* containing the *engA, eno, gmk* and *groES* (*mg329, mg407, mg107* and *mg393*) genes from WT_{G37} under the control of approximately 100 bp-own promoter region of their orthologous genes from WT_{M129} (*mpn475, mpn606, mpn246* and *mpn574*).	This study
pMT*TcLox-5metab*	pMT*TcLox-4metab* containing also the *ptsH* (*mg041*) gene from WT_{G37} under the control of the 86 bp-promoter region of its orthologous gene from WT_{M129} (*mpn053*).	This study

cated upstream of the *mg438* translational start codon, and flanked by *lox66* and *lox71* at the 5' and 3' ends, respectively, was amplified from plasmid pMT*GmLox* by PCR using primers F_pMTGmLox and R_pMTGmLox . This last primer contained the 26-bp IS256 inverted repeat. The resulting 1.7 kb PCR-generated fragment and the 12.6 kb SalI-digested pUC57-RpoBCD were added in the Gibson assembly master mix to create plasmid pUC57*GmLox-RpoBCD*. The resulting plasmid pUC57*GmLox-RpoBCD* was again linearised with the AsiSI enzyme and purified. The second 26-bp IS256 inverted repeat was generated by oligo assembly and polymerase action using primers F_IR and R_IR, which contained the second 26-bp IS256 inverted repeat. The resulting 94 bp oligo assembly-generated fragment and the 14.3 kb AsiSI-digested pUC57*GmLox-RpoBCD* were added in the Gibson assembly master mix to create plasmid pUC57IR*GmLox-RpoBCD*.

pMT*CmLox-RpoA* Plasmid pMT*CmLox* was linearised with the PstI enzyme and purified. The 73 bp-promoter region of its orthologous operon from WT_{M129}, which begins with the *rpmj* (*mpn188*) gene was generated by oligo assembly and polymerase action using primers F_p_rpmj and R_p_rpmj. The *rpoA* (*mg177*) gene was amplified from genomic DNA of WT_{G37} by PCR using primers F_rpoA and R_rpoA. The resulting 110 bp oligo assembly-generated fragment, 987 bp PCR-generated fragment and the 5.0 kb PstI-digested pMT*CmLox* were added in the Gibson assembly master mix to create plasmid pMT*CmLox-RpoA*.

pMT*PacLox-5transla* *rpsJ, rplC* and *rpmC* (*mg150, mg151* and *mg159*) genes from WT_{G37} were placed downstream of the 150 bp-promoter region of their orthologous genes from WT_{M129} (*mpn164, mpn165* and *mpn173*).*trpS* and *nusB* (*mg126* and *mg027*) genes from WT_{G37} were placed downstream of the 133 and 227 bp-own promoter regions of their orthologous genes from WT_{M129} (*mpn265* and *mpn030*), respectively.The the 54-bp terminator of the rRNA operon of WT_{M129} was placed downstream of each of the three subfragments mentioned above. One, two and one opal (TGA) stop codons, found in *rplC*, *rpmC* and *nusB* genes, respectively, were modified for thryptophan (TGG) codon. The total 3.6 kb fragment was synthesized by GenScript with the ApaI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with ApaI and SalI and

cloned into ApaI-SalI-digested pMT*PacLox*, generating the pMT*PacLox-5transla* plasmid.

pMT*TcLox-4metab* *engA* (*mg329*) gene from WT_{G37} was placed downstream of the 200 bp-promoter region of the *tuf* gene from WT_{M129} (*mpn665*) and the flag-tag. *eno, gmk* and *groES* (*mg407, mg107* and *mg393*) genes from WT_{G37} were placed downstream of the 109, 62 and 70 bp-own promoter regions of their orthologous genes from WT_{M129} (*mpn606, mpn246* and *mpn574*), respectively. The the 54-bp terminator of the rRNA operon of WT_{M129} was placed downstream of each of the four genes mentioned above. Three and two opal (TGA) stop codons, found in *eno* and *gmk* genes, respectively, were modified for thryptophan (TGG) codon. The total 4.3 kb fragment was synthesized by GenScript with the ApaI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with ApaI and SalI and cloned into ApaI-SalI-digested pMT*TcLox*, generating the pMT*TcLox-4metab* plasmid.

pMT*TcLox-5metab* *engA, eno, gmk* and *groES* were designed as indicated for pMT*TcLox-4metab*. *ptsH* (*mg041*) gene from WT_{G37} was placed downstream of the 86 bp-promoter region of its orthologous gene from WT_{M129} (*mpn053*) and downstream the 54-bp terminator of the rRNA operon of WT_{M129}. The total 4.8 kb fragment was synthesized by GenScript with the ApaI and SalI restriction sites at its 5' and 3' ends, respectively. The fragment was digested with ApaI and SalI and cloned into ApaI-SalI-digested pMT*TcLox*, generating the pMT*TcLox-5metab* plasmid.

4.2.3. Transformation of *M. pneumoniae*

WT_{M129} cells were transformed with some of the plasmids from Table 4.3 by electroporation (Hedreyda et al., 1993) as previously described in section 2.2.2. The transformants were selected and cultured in broth medium supplemented with the correspondent antibiotic, namely, 200 μg/mL gentamicin, 2 μg/mL tetracycline, 3.3 μg/mL puromycin, or 20 μg/mL chloramphenicol.

4.2.4. Growth curves conditions

Growth estimation for the medium-throughput screening were performed by means of the medium color (pH) change curve method and intracellular protein at 48 h in batch cultures, described below and for the validation by intracellular protein concentrations method in semicontinuous cultures (section 2.3.2.1).

4.2.4.1. Medium color (pH) change curve and intracellular protein at 48 h in batch cultures

Some modification for the medium color (pH) change method were carried out as previously described (Yus et al., 2017). Briefly, fresh prepared seed cultures, grown to late exponential phase (for 60 h) in 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL fresh Hayflick broth supplemented with gentamicin, were frozen and stored at -70°C Fresh prepared seed cultures, grown to late exponential phase (for 60 h) in 1:100 dilution in 25 cm^2 tissue culture flasks containing 5 mL fresh Hayflick broth supplemented with the corresponding antibiotic, were frozen and stored at -70°C and a 100 µL aliquot thereof was pelleted at 14100 ×g for 10 min, medium was removed and pelleted cells were washed with PBS (for 1 liter: 8 g NaCl, 0.2 g KCl, 0.24 g KH_2PO_4, 1.44 g $Na_2HPO_4 \cdot 2H_2O$, pH 7.4). A second cycle of centrifugation and washing was performed, After a third centrifugation, PBS was removed and pelleted cells were suspended in 100 µL lysis buffer (10 mM Tris·HCl, 6 mM $MgCl_2$, 1 mM EDTA, 100 mM NaCl, 0.1% Tx-100, pH 8, and 1× Protease Inhibitor Cocktail (Roche)). Cell lysates were kept on ice and spun down, pipetted up and down to complete lysis and extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225).

1 µg of these quantified frozen stocks was inoculated per well and in duplicates and in two 96-well plates, excluding the peripheral wells to avoid the edge effect, containing 200 µL Hayflick broth without antibiotic per well. The first 96-well plate was incubated at 37°C under 5% CO_2 and after 48 h of growth, medium was sucked out, cells were carefully washed twice with 200 µL PBS and lised with 100 µL lysis buffer. In the same first 96-well plate, cell lysates were kept

on ice and extracted protein was quantified by Pierce^{TM} BCA Protein Assay Kit. The second 96-well plate was incubated at 37°C in the Infinite^{TM} M200 Tecan plate reader. The absorbance at 430 and 560 were automatically taken, with a settle time at 300 ms and 25 flashes, every 20 min for 5 days. Medium color (pH) change was measured by calculating the ratio between the absorbance at 430 and 560 nm ($A_{430/560}$), as previously described (Yus et al., 2009).

To quantify growth, the intracellular protein concentrations taken at 48 h and the value of $A_{430/560}$ at which the medium color (pH) change curve reached its midpoint and two slopes (early and late) of such curve were determined. The early slope was calculated between 10 and 30 h of the curve and the late slope during the complete curve. The early (equation 4.1) and late (equation 4.2) slopes were calculated by determining the maximum median value of the slope between two time points separated by three and four time measurements, respectively, over successive periods of 30 time points.

$$Early\ slope = \frac{(A_{430/560})_{i+3} - (A_{430/560})_i}{t_{i+3} - t_i} \tag{4.1}$$

$$Late\ slope = \frac{(A_{430/560})_{i+4} - (A_{430/560})_i}{t_{i+4} - t_i} \tag{4.2}$$

The protein concentrations at 48 h and early slope are more representatives of growth, while the late slope and the value of $A_{430/560}$ at midpoint reflect the metabolic activity. These four parameters of growth and metabolism were analyzed for each batch of experiments. Outliers (larger than quartile 3 (Q3) by at least 1.5 times the interquartile range (IQR), or smaller than Q1 by at least 1.5 times the IQR) were removed to calculate the mean and the standard deviation of each of the parameters for each batch. Values larger or smaller than the mean by at least 2 times the standard deviation of each parameter were considered to determine as fast- or slow-growing/metabolizing transformants, respectively.

4.3. Results and discussion

To find alternative genes whose overexpression, disruption or other modifications have a favorable effect on *M. pneumoniae* growth, the *M. pneumoniae* transformants's collection was used to make a first medium-throughput screening of transformants having an effect on growth. The fast-growing transformants found in the screening were validated by a different method to estimate doubling times as described below.

4.3.1. Medium-throughput screening of transformants having an effect on growth

M. pneumoniae transformants's collection consisted of 236 single transformants, encompassing 148 genes to which certain modifications have been applied with the mini-transposon plasmids described in section 4.2.1 and detailed in Appendix A, Table A.2. There are 179 gene overexpressions (OE), 25 gene disruptions (DR), 25 gene point mutations (Mut), 6 point deletions (Del) and 1 fusion (Fus).

The estimation of growth of these transformants was performed by medium color (pH) change curves and intracellular protein at 48 h in batch cultures (see methods in section 4.2.4.1). Appendix A, Table A.2 indicates the transformants to which the growth curves were carried out and Appendix A, Table A.3 shows the transformants that exhibit an effect in growth and metabolic phenotype. Due to batch effects, the four parameters of growth (early slope and protein at 48 h) and metabolism (late slope and the value of $A_{430/560}$ at midpoint) were analyzed per batch experiment.

4.3.2. Validation of fast-growing transformants

From the transformants exhibiting an effect in growth and metabolic phenotypes, determined with the above methods, we validated the phenotype by determining intracellular protein concentration at two time points in semicontinuous cultures (see methods in section 2.3.2.1). Appendix A, Table A.2 indicates the transfor-

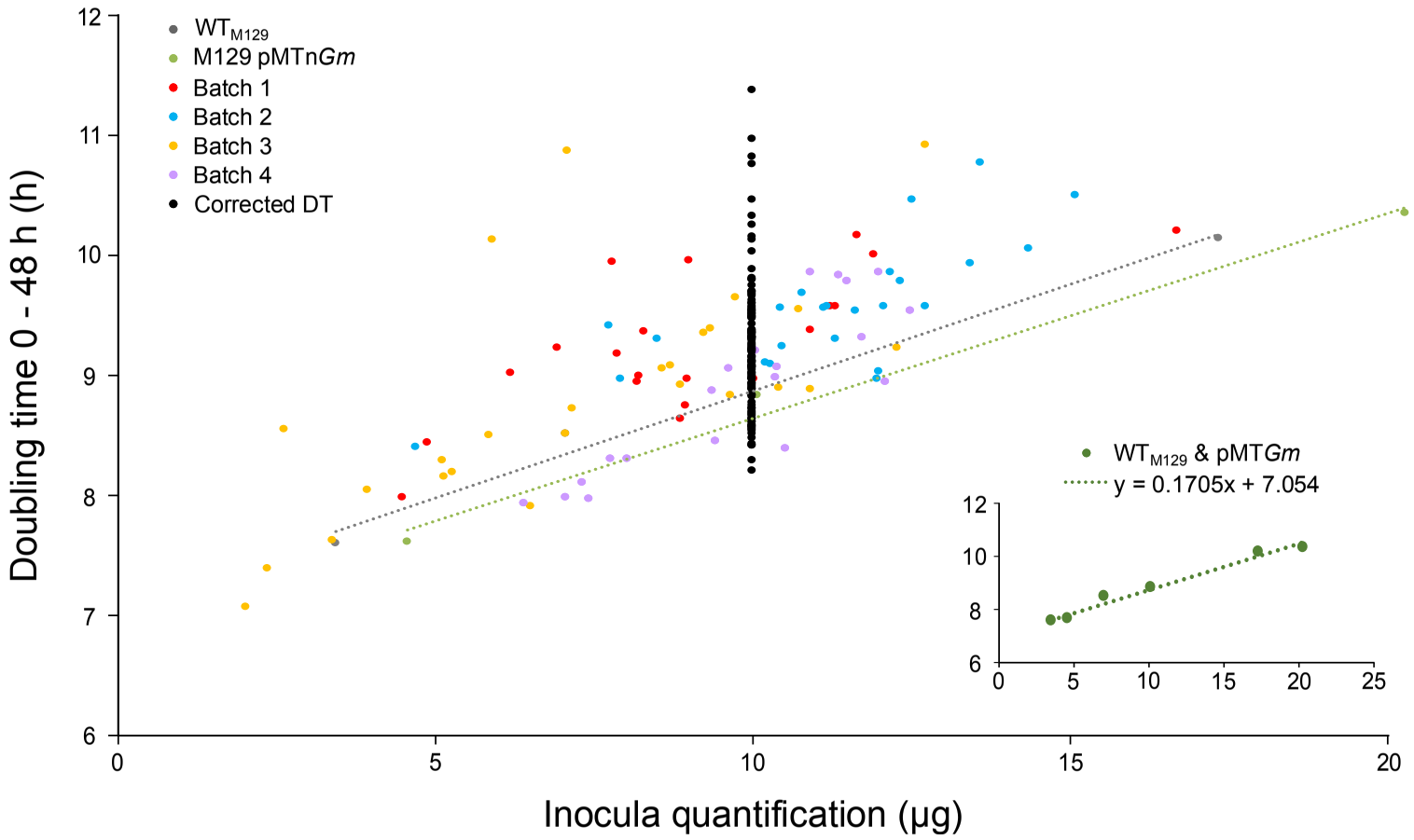


Figure 4.1: Correlation observed between calculated doubling times and inocula with intracellular protein concentrations in semicontinuous cultures. Doubling times and inocula are showed color coded by batch experiment. The Inset shows the correlation between doubling times and inocula for WT_{M129} and M129 pMTn*Gm* and the equation 2.5 for the correction of doubling time calculation. $n = 4$ (two independent experiments with two technical replicates each).

mants to which growth was estimated in the validation and Appendix A, Table A.3 shows estimated doubling times, including the correction for their determination due to the correlation observed between inocula and doubling times (Figure 4.1), as it was previously described for batch 1 in section 2.3.2.1).

The validation was performed mostly for single transformants showing fast growth, and also for some transformants showing slow growth. This way we could determine the dynamic range and the sensitivity of the methods. Also some mutants were measured more than once in different batches to evaluate the reproducibility. The WT_{M129} and M129 pMT85 transformants were measured in each batch as controls. The reproducibility was high within the batch, between batches it was low but consistent to each other, indicating that the calculated doubling times are relative measures that need to be compared with respect to the controls.

Among the slowest and comparing with respect to the negative control transformants for each batch, we found, from the slowest to the fastest, *glpD* DR, *ymdB* OE, *hprK* DR, *spoT* OE, *hepA* OE and interestingly both *whiA* DR and *whiA* OE, being the OE transformants significantly faster than the DR transformants (*P*-value = 0.0016) (Appendix A, Table A.3). These results coincide with the results from the medium-throughput screening, with both growth and metabolic phenotypes, except for *spoT* OE, which were found to grow fast by both growth and metabolic parameters.

Among the 20 fastest and comparing with respect to the negative control transformants for each batch, we found, from the fastest to the slowest, *rpmC* OE, *nusB* OE, *mpn148* OE, *rpsJ* OE, *ptsI* OE, *trpS* OE, *groES* OE, *glpQ* DR, *eno* OE, *spxA* Mut^{G63R}, *rplC* OE, *rpoB* OE, *rpoA* OE, *rimK* OE, *engA* OE, *mraW* OE, *gmk* OE, *gap* OE, *spoT* DR and *deoA* OE (Appendix A, Table A.3). These results coincide with the results from the medium-throughput screening, with both growth and metabolic phenotypes, except for *engA* OE, which were found to grow slow by metabolic parameters.

4.3.3. Generation of multiple transformants

Table 4.4: Prediction of promoter strengths to overexpress *rpoA*. The prediction was calculated using the naive Bayes classifier by Yus et al. (2017)

Promoter	Sequence (5'-3')	Strength prediction
$clpB_{M129}$ _TAPtag	AAATTATGCCTCAAACACAGCAAATTATTGATTACAAAAAGTGGG TGCAAAAGCGCTTGAACTACTTTTTAAAGGCACTGTTCCCTAACA CTTGGAAGATTGCCAAAACAATTCCGTTGTACTAGTTAAGTAACT CAAAACCGTCTTTATAAATGAAAGAGAGCAGCACTTGGGGGTTG GCTCTCTTTTACTTTATAACCACCAAAAAAGGTAAATCTTTAGCCC ACTCGTTAGCACTCAAGCCATTCGAGTGCTAATTTTATATAATTGG CCTATTAACAAAGAAAGGGGATAAAATAGCGGCCGCTA	0.8558
tuf_{M129} _pMT85	TCAGCAATTACAAAAACAAAACAAATAAAAAATAAGGGAATTACC CCCAAGAAGACCTTTTGTGCTAACGCCAGTTTGGCAAATCAAGTT CTGATTTTGCAATTATTTTGCTCCATATGAATTACACTACTCCAAGA ATTATAAGCCTCTCTACAGCTTTATCTCAAACTTATGTAAAATTAGA GACGTAATTCAAACAC	1.4512
$mg438_{G37}$	TAGTATTTAGAATTAATAAAGT	1.3219
$rpmJ_{M129}$	TTAGTTGAAGTGGAATTTTCCCCTTATGATTTAAAGTTAGGTCGAA TAATAGGTAGAAAATAATTAAAATACT	-05320
$rpoBC_{M129}$	CTATCATGATGGCTTTGATTTTCAGCCATTTATCACTGACGTACAAA AATTGATATAATGTAGCAAACTTTACAAAGCCCATTAGCATGCGGC CCTTTTGTCGCTGTTAGTGGCTTTTTACTGTTATTTTTAGCCCTTTT TTGTACACGA	0.6411
$rpoD_{M129}$	AAACAAAAAGAGCTTTTAAAATATTTTAAAAGCATGGTTGAGCAC TAAATAATGTACAATTCCACTTGGCGTGCGTGTATACGCGTGCGCA CTAGACTGA	0.9460

Those genes found to increase the growth rate of *M. pneumoniae* were grouped into three functional categories, namely, transcriptional, translational and metabolic genes, and design the corresponding constructs.

The RNA polymerase (RNAP) of *M. pneumoniae* is constituted by three different subunits, with the composition $\alpha_2\beta\beta$', which are encoded by the *rpoABC* genes, respectively. Since the OE of *rpoB* and *rpoA* in *M. pneumoniae* showed to increase its growth rate, we thought to overexpress the whole RNAP and the sigma factor, despite that the OE of *rpoC* was not tested and that the OE of only the sigma factor, encoded by the *rpoD* gene, decreased the growth rate. To avoid possible toxicity in *M. pneumoniae* due to the excessive overexpression of the whole RNAP, the genes were placed under the control of their own promoters instead of

the *tuf* promoter as was the case for the single transformants. Moreover, we used the orthologous genes from *M. genitalium* G37 in order to differentiate the expression of the endogenous from the exogenous copies by mass spectometry thereby avoiding using tags that may affect the conformation and stability of the proteins (see pBLAST alignments in Appendix A, Figure A.1 - A.4). Furthermore, to avoid toxicity in *E. coli* due to the possible expression of truncated proteins sequestering RNAP subunits, opal (TGA) stop codons were modified for thryptophan (TGG) codons. With these specifications, we initially designed a construct containing the *rpoABCD* genes but this was impossible to build, presumably due to toxicity in *E. coli*, which is used for the cloning steps. Therefore, we had to make some modifications and separate *rpoA* from *rpoBCD*. Thus, the *rpoBCD* genes were cloned into pUC57, to create construct pUC57*RpoBCD*, but these genes were impossible to be subcloned into the minitransposon pMT*GmLox* to be able to transform it into *M. pneumoniae*. Since it was not possible to subclone the pUC57*RpoBCD* in a minitransposon vector, we managed to modify this former vector by adding the gentamicin resistance gene, *aac(6')-aph(2")*, and the inverted repeats flanking the antibiotic resistance gene and *rpoBCD* genes, in this way we obtained the plasmid pUC57IR*GmLox-RpoBCD*, for a detailed description of this molecular cloning, see methods in section 4.2.2.1.

Since *rpoA* had to be separated from *rpoBCD*, we had to clone it in another minitransposon vector. Additionally, we tested different promoter strengths to see which promoter increases more the growth rate of *M. pneumoniae*. First, we predicted the strengths of six promoter regions that we generally use and the own promoters of the RNAP and the sigma factor, *i.e.* $clpB_{\mathrm{M129}}$, tuf_{M129}, $mg438_{\mathrm{G37}}$, $rpmJ_{\mathrm{M129}}$, $rpoBC_{\mathrm{M129}}$, and $rpoD_{\mathrm{M129}}$ (Table 4.4) (Yus et al., 2017). Second, with these predictions, we constructed three plasmids containing the $rpoA_{\mathrm{G37}}$ under the control of the strongest promoter, tuf_{M129}, which is used in most of the constructs with plasmid pMT85; the medium-strength promoter, $clpB_{\mathrm{M129}}$, used in the TAPtag constructs; or the own promoter of $rpoA_{\mathrm{M129}}$, which is the region upstream of $rpmJ_{\mathrm{M129}}$, the first gene of the operon. Finally, $\mathrm{WT}_{\mathrm{M129}}$ was transformed with those three plasmids independently and the doubling times of the resulting transformants were estimated using the intracellular protein at two time

points in semicontinuous cultures method (see methods in section 2.3.2.1). In Appendix A, Table A.3, the corrected doubling times indicate that the OE of rpmJ-rpoA$_{G37}$ leads to significantly faster transformants (*P*-value = 0.0098) than those of clpB-rpoA$_{G37}$ OE and these in turn significantly faster (*P*-value = 6.6×10^{-27}) than tuf-rpoA$_{G37}$ OE transformants.

In order to obtain transformants overexpressing the whole RNAP and the sigma factor, WT$_{M129}$ was transformed with plasmids pUC57IR*GmLox-RpoBCD* and pMT*CmLox-RpoA*. Currently, we are estimating the doubling times of the resulting double transformants.

The OE of genes, which were found to increase the growth rate of *M. pneumoniae* and were grouped into the translational construct were *rpsJ*, *rplC*, *rpmC*, *trpS* and *nusB*. As for the RNAP and sigma factor constructs, the genes were placed under the control of their own promoters, the orthologous genes from *M. genitalium* G37 were used and opal (TGA) stop codons were modified for thryptophan (TGG) codons. With these specifications (see methods in section 4.2.2.1, for a detailed description), we obtained the plasmid pMT*PacLox-5transla*.

Similarly, the OE of genes found to increase the growth rate of *M. pneumoniae* and grouped into the metabolic construct were *engA*, *eno*, *gmk* and *groES*. With the same previous specifications, we obtained the plasmid pMT*TcLox-4metab*. Although, *ptsH* OE was found in the last analysis to grow slower than the negative control, in out first analysis (not shown), it was found to grow faster, so we obtained another plasmid, pMT*TcLox-5metab*, including the *ptsH* OE.

To test whether overexpressing the genes, found to increase the growth rate of *M. pneumoniae*, together would have an additive effect on growth, WT$_{M129}$ was transformed with either plasmid pMT*PacLox-5transla*, pMT*TcLox-4metab* or pMT*TcLox-5metab* and with their negative controls pMT*PacLox* or pMT*TcLox*. Growth curves from these resultant transformants were performed by medium color (pH) change in batch cultures (see methods in section 3.2.4.1) (Figure 4.2) and intracellular protein at two time points in semicontinuous cultures (see methods in section

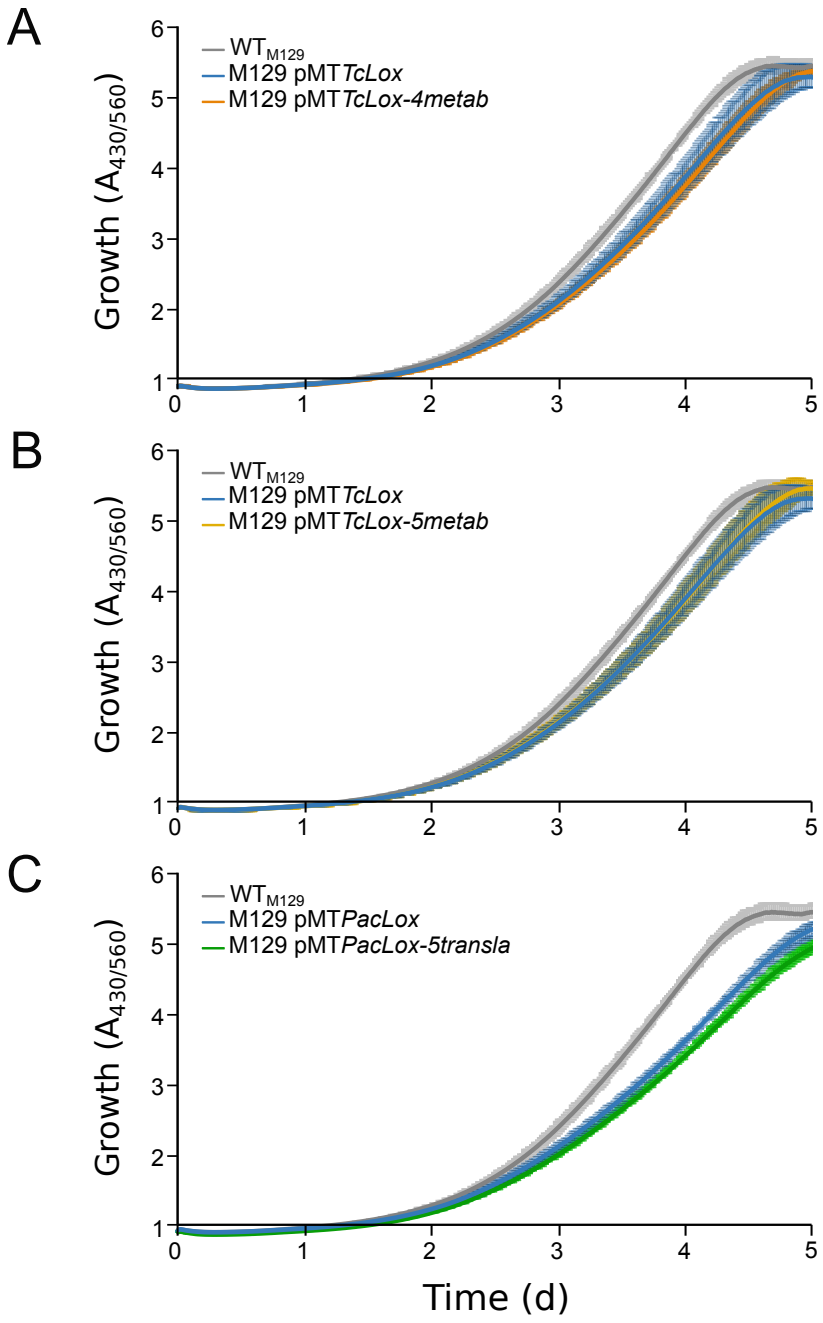


Figure 4.2: Growth curves (GC) of *M. pneumoniae* expressing the *4metab*, *5metab* and *5transla* constructs and their controls. Medium color (pH) change GC are measured by $A_{430/560}$ in batch cultures. $n = 2$ (two independent experiments). Error bars = SD. **(A)** M129 pMT*TcLox-4metab* transformants are shown with their negative control M129 pMT*TcLox* transformants and WT_{M129}. **(B)** M129 pMTn*TcLox-5metab* transformants are shown with their negative control M129 pMT*TcLox* transformants and WT_{M129}. **(C)** M129 pMT*PacLox-5transla* transformants are shown with their negative control M129 pMT*PacLox* transformants and WT_{M129}

Table 4.5: Estimated doubling times (DT) of *M. pneumoniae* expressing the *4metab*, *5metab* and *5transla* constructs and their controls. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5. Values indicate the mean ± SD of two independent experiments with two technical replicates each (n = 4). Statistical significant differences compared to the corresponding control/WT_{M129} or WT_{M129} are indicated: NS: non-significant P-value, * P-value < 0.05, ** P-value < 0.01, *** P-value < 0.001.

	Protein concentration in semicontinuous cultures	
Type of cells	**Doubling time (h)**	**(Corrected DT)**
WT_{M129}	11.04 ± 0.48	(11.58 ± 0.33)
M129 pMT*TcLox* (- ctrl.)	$11.56 \pm 0.13^{***}$	$(11.66 \pm 0.10)^{NS}$
M129 pMT*TcLox-4metab* (4metab)	$11.32 \pm 0.32^{**/NS}$	$(11.46 \pm 0.31)^{*/NS}$
M129 pMT*TcLox-5metab* (5metab)	$11.09 \pm 0.32^{***/NS}$	$(11.41 \pm 0.24)^{***/NS}$
M129 pMT*PacLox* (- ctrl.)	$10.61 \pm 0.25^{***}$	$(10.85 \pm 0.16)^{***}$
M129 pMT*PacLox-5transla* (5transla)	$11.07 \pm 0.15^{***/NS}$	$(11.37 \pm 0.13)^{***/*}$

2.3.2.1), from which the doubling times were estimated (Table 4.5).

The medium color (pH) change, measured at $A_{430/560}$ (Figure 4.2.A) did not show any significant difference with M129 pMT*TcLox-4metab* (4metab) or M129 pMT*TcLox-5metab* (5metab) transformants with respect to their controls. M129 pMT*PacLox-5transla* (5transla) transformants were instead slightly faster than its negative control.

In contrast with the former method, with the intracellular protein concentrations in semicontinuous cultures, the corrected doubling times showed that 4metab and 5metab transformants grow significantly faster than its negative control transformants (P-value = 0.026 and 0.00072, respectively) (Table 4.5). However, 5transla transformants grow significantly slower than its negative control transformants (P-value = 8.1×10^{-11}).

The low additive increase could be the result of using the incorrect negative control plasmids, while four or five genes are overexpressed in the 4metab, 5metab and 5transla transformants, none is overexpressed with the negative control plasmids. This is why, currently, additional negative control plasmids overexpressing fluorescent proteins were constructed and the doubling times of the resulting

transformants will be estimated.

Chapter 5

COMPARATIVE -OMICS DATA OF *MOLLICUTES* SPECIES

5.1. Introduction

In this Chapter, we aimed to identify differently expressed genes of fast- and slow-growing *Mollicutes* species that correlate with fast growth by taking a data-driven approach, in which we compared transcriptomics and proteomics data of fast- and slow-growing *Mollicutes* species. We found the homologous genes among those species and grew them in the same culture media to prepare the samples for transcriptomics and proteomics procedures. I performed different normalization strategies in order to analyze the data and infer relevant conclusion.

5.2. Materials and methods

5.2.1. Bacterial strains and culture conditions

The ten species of the three main phylogenetic groups from the class *Mollicutes* used in this study are shown in Table 5.1. *Mycoplasma* cells were adapted for one passage and grown in SP-4 medium (Tully et al., 1979). Cell cultures were grown at 37°C under agitation (120 rpm).

Table 5.1: *Mollicutes* species used in this study. Some of the genomic features are indicated. The doubling time (DT) is indicated for the same or a closely related strain, calculated from CFU determinations in SP-4 medium unless indicated otherwise.

