Study of a master regulator of recombination in *Mycoplasma genitalium*

Sergi Torres Puig
2017

Doctoral thesis submitted in the Biochemistry, Molecular Biology and Biomedicine PhD program at the Universitat Autònoma de Barcelona.

This work was developed at the Institut de Biotecnologia i Biomedicina (IBB) and at the Biochemistry and Molecular Biology Department of the Universitat Autònoma de Barcelona, supervised by Dr. Òscar Quijada Pich and Dr. Jaume Pinyol Ribas.

Dr. Òscar Quijada Pich   Dr. Jaume Pinyol Ribas

08193-Cerdanyola del Vallès   Phone +34 935868952   sergi.torres@uab.cat

Sergi Torres Puig
PhD candidate

Laboratori de Biologia Molecular
Institut de Biotecnologia i Biomedicina (IBB)
Departament de Bioquímica i Biologia Molecular
Universitat Autònoma de Barcelona (UAB)

Every individual matters. Every individual has a role to play. Every individual makes a difference.

Jane Goodall
Primatologist

Als meus pares
Al meu avi

Table of contents

AGRAÏMENTS/ACKNOWLEDGEMENTS

Això no només va de ciència...

En aquest espai vull agrair a totes aquelles persones que han permès que jo pugui arribar fins aquí. En primer lloc, vull destacar l'ajut dels meus directors de tesi, l'Òscar Quijada i en Jaume Pinyol. Òscar, quan vaig començar aquesta aventura no sabia el que em vindria a sobre. Han estat molts moments d'alts i baixos, de sorpreses i fracassos, però sempre he comptat amb el teu suport. Moltes gràcies per ensenyar-me tant i per donar forma al científic que sóc avui dia. A tu, Jaume, vull agrair-te la confiança que sempre has tingut en mi des d'aquell Juliol de 2010 que em vaig plantar en aquest laboratori. Gràcies per ensenyar-me la passió per la professió i per la ciència, pels consells i per la teva ajuda tots aquests anys. També dono gràcies a l'Enrique, al Pepo i a l'Ángel, per donar-me tot el suport.

No podia presentar aquesta tesi doctoral sense donar les gràcies als MoleK@s d'ahir i d'avui, és per vosaltres que vaig escollir la UAB per a fer el doctorat. A tu, Alis, per ensenyar-me tantes coses durant les meves pràctiques, gran part de com faig les coses en un laboratori és gràcies a tu. Noe, gràcies per la teva paciència els primers dies de tesi i per ensenyar-me tot el que calia per a fer V.F. Xavi, no em sé imaginar el laboratori sense tu, i mira que ja fa uns anys que has marxat (no massa lluny), però trobem a faltar les teves normes de l'Äkta o els teus interminables lab meetings... A ti, Luis González, por transmitirme tu pasión por el laboratorio por tus ideas locas y por el respeto que siempre me has tenido. Lo siento, no terminé el proyecto del MG_386 ☹ Luis Garcia, has estat un exemple a seguir per mi en molts nivells, ets molt gran! Espero que seguim sent amics molts anys més! I com no, no podia oblidar-me de tu, Ana, vam començar quasi en paral·lel aquesta sogeria i sempre he pogut comptar amb tu, sempre. No ho podia haver aconseguit sense tu, recorda que ets una gran científica, no ho oblidis mai! En aquests darrers anys hem tingut la incorporació de tres membres importantíssims: la Jara, l'Antino i en Carlos. Jara, no eres oficialmente BioMol, pero has fecho tantos méritos que ya yes una d'as nuestras. Ha estat un zaguer any de tesi brutal, en gran part gràcies a tu. Antino, gracias por ser el postdoc de este lab, gracias por tu sentido del humor, por tu experiencia y por escucharme y entenderme en esos momentos cuando las cosas no estaban bien. Eres el "experto" en tesis, y lo sabes. Carlos, ets i has estat un dels grans dins els MoleK@s. Gràcies per fer més lleuger el final del viatge, per ser un gran company de lab, un bon entrenador de running, un gran cinèfil (encara que no preparis bé les crispetes), un artista de l'absurd, un fantàstic company de congressos (ehem, ehem)... En definitiva, un gran amic. A tots, 10^3 gràcies!

En aquests darrers anys també han passat pel laboratori els nostres "becarios" (carinyosament), que ens han fet (quasi sempre) la vida més fàcil. Gràcies Ruben, Marta Hernández, Ari, Carmen, Marta Huguet, Ignasi, Christian, Marina, Esther, Santi, Miguel i Lucía. Sense vosaltres hauríem estat molt avorrits tots aquests anys. Vull agrair especialment la teva feina, Ignasi, per involucrar-te tant i currar de valent. Com ja et vaig dir fa un temps, m'enduc un gran amic, espero tenir l'oportunitat de treballar junts un altre cop. También darte las gracias a ti, Lucía, por echarme un cable este último año y por ser indispensable para el lab a muchos niveles. Llegarás lejos, date tiempo, ¡porque vales mucho!

Christine, je vous remercie de me avoir accueilli à Toulouse. Emilie, malgré le sabotage, je vous remercie de me apprendre tant de choses. Eric, Gabriela, Eveline, les deux Xaviers, Marie-Claude et Marion : sans vous mon séjour aurait été très ennuyeux, merci pour les bières et une grande soirée. Merci à tous! Also thank you, Scotty, for helping me through ASM and for showing me your lab. And Beth, for the revision of my results!

Vull donar les gràcies també a tota la gent de l'IBB que m'ha acompanyat tots aquests anys. Especialment agrair-vos a vosaltres, Miguel, Almudena, Anaís, Francesca, Fran, Olga, Sílvia i a la gent del Servei de Genòmica, Roger, Anna, Guillem i Oliva. Almu, gràcies per arreglar tots els desastres i per les teves "mans miraculoses". Sense vosaltres, no hagués estat igual.

En aquesta etapa també he d'agrair als de sempre. Jess, Humbert, Dani, Aldo, Lau, Luna, Xavi, Paula, Sílvia i Teo, sou enormes, tots. Fa molts anys que ens coneixem i sé que aquest cop ha estat dur, no he pogut estar al vostre costat tot el temps que m'hagués agradat. No em sé imaginar la vida sense vosaltres, gràcies per aplanar-me el camí, gràcies per seguir caminant amb mi. Aquesta gesta no és només meva, també és vostra.

No m'oblido tampoc de vosaltres, Estrella, Albert i Marta. Estrella, ets el meu exemple de que amb esforç i ganes un pot aconseguir el que vulgui. Survive! Marta, gràcies per mostrar-me el meu jo més "irracional", m'hi seguiré esforçant ☺. A tu Albert, gràcies per donar-me oxigen durant els inicis i pel suport que tu i la teva família sempre m'heu donat. Eskerrak eman nahi dut, Leire, por seguir a mi lado a pesar del océano. Imaginatzen zaitut, obituz olatu, bonokatuz...

I a tu, Marta, per ser providencial en els inicis d'aquesta aventura, no ho hagués aconseguit sense tu. Sé que sovint no he estat a l'alçada i que el repte de vegades ha pogut més amb mi que jo amb ell. Espero poder-ho compensar amb el temps. 我们现在和将来. Xiè Xiè!

En darrer lloc, vull agrair als meus pares per donar-me la oportunitat d'arribar fins aquí. Gràcies per tot el suport que he rebut, per ajudar-me quan he trobat pedres al camí i per exigir sempre el millor de mi. Aquesta tesi també va per als mestres de la meva vida, sobretot per a tu, avi, que seguir que avui n'estaries molt orgullós.

Gràcies de tot cor!

ABBREVIATIONS

6-Fam	6-carboxyfluorescein
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BER	Base excision repair
bp	Base pairs
CARDS	Community acquired respiratory distress syndrome
cat	Chloramphenicol acetyl transferase
CDS	Coding DNA sequence
CIRCE	Chaperon expression
Cm	Chloramphenicol
CT	Chromosomal transfer
CPX	Ciprofloxacin
DCT	Distributive conjugal transfer
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonucleic acid nuclease
DR	Downstream region
DSB	Double-strand break
dsRNA	Double stranded RNA
DTT	Dithiothreitol
ECF	Extracytoplasmic function
EDTA	Ethylenediaminetetraacetic acid
EGTA	Egtazic acid
eYFP	Enhanced yellow fluorescent protein
FBS	Fetal bovine serum
gDNA	Genomic DNA
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
HJ	Holliday junction
HP	Hypothetical protein
HPV	Human papillomavirus
HR	Homologous recombination
HTH	Helix-turn-helix
ICE	Integrative conjugative element
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Inverted repeat
kb	Kilobase
KO	Knock-out
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
MgPar	MgPa repeat
MiniTnp	Minitransposon
MMC	Mitomycin C
mRNA	Messenger RNA
MS	Mass spectrometry
ncRNA	Non-coding RNA

NER	Nucleotide excision repair
NGU	Non-gonococcal urethritis
NRT	Non-retrotranscribed control
nt	Nucleotides
NTC	Non-template control
O/N	Overnight
pac	Puromycin acetyl transferase
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIPES	Piperazine-N,N′-bis(2-ethanesulfonic acid)
Pm	Puromycin
PMSF	Phenylmethane sulfonyl fluoride
ppm	Parts per million
PVDF	Polyvinylidene fluoride
qRT-PCR	Quantitative real-time PCR
R-M	Restriction-modification
RNA	Ribonucleic acid
RNase	Ribonucleic acid nuclease
RNAP	RNA polymerase
ROS	Reactive oxygen species
RPKM	Reads per kilobase million
rpm	Revolutions per minute
RrlA	Recombination regulatory loci A (MG_220)
RrlB	Recombination regulatory loci B (MG_RS02200)
rRNA	Ribosomal RNA
RT	Retrotranscriptase
RT (ºC)	Room temperature
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SOB	Super optimal broth
SOE-PCR	Splicing by overlap extension-PCR
SP4	Spiroplasma-4 broth
SSR	Simple sequence repeats
Supp.	Supplementary
TAP	Tandem affinity purification
TC	Total counts
TCA	Trichloroacetic acid
Tet	Tetracycline
tetM438	Tetracycline resistance marker
TEV	Tobacco etch virus
TSS	Transcriptional start site
UR	Upstream region
UV	Ultraviolet light
WT	Wild-type strain
YAC	Yeast artificial chromosome
σ^{20}	Alternative sigma factor-20 (MG_428)

GENERAL INTRODUCTION

i.1. *MYCOPLASMA*, A UNIQUE BACTERIAL GENUS

i.2. *MYCOPLASMA GENITALIUM*, AN INFECTIOUS MINIMAL GENOME BACTERIUM

i.1. *MYCOPLASMA*, A UNIQUE BACTERIAL GENUS

Mycoplasma is a bacterial genus that belongs to the class Mollicutes. One of the major features of mycoplasma cells is the absence of a cell wall, which leads to cell plomorphism and imposes membrane rigidness. Since their discovery in 1843 by Louis Pasteur as possible causal agents of pleuropneumonia in cattle, mycoplasmas have been misclassified as viruses, eukaryotes or even bacterial L-forms (Klieneberger, 1935). In 1956, Edward and Freundt proposed a separate classification and nomenclature for these organisms (Edward and Freundt, 1956), but it wasn't until 1967 that the class Mollicutes was created by the Subcommittee of Taxonomy of the Mycoplasmatae (Edward and Freundt, 1967). Since then, more than 120 different species have been identified.

Mycoplasmas are parasites of animals and humans, while Mollicutes that infect plants (normally through insect vectors) are classified in the *Phytoplasma*, *Acholeplasma, Mesoplasma* or *Spiroplasma* genus. Mycoplasmas are widely distributed as pathogens or commensals and usually exhibit strict host and tissue specificity, probably due to their obligate parasitic lifestyle. The list of hosts harboring mycoplasmas is continuously increasing as more advanced molecular tools are available for taxonomy purposes (Razin, 1994). The number of known mycoplasma species can widely increase as new cell wall-less bacteria are discovered, such as *Eperythrozoon sp.* (Neimark and Kocan, 1997) or *Haemobartonella sp.* (Rikihisa *et al.*, 1997), two pathogenic bacteria that infect erythrocytes of their vertebrate hosts, as well as other hemotrophic mycoplasmas (hemoplasmas) (Messick, 2004; Biondo *et al.*, 2009).

Mycoplasmas are well known as a frequent source of cell culture contamination. Their small size and innate resistance to many antibiotics makes them difficult to detect using conventional microscopy techniques, and difficult to eradicate. Contamination occurs mainly from laboratory personnel or cell culture ingredients such as contaminated bovine serum (Drexler and Uphoff, 2002). It is estimated that ~15% of laboratory cell cultures were contaminated with mycoplasmas in the mid-90's (source: Food and Drug Administration, FDA) and that ~1% of published Gene Expression Omnibus (GEO) data is in fact mycoplasma cDNA (Miller *et al.*, 2003; Aldecoa-Otalora *et al.*, 2009).

i.1.1. Genome reduction and plasticity

Mycoplasma genomes figures among the smallest of all bacterial domain. They range from the 580kb of *Mycoplasma genitalium* the smallest to the 1360kb of *Mycoplasma penetrans* the largest. Thus, mycoplasmas have evolved from Gram-positive clostridia through an extensive genome reduction and loss of gene redundancy (Woese, 1987). Mycoplasmas with larger genomes are considered "older" Mollicutes than mycoplasmas with smaller genomes. Moreover, their genomes have a characteristic low G+C content (see Table i.1), ranging between 25-35% average; and intergenic regions are particularly rich in A+T content (often about 90%). Besides, mycoplasmas have an AT-biased codon usage and use UGA as tryptophan codon. Consequently, they lose the stop codon-recognizing peptide chain release factor 2 (RF2) (Inagaki *et al.*, 1993, 1996).