Species	Genome size (Mb)	rRNA	tRNA	Other RNA	Proteins	Doubling time (h)	Reference	Source/ Collaborator
Pneumoniae phylogenetic group								
1. *M. genitalium* G37	0.58	3	36	4	507	11.9[a]	Lluch-Senar et al. (2010)	-
2. *M. peumoniae* M129	0.82	3	37	315	691	6.5	Furness et al. (1968a)[b]	-
3. *M. gallisepticum* R_{low}	1.01	6	32	1	747	1.2	Beaman and Pollack (1983)[c]	Michael Szostak
4. *M. gallisepticum* R_{high}	1.01	6	32	1	739	1.2	Beaman and Pollack (1983)[c]	Michael Szostak
Hominis phylogenetic group								
5. *M. hyopneumoniae* 232	0.89	3	30	2	657	4.8 - 7.8 h[d]	Calus et al. (2010)[d]	Chris Minion
6. *M. agalactiae* 7784	0.88	6	34	1	666	-	-	Christine Citti
7. *M. bovis* JF4278	1	6	34	1	788	1.9	Beaman and Pollack (1983)[e]	Jetta Bijlsma
Spiroplasma phylogenetic group								
8. *M. capricolum* subsp. *capricolum* California kid	1.01	6	30	6	793	3	Jores et al. (2013)[f]	John Glass
9. *M. mycoides* subsp. *capri* GM12	1.08	6	30	3	850	3.7	Jores et al. (2013)[f]	Carole Lartigue
10. *M. feriruminatoris* G5847	2.03	3	29	3	875	0.5	Jores et al. (2013)[f]	Carole Lartigue

[a] DT calculated from ATP determinations.
[b] *M. pneumoniae* (M. F. Barile, National Institutes of Health) grown in Eaton agent broth enriched with 20% uninactivated horse serum.
[c] *M. gallisepticum* S6 grown in modified Edward medium supplemented with 5% heat-inactivated horse serum and 0.1% L-arginine.
[d] DT calculated from ATP and CCU determinations of various *M. hyopneumoniae* isolates grown in Friis medium.
[e] *M. bovis* PG45 grown in modified Edward medium supplemented with 5% heat-inactivated horse serum and 0.1% L-arginine.
[f] *M. capricolum* subsp. *capricolum*, *M. mycoides* subsp. *capri* and *M. feriruminatoris* G5847 were grown in Mycoplasma standard medium (Mycoplasma experience, Reigate, Surrey, UK).

5.2.2. Growth curves conditions

Growth curves were performed following the intracellular protein concentrations method in batch cultures as previously described in section 2.2.3.1 with some modifications. Briefly, frozen stocks from the 10 *Mycoplasma* species were inoculated in 1:100 dilution in 125 mL sterile polycarbonate erlenmeyer flasks, with vent cap, containing 20 mL fresh SP-4 medium and grown at 37^{o}C under agitation (120 rpm). From five to ten samples were taken at selected time points, depending on the doubling time reported in the literature (Table 5.1). At selected times, 1 mL of cell suspension was taken from the growing cultures, frozen and stored at -70°C for further processing. After collecting all samples, frozen sampled cells were pelleted at 14100 $\times$g for 10 min, medium was removed and pelleted cells were washed with PBS (for 1 liter: 8 g NaCl, 0.2 g KCl, 0.24 g KH_2PO_4, 1.44 g $Na_2HPO_4{\cdot}2H_2O$, pH 7.4). A second cycle of centrifugation and washing was performed, After a third centrifugation, PBS was removed and pelleted cells were suspended in variable volumes of lysis buffer (4% Sodium Dodecyl Sulfate (SDS), 0.1M HEPES), depending on the amount of cells at the selected time points. Cell lysates were kept on ice and disrupted by using a Bioruptor®sonication system (Diagenode, B01010004) with an On/Off interval time of 30/30 sec at high frequency for 10 min. Finally, cell lysates were spun down, pipetted up and down to complete lysis and extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225).

5.2.3. De-novo genome sequencing, assembling and automatic annotations

In order to de-novo sequence the genome of *M. agalactiae* 7784, cells were diluted 1:130 in 125 mL sterile polycarbonate erlenmeyer flasks, with vent cap, containing 20 mL fresh SP-4 medium supplemented with 0.5% sodium pyruvate, pH 7.4 and grown for 4 days at 37^{o}C under agitation (120 rpm). After growth, cells were pelleted at 14100 $\times$g for 10 min, medium was removed and pelleted cells were frozen and store at -70^{o}C. Genomic DNA of was isolated by using the MAsterPure DNA purification kit (Epicentre, Cat no. MCD85201). Library preparation and

DNA sequencing (DNA-seq) were carried out by the Genomics CRG facility. A library of mate-pair of 10 Kb was prepared, following manufacturer's instructions, and validated running a Bioanalyzer (Agilent) DNA 1000 series to check size and quantity of sample. The sequencing run was performed in a Mi-Seq (Illumina) by the Genomics CRG facility.

Genome assembly and automatic annotations were carried out by the Bioinformatics CRG facility. Briefly, initial reads have been trimmed using the program skewer (version 0.2.2) (Jiang et al., 2014) considering trimming mode as mate pair, redistributing reads based on junction information, filtering out very degenerative reads and undetermined pairs, trimming reads to the 3 prime until reaching a quality of 20, removing reads with an average quality below 20 and discarding trimmed reads shorter than 100 bases. Trimmed reads have been then assembled by using SPAdes assembler (version 3.8.2) (Bankevich et al., 2012) and resulting contigs were grouped together into scaffolds with SSPACE (version 3.0) (Boetzer et al., 2011). Finally gaps between assembled contigs were filled by using GapCloser tool from the SOAPdenovo2 package (Luo et al., 2015). One scaffold of 871,489 bases (including 4000 Ns) and 21 pieces of mostly highly covered repeats, whose size range from 128 to 576 bases from the genome of *M. agalactiae* 7784 was obtained.

The complete genome of *M. bovis* JF4278 was gently provided by our collaborators at MSD Animal Health, Boxmeer, The Netherlands.

Automatic annotation of the de-novo assembled genome of *M. agalactiae* 7784 and the provided genome of *M. bovis* JF4278 was performed using GRC tool (Warren and Setubal, 2009) and using *M. agalactiae* 5632 (NC_013948.1) and *M. bovis* PG45 clone MU clone A2 (NC_014760.1) as reference proteomes. Keeping only proteins with a coverage of 90%, the automatic annotation was able to get 645 proteins for *M. agalactiae* 7784, with respect to 791 proteins from *M. agalactiae* 5632 and 695 proteins for *M. bovis* JF4278, out of 779 proteins from *M. bovis* PG45.

5.2.4. Comparative genomics: homology determination

In order to compare transcriptomics and proteomics data of the 10 *Mycoplasma* species used in this study, a homology database with the gene names (with locus_tag as identifiers) of the 10 genomes was performed by our collaborator Pascal Sirand-Pugnet at the Institut National de la Recherche Agronomique (INRA), France.

The gene sets of *M. genitalium* str. G37 (NC_000908.2), *M. peumoniae* str. M129 (NC_000912.1), *M. gallisepticum* str. R_{low} (NC_004829.2), *M. gallisepticum* str. R_{high} (NC_017502.1), *M. hyopneumoniae* str. 232 (NC_006360.1), *M. mycoides* subsp. *capri* str. GM12 (NZ_CP001668.1), *M. capricolum* subsp. *capricolum* str. California kid ATCC 27343 (NC_007633.1), *M. feriruminatoris* str. G5847 (NZ_ANFU00000000) and the closely related strains *M. agalactiae* str. 5632 (NC_013948.1) and *M. bovis* str. PG45 (NC_014760.1) to the species cultured in this study were compared using the MBGD database (http://mbgd.genome.ad.jp) (Uchiyama, 2003).

5.2.5. Growth culture conditions for transcriptomics and proteomics

Frozen stocks from the 10 *Mycoplasma* species were inoculated in 1:100 dilution in either 50 mL Falcon tubes containing 2.5 mL fresh SP-4 medium in duplicate for transcriptomics processing or 125 mL sterile polycarbonate erlenmeyer flasks, with vent cap, containing 20 mL fresh SP-4 medium for proteomics processing and were grown at 37^{o}C under agitation (120 rpm). Samples were taken at the exponential growth phase of each *Mycoplasma* species according to the growth curves performed for each one, obtained in section 5.2.2. Thus, cells were grown for the following times: 5 h for *M. feriruminatoris*, *M. mycoides* subsp. *capri* and *M. capricolum* subsp. *capricolum*; 24 h for *M. gallisepticum* str. R_{low} and str. R_{high}; 48 h for *M. peumoniae* and *M. agalactiae*; 80 h *M. bovis* and *M. hyopneumoniae*; and 96 h for *M. genitalium*.

5.2.6. Transcriptomics

RNA abundances were measured at the determined times of exponential growth phase for each *Mycoplasma* species. At the determined times, the 2.5 mL of cell culture, prepared in section 5.2.5, were collected in 2 mL microcentrifuge tubes and centrifuged at 14100 $\times$g for 10 min. Supernatant was removed and pelleted cells were washed with PBS without disgregating the pellet, PBS was removed and pelleted cells were immediately lysed by resuspension in 500 µL of QIAzol lysis reagent (QIAGEN, Cat. No. 79306). Cell lysates were frozen and stored at -70°C for further processing as previously described in section 3.2.6.

5.2.7. Proteomics

Protein abundances were measured at the determined times of exponential growth phase for each *Mycoplasma* species. At the determined times, the 20 mL of cell culture, prepared in section 5.2.5, were passed through a 25-gauge (G25) syringe needle twice and aliquoted into 1.8 mL portions. Aliquotes were frozen and stored at -70°C for further processing. Four 1.8 mL frozen aliquotes, were thawed and centrifuged at 14100 $\times$g for 10 min, medium was removed and pelleted cells were washed with PBS and two aliquotes were combined to obtain two replicates. A second cycle of centrifugation and washing was performed, After a third centrifugation, PBS was removed and pelleted cells were resuspended in variable volumes of lysis buffer (4% SDS, 0.1M HEPES) for the first replicate and freshly prepared and precooled urea buffer (6M urea, 0.2 M NH4HCO3) for the second replicate. The volumes per species were 60 (for *M. capricolum* subsp. *capricolum*, *M. gallisepticum* str. R_{low} and str. R_{high}, *M. peumoniae*, *M. agalactiae*, *M. hyopneumoniae*, and *M. genitalium*), 100 (for *M. feriruminatoris* and *M. bovis*), or 200 µL (for *M. mycoides* subsp. *capri*). Cell lysates were kept on ice and disrupted by using a Bioruptor®sonication system (Diagenode, B01010004) with an On/Off interval time of 30/30 sec at high frequency for 10 min. Finally, cell lysates were spun down, pipetted up and down to complete lysis and extracted protein was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). Total protein extracts from the replicate lysed with SDS were analyzed by mass spectrometry (MS) by the CRG/UPF Proteomics Unit as

previously described in section 3.2.7.

Sample preparation - In solution-digestion (with urea)

Total protein extracts from the replicate lysed with urea were analyzed by MS by the CRG/UPF Proteomics Unit as following. 15 μg of samples were reduced with 10 mM dithiothreitol (30 nmols, 1 h, 37°C) and alkylated in the dark with 20 mM iodoacetamide (60 nmol, 30 min, 25 °C). The resulting protein extract was first diluted 1:3 with 200 mM NH4HCO3 and digested with 1 μg LysC (Wako, Cat. No. 129-02541) overnight at 37°C and then diluted 1:2 and digested with 1 μg of trypsin (Promega, Cat. No. V5113) for eight hours at 37°C. Finally, the peptide mix was acidified with formic acid and desalted with a MicroSpin C18 column (The Nest Group, Inc) prior to LC-MS/MS analysis. LC-MS/MS analysis was performed as previously described in section 3.2.7.

Data Analysis

Data analysis was performed as previously described in section 3.2.7. Samples were searched against the *Mycoplasma pneumoniae* str. M129 HomoConTrans19 database and the databases generated for each one of the nine remaining *Mycoplasma* species. The databases were generated for all putative proteins longer than 19 amino acids (after in silico translation of each genome in the six putative frames) with a list of common contaminants and all the corresponding decoy entries.

Normalization of proteomics data

In order to compare the abundance of proteins between fast- and slow-growing *Mollicutes* species, the log2 values of the areas were normalized by eight different ways. The log2 of the areas were normalized by (1) the sum of the log2 of all detected proteins, (2) the mean of the log2 of all detected proteins, (3) the median of the log2 of all detected proteins, (4) the sum of log2 corresponding to the components of the RNA polymerase, (5) the sum of the log2 of the ATPase complex proteins, (6) the sum of the log2 of all chaperon proteins, (7) the sum of the log2 of all ribosomal proteins, and (8) the sum of the log2 of proteins in the glycolysis pathway. Data analysis was done by using a homemade R studio-based script.

5.3. Results and discussion

5.3.1. Growth curves of *Mollicutes* species

The species selected for this study belong to the three main phylogenetic groups from the class *Mollicutes*. According to literature data, we classified them as fast growers, with doubling times greater than 4 hours, or as slow growers, with doubling times lower than 4 hours (Table 5.1). Thus, we have three slow growers and seven fast growers (*M. agalactiae* is supposed to grow very similar to *M. bovis*). To identify differently expressed genes of fast- and slow-growing *Mollicutes* species that correlate with fast growth, we first determined experimentally the exponential growth phase in SP-4 medium of all the ten *Mycoplasma* species. For this purpose, we cultured them in the laboratory with the same culture conditions and quantified intracellular protein at selected time points (Figure 5.1).

The resulting growth curves of *M. bovis* and *M. agalactiae* showed that these two species grow slow, presumably for the lack of sodium pyruvate, which is reported as a medium supplement for their growth (Hegde et al., 2015a). In contrast, *M. mycoides* and *M. capricolum* turned out to grow faster than expected, as fast as *M. feriruminatoris*. The determined times for the exponential growth phase were 5 h for *M. feriruminatoris*, *M. mycoides* and *M. capricolum*; 24 h for *M. gallisepticum* str. R_{low} and str. R_{high}; 48 h for *M. peumoniae* and *M. agalactiae*; 80 h *M. bovis* and *M. hyopneumoniae*; and 96 h for *M. genitalium*. However, these exponential times are approximate as we have only one biological replicate, and we used them only as guides to decide when to measure the corresponding samples.

5.3.2. Comparative genomics: homology determination of *Mollicutes* species

The genomes of *M. agalactiae* str. 7784 and *M. bovis* str. JF4278 had not been sequenced before. Therefore, the genome of *M. agalactiae* 7784 was de-novo sequenced and assembled and the sequence of *M. bovis* JF4278 genome was obtained from our collatorators in MSD Animal Health. Both *M. agalac-*

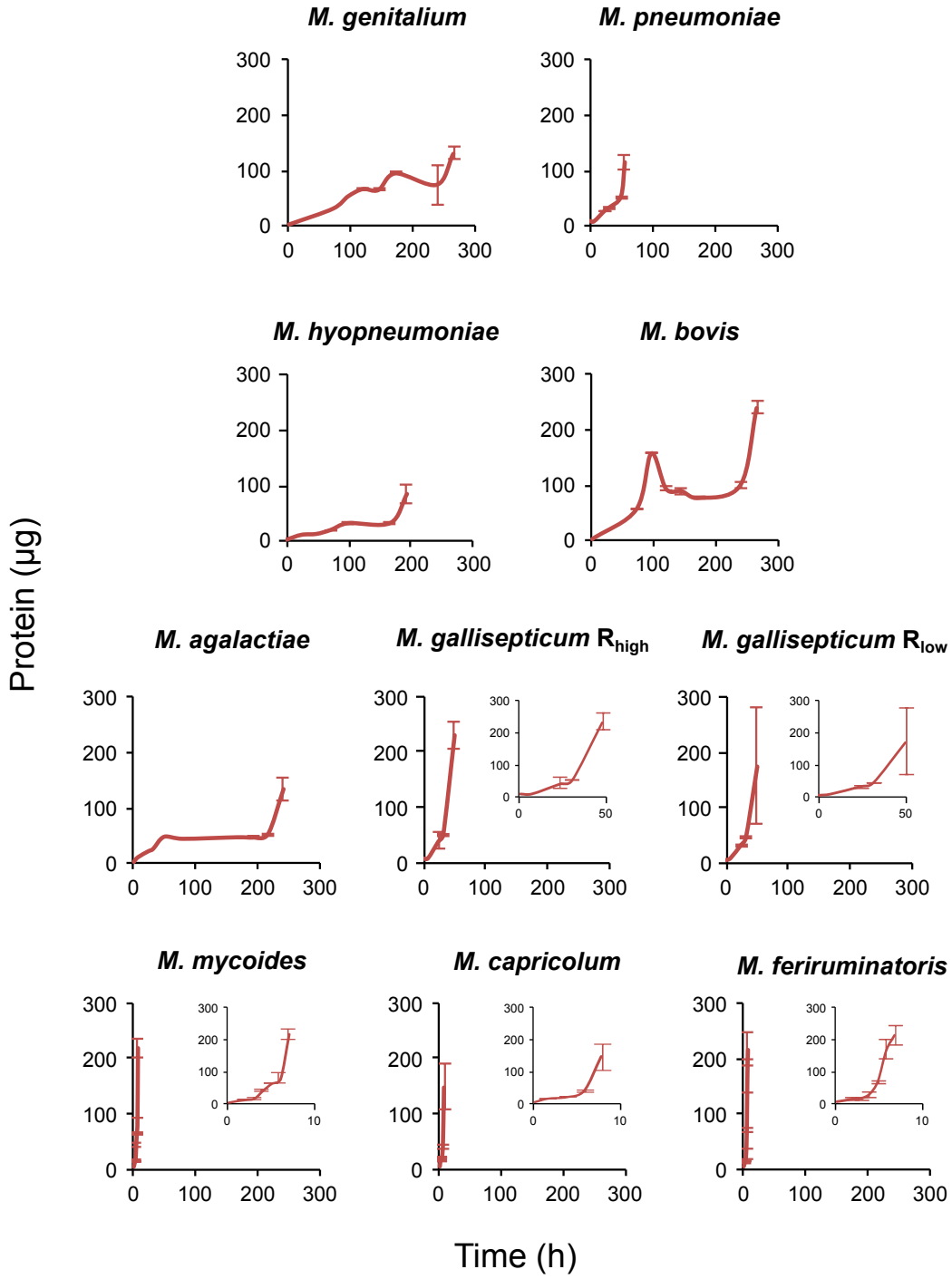


Figure 5.1: Growth curves (GC) of *Mollicutes* species. Intracellular protein GC, cells were grown in SP-4 medium in batch cultures at 37°C under agitation (120 rpm). The insets show the GC in a shorter time scale for the fast-growing species. $n = 2$ (one independent experiment with two technical replicates). Error bars = SD.

tiae 7784 and *M. bovis* JF4278 complete genomes were automatically annotated using *M. agalactiae* 5632 (NC_013948.1) and *M. bovis* PG45 clone MU clone A2 (NC_014760.1) as reference proteomes. For *M. agalactiae* 7784, 645 proteins were obtained out of 791 proteins from *M. agalactiae* 5632 and for *M. bovis* JF4278, 695 out of 779 proteins from *M. bovis* PG45.

In order to compare transcriptomics and proteomics data of the 10 *Mycoplasma* species used in this study, a homology database with the genes of the 10 genomes was performed. A total of 2069 clusters of orthologs were identified corresponding to the pan genome of this group of mycoplasmas. A core genome of 257 gene clusters was identified for the ten species.

5.3.3. Comparative transcriptomics of *Mollicutes* species

The samples for transcriptomics were obtained at the exponential growth phases of each *Mollicutes* species, at the determined times described in section 5.3.1 and are currently being analyzed.

5.3.4. Comparative proteomics of *Mollicutes* species

The samples for proteomics were obtained at the determined times of the exponential growth phases of each *Mollicutes* species. The samples were processed either with SDS or in solution with urea prior to mass spectrometry. The advantage of using SDS is that it can solubilize better membrane proteins, but the main drawback is that its processing involves the use of a 30 kDa molecular weight cut-off filter during the Filter-Aided Sample Preparation (FASP). In solution with urea on the other hand, does not require a MW filter but could do less well in solubilizing integral membrane proteins. Thus we decided to performed both preparations and analyze both results in parallel. Comparing the number of detected proteins in each procedure, we found that the detection level was very similar for all species except for *M. hyopneumoniae* and *M. agalactiae* (Figure 5.2), to which the number of detected proteins was very low (15% and 48% of the total number of annotated proteins, respectively). In contrast, for the other *Mollicutes* species, the percentage of detected proteins ranged between 72% (for *M. genitalium*) and

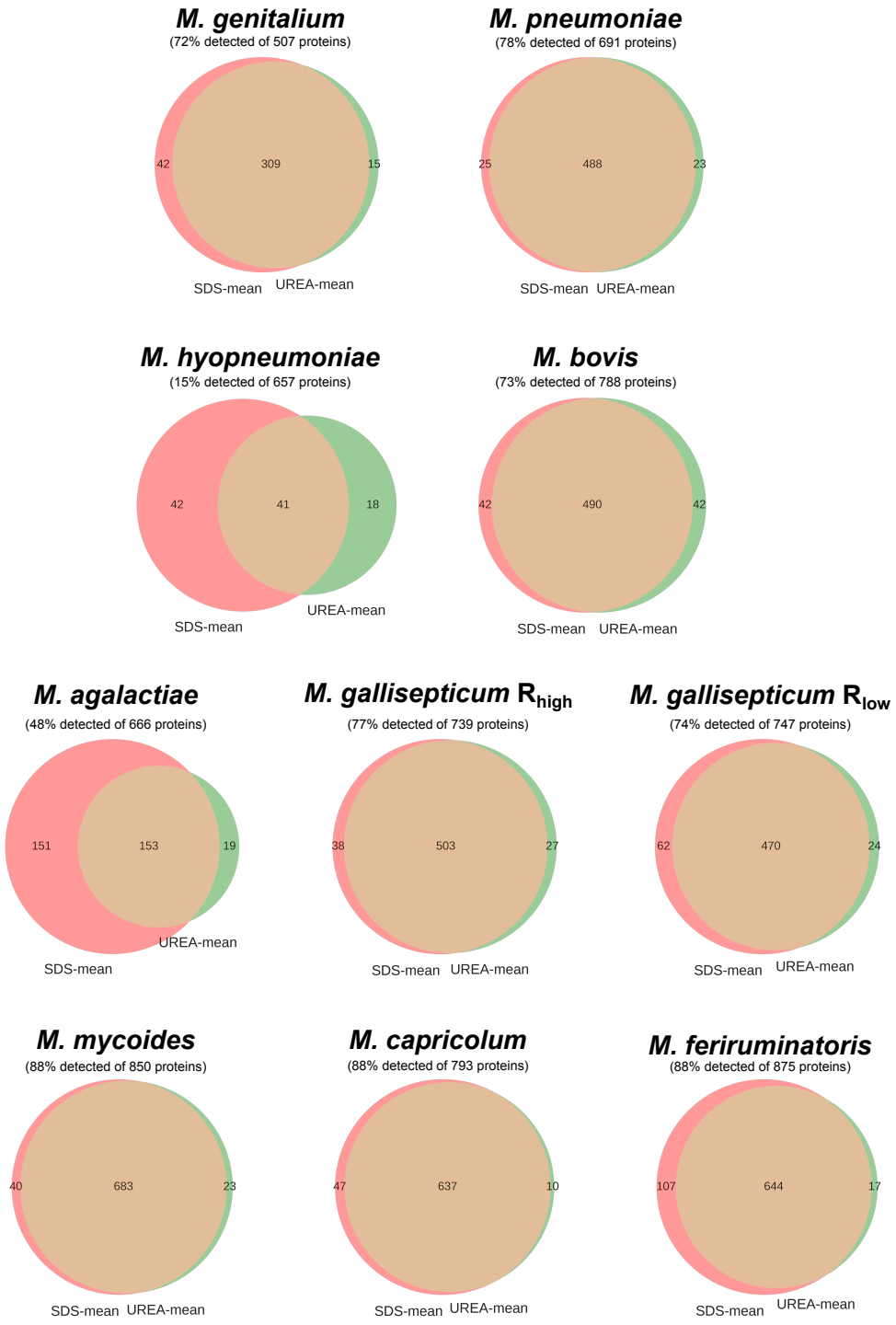


Figure 5.2: Venn diagrams of proteomics data of *Mollicutes* species. Mean of number of detected proteins of two replicates processed with 4% SDS, 0.1M HEPES (SDS-mean) and 6M urea, 0.2 M NH4HCO3 (UREA-mean). Percentage of detected protein with respect to the total annotated proteins of each *Mollicutes* species is indicated in brackets.

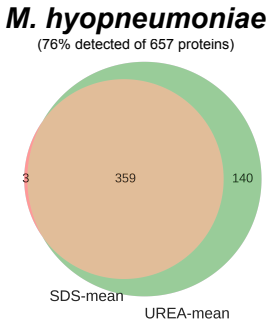


Figure 5.3: Venn diagram of proteomics data of *M. hyopneumoniae*. Mean of number of detected proteins of two replicates processed with 0.1% SDS, 0.01M PBS (SDS-mean) and 6M urea, 8% triton (UREA-mean). Percentage of detected protein with respect to the total annotated proteins of each *Mollicutes* species is indicated in brackets.

88% (for *M. mycoides*, *M. capricolum*, and *M. feriruminatoris*) of the total number of annotated proteins. The low number of detected proteins in *M. agalactiae* might be caused to the lack of sodium pyruvate in the medium, which is required for efficient growth. However, it was not clear the reason behind the low number of detected proteins for *M. hyopneumoniae*, since the protein yield was very similar to that of *M. capricolum* and even higher to that of the two *M. gallisepticum* strains (R_{high} and R_{low}). Due to the extremely low number of detected proteins for *M. hyopneumoniae*, we took other proteomics data processed similarly to the other species (Figure 5.3). In this data, 76% of the total number of annotated proteins were detected. Comparing log2 of detected proteins areas of SDS and urea samples, we found that in general, the abundances of detected proteins is higher in samples processed with urea than with SDS (Figure 5.4.A.). The difference of areas in the two preparations was especially high in the data from *M. hyopneumoniae*, which are caused by the difference in the number of detected proteins, 77% of detected proteins in samples with urea against 55.1% with SDS, from which 54.6% are common.

In order to compare the abundance of proteins among the fast- and slow-growing *Mollicutes* species, we performed several types of normalizations of the log2 of all detected protein areas and then took only the proteins encoded by the 257 orthologous genes found among the ten *Mollicutes* species (Figure 5.4.B-I). Certain nor-

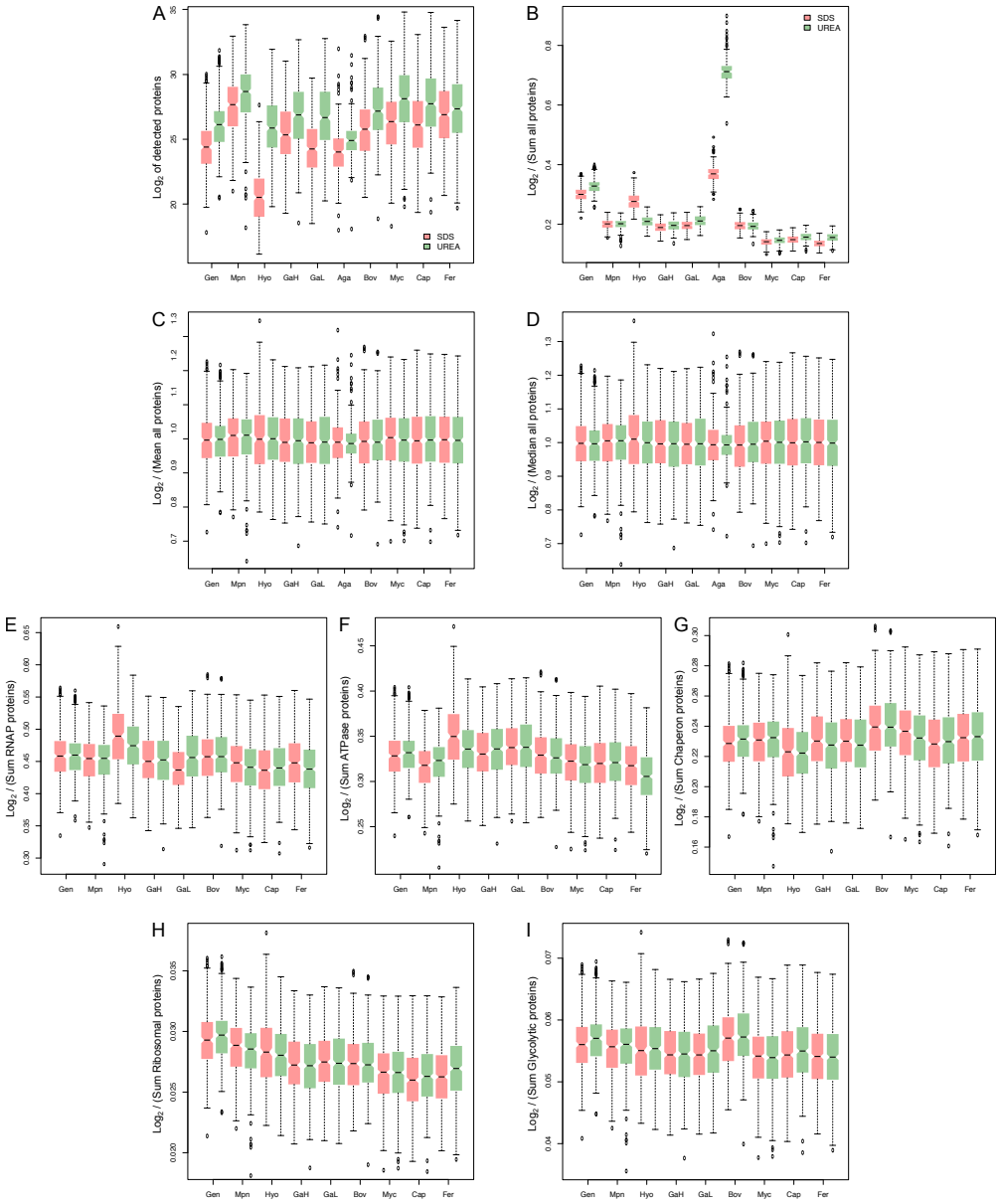


Figure 5.4: Box plots of proteomics data encoded by the 257 orthologous genes among the 10 *Mollicutes* species. Every box corresponds to the mean of two replicates processed with 4% SDS, 0.1M HEPES (SDS), indicated in red, and 6M urea, 0.2 M NH4HCO3 (UREA), indicated in green. *Mollicutes* species are indicated as Gen for *M. genitalium*, Mpn for *M. peumoniae*, Hyo for *M. hyopneumoniae*, GaH for *M. gallisepticum* R_{high}, GaL for *M. gallisepticum* R_{low}, Aga for *M. agalactiae*, Bov for *M. bovis*, Myc for *M. mycoides* subsp. *capri*, Cap for *M. capricolum* subsp. *capricolum*, and Fer for *M. feriruminatoris*. **(A)** Log2 of the areas of detected proteins of the 257 orthologous genes. **(B-I)** Different normalizations of those log2 areas by: **(B)** the sum of all detected proteins, **(C)** the mean of all detected proteins, **(D)** the median of all detected proteins, **(E)** the sum of proteins of the RNP polymerase, **(F)** the sum of proteins of the ATPase, **(G)** the sum of chaperon proteins, **(H)** the sum of ribosomal proteins, and **(I)** the sum of proteins of the glycolysis pathway.

malizations cannot be used to compare between species because they are highly affected by number of detected proteins and their abundances. For instance, the normalized data by the sum of all detected proteins of *M. agalactiae* are significantly high (Figure 5.4.B) because the sum of all detected proteins is very low due to the low number of detected proteins (48%). In contrast, the normalized data by the sum of all detected proteins of *M. mycoides*, *M. capricolum*, and *M. feriruminatoris* is very low because the sum of all detected proteins is high due to the high number of detected proteins (88%). The normalized data by the mean or the median make the distributions of the proteins much more similar (Figure 5.4.C-D) and thereby allowing comparison between the abundances of determined proteins.

We also performed unusual normalizations by the sum of proteins that make part of complexes or pathways, *i.e.* the RNA polymerase complex (Figure 5.4.E), the ATPase complex (Figure 5.4.F), the chaperon complex (Figure 5.4.G), the large and small subunit of the ribosome (Figure 5.4.H) and the glycolytic pathway (Figure 5.4.I). Interestingly, we observed some tendencies in these normalizations. The normalized data by the ribosomal proteins decreases as the doubling time decreases meaning that ribosomal proteins are more abundant in fast-growing species than in slow-growing species, hence presumably more number of ribosomes. Interestingly, this tendency is also slightly observed in the normalized data by the glycolytic proteins, except that *M. bovis* breaks such tendency, although the glycolytic proteins of *M. bovis* could have been downregulated due to the lack of sodium pyruvate, which has also been reported to be an important medium supplement for its growth (Sulyok et al., 2014).

Regarding other proteins we are currently evaluating their relative abundance.

Chapter 6

DISCUSSION AND CONCLUDING REMARKS

Understanding how cells work, how they grow and divide has been one of the fundamental questions in bacterial physiology and nowadays in systems biology and not only in academy but also in industry. The ability to control bacteiral growth is critical for biotechnological applications. Despite of the great advances achieved in the last years on the knowledge of external (environmental) and internal (genetic) factors that regulate growth rates, we lack a full understanding on the mechanisms underlying such regulation. Therefore, we are unable to genetically modified an organisms in a rational manner to change its growth rate and as Richard Feynman said "What I cannot create, I do not understand".

Systems biology studies on *M. pneumoniae* have considerably increased our knowledge on this genome-reduced bacterium, however, more questions need to be addressed and further investigated to have a clearer picture of its cellular activities thereby helping us to tune its genome to reach maximal growth.

The main objective of this thesis was to identify the genetic factors that regulate the growth rate in *M. pneumoniae* and in other *Mollicutes* species to ultimately increase the growth rate of *M. pneumoniae*. Nevertheless, I dedicated Chapter 2 to describe the difficulties, discrepancies and challenges in measuring growth

rates in *M. pneumoniae*, since in order to determine reliable doubling times, we must have a consistent, rapid and reproducible method and specially capable of detecting statistical significance in small differences of growth rates. The difficulty in determining cell counts has, for long time, been known in the field of Mycoplasmology and, in the literature, there are different opinions on what is the best/convenient method to measure growth rates in *Mollicutes* species. We have developed two different methods to measure growth rates: the first one, based on intracellular protein measurements in semicontinuous cultures and the second one, based on the microcolony growth by automated high-throughput time-lapse microscopy. The first method, the intracellular protein measurements in semicontinuous cultures, was developed in response to the necessity of balanced, steady state exponential growth in growth rate determinations (Schaechter, 2006, 2015; Thomas, 2015) and was used throughout the thesis to obtain specific growth rates, together with additional regularly used methods to have higher certainty of our findings. The second method was developed to measure growth rates in a high-throughout manner, in consistency with systems biology approaches, since such a method has not been implemented in *Mollicutes* species. This second method was described as a proof of concept to demonstrate its feasibility and its great potential, yet further optimization needs to be done before it can be used in a high-throughput manner.

In the subsequent chapters, I entered in the matter of the main goal of this thesis: how to increase growth rate in M. pneumoniae. We used two approaches, a rational one based on data in the literature, and an unbiased one using a large collection of mutants stored in the group. Throughout the thesis, we found that the factors that successfully increased the growth rate of *M. pneumoniae* were the widely studied genes involved in protein synthesis, namely, rRNA operons, tRNA and r-proteins genes, from which their copy numbers or abundances are also well known to correlate with growth rates. We have found that we could increase the growth rate of *M. pneumoniae* by increasing rRNA gene copy number up to three copies, which seems to result in an increase in expression of some of the lowly expressed ribosomal proteins. However, the increase in growth rate is not lineal with the number of rRNA operon copies and above three, in fact, we see a decline

in growth rate. This could be because the absolute rRNA expression level is reduced in transformants containing more than three rRNA operon copies (*i.e. 4rrn* and *5rrn*) and therefore there is a negative feedback mechanism of rRNA synthesis, which would be activated in these transformants. Alternatively, there could be a limiting factor that prevents an increase in r-proteins when rRNA increases above a certain threshold. In *E. coli*, there is a translational negative feedback autogenous regulation of r-proteins (for a review see Aseev and Boni (2011)). In *E. coli*, the regulation of ribosome biogenesis occurs at the level of rRNA synthesis, thus, the synthesis of r-proteins is adjusted to the accumulating rRNA. R-protein operons encode their own translational repressor acting on the polycistronic r-protein mRNA, thereby inhibiting the synthesis of all proteins encoded by the mRNA (Nomura, 1999). The accumulation of rRNA releases the repression and the repressor r-proteins are incorporated into new synthesized ribosomes whereby operons continue to express. When rRNA synthesis drops, the repressor concentrations increase and the operons are repressed. If this model applied in *M. pneumoniae* and assuming that the absolute rRNA expression increases with increasing rRNA operon copies, the rRNA of transformants containing more than 3 copies would assemble with the repressor r-proteins thereby upregulating their translation to be incorporated in new ribosomes. Since this is not what we observed with our results, there could be, however, a possible reasoning to explain why *E. coli*'s model does not fit in the genome-reduced bacterium *M. pneumoniae*. The possible explanation is the deficiency of the nucleotide (NTP) pool, specially that of purines, and the relatively low adenylate energy charge and ATP concentration found in *Mollicutes* species (Beaman and Pollack, 1983), characterized by the lack of tricarboxylic acid cycle enzymes, quinones, and cytochromes. The deficiency in GTP and ATP concentrations is well known to decrease rRNA transcription (Gaal et al., 1997). Beaman and Pollack (1983) found that the ATP concentration ranged from 45% to 63% of the total adenylate nucleotide pool in *Mollicutes* species, in contrast to 97% in *Acholepalsma laidlawii* and 92% in *E. coli*. Particularly in *M. pneumoniae*, intracellular ADP and ATP concentrations are on average three times lower than in *E. coli* (Wodke, 2012). Consequently, it could be that *M. pneumoniae* cannot afford synthesis and upregulation of r-proteins in *4rrn* and *5rrn* transformants, since protein synthesis is estimated to

account for 50% of the energy consumption of a rapidly growing bacterial cell (Russell and Cook, 1995). Alternatively, we could contemplate that some of the r-protein operons in *M. pneumoniae* have constant expression that is not autoregulated and therefore at a certain increase in rRNA, the available r-proteins for assembly will not be enough. Finally, even in a model were all r-proteins were autoregulated and there was no limitation of ATP and GTP, we could contemplate that the transcription machinery of *M. pneumoniae* cannot cope with the higher demand required by having five rRNA operon copies.

Another important confirmation of the genetic factors regulating growth rates was the concentration of the RNA polymerase (RNAP) subunits. The *rpoB* OE and *rpoA* OE were found to increase *M. pneumoniae* growth rate (see Chapter 4). In *E. coli*, the average numbers of core RNAP ($\alpha_2\beta\beta$') molecules per cell were found to increase from about 1,800 to 10,200 between growth rates of 0.6 and 3.0 doublings/h (Dennis and Bremer, 2008). In *E. coli*, the α subunit is in excess and the amount of core enzyme is limited by the amount of β and β' subunits (Engbaek et al., 1976). The mechanism of synthesis of the β and β' subunit is autogenously regulated at both the transcriptional and translational level (Dykxhoorn et al., 1996). *rpoB* and *rpoC* genes are located in the *rplKAJLrpoBC* r-protein-RNA polymerase operon, thus, the translational negative feedback mechanism is the one described before for the expression of the r-proteins: increasing concentrations of either β and β' were found to repress the expression of both *rpoB* and *rpoC* genes. At the level of transcription, its synthesis is controlled by the termination-antitermination at an attenuator in front of *rpoB*, this attenuator stops 80% of the transcripts coming from the promoters upstream of *rplK* and *rplJ*, therefore *rpoBC* gene expression is the readthrough at the attenuator (Downing and Dennis, 1991). Moreover, it was found that the increase in the expression of the holoenzyme ($\alpha_2\beta\beta'\sigma^{70}$) repressed the transcription of *rpoBC* (Dykxhoorn et al., 1996). In *M. pneumoniae*, the operon organization is different, *rpoB* and *rpoC* genes consitute their own operon and *rpoA* is found in the *rpmJMKrpoArplQ* r-protein operon. Our results suggest that both α and β subunits are limiting the number of core RNAP enzyme molecules as either the overexpression of α or β subunit increased the growth rate of *M. pneumoniae* by presumably upregulating

the core RNAP enzyme. Increasing RNA polymerase concentrations in *E. coli* was shown to switch bacterial growth from zero to the maximal growth rate (Izard et al., 2015).

Following the same idea of transcriptional regulation, *spoT* DR (gene disruption) transformants were found as well to increase slighly but significantly *M. pneumoniae* growth rate. In Gram-positive bacteria, the RelA/SpoT protein homologue synthesizes and degrades the alarmones guanosine tetraphosphate and guanosine pentaphosphate (referred as (p)ppGpp). (p)ppGpp are involved in regulating growth, the stringent response to amino acid starvation stress as well as other different stress responses in bacteria (Hauryliuk et al., 2015). During amino acid starvation, (p)ppGpp are synthesized and inhibit transcription of rRNA and r-protein genes, but activate transcription of amino acid biosynthesis genes. To regulate transcription, (p)ppGpp interact directly with RNAP in *E. coli* while in other bacteria, including *B. subtilis*, (p)ppGpp downregulate transcription without directly interacting with RNAP but rather by reducing the cellular pool of GTP. This reduction occurs because the production of the alarmone consumes GTP and because the alarmone directly inhibits the enzymes responsible for GTP synthesis, among them the GMP kinase (Gmk), which is strongly inhibited (Hauryliuk et al., 2015). This brings into the picture the *gmk* OE transmormants, which were also found to increase *M. pneumoniae* growth rate. These transformans turned out to grow faster because the overexpresion of *gmk* would presumably lead to an increase in the cellular GTP pool, which allows the transcription of rRNA and r-protein genes thereby boosting growth. However, the slight increase of growth rate in *spoT* DR transformants, might be an artifact, because the absence of alarmone in these transformants would only keep constant the growth rate and not increase it. Since the estimated doubling times were relative to the negative control transformants, a logical explanation is that those control transformants decreased its growth rate in response to a stress, meaning that they were not in balanced growth conditions (steady state exponential growth), and for this reason we observed such an increase in growth rate in *spoT* DR transformants. In constrast, our results of *spoT* OE, which reduce the growth rate, are consistent with the literature (Schreiber et al., 1991), indicating that the overexpression elevates the

(p)ppGpp levels thereby decreasing the growth rate.

Interestingly both *whiA* OE and DR were found to grow slow. WhiA has been characterized in *M. pneumoniae* as a transcription factor that represses the transcription of the main r-protein operon, which contains 23 r-proteins (Llorens Rico, 2016). Therefore, it was expected that *whiA* OE transformants decreased the growth rate, because they the transcriptional repression of the main r-proteins operon would decrease significantly the number of ribosomes synthesized. However, it was surprising that *whiA* DR transformants had the same effect in growth rate.

The knowledge acquired in this thesis have contributed to better understand the determinants of growth in genome-reduced bacteria, which are not that simple as Balish (2014) stated "Mycoplasmas have a long-standing reputation for being something they are not: exceptionally simple".

These determinants of faster growth found in *M. pneumoniae* will be experimentally implemented in a *M. pneumoniae* chassis that is being designed and developed in the laboratory to use as an attenuated live vaccine. Moreover, all the pathogenic genes will be deleted in this chassis in order to be used in animals. Also, these determinants will be used *in silico* to complement the first whole-cell mathematical model for *M. pneumoniae* that is currently been developed in the team. The development of these kind of models is of great importance because of their predictive power of emergent cellular behaviors, which will contribute for the understanding of bacterial growth.

Appendix A

SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Table A.1: Gene product functions of *M. pneumoniae*

ID	Gene Name	Protein Name	Functional category
MPN001	dnaN	DNA polymerase III subunit beta (EC 2.7.7.7)	DNA synthesis/modifications/repair/conformation
MPN003	gyrB	DNA gyrase subunit B (EC 5.99.1.3)	DNA synthesis/modifications/repair/conformation
MPN006	tmk	Thymidylate kinase (EC 2.7.4.9) (dTMP kinase)	Nucleotide metabolism/Salvage
MPN007	holB	DNA polymerase III subunit delta'	DNA synthesis/modifications/repair/conformation
MPN009	yabD	Uncharacterized deoxyribonuclease MPN_009 (EC 3.1.21.-)	DNA synthesis/modifications/repair/conformation
MPN011	mpn011	Conserved hypothetical lipoprotein MPN_011	Lipoprotein
MPN012	mpn012	Conserved hypothetical protein MPN_012	Cytoskeleton
MPN016	rimK2	Ribosomal protein S6 modification enzyme	rRNA and rProtein modifications/processing
MPN017	mtd1	Methylenetetrahydrofolate dehydrogenase (EC 1.5.1.5); Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)	Vitamin/Folate metabolism + Redox balance

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN018	pmd1	Putative ABC transporter ATP-binding protein MPN_018	Lipid/CoA metabolism
MPN019	msbA	Putative ABC transporter ATP-binding protein MPN_019	Lipid/CoA metabolism
MPN022	pip	Putative proline iminopeptidase (PIP) (EC 3.4.11.5) (Prolyl aminopeptidase) (PAP)	Proteases
MPN023	metS	Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine–tRNA ligase) (MetRS)	tRNA synthases
MPN024	rpoE	DNA-directed RNA polymerase delta subunit	RNA synthesis
MPN025	tsr	Fructose-bisphosphate aldolase (FBP aldolase) (FBPA) (EC 4.1.2.13)	Glycolysis
MPN026	engD	GTP-dependent nucleic acid-binding protein engD	Protein synthesis factors
MPN029	efp	Elongation factor P (EF-P)	Protein synthesis factors
MPN030	nusB	Transcription termination/antitermination protein NusB	RNA synthesis
MPN036	mpn036	Conserved hypothetical protein MPN_036	Conserved hypothetical proteins
MPN043	glpF	Glycerol uptake facilitator protein	Lipid/CoA metabolism
MPN045	hisS	Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine–tRNA ligase) (HisRS)	tRNA synthases
MPN047	pncB	Nicotinate phosphoribosyl transferase (EC 2.4.2.12)	Vitamin/Folate metabolism
MPN052	bmpA	Basic membrane protein, probable purine/cytidine ABC transporter substrate-binding protein	Nucleotide metabolism/ Salvage
MPN053	ptsH	Phosphocarrier protein HPr (EC 2.7.11.-) (Histidine-containing protein)	Glycolysis
MPN056	potB	Spermidine/putrescine transport system permease protein potB homolog	DNA synthesis/modifications/repair/conformation
MPN057	potC	Spermidine/putrescine transport system permease protein potC homolog	DNA synthesis/modifications/repair/conformation
MPN058	potD	Spermidine/putrescine transport system permease protein potD homolog	DNA synthesis/modifications/repair/conformation

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN059	tsaD	tRNA N6-adenosine threonylcarbamoyltransferase; homology to endoglycopeptidase	tRNA modifications/processing
MPN060	metX	S-adenosylmethionine synthetase (EC 2.5.1.6) (Methionine adenosyltransferase) (AdoMet synthetase) (MAT)	DNA synthesis/modifications/repair/conformation
MPN062	deoD	Purine nucleoside phosphorylase deoD-type (PNP) (EC 2.4.2.1)	Nucleotide metabolism/ Salvage
MPN065	cdd	Cytidine deaminase (CDA) (EC 3.5.4.5) (Cytidine aminohydrolase)	Nucleotide metabolism/ Salvage
MPN068	secE	Preprotein translocase subunit secE	Protein secretion
MPN070	mpn070	Uncharacterized protein MPN_070	Conserved hypothetical proteins
MPN071	rsmI	Ribosomal RNA small subunit methyltransferase (EC 2.1.1.198) (16S rRNA 2'-O-ribose C1402 methyltransferase) (rRNA (cytidine-2'-O-)-methyltransferase RsmI)	rRNA and rProtein modifications/processing
MPN075	ywdF	Glycosyltransferase, group 2 family protein (EC:2.4.-.-)	Glycocalyx
MPN076	uhpT	Hexose phosphate transport protein	Lipid/CoA metabolism
MPN078	fruA	PTS system fructose-specific EIIABC component (EIIABC-Fru) (2.7.1.69)	Glycolysis
MPN079	fruK	Fructose 1-phosphate kinase (EC 2.7.1.56)	Glycolysis
MPN084	mpn084	Conserved hypothetical lipoprotein MPN_084	Lipoprotein
MPN090	mpn090	Uncharacterized protein MPN_090	Conserved hypothetical proteins
MPN092	mpn092	Putative mgpC-like protein MPN_092	Adhesins/attachment organelle
MPN093	mpn093	Putative mgpC-like protein MPN_093	Adhesins/attachment organelle
MPN094	mpn094	UPF0134 protein MPN_094	Cytoskeleton
MPN095	mpn095	Uncharacterized amino acid permease	AA metabolism
MPN099	mpn099	Putative adhesin P1-like protein MPN_099	Adhesins/attachment organelle
MPN101	mpn101	Uncharacterized protein MPN_101	Conserved hypothetical proteins
MPN102	mpn102	Putative mgpC-like protein MPN_102	Adhesins/attachment organelle
MPN104	mpn104	Uncharacterized protein MPN_104	Cytoskeleton
MPN109	mpn109	Uncharacterized protein MPN_109	DNA synthesis/modifications/repair/conformation

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN118	rnhC	Ribonuclease HIII	RNA degradation + DNA synthesis/modifications/repair/conformation
MPN120	grpE	Heat shock protein GrpE	Protein homeostasis
MPN121	mpn121	Uncharacterized protein MPN_121	Conserved hypothetical proteins
MPN126	ysnB	Putative metallophosphoesterase YsnB (EC 3.1.4.-)	Metabolism homeostasis
MPN128	mpn128	Uncharacterized protein MPN_128	Conserved hypothetical proteins
MPN130	mpn130	Uncharacterized protein MPN_130	Conserved hypothetical proteins
MPN132	mpn132	Putative adhesin P1-like protein MPN_132	Adhesins/attachment organelle
MPN133	mpn133	Ca2+-dependent cytotoxic nuclease of Mycoplasma	Nucleotide metabolism/ Salvage
MPN135	ugpA	sn-glycerol-3-phosphate transport system permease protein ugpA	Nucleotide metabolism/salvage
MPN136	ugpE	sn-glycerol-3-phosphate transport system permease protein ugpE	Nucleotide metabolism/salvage
MPN142	orf6	Mgp-operon protein 3 (Mgp3) (ORF-3 protein)	Adhesins/attachment organelle
MPN144	mpn144	Putative adhesin P1-like protein MPN_144	Adhesins/attachment organelle
MPN148	mpn148	Conserved hypothetical protein MPN_148	Conserved hypothetical proteins
MPN149	mpn149	Putative mgpC-like protein MPN_149	Adhesins/attachment organelle
MPN152	mpn152	Polynucleotide binding protein Infect Immun. 2013 Sep; 81(9): bound single- and double-stranded DNA, as well as single-stranded RNA, with a predicted binding site of greater than 1 nucleotide but less than or equal to 5 nucleotides in length	Lipid/CoA metabolism related + DNA synthesis/modifications/repair/conformation
MPN155	infB	Translation initiation factor IF-2	Protein synthesis factors
MPN158	yaaC	Putative riboflavin biosynthesis protein ribF [Includes: Riboflavin kinase (EC 2.7.1.26) (Flavokinase); FMN adenylyltransferase (EC 2.7.7.2) (FAD pyrophosphorylase) (FAD synthetase)]	Vitamin/Folate metabolism
MPN159	corB	Hemolysisn, TlyC-like protein, Putative Mg2+ and Co2+ transporter CorB	Ions/Phosphate transport + Lipid/CoA metabolism
MPN162	mpn162	Zinc Transporter/ Ni,Co efflux protein	Ions/Phosphate transport

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN163	mpn163	Conserved hypothetical protein MPN_163	Conserved hypothetical proteins
MPN184	secY	Preprotein translocase subunit secY	Protein secretion
MPN185	adk	Adenylate kinase (AK) (EC 2.7.4.3) (ATP-AMP transphosphorylase)	Nucleotide metabolism/ Salvage
MPN186	map	Methionine aminopeptidase (MAP) (EC 3.4.11.18) (Peptidase M)	Protein synthesis factors
MPN194	ecfA1	Energy-coupling factor transporter ATP-binding protein ecfA1	Unassigned transporters
MPN195	ecfT	Energy-coupling factor transport system permease protein ecfT	Unassigned transporters
MPN199	mpn199	Conserved hypothetical lipoprotein MPN_199	Lipoprotein
MPN200	mpn200	Conserved hypothetical lipoprotein MPN_200	Lipoprotein
MPN202	mpn202	Putative adhesin P1-like protein MPN_202	Adhesins/attachment organelle
MPN204	mpn204	Conserved hypothetical protein MPN_204	Cytoskeleton
MPN205	mpn205	Uncharacterized protein MPN_205	Conserved hypothetical proteins
MPN209	mgtA	Probable cation-transporting P-type ATPase (EC 3.6.3.-)	Ions/Phosphate transport
MPN213	mpn213	Conserved hypothetical protein MPN_213	Conserved hypothetical proteins
MPN214	mpn214	Conserved hypothetical protein MPN_214	Conserved hypothetical proteins
MPN215	oppB	Oligopeptide transport system permease protein oppB	AA metabolism
MPN216	oppC	Oligopeptide transport system permease protein oppC	AA metabolism
MPN223	hprK	HPr kinase/phosphorylase (HPrK/P) (EC 2.7.11.-) (EC 2.7.4.-) (HPr(Ser) kinase/phosphorylase)	Glycolysis
MPN224	lgt	Prolipoprotein diacylglyceryl transferase (EC 2.4.99.-)	Protein secretion
MPN227	fus	Elongation factor G (EF-G)	Protein synthesis factors
MPN232	dnaB	Replicative DNA helicase dnaB (EC 3.6.1.-)	DNA synthesis/modifications/repair/conformation
MPN233	mpn233	Uncharacterized lipoprotein MPN_233	Lipoprotein
MPN234	mpn234	Conserved hypothetical lipoprotein MPN_234	Lipoprotein

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN240	trxB	Thioredoxin reductase (TRXR) (EC 1.8.1.9)	Nucleotide metabolism/ Salvage + Redox balance
MPN244	cdaS	Diadenylate cyclase CdaS	DNA synthesis/modifications/repair/conformation + Signal transduction
MPN245	def	Peptide deformylase (PDF) (EC 3.5.1.88) (Polypeptide deformylase)	Protein synthesis factors
MPN246	gmk	Guanylate kinase (EC 2.7.4.8) (GMP kinase)	Nucleotide metabolism/ Salvage
MPN251	cfxE	Probable ribulose-phosphate 3-epimerase (EC 5.1.3.1) (Pentose-5-phosphate 3-epimerase) (PPE) (R5P3E)	Pentose Pathway
MPN253	pgsA	CDP-diacylglycerol–glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5) (Phosphatidylglycerophosphate synthase) (PGP synthase)	Lipid/CoA metabolism
MPN254	cinA	Putative competence-damage inducible protein CinA	DNA synthesis/modifications/repair/conformation + DNA recombination
MPN257	galE	UDP-glucose 4-epimerase (EC 5.1.3.2) (UDP-galactose 4-epimerase)	Lipid/CoA metabolism + Glycocalyx
MPN258	nupA	Purine/cytidine ABC transporter permease protein nupA	Nucleotide metabolism/ Salvage
MPN259	nupB	Purine/cytidine ABC transporter permease protein nupB	Nucleotide metabolism/ Salvage
MPN262	mpn262	Uncharacterized protein, similar to reticulocyte binding protein	Conserved hypothetical proteins
MPN268	ptsG1	Putative phosphotransferase enzyme IIB component MPN_268 (EC 2.7.1.69). Glucose or acetyl-glucosamine transporter	Glycolysis + Glycocalix
MPN269	ymdA	Endoribonuclease Y (Rnase Y)	RNA degradation
MPN271	mpn271	Conserved hypothetical lipoprotein MPN_271	Lipoprotein
MPN273	hit1	Diadenosine tetraphosphate Ap4A hydrolase (EC 3.6.1.41)	Signal transduction
MPN274	mpn274	Conserved hypothetical proteins	Conserved hypothetical proteins

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN275	ybaB	DNA-binding protein, YbaB/EbfC family	DNA synthesis/modifications/repair/conformation
MPN276	mpn276	Conserved hypothetical protein MPN_276	Conserved hypothetical proteins
MPN278	glf	UDP-galactopyranose mutase (EC 5.4.99.9)	Glycocalyx
MPN281	mpn281	Conserved hypothetical lipoprotein MPN_281; ABC transporter substrate-binding protein	Unassigned transporters
MPN284	mpn284	Polynucleotide binding protein Infect Immun. 2013 Sep; 81(9): bound single- and double-stranded DNA, as well as single-stranded RNA, with a predicted binding site of greater than 1 nucleotide but less than or equal to 5 nucleotides in length	Lipid/CoA metabolism related + DNA synthesis/modifications/repair/conformation
MPN286	mpn286	Putative adhesin P1-like protein MPN_286	Adhesins/attachment organelle
MPN288	mpn288	Polynucleotide binding protein Infect Immun. 2013 Sep; 81(9): bound single- and double-stranded DNA, as well as single-stranded RNA, with a predicted binding site of greater than 1 nucleotide but less than or equal to 5 nucleotides in length	Lipid/CoA metabolism related + DNA synthesis/modifications/repair/conformation
MPN291	yeaZ	Universal bacterial protein YeaZ ; tRNA threonylcarbamoyl adenosine modification protein YeaZ / ribosomal-protein-alanine acetyltransferase multi-domain protein	Glycocalyx + Proteases
MPN292	rluD	Ribosomal large subunit pseudouridine synthase D (EC 5.4.99.-)	rRNA and rProtein modifications/processing
MPN293	lsp	Lipoprotein signal peptidase (EC 3.4.23.36) (Prolipoprotein signal peptidase) (Signal peptidase II) (SPase II)	Protein secretion
MPN298	acpS	Holo-[acyl-carrier-protein] synthase (Holo-ACP synthase) (EC 2.7.8.7) (4'-phosphopantetheinyl transferase acpS)	Lipid/CoA metabolism
MPN302	pfk	6-phosphofructokinase (Phosphofructokinase) (EC 2.7.1.11) (Phosphohexokinase)	Glycolysis

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN305	arcA1	Putative arginine deiminase (ADI) (EC 3.5.3.6) (Arginine dihydrolase) (AD) C-terminal fragment	Arg metabolism
MPN306	argI	Ornithine carbamoyltransferase, catabolic (OTCase) (EC 2.1.3.3)	Arg metabolism
MPN307	arcC	Carbamate kinase-like protein (EC 2.7.2.2)	Arg metabolism
MPN308	mpn308	Uncharacterized amino acid pernease	AA metabolism
MPN309	p65	Proline-rich P65 protein	Adhesins/attachment organelle
MPN316	ftsA	Cell division protein ftsA	Cell division
MPN318	mpn318	Putative amino acid pernease	AA metabolism
MPN319	gap1	Putative amino acid pernease	AA metabolism
MPN320	thyA	Thymidylate synthase (TS) (TSase) (EC 2.1.1.45)	Nucleotide metabolism/ Salvage
MPN321	dhfr	Dihydrofolate reductase (EC 1.5.1.3)	Nucleotide metabolism/ Salvage + Redox balance
MPN322	nrdF	Ribonucleoside-diphosphate reductase subunit beta (EC 1.17.4.1) (Ribonucleotide reductase small subunit)	Nucleotide metabolism/ Salvage + Redox balance
MPN323	nrdI	Ribonucleotide reductase stimulatory protein nrdI	Nucleotide metabolism/ Salvage + Redox balance
MPN324	nrdE	Ribonucleoside-diphosphate reductase subunit alpha (EC 1.17.4.1) (Ribonucleotide reductase)	Nucleotide metabolism/ Salvage + Redox balance
MPN326	ysxB	Predicted ribosomal protein	Ribosome
MPN333	mpn333	Putative ABC transport system permease protein	Unassigned transporters
MPN337	mpn337	Conserved hypothetical protein MPN_337	Conserved hypothetical proteins
MPN338	mpn338	Conserved hypothetical protein MPN_338	Conserved hypothetical proteins
MPN339	mpn339	Conserved hypothetical protein MPN_339 (RDD family protein)	Conserved hypothetical proteins
MPN340	uvrD1	Probable DNA helicase I homolog (EC 3.6.1.-)	DNA synthesis/modifications/repair/conformation
MPN342	hsdM	Putative type I restriction enzyme HsdM (EC 2.1.1.72)	DNA synthesis/modifications/repair/conformation
MPN344	mpn344	Uncharacterized protein MPN_344	Cytoskeleton

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN348	mthfs	5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)	Vitamin/Folate metabolism
MPN350	plsY	Putative glycerol-3-phosphate acyltransferase PlsY (EC 2.3.1.15)	Lipid/CoA metabolism
MPN351	trmK	tRNA: m1A22 methyltransferase TmrK	tRNA modifications/processing
MPN354	glyS	Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine–tRNA ligase) (GlyRS)	tRNA synthases
MPN358	mpn358	Uncharacterized protein MPN_358	Conserved hypothetical proteins
MPN361	prfA	Peptide chain release factor 1 (RF-1)	Protein synthesis factors
MPN364	mpn364	Conserved hypothetical protein MPN_364	Conserved hypothetical proteins
MPN369	mpn369	Uncharacterized lipoprotein MPN_369	Lipoprotein
MPN370	mpn370	Putative adhesin P1-like protein MPN_370	Adhesins/attachment organelle
MPN377	mpn377	Uncharacterized protein MPN_377	Conserved hypothetical proteins
MPN381	yidA	Putative Cof-like hydrolase acting on phosphorylated sugars and small metabolites	Metabolism homeostasis
MPN382	coaE	Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase)	Lipid/CoA metabolism
MPN383	yidA1	Nucleotidase, also common to HADs (works well also with CoA and NADP)	Metabolism homeostasis
MPN384	leuS	Leucyl-tRNA synthetase (EC 6.1.1.4) (Leucine–tRNA ligase) (LeuRS)	tRNA synthases
MPN385	mpn385	Conserved hypothetical protein MPN_385	Conserved hypothetical proteins
MPN388	mpn388	Conserved hypothetical protein MPN_388	Conserved hypothetical proteins
MPN389	lplA	Probable lipoate-protein ligase A (Lipoate–protein ligase) (EC 2.7.7.63)	Fermentation
MPN395	apt	Adenine phosphoribosyltransferase (APRT) (EC 2.4.2.7)	Nucleotide metabolism/ Salvage
MPN398	mpn398	Uncharacterized protein MPN_398	Conserved hypothetical proteins
MPN399	mpn399	Conserved hypothetical protein MPN_399	Conserved hypothetical proteins
MPN400	mpn400	Conserved hypothetical protein MPN_400	Pathogenesis
MPN401	greA	Transcription elongation factor greA (Transcript cleavage factor greA)	Protein synthesis factors
MPN402	proS	Prolyl-tRNA synthetase (EC 6.1.1.15) (Proline–tRNA ligase) (ProRS)	tRNA synthases
MPN406	acpA	Acyl carrier protein homolog	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN407	lipC	GDSL-like Lipase / Acylhydrolase (EC 3.1.1.5)	Lipid/CoA metabolism
MPN408	mpn408	Conserved hypothetical lipoprotein MPN_408	Lipoprotein
MPN412	mpn412	Uncharacterized protein MPN_412	Conserved hypothetical proteins
MPN414	mpn414	Putative mgpC-like protein MPN_414	Adhesins/attachment organelle
MPN415	thiB	ABC thiamine importer, substrate-binding subunit, thiB	Vitamin/Folate metabolism
MPN416	thiQ	ABC thiamine importer, ATPase subunit, thiQ	Vitamin/Folate metabolism
MPN418	alaS	Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine–tRNA ligase) (AlaRS)	tRNA synthases
MPN421	GlpU	Glycerophosphodiester transporter	Lipid/CoA metabolism
MPN422	mnmA	tRNA-specific 2-thiouridylase mnmA (EC 2.8.1.-)	tRNA modifications/processing
MPN423	mpn423	Conserved hypothetical protein MPN_423	Conserved hypothetical proteins
MPN427	yidA2	P-sugar phosphatase YidA	Metabolism homeostasis
MPN429	pgk	Phosphoglycerate kinase (EC 2.7.2.3)	Glycolysis
MPN431	ecfT1	Energy-coupling factor transport system permease protein ecfT1	Unassigned transporters
MPN432	ecfA2	Energy-coupling factor transporter ATP-binding protein ecfA2	Unassigned transporters
MPN433	ecfA3	Energy-coupling factor transporter ATP-binding protein ecfA3	Unassigned transporters
MPN435	mpn435	Predicted lysylphosphatidylglycerol synthetase	Pathogenesis + Lipid/CoA metabolism
MPN437	mpn437	Uncharacterized protein MPN_437	Conserved hypothetical proteins
MPN440	mpn440	Uncharacterized protein MPN_440	Conserved hypothetical proteins
MPN444	mpn444	Conserved hypothetical lipoprotein MPN_444	Lipoprotein
MPN445	lip3	Triacylglycerol lipase (lip) 3 (EC 3.1.-.-)	Lipid/CoA metabolism
MPN449	orf8	Conserved hypothetical protein MPN_449	Conserved hypothetical proteins
MPN450	holA	DNA polymerase III subunit delta (EC 2.7.7.7)	DNA synthesis/modifications/repair/conformation
MPN455	ctaD	Putative phosphatidic acid phosphatase family protein	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN456	oppA	Oligopeptide transport system permease protein oppA	AA metabolism
MPN457	mpn457	Uncharacterized protein MPN_457	AA metabolism
MPN458	mpn458	Conserved hypothetical protein MPN_458	AA metabolism
MPN459	mpn459	Conserved hypothetical lipoprotein MPN_459	AA metabolism
MPN460	ktrB	Ktr system potassium uptake protein B	Ions/Phosphate transport
MPN461	ktrA	Ktr system potassium uptake protein A	Ions/Phosphate transport
MPN462	mpn462	Uncharacterized protein MPN_462	Conserved hypothetical proteins
MPN464	mpn464	Putative mgpC-like protein MPN_464	Adhesins/attachment organelle
MPN465	mpn465	Conserved hypothetical protein MPN_465	Conserved hypothetical proteins
MPN466	mpn466	Conserved hypothetical protein MPN_466	Conserved hypothetical proteins
MPN467	mpn467	Conserved hypothetical lipoprotein MPN_467	Lipoprotein
MPN468	mpn468	Putative adhesin P1-like protein MPN_468	Adhesins/attachment organelle
MPN469	mpn469	Conserved hypothetical protein MPN_469	Conserved hypothetical proteins
MPN472	fakB1	Putative fatty acid kinase subunit FakB1	Lipid/CoA metabolism
MPN476	cmk	Cytidylate kinase (CK) (EC 2.7.4.14) (Cytidine monophosphate kinase) (CMP kinase)	Nucleotide metabolism/ Salvage
MPN477	mpn477	Conserved hypothetical protein MPN_477	Conserved hypothetical proteins
MPN478	yrbC	YebC family protein (transcription factor of the tetR family)	Transcription factors
MPN479	acpH	Acyl carrier protein phosphodiesterase (ACP phosphodiesterase) (EC 3.1.4.14)	Lipid/CoA metabolism
MPN483	yibD	Glycosyltransferase (EC 2.4.1.157, EC 2.4.1.46)	Glycocalyx
MPN485	mpn485	Uncharacterized protein MPN_485	Conserved hypothetical proteins
MPN486	mpn486	Uncharacterized protein MPN_486	Conserved hypothetical proteins
MPN488	nifU	NifU-like protein	Protein homeostasis
MPN489	mpn489	Uncharacterized lipoprotein MPN_489	Lipoprotein
MPN492	ulaE	Probable L-ribulose-5-phosphate 3-epimerase ulaE (EC 5.1.3.22) (L-xylulose-5-phosphate 3-epimerase) (L-ascorbate utilization protein E)	Pentose Pathway