Table i.1. Properties of representative Mycoplasmas. Adapted from Dybvig and Voelker, 1996.

Organism	Genome size (kb)	Mol% G+C	Host	Phylogenetic group
M. arthritidis	840	31	Rodent	Hominis
M. bovis	1100	28–33	Cow	Fermentans
M. capricolum	1100	25	Goat	Mycoides
M. fermentans	1160	27	Human	Fermentans
M. flocculare	1200	ND	Pig	Hyorhinis
M. gallisepticum	1000	31	Bird	Pneumoniae
M. genitalium	580	32	Human	Pneumoniae
M. ominis	700	30	Human	Hominis
M. hyopneumoniae	1100	28	Pig	Hyorhinis
M. hyorhinis	820	27	Pig	Hyorhinis
M. iowae	1300	25	Bird	Pneumoniae
M. mobile	780	24	Fish	Hyorhinis
M. mycoides	1300	25	Cow, goat	Mycoides
M. pneumoniae	800	40	Human	Pneumoniae
M. pulmonis	950	28	Rodent	Fermentans
M. salivarium	900	29	Human	Hominis

ND, not determined

Despite this extreme genome size reduction, most mycoplasma genomes contain repetitive elements such as insertion-like sequence elements (IS-like) or genomic repeats. In some species, these genomic repeats have been associated with antigenic variation and pathogenicity (Kenri *et al.*, 1999; Rocha and Blanchard, 2002). On the other hand, IS-like elements in mycoplasmas have shown to be capable of transposition within the chromosome, with the notable exception of IS1138 of *M. pulmonis* (Bhugra and Dybvig, 1993).

Extrachromosomal DNA is rarely found in mycoplasmas, although several plasmids and replicative phage genomes have been identified (Maniloff and Dybvig, 1988). The most studied mycoplasma extrachromosomal plasmids are found in *Mycoplasma mycoides* and they code for factors implicated in plasmid maintenance (King and Dybvig, 1992). The most characterized mycoplasma viruses are P1 (infecting *Mycoplasma pulmonis*) (Tu *et al.*, 2001) and MAV1 (infecting *Mycoplasma arthritidis*). Of note, the latter encodes a viral factor that increases bacterial virulence (Voelker and Dybvig, 1999).

As in many other prokaryotes, mycoplasmas encode restriction and modification (R-M) systems in their genomes (Brocchi *et al.*, 2007). Type I R-M systems are largely widespread, but some species like *M. genitalium* only have the gene coding for the S subunit of a vestigial R-M system (HsdS, MG_438) and perhaps suggesting a moonlight function (Calisto *et al.*, 2005). Type II and III R-M systems are less common but have also been found in some mycoplasmas (Dybvig *et al.*, 2007). It is thought that DNA methylation in mycoplasmas protects the chromosome against possible DNA intrusions, but unique Gm^{6}ATC methylase activity not related to DNA protection has been observed in several species as well (Bergemann *et al.*, 1990), with no defined role so far. As with other enzymes purified from many bacterial species, *Mycoplasma fermentans Mfe*I endonuclease (which recognizes CAATTG as the target sequence) has been included in the common restriction enzymes available on the market (Halden *et al.*, 1989).

i.1.2. Metabolism of mycoplasmas

The extreme genome reduction and the parasitic lifestyle of mycoplasmas has led to a dramatic decrease in the number of genes related to cell metabolism and catabolism. All mycoplasmas have truncated respiratory systems, relying on fermentation of several substrates or arginine hydrolysis for energy production (Table i.2). Their lack of a complete tricarboxylic acid cycle or quinones results in poor ATP yields and relatively large quantities of metabolic end products (Miles, 1992). Some mycoplasmas are capable of oxidation of organic acids using alternative metabolic pathways (Taylor *et al.*, 1994), but the contribution of these substrates to the energy provision is unclear.

Genome streamlining has also a big impact on genes from the biosynthetic pathways. Mycoplasmas have lost most genes involved in amino acid and cofactor biosynthesis. Lack of all these genes has resulted in a small genome, but it also has increased their dependence on robust metabolite transport systems (Razin *et al.*, 1998). Moreover, mycoplasmas incorporate host phospholipids and large quantities of exogenous cholesterol

into their cell membranes to regulate membrane fluidity despite their lack of endogenous fatty acid biosynthesis networks (Dahl, 1993).

Table i.2. Primary sites of colonization, metabolism and pathogenicity of mycoplasmas isolated from humans. Adapted from Taylor-Robinson, 1997.

| Species | Primary site of colonization | | Metabolism of | | Pathogenicity |
	Oropharynx	Genitourinary tract	Glucose	Arginine	
M. buccale	+	-	-	+	Non-pathogenic[a]
M. faucium	+	-	-	+	Non-pathogenic
M. fermentans	+	±[b]	+	+	Detected in joints in inflammatory arthritides and in lungs in HIV infection
M. genitalium	±	+	+	-	Acute NGU
M. hominis	±	+	-	+	A possible cause of pelvic inflammatory disease
M. liophilum	+	-	-	+	Non-pathogenic
M. orale	+	-	-	+	Non-pathogenic
M. penetrans	-	+	+	+	Associated with HIV infection
M. pirum	?	?	+	+	Non-pathogenic
M. pneumoniae	+	±	+	-	Atypical pneumonia and sequelae
M. primatum	-	+	-	+	Non-pathogenic
M. salivarium	+	-	-	+	Non-pathogenic, causes arthritis in hypogammaglobulinaemia
M. spermatophilum	-	+	-	+	Non-pathogenic
U. urealyticum[c]	+	+	-	-	Chronic NGU
A. laidlawii	+	-	+	-	Non-pathogenic
A. oculi	?	-	+	-	Non-pathogenic

[a] "Non-pathogenic" means that no evidence for pathogenicity is available.
[b] ± = primary use occasionally.
[c] Metabolizes urea.

Recently, large studies on the metabolome of *Mycoplasma pneumoniae* have been conducted (Yus *et al.*, 2009; Maier *et al.*, 2013). Despite the data integration of multiple metabolic pathways and reactions, several enzymatic activities are still carried out by "orphan enzymes" or by uncharacterized moonlighting proteins.

i.1.3. Mycoplasma infection and disease

Mycoplasma infections are typically very fastidious and persistent in time. The initial steps of mycoplasma infection involve the attachment and possible internalization into the host cells. Adhesion to host tissue is an essential step for colonization and infection and it relies on cytadhesins or mycoplasma specialized lipoproteins as well as other accessory membrane proteins (Krause and Baseman, 1982; Fisseha *et al.*, 1999; Burgos *et al.*, 2006;

Pich *et al.*, 2008; Indikova *et al.*, 2013). It has been shown that mycoplasmas need sialidated lipids or sialidated glycoproteins present in host cell membranes for attachment (Razin *et al.*, 1981; Baseman *et al.*, 1982). Although it has been clearly established in some cases, whether most animal mycoplasmas can invade epithelial cells during infection is still a matter of debate (Taylor-Robinson *et al.*, 1991; Lo *et al.*, 1993; Jensen *et al.*, 1994; Winner *et al.*, 2000; Blaylock *et al.*, 2004). The exact mechanism used to enter the host cells is not well understood, though some mycoplasmas such as *M. genitalium* or *M. penetrans* use a specialized tip structure (known as terminal organelle) to facilitate this process (Ueno *et al.*, 2008). In some species, internalization in target cells can be halted using actin polymerization inhibitors such as cytochalasin D (Andreev *et al.*, 1995). This internalized reservoir could have major implications in infection persistence and immune evasion (Baseman and Tully, 1997; McGowin *et al.*, 2009b).

While most mycoplasmas are commensals living innocuously within their host, some species have been associated to several diseases. Pathogenic mycoplasmas infect and colonize the respiratory membranes, the genitourinary tract, the conjunctiva, the mammary glands or joints, causing both acute and chronic diseases. In humans, the most common mycoplasma infections are caused by *M. pneumoniae*, *M. genitalium*, *Mycoplasma hominis*, *M. penetrans*, *Ureaplasma urealyticum* and *Ureaplasma parvum*. *M. pneumoniae* is the main agent causing primary atypical pneumonia or "walking pneumonia" as well as chronic lung conditions due to vigorous immune responses (Waites and Talkington, 2004). *M. genitalium* infections are sometimes asymptomatic (Clarivet *et al.*, 2014; Henning *et al.*, 2014; Philibert *et al.*, 2014), but can cause non-gonococcal urethritis (NGU), cervicitis, vaginal discharge and pelvic inflammatory diseases (Manhart, 2013). *M. hominis* can cause bacterial vaginosis, pelvic inflammatory disease, post-partum fevers and male infertility (Mårdh, 1983). *U. urealyticum* and *U. parvum* infections have been linked as causants of NGU as well as abortions or preterm birth, although their exact role in developing these illnesses is not yet clear (Ireland and Keelan, 2014; Zhang *et al.*, 2014; Sweeney *et al.*, 2017). *M. penetrans* pathogenesis is still a matter for debate, although some studies have established a link between *M. penetrans* infection and HIV progression (Blanchard and Montagnier, 1994; Grau *et al.*, 1995) or nephropathies (Jiang *et al.*, 2013). Some mycoplasmas infecting humans have also oncogenic potential and can induce chromosome aberrations in prolonged infections (Tsai *et al.*, 1995).

Mycoplasma infections have also an impact in the stockbreeding industry, affecting pigs (*Mycoplasma hyopneumoniae*), cattle (*Mycoplasma bovis*) or sheep and goats

(*Mycoplasma agalactiae*). Outbreaks of these pathogens have derived in losses of whole herds and huge economic setbacks (DaMassa *et al.*, 1992).

Damage to mucosal surfaces due to mycoplasma infection is caused mainly by reactive oxygen species (ROS) derived from bacterial metabolism. Hydrogen peroxide (H_2O_2), the end product of catabolism of many mycoplasmas, is responsible for most tissue damage (Cole *et al.*, 1968; Cherry and Taylor-Robinson, 1970; Pilo *et al.*, 2005; Schott *et al.*, 2014) and hydrogen peroxide-producing enzymes are considered virulence factors in some cases (Zhao *et al.*, 2017). However, production of H_2O_2 is not the only toxic factor secreted by mycoplasmas, as non-ROS producing mutants are still virulent (Hames *et al.*, 2009; Szczepanek *et al.*, 2014). Despite reactive oxygen species are secreted, lack of catalases coded in mycoplasma genomes causes an important fitness setback in these small pathogens (Simmons and Dybvig, 2015).

M. pneumoniae possesses another virulence factor that is unique among other mycoplasmas. The community-acquired respiratory distress syndrome (CARDS) toxin is an ADP-ribosyltransferase that interferes with transduction pathways within the host and induces strong vacuolization and damage to respiratory cells (Kannan and Baseman, 2006).
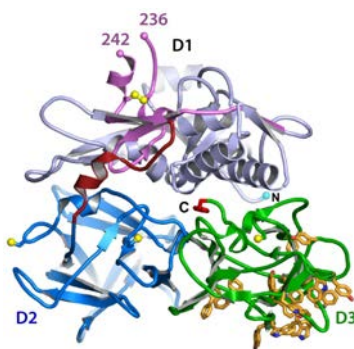
Figure i.1. CARDS toxin tertiary structure. It has three distinctive domains, marked as D1, D2 and D3. Extracted from Becker *et al.*, 2015.

Expression of CARDS toxin is up-regulated during infection, exacerbating cytotoxicity in host cells (Kannan *et al.*, 2010). CARDS toxin causes "hyperinflammation" of the targeted tissue via IL-1β (Bose *et al.*, 2014), causing acute and chronic airway diseases such as asthma. As it is well described in CARDS-derived pathogenesis, mycoplasma infections often cause the hyperactivation of the innate immune response displayed by their hosts. Inflammation caused by mycoplasma lipoproteins can cause strong inflammatory responses and thus increase and worsen infection symptoms (Into *et al.*, 2004; McGowin *et al.*, 2009a; Shimizu, 2016).

i.1.4. Treatment and vaccine development

Mycoplasmas are resistant to antibiotics that specifically target bacterial wall synthesis. Moreover, DNA-dependent RNA polymerases from mycoplasmas are not targeted by rifampicin; at the same time nalidixic acid cannot inhibit DNA gyrase or topoisomerases of these bacteria (Table i.3). These facts limit antibiotic treatment to the use of tetracyclines, macrolides and the new-generation quinolones.

Table i.3. Susceptibilities of *M. pneumoniae*, *M. genitalium*, *M. hominis*, *M fermentans* and *U. urealyticum* to various antibiotics[a]. Adapted from Taylor-Robinson, 1997.

Antibiotic(s)	*M. pneumoniae*	*M. genitalium*	*M. hominis*	*M. fermentans*	*U. urealyticum*
Tetracyclines	++	++	++[b]	++	++[b]
Erythromycin	++	++	-	+	++
Clarithromycin	++	++	-	-	++
Azythromycin	++	++	-	++	+
Pristinamycin	++	++	++	?	++
Streptomycin	++	++	-	-	+
Spectinomycin	++	?	+	+	+
Gentamicin	+	++[c]	+	+	+
Chloramphenicol	+	+	+	+	+
Clindamycin	+	+	++	++	-
Lincomycin	+	+	++	++	-
Sparfloxacin	++	++	++	++	++
Ciprofloxacin	+	+	++	++	+
Difloxacin	+	?	++	?	+
Nalidixic acid	-	-	-	-	-
Cephalosporins	-	-	-	-	-
Penicillins	-	-	-	-	-
Rifampicin	-	-	-	-	-

[a] ++, susceptible (MIC< 1mg/mL); +, partially susceptible (MIC = 1-10mg/mL); -, resistant (MIC> 10mg/mL). Results are presented mostly in order of diminishing activity for *M. pneumoniae*.
[b] Organisms within this species that carry the TetM determinant are not susceptible to tetracyclines.
[c] Experimental data from our lab and others demonstrates gentamicin susceptibility of *M. genitalium* G37 strain (Dhandayuthapani *et al.*, 1999; Pich *et al.*, 2006b), although this fact was unknown (?) in the original work published in 1997.

Nowadays, infections of *M. pneumoniae* are often treated with azythromycin or erythromycin, which have a bacteriostatic effect that reduce the longevity and intensity of infections (Waites and Talkington, 2004). *M. genitalium* infections are treated similarly, although doxycycline and moxifloxacin are also used (Bradshaw *et al.*, 2006).

As for many other bacterial infections, mycoplasmas are acquiring antibiotic resistances and treatment failures are escalating worldwide (Jensen and Bradshaw, 2015; Beeton and Spiller, 2017). Mutations at the 23S rRNA subunit arise and reduce the efficacy of macrolides, while mutations at DNA gyrases and DNA topoisomerases make the use of classic quinolones ineffective (Antunes *et al.*, 2015; Sulyok *et al.*, 2016; Piccinelli *et al.*, 2017).

As the antibiotic resistance in bacteria arises, regulations on the use of antibiotics in livestock are becoming more restrictive and the need for alternative treatments and protective vaccines has become crucial. However, development of effective vaccines for mycoplasma infections has proven to be extremely difficult, as most of current available vaccines consist on mildly attenuated strains or bacterin-derived vaccines with modest results (Nicholas *et al.*, 2009). A total of 26 vaccines are licensed or under late research stages (2014), most of them focused on treating *M. hyopneumoniae*, *M. gallisepticum* and *M. synoviae* infections (He *et al.*, 2014).

Table i.4. Currently used mycoplasma vaccines (2009). Extracted from Nicholas, Ayling and McAuliffe, 2009.

Disease	Causative mycoplasma	Main host	Vaccine type
Contagious bovine pleuropneumonia	*M. mycoides* subs. *mycoides* SC	Cattle	T1/44: live semivirulent passaged in eggs 44 times (Africa)
Contagious caprine pleuropneumonia	*M. capricolum* subs. *capripneumoniae*	Goats	Whole cells inactivated by saponin (Kenya, Ethiopia)
Contagious agalactia	*M. agalactiae*	Sheep, goats	Formalin-inactivated (Europe) Live multipassaged strain (Turkey)
Contagious agalactia	*M. agalactiae*, *M. mycoides* subs. *mycoides* LC, *M. capricolum*	Sheep, goats	Whole cell inactivated, aluminum hydroxide adjuvant
Contagious agalactia, pleuropneumonia	*M. mycoides* subs. *mycoides* LC	Goats	Live multipassaged strain (Turkey)
Calf pneumonia	*M. bovis*	Cattle	Autogenous vaccines (killed) (USA, UK)
Enzootic pneumonia	*M. hyopneumoniae*	Pigs	Inactivated vaccine (worldwide)
Chronic respiratory disease	*M. gallisepticum*	Poultry	Live TS mutant (USA, Europe)
Infectious synovitis	*M. synoviae*	Poultry	Live attenuated (USA, Europe)

Human vaccines for mycoplasma infections have mainly targeted atypical pneumoniae derived from *M. pneumoniae* infections. Although several strategies have been attempted (Barile *et al.*, 1988), vaccines did not reduce the number of infections or the severity of the symptoms snf some of them even exacerbated the disease (Jacobs *et al.*, 1988). Due to these discouraging results, *M. pneumoniae* vaccine development was discontinued (Linchevski *et al.*, 2009).

In mid-2015, a European project (MycoSynVac) coordinated by several laboratories received funding for the development of vaccines using an *M. pneumoniae* engineered chassis.

i.1.5. Immune evasion strategies of minimal cells

Despite their reduced genomes and limited set of genes, mycoplasmas have developed a wide range of strategies to generate variants of their surface exposed proteins. Generally, all strategies are based on rapid and reversible genetic changes that can produce variants in a combinatorial manner. These strategies include but are not limited to DNA slippage, site-specific recombination, gene rearrangement, gene conversion and reciprocal recombination (Citti *et al.*, 2010) (Table i.5).

Table i.5. Main mechanisms governing phase- and antigenic-variation in mycoplasmas. Adapted from Citti *et al.*, 2010.

Type of variation	Genetic event	Mycoplasma species (gene family)
ON/OFF switching	DNA slippage involving SSR in promoter regions	*M. hyorhinis* (*vlp*) *M. gallisepticum* (*vlhA*)[a] *M. capricolum* subs. *capricolum* (*vmc*)
	Site-specific recombination (gene rearrangement)	*M. pulmonis* (*vsa*) *M. bovis* (*vsp*) *M. agalactiae* (*vpma*)[b] *U. parvum* (*mba*)
	Site-specific recombination (promotor inversion)	*M. penetrans* (*mpl*)
Size variation	DNA slippage involving short direct repeats within CDSs	*M. agalactiae* (*vpma*) *M. hyorhinis* (*vlp*) *M. pulmonis* (*vsa*) *M. bovis* (*vsp*) *U. urealyticum* (*mba*)
	Gene conversion (unidirectional)	*S. synoviae* (*vlhA*)
Domain shuffling	Reciprocal recombination	*M. genitalium* (*mgp*) *M. pneumoniae* (*P1*)
	DNA recombination[c]	*M. agalactiae* (*vpma*) *M. bovis* (*vsp*)
Other	Gene or locus duplication[c]	*M. agalactiae* (*vpma*)

[a] The *vlhA* gene family was previously designated as pMGA.
[b] Some of the *vpma* genes were also designated as *avg*.
[c] Expected to occur at low frequency (rare events).
CDS, coding sequence; SSR, single sequence repeat

The *vlhA* gene family of the avian pathogen *Mycoplasma gallisepticum* contains 33 different proteins coding for surface lipoproteins. Apparently, only one *vlhA* gene is expressed dominantly in a population, while the other gene variants remain dormant. It has been shown that *in* vitro, *M. gallisepticum* can switch the expression to another *vlhA* variant in a reversible fashion through the modification of a (GAA)$_n$ trinucleotide repeat present upstream of the -35 element of *vlhA* promoters (Glew *et al.*, 1998). However, this hypothesis was challenged in an *in vivo* study, where a more complex phase variation regulation of the *vlhA* genes dependent on the infection stage is suggested (Pflaum *et al.*, 2016).

On the other hand, some mycoplasmas rely on site-specific recombinases to rearrange surface lipoproteins and alter their expression. This is the case of the *vpma* family of proteins of *M. agalactiae*. Vpma protein coding genes are clustered in the genome and share highly conserved 5' untranslated regions and a unique active promoter sequence. A site-specific recombinase (*xer1*) is responsible for linking silent *vpma* genes to this active promoter, altering the expression pattern of expressed surface lipoproteins of *M. agalactiae* (Glew *et al.*, 2002; Flitman-Tene *et al.*, 2003). Mutants for *xer1* had similar colonization capabilities than the WT strain, though they showed dissemination and invasion defects (Chopra-Dewasthaly *et al.*, 2012). In the case of *M. penetrans*, a similar strategy is followed to generate antigenic diversity changing *mpl* surface antigens. Promoter inversions induced by a site-specific tyrosine recombinase mediate *mpl* surface lipoprotein expression (Röske *et al.*, 2001; Horino *et al.*, 2003, 2009).

M. hyorhinis can generate Vlp surface lipoprotein variants using two distinct strategies. First, extension or reduction of a polyA tract located between the -35 and -10 elements of *vlp* promoter regions can switch off the expression of *vlp* genes (Citti and Wise, 1995). Moreover, intragenic expansion or contraction of *vlp* genes can occur via recombination of the 3' region containing tandem repeats, generating a large combination of possible *vlp* variants that allow immune evasion (Rosengarten and Wise, 1991; Citti *et al.*, 1997).

Another avian mycoplasma, *M. synoviae*, relies in gene conversion via unidirectional recombination to generate antigenic variation. *M. synoviae* has several pseudogenes homologous to the lipoprotein hemagglutinin (*vlhA*) encoded throughout the genome. By unidirectional recombination, pseudogenes are duplicated and previously expressed variants are lost from the final expressed *vlhA* gene (Noormohammadi *et al.*, 2000). This process has also been observed *in vivo* during infections in chickens (Slavec *et al.*, 2011).

The human pathogens *M. genitalium* and *M. pneumoniae* carry several genomic repeats homologous to the main cytadhesins P1 (*M. pneumoniae*) and *mgpB* and *mgpC* (*M. genitalium*). Upon homologous recombination, several of these repeats can be exchanged with the cytadhesin coding sequence deriving into functional genetic variants in most cases (Kenri *et al.*, 1999; Iverson-Cabral *et al.*, 2007). In *M. genitalium*, this process is tightly regulated by the alternative sigma factor MG428 (this work).

In addition to these sophisticated strategies utilized by the different species, mycoplasmas have also developed an active system to fight antibody recognition from the host immune system. MIB-MIP system (*Mycoplasma* Ig Binding protein – *Mycoplasma* Ig Protease) is a very specialized binomial system that is able to recognize and cleave immunoglobulins (Arfi *et al.*, 2016). The molecular mechanism proposed states that MIB is able to tightly bind to the variable domain (F_v) of many IgGs, then MIP is recruited to the MIB-IgG complex and cleaves off the heavy chain (V_H) of IgG. This system is widespread among Mollicutes, but it is absent in some important species like *M. genitalium*, *M. pneumoniae* or *M. suis* (Figure i.2).
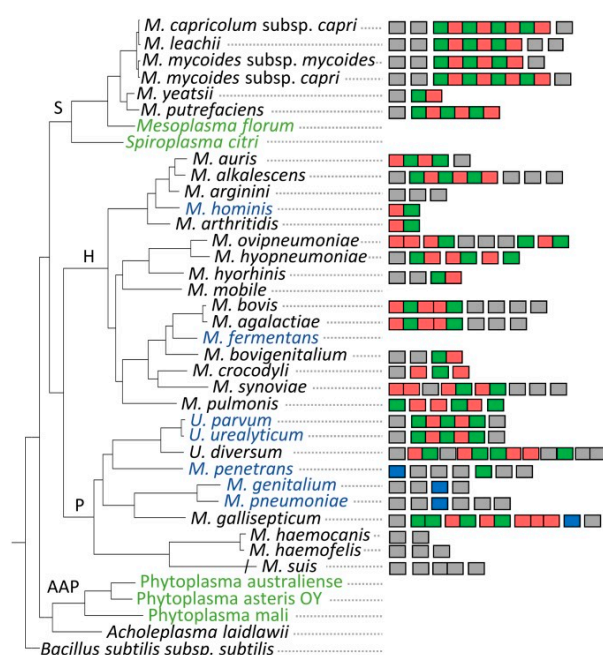
Figure i.2. Distribution of the MIB and MIP homologs in Mollicutes. The occurrence of a gene homolog of MIB, MIP, protein M, or a gene annotated DUF31 is indicated by rectangles colored in red, green, dark blue, and light gray, respectively. Adjoining rectangles denote genes located on the same loci. Vertical alignment of the symbols does not infer synteny. The host specificity of the Mollicutes species is indicated as follows: black, animal host; cyan, human host; green, plant host. Main phylogenetic groups are indicated: S, spiroplasma; H, hominis; P, pneumoniae; and AAP, acholeplasma/ phytoplasma. Extracted from Arfi *et al.*, 2016.

Nevertheless, *M. pneumoniae* and *M. genitalium* possess a MIB-related protein, named protein M, which also binds the variable region of several IgGs (Grover *et al.*, 2014). In this case, protein M is able to block antigen-antibody union, but no complementary cleavage system has been yet identified.