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN493	ulaD	Probable 3-keto-L-gulonate-6-phosphate decarboxylase (KGPDC) (EC 4.1.1.85) (3-dehydro-L-gulonate-6-phosphate decarboxylase) (L-ascorbate utilization protein D)	Pentose Pathway
MPN494	ulaC	Ascorbate-specific phosphotransferase enzyme IIA component (EC 2.7.1.-) (PTS system ascorbate-specific EIIA component)	Pentose Pathway
MPN495	ulaB	Ascorbate-specific phosphotransferase enzyme IIB component (EC 2.7.1.69) (Ascorbate-specific PTS system EIIB component)	Pentose Pathway
MPN496	ulaA	Ascorbate-specific permease IIC component ulaA (Ascorbate-specific PTS system EIIC component)	Pentose Pathway
MPN497	ulaG	Probable L-ascorbate-6-phosphate lactonase ulaG (EC 3.1.1.-)	Pentose Pathway
MPN498	araD	Probable L-ribulose-5-phosphate 4-epimerase ulaF (EC 5.1.3.4) (Phosphoribulose isomerase) (L-ascorbate utilization protein F)	Pentose Pathway
MPN500	mpn500	Putative adhesin P1-like protein MPN_500	Adhesins/attachment organelle
MPN502	mpn502	Uncharacterized protein MPN_502	Conserved hypothetical proteins
MPN503	mpn503	Putative mgpC-like protein MPN_503	Adhesins/attachment organelle
MPN505	mpn505	Uncharacterized protein MPN_505	Conserved hypothetical proteins
MPN506	mpn506	Polynucleotide binding protein Infect Immun. 2013 Sep; 81(9): bound single- and double-stranded DNA, as well as single-stranded RNA, with a predicted binding site of greater than 1 nucleotide but less than or equal to 5 nucleotides in length	Lipid/CoA metabolism related + DNA synthesis/modifications/repair/conformation
MPN508	mpn508	Conserved hypothetical proteins	Conserved hypothetical proteins
MPN509	mpn509	Uncharacterized protein MPN_509	Conserved hypothetical proteins
MPN517	yhdA	Putative NADPH-dependent FMN reductase	Vitamin/Folate metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN519	lip3B	Putative esterase/lipase 3 (EC 3.1.-.-)	Lipid/CoA metabolism
MPN522	trmB	tRNA (guanine-N(7)-)-methyltransferase (EC 2.1.1.33) (tRNA(m7G46)-methyltransferase)	tRNA modifications/processing
MPN527	mpn527	Uncharacterized protein MPN_527	Conserved hypothetical proteins
MPN528	ppa	Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-hydrolase) (PPase)	Phosphate metabolism
MPN530	mpn530	Conserved hypothetical protein MPN_530	Conserved hypothetical protein
MPN531	clpB	Chaperone protein clpB; ATP-dependent protease.	Protein homeostasis
MPN532	licA	Predicted choline kinase (EC 2.7.1.32)	Lipid/CoA metabolism
MPN533	ackA	Acetate kinase (EC 2.7.2.1) (Acetokinase)	Fermentation
MPN537	mucB	UV protection protein MucB	DNA synthesis/modifications/repair/conformation
MPN538	rplJ	50S ribosomal protein L10	Ribosome
MPN539	rplL	50S ribosomal protein L7/L12	Ribosome
MPN543	fmt	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	tRNA synthases
MPN544	mpn544	Conserved hypothetical protein MPN_544	Conserved hypothetical protein
MPN547	fakA	Fatty acid kinase subunit FakA	Lipid/CoA metabolism
MPN550	thiI	Probable thiamine biosynthesis protein thiI or tRNA sulfur transferase modification enzyme (EC 2.8.1.4)	tRNA modifications/processing
MPN552	mpn552	Conserved hypothetical protein MPN_552	Conserved hypothetical proteins
MPN555	mpn555	Putative molecular chaperone	Ribosome
MPN556	argS	Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine–tRNA ligase) (ArgRS)	tRNA synthases
MPN558	gidB	Ribosomal RNA small subunit methyltransferase G (EC 2.1.1.-) (16S rRNA 7-methylguanosine methyltransferase) (16S rRNA m7G methyltransferase) (Glucose-inhibited division protein B)	rRNA and rProtein modifications/processing
MPN561	udk	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase)	Nucleotide metabolism/ Salvage

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN562	outB	Probable NH(3)-dependent NAD(+) synthetase (EC 6.3.1.5)	Vitamin/Folate metabolism
MPN563	obg	GTP-binding protein Obg	Signal transduction
MPN564	adh	Probable NADP-dependent alcohol dehydrogenase (EC 1.1.1.2)	Fermentation + Redox balance + Nucleotide metabolism/Salvage
MPN565	mpn565	Conserved hypothetical protein MPN_565	Conserved hypothetical protein
MPN573	groEL	60 kDa chaperonin (Protein Cpn60) (GroEL protein)	Protein homeostasis
MPN575	mpn575	Conserved hypothetical protein MPN_575	Conserved hypothetical proteins
MPN582	mpn582	Conserved hypothetical lipoprotein MPN_582	Lipoprotein
MPN585	mpn585	Conserved hypothetical lipoprotein MPN_585	Lipoprotein
MPN588	mpn588	Uncharacterized lipoprotein MPN_588	Lipoprotein
MPN590	mpn590	Conserved hypothetical lipoprotein MPN_590	Lipoprotein
MPN591	mpn591	Conserved hypothetical protein MPN_591	Lipoprotein
MPN592	mpn592	Peptidase S7 family-like protein	Membrane Protease
MPN593	mpn593	Conserved hypothetical protein MPN_593	Conserved hypothetical protein
MPN595	rpiB	Probable ribose-5-phosphate isomerase B (EC 5.3.1.6) (Phosphoriboisomerase B)	Pentose Pathway
MPN596	erzA	Negative regulator of FtsZ ring formation	Cell Division
MPN598	atpD	ATP synthase subunit beta (EC 3.6.3.14) (F-ATPase subunit beta) (ATP synthase F1 sector subunit beta)	ATPase
MPN599	atpG	ATP synthase gamma chain (ATP synthase F1 sector gamma subunit) (F-ATPase gamma subunit)	ATPase
MPN600	atpA	ATP synthase subunit alpha (EC 3.6.3.14) (F-ATPase subunit alpha) (ATP synthase F1 sector subunit alpha)	ATPase
MPN601	atpH	ATP synthase delta chain (F-ATPase delta chain)	ATPase
MPN602	atpF	ATP synthase B chain	ATPase

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN603	atpE	ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein)	ATPase
MPN604	atpB	ATP synthase A chain (F-ATPase subunit 6) (ATP synthase F0 sector subunit A)	ATPase
MPN605	atpI	ATP synthase protein I	ATPase
MPN606	eno	Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase)	Glycolysis
MPN607	pmsR	Peptide methionine sulfoxide reductase msrA (Protein-methionine-S-oxide reductase) (EC 1.8.4.11) (Peptide-methionine (S)-S-oxide reductase) (Peptide Met(O) reductase)	Protein homeostasis
MPN608	phoU	Transcriptional regulator involved in phosphate transport system	Ions/Phosphate transport + Phosphate metabolism
MPN610	pstA	Phosphate transport system permease protein pstA homolog	Ions/Phosphate transport + Phosphate metabolism
MPN611	pstS	Phosphate-binding protein pstS (PBP)	Ions/Phosphate transport + Phosphate metabolism
MPN620	mpn620	Conserved hypothetical protein MPN_620	Conserved hypothetical proteins
MPN628	pgm	2,3-bisphosphoglycerate-independent phosphoglycerate mutase (Phosphoglyceromutase) (BPG-independent PGAM) (iPGM) (EC 5.4.2.1)	Glycolysis
MPN629	tim	Triosephosphate isomerase (TIM) (EC 5.3.1.1) (Triose-phosphate isomerase)	Glycolysis
MPN630	yfiB	Uncharacterized ABC transporter ATP-binding protein	Lipid/CoA metabolism
MPN631	tsf	Elongation factor Ts (EF-Ts)	Protein synthesis factors
MPN632	pyrH	Uridylate kinase (UK) (EC 2.7.4.22) (Uridine monophosphate kinase) (UMP kinase) (UMPK)	Nucleotide metabolism/ Salvage
MPN636	frr	Ribosome-recycling factor (RRF) (Ribosome-releasing factor)	Protein synthesis factors
MPN639	mpn639	Conserved hypothetical lipoprotein MPN_639	Lipoproteins

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN640	mpn640	Conserved hypothetical lipoprotein MPN_640	Lipoproteins
MPN641	mpn641	Conserved hypothetical lipoprotein MPN_641	Lipoproteins
MPN642	mpn642	Conserved hypothetical lipoprotein MPN_642	Lipoproteins
MPN643	mpn643	Conserved hypothetical lipoprotein MPN_643	Lipoproteins
MPN644	mpn644	Conserved hypothetical lipoprotein MPN_644	Lipoproteins
MPN645	mpn645	Conserved hypothetical lipoprotein MPN_645	Lipoproteins
MPN646	mpn646	Conserved hypothetical lipoprotein MPN_646	Lipoproteins
MPN647	mpn647	Conserved hypothetical lipoprotein MPN_647	Lipoproteins
MPN652	mtlD	Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17)	Glycolysis
MPN653	mtlF	Mannitol-specific phosphotransferase enzyme IIA component (EC 2.7.1.-) (PTS system mannitol-specific EIIA component) (EIIA-Mtl) (EIII-Mtl)	Glycolysis
MPN654	mpn654	Conserved hypothetical lipoprotein MPN_654	Lipoproteins
MPN656	rbgA	Ribosome biogenesis GTPase A	Ribosome
MPN657	yitT	Putative YitT family protein; general stress protein, required for protection against paraquat stress	Oxidative homeostasis
MPN659	trmD	tRNA (guanine-N(1)-)-methyltransferase (EC 2.1.1.31) (M1G-methyltransferase) (tRNA [GM37] methyltransferase)	tRNA modifications/processing
MPN661	norM	Na+ -driven efflux pump	Ions/Phosphate transport
MPN662	pilB	Peptide methionine sulfoxide reductase msrB (EC 1.8.4.12) (Peptide-methionine (R)-S-oxide reductase)	Oxidative homeostasis
MPN664	fakB2	Putative fatty acid kinase subunit FakB2	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN667	gtaB	UTP–glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) (UDP-glucose pyrophosphorylase) (UDPGP) (Alpha-D-glucosyl-1-phosphate uridylyltransferase) (Uridine diphosphoglucose pyrophosphorylase)	Lipid/CoA metabolism + Glycocalyx
MPN668	ohr	Organic hydroperoxide resistance protein-like	Oxidative homeostasis
MPN672	hpt	Hypoxanthine-guanine phosphoribosyltransferase (HGPRTase) (HGPRT) (EC 2.4.2.8)	Nucleotide metabolism/ Salvage
MPN678	gltX	Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate–tRNA ligase) (GluRS)	tRNA synthases
MPN680	yidC	Preprotein translocase, YidC component	Protein secretion
MPN027	rimL	Ribosomal-protein alanineacetyltransferase	rRNA and rProtein modifications/processing
MPN031	mpn031	Uncharacterized protein MPN031	Conserved hypothetical proteins
MPN044	tdk	Thymidine kinase (EC 2.7.1.21)	Nucleotide metabolism/ Salvage
MPN050	glpK	Glycerol kinase (EC 2.7.1.30) (ATP:glycerol 3-phosphotransferase) (GK)	Lipid/CoA metabolism
MPN061	ffh	Signal recognition particle protein (Fifty-four homolog)	Protein synthesis factors
MPN066	cpsG	Phosphomannomutase (PMM) (EC 5.4.2.8)	Pentose Pathway
MPN067	nusG	Transcription antitermination factor	RNA synthesis
MPN077	uhpT2	Hexosephosphate transport protein	Lipid/CoA metabolism
MPN080	ftsX	FtsX-like protein	Cell division
MPN083	mpn083	Uncharacterized lipoprotein MPN_083	Lipoprotein
MPN119	topJ	Attachment organelle-associated co-chaperone	Adhesins/attachment organelle
MPN139	mpn139	Uncharacterized protein MPN_139	Cytoskeleton
MPN141	mgpA	Adhesin P1 (Cytadhesin P1) (Attachment protein)	Adhesins/attachment organelle
MPN145	mpn145	Uncharacterized protein MPN_145	Cytoskeleton

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN151	mpn151	Uncharacterized protein MPN_151	Cytoskeleton
MPN156	rbfA	Ribosome-binding factor A	Ribosome
MPN164	rpsJ	30S ribosomal protein S10- NusE termination antitermination factor	Ribosome + RNA synthesis
MPN187	infA	Translation initiation factor IF-1	Protein synthesis factors
MPN188	rpmJ	50S ribosomal protein L36	Ribosome
MPN193	ecfA	Energy-coupling factor transporter ATP-binding protein ecfA	Unassigned transporters
MPN207	ptsG	PTS system glucose-specific EIICBA component (EIICBA-Glc) (EII-Glc/EIII-Glc) (EC 2.7.1.69)	Glycolysis
MPN210	secA	Protein translocase subunit secA	Protein secretion
MPN211	uvrB	UvrABC system protein B (Protein uvrB) (Excinuclease ABC subunit B)	DNA synthesis/modifications/repair/conformation
MPN219	rplK	50S ribosomal protein L11	Ribosome
MPN221	pth	Peptidyl-tRNA hydrolase (PTH) (EC 3.1.1.29)	Protein synthesis factors
MPN236	gatC	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C (Asp/Glu-ADT subunit C) (EC 6.3.5.-)	tRNA synthases
MPN237	gatA	Glutamyl-tRNA(Gln) amidotransferase subunit A (Glu-ADT subunit A) (EC 6.3.5.-)	tRNA synthases
MPN238	gatB	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B (Asp/Glu-ADT subunit B) (EC 6.3.5.-)	tRNA synthases
MPN243	vacB	Ribonuclease R (RNase R) (EC 3.1.-.-) (VacB protein homolog)	RNA degradation
MPN248	prkC	Putative serine/threonine-protein kinase PrkC (EC 2.7.11.1)	Signal transduction
MPN260	nupC	Purine/cytidine ABC transporter permease protein nupC	Nucleotide metabolism/ Salvage
MPN263	trx	Thioredoxin (Trx)	Nucleotide metabolism/ Salvage + Redox balance

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN264	mpn264	Uncharacterized hydrolase, haloacid dehalogenase-like family protein (EC 3.1.-.-)	Metabolism homeostasis
MPN267	ppnK	Probable inorganic polyphosphate/ATP-NAD kinase (Poly(P)/ATP NAD kinase) (EC 2.7.1.23)	Phosphate metabolism
MPN272	mpn272	Uncharacterized protein MPN_272	Conserved hypothetical proteins
MPN277	lysS	Lysyl-tRNA synthetase (EC 6.1.1.6) (Lysine–tRNA ligase) (LysRS)	tRNA synthases
MPN279	lepA	GTP-binding protein lepA	Protein synthesis factors
MPN297	gpsB	Probable cell-division initiation protein DivIVA	Cell division
MPN299	plsC	Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (1-AGP acyltransferase) (1-AGPAT) (EC 2.3.1.51) (Lysophosphatidic acid acyltransferase) (LPAAT)	Lipid/CoA metabolism
MPN300	scpA	Dihydrofolate reductase (EC 1.5.1.3) and segregation and condensation protein A, ScpA	Cell division
MPN311	p41	Cytadherence-related protein P41	Adhesins/attachment organelle
MPN312	p24	Cytadherence-related protein P24	Adhesins/attachment organelle
MPN315	mraW	S-adenosyl-L-methionine-dependent methyltransferase mraW (EC 2.1.1.-)	Cell division
MPN317	ftsZ	Cell division protein ftsZ	Cell division
MPN330	mpn330	Conserved hypothetical protein MPN_330	Conserved hypothetical proteins
MPN331	tig	Trigger factor (TF) (EC 5.2.1.8)	Protein synthesis factors
MPN335	mpn335	Putative ABC transport system permease protein	Unassigned transporters
MPN336	nadD	Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18), phosphopantetheine adenylyltransferase (EC 2.7.7.3) and choline-phosphate cytidylyltransferase (EC 2.7.7.15)	Vitamin/Folate metabolism
MPN341	pcrA	DNA helicase II / ATP-dependent DNA helicase PcrA (EC 3.6.4.12)	DNA synthesis/modifications/repair/conformation

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN353	dnaG	DNA primase (EC 2.7.7.-)	DNA synthesis/modifications/repair/conformation
MPN359	mpn359	Conserved hypothetical protein MPN_359	Conserved hypothetical proteins
MPN360	rpmE	50S ribosomal protein L31	Ribosome
MPN366	mpn366	Putative mgpC-like protein MPN_366	Adhesins/attachment organelle
MPN368	mpn368	Uncharacterized protein MPN_368	Cytoskeleton
MPN372	ptxA	ADP-ribosylating toxin CARDS (EC 2.4.2.-) (CARDX TX) (ADP-ribosyltransferase CARDS)	Pathogenesis
MPN376	mpn376	Uncharacterized protein MPN_376	Conserved hypothetical proteins
MPN378	dnaE	DNA polymerase III subunit alpha (EC 2.7.7.7)	DNA synthesis/modifications/repair/conformation
MPN387	mpn387	Conserved hypothetical protein MPN_387	Conserved hypothetical proteins
MPN390	pdhD	Dihydrolipoyl dehydrogenase (EC 1.8.1.4) (Dihydrolipoamide dehydrogenase) (E3 component of pyruvate complex)	Fermentation
MPN391	pdhC	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12) (E2) (Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex)	Fermentation
MPN392	pdhB	Pyruvate dehydrogenase E1 component subunit beta (EC 1.2.4.1)	Fermentation
MPN393	pdhA	Pyruvate dehydrogenase E1 component subunit alpha (EC 1.2.4.1)	Fermentation
MPN394	nox	Probable NADH oxidase (NOXase) (EC 1.6.99.3)	Fermentation
MPN396	secD	Protein-export membrane protein secD	Protein secretion
MPN397	spoT	Probable guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2) (Penta-phosphate guanosine-3'-pyrophosphohydrolase) ((ppGpp)ase)	Signal transduction
MPN409	mpn409	Putative adhesin P1-like protein MPN_409	Adhesins/attachment organelle
MPN410	mpn410	Uncharacterized protein MPN_410	Cytoskeleton

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN420	glpQ	Glycerophosphoryl diester phosphodiesterase	Lipid/CoA metabolism
MPN425	ftsY	Cell division protein ftsY homolog	Cell division
MPN426	smc	SMC family, chromosome/DNA binding/protecting functions	Cell division + DNA synthesis/modifications/repair/conformation
MPN428	pta	Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase)	Glycolysis
MPN434	dnaK	Chaperone protein DnaK (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70)	Protein homeostasis
MPN436	mpn436	Uncharacterized lipoprotein MPN_436	Lipoprotein
MPN447	hmw1	Cytadherence high molecular weight protein 1 (Cytadherence accessory protein 1)	Adhesins/attachment organelle
MPN453	p30	P30 adhesin (Cytadhesin P30) (30 kDa adhesin-related protein)	Adhesins/attachment organelle
MPN471	rpmG	50S ribosomal protein L33 1	Ribosome
MPN473	lip2	Putative esterase/lipase 2 (EC 3.1.-.-)	Lipid/CoA metabolism
MPN482	yneF	Uncharacterized protein YneF	Conserved hypothetical proteins
MPN499	mpn499	Uncharacterized protein MPN_499	Conserved hypothetical proteins
MPN504	mpn504	Uncharacterized protein MPN_504	Cytoskeleton
MPN523	mpn523	Uncharacterized lipoprotein MPN_523	Lipoprotein
MPN524	mpn524	Uncharacterized protein MPN_524	Cytoskeleton
MPN525	dnaB1	Replication initiation and membrane attachment protein	DNA synthesis/modifications/repair/conformation + Adhesins/attachment organelle
MPN529	ihf	Histone-like bacterial DNA-binding protein	DNA synthesis/modifications/repair/conformation
MPN546	plsX	Fatty acid/phospholipid synthesis protein plsX (EC 2.3.1.15)	Lipid/CoA metabolism
MPN551	yqaJ	Exonuclease, YqaJ recombinase family	DNA recombination
MPN557	gidA	tRNA uridine 5-carboxymethylaminomethyl modification enzyme mnmG (Glucose-inhibited division protein A)	tRNA modifications/processing
MPN560	arcA2	Arginine deiminase-like protein	Arg metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN567	p200	Protein P200	Adhesins/attachment organelle
MPN568	spg	GTP-binding protein Era homolog	Ribosome
MPN569	ybeY	Endoribonuclease YbeY; probable rRNA maturation factor (EC 3.4.24.-)	RNA degradation + rRNA and rProtein modifications/processing
MPN597	atpC	ATP synthase epsilon chain (ATP synthase F1 sector epsilon subunit) (F-ATPase epsilon subunit)	ATPase
MPN622	rpsO	30S ribosomal protein S15	Ribosome
MPN625	osmC	Osmotically inducible protein C	Oxidative homeostasis
MPN665	tuf	Elongation factor Tu (EF-Tu)	Protein synthesis factors
MPN670	mpn670	Conserved hypothetical protein MPN_670	Conserved hypothetical proteins
MPN671	ftsH	Cell division protease ftsH homolog (EC 3.4.24.-)	Proteases
MPN673	ygbB	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase-like protein	Conserved hypothetical proteins
MPN674	ldh	L-lactate dehydrogenase (L-LDH) (EC 1.1.1.27)	Fermentation + Oxidative homeostasis + Redox balance
MPN677	ywfO	DGT (deoxyguanosinetriphosphate triphosphohydrolase) (EC 3.1.5.1)	Nucleotide metabolism/ Salvage
MPN684	mpn684	Uncharacterized ABC transporter permease MPN_684	Unassigned transporters
MPN687	mpn687	Uncharacterized protein MPN_687	Conserved hypothetical proteins
MPN002	cbpA	Curved DNA-binding protein CbpA	DNA synthesis/modifications/repair/conformation
MPN005	serS	Seryl-tRNA synthetase (EC 6.1.1.11) (Seryl-tRNA(Ser/Sec) synthetase) (Serine–tRNA ligase) (SerRS)	tRNA synthases
MPN008	thdF	tRNA modification GTPase mnmE (EC 3.6.-.-)	tRNA modifications/processing
MPN021	dnaJ	Chaperone protein dnaJ	Protein homeostasis
MPN063	deoC	Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyriboaldolase) (Deoxyriboaldolase) (DERA)	Nucleotide metabolism/ Salvage
MPN069	rpmG2	50S ribosomal protein L33 type 2	Ribosome

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN073	prs	Ribose-phosphate pyrophosphokinase (RPPK) (EC 2.7.6.1) (Phosphoribosyl pyrophosphate synthetase) (P-Rib-PP synthetase) (PRPP synthetase)	Pentose Pathway + Nucleotide metabolism/ Salvage
MPN074	smpB	tmRNA-binding protein	Protein synthesis factors + Protein homeostasis
MPN082	tklB	Transketolase (TK) (EC 2.2.1.1)	Pentose Pathway
MPN116	rpmI	50S ribosomal protein L35	Ribosome
MPN123	parC	DNA topoisomerase 4 subunit A (EC 5.99.1.-) (Topoisomerase IV subunit A)	DNA synthesis/modifications/repair/conformation
MPN127	mpn127	Uncharacterized protein MPN_127	Cytoskeleton
MPN134	ugpC	Putative ABC sn-glycerol-3-phosphate transporter ATP-binding protein ugpC	Nucleotide metabolism/salvage
MPN137	mpn137	Uncharacterized protein MPN_137	Cytoskeleton
MPN138	mpn138	Uncharacterized protein MPN_138	Cytoskeleton
MPN161	mpn161	Conserved hypothetical protein MPN_161	Conserved hypothetical proteins
MPN173	rpmC	50S ribosomal protein L29	Ribosome
MPN182	rpsE	30S ribosomal protein S5	Ribosome
MPN196	truA	tRNA pseudouridine synthase A, TruA (EC 5.4.99.12) (tRNA-uridine isomerase I) (tRNA pseudouridylate synthase I)	tRNA modifications/processing
MPN198	mte1	Uncharacterized adenine-specific methylase MPN_198 (EC 2.1.1.72) (M.mpnIP)	DNA synthesis/modifications/repair/conformation
MPN208	rpsB	30S ribosomal protein S2	Ribosome
MPN218	oppF	Oligopeptide transport ATP-binding protein oppF	AA metabolism
MPN220	rplA	50S ribosomal protein L1	Ribosome
MPN222	yacA	tRNA(Ile)-lysidine synthase (EC 6.3.4.-) (tRNA(Ile)-lysidine synthetase) (tRNA(Ile)-2-lysyl-cytidine synthase)	tRNA modifications/processing
MPN235	ung	Uracil-DNA glycosylase (UDG) (EC 3.2.2.-)	DNA synthesis/modifications/repair/conformation
MPN241	whiA	Transcription factor with WhiA C-terminal domain	Transcription factors

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN250	pgiB	Glucose-6-phosphate isomerase (GPI) (EC 5.3.1.9) (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase) (PHI)	Glycolysis
MPN256	mpn256	Conserved hypothetical protein MPN_256	Conserved hypothetical proteins
MPN265	trpS	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan–tRNA ligase) (TrpRS)	tRNA synthases
MPN285	prrB	Putative type-1 restriction enzyme specificity protein MPN_285 (S.mpnORFGP) (Type I restriction enzyme specificity protein MPN_285) (S protein)	DNA synthesis/modifications/repair/conformation
MPN296	rpsU	30S ribosomal protein S21	Ribosome
MPN301	scpB	Segregation and condensation protein B	Cell division
MPN310	hmw2	Cytadherence high molecular weight protein 2 (Cytadherence accessory protein 2)	Adhesins/attachment organelle
MPN328	nfo	Probable endonuclease 4 (EC 3.1.21.2) (Endonuclease IV) (Endodeoxyribonuclease IV)	DNA synthesis/modifications/repair/conformation
MPN329	fur	Ferric uptake regulation protein	Transcription factors
MPN343	mpn343	Putative type-1 restriction enzyme specificity protein MPN_343 (S.mpnORFDP) (Type I restriction enzyme specificity protein MPN_343) (S protein)	DNA synthesis/modifications/repair/conformation
MPN349	ymdB	Putative bi-c-amp phosphodiesterase / M-pesterase	Signal transduction
MPN355	trmH	RNA methyltransferase, TrmH family, group 3 (Probable 23S rRNA (guanosine2251-2'-O-)-methyltransferase rlmB) (EC 2.1.1.185)	rRNA and rProtein modifications/processing
MPN356	cysS	Cysteinyl-tRNA synthetase (EC 6.1.1.16) (Cysteine–tRNA ligase) (CysRS)	tRNA synthases
MPN362	hemK	Protein methyltransferase hemK (EC 2.1.1.-) (Protein-glutamine N-methyltransferase hemK) (Protein-(glutamine-N(5)) MTase hemK)	rRNA and rProtein modifications/processing

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN365	mpn365	Putative type-1 restriction enzyme specificity protein MPN_365 (S.mpnORFCP) (Type I restriction enzyme specificity protein MPN_365) (S protein)	DNA synthesis/modifications/repair/conformation
MPN367	mpn367	Putative mgpC-like protein MPN_367	Adhesins/attachment organelle
MPN386	yaaF	Putative deoxynucleoside kinase protein (EC 2.7.1.76, EC 2.7.1.113)	Nucleotide metabolism/ Salvage
MPN443	deaD	Probable ATP-dependent RNA helicase (EC 3.6.1.-)	RNA synthesis
MPN446	rpsD	30S ribosomal protein S4	Ribosome
MPN452	hmw3	Cytadherence high molecular weight protein 3 (Cytadherence accessory protein 3) (Accessory adhesin protein 3)	Adhesins/attachment organelle
MPN481	engB	Probable GTP-binding protein engB	Ribosome
MPN484	mpn484	Uncharacterized protein MPN_484	Cytoskeleton
MPN507	mpn507	Putative type-1 restriction enzyme specificity protein MPN_507 (S.mpnORFBP) (Type I restriction enzyme specificity protein MPN_507) (S protein)	DNA synthesis/modifications/repair/conformation
MPN526	mpn526	Conserved hypothetical protein MPN_526	Conserved hypothetical proteins
MPN540	rpmF	50S ribosomal protein L32	Ribosome
MPN541	rpsT	30S ribosomal protein S20	Ribosome
MPN542	mpn542	Uncharacterized protein MPN_542	Conserved hypothetical protein
MPN548	truC	RNA pseudouridine synthase TruC (EC 5.4.99.-) (RNA-uridine isomerase) (RNA pseudouridylate synthase)	rRNA and rProtein modifications/processing
MPN553	thrSv	Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine–tRNA ligase) (ThrRS)	tRNA synthases
MPN554	ssbB	Putative single-stranded DNA-binding protein (SSB)	DNA recombination + DNA synthesis/modifications/repair/conformation
MPN559	mpn559	Conserved hypothetical protein MPN_559	Conserved hypothetical protein
MPN566	glpQ1	Glycerophosphoryl diester phosphodiesterase. [EC:3.1.4.46]	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN609	pstB	Phosphate import ATP-binding protein pstB (EC 3.6.3.27) (Phosphate-transporting ATPase) (ABC phosphate transporter)	Ions/Phosphate transport + Phosphate metabolism
MPN615	hsdS1	Putative type I restriction enzyme S protein (HsdS)	DNA synthesis/modifications/repair/conformation
MPN616	rpsI	30S ribosomal protein S9	Ribosome
MPN617	rplM	50S ribosomal protein L13	Ribosome
MPN618	dnaX	DNA polymerase III subunit gamma/tau (EC 2.7.7.7)	DNA synthesis/modifications/repair/conformation
MPN624	rpmB	50S ribosomal protein L28	Ribosome
MPN655	mpn655	Uncharacterized protein MPN_655	Cytoskeleton
MPN675	mpn675	Uncharacterized protein MPN_675	Conserved hypothetical proteins
MPN685	bcaA2	Similar to bacitracin export ATP-binding protein B	Unassigned transporters
MPN686	dnaA	Chromosomal replication initiator protein dnaA	DNA synthesis/modifications/repair/conformation
MPN688	soj	ParA family protein MPN_688	DNA synthesis/modifications/repair/conformation + Cell Division
MPN013	mpn013	Uncharacterized protein MPN_013	Cytoskeleton
MPN015	rimK	Ribosomal protein S6 modification enzyme	rRNA and rProtein modifications/processing
MPN033	upp	Uracil phosphoribosyltransferase (EC 2.4.2.9) (UMP pyrophosphorylase) (UPRTase)	Nucleotide metabolism/ Salvage
MPN034	polC	DNA polymerase III polC-type (PolIII) (EC 2.7.7.7)	DNA synthesis/modifications/repair/conformation
MPN046	aspS	Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate–tRNA ligase) (AspRS)	tRNA synthases
MPN055	potA	Spermidine/putrescine import ATP-binding protein potA (EC 3.6.3.31)	DNA synthesis/modifications/repair/conformation
MPN072	rnmV	Ribonuclease M5 (EC 3.1.26.8)	RNA degradation
MPN081	glnQ	Glutamine transport ATP-binding protein glnQ	AA metabolism
MPN100	mpn100	Uncharacterized protein MPN_100	Conserved hypothetical proteins