Finally, some mycoplasma species have integrative conjugative elements (ICE) that confer the ability to transfer DNA horizontally among individuals (Calcutt *et al.*, 2002; Dordet Frisoni *et al.*, 2013). Moreover, some mycoplasmas can use this specialized machinery to recombine large DNA sequences and thus increase genomic variability (Dordet-Frisoni *et al.*, 2014). However, the impact of these systems in immune evasion or antigenic variation *in vivo* has yet to be determined.

i.1.6. Isolation of mycoplasmas and culture

Mycoplasmas have intricate nutritional requirements that challenge *in vitro* culture. In fact, most mycoplasmas are isolated using cell-assisted co-cultivation with eukaryotic cells. This is particularly useful for identification and later classification of mycoplasma clinical isolates (Hamasuna *et al.*, 2007). Detailed protocols on mycoplasma recovery, isolation and cultivation strategies can be found at Bradbury, 1998; Nicholas and Baker, 1998 and Taylor, 1998. In the other hand cultivable mycoplasmas grow very slowly on optimal medium, possibly due to their lack of most anabolic pathways. Mycoplasma culture media are usually based on heart infusion, yeast extract, peptone, salts, glucose and serum with several supplements. High concentrations of fetal calf serum or horse serum are essential for mycoplasma growth, as they are the main source of lipids and cholesterol required for membrane synthesis. To prevent the overgrowth of other bacteria that usually accompany mycoplasmas in clinical samples, penicillin derivatives and thallium acetate can be added as selective agents (Razin, 1996). Despite these nutritional requirements, it has been proposed that some mycoplasmas may not grow in rich complex mediums because of the presence of toxic components or inhibitors. Therefore, mycoplasma cultivation is exceptionally fastidious.
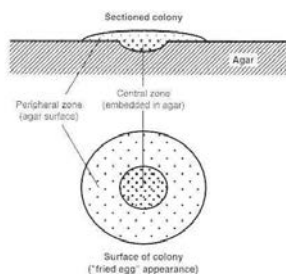
A particular feature of mycoplasmas is that they grow in solid surfaces as fried-egg shaped colonies, due to the embedding of the central zone of the colony into the agar.

Figure i.3. Morphology of a typical fried-egg mycoplasma colony. Extracted from Razin, 1996.

i.1.7. Transformation of mycoplasmas and available genetic tools

Despite their lack of cell wall, transformation of mycoplasmas with exogenous DNA has proven to be problematic, probably due to the high activity of mycoplasma membrane nucleases (Minion *et al.*, 1993) and the presence of R-M systems (Dybvig *et al.*, 1998). First reported transformation of mycoplasmas was achieved via PEG in *M. pulmonis* using the Tn*916* transposon derived from *Enterococcus faecalis* encoding the tetracycline resistance marker *tetM* (Dybvig and Cassell, 1987). Since then, several gene delivery systems have been developed, for instance calcium chloride treatment followed by PEG (King and Dybvig, 1991), liposome-mediated delivery (Minion and Kapke, 1998) or electroporation (Hedreyda

et al., 1993). Transformation protocols vary depending on the mycoplasma species and usually require relatively high amounts of DNA.

The main available tool for genetic studies in mycoplasmas is random mutagenesis by transposition. The first transposon-based plasmids available were derived from Tn*916* (Franke and Clewell, 1981) and Tn*4001* (Lyon *et al.*, 1984), but they were highly unstable once transposed. The generation of a minitransposon derived from Tn*4001* (Pour-El *et al.*, 2002) solved this problem and has become the scaffold to create new minitransposons for general use in mycoplasmas (Pich *et al.*, 2006; Algire *et al.*, 2009).

Although natural plasmids have only been detected in *M. mycoides* (see i.1.1), *Mycoplasma leachii* (Djordjevic *et al.*, 2001) and *Mycoplasma yeatsii* (Kent *et al.*, 2012), several laboratories have developed replicative plasmids using oriC sequences from mycoplasma genomes (Lartigue *et al.*, 2003). This strategy has been applied to transform and maintain extrachromosomal plasmids in *M. pulmonis* (Cordova *et al.*, 2002), *M. capricolum* subs. *capricolum* (Janis *et al.*, 2005)*, M. hyopneumoniae* (Maglennon *et al.*, 2013b; Ishag *et al.*, 2016), *M. gallisepticum* and *M. imitans* (Lee *et al.*, 2008), *M. synoviae* (Shahid *et al.*, 2014), *M. agalactiae* and *M. bovis* (Chopra-Dewasthaly *et al.*, 2005; Sharma *et al.*, 2015). Replicable oriC plasmids have also been developed for *M. genitalium* and *M. pneumoniae*, but with high chromosome integration rates (Torres-Puig *et al.*, unpublished data).

Targeted mutagenesis can only be achieved in a few mycoplasma species, such as *M. genitalium* (Dhandayuthapani *et al.*, 1999) and some strains of *M. pneumoniae* (Krishnakumar *et al.*, 2010). Mutations to targeted genes are inserted by gene replacement in a double crossover event using the recombination machinery of the cell. In some mycoplasmas, the inefficiency of this machinery can be overcome by the addition of heterologous expression of the *recA* gene from *E. coli* to generate site-specific recombinations (Allam *et al.*, 2010; Ishag *et al.*, 2017). Finally, targeted disruption of genes can also be achieved by replicative oriC plasmid curation (Chopra-Dewasthaly *et al.*, 2008; Lee *et al.*, 2008), although insertion of the replicative plasmids at the chromosomal oriC site is highly frequent (Renaudin *et al.*, 1995; Cordova *et al.*, 2002; Chopra-Dewasthaly *et al.*, 2005).

Generation of mutants in mycoplasmas is highly dependent on the use of antibiotic markers. To date, the most successful marker used in mycoplasmas is the *tetM* determinant (Dybvig and Cassell, 1987), coding for a ribosomal protection protein that reduces the susceptibility

of ribosomes to the action of tetracyclines (Chopra and Roberts, 2001). *aac(6')-aph(2")* marker, which confers resistance to aminoglycosides (Dowding, 1977), is also widely used (Mahairas and Minion, 1989), although its utilization has some detrimental effects in mutant cell growth in certain mycoplasmas (Pich *et al.*, 2006b). Despite some mycoplasmas have certain natural resistance to chloramphenicol, *cat* gene coding for a type A-9 chloramphenicol acetyl transferase (Horinouchi and Weisblum, 1982) can also be used for mutant selection (Hahn *et al.*, 1999; Calisto *et al.*, 2012). Puromycin acetyl transferase from *Streptomyces alboniger* (Pérez-González *et al.*, 1983) is the newest antibiotic marker that has been adapted for use in mycoplasmas (Algire *et al.*, 2009). Some laboratories have also adapted spontaneous mutations at DNA gyrases or topoisomerases to generate fluoroquinolone-resistant mycoplasma strains and use antibiotics such as enrofloxacin for mutant selection (Dordet-Frisoni *et al.*, 2014).

The tetracycline-inducible promoter P*xyl/tetO$_2$* has been successfully tested in *S. citri* and *M. agalactiae* inducing gene expression with sub-lethal concentrations of tetracycline in these organisms (Breton *et al.*, 2010). Using this tool, the Cre-lox technology was effectively established in *M. genitalium* using tetracycline-inducible promoters to tightly control Cre recombinase expression (Mariscal *et al.*, 2016). Cre-lox technology can allow multiple rounds of genome editing removing antibiotic markers.

The recent discovery of the CRISPR/Cas system in prokaryotes and archaea has provided a quantum leap forward in terms of genome editing and control of gene expression (Ledford, 2015). Most mycoplasmas lack the non-homologous end joining (NHEJ) repair machinery needed for genome editing using CRISPR/Cas9 technology, but this system can still be useful in gene knockdown strategies. Adaptation of the CRISPR/Cas9 technology for these purposes is currently being developed in several mycoplasma strains (Mariscal *et al.*, manuscript in preparation).

i.2. *MYCOPLASMA GENITALIUM*, AN INFECTIOUS MINIMAL GENOME BACTERIUM

Mycoplasma genitalium was first isolated in 1980 from two men with non-gonococcal urethritis (Tully *et al.*, 1981). In 1983, one of the two strains isolated (G37) was designated as the type strain of this mycoplasma species (Tully *et al.*, 1983). *M. genitalium* colonizes the genital and urinary tracts of humans and most infections are asymptomatic (Uno *et al.*, 1997; Henning *et al.*, 2014). Untreated infections though, have shown to produce urethritis in both men and women, cervicitis, pelvic inflammatory diseases, infertility, salpingitis and endometritis in women (Manhart, 2013). Although *M. genitalium* targets the urogenital tract, infections are not involved in bacterial vaginosis and have little impact on adverse pregnancy outcomes (Oakeshott *et al.*, 2004). Despite extra-genital infections are rare, it has been reported in knee joints and it is envisioned as a plausible cause of arthritis (Taylor-Robinson *et al.*, 1994). Prevalence of *M. genitalium* infections is highly variable because no convenient test methods were available until PCR detection was developed (Palmer *et al.*, 1991), but rates oscillate between 2% and 7% in the population (McGowin *et al.*, 2011). Prevalence increases among population with other sexually transmitted diseases (Daley *et al.*, 2014).
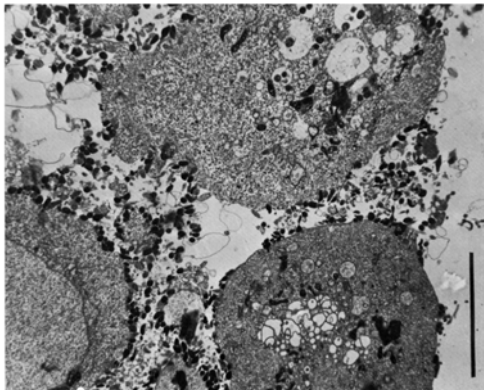
Figure i.4. Electron micrograph of a section of a Vero cell culture infected with *M. genitalium* for 5 days. Note the small dark mycoplasmas and the stages of cell destruction. Scale bar is 5µm. Extracted from Tully *et al.*, 1983.

The evidence that *M. genitalium* is a sexually transmitted pathogen is unquestionable (Ross and Jensen, 2006). However, the lack of large clinical studies and detection protocols have often underestimated the impact of *M. genitalium* infections (see Table i.6). Mycoplasma cells could be carried by motile sperm during infection, as it has been demonstrated that they can attach to human spermatozoa (Svenstrup *et al.*, 2003). Some studies state that in ~58% of cases, both partners in a heterosexual couple were tested positive for *M. genitalium* (Keane *et al.*, 2000) and often a concordance of genotypes is detected in both

partners (Hjorth *et al.*, 2006). Moreover, another study from the USA concluded that *M. genitalium* infections were strongly associated with sexual activities and behaviors, and that it was more prevalent than *Neisseria gonorrhoeae*, the causative agent of gonorrhea (Manhart *et al.*, 2007). Among homosexual men, infections of *M. genitalium* are less prevalent and often asymptomatic (Taylor-Robinson *et al.*, 2003; Gottesman *et al.*, 2017). However, asymptomatic rectal infections might serve as important reservoirs for *M. genitalium* spreading to other individuals in rectal-to-urethral transmission (Edlund *et al.*, 2012).

Table i.6. Comparison of assay methods for detection of *M. genitalium*.

Assay	Target of assay	Detection limit	Assay time	Developed since
Multiple steps: including cell culture & radiolabeled probes	256bp of an unknown DNA region	$\sim 1.2 \times 10^4$ organisms	Several weeks	1988
Multiple steps: including cell culture & PCR	MgPa	$>10^6$ dilution	~20 days	1996
PCR	16S rRNA	<50 to >6 gene copies	4h	2002
PCR	MgPa	<23 or >5 gene copies	60min	2004
LCRT-PCR	MG_301 (G3PDH)	<2.5 gene copies	50min	2005
PCR	16S rRNA	0.5-5 CFU	90min	2006
PCR	MgPa	<1 CFU	90min	2006
Multiplex PCR reverse line blot	MgPa	2×10^{-8}ng DNA	4h	2009
PCR	MG_219	0.5pg DNA or 825 gene copies	60min	2009

LCRT-PCR, LightCycler quantitative real-time PCR

Adapted from Daley et al., 2014

Whether *M. genitalium* is linked to HIV acquisition and transmission is a matter of debate. In 1990, *M. genitalium* was detected in blood of an HIV-positive patient (Montagnier *et al.*, 1990). A few years later, it was demonstrated that adherence of mycoplasmas to HIV-infected cells could trigger viral release and thus accelerate HIV progression (Phillips *et al.*, 1992; Sasaki *et al.*, 1993; Chowdhury *et al.*, 1994). More recently, a study concluded that *M. genitalium* could promote HIV-1 crossing the epithelial cell barrier across endocervical cells by reducing membrane integrity of epithelial cells and by activating HIV-infected cells (Das *et al.*, 2014). Despite this, it is unclear wether mycoplasmas support HIV-1 infections or HIV-infected cells induce changes in cell membranes that favour mycoplasma adherence and/or invasion. Taking together that both HIV and *M. genitalium* are sexually transmitted infections and that HIV-1 induced immunodefficiency could favor *M. genitalium* infections,

it is difficult to state wether *M. genitalium* is a risk factor for HIV acquisition (Napierala and Weiss, 2009; Manhart, 2012).

Similarly, although malignant transformation in long term mycoplasma infected cell cultures seems clear (Zhang *et al.*, 1998, 2004; Feng *et al.*, 1999), few studies have assessed the association between *M. genitalium* and cancer. Some studies carried out by Idahl and colleagues suggest a possible association between ovarian cancer and infection by *M. genitalium* (Idahl *et al.*, 2011). In addition, an association between urogenital mycoplasmas and human papiloma virus (HPV) infections has been shown, although the exact role of *M. genitalium* in these infections is not yet clear (Biernat-Sudolska *et al.*, 2011). Finally, *in vitro* and *in vivo* assays with *M. genitalium* have proven a role for this pathogen in malignancy switch and tumor progression of bening human prostate cells. However, so far, the ability of *M. genitalium* to influence genetically unmodified human cell lines remains unreported (Namiki *et al.*, 2009). Overall, the role of *M. genitalium* in cancer development is still speculative (Zarei *et al.*, 2013).

i.2.1. *M. genitalium* as a systems biology model

Since the isolation of *M. genitalium* in 1980, its minimal genome has awakened the curiosity as to how many genes and cell functions are necessary to keep a cell alive. Despite the discovery of bacteria with smaller genomes (Nakabachi *et al.*, 2006; Perez-Brocal *et al.*, 2006), *M. genitalium* remains to be the natural bacterium with the smallest known genome (580kb) of any free-living bacteria. In 1995, it was the second genome the be fully sequenced (Fraser *et al.*, 1995) and a first glance in gene essentiality was published only four years later (Hutchison *et al.*, 1999). To date, the most complete study on gene essentiality in *M. genitalium* using global transposon mutagenesis was published by Glass and colleagues (Glass *et al.*, 2006). The study revealed that at least 100 genes were dispensable for cell growth *in vitro* (Figure i.5). Since then, more genes have either been found interrupted by transposons (Lluch-Senar *et al.*, 2007) or replaced by allelic exchange (Dhandayuthapani *et al.*, 1999; Lluch-Senar *et al.*, 2010; Torres-Puig *et al.*, manuscript in preparation; Martínez-Torró *et al*, manuscript in preparation). Moreover, a recent study on gene essentiality conducted on *M. pneumoniae* (Lluch-Senar *et al.*, 2015), a closely related mycoplasma species, stated that the core number of essential genes could even be slightly reduced.

Figure i.5. Global transposition mutagenesis of _M. genitalium_. The location of transposon insertions from Glass _et al._, 2006 study is noted by a △ below the insertion site on the map. Insertions mapped in Hutchison _et al._, 1999 are noted with a ▽. Sites with 10 or more insertions are noted by a red filled triangle. Extracted from Glass _et al._, 2006

Extensive studies on _M. genitalium_ finally conducted to the construction of _Mycoplasma genitalium_ JCVI-1.0: the first synthetic genome built using BACs in _E. coli_ and assembled in YACs in _Saccharomyces cerevisiae_ (Gibson _et al._, 2008). Despite achieving this milestone, this artificial genome could not be transplanted to a new cell as initially planned. Years later, genome synthesis and transplantation were achieved using artificially-made _M. mycoides_ chromosome in a DNA-free _M. capricolum_ cell chassis. The resulted artificial bacterium, named _Mycoplasma mycoides_ JCVI-syn1.0, was capable of self-replication (Gibson _et al._, 2010). Later, a new minimal synthetic bacterium (_Mycoplasma mycoides_ JCVI-syn3.0) was created, containing only 473 genes (Hutchison _et al._, 2016) and capable of self-replication. This minimal genome was created using _M. mycoides_ JCVI-syn1.0 as a template instead of the reduced genome of _M. genitalium_, probably due to the slow doubling time of _M. genitalium_ as compared to _M. mycoides_ (8.5h vs. 180min) (Pennisi E., 2010).

Despite this breakthrough, it is important to state that minimal genomes are environment-dependent, a genome set is only one of many components of living systems (Coyle *et al.*, 2016).

Despite these changes in the model microorganism chosen for synthetic biology, *M. genitalium* and *M. pneumoniae* remain two of the most characterized mycoplasmas and they have become model organisms for systems biology (Balish, 2014). Several large studies have been conducted to better understand these minimal cells in the fields of transcriptomics (Weiner *et al.*, 2000; Güell *et al.*, 2009; Maier *et al.*, 2011; Yus *et al.*, 2012), proteomics (Kuhner *et al.*, 2009; Schmidl *et al.*, 2010; Párraga-Niño *et al.*, 2012; van Noort *et al.*, 2012), metabolomics (Yus *et al.*, 2009; Maier *et al.*, 2013; Wodke *et al.*, 2013) and genetic methylome (Lluch-Senar *et al.*, 2013). Integration of all these data permitted the generation of a whole-cell computational model that aimed at providing accurate predictions on *M. genitalium* cell behavior and biological functions (Karr *et al.*, 2012). However, the fact that there are still more than 100 genes of *M. genitalium* with unknown function enhances the importance of single-gene characterization studies in this organism.

i.2.2. *M. genitalium* as a singularly-shaped bacterium

M. genitalium, as well as other mycoplasmas, has the ability to move over solid surfaces using a characteristic sliding movement known as gliding (Miyata, 2010; García-Morales *et al.*, 2016; Miyata and Hamaguchi, 2016). A differenciated polar structure known as terminal organelle carries the entire gliding machinery required for movement as well as the proteins necessary for adherence to the host cells(Tully *et al.*, 1983; Jensen *et al.*, 1994; Krause and Balish, 2001; Svenstrup *et al.*, 2002; Hasselbring and Krause, 2007). Thus, this tip structure is of vital importance for many biological functions in these bacteria.

Terminal organelle of *M. genitalium* and *M. pneumoniae* is supported by a complex cytoskeleton formed by several proteins arranged in separate macromolecular structures. Electron cryotomography studies have determined that the terminal organelle can be divided in four distinct regions: the nap, the terminal button or knob, the rod or segmented plates and the wheel-like complex (see Figure i.6). Clusters of P140 and P110, the main adhesins of *M. genitalium*, compose the nap structure. Both proteins are required for the terminal organelle formation and regulate the generation of new terminal organelles in this bacterium (Burgos *et al.*, 2006; Pich *et al.*, 2009). Adhesion of *M. genitalium* to eukaryotic cells can be monitored by qualitative methods (Miyata *et al.*, 2000; Pich *et al.*, 2008) or quantified using flow cytometry (García-Morales *et al.*, 2014).
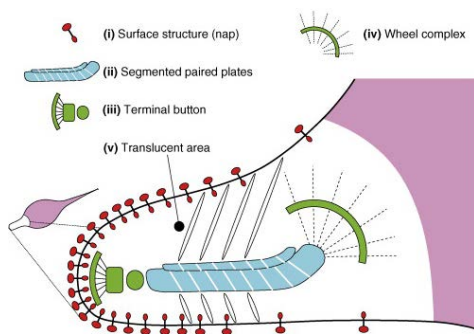
Figure i.6. Architecture of the attachment organelle of *M. pneumoniae*. An outline of the entire cell is presented on the left-hand side, alongside an enlargement of the attachment organelle. The cytoplasm has a normal electron density and is colored magenta. Extracted from Miyata, 2008.

The terminal button is located at the distal end of the tip structure of *M. genitalium* (see Figure i.6). At least three proteins localize to this structure: MG318 (P32) (Relich and Balish, 2011), MG317 (HMW3) (Seto *et al.*, 2001) and MG217 (P65) (Seto *et al.*, 2001; Burgos *et al.*, 2008). This structure has been related to the curvature of the terminal organelle, the direction of movement and nap anchoring to the rod structure (Burgos *et al.*, 2008; Pich *et al.*, 2008).

The rod, or segmented paired plates, is located at the central part of the electron-dense core of the terminal organelle (see Figure i.6). That is between the terminal button and the wheel-complex. The subset of proteins forming this structure are MG218 (HMW2) (Balish *et al.*, 2003; Pich *et al.*, 2008), MG312 (HMW1) (Burgos *et al.*, 2007) and MG527, which represents a small peptide corresponding to the C-terminal part of MG218 (Boonmee *et al.*, 2009; Broto *et al.*, unpublished). This is a key structure for correct terminal organelle development as it prevents the proteolysis of the other cytoskeleton proteins (Krause and Balish, 2004; Pich *et al.*, 2008).

Finally, the wheel-like complex is the most distal structure of the terminal organelle (see Figure i.6). It is composed by MG386 (Pich *et al.*, 2006a), MG200 (TopJ) (Calisto *et al.*, 2012), MG491 (P41) (Martinelli *et al.*, 2015, 2016) and MG219 (P24-like) (González-González, unpublished) proteins. These proteins are involved in adhesin folding and trafficking to the membrane as well as in anchoring the organelle to the cell body (Hasselbring and Krause, 2007; Cloward and Krause, 2011).

Terminal organelle duplication is synchronized with cell division (Hasselbring *et al.*, 2006) Upon duplication, the new terminal organelle migrates to the opposite cell pole (Figure i.7A and 7B). During this process, cells are thought to be motile only through the old organelle and this motile force actually propels the new terminal organelle in the direction opposite to the movement (Hasselbring *et al.*, 2006). It has been suggested that chromosome

replication occurs in a coordinated fashion with terminal organelle duplication and migration (Seto *et al.*, 2001). However, mutants with an increased frequency of terminal organelle duplication do not seem to initiate DNA replication at higher rates (Pich *et al.*, 2009). Once the new terminal organelle reaches the opposite pole, it gains motility and starts pulling the daughter cell until cytokinesis ends (see Figure i.7C) (Hasselbring *et al.*, 2006). Although it is possible that the FtsZ protein participates in the final steps of cytokinesis, the *ftsZ* gene has shown to be non-essential for *in vitro* growth in *M. genitalium* and *M. pneumoniae* (Lluch-Senar *et al.*, 2010, 2015).
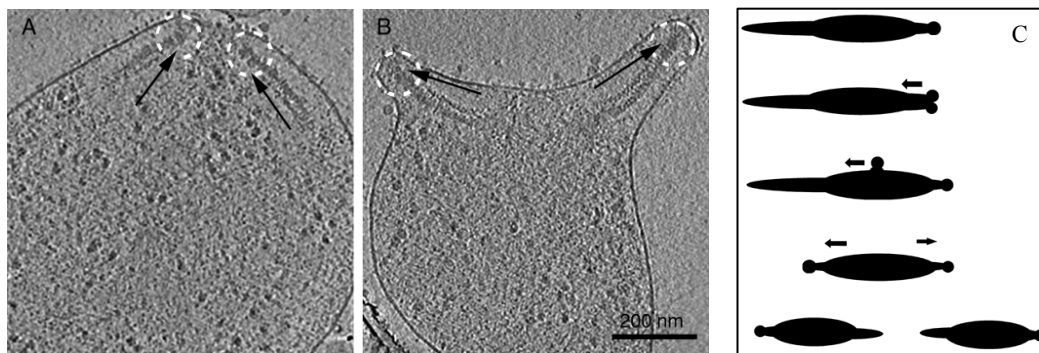
Figure i.7. Probable sequence of events in *M. pneumoniae* cell division. Cells with multiple electron-dense cores are shown where the terminal buttons (highlighted with dotted circles) are oriented either towards **(A)** or away **(B)** from each other. The black arrows point from the bowl complex to the terminal button alongside the thin rods. **(C)** Duplication of the terminal organelle is followed by movement of the original or nascent organelle to the opposite end of the cell and cytokinesis. This model is also valid for *M. genitalium*. Extracted from Henderson and Jensen, 2006 (A and B) and Krause and Balish, 2001 (C).

i.2.3. Gene regulation complexity in *M. genitalium*

i.2.3.1. Transcription response to stress

Due to its reduced genome size, for a long time it was thought that mycoplasmas did not have coordinated transcription response to environmental stimuli or stress (Razin *et al.*, 1998). However, years later these impressions were challenged with the transcriptional profiling of mycoplasmas subjected to different stresses (Weiner *et al.*, 2003; Madsen *et al.*, 2006a, 2006b; Schafer *et al.*, 2007) or during infection (Cecchini *et al.*, 2007; Madsen *et al.*, 2008). *M. genitalium* and close mycoplasmas respond to heat shock stress using the transient induction of the heat shock proteins (Hsp), which can protect cells from the accumulation of misfolded proteins. Generally, this classic response is activated by alternative sigma factors (*rpo*H, σ^{32}) (Grossman *et al.*, 1984) or transcription factors, such as *hrc*A (Schulz and Schumann, 1996) and its associated CIRCE (controlling inverted repeat of chaperon expression) element. In *M. genitalium*, CIRCE elements have been

detected upstream of the chaperone *dnaK* (MG_305), and the proteases *clpB* and *lon* (MG_355 and MG_239, respectively) (Figure i.8), although additional Hsp genes are present in the chromosome (Musatovova *et al.*, 2006).


```
                          TTAGCACTCNNNNNNNNNNGAGTGCTAA
dnaK    AGATAAATAAAACAAACTTTAATTAGCACTTTTAGTGTTTGAGTGCTAATTTTTGTTTATAATTACTG*AGTAATTGTTATTGATCTAAAAAAAAAGTATG
clpB    AAACAAAAAACCTAGTTTTTAATTAGCACTCGAATGCCTTGAGTGCTAATTTTATATAATTTTTGC*ATTAAAAGGAGTGAGCAATTAAAAATG
lon     AAATCGTTAGCTATTCAACTCATTAGCACTCAAAGCTTGTGAGTGCTAAGAAATGTGTTAAAATTTATTG*AAATTCCCTAATTAACTTTTAAATATG

hrcA    AGGTGCAGTTTTTAACTTTTGTTTTAGATTAGTGGTCAAGCTGGTTATTGTTAAAAGATAAAATTAGATA*ACTAATG

dnaJ2   TTTCTTTAAATCTCTTTCACATGAAGAACTCTTAAAGCTTTTTGAATAAAGCAAGAATTATAATTAACACTC*TAAGGATGCAATGTTAAATG
dnaJ3   ACTGCCAATTAAATAAAAATCAGCTGGTTTGTAATGTTGCAACTATCTAGTTAATTAATAAAATTATACTTA*CTCTGATTGATTCTTAATG
groES   CTAATATCAGATAAATTTAAGATAACTTTTACTTTCTTTTAAATTAGTTTTTAATTATTAAAATATTTTTA*TTATG
tig     TAAACTAATTGTTGATACACTTTATGAACAGTTAGATAGAGCAATTGGTGAGGAAAATTAAAATTAAGTTA*GCACTAGTAGATACAAAAGATG
ftsH    TATTAACACTTATTAGTTTCATTACCATTGGCAGTATAATTACAGT*T*AATGTAGTTCAAAACCAAAAAAACCCAAAATCGGTCAAAAACTAACATAGATG
```

Figure i.8. Promoter regions and transcriptional start sites of *M. genitalium* heat shock genes. Asterisks designate transcriptional start sites (bold letters) determined experimentally. Shaded areas indicate putative -10 regions. The bold ATG in each sequence indicates the putative translational start site. Underlined sequences represent CIRCE elements are shown in bold. The dashed underline indicates a CIRCE-like sequence upstream of *dnaJ2* (MG_019), and only conserved nucleotides of the IR are bolded. The CIRCE sequence is indicated above the *dnaK* promoter region. Adapted from Musatovova, Dhandayuthapani and Baseman, 2006.