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN105	pheS	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20) (Phenylalanine–tRNA ligase alpha chain) (PheRS)	tRNA synthases
MPN106	pheT	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine–tRNA ligase beta chain) (PheRS)	tRNA synthases
MPN117	rpLT	50S ribosomal protein L20	Ribosome
MPN124	hrcA	Heat-inducible transcription repressor hrcA	Transcription factors
MPN125	uvrC	UvrABC system protein C (Protein uvrC) (Excinuclease ABC subunit C)	DNA synthesis/modifications/repair/conformation
MPN140	nrnA	Bifunctional oligoribonuclease and PAP phosphatase NrnA EC=3.1.-.-	RNA degradation
MPN153	uvrD	Probable DNA helicase I homolog (EC 3.6.1.-)	DNA synthesis/modifications/repair/conformation
MPN154	nusA	Transcription elongation protein nusA	RNA synthesis
MPN157	mpn157	Conserved hypothetical protein MPN_157	Conserved hypothetical proteins
MPN166	rplD	50S ribosomal protein L4	Ribosome
MPN169	rpsS	30S ribosomal protein S19	Ribosome
MPN170	rplV	50S ribosomal protein L22	Ribosome
MPN172	rplP	50S ribosomal protein L16	Ribosome
MPN174	rpsQ	30S ribosomal protein S17	Ribosome
MPN175	rplN	50S ribosomal protein L14	Ribosome
MPN177	rplE	50S ribosomal protein L5	Ribosome
MPN178	rpsN	30S ribosomal protein S14 type Z	Ribosome
MPN179	rpsH	30S ribosomal protein S8	Ribosome
MPN180	rplF	50S ribosomal protein L6	Ribosome
MPN181	rplR	50S ribosomal protein L18	Ribosome
MPN183	rplO	50S ribosomal protein L15	Ribosome
MPN189	rpsM	30S ribosomal protein S13	Ribosome
MPN190	rpsK	30S ribosomal protein S11	Ribosome
MPN191	rpoA	DNA-directed RNA polymerase subunit alpha (RNAP subunit alpha) (EC 2.7.7.6) (Transcriptase subunit alpha)	RNA synthesis
MPN217	oppD	Oligopeptide transport ATP-binding protein oppD	AA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN225	rpsL	30S ribosomal protein S12	Ribosome
MPN229	ssbA	Single-stranded DNA-binding protein (SSB) (Helix-destabilizing protein)	DNA recombination + DNA synthesis/modifications/repair/conformation
MPN230	rpsR	30S ribosomal protein S18	Ribosome
MPN231	rplI	50S ribosomal protein L9	Ribosome
MPN239	gntR	Probable HTH-type transcriptional regulator gntR	Transcription factors
MPN247	ptc1	Putative protein phosphatase PrpC (EC 3.1.3.16)	Signal transduction
MPN252	asnS	Asparaginyl-tRNA synthetase (EC 6.1.1.22) (Asparagine–tRNA ligase) (AsnRS)	tRNA synthases
MPN255	ygbP	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase-like	Conserved hypothetical proteins
MPN287	mpn287	Uncharacterized protein MPN_287	Cytoskeleton
MPN314	mraZ	Cell division protein mraZ	Cell division
MPN327	rpl27	50S ribosomal protein L27	Ribosome
MPN332	lon	ATP-dependent protease La (EC 3.4.21.53)	Proteases + Protein homeostasis
MPN357	lig	DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+])	DNA synthesis/modifications/repair/conformation
MPN379	polA	5'-3' exonuclease (EC 3.1.11.-)	DNA synthesis/modifications/repair/conformation
MPN380	fpg	Formamidopyrimidine-DNA glycosylase (Fapy-DNA glycosylase) (EC 3.2.2.23) (DNA-(apurinic or apyrimidinic site) lyase mutM) (AP lyase mutM) (EC 4.2.99.18)	DNA synthesis/modifications/repair/conformation
MPN424	ylxM	Putative helix-turn-helix protein, YlxM/p13-like protein	Signal transduction
MPN454	mpn454	Conserved hypothetical protein MPN_454	Conserved hypothetical proteins
MPN470	pepX	Putative Xaa-Pro aminopeptidase (X-Pro aminopeptidase) (EC 3.4.11.9) (Aminopeptidase P) (APP) (Aminoacylproline aminopeptidase)	Proteases
MPN474	mpn474	Uncharacterized protein MPN_474	Conserved hypothetical proteins

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN480	valS	Valyl-tRNA synthetase (EC 6.1.1.9) (Valine–tRNA ligase) (ValRS)	tRNA synthases
MPN487	nifS	Probable cysteine desulfurase (EC 2.8.1.7)	Protein homeostasis
MPN490	recA	Protein recA (Recombinase A)	DNA recombination
MPN491	mnuA	Membrane nuclease A	Nucleotide metabolism/ Salvage
MPN501	mpn501	Uncharacterized protein MPN_501	Cytoskeleton
MPN515	rpoC	DNA-directed RNA polymerase subunit beta' (RNAP subunit beta') (EC 2.7.7.6) (Transcriptase subunit beta') (RNA polymerase subunit beta')	RNA synthesis
MPN518	mpn518	Conserved hypothetical protein MPN_518	Conserved hypothetical proteins
MPN521	ygl3	Probable tRNA/rRNA methyltransferase SpoU (EC 2.1.1.-)	tRNA modifications/processing
MPN545	rnc	Ribonuclease 3 (EC 3.1.26.3) (Ribonuclease III) (RNase III)	RNA degradation + Ribosome
MPN572	pepA	Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase) (LAP) (Leucyl aminopeptidase); DNA binding protein	Proteases
MPN574	groES	10 kDa chaperonin (Protein Cpn10) (GroES protein)	Protein homeostasis
MPN576	glyA	Serine hydroxymethyltransferase (Serine methylase) (SHMT) (EC 2.1.2.1)	Vitamin/Folate metabolism
MPN623	deaD1	Probable ATP-dependent RNA helicase (EC 3.6.1.-)	RNA synthesis
MPN627	ptsI	Phosphoenolpyruvate-protein phosphotransferase (EC 2.7.3.9) (Phosphotransferase system, enzyme I)	Glycolysis
MPN638	hsdS2	Putative type I restriction enzyme S protein (HsdS)	DNA synthesis/modifications/repair/conformation
MPN660	rpsP	30S ribosomal protein S16	Ribosome
MPN663	ytpR	Putative tRNA-binding protein ytpR	Protein synthesis factors
MPN669	tyrS	Tyrosyl-tRNA synthetase (EC 6.1.1.1) (Tyrosine–tRNA ligase) (TyrRS)	tRNA synthases

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN679	ksgA	Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N', N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High level kasugamycin resistance protein ksgA) (Kasugamycin dimethyltransferase)	rRNA and rProtein modifications/processing
MPN682	rpmH	50S ribosomal protein L34	Ribosome
MPN683	bceA	Similar to bacitracin export ATP-binding protein B	Unassigned transporters
MPN004	gyrA	DNA gyrase subunit A (EC 5.99.1.3)	DNA synthesis/modifications/repair/conformation
MPN020	hepA	Uncharacterized ATP-dependent helicase (EC 3.6.1.-)	RNA synthesis + DNA synthesis/modifications/repair/conformation
MPN051	glpD	Glycerol-3-phospate dehydrogenase (EC 1.1.99.5)	Lipid/CoA metabolism + Pathogenesis
MPN064	deoA	Thymidine phosphorylase (EC 2.4.2.4) (TdRPase)	Nucleotide metabolism/ Salvage
MPN115	infC	Translation initiation factor IF-3	Protein synthesis factors
MPN122	parE	DNA topoisomerase 4 subunit B (EC 5.99.1.-) (Topoisomerase IV subunit B)	DNA synthesis/modifications/repair/conformation
MPN165	rplC	50S ribosomal protein L3	Ribosome
MPN167	rplW	50S ribosomal protein L23	Ribosome
MPN168	rplB	50S ribosomal protein L2	Ribosome
MPN171	rpsC	30S ribosomal protein S3	Ribosome
MPN176	rplX	50S ribosomal protein L24	Ribosome
MPN192	rplQ	50S ribosomal protein L17	Ribosome
MPN197	pepF	Oligoendopeptidase F homolog (EC 3.4.24.-)	Proteases
MPN226	rpsG	30S ribosomal protein S7	Ribosome
MPN228	rpsF	30S ribosomal protein S6	Ribosome
MPN261	topA	DNA topoisomerase 1 (EC 5.99.1.2) (DNA topoisomerase I) (Omega-protein) (Relaxing enzyme) (Untwisting enzyme) (Swivelase)	DNA synthesis/modifications/repair/conformation
MPN266	spxA	Transcriptional regulator Spx	Transcription factors
MPN280	rnjA	Ribonuclease J1 (RNase J1) (EC 3.1.-.-)	RNA degradation

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN294	yajL2	DJ-1/PfpI family protein, protein deacylase (EC:3.5.1.124)	Protein homeostasis + Oxidative homeostasis
MPN295	mpn295	Conserved hypothetical protein MPN_295	Glycocalyx + Proteases
MPN303	pyk	Pyruvate kinase (PK) (EC 2.7.1.40)	Glycolysis
MPN325	rplU	50S ribosomal protein L21	Ribosome
MPN352	sigA	RNA polymerase sigma factor rpoD (Sigma-A) (EC 2.7.7.6)	RNA synthesis
MPN430	gap	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (EC 1.2.1.12)	Glycolysis
MPN475	engA	Ribosome assembly GTP-binding protein engA	Ribosome
MPN516	rpoB	DNA-directed RNA polymerase subunit beta (RNAP subunit beta) (EC 2.7.7.6) (Transcriptase subunit beta) (RNA polymerase subunit beta)	RNA synthesis
MPN520	ileS	Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine–tRNA ligase) (IleRS)	tRNA synthases
MPN549	recJ	Probable single-stranded-DNA-specific exonuclease recJ (EC 3.1.-.-)	DNA recombination
MPN619	uvrA	UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A)	DNA synthesis/modifications/repair/conformation
MPN621	rnjB	Non-catalytic ribonuclease J2	RNA degradation
MPN658	rplS	50S ribosomal protein L19	Ribosome
MPN010	mpn010	Uncharacterized protein MPN_010	Conserved hypothetical proteins
MPN014	mpn014	DNA primase-related protein (EC:2.7.7.-)	DNA recombination
MPN028	trsB	Uncharacterized glycosyl transferase	Glycocalyx
MPN032	yajL	DJ-1/PfpI family protein, protein deglycase (EC:3.5.1.124)	Oxidative homeostasis
MPN035	mpn035	Conserved hypothetical protein MPN_035	Conserved hypothetical proteins
MPN037	mpn037	Uncharacterized protein MPN_037	Conserved hypothetical proteins
MPN038	mpn038	Uncharacterized protein MPN_038	Cytoskeleton
MPN039	mpn039	Conserved hypothetical protein MPN_039	Conserved hypothetical proteins
MPN040	mpn040	Conserved hypothetical protein MPN_040	Conserved hypothetical proteins
MPN041	mpn041	Conserved hypothetical protein MPN_041	Conserved hypothetical proteins
MPN042	mpn042	Conserved hypothetical protein MPN_042	Conserved hypothetical proteins
MPN048	mpn048	Uncharacterized protein MPN_048	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN049	mpn049	Uncharacterized protein MPN_049	Lipid/CoA metabolism
MPN054	mpn054	Conserved hypothetical lipoprotein MPN_054	Lipoprotein
MPN085	mpn085	Uncharacterized protein MPN_085	Conserved hypothetical proteins
MPN086	mpn086	Uncharacterized protein MPN_086	Conserved hypothetical proteins
MPN087	mpn087	Uncharacterized protein MPN_087	Conserved hypothetical proteins
MPN088	mpn088	Uncharacterized protein MPN_088	Conserved hypothetical proteins
MPN089	hsdS	Putative type-1 restriction enzyme specificity protein MPN_089 (S.mpnORFAP) (Type I restriction enzyme specificity protein MPN_089) (S protein)	DNA synthesis/modifications/repair/conformation
MPN091	mpn091	Conserved hypothetical protein protein MPN_091	Conserved hypothetical proteins
MPN096	mpn096	Uncharacterized amino acid permease	AA metabolism
MPN097	mpn097	Conserved hypothetical lipoprotein MPN_097	Lipoprotein
MPN098	mpn098	Conserved hypothetical lipoprotein MPN_098	Lipoprotein
MPN103	mpn103	Uncharacterized protein MPN_103	Conserved hypothetical proteins
MPN107	mpn107	Possible uncharacterized DNA methylase	DNA synthesis/modifications/repair/conformation
MPN108	mpn108	Uncharacterized adenine-specific methylase MPN_108 (EC 2.1.1.72)	DNA synthesis/modifications/repair/conformation
MPN110	mpn110	Conserved hypothetical protein MPN_110	DNA synthesis/modifications/repair/conformation
MPN111	mpn111	Uncharacterized adenine-specific methylase MPN_111 (EC 2.1.1.72)	DNA synthesis/modifications/repair/conformation
MPN112	cglT	Major facilitator family protein CglT	Fermentation
MPN113	mpn113	Putative uncharacterized permease MPN_113 (major facilitator family protein)	Fermentation
MPN114	cpt2	Putative acetyltransferase MPN_114 (EC 2.3.1.-)	Signal transduction
MPN129	mpn129	Uncharacterized protein MPN_129	Conserved hypothetical proteins
MPN131	mpn131	Putative adhesin P1-like protein MPN_131	Adhesins/attachment organelle
MPN143	mpn143	Uncharacterized protein MPN_143	Conserved hypothetical proteins

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN146	mpn146	Conserved hypothetical protein MPN_146	Conserved hypothetical proteins
MPN147	mpn147	Conserved hypothetical protein MPN_147	Conserved hypothetical proteins
MPN150	mpn150	Putative mgpC-like protein MPN_150	Adhesins/attachment organelle
MPN160	mpn160	Conserved hypothetical protein MPN_160	Conserved hypothetical proteins
MPN201	mpn201	Putative type-1 restriction enzyme specificity protein MPN_201 (S.mpnORFFP) (Type I restriction enzyme specificity protein MPN_201) (S protein)	DNA synthesis/modifications/repair/conformation
MPN203	mpn203	Putative adhesin P1-like protein MPN_203	Adhesins/attachment organelle
MPN206	mpn206	Uncharacterized protein MPN_206	Conserved hypothetical proteins
MPN212	mpn212	Conserved hypothetical protein MPN_212	Conserved hypothetical proteins
MPN242	secG	Probable protein-export membrane protein secG	Protein secretion
MPN249	rsgA	Putative ribosome biogenesis GTPase rsgA (EC 3.6.1.-)	Ribosome
MPN270	mpn270	Conserved hypothetical proteins	Conserved hypothetical proteins
MPN282	mpn282	Conserved hypothetical protein MPN_282	Conserved hypothetical proteins
MPN283	mpn283	Uncharacterized protein MPN_283	Cytoskeleton
MPN289	hsdS1B	Putative type-1 restriction enzyme specificity protein MPN_289 (S.mpnORFEBP) (Type I restriction enzyme specificity protein MPN_289) (S protein)	DNA synthesis/modifications/repair/conformation
MPN290	mpn290	Putative type-1 restriction enzyme specificity protein MPN_290 (S.mpnORFEAP) (Type I restriction enzyme specificity protein MPN_290) (S protein)	DNA synthesis/modifications/repair/conformation
MPN304	arcA	Putative arginine deiminase (ADI) (EC 3.5.3.6) (Arginine dihydrolase) (AD) N-terminal fragment	Arg metabolism
MPN313	mpn313	Conserved hypothetical protein MPN_313	DNA recombination
MPN334	bcrA	Putative ABC transporter ATP-binding protein MPN_334	Unassigned transporters
MPN345	hsdR	Putative type-1 restriction enzyme mpnORFDP R protein part 2 (EC 3.1.21.3) (Putative type I restriction enzyme mpnORFDP R protein part 2) (mpnORFDBP)	DNA synthesis/modifications/repair/conformation

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN346	mpn346	Uncharacterized protein MPN_346	Conserved hypothetical proteins
MPN347	hsdR1	Putative type I restriction enzyme mpnORFDP R protein part 1 (EC 3.1.21.3) (mpnORFDAP)	DNA synthesis/modifications/repair/conformation
MPN363	mpn363	Conserved hypothetical lipoprotein MPN_363	Lipoprotein
MPN371	mpn371	Uncharacterized protein MPN_371	Conserved hypothetical proteins
MPN373	mpn373	Uncharacterized protein MPN_373	Conserved hypothetical proteins
MPN374	mpn374	Uncharacterized protein MPN_374	Conserved hypothetical proteins
MPN375	mpn375	Uncharacterized protein MPN_375	Conserved hypothetical proteins
MPN403	mpn403	Conserved hypothetical protein MPN_403	Conserved hypothetical proteins
MPN404	mpn404	Conserved hypothetical protein MPN_404	Conserved hypothetical proteins
MPN405	mpn405	Conserved hypothetical protein MPN_405	Conserved hypothetical proteins
MPN411	mpn411	Conserved hypothetical lipoprotein MPN_411	Lipoprotein
MPN413	mpn413	Uncharacterized protein MPN_413	Conserved hypothetical proteins
MPN417	thiP	ABC thiamine importer, permease subunit, thiP	Vitamin/Folate metabolism
MPN419	ruvX	Putative Holliday junction resolvase (EC 3.1.-.-)	DNA synthesis/modifications/repair/conformation + DNA recombination
MPN438	mpn438	Uncharacterized protein MPN_438	Conserved hypothetical proteins
MPN439	mpn439	Uncharacterized lipoprotein MPN_439	Lipoprotein
MPN441	mpn441	Conserved hypothetical protein MPN_441	Conserved hypothetical proteins
MPN442	mpn442	Uncharacterized lipoprotein MPN_442	Lipoprotein
MPN448	ecfS	Energy-coupling factor transport system substrate-binding protein ecfS, putative riboflavin transporter	Vitamin/Folate metabolism
MPN451	comE3	ComE operon protein 3 related protein	Nucleotide metabolism/Salvage
MPN463	mpn463	Conserved hypothetical protein MPN_463	Conserved hypothetical proteins
MPN510	mpn510	Uncharacterized protein MPN_510	Conserved hypothetical proteins
MPN511	mpn511	Uncharacterized protein MPN_511	Conserved hypothetical proteins
MPN512	mpn512	Uncharacterized protein MPN_512	Conserved hypothetical proteins
MPN513	mpn513	Conserved hypothetical protein MPN_513	Conserved hypothetical proteins
MPN514	mpn514	Conserved hypothetical protein MPN_514	Conserved hypothetical proteins
MPN528a	recU	Recombination protein U	DNA recombination

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN534	mpn534	Conserved hypothetical protein MPN_534	DNA recombination
MPN535	ruvA	Holliday junction ATP-dependent DNA helicase ruvA (EC 3.6.1.-)	DNA recombination + DNA synthesis/modifications/repair/conformation
MPN536	ruvB	Holliday junction ATP-dependent DNA helicase ruvB (EC 3.6.1.-)	DNA recombination + DNA synthesis/modifications/repair/conformation
MPN570	mpn570	Conserved hypothetical proteinMPN_570	Conserved hypothetical protein
MPN571	lcnDR3	ABC-bacteriocin transporter-related extracellular peptidase; ATP-binding protein	Protein secretion + DNA recombination
MPN577	mpn577	Conserved hypothetical protein MPN_577	Conserved hypothetical protein
MPN578	mpn578	Conserved hypothetical protein MPN_578	Conserved hypothetical protein
MPN579	mpn579	Conserved hypothetical protein MPN_579	Conserved hypothetical protein
MPN580	mpn580	Putative protease MPN_580	Membrane Protease
MPN581	mpn581	Conserved hypothetical protein MPN_581	Conserved hypothetical protein
MPN583	mpn583	Conserved hypothetical protein MPN_583	Conserved hypothetical protein
MPN584	mpn584	Conserved hypothetical protein MPN_584	Conserved hypothetical protein
MPN586	mpn586	Conserved hypothetical protein MPN_586	Lipoprotein
MPN587	mpn587	Conserved hypothetical lipoprotein MPN_587	Lipoprotein
MPN589	mpn589	Conserved hypothetical protein MPN_589	Conserved hypothetical protein
MPN594	mpn594	Conserved hypothetical protein MPN_594	Conserved hypothetical protein
MPN612	mpn612	Conserved hypothetical protein MPN_612	DNA recombination
MPN613	mpn613	Conserved hypothetical protein MPN_613	DNA recombination
MPN614	mpn614	Conserved hypothetical protein MPN_614	DNA recombination
MPN626	mpn626	Alternative sigma factor	DNA recombination + Transcription factor
MPN633	mpn633	Conserved hypothetical protein MPN_633	Conserved hypothetical proteins
MPN634	mpn634	Conserved hypothetical protein MPN_634	Conserved hypothetical proteins
MPN635	mpn635	Conserved hypothetical protein MPN_635	Conserved hypothetical proteins
MPN637	cdsA	Putative phosphatidate cytidylyltransferase (EC 2.7.7.41) (CDP-diacylglycerol synthase) (CDS) (CTP:phosphatidate cytidylyltransferase) (CDP-DG synthetase) (CDP-DAG synthase)	Lipid/CoA metabolism

Continued on next page

Table A.1 – *Continued from previous page*

ID	Gene Name	Protein Name	Functional category
MPN648	mpn648	Conserved hypothetical lipoprotein MPN_648	Lipoproteins
MPN649	mpn649	Uncharacterized protein MPN_649	Lipoproteins
MPN650	mpn650	Uncharacterized lipoprotein MPN_650	Lipoproteins
MPN651	mtlA	PTS system mannitol-specific EIICB component (EIICB-Mtl) (EII-Mtl)	Glycolysis
MPN666	mpn666	Conserved hypothetical protein MPN_666	Conserved hypothetical proteins
MPN676	mpn676	Uncharacterized protein MPN_676	Conserved hypothetical proteins
MPN681	rnpA	Ribonuclease P protein component (RNaseP protein) (RNase P protein) (EC 3.1.26.5) (Protein C5)	RNA degradation

Table A.2: Cell lines and plasmids used in this study. In type of modification: OE means gene overexpression, DR means gene disruption, Mut means point mutation (the two letters after the gene number indicates the wild type and mutated amino acid in one letter code), Del means deletion mutant, and Fus means fusion. Note: TF149 had a mutation in the flag-tag. The essentiality categories (EssC) were taken from Lluch-Senar et al. (2015b). Growth curves (GC) were done in duplicates, a first screening (S) was performed by medium color (pH) change method and intracellular protein at 48 h in batch cultures and a validation (V) was performed by intracellular protein in semicontinuous culture method.

	Gene	Cell				GC	
ID	Name	line	Plasmid	Type	EssC	S	V
		WT				1	1
		TF167	pMT85-tuf->0-flag			1	1
		TAP	pMT85-clpB->0-TAP			1	1
	YFP	TF42	pMT85-tuf->YFP			1	
MPN002	cbpA	TF201	pMT85-tuf->flag-002	OE	F	1	
MPN004	gyrA	TF267	pMT85-tuf->flag-004	OE	E	1	
MPN015	rimK	TF246	pMT85-tuf->flag-015	OE	F	1	1
MPN020	hepA	TF210	pMT85-tuf->flag-020	OE	E	1	1
MPN024	rpoE	TF4	pMT85-tuf->rpoE-flag	OE	E	1	
		TF499	pMT85-tuf->024	OE		1	
		TF500	pMT85-tuf->024-P	Mut		1	
		TF501	pMT85-tuf->024-A	Mut		1	
MPN025	tsr	TF462	pMT85-tuf->025	OE	E	1	1
MPN027	rimL	TF291	pMT85-tuf->flag-027	OE	F	1	
		Tn027		DR		1	
MPN030	nusB	TF202	pMT85-tuf->flag-030	OE	NE	1	1
MPN032	yfkM	TF21	pMT85-tuf->yfkM-flag	OE	E	1	
MPN033	upp	TF169	pMT85-tuf->flag-upp	OE	NE	1	
MPN038		TF247	pMT85-tuf->flag-038	OE	NE	1	1
MPN051	glpD	TF413	pMT85-tuf->051-flag	OE	NE	1	
		Tn051		DR		1	1
							1
MPN053	ptsH	TF482	pMT85-tuf->053	OE	E	1	1
		TF483	pMT85-tuf->053SE	Mut		1	
		TF484	pMT85-tuf->053SA	Mut		1	
MPN055	potA	TF406	pMT85-tuf->flag-055	OE	E	1	
MPN063	deoC	TF170	pMT85-tuf->flag-deoC	OE	E	1	
MPN064	deoA	TF171	pMT85-tuf->flag-deoA	OE	F	1	1
MPN067	nusG	TF203	pMT85-tuf->flag-067	OE	E	1	

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN069	rpmG2	TF208	pMT85-tuf->flag-069	OE	E	1	
MPN076	uhpT	Tn076		DR	F	1	
MPN077	uhpT2	Tn077		DR	NE	1	
MPN081	glnQ	TF407	pMT85-tuf->flag-081	OE	E	1	
MPN082	tklB	TF172	pMT85-tuf->flag-tlkB	OE	F	1	1
MPN106	pheT	TF188	pMT85-tuf->flag-pheT	OE	E	1	
MPN114	cpt2	TF292	pMT85-tuf->flag-114	OE	NE	1	
		Tn114		DR		1	1
MPN119	topJ	TF477	pMT85-tuf->119-flag	OE	E	1	
MPN122	parE	TF212	pMT85-tuf->flag-122	OE	E	1	
MPN124	hrcA	TF5	pMT85-tuf->hrcA-flag	OE	E	1	
		TF159	pMT85-tuf->flag-hrcA	OE		1	
		TF298	pMT85-tuf->hrcATA-flag	Mut		1	
MPN127		TF249	pMT85-tuf->flag-127	OE	NE*	1	
MPN133		Tn133		DR	F	1	
MPN140	nrnA	TF250	pMT85-tuf->flag-140	OE	E	1	
MPN148		TF224	pMT85-tuf->flag-148	OE	NE*	1	1
MPN154	nusA	TF39	pMT85-p->nusA-2xflag	OE	E	1	
MPN159	tlyC	TF225	pMT85-tuf->flag-C159	Del	F	1	1
		Tn159		DR		1	
MPN162		Tn162		DR	E	1	1
MPN164	rpsJ	TF204	pMT85-tuf->flag-164	OE	E	1	1
MPN165	rplC	TF310	pMT85-tuf->flag-165	OE	E	1	1
MPN166	rplD	TF311	pMT85-tuf->flag-166	OE	E	1	
MPN168	rplB	TF312	pMT85-tuf->flag-168	OE	E	1	
MPN173	rpmC	TF205	pMT85-tuf->flag-173	OE	E	1	1
MPN178	rpsN	TF211	pMT85-tuf->flag-178	OE	E	1	
MPN191	rpoA	TAP191	pMT85-clpB->191-TAP	OE	E	1	1
		TAP191R	pMT85-clpB->191-TAP	OE		1	1
		clpB-191	pMTCmLox-clpB->191-TAP	OE		1	1
		tuf-191	pMTCmLox-tuf->191-TAP	OE		1	1
		rpmJ-191	pMTCmLox-rmpJ->191-TAP	OE		1	1
MPN192	rplQ	TF313	pMT85-tuf->flag-192	OE	E	1	
MPN194	CbiO 2	TF242	pMT85-tuf->flag-194	OE	E	1	1
MPN197	pepF	TF232	pMT85-tuf->flag-197	OE	E	1	
MPN208	rpsB	TF314	pMT85-tuf->flag-208	OE	E	1	

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN222	yacA	TF412	pMT85-tuf->flag-222	OE	E	1	
MPN223	hprK	TF176	pMT85-tuf->hprK-flag	OE	F	1	
		Tn223		DR		1	1
MPN229	ssbA	TF230	pMT85-tuf->flag-229	OE	E	1	
MPN239	gntR	TF6	pMT85-tuf->gntR-flag	OE	E	1	
		TF160	pMT85-tuf->flag-gntR	OE		1	
		TF392	pMT85-tuf->-239	OE		1	
		TF403	pMT85-tuf->239RE	Mut		1	
MPN241	whiA	TF29	pMT85-tuf->whiA-flag	OE	F	1	1
		TF240	pMT85-tuf->flag-N241	Del		1	1
		TF241	pMT85-tuf->flag-C241	Del		1	1
		Tn241		DR		1	1
MPN243	vacB	TF293	pMT85-tuf->flag-243	OE	E	1	
MPN244	cdaS	TF168	pMT85-tuf->disA-flag	OE	E	1	
MPN246	gmk	TF297	pMT85-tuf->flag-246	OE	E	1	1
MPN247	ptc1	TF25	pMT85-tuf->pp2c-flag	OE	F	1	1
		Tn247		DR		1	
MPN248	prkC	TF22	pMT85-tuf->prkC-flag	OE	F	1	
		Tn248		DR		1	
MPN250	pgiB	TF178	pMT85-tuf->flag-pgiB	OE	E	1	
MPN252	asnS	TF206	pMT85-tuf->flag-252	OE	E	1	
MPN255	ygbP	TF179	pMT85-tuf->flag-ygbP	OE	F	1	1
MPN263	trx	TF510	pMT85-tuf->263	OE	E	1	
MPN265	trpS	TF207	pMT85-tuf->flag-265	OE	E	1	1
MPN266	spxA	TF7	pMT85-tuf->spxA-flag	OE	E	1	
		TF299	pMT85-tuf->spxACS-flag	Mut		1	1
		TF393	pMT85-266->266	OE		1	
		TF394	pMT85-266->266CS	Mut		1	
		TF395	pMT85-266->266GR	Mut		1	1
		TF396	pMT85-266->266RE	Mut		1	
MPN269	ymdA	TF294	pMT85-tuf->flag-269	OE	E	1	
MPN273	hit1	TF27	pMT85-tuf->hit1-flag	OE	E	1	
MPN275	ybaB	TF397	pMT85-tuf->flag-275	OE	F	1	
		TF398	pMT85-tuf->275	OE		1	
		TF409	pMT85-tuf->275DN	Mut		1	1
		Tn275		DR		1	

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN280	rnjA	TF295	pMT85-tuf->flag-280	OE	E	1	
MPN284		Tn284		DR	NE*	1	
MPN287		TF259	pMT85-tuf->flag-287	OE	NE*	1	
MPN294	pfpI	TF194	pMT85-tuf->flag-294	OE	NE	1	
		TF517	pMT85-tuf->294-flag	OE		1	
		TF399	pMT85-tuf->294	OE		1	
		TF433	pMT85-tuf->294DN	Mut		1	
		Tn294		DR		1	
MPN295		TF8	pMT85-tuf->295-flag	OE	NE	1	
		TF404	pMT85-tuf->295	OE		1	
		TF410	pMT85-tuf->295DN	Mut		1	
		TF502	pMT85-tuf->295-P	Mut		1	
		TF503	pMT85-tuf->295-A	Mut		1	
MPN300	scpA	TF251	pMT85-tuf->flag-300	OE	E	1	1
MPN301	scpB	TF217	pMT85-tuf->flag-301	OE	E	1	1
		TF265	pMT85-tuf->301-flag	OE		1	
MPN303	pyk	TF173	pMT85-tuf->flag-pyk	OE	E	1	1
MPN314	mraZ	TF9	pMT85-tuf->mraz-flag	OE	E	1	
		TF164	pMT85-tuf->mraZR->A-flag	Mut		1	
		TF275	pMT85-tuf->flag-314	OE		1	
		E10	pTnTc-flag-mraZ-delC	Del			
MPN315	mraW	TF243	pMT85-tuf->flag-315	OE	F	1	1
		B10	pTnTc-mraW-flag	OE			
MPN316	ftsA	TF226	pMT85-tuf->flag-316	OE	NE	1	
MPN329	fur	TF10	pMT85-tuf->fur-flag	OE	F	1	
		TF161	pMT85-tuf->flag-fur	OE		1	1
		TF480	pMT85-tuf->329	OE		1	
		TF481	pMT85-tuf->329DN	Mut		1	
		TAP329	pMT85-clpB->329-TAP	OE			
MPN330		TF195	pMT85-tuf->flag-330	OE	F	1	
MPN332	lon	TF186	pMT85-tuf->flag-lon	OE	E	1	
MPN348	mthfs	TF227	pMT85-tuf->flag-348	OE	F	1	
MPN349	ymdB	TF24	pMT85-tuf->ymdB-flag	OE	E	1	1
MPN352	sigA	TF14	pMT85-tuf->sigA-flag	OE	E	1	1
		TF200	pMT85-tuf->N352-flag	Del		1	
MPN368		TF260	pMT85-tuf->flag-368	OE	NE*	1	

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN372	ptxA	Tn372		DR	NE	1	
MPN397	spoT	TF220	pMT85-p397->397-flag	OE	NE	1	1
							1
		TAP397	pMT85-clpB->397-TAP	OE			
		Tn397		DR		1	1
		TF488	pMT85-p674.1->397	OE		1	
		TF489	pMT85-p397->397	OE		1	
		TF504	pMT85-p674.1->397HY	Mut		1	
		TF505	pMT85-p397->397HY	Mut		1	
		TF506	pMT85-p674.17->397SY	Mut		1	
		TF507	pMT85-p397->397SY	Mut		1	
MPN400		TF40	pMT85-tuf->400-flag	OE	F	1	
MPN401	greA	TF177	pMT85-tuf->flag-greA	OE	E	1	
MPN420	glpQ	TF180	pMT85-tuf->flag-glpQ	OE	F	1	
		TF518	pMT85-tuf->420-flag	OE		1	
		Tn420		DR		1	
MPN421	cglT1	Tn421		DR	F	1	
MPN424	ylxM	TF11	pMT85-tuf->ylxM-flag	OE	NE	1	
		TF400	pMT85-tuf->424	OE		1	
		TF411	pMT85-tuf->424DN	Mut		1	
		TF487	pMT85-tuf->flag-424	OE		1	
		Tn424		DR		1	
MPN426	smc	*TF149*	*pMT85-tuf->smc-flag*	OE	E	1	
		TF166	pMT85-tuf->smc-flag	OE		1	
		TF266	pMT85-tuf->flag-426	OE		1	
MPN428	pta	TF174	pMT85-tuf->flag-pta	OE	F	1	1
MPN430	gap	TF414	pMT85-tuf->430-flag	OE	E	1	1
MPN440		TF261	pMT85-tuf->flag-440	OE	NE*	1	
MPN443	deaD	TF209	pMT85-tuf->flag-443	OE	E	1	
		TF214	pMT85-tuf->flag-443	OE		1	
MPN446	rpsD	TF315	pMT85-tuf->flag-446	OE	E	1	
MPN473	lip2	TF196	pMT85-tuf->flag-473	OE	E	1	1
MPN475	engA	TF229	pMT85-tuf->flag-475	OE	E	1	1
MPN478	YebC	TF162	pMT85-tuf->flag-478	OE	E	1	
		TF419	pMT85-tuf->478	OE		1	1
							1

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN481	engB	TF233	pMT85-tuf->flag-481	OE	E	1	
MPN482	yneF	TF234	pMT85-tuf->flag-C482	Del	E	1	
MPN484		TF197	pMT85-tuf->flag-484	OE	NE*	1	
MPN485		TF264	pMT85-tuf->flag-485	OE	NE*	1	
MPN487	nifS	TF175	pMT85-tuf->487-flag	OE	E	1	
MPN490	recA	TF12	pMT85-tuf->recA-flag	OE	F	1	
		TF244	pMT85-tuf->flag-490	OE		1	1
		Tn490		DR		1	
MPN499		TF254	pMT85-tuf->flag-499	OE	NE	1	
MPN506		Tn506		DR	NE*	1	
MPN507		TF221	pMT85-tuf->flag-507	OE	NE	1	
MPN516	rpoB	TAP516	pMT85-clpB->516-TAP	OE	E	1	1
		TAP516R	pMT85-clpB->516-TAP	OE		1	1
MPN518		TF41	pMT85-tuf->518-flag	OE	NE	1	
MPN525	dnaB1	TF252	pMT85-tuf->flag-525	OE	E	1	
MPN526		TF18	pMT85-tuf->526-flag	OE	E	1	
MPN529	ihf	TF19	pMT85-tuf->Ihf-flag	OE	F	1	
		TF253	pMT85-tuf->flag-529	OE		1	
		TF415	pMT85-tuf->529	OE		1	
		TF416	pMT85-tuf->529DN	Mut		1	
MPN545	rnc	TF296	pMT85-tuf->flag-545	OE	F	1	1
		Tn545		DR		1	
MPN547	fak	TF263	pMT85-tuf->flag-547	OE	E	1	
MPN549	recJ	TF255	pMT85-tuf->flag-549	OE	E	1	
MPN554	ssbB	TF231	pMT85-tuf->flag-554	OE	E	1	
MPN555		TF17	pMT85-tuf->555-flag	OE	E	1	
MPN559		TF262	pMT85-tuf->flag-559	OE	E	1	
MPN563	obg	TF26	pMT85-tuf->obg-flag	OE	E	1	1
MPN566	glpQ1	TF181	pMT85-tuf->flag-glpQ	OE	NE	1	
	glpQ-2	Tn566		DR		1	1
MPN568	spg	TF245	pMT85-tuf->flag-568	OE	E	1	1
MPN569	ybeY	TF235	pMT85-tuf->flag-569	OE	E	1	1
MPN572	pepA	TF185	pMT85-tuf->flag-pepA	OE	E	1	
		TF236	pMT85-tuf->flag-572	OE		1	
		TF446	pMT85-tuf->572	OE		1	
		TF447	pMT85-tuf->572DN	Mut		1	

Continued on next page

Table A.2 – *Continued from previous page*

ID	Gene Name	Cell line	Plasmid	Type	EssC	GC S	GC V
MPN574	groES	TF237	pMT85-tuf->flag-574	OE	E	1	
		TAP574R	pMT85-clpB->574-TAP	OE		1	1
MPN576	glyA	TF182	pMT85-tuf->flag-glyA	OE	E	1	
		TF198	pMT85-tuf->flag-576	OE		1	
MPN590		TF238	pMT85-tuf->flag-590	OE	NE	1	
MPN606	eno	TF183	pMT85-tuf->flag-eno	OE	E	1	1
MPN608	phoU	TF23	pMT85-tuf->phoU-flag	OE	E	1	
		TF417	pMT85-tuf->608	OE		1	
		TF418	pMT85-tuf->608DN	Mut		1	
MPN615	hsdS1	TF222	pMT85-tuf->flag-615	OE	E	1	
MPN617	rplM	TF316	pMT85-tuf->flag-617	OE	E	1	
MPN621	rnjB	TF199	pMT85-tuf->flag-621	OE	E	1	
MPN626		TF16	pMT85-tuf->626-flag	OE	NE	1	
		TF163	pMT85-tuf->flag-626	OE		1	
		TF405	pMT85-tuf->626	OE		1	1
MPN627	ptsI	TF184	pMT85-tuf->flag-pstI	OE	E	1	1
MPN629	tim	TF465	pMT85-tuf->629	OE	E	1	1
MPN633		TF256	pMT85-tuf->flag-633	OE	NE*	1	
MPN634		TF257	pMT85-tuf->flag-633-4	OE	NE	1	
		TF486	pMT85-tuf->flag-633Q4	Fus		1	
		TF13	pMT85-tuf->relA-flag	OE		1	
MPN635		TF258	pMT85-tuf->flag-635	OE	NE	1	
MPN638	hsdS2	TF219	pMT85-tuf->flag-638	OE	NE	1	
MPN651	mtlA	TF2	pMT85-Mtl-flag	OE	NE	1	
MPN652	mtlD	”		OE	NE		
MPN653	mtlF	”		OE	NE		
MPN663	ytpR	TF239	pMT85-tuf->flag-663	OE	E	1	
MPN667	gtaB	TF213	pMT85-tuf->flag-667	OE	E	1	
MPN673	ygbB	TF268	pMT85-tuf->flag-673	OE	E	1	
MPN674	ldh	TF187	pMT85-tuf->flag-674	OE	E	1	
		Tn674		DR		1	
MPN677	ywfO	TF248	pMT85-tuf->flag-677	OE	E	1	
MPN683	bceA	TF408	pMT85-tuf->flag-683	OE	F	1	
MPN686	dnaA	TF165	pMT85-p688->dnaA-flag	OE	E	1	

Table A.3: Growth phenotypes and doubling times (DT) of cell lines showing an effect in growth. In type of modification: OE means gene overexpression, DR means gene disruption, Mut means point mutation (the two letters after the gene number indicates the wild type and mutated amino acid in one letter code), and Del means deletion mutant. Growth curves were done in duplicates, a first screening was performed by medium color (pH) change curve method and intracellular protein at 48 h in batch cultures, in which there are two indicators of growth (G) (protein and the early slope of the medium color curve) and two of metabolism (M) (the $A_{430/560}$ value of the medium at the highest slope of the curve and the the value of this last slope (late)). A validation was performed by intracellular protein in semicontinuous culture method and is shown per batch experiment. The DT measured by protein at time points 0 and 48 h was corrected using the equation 2.5.

				Screening		Validation				
							DT		Corrected DT	
ID	Gene Name	Cell line	Type	G	M	Batch	Mean	SD	Mean	SD
		WT				1	8.99	0.0928	9.30	0.0469
		WT				2	9.30	0.1512	9.07	0.1051
		WT				3	8.51	0.0568	9.01	0.0551
		WT				4	8.11	0.2576	8.57	0.1003
		TF167				1	10.16	0.0662	9.88	0.0494
		TF167				2	9.58	0.0962	9.22	0.0250
		TF167				3	8.83	0.1476	8.93	0.0820
		TF167				4	8.86	0.0508	8.97	0.0403
		TAP				4	8.29	0.1879	8.67	0.0941
MPN002	cbpA	TF201	OE	Slow						
MPN004	gyrA	TF267	OE	Slow						
MPN015	rimK	TF246	OE	Fast		1	9.17	0.1299	9.53	0.0941
MPN020	hepA	TF210	OE		Slow	3	9.65	0.1490	9.69	0.0839
MPN025	tsr	TF462	OE			4	9.07	0.0562	9.00	0.0492
MPN030	nusB	TF202	OE			1	8.75	0.0537	8.93	0.0352
MPN038		TF247	OE			2	9.56	0.2252	9.48	0.1684
MPN051	glpD	TF413	OE		Slow					
		Tn051	DR	Slow	Slow	2	9.85	0.1585	9.48	0.0734
		Tn051	DR	Slow	Slow	3	10.12	0.2026	10.82	0.1788
MPN053	ptsH	TF482	OE			4	9.31	0.1243	9.01	0.1211
MPN064	deoA	TF171	OE	Fast		1	9.22	0.1105	9.74	0.0835
MPN076	uhpT	Tn076	DR		Slow					
MPN082	tklB	TF172	OE	Fast	Slow	2	8.97	0.1009	9.32	0.0899
MPN106	pheT	TF188	OE	Slow						
MPN114	cpt2	TF292	OE		Slow					
		Tn114	DR		Fast	2	9.41	0.1341	9.79	0.0986

Continued on next page

Table A.3 – *Continued from previous page*

ID	Gene Name	Cell line	Type	Screening		Validation				
							DT		Corrected DT	
				G	M	Batch	Mean	SD	Mean	SD
MPN122	parE	TF212	OE		Slow					
MPN124	hrcA	TF298	Mut		Fast					
MPN148		TF224	OE	Fast	Fast	1	10.20	0.1263	9.06	0.0995
MPN159	tlyC	TF225	Del			3	8.92	0.1148	9.11	0.0616
MPN162		Tn162	DR		Fast	4	9.78	0.1283	9.52	0.1099
MPN164	rpsJ	TF204	OE		Fast	1	8.97	0.2118	9.14	0.0850
MPN165	rplC	TF310	OE			3	7.06	0.1707	8.42	0.1447
MPN173	rpmC	TF205	OE			1	8.64	0.1490	8.83	0.1244
MPN191	rpoA	TAP191	OE	Fast	Fast	1	7.98	0.1116	8.92	0.0761
		TAP191R	OE			4	7.93	0.1926	8.54	0.1151
		clpB-rpoA$_{G37}$	OE			4	8.94	0.0440	8.58	0.0346
		tuf-rpoA$_{G37}$	OE			4	9.53	0.0828	9.10	0.0423
		rpmJ-rpoA$_{G37}$	OE			4	8.44	0.1595	8.54	0.0542
MPN194	CbiO 2	TF242	OE			2	9.68	0.0842	9.55	0.0535
MPN223	hprK	TF176	OE	Fast						
		Tn223	DR	Slow		3	8.55	0.5146	9.81	0.5030
MPN239	gntR	TF6	OE	Slow	Fast					
MPN241	whiA	TF29	OE	Slow		4	9.83	0.1409	9.60	0.0642
		TF240	Del			4	9.05	0.0446	9.11	0.0211
		TF241	Del			4	8.98	0.1639	8.92	0.0467
		Tn241	DR	Slow	Slow	4	9.85	0.1937	9.70	0.0967
MPN246	gmk	TF297	OE			3	7.62	0.1174	8.75	0.0862
MPN247	ptc1	TF25	OE			3	8.15	0.0973	8.98	0.0666
MPN252	asnS	TF206	OE	Fast						
MPN255	ygbP	TF179	OE	Slow	Slow	3	9.05	0.1017	9.29	0.0678
MPN265	trpS	TF207	OE			2	8.97	0.1661	8.64	0.0708
MPN266	spxA	TF299	Mut			3	8.05	0.4269	9.08	0.2779
		TF394	Mut	Slow	Slow					
		TF395	Mut			1	8.44	0.1140	9.31	0.0778
MPN275	ybaB	TF409	Mut			2	9.56	0.3190	9.10	0.0955
MPN300	scpA	TF251	OE	Slow	Slow	4	9.20	0.1462	9.19	0.0806
MPN301	scpB	TF217	OE			2	10.06	0.1451	9.31	0.1333
MPN303	pyk	TF173	OE		Slow	2	9.23	0.1880	9.15	0.0755
MPN315	mraW	TF243	OE			1	9.02	0.2411	9.67	0.1815
MPN329	fur	TF161	OE			3	8.50	0.1449	9.21	0.1096

Continued on next page

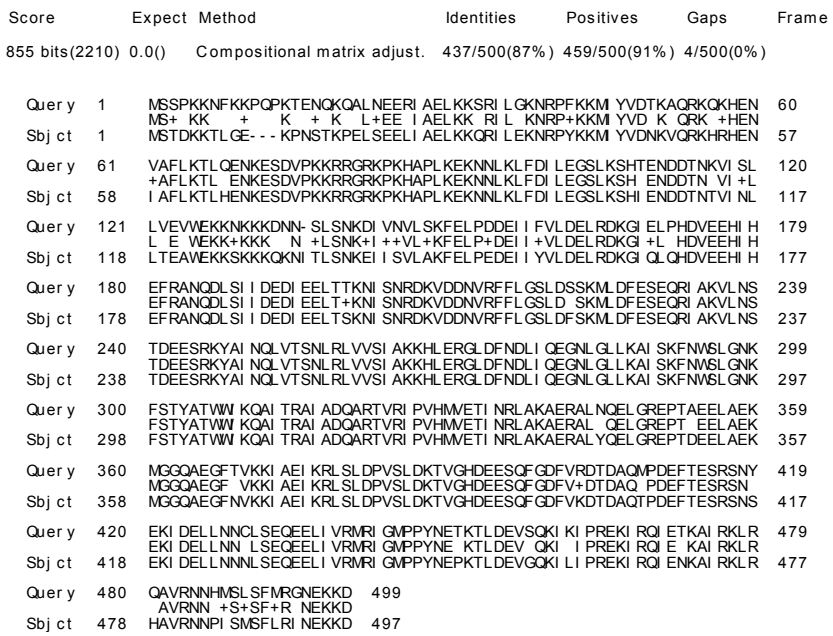
Table A.3 – *Continued from previous page*

				Screening		Validation				
							DT		Corrected DT	
ID	Gene Name	Cell line	Type	G	M	Batch	Mean	SD	Mean	SD
MPN349	ymdB	TF24	OE			3	10.92	0.2144	10.45	0.0943
MPN352	sigA	TF14	OE			2	8.40	0.2238	9.31	0.1631
MPN397	spoT	TF220	OE	Fast	Fast	2	10.46	0.1815	10.03	0.1719
		TF220	OE			3	9.39	0.4000	9.50	0.1716
		Tn397	DR			2	9.10	0.1789	9.06	0.1435
		TF505	Mut		Slow					
MPN420	glpQ	TF180	OE	Fast						
		Tn420	DR	Fast	Slow					
MPN421	cglT1	Tn421	DR	Slow	Slow					
MPN428	pta	TF174	OE	Fast	Fast	1	9.94	0.3385	10.32	0.2945
MPN430	gap	TF414	OE			2	9.09	0.2427	9.04	0.1514
MPN473	lip2	TF196	OE	Slow	Slow	3	9.08	0.1968	9.30	0.1694
MPN475	engA	TF229	OE		Slow	3	7.39	0.2046	8.70	0.1554
MPN478	YebC	TF419	OE			2	9.30	0.2775	9.55	0.1289
		TF419	OE			3	8.72	0.1805	9.21	0.0914
MPN490	recA	TF12	OE		Slow					
		TF244	OE			1	9.96	0.0770	10.12	0.0511
		Tn490	DR	Fast	Fast					
MPN507		TF221	OE		Slow					
MPN516	rpoB	TAP516	OE			1	9.36	0.1413	9.65	0.1302
		TAP516R	OE			4	7.97	0.1174	8.47	0.0662
MPN545	rnc	TF296	OE			3	8.29	0.1012	9.12	0.0748
		Tn545	DR		Fast					
MPN563	obg	TF26	OE			3	8.19	0.1107	9.00	0.0764
MPN566	glpQ-2	Tn566	DR	Fast		1	8.94	0.1255	9.26	0.0878
MPN568	spg	TF245	OE			2	9.58	0.2920	9.37	0.1504
MPN569	ybeY	TF235	OE			2	10.49	0.1961	9.62	0.0909
MPN574	groES	TAP574R	OE			4	7.97	0.0647	8.41	0.0487
MPN606	eno	TF183	OE			2	9.03	0.0615	8.69	0.0240
MPN608	phoU	TF23	OE		Slow					
MPN615	hsdS1	TF222	OE		Slow					
MPN626		TF405	OE			2	9.92	0.2002	9.34	0.1222
MPN627	ptsI	TF184	OE		Fast	1	9.37	0.0934	9.22	0.0478
MPN629	tim	TF465	OE			4	9.86	0.2660	9.51	0.1593
MPN651	mtlA	TF2	OE		Fast					

Continued on next page

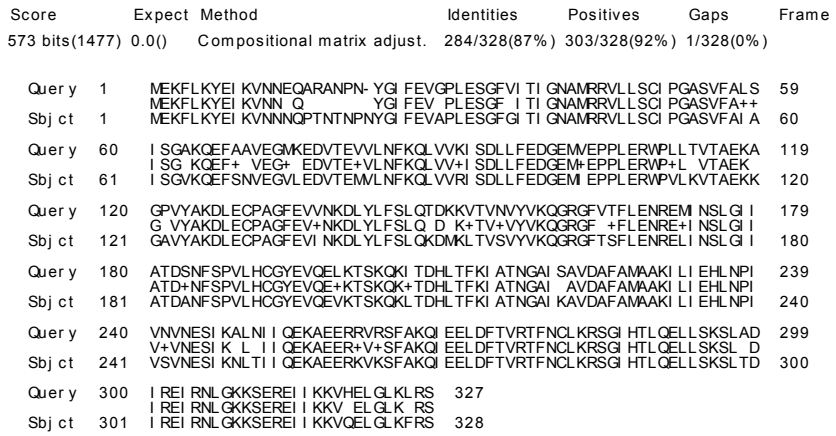
Table A.3 – *Continued from previous page*

ID	Gene Name	Cell line	Type	Screening G	Screening M	Validation Batch	Validation DT Mean	Validation DT SD	Validation Corrected DT Mean	Validation Corrected DT SD
MPN667	gtaB	TF213	OE		Slow					
MPN674	ldh	TF187	OE		Slow					
MPN686	dnaA	TF165	OE		Slow					



```
 Score          Expect  Method                      Identities      Positives      Gaps      Frame

 855 bits(2210) 0.0()   Compositional matrix adjust.  437/500(87%)  459/500(91%)  4/500(0%)


  Query  1    MSSPKKNFKKPQPKTENQKQALNEERIAELKKSRILGKNRPFKKMIYVDTKAQRKQKHEN  60
              MS+ KK   +   K  + K   L+EE IAELKK RIL KNRP+KKMIYVD K QRK +HEN
  Sbjct  1    MSTDKKTLGE---KPNSTKPELSEELIAELKKQRILEKNRPYKKMIYVDNKVQRKHRHEN  57

  Query  61   VAFLKTLQENKESDVPKKRRGRKPKHAPLKEKNNLKLFDILEGSLKSHTENDDTNKVISL  120
              +AFLKTL ENKESDVPKKRRGRKPKHAPLKEKNNLKLFDILEGSLKSH ENDDTN VI+L
  Sbjct  58   IAFLKTLHENKESDVPKKRRGRKPKHAPLKEKNNLKLFDILEGSLKSHIENDDTNTVINL  117

  Query  121  LVEVWEKKNKKKDNN-SLSNKDIVNVLSKFELPDDEIIFVLDELRDKGIELPHDVEEHIH  179
              L E WEKK+KKK  N +LSNK+I ++VL+KFELP+DEII+VLDELRDKGI+L HDVEEHIH
  Sbjct  118  LTEAWEKKSKKKQKNITLSNKEIISVLAKFELPEDEIIYVLDELRDKGIQLQHDVEEHIH  177

  Query  180  EFRANQDLSIIDEDIEELTTKNISNRDKVDDNVRFFLGSLDSSKMLDFESEQRIAKVLNS  239
              EFRANQDLSII DEDI EELT+KNI SNRDKVDDNVRFFLGSLD SKMLDFESEQRI KVLNS
  Sbjct  178  EFRANQDLSIIDEDIEELTSKNIISNRDKVDDNVRFFLGSLDFSKMLDFESEQRIVKVLNS  237

  Query  240  TDEESRKYAINQLVTSNLRLVVSIAKKHLERGLDFNDLIQEGNLGLLKAISKFNWSLGNK  299
              TDEESRKYAI NQLVTSNLRLVVSI AKKHLERGLDFNDLI QEGNLGLLKAI SKFNWSLGNK
  Sbjct  238  TDEESRKYAINQLVTSNLRLVVSIAKKHLERGLDFNDLIQEGNLGLLKAISKFNWSLGNK  297

  Query  300  FSTYATWWIKQAITRAIADQARTVRIPVHMVETINRLAKAERALNQELGREPTAEELAEK  359
              FSTYATWWI KQAI TRAI ADQARTVRI PVHMVETI NRLAKAERAL QELGREPT EELAEK
  Sbjct  298  FSTYATWWIKQAIIRAIADQARTVRIPVHMVETIKNRLAKAERALYQELGREPTDEELAEK  357

  Query  360  MGGQAEGFTVKKIAEIKRLSLDPVSLDKTVGHDEESQFGDFVRDTDAQMPDEFTESRSNY  419
              MGGQAEGF VKKI AEI KRLSLDPVSLDKTVGHDEESQFGDFV+DTDAQ PDEFTESRSN
  Sbjct  358  MGGQAEGFNVKKIAEIKRLSLDPVSLDKTVGHDEESQFGDFVKDTDAQTPDEFTESRSNS  417

  Query  420  EKIDELLNNCLSEQEELIVRMRIGMPPYNETKTLDEVSQKIKIPREKIRQIETKAIRKLR  479
              EKI DELLNN LSEQEELI VRMRI GMPPYNE KTLDEV QKI  I PREKI RQI E KAI RKLR
  Sbjct  418  EKIDELLNNNLSEQEELIVRMRIGMPPYNEPKTLDEVGQKILIPREKIRQIENKAIRKLR  477

  Query  480  QAVRNNHMSLSFMRGNEKKD  499
               AVRNN +S+SF+R NEKKD
  Sbjct  478  HAVRNNPISMSFLRINEKKD  497
```

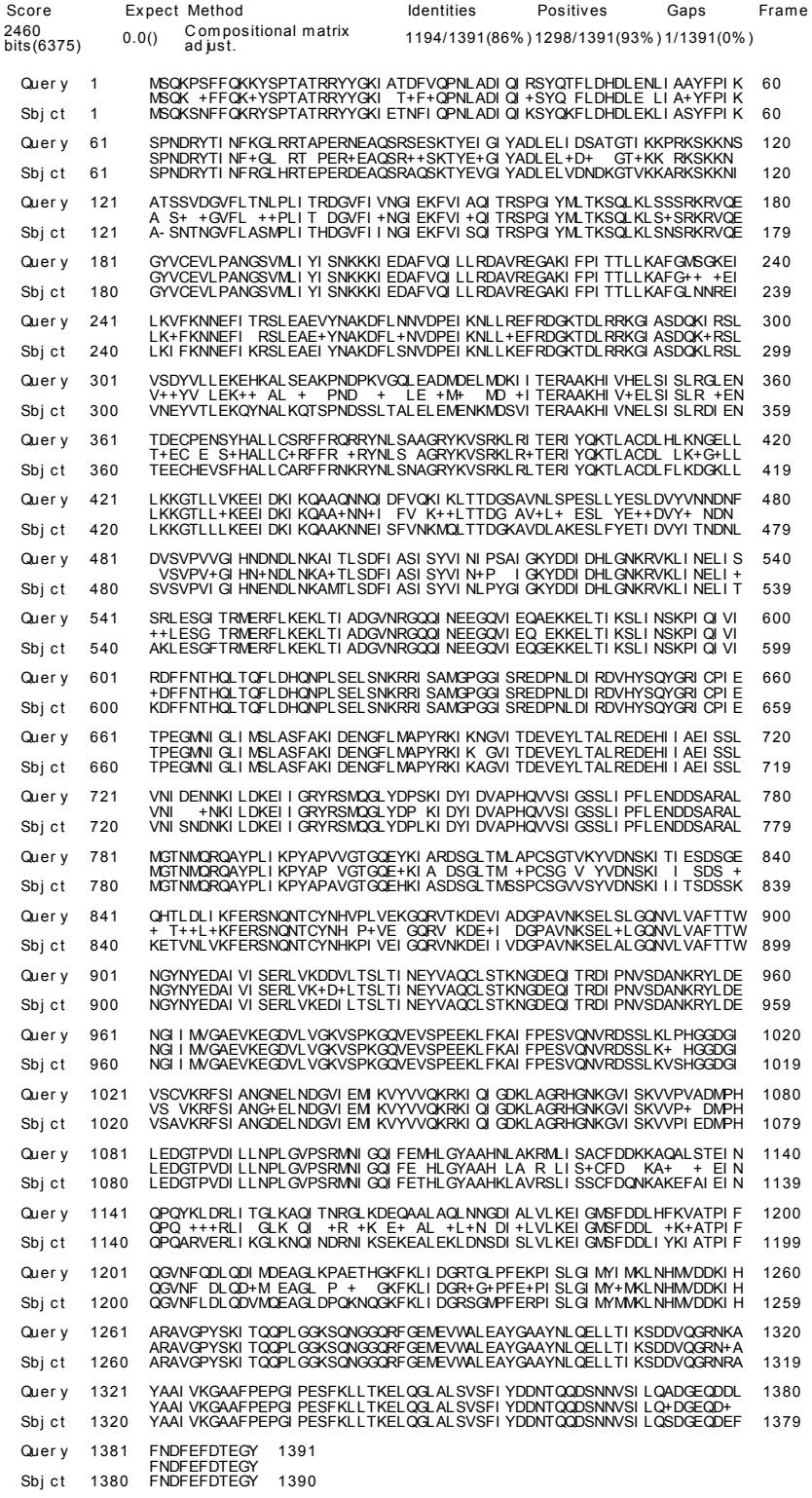
Figure A.1: Alignment of *rpoD*. Query sequence from *M. pneumoniae* M129 and subject sequence from *M. genitalium* G37. Alignment using pBLAST.



Score	Expect	Method	Identities	Positives	Gaps	Frame
573 bits(1477)	0.0()	Compositional matrix adjust.	284/328(87%)	303/328(92%)	1/328(0%)	

```
Query  1    MEKFLKYEIKVNNEQARANPN-YGIFEVGPLESGFVITIGNAMRRVLLSCIPGASVFALS  59
            MEKFLKYEIKVNN Q        YGIFEV PLESGF ITIGNAMRRVLLSCIPGASVFA++
Sbjct  1    MEKFLKYEIKVNNNQPTNTNPNYGIFEVAPLESGFGITIGNAMRRVLLSCIPGASVFAIA  60

Query  60   ISGAKQEFAAVEGMKEDVTEVVLNFKQLVVKISDLLFEDGEMVEPPLERWPLLTVTAEKA  119
            ISG KQEF+ VEG+ EDVTE+VLNFKQLVV+ISDLLFEDGEM+EPPLERWP+L VTAEK
Sbjct  61   ISGVKQEFSNVEGVLEDVTEMVLNFKQLVVRISDLLFEDGEMIEPPLERWPVLKVTAEKK  120

Query  120  GPVYAKDLECPAGFEVVNKDLYLFSLQTDKKVTVNVYVKQGRGFVTFLENREMINSLGII  179
            G VYAKDLECPAGFEV+NKDLYLFSLQ D K+TV+VYVKQGRGF +FLENRE+INSLGII
Sbjct  121  GAVYAKDLECPAGFEVINKDLYLFSLQKDMKLTVSVYVKQGRGFTSFLENRELINSLGII  180

Query  180  ATDSNFSPVLHCGYEVQELKTSKQKITDHLTFKIATNGAISAVDAFAMAAKILIEHLNPI  239
            ATD+NFSPVLHCGYEVQE+KTSKQK+TDHLTFKIATNGAI AVDAFAMAAKILIEHLNPI
Sbjct  181  ATDANFSPVLHCGYEVQEVKTSKQKLTDHLTFKIATNGAIKAVDAFAMAAKILIEHLNPI  240

Query  240  VNVNESIKALNIIQEKAEERRVRSFAKQIEELDFTVRTFNCLKRSGIHTLQELLSKSLAD  299
            V+VNESIK L IIQEKAEER+V+SFAKQIEELDFTVRTFNCLKRSGIHTLQELLSKSL D
Sbjct  241  VSVNESIKNLTIIQEKAEERKVKSFAKQIEELDFTVRTFNCLKRSGIHTLQELLSKSLTD  300

Query  300  IREIRNLGKKSEREIIKKVHELGLKLRS  327
            IREIRNLGKKSEREIIKKV ELGLK RS
Sbjct  301  IREIRNLGKKSEREIIKKVQELGLKFRS  328
```

Figure A.2: Alignment of *rpoA*. Query sequence from *M. pneumoniae* M129 and subject sequence from *M. genitalium* G37. Alignment using pBLAST.