Despite the existence of this classic heat shock response, *hrcA*-independent activation of transcription during heat-shock has been described in *M. genitalium* and other related mycoplasmas (Musatovova *et al.*, 2006; Gorbachev *et al.*, 2013; Mazin *et al.*, 2014), suggesting the presence of a more complex regulatory network.

Cellular response to osmotic stress has also been addressed in *M. genitalium*. Major transcriptional changes were observed upon high osmolality, mostly on lipoprotein expression and other membrane-related proteins (Zhang and Baseman, 2011b). Although no transcriptional regulator responsible for these changes could be identified, osmoinduction was proven to be activated by changes in DNA supercoiling (Zhang and Baseman, 2011a).

Despite that *M. genitalium* has several proteins to face oxidative stress (Dhandayuthapani *et al.*, 2001; Saikolappan *et al.*, 2009; Zhang and Baseman, 2014) no transcriptional response to oxidizing reagents has been identified. The development of high throughput technologies such as tiling arrays or RNAseq has increased the number of transcriptomic studies in mycoplasmas. These studies suggest major transcriptional changes that cannot be explained only with the use of the limited amount of transcription factors present in these small bacteria (Güell *et al.*, 2009; Mazin *et al.*, 2014; Semashko *et al.*, 2017).

i.2.3.2. Putative transcriptional regulators of *M. genitalium*

In the first bioinformatic analyses of the *M. genitalium* genome (Fraser *et al.*, 1995), researchers pointed out the presence of few transcriptional regulators as compared to other sequenced organisms such as *Haemophilus influenzae*. As more genomes are sequenced and more data on transcriptional regulators and protein domains are available, different hypothetical proteins have been annotated as putative transcription factors of *M. genitalium*. However, very few have been studied and their function can only be hypothesized assuming similar functions to other well-characterized bacteria. To date, 11 proteins from *M. genitalium* are thought to have a potential role in gene regulation (see Table i.7), although some studies in *M. pneumoniae* suggest that more proteins might be implicated in gene regulation (Trussart *et al.*, 2017).

RpoE or sigmaE

RpoE or sigmaE is a conserved protein among Firmicutes. It has been considered an integral part of the RNA polymerase (RNAP) complex, but its exact role is not well understood despite a history of more than 40 years of research (Weiss and Shaw, 2015). It seems clear that RpoE interacts with the other components of the RNAP and it is necessary for promoter recognition (Hyde *et al.*, 1986). However, the presence of a helix-turn-helix motif (HARE-HTH) at the N-terminus of this protein (Aravind and Iyer, 2012) suggests a putative role in DNA-binding and specific promoter recognition, analogously to an alternative sigma factor (Weiss and Shaw, 2015). It is thought that one of the main roles of this protein is to increase RNAP recycling (Juang and Helmann, 1994) and transcriptional specificity to certain promoter sequences (Achberger and Whiteley, 1981). Despite being dispensable in many bacteria such as *Bacillus subtilis* or *Staphyloccocus areus*, it is essential for growth in *M. pneumoniae* and *M. genitalium*.

WhiA

WhiA is a poorly characterized transcription factor present in some Gram-positive bacteria. It has been studied in *Streptomyces coelicolor* as a transcriptional activator of sporulation determinants, as well as an inhibitor of its own expression at certain protein levels (Kaiser and Stoddard, 2011). As mycoplasmas do not carry out sporulation, the role of this transcription factor remains unclear. Nevertheless, it is dispensable for growth in both *M. genitalium* and *M. pneumoniae*. Experiments performed in our lab reveal a role for this WhiA-like protein of *M. genitalium* in the regulation of transcription of genes coding for ribosomal proteins (Torres-Puig *et al.*, unpublished data).

SpxA

The Spx protein is an anti-RpoA factor that has shown to activate transcription of genes encoding products related to thiol homeostasis in *B. subtilis* (Nakano *et al.*, 2003; Newberry *et al.*, 2005). Moreover, it induces transcriptional repression of growth-related cellular functions in order to cope with the periods of disulfide stress (Nakano *et al.*, 2003). As in the case of *rpoE*, the *spx* gene is essential for growth in both *M. pneumoniae* and *M. genitalium*.

YbaB or Nucleoid-associated protein

YbaB protein belongs to a nucleoid-associated family of proteins. The homolog in *H. influenzae* has been crystalized (Lim *et al.*, 2002) and its structure revealed that it can form homodimers as a pair of tweezers that induce bending in DNA (Riley *et al.*, 2009). YbaB of *Borrelia burgdorferi* globally controls transcription of approximately 4.5% of borrelial genes (Jutras *et al.*, 2012). Despite being essential for growth in *B. burgdorferi*, *ybaB* gene is dispensable in *M. genitalium* and *M. pneumoniae*.

HrcA

Heat shock response regulator HrcA is the best characterized transcription factor of *M. genitalium* (Musatovova *et al.*, 2006). It is essential for growth in *M. pneumoniae* and it is also presumed essential in *M. genitalium*. See Transcriptional response to stress in this text for more information (i.2.3.1).

MraZ

mraZ is the first gene of the *dcw* (division and cell wall) cluster in *M. genitalium*. Recently, MraZ has been defined as a transcriptional regulator in *E. coli* (Eraso *et al.*, 2014) as well as in *M. gallisepticum* (Fisunov *et al.*, 2016). In these bacteria, it regulates the expression of the genes in the *dcw* cluster. Work is in progress to determine the role of this protein in *M. genitalium* (Martínez-Torró *et al.*, manuscript in preparation).

MG236 or Fur-family protein

MG_236 gene codes for a protein of the ferric uptake regulator (Fur)-family. This family of regulators are essential for metal acquisition and control of oxidative stress in the cells. MG236 homolog in *M. pneumoniae* (MPN329) has been overexpressed and changes in 8 genes were detected, most of them hypothetical proteins (Güell *et al.*, 2009). Transcriptional analysis of a Fur mutant of *M. genitalium* confirmed the predicted role of this protein in metal homeostasis (Torres-Puig *et al.,* manuscript in preparation).

YebC or PmpR-family protein

Little is known about this putative regulator. It has only been studied in *Pseudomonas aeruginosa* and it has a role in quorum sensing (Liang *et al.*, 2008). It has been associated with Holliday junction resolution in an *in silico* study (Zhang *et al.*, 2012), but no experimental data is available to support this hypothesis. It is essential for *in vitro* growth in *M. genitalium*.

HimA-family or HU-like protein

MG_353 codes for a small oligopeptide related to the histone-like proteins. This family of proteins play a role in chromosome compaction (Le Hégarat *et al.*, 1993) as well as in gene regulation (Li *et al.*, 2004; Suzuki-Minakuchi *et al.*, 2015). In a global transposon mutagenesis study in *M. pneumoniae*, gene coding for HU-like protein was disrupted but mutants were classified as fitness-deficient for growth (Lluch-Senar *et al.*, 2015).

MG428 or LuxR-family protein

MG428 was previously annotated as LuxR bacterial regulatory protein. In this work, it is demonstrated that it is actually an alternative sigma factor that regulates activation of recombination as well as other unknown cellular processes.

Table i.7. Summary of putative transcription factors present in *M. genitalium*

Gene	Annotation	Essentiality[a]	Function
MG_022	*rpoE*	E	Unknown, related to transcriptional specificity in other bacteria, as well as many cellular processes
MG_103	*whiA*	NE[b]	Unknown, associated with cell division and sporulation regulation in other bacteria
MG_127	*spxA*	E	Unknown, anti-RpoA factor associated to oxidative stress response in other bacteria
MG_134	*ybaB/ebfC*	NE[b]	Unknown, associated with DNA conformation in other bacteria
MG_205	*hrcA*	E	Heat shock response
MG_221	*mraZ*	NE[b,c]	Regulation of the cell division operon
MG_236	*fur*	NE[b]	Unknown, *fur*-family related regulation
MG_249	*sigA*	E	Primary sigma-70 factor, general transcription
MG_332	*yebC*	E	Unknown, associated to quorum sensing/Holliday junction resolution in other bacteria
MG_353	*himA*	F	Unknown, associated with DNA architecture maintenance in other bacteria
MG_428	*sig20*	NE[d]	Alternative sigma factor, see chapter I and II

[a] Essentiality as determined in *M. pneumoniae* homologues (Lluch-Senar *et al.*, 2015), detected in global mutagenesis studies (Glass *et al.*, 2006) and/or generated in the lab. E, essential; NE, non-essential; F, fitness (as defined in Lluch-Senar *et al.*, 2015)
[b] Mutants of these genes have been created by allelic exchange in the lab.
[c] Martínez-Torró *et al.*, manuscript in preparation.
[d] This work.

i.2.3.3. Initiation of transcription and translation in a minimal cell

For a long time, prediction of transcription promoters has been based on available data in *E. coli*. However, A+T rich mycoplasma sequences cloned in *E. coli* result in accidental promotion of transcription (Dhandayuthapani *et al.*, 1998), probably due to the lack of -35 elements in mycoplasma promoter sequences (Weiner *et al.*, 2000). Classic Pribnow box TATAAT consensus sequence can be identified upstream of most identified transcription start sites, but no conservation of sequences immediately 5' can be observed (Weiner *et al.*, 2000). This apparent simplicity of promoter structure and the lack of a major network of transcription factors is unusual in the bacterial world and suggests a complex transcription initiation process. High throughput analyses of the transcriptome have increased the understanding of transcription initiation in mycoplasmas (Güell *et al.*, 2009; Mazin *et al.*, 2014; Lloréns-Rico *et al.*, 2015). Compiling all available transcriptomic data, algorithms have been created to predict the presence of promoter sequences in *M. pneumoniae* (Lloréns-Rico *et al.*, 2015). Pribnow box has been redefined to TANAAT consensus sequence and no other upstream sequence has been identified for promoter prediction, despite the presence of a putative degenerated -35 sequence (TTGANN) in some promoter regions (Lloréns-Rico *et al.*, 2015).

Extreme genome size reduction has led to promoter and 5'untranslated region length optimization in mycoplasmas. A significant amount of *M. genitalium* and *M. pneumoniae* transcripts are leaderless, meaning that transcription and translation start sites are featured in the AUG codon (Weiner *et al.*, 2000). In consequence, most transcribed genes completely lack well-defined Shine-Delgarno sequences or even some genes might bear novel undefined translation initiation sequences, such as the *tuf* gene of *M. genitalium* (Loechel *et al.*, 1991). Some studies in *E. coli* suggest that, although they are translated poorly, efficient translation of leaderless transcripts occurs when the first base of the transcript is the adenosine of the AUG start codon (van Etten and Janssen, 1998), a very useful option for bacterial species with minimal genomes.

i.2.3.4. Transcription termination, anti-termination and organization of supra-operons

M. genitalium has a relatively large subset of proteins to control transcription termination and anti-termination (Fraser *et al.*, 1995). N-utilization substance (Nus) A, B and G and transcription elongation factor GreA are present in the *M. genitalium* genome, as well as ribosomal protein S10 that serves as NusE antitermination factor (Friedman *et al.*, 1981).

However, no rho termination factor has been found, suggesting that transcription termination is dependent on intrinsic terminator sequences (Lesnik *et al.*, 2001). This hypothesis has been confirmed in other mycoplasmas (Fritsch *et al.*, 2015) and its impact on transcriptional regulation has also been addressed (Mazin *et al.*, 2014).