Score	Expect	Method	Identities	Positives	Gaps	Frame
2460 bits(6375)	0.0()	Compositional matrix adjust.	1194/1391(86%)	1298/1391(93%)	1/1391(0%)	

```
Query  1     MSQKPSFFQKKYSPTATRRYYGKIATDFVQPNLADIQIRSYQTFLDHDLENLIAAYFPIK  60
             MSQK +FFQK+YSPTATRRYYGKI  T+F+QPNLADIQI+SYQ FLDHDLE LIA+YFPIK
Sbjct  1     MSQKSNFFQKRYSPTATRRYYGKIETNFIQPNLADIQIKSYQKFLDHDLEKLIASYFPIK  60

Query  61    SPNDRYTINFKGLRRTAPERNEAQSRSESKTYEIGIYADLELIDSATGTIKKPRKSKKNS  120
             SPNDRYTI NF+GL RT PER+EAQSR++SKTYE+GI YADLEL+D+  GT+KK RKSKKN
Sbjct  61    SPNDRYTINFRGLHRTEPERDEAQSRAQSKTYEVGIYADLELVDNDKGTVKKARKSKKNI  120

Query  121   ATSSVDGVFLTNLPLITRDGVFIVNGIEKFVIAQITRSPGIYMLTKSQLKLSSSRKRVQE  180
             A S+ +GVFL ++PLI T DGVFI +NGI EKFVI +QITRSPGIYMLTKSQLKLS+SRKRVQE
Sbjct  121   A-SNTNGVFLASMPLITHDGVFIINGIEKFVISQITRSPGIYMLTKSQLKLSNSRKRVQE  179

Query  181   GYVCEVLPANGSVMLIYISNKKKIEDAFVQILLRDAVREGAKIFPITTLLKAFGMMSGKEI  240
             GYVCEVLPANGSVMLIYISNKKKIEDAFVQILLRDAVREGAKIFPITTLLKAFG++ +EI
Sbjct  180   GYVCEVLPANGSVMLIYISNKKKIEDAFVQILLRDAVREGAKIFPITTLLKAFGLNNREI  239

Query  241   LKVFKNNEFITRSLEAEVYNAKDFLNNVDPEIKNLLREFRDGKTDLRRKGIASDQKIRSL  300
             LK+FKNNEFI RSLEAE+YNAKDFL+NVDPEIKNLL+EFRDGKTDLRRKGIASDQK+RSL
Sbjct  240   LKIFKNNEFIKRSLEAEIYNAKDFLSNVDPEIKNLLKEFRDGKTDLRRKGIASDQKLRSL  299

Query  301   VSDYVLLEKEHKALSEAKPNDPKVGQLEADMDELMDKIITERAAKHIVHELSISLRGLEN  360
             V++YV LEK++ AL + PND + LE +M+ MD +I TERAAKHIV+ELSISLR +EN
Sbjct  300   VNEYVTLEKQYNALKQTSPNDSSLTALELEMENKMDSVITERAAKHIVNELSISLRDIEN  359

Query  361   TDECPENSYHALLCSRFFRQRRYNLSAAGRYKVSRKLRITERIYQKTLACDLHLKNGELL  420
             T+EC E S+HALLC+RFFR +RYNLS AGRYKVSRKLR+TERIYQKTLACDL LK+G+LL
Sbjct  360   TEECHEVSFHALLCARFFRNKRYNLSNAGRYKVSRKLRLTERIYQKTLACDLFLKDGKLL  419

Query  421   LKKGTLLVKEEIDKIKQAAQNNQIDFVQKILKTTDGSAVNLSPESLLYESLDVYVNNDNF  480
             LKKGTLL+KEEIDKIKQAA+NN+I  FV K++LTTDG AV+L+ ESL YE++DVY+ NDN
Sbjct  420   LKKGTLLLKEEIDKIKQAAKNNEISFVNKMQLTTDGKAVDLAKESLFYETIDVYITNDNL  479

Query  481   DVSVPVVGIHNDNDLNKAITLSDFIASISYVINIPSAIGKYDDIDHLGNKRVKLINELIS  540
               VSVPV+GIHN+NDLNKA+TLSDFIASISYVIN+P  IGKYDDIDHLGNKRVKLINELI+
Sbjct  480   SVSVPVIGIHNENDLNKAMTLSDFIASISYVINLPYGIGKYDDIDHLGNKRVKLINELIT  539

Query  541   SRLESGITRMERFLKEKLTIADGVNRGQQINEEGQVIEQAEKKELTIKSLINSKPIQIVI  600
             ++LESG TRMERFLKEKLTIADGVNRGQQINEEGQVIEQ EKKELTIKSLINSKPIQIVI
Sbjct  540   AKLESGFTRMERFLKEKLTIADGVNRGQQINEEGQVIEQGEKKELTIKSLINSKPIQIVI  599

Query  601   RDFFNTHQLTQFLDHQNPLSELSNKRRISAMGPGGISREDPNLDIRDVHYSQYGRICPIE  660
             +DFFNTHQLTQFLDHQNPLSELSNKRRISAMGPGGISREDPNLDIRDVHYSQYGRICPIE
Sbjct  600   KDFFNTHQLTQFLDHQNPLSELSNKRRISAMGPGGISREDPNLDIRDVHYSQYGRICPIE  659

Query  661   TPEGMNIGLIMSLASFAKIDENGFLMAPYRKIKNGVITDEVEYLTALREDEHIIAEISSL  720
             TPEGMNIGLIMSLASFAKIDENGFLMAPYRKIK GVITDEVEYLTALREDEHIIAEISSL
Sbjct  660   TPEGMNIGLIMSLASFAKIDENGFLMAPYRKIKAGVITDEVEYLTALREDEHIIAEISSL  719

Query  721   VNIDENNKILDKEIIGRYRSMQGLYDPSKIDYIDVAPHQVVSIGSSLIPFLENDDSARAL  780
             VNI  +NKILDKEIIGRYRSMQGLYDP KIDYIDVAPHQVVSIGSSLIPFLENDDSARAL
Sbjct  720   VNISNDNKILDKEIIGRYRSMQGLYDPLKIDYIDVAPHQVVSIGSSLIPFLENDDSARAL  779

Query  781   MGTNMQRQAYPLIKPYAPVVGTGQEYKIARDSGLTMLAPCSGTVKYVDNSKITIESDSGE  840
             MGTNMQRQAYPLIKPYAP VGTGQE+KIA DSGLTM +PCSG V YVDNSKI I SDS +
Sbjct  780   MGTNMQRQAYPLIKPYAPAVGTGQEHKIASDSGLTMSSPCSGVVSYVDNSKIITSDSSK  839

Query  841   QHTLDLIKFERSNQNTCYNHVPLVEKGQRVTKDEVIADGPAVNKSELSLGQNVLVAFTTW  900
             + T++L+KFERSNQNTCYNH P+VE GQRV KDE+I  DGPAVNKSEL+LGQNVLVAFTTW
Sbjct  840   KETVNLVKFERSNQNTCYNHKPIVEIGQRVNKDEIIVDGPAVNKSELALGQNVLVAFTTW  899

Query  901   NGYNYEDAIVISERLVKDDVLTSLTINEYVAQCLSTKNGDEQITRDIPNVSDANKRYLDE  960
             NGYNYEDAIVISERLVK+D+LTSLTINEYVAQCLSTKNGDEQITRDIPNVSDANKRYLDE
Sbjct  900   NGYNYEDAIVISERLVKEDILTSLTINEYVAQCLSTKNGDEQITRDIPNVSDANKRYLDE  959

Query  961   NGIIMVGAEVKEGDVLVGKVSPKGQVEVSPEEKLFKAIFPESVQNVRDSSLKLPHGGDGI  1020
             NGIIMVGAEVKEGDVLVGKVSPKGQVEVSPEEKLFKAIFPESVQNVRDSSLK+ HGGDGI
Sbjct  960   NGIIMVGAEVKEGDVLVGKVSPKGQVEVSPEEKLFKAIFPESVQNVRDSSLKVSHGGDGI  1019

Query  1021  VSCVKRFSIANGNELNDGVIEMIKVYVVQKRKIQIGDKLAGRHGNKGVISKVVPVADMPH  1080
             VS VKRFSIANG+ELNDGVIEMIKVYVVQKRKIQIGDKLAGRHGNKGVISKVVP+ DMPH
Sbjct  1020  VSAVKRFSIANGDELNDGVIEMIKVYVVQKRKIQIGDKLAGRHGNKGVISKVVPIEDMPH  1079

Query  1081  LEDGTPVDILLNPLGVPSRMNIGQIFEMHLGYAAHNLAKRMLISACFDDKKAQALSTEIN  1140
             LEDGTPVDILLNPLGVPSRMNIGQIFE HLGYAAH LA R LIS+CFD  KA+  + EIN
Sbjct  1080  LEDGTPVDILLNPLGVPSRMNIGQIFETHLGYAAHKLAVRSLISSCFDQNKAKEFAIEIN  1139

Query  1141  QPQYKLDRLITGLKAQITNRGLKDEQAALAQLNNGDIALVLKEIGMSFDDLHFKVATPIF  1200
             QPQ +++RLI GLK QI +R +K E+ AL +L+N DI+LVLKEIGMSFDDL +K+ATPIF
Sbjct  1140  QPQARVERLIKGLKNQINDRNIKSEKEALEKLDNSDISLVLKEIGMSFDDLIYKIATPIF  1199

Query  1201  QGVNFQDLQDIMDEAGLKPAETHGKFKLIDGRTGLPFEKPISLGIMYIMKLNHMVDDKIH  1260
             QGVNF DLQD+M EAGL P +  GKFKLIDGR+G+PFE+PISLGIMY+MKLNHMVDDKIH
Sbjct  1200  QGVNFLDLQDVMQEAGLDPQKNQGKFKLIDGRSGMPFERPISLGIMYMMKLNHMVDDKIH  1259

Query  1261  ARAVGPYSKITQQPLGGKSQNGGQRFGEMEVWALEAYGAAYNLQELLTIKSDDVQGRNKA  1320
             ARAVGPYSKITQQPLGGKSQNGGQRFGEMEVWALEAYGAAYNLQELLTIKSDDVQGRN+A
Sbjct  1260  ARAVGPYSKITQQPLGGKSQNGGQRFGEMEVWALEAYGAAYNLQELLTIKSDDVQGRNRA  1319

Query  1321  YAAIVKGAAFPEPGIPESFKLLTKELQGLALSVSFIYDDNTQQDSNNVSILQADGEQDDL  1380
             YAAIVKGAAFPEPGIPESFKLLTKELQGLALSVSFIYDDNTQQDSNNVSILQ+DGEQD+
Sbjct  1320  YAAIVKGAAFPEPGIPESFKLLTKELQGLALSVSFIYDDNTQQDSNNVSILQSDGEQDEF  1379

Query  1381  FNDFEFDTEGY  1391
             FNDFEFDTEGY
Sbjct  1380  FNDFEFDTEGY  1390
```

Figure A.3: Alignment of *rpoB*. Query sequence from *M. pneumoniae* M129 and subject sequence from *M. genitalium* G37. Alignment using NCBI pBLAST.

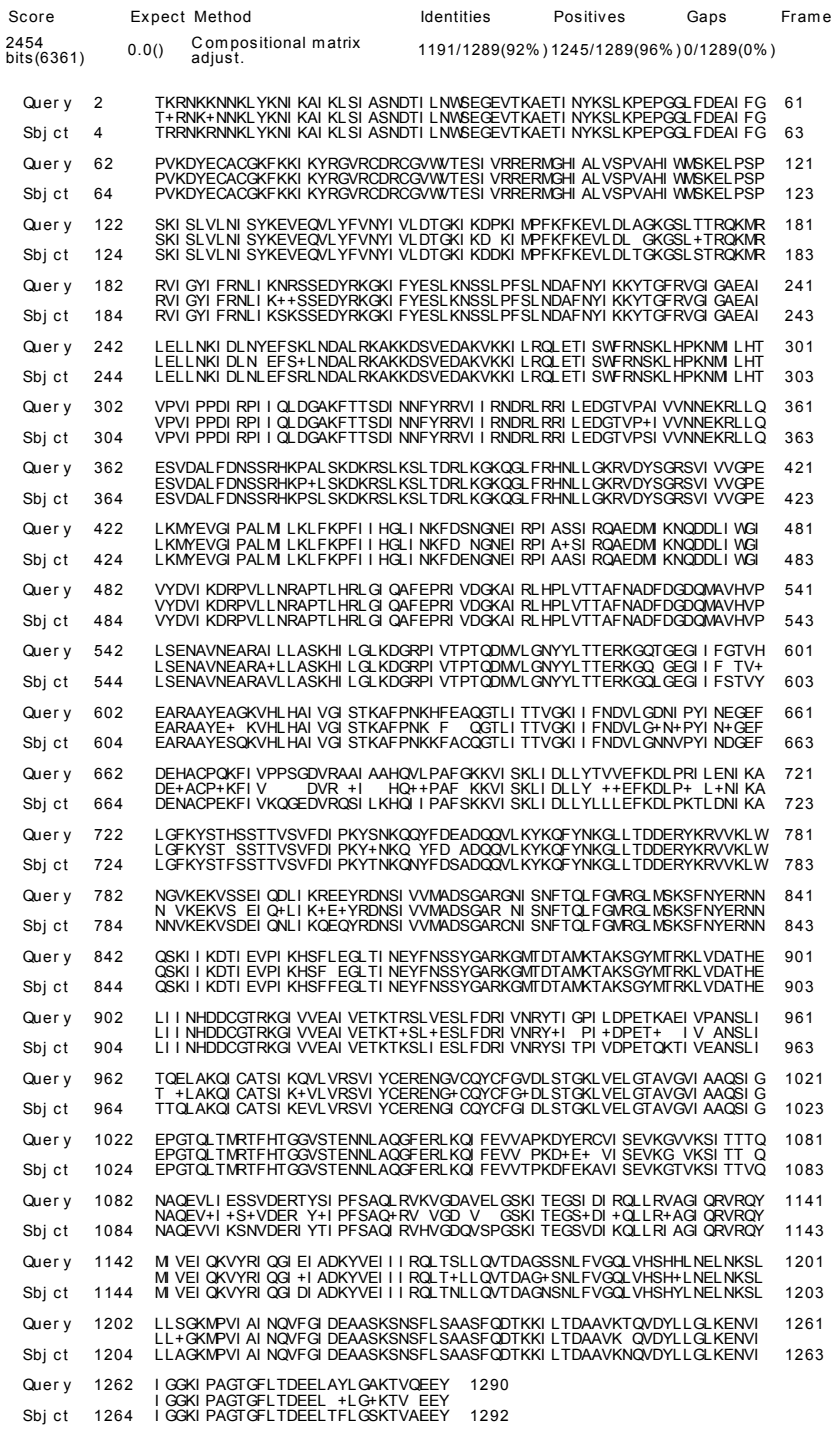


Figure A.4: Alignment of *rpoC*. Query sequence from *M. pneumoniae* M129 and subject sequence from *M. genitalium* G37. Alignment using pBLAST.

Bibliography

Adams, D. W. and Errington, J. (2009). Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat Rev Micro*, 7(9):642–653.

Adams, M. A., Udell, C. M., Pal, G. P., and Jia, Z. (2005). MraZ from Escherichia coli: Cloning, purification, crystallization and preliminary X-ray analysis. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 61(4):378–380.

Alarcón, F., de Vasconcelos, A. T. R., Yim, L., and Zaha, A. (2007). Genes involved in cell division in mycoplasmas. *Genetics and Molecular Biology*, 30(SUPPL. 1):174–181.

Amikam, D., Glaser, G., and Razin, S. (1984). Mycoplasmas (Mollicutes) have a low number of rRNA genes. *Journal of Bacteriology*, 158(1):376–378.

Asai, T., Condon, C., Voulgaris, J., Shen, B., Al-omar, M., Squires, C., and Squires, C. L. (1999a). Construction and Initial Characterization of Escherichia coli Strains with Few or No Intact Chromosomal rRNA Operons Construction and Initial Characterization of Escherichia coli Strains with Few or No Intact Chromosomal rRNA Operons. *Journal of Bacteriology*, 181(12):3803.

Asai, T., Zaporojets, D., Squires, C., and Squires, C. L. (1999b). An Escherichia coli strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 96(5):1971–1976.

Aseev, L. V. and Boni, I. V. (2011). Extraribosomal functions of bacterial ribosomal proteins. *Molecular Biology*, 45(5):739–750.

Atkinson, T. P., Balish, M. F., and Waites, K. B. (2008). Epidemiology, clinical manifestations, pathogenesis and laboratory detection of Mycoplasma pneumoniae infections. *FEMS Microbiology Reviews*, 32(6):956–973.

Badrinarayanan, A., Le, T. B. K., and Laub, M. T. (2015). Bacterial Chromosome Organization and Segregation. *Annual review of cell and developmental biology*, 31:171–99.

Bailey, M. W., Bisicchia, P., Warren, B. T., Sherratt, D. J., and M??nnik, J. (2014). Evidence for Divisome Localization Mechanisms Independent of the Min System and SlmA in Escherichia coli. *PLoS Genetics*, 10(8).

Balish, M. F. (2006). Subcellular structures of mycoplasmas. *Frontiers in Bioscience*, 11:2017–2027.

Balish, M. F. (2014). Mycoplasma pneumoniae, an underutilized model for bacterial cell biology. *Journal of Bacteriology*, 196(21):3675–3682.

Balish, M. F. and Krause, D. C. (2006). Mycoplasmas: A distinct cytoskeleton for wall-less bacteria. *Journal of Molecular Microbiology and Biotechnology*, 11(3-5):244–255.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5):455–477.

Barré, A., de Daruvar, A., and Blanchard, A. (2004). MolliGen, a database dedicated to the comparative genomics of Mollicutes. *Nucleic acids research*, 32(Database issue):D307–10.

Beaman, K. D. and Pollack, J. D. (1981). Adenylate energy charge in Acholeplasma laidlawii. *Journal of Bacteriology*, 146(3):1055–1058.

Beaman, K. D. and Pollack, J. D. (1983). Synthesis of adenylate nucleotides by Mollicutes (mycoplasmas). *Journal of general microbiology*, 129(10):3103–3110.

Berghuis, B. a., Dulin, D., Xu, Z.-Q., van Laar, T., Cross, B., Janissen, R., Jergic, S., Dixon, N. E., Depken, M., and Dekker, N. H. (2015). Strand separation establishes a sustained lock at the Tus-Ter replication fork barrier. *Nature chemical biology*, 11(8):579–85.

Bernander, R. and Ettema, T. J. G. (2010). FtsZ-less cell division in archaea and bacteria. *Current Opinion in Microbiology*, 13(6):747–752.

Biberfield, G. and Biberfield, P. (1970). Ultrastructural features of Mycoplasma pneumoniae. *J Bacteriol*, 102(3):855–861.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science (New York, NY)*, 277(5331):1453–1474.

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, 27(4):578–579.

Borràs, E., Espadas, G., Mancuso, F. M., Maier, T., Chiva, C., and Sabidó, E. (2013). Integrative quantitation enables a comprehensive proteome comparison of two Mycoplasma pneumoniae genetic perturbations. *Molecular bioSystems*, 9(6):1249–1256.

Bové, J. M. (1993). Molecular features of mollicutes. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, 17 Suppl 1(Suppl 1):S10–S31.

Bredt, W. (1968). Growth Morphology of Mycoplasma pneumoniae Strain FH on Glass Surface. *Proceedings of the Society for Experimental Biology and Medicine*, 128(2):338–40.

Bredt, W. (1976). Estimation of Mycoplasma pneumoniae inoculum size by rate of tetrazolium reduction. *Journal of Clinical Microbiology*, 4(1):92–94.

Bremer, H. and Dennis, P. P. (1996). Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate. In Neidhardt, F. C. and Curtiss, R., editors, *Escherichia coli and Salmonella: cellular and molecular biology*, volume 2, pages 1527–1542. Second edition.

Breton, M., Sagné, E., Duret, S., Béven, L., Citti, C., and Renaudin, J. (2010). First report of a tetracycline-inducible gene expression system for mollicutes. *Microbiology (Reading, England)*, 156(Pt 1):198–205.

Breton, M., Tardy, F., Dordet-Frisoni, E., Sagne, E., Mick, V., Renaudin, J., Sirand-Pugnet, P., Citti, C., and Blanchard, A. (2012). Distribution and diversity of mycoplasma plasmids: lessons from cryptic genetic elements. *BMC microbiology*, 12(1):257.

Briggs, G. S., Smits, W. K., and Soultanas, P. (2012). Chromosomal replication initiation machinery of low-G+C-content firmicutes. *Journal of Bacteriology*, 194(19):5162–5170.

Brown, D. R., Whitcomb, R. F., and Bradbury, J. M. (2007). Revised minimal standards for description of new species of the class Mollicutes (division Tenericutes). *International Journal of Systematic and Evolutionary Microbiology*, 57(11):2703–2719.

Burgos, R. and Totten, P. A. (2014). Characterization of the operon encoding the holliday junction helicase RuvAB from Mycoplasma genitalium and its role in mgpB and mgpC gene variation. *Journal of Bacteriology*, 196(8):1608–1618.

Burgos, R., Wood, G. E., Young, L., Glass, J. I., and Totten, P. A. (2012). RecA mediates MgpB and MgpC phase and antigenic variation in Mycoplasma genitalium, but plays a minor role in DNA repair. *Molecular Microbiology*, 85(4):669–683.

Calus, D., Maes, D., Vranckx, K., Villareal, I., Pasmans, F., and Haesebrouck, F. (2010). Validation of ATP luminometry for rapid and accurate titration of Mycoplasma hyopneumoniae in Friis medium and a comparison with the color changing units assay. *Journal of Microbiological Methods*, 83(3):335–340.

Cambridge, J., Blinkova, A., Magnan, D., Bates, D., and Walker, J. R. (2014). A Replication-inhibited unsegregated nucleoid at mid-cell blocks Z-ring formation and cell division independently of SOS and the SlmA nucleoid occlusion protein in Escherichia coli. *Journal of Bacteriology*, 196(1):36–49.

Campos, M., Surovtsev, I. V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S. E., and Jacobs-Wagner, C. (2014). A constant size extension drives bacterial cell size homeostasis. *Cell*, 159(6):1433–1446.

Carrión, M., Gómez, M. J., Merchante-Schubert, R., Dongarrá, S., and Ayala, J. A. (1999). mraW, an essential gene at the dcw cluster of Escherichia coli codes for a cytoplasmic protein with methyltransferase activity. *Biochimie*, 81(8-9):879–888.

Carvalho, F. M., Fonseca, M. M., De Medeiros, S. B., Scortecci, K. C., Galindo Blaha, C. A., and Agnez-Lima, L. F. (2005). DNA repair in reduced genome: The Mycoplasma model. *Gene*, 360(2):111–119.

Chanock, R. M., Hayflick, L., and Barile, M. F. (1962). Growth on artificial medium of an agent associated with atypical pneumonia and its identification as a PPLO. *Proceedings of the National Academy of Sciences of the United States of America*, 48:41–9.

Chen, W. H., Van Noort, V., Lluch-Senar, M., Hennrich, M. L., Wodke, J. A., Yus, E., Alibés, A., Roma, G., Mende, D. R., Pesavento, C., Typas, A., Gavin, A. C., Serrano, L., and Bork, P. (2015). Integration of multi-omics data of a genome-reduced bacterium: Prevalence of post-transcriptional regulation and its correlation with protein abundances. *Nucleic Acids Research*, 44(3):1192–1202.

Cloward, J. M. and Krause, D. C. (2009). Mycoplasma pneumoniae J-domain protein required for terminal organelle function. *Molecular Microbiology*, 71(5):1296–1307.

Cooper, S. (1993). The origins and meaning of the Schaechter-Maaloe-Kjeldgaard experiments. *Journal of General Microbiology*, 139:1117–1124.

Cooper, S. and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of Escherichia coli B/r. *Plasmid*, 31:519–540.

Cordova, C. M. M., Lartigue, C., Sirand-pugnet, P., Cunha, R. A. F., and Blanchard, A. (2002). Identification of the Origin of Replication of the Mycoplasma pulmonis Chromosome and Its Use in oriC Replicative Plasmids. *Journal of Bacteriology*, 184(19):5426–5435.

Couturier, E. and Rocha, E. P. C. (2006). Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Molecular microbiology*, 59(5):1506–1518.

Dandekar, T., Huynen, M., Regula, J. T., Ueberle, B., Zimmermann, C. U., Andrade, M. a., Doerks, T., Sánchez-Pulido, L., Snel, B., Suyama, M., Yuan, Y. P., Herrmann, R., and Bork, P. (2000). Re-annotating the Mycoplasma pneumoniae genome sequence: adding value, function and reading frames. *Nucleic acids research*, 28(17):3278–3288.

de Boer, P., Crossley, R., and Rothfield, L. (1992). The essential bacterial cell-division protein FtsZ is a GTPase. *Nature*, 359(6392):254–256.

Dennis, P. P. and Bremer, H. (2008). Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*, 3(1).

Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Jannière, L., and Ehrlich, S. (2001). Two Essential DNA Polymerases at the Bacterial Replication Fork. *Science*, 294(5547):1716–1719.

Desantis, D., Tryon, V. V., and Pollack, J. D. (1989). Metabolism of Mollicutes: the Embden-Meyerhof-Parnas Pathway and the Hexose Monophosphate Shunt. *Journal of General Microbiology*, 135(1989):683–691.

Domermuth, C. H., Nielsen, M. H., Freundt, E. A., and Birch-Andersen, A. (1964). Ultrastructure of Mycoplasma Species. *Journal of bacteriology*, 88(3):727–744.

Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwińska, J., and Zawilak-Pawlik, A. (2012). Helicobacter pylori oriC-the first bipartite origin of chromosome replication in Gram-negative bacteria. *Nucleic Acids Research*, 40(19):9647–9660.

Dong, H., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA Abundance and Codon Usage inEscherichia coliat Different Growth Rates. *Journal of Molecular Biology*, 260(5):649–663.

Downing, W. and Dennis, P. P. (1991). RNA polymerase activity may regulate transcription initiation and attenuation in the rplKAJLrpoBC operon in Escherichia coli. *The Journal of biological chemistry*, 266(2):1304–11.

Duret, S., Danet, J.-l., Garnier, M., and Renaudin, J. (1999). Gene Disruption through Homologous Recombination in Spiroplasma citri : an scm1 -Disrupted Motility Mutant Is Pathogenic. *Journal of Bacteriology*, 181(24):7449–7456.

Dybvig, K., Simecka, J. W., Watson, H. L., and Cassell, G. H. (1989). High-frequency variation in Mycoplasma pulmonis colony size. *Journal of Bacteriology*, 171(9):5165–5168.

Dybvig, K. and Voelker, L. L. (1996). Molecular Biology of Mycoplasmas. *Annual Review of Microbiology*, 50:25–57.

Dykxhoorn, D. M., Pierre, R. S., Linn, T., and St Pierre, R. (1996). Synthesis of the beta and beta' subunits of Escherichia coli RNA polymerase is autogenously regulated in vivo by both transcriptional and translational mechanisms. *Mol Microbiol*, 19(3):483–493.

Ecker, R. E. and Schaechter, M. (1963). Ribosome content and the rate of growth of Salmonella typhimurium. *Biochimica et Biophysica Acta (BBA) - Specialized Section on Nucleic Acids and Related Subjects*, 76:275–279.

Engbaek, F., Gross, C., and Burgess, R. R. (1976). Quantitation of RNA polymerase subunits in Escherichia coli during exponential growth and after bacteriophage T4 infection. *Molecular and General Genetics*, 143(3):291–295.

Eraso, J. M., Markillie, L. M., Mitchell, H. D., Taylor, R. C., Orr, G., and Margolin, W. (2014). The highly conserved MraZ protein is a transcriptional regulator in Escherichia coli. *Journal of Bacteriology*, 196(11):2053–2066.

Errington, J., Daniel, R. A., and Scheffers, D.-J. (2003). Cytokinesis in Bacteria. *Microbiology and Molecular Biology Reviews*, 67(1):52–65.

Fisunov, G. Y., Evsyutina, D. V., Semashko, T. A., Arzamasov, A. A., Manuvera, V. A., Letarov, A. V., and Govorun, V. M. (2016). Binding site of MraZ transcription factor in Mollicutes. *Biochimie*, 125:59–65.

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-f. J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C. P.-c., Lucier, T. S., Peterson, S. N., Smith, H., Hutchison, C. a., and Venter, J. C. (1995). The Minimal Gene Complement of Mycoplasma genitalium. *Science*, 270(5235):397–403.

Fukuda, Y., Washio, T., and Tomita, M. (1999). Comparative study of overlapping genes in the genomes of Mycoplasma genitalium and Mycoplasma pneumoniae. *Nucleic Acids Research*, 27(8):1847–1853.

Furness, G. (1968). Analysis of the Growth Cycle of Mycoplasma Orale. *The Journal of Infectious Diseases*, 118(4):436–442.

Furness, G., Pipes, F. J., and McMurtrey, M. J. (1968a). Analysis of the life cycle of Mycoplasma pneumoniae by synchronized division and by ultraviolet and x irradiations. *The Journal of Infectious Diseases*, 118(1):7–13.

Furness, G., Pipes, F. J., and McMurtrey, M. J. (1968b). Susceptibility of human mycoplasmata to ultraviolet and X irradiations. *The Journal of Infectious Diseases*, 118(1):1–6.

Gaal, T., Bartlett, M. S., Ross, W., Jr, C. L. T., and Gourse, R. L. (1997). Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. *Science*, 278:2092–2097.

Gausing, K. (1977). Regulation of ribosome production in< i> Escherichia coli</i>: Synthesis and stability of ribosomal RNA and of ribosomal protein messenger RNA at different growth .... *Journal of molecular biology*, pages 335–354.

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2010). Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science*, 329(5987):52–56.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. a., Smith, H. O., Iii, C. A. H., and America, N. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, 6(5):343–5.

Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchison, C. a., Smith, H. O., and Venter, J. C. (2006). Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2):425–430.

Glass, J. I., Hutchison, C. a., Smith, H. O., and Venter, J. C. (2009). A systems biology tour de force for a near-minimal bacterium. *Molecular systems biology*, 5(330):330.

Grosjean, H., Breton, M., Sirand-Pugnet, P., Tardy, F., Thiaucourt, F., Citti, C., Barré, A., Yoshizawa, S., Fourmy, D., de Crécy-Lagard, V., and Blanchard, A. (2014). Predicting the Minimal Translation Apparatus: Lessons from the Reductive Evolution of Mollicutes. *PLoS Genetics*, 10(5):1–21.

Güell, M., van Noort, V., Yus, E., Chen, W.-H., Leigh-Bell, J., Michalodimitrakis, K., Yamada, T., Arumugam, M., Doerks, T., Kühner, S., Rode, M., Suyama, M., Schmidt, S., Gavin, A.-C., Bork, P., and Serrano, L. (2009). Transcriptome complexity in a genome-reduced bacterium. *Science (New York, N.Y.)*, 326(5957):1268–1271.

Guo, T., Kong, J., Zhang, L., Zhang, C., and Hu, S. (2012). Fine tuning of the lactate and diacetyl production through promoter engineering in lactococcus lactis. *PLoS ONE*, 7(4).

Halbedel, S., Hames, C., and Stülke, J. (2004). In Vivo Activity of Enzymatic and Regulatory Components of the Phosphoenolpyruvate: Sugar Phosphotransferase System in Mycoplasma pneumoniae In Vivo Activity of Enzymatic and Regulatory Components of the Phosphoenolpyruvate : Sugar Phosphotransferase Sys. *Journal of bacteriology*, 186(23):7936–7943.

Halbedel, S. and Stülke, J. (2005). Dual phosphorylation of Mycoplasma pneumoniae HPr by Enzyme I and HPr kinase suggests an extended phosphoryl group susceptibility of HPr. *FEMS Microbiology Letters*, 247(2):193–198.

Halbedel, S. and Stülke, J. (2007). Tools for the genetic analysis of Mycoplasma. *International Journal of Medical Microbiology*, 297(1):37–44.

Hardy, R. D., Coalson, J. J., Peters, J., Chaparro, A., Techasaensiri, C., Cantwell, A. M., Kannan, T. R., Baseman, J. B., and Dube, P. H. (2009). Analysis of pulmonary inflammation and function in the mouse and baboon after exposure to Mycoplasma pneumoniae CARDS toxin. *PLoS ONE*, 4(10).

Hasselbring, B. M., Jordan, J. L., Krause, R. W., and Krause, D. C. (2006). Terminal organelle development in the cell wall-less bacterium Mycoplasma pneumoniae. *Proceedings of the National Academy of Sciences*, 103(44):16478–16483.

Hatchel, J. M. and Balish, M. F. (2008). Attachment organelle ultrastructure correlates with phylogeny, not gliding motility properties, in Mycoplasma pneumoniae relatives. *Microbiology*, 154(1):286–295.

Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology*, 13(5):298–309.

Hayflick, L. (1965). Tissue Cultures and Mycoplasmas. *Texas Reports on Biology and Medicine*, 23(1, Suppl):285–303.

Hedreyda, C. T., Lee, K. K., and Krause, D. C. (1993). Transformation of Mycoplasma pneumoniae with Tn4001 by electroporation.

Hegde, S., Hegde, S., Zimmermann, M., Flöck, M., Spergser, J., Rosengarten, R., and Chopra-Dewasthaly, R. (2015a). Simultaneous identification of potential pathogenicity factors of Mycoplasma agalactiae in the natural ovine host by negative selection. *Infection and Immunity*, 83(7):2751–2761.

Hegde, S., Rosengarten, R., and Chopra-Dewasthaly, R. (2015b). Disruption of the pdhB pyruvate dehyrogenase gene affects colony morphology, in vitro growth and cell invasiveness of mycoplasma agalactiae. *PLoS ONE*, 10(3):1–13.

Helmstetter, C. E. and Cooper, S. (1968). DNA synthesis during the division cycle of rapidly growing Escherichia coli B/r. *Journal of Molecular Biology*, 31(3):507–518.

Hill, T. M. (1992). Arrest of Bacterial DNA Replication. *Annual Review of Microbiology*, 46(1):603–633.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B. C., and Herrmann, R. (1996a). Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. *Nucleic acids research*, 24(22):4420–4449.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.-C. C., and Herrmann, R. (1996b). Complete Sequence Analysis of the Genome of the Bacterium Mycoplasma Pneumoniae. *Nucleic Acids Research*, 24(22):4420–49.

Himmelreich, R., Plagens, H., Hilbert, H., Reiner, B., and Herrmann, R. (1997). Comparative analysis of the genomes of the bacteria Mycoplasma pneumoniae and Mycoplasma genitalium. *Nucleic Acids Research*, 25(4):701–712.

Hiratsuka, Y., Miyata, M., and Uyeda, T. Q. P. (2005). Living microtransporter by uni-directional gliding of Mycoplasma along microtracks. *Biochemical and Biophysical Research Communications*, 331(1):318–324.

Hutchison, C. A., Chuang, R.-Y. R.-Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z.-Q. Z.-Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y. Y., Tsvetanova, B., KimS.Wise, Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., Venter, J. C., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., and Venter, J. C. (2016). Design and synthesis of a minimal bacterial genome. *Science*, 351(6280):aad6253–aad6253.

Hutchison, C. A., Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O., and Venter, J. C. (1999). Global transposon mutagenesis and a minimal Mycoplasma genome. *Science (New York, N.Y.)*, 286(5447):2165–2169.

Itaya, M. (1995). An estimation of minimal genome size required for life. *FEBS Letters*, 362(3):257–260.

Izard, J., Gomez Balderas, C. D., Ropers, D., Lacour, S., Song, X., Yang, Y., Lindner, A. B., Geiselmann, J., and de Jong, H. (2015). A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular Systems Biology*, 11(11):840–840.

Jameson, K. H. and Wilkinson, A. J. (2017). Control of initiation of DNA replication in Bacillus subtilis and Escherichia coli. *Genes*, 8(22):1–32.

Janis, C., Lartigue, C., Frey, J., Wroblewski, H., Thiaucourt, F., Blanchard, A., and Sirand-pugnet, P. (2005). Versatile use of oriC plasmids for functional genomics of Mycoplasma capricolum subsp. capricolum. *Applied and Environmental Microbiology*, 71(6):2888–2893.

Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, 15(1):182.

Jores, J., Fischer, A., Sirand-Pugnet, P., Thomann, A., Liebler-Tenorio, E. M., Schnee, C., Santana-Cruz, I., Heller, M., and Frey, J. (2013). Mycoplasma feriruminatoris sp. nov., a fast growing Mycoplasma species isolated from wild Caprinae. *Systematic and applied microbiology*, 36(8):533–538.

Junier, I., Unal, E. B., Yus, E., Lloréns-Rico, V., and Serrano, L. (2016). Insights into the Mechanisms of Basal Coordination of Transcription Using a Genome-Reduced Bacterium. *Cell Systems*, 2(6):391–401.

Kammer, G. M., Pollack, J. D., and Klainer, A. S. (1970). Scanning-beam electron microscopy of Mycoplasma pneumoniae. *Journal of Bacteriology*, 104(1):499–502.

Kannan, T. R. and Baseman, J. B. (2006). ADP-ribosylating and vacuolating cytotoxin of Mycoplasma pneumoniae represents unique virulence determinant among bacterial pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17):6724–9.

Kannan, T. R., Musatovova, O., Balasubramanian, S., Cagle, M., Jordan, J. L., Krunkosky, T. M., Davis, A., Hardy, R. D., and Baseman, J. B. (2010). Mycoplasma pneumoniae Community Acquired Respiratory Distress Syndrome toxin expression reveals growth phase and infection-dependent regulation. *Molecular Microbiology*, 76(5):1127–1141.

Kannan, T. R., Provenzano, D., Wright, J. R., and Baseman, J. B. (2005). Identification and Characterization of Human Surfactant Protein A Binding Protein of Mycoplasma pneumoniae. *Infection and immunity*, 73(5):2828–2834.

Karas, B. J., Wise, K. S., Sun, L., Craig Venter, J., Glass, J. I., Hutchison, C. A., Smith, H. O., and Suzuki, Y. (2014). Rescue of mutant fitness defects using in vitro reconstituted designer transposons in Mycoplasma mycoides. *Frontiers in Microbiology*, 5(JUL):1–9.

Karr, J. R., Sanghvi, J. C., MacKlin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401.

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nature reviews. Microbiology*, 8(3):163–170.

Kihara, K., Ishida, S., Shintani, M., and Sasaki, T. (1983). Statistical consideration of using colony diameter as a measure of mycoplasmal growth. *FEMS Microbiology Letters*, 19:7–9.

Kim, K. S., Clyde, W. A., and Denny, F. W. (1966). Physical properties of human Mycoplasma species. *Journal of Bacteriology*, 92(1):214–219.

Kimura, S. and Suzuki, T. (2009). Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the Escherichia coli 16S rRNA. *Nucleic Acids Research*, 38(4):1341–1352.

Kirchhoff, H., Maass, C., Runge, M., Franz, B., Schmidt, R., Quentmeier, H., and Peter, F. M. (1992). Tetrazolium [3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide] reduction by Mycoplasmas. *International Journal of Systematic Bacteriology*, 42(3):506–508.

Kishaba, T. (2016). Community-Acquired Pneumonia Caused by Mycoplasma pneumoniae: How Physical and Radiological Examination Contribute to Successful Diagnosis. *Frontiers in medicine*, 3(June):28.

Kjeldgaard, B. Y. N. and Schaechter, M. (1958). The Transition Between Different Physiological States During Balanced Growth of Salmonella typhimurium. *J. Gen. Microbiol.*, 19(1958):607–616.

Klumpp, S., Zhang, Z., and Hwa, T. (2009). Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7):1366–1375.

Krause, D. C. and Balish, M. F. (2001). Structure, function, and assembly of the terminal organelle of Mycoplasma pneumoniae.

Krishnakumar, R., Assad-Garcia, N., Benders, G. A., Phan, Q., Montague, M. G., and Glass, J. I. (2010). Targeted chromosomal knockouts in mycoplasma pneumoniae. *Applied and Environmental Microbiology*, 76(15):5297–5299.

Kühner, S., van Noort, V., Betts, M. J., Leo-Macias, A., Batisse, C., Rode, M., Yamada, T., Maier, T., Bader, S., Beltran-Alvarez, P., Castaño-Diez, D., Chen, W.-H., Devos, D., Güell, M., Norambuena, T., Racke, I., Rybin, V., Schmidt, A.,

Yus, E., Aebersold, R., Herrmann, R., Böttcher, B., Frangakis, A. S., Russell, R. B., Serrano, L., Bork, P., and Gavin, A.-C. (2009). Proteome organization in a genome-reduced bacterium. *Science (New York, N.Y.)*, 326(5957):1235–1240.

Lartigue, C., Blanchard, A., Renaudin, J., Thiaucourt, F., and Sirand-Pugnet, P. (2003). Host specificity of mollicutes oriC plasmids: Functional analysis of replication origin. *Nucleic Acids Research*, 31(22):6610–6618.

Lartigue, C., Duret, S., Garnier, M., and Renaudin, J. (2002). New plasmid vectors for specific gene targeting in spiroplasma citri. *Plasmid*, 48(2):149–159.

Lee, S. W., Browning, G. F., and Markham, P. F. (2008a). Development of a replicable oriC plasmid for Mycoplasma gallisepticum and Mycoplasma imitans, and gene disruption through homologous recombination in M. gallisepticum. *Microbiology*, 154(9):2571–2580.

Lee, S. W., Browning, G. F., and Markham, P. F. (2008b). Development of a replicable oriC plasmid for Mycoplasma gallisepticum and Mycoplasma imitans, and gene disruption through homologous recombination in M. gallisepticum. *Microbiology*, 154(9):2571–2580.

Leland, D. s., Lapworth, M. A., Jones, R. B., and French, M. L. V. (1982). Comparative evaluation of media for isolation of Ureaplasma urealyticum and genital Mycoplasma species. *Journal of Clinical Microbiology*, 16(4):709–714.

Levy, S. F., Ziv, N., and Siegal, M. L. (2012). Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biology*, 10(5).

Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping. *Genome research*, pages 1851–1858.

Lin, Y. C., Agbanyim, C. N. M., Miles, R. J., Nicholas, R. A. J., Kelly, D. P., and Wood, A. P. (2008). Tetrazolium reduction methods for assessment of substrate oxidation and strain differentiation among mycoplasmas, with particular reference to Mycoplasma bovigenitalium and some members of the Mycoplasma mycoides cluster. *Journal of Applied Microbiology*, 105(2):492–501.

Llorens Rico, V. (2016). *Integrative understanding of transcription in a minimal cell model Verónica Lloréns Rico*. Doctoral thesis, Univesty Pompeu Fabra.

Lloréns-Rico, V., Cano, J., Kamminga, T., Gil, R., Latorre, A., Chen, W.-H., Bork, P., Glass, J. I., Serrano, L., and Lluch-Senar, M. (2016). Bacterial antisense RNAs are mainly the product of transcriptional noise. *Science Advances*, 2(3):1–10.

Lloréns-Rico, V., Lluch-Senar, M., and Serrano, L. (2015). Distinguishing between productive and abortive promoters using a random forest classifier in Mycoplasma pneumoniae. *Nucleic Acids Research*, 43(7):3442–3453.

Lluch-Senar, M., Cozzuto, L., Cano, J., Delgado, J., Llórens-Rico, V., Pereyre, S., Bebear, C., and Serrano, L. (2015a). Comparative "-omics" in mycoplasma pneumoniae clinical isolates reveals key virulence factors. *PLoS ONE*, 10(9):1–20.

Lluch-Senar, M., Delgado, J., Chen, W.-H., Lloréns-Rico, V., O'Reilly, F. J., Wodke, J. A. H., Unal, E. B., Yus, E., Martínez, S., Nichols, R. J., Ferrar, T., Vivancos, A., Schmeisky, A., Stülke, J., van Noort, V., Gavin, A.-C., Bork, P., and Serrano, L. (2015b). Defining a minimal cell: essentiality of small ORFs and ncRNAs in a genome-reduced bacterium. *Molecular Systems Biology*, 11(1).

Lluch-Senar, M., Luong, K., Lloréns-Rico, V., Delgado, J., Fang, G., Spittle, K., Clark, T. a., Schadt, E., Turner, S. W., Korlach, J., and Serrano, L. (2013). Comprehensive methylome characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. *PLoS genetics*, 9(1):e1003191–e1003191.

Lluch-Senar, M., Mancuso, F. M., Climente-González, H., Peña-Paz, M. I., Sabido, E., and Serrano, L. (2016). Rescuing discarded spectra: Full comprehensive analysis of a minimal proteome. *Proteomics*, 16(4):554–563.

Lluch-Senar, M., Querol, E., and Piñol, J. (2010). Cell division in a minimal bacterium in the absence of ftsZ. *Molecular Microbiology*, 78(2):278–289.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D. W., Yiu, S.-M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.-W., and Wang, J. (2015). Erratum: SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*, 4(1):30.

Lutkenhaus, J. (2007). Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. *Annual review of biochemistry*, 76:539–562.

Maglennon, G. A., Cook, B. S., Matthews, D., Deeney, A. S., Bossé, J. T., Langford, P. R., Maskell, D. J., Tucker, A. W., Wren, B. W., Rycroft, A. N., and Consortium, B. (2013). Development of a self-replicating plasmid system for Mycoplasma hyopneumoniae. *Veterinary Research*, 44(1).

Maier, T., Marcos, J., Wodke, J. a. H., Paetzold, B., Liebeke, M., Gutiérrez-Gallego, R., and Serrano, L. (2013). Large-scale metabolome analysis and quantitative integration with genomics and proteomics data in Mycoplasma pneumoniae. *Molecular bioSystems*, 9(7):1743–1755.

Maier, T., Schmidt, A., Güell, M., Kühner, S., Gavin, A.-C., Aebersold, R., and Serrano, L. (2011). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Molecular systems biology*, 7(511):511.

Maniloff, J. (1992). *Mycoplasmas : molecular biology and pathogenesis*. American Society for Microbiology, Washingon, D.C.

Maniloff, J. (1996). The minimal cell genome: ”on being the right size”. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10004–6.

Margolin, W. (2000). Themes and variations in prokaryotic cell division. *FEMS Microbiology Reviews*, 24(4):531–548.

Margolin, W. (2005). FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol*, 6(11):862–871.

Mariscal, A. M., González-González, L., Querol, E., and Piñol, J. (2016). All-in-one construct for genome engineering using Cre-lox technology. *DNA Research*, 23(3):263–270.

Medina, J. L., Coalson, J. J., Brooks, E. G., Winter, V. T., Chaparro, A., Principe, M. F. R., Kannan, T. R., Baseman, J. B., and Dube, P. H. (2012). Mycoplasma pneumoniae CARDS toxin induces pulmonary eosinophilic and lymphocytic inflammation. *American Journal of Respiratory Cell and Molecular Biology*, 46(6):815–822.

Mengin-Lecreulx, D., Ayala, J., Bouhss, A., Van Heijenoort, J., Parquet, C., and Hara, H. (1998). Contribution of the P(mra) promoter to expression of genes in the Escherichia coli mra cluster of cell envelope biosynthesis and cell division genes. *Journal of Bacteriology*, 180(17):4406–4412.

Meur, S. K., Sikdar, A., Srivastava, N. C., and Srivastava, S. K. (1988). Rapid photometric assay of growth of Mycoplasma mycoides subsp. capri. *Journal of Applied Bacteriology*, 66:301–302.

Miyata, M. and Hamaguchi, T. (2016). Integrated information and prospects for gliding mechanism of the pathogenic bacterium Mycoplasma pneumoniae. *Frontiers in Microbiology*, 7(JUN).

Miyata, M. and Seto, S. (1999). Cell reproduction cycle of mycoplasma. *Biochimie*, 81(8-9):873–878.

Mott, M. L. and Berger, J. M. (2007). DNA replication initiation: mechanisms and regulation in bacteria. *Nature reviews. Microbiology*, 5(5):343–354.

Mushegian, A. R. and Koonin, E. V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10268–10273.

Neidhardt, F. C. (1999). Bacterial Growth : Constant Obsession with dN/dt. *Journal of Bacteriology*, 181(24):7405–7408.

Nomura, M. (1999). Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae : Diversity and Common GUEST COMMENTARY Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae : Diversity and Common Principles. *J Bacteriol*, 181(22):6857–6864.

Oshima, K., Maejima, K., and Namba, S. (2013). Genomic and evolutionary aspects of phytoplasmas. *Frontiers in Microbiology*, 4(AUG):1–8.

Pachkov, M., Dandekar, T., Korbel, J., Bork, P., and Schuster, S. (2007). Use of pathway analysis and genome context methods for functional genomics of Mycoplasma pneumoniae nucleotide metabolism. *Gene*, 396(2):215–225.

Paetzold, B., Carolis, C., Ferrar, T., Serrano, L., and Lluch-Senar, M. (2013). In situ overlap and sequence synthesis during DNA assembly. *ACS Synthetic Biology*, 2(12):750–755.

Parrott, G. L., Kinjo, T., and Fujita, J. (2016). A compendium for Mycoplasma pneumoniae. *Frontiers in Microbiology*, 7(APR):1–16.

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18):3551–3567.

Pich, O. Q., Burgos, R., Planell, R., Querol, E., and Piñol, J. (2006). Comparative analysis of antibiotic resistance gene markers in Mycoplasma genitalium: Application to studies of the minimal gene complement. *Microbiology*, 152(2):519–527.

Pollack, J. D., Myers, M. a., Dandekar, T., and Herrmann, R. (2002). Suspected utility of enzymes with multiple activities in the small genome Mycoplasma species: the replacement of the missing "household" nucleoside diphosphate kinase gene and activity by glycolytic kinases. *Omics : a journal of integrative biology*, 6(3):247–258.

Pollack, J. D., Williams, M. V., Banzon, J., Jones, M. A., Harvey, L., and Tully, J. G. (1996). Comparative Metabolism of Mesoplasma, Entornoplasma, My-

coplasma, and Acholeplasma. *International Journal of Systematic Bacteriology*, 46(4):885–890.

Pollack, J. D., Williams, M. V., and McElhaney, R. N. (1997). The comparative metabolism of the Mollicutes (Mycoplasmas): the utility for taxonomic classification and the relationship of putative gene annotation and phylogeny to enzymatic function in the smallest free-living cells. *Critical reviews in microbiology*, 23(4):269–354.

Posfai, G., Plunkett, G., Feher, T., Frisch, C., Keil, G. M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S. S., de Arruda, M., Burland, V., Harcum, S. W., and Blattner, F. R. (2006). Emergent properties of reduced-genome Escherichia coli. *Science*, 312(May):1044–1047.

Pour-El, I., Adams, C., and Minion, F. C. (2002). Construction of mini-Tn4001tet and its use in Mycoplasma gallisepticum. *Plasmid*, 47(2):129–137.

Radestock, U. and Bredt, W. (1977). Motility of Mycoplasma pneumoniae. *Journal of Bacteriology*, 129(3):1495–1501.

RayChaudhuri, D. and Park, J. T. (1992). Escherichia coli cell-division gene ftsZ encodes a novel GTP-binding protein. *Nature*, 359(6392):251–254.

Razin, S. (2010). Time-line of significant contributions to mycoplasmology. *Biologicals*, 38(2):191–192.

Razin, S. and Hayflick, L. (2010). Highlights of mycoplasma research-An historical perspective. *Biologicals*, 38(2):183–190.

Razin, S., Masover, G. K., Palant, M., and Hayflick, L. (1977). Morphology of Ureaplasma urealyticum (T-mycoplasma) organisms and colonies. *Journal of bacteriology*, 130(1):464–471.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of mycoplasmas. *Microbiology and molecular biology reviews : MMBR*, 62(4):1094–156.

Regula, J. T., Ueberle, B., Boguth, G., Görg, A., Schnölzer, M., Herrmann, R., and Frank, R. (2000). Towards a two-dimensional proteome map of Mycoplasma pneumoniae. *Electrophoresis*, 21(17):3765–3780.

Reich, K. A. (2000). The search for essential genes. *Science*, 151:319–320.

Renaudin, J., Marais, A., Verdin, E., Duret, S., Foissac, X., Laigret, F., and Bove, J. (1995). Integrative and free Spiroplasma citri oriC plasmids: Expression of the Spiroplasma phoeniceum spiralin in Spiroplasma citri. *Journal of Bacteriology*, 177(10):2870–2877.

Reuß, D. R., Commichau, F. M., Gundlach, J., Zhu, B., and Stülke, J. (2016). The Blueprint of a Minimal Cell: MiniBacillus. *Microbiology and molecular biology reviews : MMBR*, 80(4):955–987.

Rivas-Marín, E., Canosa, I., and Devos, D. P. (2016). Evolutionary cell biology of division mode in the bacterial Planctomycetes-verrucomicrobia- Chlamydiae superphylum. *Frontiers in Microbiology*, 7(DEC):1–11.

Rocha, E. P. C. (2004). Codon usage bias from tRNA's point of view: Redundancy, specialization, and efficient decoding for translation optimization. *Genome Research*, 14(11):2279–2286.

Rocha, E. P. C. and Blanchard, A. (2002). Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution. *Nucleic acids research*, 30(9):2031–42.

Rodrigues, C. D. A. and Harry, E. J. (2012). The min system and nucleoid occlusion are not required for identifying the division site in bacillus subtilis but ensure its efficient utilization. *PLoS Genetics*, 8(3).

Romano, N., Tolone, G., Ajello, F., and Licatai, R. L. A. (1980). Adenosine 5 ' -Triphosphate Synthesis Induced by Urea Hydrolysis in Ureaplasma urealyticum. *Journal of Bacteriology*, 144(2):830–832.

Rothfield, L., Taghbalout, A., and Shih, Y.-L. (2005). Spatial control of bacterial division-site placement. *Nature Reviews Microbiology*, 3(12):959–968.

Rowlett, V. W. and Margolin, W. (2015). The bacterial divisome: ready for its close-up. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370:20150028.

Russell, J. B. and Cook, G. M. (1995). Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiological reviews*, 59(1):48–62.

Saglio, P. H. M., Daniels, M. J., and Pradet, A. (1979). ATP and Energy Charge as Criteria of Growth and Metabolic Activity of Mollicutes: Application to Spiroplasma citri. *Journal of General Microbiology*, 110(1):13–20.

Sanders, G. M., Dallmann, H. G., and McHenry, C. S. (2010). Reconstitution of the B. subtilis Replisome with 13 Proteins Including Two Distinct Replicases. *Molecular Cell*, 37(2):273–281.

Sasaki, Y., Ishikawa, J., Yamashita, A., Oshima, K., Kenri, T., Furuya, K., Yoshino, C., Horino, A., Shiba, T., Sasaki, T., and Hattori, M. (2002). The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans. *Nucleic Acids Research*, 30(23):5293–5300.

Schaechter, M. (2006). From growth physiology to systems biology. *International Microbiology*, 9:157–161.

Schaechter, M. (2015). A brief history of bacterial growth physiology. *Frontiers in Microbiology*, 6(MAR):1–5.

Schaechter, M., MaalOe, O., and Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *Journal of General Microbiology*, 19(3):592–606.

Schiefer, H. G., Krauss, H., Brunner, H., and Gerhardt, U. (1975). Ultrastructural visualization of surface carbohydrate structures on mycoplasma membranes by concanavalin A. *Journal of Bacteriology*, 124(3):1598–1600.

Schiefer, H. G., Krauss, H., Schummer, U., Brunner, H., and Gerhardt, U. (1978). Studies with ferritin-conjugated concanavalin A on carbohydrate structures of mycoplasma membranes. *FEMS Microbiology Letters*, 3:183–185.

Schmidl, S. R., Gronau, K., Pietack, N., and Hecker, M. (2010). The Phosphoproteome of the Minimal Bacterium Mycoplasma pneumoniae. *Molecular & Cellular Proteomics*, pages 1228–1242.

Schreiber, G., Metzger, S., Aizenman, E., Roza, S., Cashel, M., and Glaser, G. (1991). Overexpression of the relA Gene in Escherichia coli. *The Journal of biological chemistry*, 266(6):3760–3767.

Seto, S., Layh-Schmitt, G., Kenri, T., and Miyata, M. (2001). Visualization of the attachment organelle and cytadherence proteins of Mycoplasma pneumoniae by immunofluorescence microscopy. *Journal of Bacteriology*, 183(5):1621–1630.

Seto, S., Murata, S., and Miyata, M. (1997). Characterization of dnaA gene expression in Mycoplasma capricolum. *FEMS Microbiology Letters*, 150(2):239–247.

Shahid, M. A., Marenda, M. S., Markham, P. F., and Noormohammadi, A. H. (2014). Development of an oriC vector for use in Mycoplasma synoviae. *Journal of Microbiological Methods*, 103:70–76.

Shepard, M. C. and Lunceford, C. D. (1965). Effect of pH on Human Mycoplasma Strains. *Journal of Bacteriology*, 89(2):265–270.

Sirand-Pugnet, P., Citti, C., Barré, A., and Blanchard, A. (2007). Evolution of mollicutes: down a bumpy road with twists and turns. *Research in Microbiology*, 158(10):754–766.

Skarstad, K. and Boye, E. (1994). The initiator protein DnaA: evolution, properties and function. *BBA - Gene Structure and Expression*, 1217(2):111–130.

Sluijter, M., Kaptein, E., Spuesens, E. B. M., Hoogenboezem, T., Hartwig, N. G., Van Rossum, A. M. C., and Vink, C. (2010). The Mycoplasma genitalium MG352-encoded protein is a Holliday junction resolvase that has a nonfunctional orthologue in Mycoplasma pneumoniae. *Molecular Microbiology*, 77(5):1261–1277.

Smith, P. F. (1956). Quantitative measurement of the growth of pleuropneumonia-like organisms. *Applied microbiology*, 4(5):254–259.

Snell, G. C. (1981). A Comparison of Alternative Methods to Viable Count for Indicating Growth of Mycoplasma gallisepticum in Liquid Culture. *Journal of Applied Bacteriology*, 50(2):275–281.

Solomon, J., Su, L., Shyn, S., and Grossman, A. D. (2003). Isolation and Characterization of Mutants of the Bacillus subtilis Oligopeptide Permease with Altered Specificity of Oligopeptide Transport. *Journal of Bacteriology*, 185(21):6425–6433.

Staats, C. C., Boldo, J., Broetto, L., Vainstein, M., and Schrank, A. (2007). Comparative genome analysis of proteases, oligopeptide uptake and secretion systems in Mycoplasma spp. *Genetics and Molecular Biology*, 30(SUPPL. 1):225–229.

Stemke, G. W. and Robertson, J. A. (1982). Comparison of Two methods for Enumeration of Mycoplasma. *Journal of Clinical Microbiology*, 16(5):959–961.

Stemler, M. E., Stemke, G. W., and Robertson, J. A. (1987). ATP measurements obtained by luminometry provide rapid estimation of Ureaplasma urealyticum growth. *Journal of Clinical Microbiology*, 25(2):427–429.

Su, H.-C., Hutchison, C. A., Giddings, M. C., Hutchison III, C. A., Giddings, M. C., Hutchison, C. A., Giddings, M. C., and Hutchison III, C. A. (2007). Mapping phosphoproteins in Mycoplasma genitalium and Mycoplasma pneumoniae. *BMC Microbiology*, 7(1):63.

Sulyok, K. M., Kreizinger, Z., Fekete, L., Hrivnák, V., Magyar, T., Jánosi, S., Schweitzer, N., Turcsányi, I., Makrai, L., Erdélyi, K., and Gyuranecz, M. (2014). Antibiotic susceptibility profiles of Mycoplasma bovis strains isolated from cattle in Hungary, Central Europe. *BMC Veterinary Research*, 10(1):256.

Suzuki, Y., Assad-Garcia, N., Kostylev, M., Noskov, V. N., Wise, K. S., Karas, B. J., Stam, J., Montague, M. G., Hanly, T. J., Enriquez, N. J., Ramon, A., Goldgof, G. M., Richter, R. A., Vashee, S., Chuang, R. Y., Winzeler, E. A., Hutchison, C. A., Gibson, D. G., Smith, H. O., Glass, J. I., and Venter, J. C.

(2015). Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. *Genome Research*, 25(3):435–444.

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. (2015). Cell-size control and homeostasis in bacteria. *Current Biology*, 25(3):385–391.

Thanbichler, M. (2010). Synchronization of Chromosome Dynamics and Cell Division in Bacteria. *Cold Spring Harbor Perspectives in Biology*, 2(a000331):1–16.

Thomas, E. (2015). Microbial growth and physiology: A call for better craftsmanship. *Frontiers in Microbiology*, 6(MAR):1–12.

Trussart, M., Yus, E., Martinez, S., Baù, D., Tahara, Y. O., Pengo, T., Widjaja, M., Kretschmer, S., Swoger, J., Djordjevic, S., Turnbull, L., Whitchurch, C., Miyata, M., Marti-Renom, M. A., Lluch-Senar, M., and Serrano, L. (2017). Defined chromosome structure in the genome-reduced bacterium Mycoplasma pneumoniae. *Nature Communications*, 8:14665.

Tully, J. G., Rose, D. L., Whitcomb, R. F., and Wenzel, R. P. (1979). Enhanced isolation of Mycoplasma pneumoniae from throat washings with a newly-modified culture medium. *The Journal of infectious diseases*, 139(4):478–82.

Tully, J. G., Taylor Robinson, D., Rose, D. L., Cole, R. M., and Bove, J. M. (1983). Mycoplasma genitalium, a new species from the human urogenital tract. *International Journal of Systematic Bacteriology*, 33(2):387–396.

Uchiyama, I. (2003). MBGD: Microbial genome database for comparative analysis. *Nucleic Acids Research*, 31(1):58–62.

Ueberle, B., Frank, R., and Herrmann, R. (2002). The proteome of the bacterium Mycoplasma pneumoniae: comparing predicted open reading frames to identified gene products. *Proteomics*, 2(6):754–764.

van Noort, V., Seebacher, J., Bader, S., Mohammed, S., Vonkova, I., Betts, M. J., Kühner, S., Kumar, R., Maier, T., O'Flaherty, M., Rybin, V., Schmeisky, A., Yus, E., Stülke, J., Serrano, L., Russell, R. B., Heck, A. J. R., Bork, P., and Gavin, A.-C. (2012). Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Molecular systems biology*, 8(571):571.

Vieira-Silva, S. and Rocha, E. P. (2010). The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genetics*, 6(1).

Vivian, J. P., Porter, C. J., Wilce, J. A., and Wilce, M. C. J. (2007). An Asymmetric Structure of the Bacillus subtilis Replication Terminator Protein in Complex with DNA. *Journal of Molecular Biology*, 370(3):481–491.

Voulgaris, J., French, S., Gourse, R. L., Squires, C., and Squires, C. L. (1999). Increased rrn gene dosage causes intermittent transcription of rRNA in Escherichia coli. *Journal of Bacteriology*, 181(14):4170–4175.

Waites, K. B. and Talkington, D. F. (2004). Mycoplasma pneumoniae and its role as a human pathogen. *Clinical Microbiology Reviews*, 17(4):697–728.

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., and Jun, S. (2010). Robust growth of escherichia coli. *Current Biology*, 20(12):1099–1103.

Wanner, U. and Egli, T. (1990). Dynamics of microbial growth and cell composition in batch culture. *FEMS Microbiology Reviews*, 75:19–44.

Warren, A. S. and Setubal, J. (2009). The Genome Reverse Compiler: an explorative annotation tool. *BMC Bioinformatics*, 10(1):35.

Weiner III, J. (2003). Transcription profiles of the bacterium Mycoplasma pneumoniae grown at different temperatures. *Nucleic Acids Research*, 31(21):6306–6320.

Weiss, D. S. (2004). Bacterial cell division and the septal ring. *Molecular Microbiology*, 54(3):588–597.

Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nature Methods*, 6(5):359–362.

Wium, M., Botes, A., and Bellstedt, D. U. (2015). The identification of oppA gene homologues as part of the oligopeptide transport system in mycoplasmas. *Gene*, 558(1):31–40.

Wodke, J. A. H. (2012). *Organization and Integration of Large-scale Datasets for Designing a Metabolic Model and Re-annotating the Genome of Mycoplasma pneumoniae*. Doctoral thesis, Humboldt-Universität zu Berlin.

Wodke, J. A. H., Puchałka, J., Lluch-Senar, M., Marcos, J., Yus, E., Godinho, M., Gutiérrez-Gallego, R., dos Santos, V. a. P. M., Serrano, L., Klipp, E., and Maier, T. (2013). Dissecting the energy metabolism in Mycoplasma pneumoniae through genome-scale metabolic modeling. *Molecular systems biology*, 9(653):653.

Wolanski, M., Donczew, R., Zawilak-Pawlik, A., and Zakrzewska-Czerwinska, J. (2015). oriC-encoded instructions for the initiation of bacterial chromosome replication. *Frontiers in Microbiology*, 6(JAN):1–14.

Xia, X. (2003). DNA methylation and Mycoplasma genomes. *Journal of molecular evolution*, 57 Suppl 1:S21–8.

Yano, K., Wada, T., Suzuki, S., Tagami, K., Matsumoto, T., Shiwa, Y., Ishige, T., Kawaguchi, Y., Masuda, K., Akanuma, G., Nanamiya, H., Niki, H., Yoshikawa, H., and Kawamura, F. (2013). Multiple rRNA operons are essential for efficient cell growth and sporulation as well as outgrowth in Bacillus subtilis. *Microbiology (United Kingdom)*, 159(PART11):2225–2236.

Yuan, X., Couto, J. M., Glidle, A., Song, Y., Sloan, W., and Yin, H. (2017). Single-cell microfluidics to study the effects of genome deletion on bacterial growth behaviour. *ACS Synthetic Biology*, page acssynbio.7b00177.

Yus, E., Güell, M., Vivancos, A. P., Chen, W.-H., Lluch-Senar, M., Delgado, J., Gavin, A.-C., Bork, P., and Serrano, L. (2012). Transcription start site associated RNAs in bacteria. *Molecular systems biology*, 8(585):585.

Yus, E., Maier, T., Michalodimitrakis, K., van Noort, V., Yamada, T., Chen, W.-H., Wodke, J. a. H., Güell, M., Martínez, S., Bourgeois, R., Kühner, S., Raineri, E., Letunic, I., Kalinina, O. V., Rode, M., Herrmann, R., Gutiérrez-Gallego, R., Russell, R. B., Gavin, A.-C., Bork, P., and Serrano, L. (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science (New York, N.Y.)*, 326(5957):1263–1268.

Yus, E., Yang, J. S., Sogues, A., and Serrano, L. (2017). A reporter system coupled with high-throughput sequencing unveils key bacterial transcription and translation determinants. *Nature Communications*, 8:1–12.

Zackrisson, M., Hallin, J., Ottosson, L.-G., Dahl, P., Fernandez-Parada, E., Ländström, E., Fernandez-Ricaud, L., Kaferle, P., Skyman, A., Stenberg, S., Omholt, S., Petrovič, U., Warringer, J., and Blomberg, A. (2016). Scan-o-matic: High-Resolution Microbial Phenomics at a Massive Scale. *G3: Genes—Genomes—Genetics*, 6(9):3003–3014.

Zaritsky, A. and Woldringh, C. L. (2015). Chromosome replication, cell growth, division and shape: A personal perspective. *Frontiers in Microbiology*, 6(AUG):1–11.

Zimmerman, C.-U. and Herrmann, R. (2005). Synthesis of a small, cysteine-rich, 29 amino acids long peptide in Mycoplasma pneumoniae. *FEMS microbiology letters*, 253(2):315–321.

Ziv, N., Siegal, M. L., and Gresham, D. (2013). Genetic and nongenetic determinants of cell growth variation assessed by high-throughput microscopy. *Molecular Biology and Evolution*, 30(12):2568–2578.

Zucker-Franklin, D., Davidson, M., and Thomas, L. (1966). The interaction of Mycoplasmas with mammalian cells. *The Journal of experimental medicine*, 124(3):521–32.