Coordination of transcription is also based on co-expression of supra-operonic units and transcription attenuation, as stated by Junier *et al.*, 2016. Transcriptional read-through among different operons and DNA supercoiling could transmit complex regulatory signals that modulate transcription in bacteria with a limited transcription factor toolbox (Zhang and Baseman, 2011a; Junier *et al.*, 2016). Moreover, a recent study provides evidence that chromosome architecture and localization in the cell can influence transcriptional regulation (Trussart *et al.*, 2017), adding another layer of transcriptomic complexity in these small pathogens.

i.2.4. Minimal recombination machinery and DNA repair systems of *M. genitalium*

Despite the diversity of DNA repair mechanisms described in bacteria, proteins related to repair systems in mycoplasmas are scarce (Carvalho *et al.*, 2005). *M. genitalium* lacks most proteins involved in direct DNA repair such as alkyltransferases or photolyases. Also, the SOS response-trigger protein LexA is missing, as well as SOS box sequences (Fraser *et al.*, 1995). These facts suggest that *M. genitalium* cannot respond to major DNA insults and relies on the basal expression of the DNA repair systems to face them.

Mismatch repair system, composed by MutS, MutL and MutH proteins in other bacteria, is also missing in *M. genitalium*. Lack of a proper mismatch repair system can increase spontaneous mutagenesis induced by replication errors, contributing in this way to mycoplasma genetic variability (Razin *et al.*, 1998).

M. genitalium can correct cross-links and pyrimidine dimers induced by photochemical reactions (Burgos *et al.*, 2012). Genes coding for the UvrABC complex (MG_421, MG_073 and MG_206) are present in *M. genitalium*, which can recognize DNA damage and cleave off the lesion in the form of an oligonucleotide. Damaged DNA is excised by 3' to 5' helicase PcrA (MG_244), an enzyme replacing UvrD in Gram-positive bacteria (Dillingham *et al.*, 2001). Therefore, the complete nucleotide excision repair (NER) system is present in *M. genitalium*.

Components of the base excision repair (BER) system in *M. genitalium* are also encoded in its chromosome. Deamination and base oxidation are repaired using uracil-DNA glycosilase (Ung) (MG_097) and formamydopyrimidine-DNA-glycosylase (MutM) (MG_498), and damaged bases excised with apurinic endonuclease IV (Nfo) (MG_235) (Vink, 2014). It has been suggested that the low G+C content in mycoplasma genomes might be enhanced by a low uracil-DNA glycosylase activity from Ung protein (Williams and Pollack, 1990). Final steps in both NER and BER repair pathways involve the activity of translesion DNA polymerases and ligases to finally generate a corrected DNA strand. DNA polymerase IV from *M. genitalium* (MG_360) could be involved in correction of chromosomal errors. Furthermore, despite *M. genitalium* lacks DNA polymerase I, a translesion DNA synthesis polymerase; a protein with 5'-3' exonuclease activity similar to DNA polymerase I (Pol-I-like) (Carvalho *et al.*, 2005) is encoded by the MG_262 gene. Nonetheless, its role in DNA repair has not been addressed yet.

One last mechanism to overcome DNA lesions is homologous recombination (HR). This mechanism is one of the most complex repair pathways in bacteria and involves more than 50 enzymes and accessory proteins (Lusetti and Cox, 2002). *M. genitalium* has an efficient rudimental machinery to perform HR between homologous strands (see Figure i.9), although it has not been studied in depth. Proteins involved in double strand break (DSB) or DNA lesion recognition are not known, but damaged DNA could be sensed by a DisA homolog (MG_105) present in *M. genitalium*, which can act as a bacterial checkpoint protein (Witte *et al.*, 2008).

Figure i.9. Proposed DNA repair mechanism by HR in *M. genitalium*. Proteins that could be involved in each step are marked in green. Question marks denote that these steps could be performed by these proteins or other unknown enzymes. Resolution of Holliday junctions can result in single or double DNA crossovers. Single crossovers are highly unstable and often result in deleterious phenotypes.
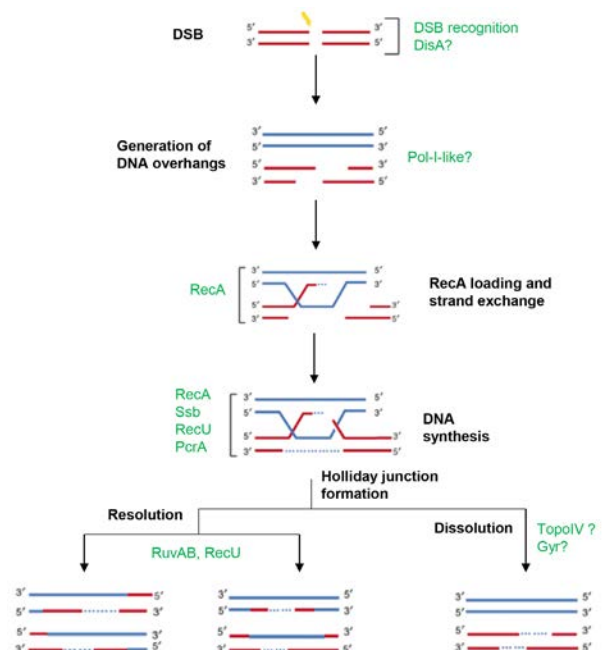
In other Gram-positive bacteria, DSB are recognized by RecN, which is absent in *M. genitalium*. After recognition, 5'-ends surrounding the lesion are resected by the RecJ exonuclease. *M. genitalium* also lacks this enzyme, but this reaction could be catalyzed by the Pol-I-like protein and stabilization of the DNA ends could be performed by SsbA (MG_091) (Sluijter *et al.*, 2008) and the PcrA helicase. Then, RecA (MG_339) can catalyze the homologous strand exchange (Sluijter *et al.*, 2009) and promote DNA synthesis with a translesion DNA polymerase. PcrA, SsbA protein and RecU (MG_352) are also required for this step. Finally, once the DNA synthesis has finished, a Holliday junction (HJ) conformation must be resolved or dissolved. Dissolution of HJ is catalyzed by topoisomerases, RecQ and other helicases such as UvrD in other bacteria (Ayora *et al.*, 2011); however, this process is not well understood. HJ resolution is more common and is dependent on RuvAB resolvases (MG_358 and MG_359) and RuvC endonuclease (Sluijter *et al.*, 2010, 2012; Estevao *et al.*, 2011; Burgos and Totten, 2014a). Most Gram-positive bacteria lack RuvC, but instead utilize RecU instead (Ayora *et al.*, 2004), a similar protein also present in *M. genitalium*.

Table i.8. Summary of DNA repair components present in M. genitalium.

Gene	Locus Tag	Repair system	Function
uvrA	MG_421	NER	Excinuclease A, recognition of distortions in DNA helices
uvrB	MG_073	NER	Excinuclease B, cleavage of DNA (downstream of the lesion)
uvrC	MG_206	NER	Excinuclease C, cleavage of DNA (upstream of the lesion)
uvrD/pcrA	MG_244	NER/HR	Helicase, removal of the excised fragment in NER
ung	MG_097	BER	Uracil recognition and removal of the glycosidic bond
mutM	MG_498	BER	Recognition of oxidized guanines and removal from DNA strand
nfo	MG_235	BER	Endonuclease IV, recognition and excision of DNA lesions
dinB	MG_360	NER/BER/HR	Translesion DNA polymerase IV, it lacks proofreading activity
Pol-I-like	MG_262	NER/BER/HR	Putative 5' to 3' exonuclease, similar to DNA polymerase I subunit
ligA	MG_254	NER/BER/HR	NAD-dependent DNA ligase
disA	MG_105	HR (putative)	DNA integrity scanning protein
recA	MG_339	HR	Recombinase A, homologous strand exchange catalysis
ruvA	MG_358	HR	HJ DNA binding protein and DNA unwinding
ruvB	MG_359	HR	HJ ATPase, emerging DNA spooling
recU	MG_352	HR	HJ resolvase, DNA cleavage
ssbA	MG_091	HR	ssDNA binding, RecA recruitment to ssDNA

Homologous recombination repair system in *M. genitalium* has been questioned. A Δ*recA* mutant of *M. genitalium* had similar DNA repair capabilities than the WT strain upon exposure to DNA-damaging agents such as mitomycin C (MMC) (Burgos *et al.*, 2012). In contrast, a Δ*uvrC* mutant was more sensitive to MMC than the WT strain, suggesting that

NER pathway is the main DNA repair system of *M. genitalium* (Burgos *et al.*, 2012). However, the fact that MMC is highly unstable in medium containing high concentrations of fetal bovine serum (Proctor and Gaulden, 1986) could be altering these results. In fact, repair of crosslinked DNA can be done by NER pathway and translesion DNA polymerases alone (Dronkert and Kanaar, 2001), but NER cannot repair DSB in DNA. This suggests that MMC treatment in *M. genitalium* induces DNA crosslinks but does not result in DSB.

i.2.5. Scrambling DNA sequences: antigenic variation in *M. genitalium*

Despite the main role of RecA in recombinational repair of DNA in *E. coli* (Cox, 1991), *M. genitalium* has taken advantage of this machinery to generate antigenic variation. Scattered around the genome there are nine repetitive DNA regions homologous to the main cytadhesins of *M. genitalium* (MgpB and MgpC), named MgPa repeats (Fraser *et al.*, 1995). Each repeat, numbered 1-9, contains DNA sequences homologous to different regions of the MG_191 and MG_192 genes, named B, EF, G, KL and LM (Iverson-Cabral *et al.*, 2007) (see Figure i.10).
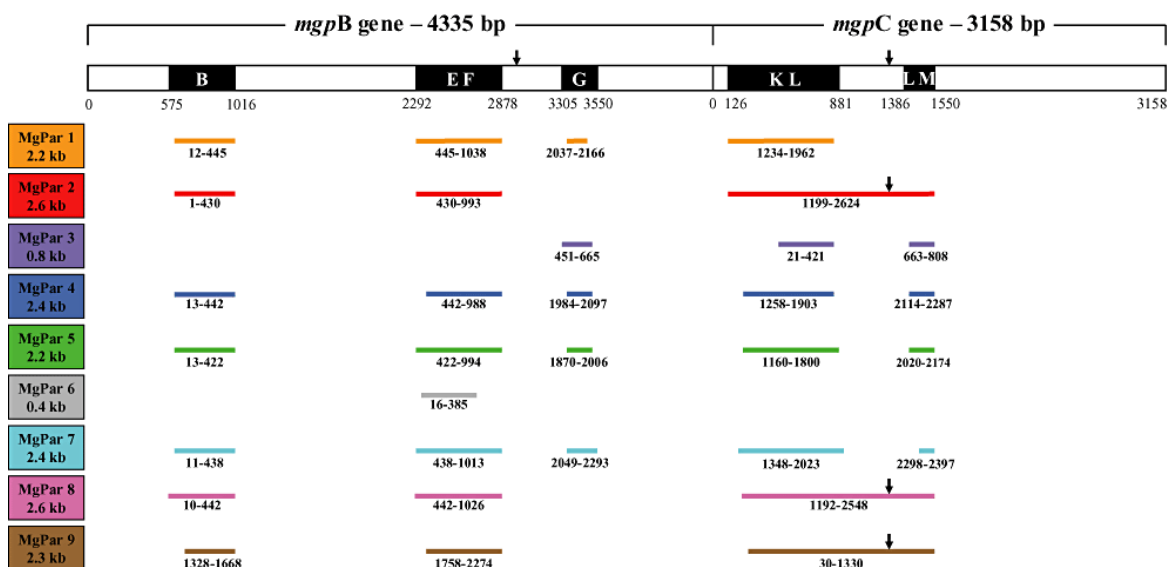
Figure i.10. Homology between the *mgpB* and *mgpC* expression sites and the nine MgPar sequences in the G37 genome. This figure depicts concentrated regions of homology within *mgpB* regions B, EF and G, and *mgpC* regions KL and LM, as well as the discontinuous and alternating homology characteristic of the majority of MgPar regions, determined by pairwise alignment. Because MgPars 2, 8 and 9 have homology to both *mgpC* region KL and LM, as well as the intervening sequences, they are referred as single, large *mgpC* repeat region KLM. Conserved portions of each gene are shown in white, repeat regions highlighted in black, and small arrows indicate the location of the 'AGT' trinucleotide repeats present within *mgpB*, *mgpC* and three MgPar sequences. Extracted from Iverson-Cabral *et al.*, 2007.

Interspaced position of the sequences homologous to *mgpB* and *mgpC* in the MgPa regions and the absence of putative functional transcriptional start sites indicate that these homologous sequences do not encode for alternative expressed peptides. Moreover, in all

MgPa repeats there are stop codons coded in all three reading frames, mostly within the A+T rich sequences (Peterson *et al.*, 1995). Instead, MgPa repeats serve as a reservoir of alternative DNA sequences to be used in genetic variation of both cytadhesins (Iverson-Cabral *et al.*, 2006, 2007) via homologous recombination mediated by RecA (Burgos *et al.*, 2012). Upon recombination, the reading frames of both *mgpB* and *mgpC* are normally maintained, as nucleotide insertions or deletions coming from MgPa regions occur in multiples of three. However, a small region of MgPar III homologous to KL region of *mgpC* has a different reading frame than the corresponding cytadhesin; as a consequence, premature stop codons would be introduced if recombination occurred at this site (Iverson-Cabral *et al.*, 2007). This mechanism of antigenic variation has been observed both in *in vitro* cultures (Burgos *et al.*, 2006; Iverson-Cabral *et al.*, 2006) and in *in vivo* infections (Wood *et al.*, 2013).

Another source of genetic variation in *mgpB* and *mgpC* adhesins is DNA slippage during replication. Both adhesins bear a trinucleotide repeat of the sequence AGT encoding successive serine residues (see Figure i.10). In *mgpB*, this trinucleotide is repeated seven times and it is located immediately downstream of the EF region. On the other hand, AGT repeat of *mgpC* is situated between regions KL and LM and it is also encoded in the MgPa regions 2, 8 and 9, which have the sequence KLM. In this case, *mgpC* wild-type sequence has 11 repeats while MgPa region 2, 8 and 9 code for 16, 10 and 9 repeats, respectively. Therefore, recombination of the different MgPa regions could modify the number of Ser residues of MgpC protein. However, slipped-strand mispairing during DNA replication could also increase or decrease the AGT repeat length and generate new variants in both *mgpB* and *mgpC* genes (Iverson-Cabral *et al.*, 2007).

In some cases, recombination between specific MgPa regions and the adhesin genes can generate hemadsorption negative mutants (Mernaugh *et al.*, 1993). These mutants were classified in two variants: class I, which completely lack MgpC protein and show trace amount of MgpB and class II mutants, which lack both MgpB and MgpC adhesins. Although some of these mutations are reversible, some class I and class II mutations are perpetual and can only be complemented by the reintroduction of *mgpB* and *mgpC* wild-type genes (Burgos *et al.*, 2006). This permanent loss of part of the MgPa operon arises as a consequence of a single recombination event between sequences of the MgPa operon and the MgPa region V, placed immediately downstream of the cytadherence operon (Figure i.11). Class I mutants are generated by the recombination of KL region of *mgpC* gene and KL sequence of the MgPa region V, which truncates the *mgpC* gene while leaving the *mgpB*

intact. In constrast, class II mutants are the result of a single recombination event between EF regions of *mgpB* and MgPa region V that eliminates part of the *mgpB* gene, the whole *mgpC* gene plus part of the MgPa region V (see Figure i.11). Despite class I mutants preserve the *mgpB* gene, they yield a hemadsorption-negative phenotype (HA⁻). This fact could be explained because both MgpB and MgpC cytadhesins are co-stabilized and required for cell adhesion (Burgos *et al.*, 2006).
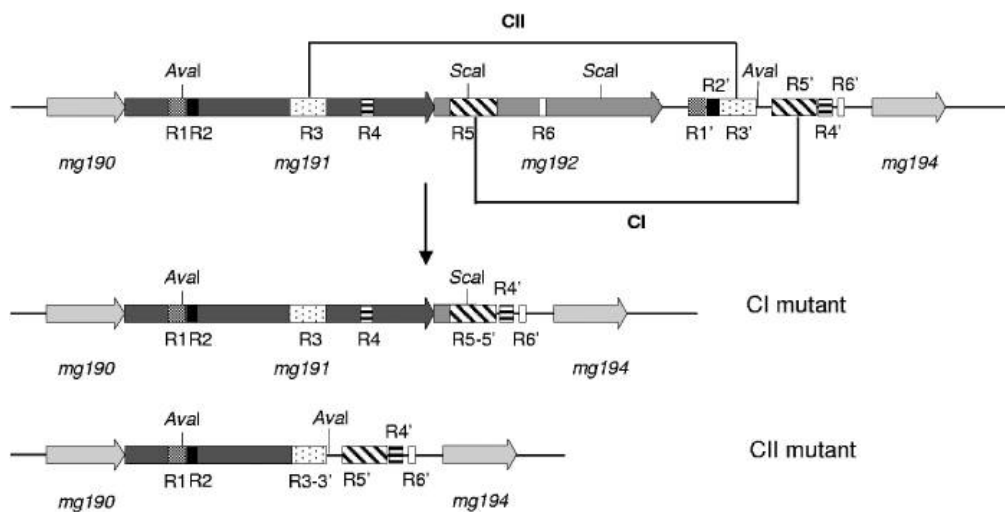
Figure i.11.Schematic representation of the DNA repetitive elements of the MgPa operon and MgPa island V located immediately downstream. A single-recombination event between the R5 and R5′ boxes (KL region) and the R3 and R3′ boxes (EF region) may generate class I and class II mutants, respectively. CI, class I mutant; CII, class II mutant. Adapted from Burgos *et al.*, 2006.

Despite HA⁻ mutants have impaired infection capabilities (Mernaugh *et al.*, 1993), some *M. genitalium* HA⁻ mutants have been isolated from clinical samples (Jensen *et al.*, 1996), like *M. genitalium* M2288 strain. However, M2288 hemadsorption-negative phenotype could be reversed by cell passaging and selection of adherent revertants, and complete genome sequencing revealed that the MgPa operon remains intact (McGowin *et al.*, 2012). Therefore, to date, there is no evidence of the presence of irreversible class I or II mutants *in vivo*, despite the high rates of appearance of these mutants *in vitro*. However, it seems unlikely that *M. genitalium* has kept an irreversible mechanism for adhesin loss with the consequent colonization impairment for no significant reason.

OBJECTIVES

Main objective

To characterize the MG428 regulator (σ^{20}) and to define its regulon and its biological role in the human pathogen *M. genitalium*.

Chapter 1

1. To obtain a null mutant for MG_428 gene of *M. genitalium* and determine the nature of the MG428 regulator.
2. To identify the whole regulon of MG428 through high throughput analyses.
3. To identify the DNA sequence recognized specifically by MG428.
4. To understand the protein expression patterns of MG428 and some of the regulated genes in single cells.
5. To examine the biological role of the MG428 pathway and its involvement in homologous recombination and antigenic variation.

Chapter 2

6. To characterize two novel proteins under the control of σ^{20}, RrlA and RrlB, and understand their unique regulatory role.
7. To identify factors controlling the activation of the σ^{20} pathway.
8. To assess the capacity of *M. genitalium* to transfer DNA horizontally using the σ^{20} pathway and the recombination machinery.

CHAPTER I

MG428 IS AN ALTERNATIVE SIGMA FACTOR THAT PROMOTES ANTIGENIC VARIATION IN *M. GENITALIUM*

The *Mycoplasma genitalium* MG428 protein shows homology to members of the sigma-70 family of sigma factors. Herein, we found that MG428 activates transcription of *recA*, *ruvA* and *ruvB* as well as several genes with unknown function. Deletion of MG_428 or some of the up-regulated unknown genes led to severe recombination defects. Single cell analyses revealed that activation of the MG428-regulon is a rare event under laboratory growth conditions. A conserved sequence with sigma-70 promoter architecture (TTGTCA-N$_{18/19}$-ATTWAT) was identified in the upstream region of all the MG428-regulated genes or operons. Primer extension analyses demonstrated that transcription initiates immediately downstream of this sigma70-type promoter in a MG428-dependent manner. Furthermore, mutagenesis of the conserved -10 and -35 elements corroborated the requirement of these regions for promoter function. Therefore, a new mycoplasma promoter directs transcription of a unique recombination regulon. Additionally, MG428 was found to interact with the RNAP core enzyme, reinforcing the predicted role of this protein as an alternative sigma factor. Finally, our results indicate that MG428 contributes to the generation of genetic diversity in this model organism. Since recombination is an important mechanism to generate antigenic variation, MG428 emerges as a novel factor contributing to *M. genitalium* virulence.

I.1. INTRODUCTION

Mycoplasma genitalium is a sexually transmitted pathogen implicated in urogenital diseases such as urethritis, cervicitis, pelvic inflammatory disease and infertility (McGowin *et al.*, 2011). In addition to the significance of this microorganism as an emerging human pathogen, research on *M. genitalium* is intimately associated with important scientific milestones such as the construction of the first synthetic chromosome (Gibson *et al.*, 2008, 2010) or the development of a whole-cell computational model (Karr *et al.*, 2012). Certainly, the massive genome reduction undergone by *M. genitalium* favors its suitability as a systems biology model. Compared with more complex bacteria, *M. genitalium* lacks most known transcription factors and regulatory pathways. However, despite the apparently reduced gene regulatory toolbox, both environmental stresses and metabolic insults induce complex, specific transcriptional responses in this bacterium (Zhang and Baseman, 2011a, 2011b). Therefore, the elements and mechanisms regulating gene expression in *M. genitalium* are essentially unexplored.

Sigma factors are fundamental components of the prokaryotic transcriptional machinery that direct the core RNA polymerase to specific promoter elements. Members of the sigma-70 family recognize promoters with two conserved motifs centered roughly 10 and 35 base pairs upstream of the transcription initiation site (Figure I.1) (Paget and Helmann, 2003). A representative member of this family is the primary sigma factor, RpoD, which recognizes the consensus sequence TTGACA-N$_{15/19}$-TATAAT and directs transcription of most genes in growing cells.

Figure I.1. Schematic representation of σ70 family protein sequence (marked 1-4). Based on sequence conservation among members of the σ70 family, σ factors can be divided into four different regions. Region 4 is placed at the carboxy-terminal and forms a helix-turn-helix motif (subregion 4.2) that binds the -35 element of the promoter sequence. Amino acids from conserved regions 2 and 3 intervene in the recognition of the -10 element and melting of the double strand. Subregions 2.3 and 2.4 interact intimately with the -10 element, while subregion 2.3 (dashed arrow) is hypothesized to interact with ssDNA once the RNApol has melted DNA strands. Typically, a single residue in the amino-terminal end of region 3 (subregion 3.0) is responsible of contacting a conserved motif situated upstream of the -10 element, normally TG. Underneath the linear structure, the three domains of the σ factor observed by X-ray crystallography (σ$_2$, σ$_3$ and σ$_4$) are boxed. Adapted from Paget and Helmann, 2003.

However, most bacteria have alternative sigma factors, which control the expression of different subsets of genes usually in response to specific environmental conditions. *Mycoplasma genitalium* and its close relative *Mycoplasma pneumoniae* possess a primary sigma-70 factor (MG249 and MPN352, respectively) but no alternative sigma factors have been characterized. Transcriptional analyses have shown that the archetypal sigma-70 promoter of *M. pneumoniae* features a conserved −10 element but it is devoid of a well-conserved −35 region (Table I.1) (Weiner, Herrmann and Browning, 2000; Halbedel *et al.*, 2007; Güell *et al.*, 2009; Lluch-Senar *et al.*, 2015).

Table I.1. Experimentally determined transcriptional start sites (TSS) and promoter regions of several genes of *M. pneumoniae* strain M129. Adapted from Weiner *et al.*, 2000.

Gene[a]	Annotation[b]	No. Starts[c]	Upstream region[d]	-10[e]	Start[f]
MPN152	lipoprotein	5	CTAGTAAATTTCAAGGG**TTAAAC**ATCTACAAACTTTGT	TAATAT	TTCTC**T**
MPN140	DHH family	2	AGTCAAACTACGA**CAACAA**CAGTTGCTGTTTGATTCTT	TAAACT	TAAACA**G**
MPN097	lipoprotein	5	AGAGGGCGAGTTTTGC**TGGTAT**GCTTAGCGCGTTTTTG	TAAGAT	TTGG**C**
MPN084	lipoprotein	5	ACAGCGGCATTAA**TCGGTT**TTTGAGTAAATTTATAAAG	TACTAT	TTTCTTT**A**
MPN083	lipoprotein	2	TGACTAACTTACTCAAG**TTGATT**AAAACTCTAAAGAGT	TAAAAT	AGAGT**A**
MPN052	lipoprotein	1	CTTATTTAAATACT**TTGAAA**ATCAACGATCTAATAAGT	TACTAT	TACG**C**
MPN592	lipoprotein	4	ACCTTTAATTAAGGG**TTCACC**TTTCAAACTTTCTTTTT	TAAAAT	TAAAGC**A**
MPN591	HP	1	TTTTCCAGCTTAACCTA**AGTACA**TTTATGCCAAACTTA	TATAAT	TGGGC**A**
MPN574	groES	1	GCACTTTAAGCATTTAA**CTGCTA**ATTGTTAATAATTGG	TAGAAT	ATTG**A**
MPN560	*arcA*	1	CCATGGAAGTAACGCAA**TTAAAG**CAATTGTTAATTATT	TAAAAT	TAATC**G**
MPN531	*clpB*	2	CACTCGTTAGCACTC**AAGCCA**TTCGAGTGCTAATTTTA	TATAAT	TGGCCT**A**
		1	CAAGCCATTCGAGTG**CTAATT**TTATATAATTGGCCTAT	TAACAA	AAGAAAG**G**
MPN491	*mnuA*	4	CTAAATTGATTCAAAAA**GGGAAA**AACGCCATTTTTTA	GATAAT	TAAGG**C**
MPN459	lipoprotein	1	CTTGAGCTTTAATT**TCGCAG**ATTAGCAAGCTCCCAAAT	TAGAAT	AAAATC**A**
MPN455	*ctaD*		ATAATTTTGCATTTA**TTAGCA**ATTAAGCGCTTACAAAT	TAAAAT	TGACCC**A**
MPN454	HP		AGTTAAGCTGTAAGTG**GGGAAT**TAAAGCTTTGAAAAGT	TAAAAT	TTTCCC**A**
MPN449	HP		TTAACGGCCTTAATTA**ATGCGA**TTTGGCAGCTGGATAT	TAAAAT	TAAGCG**C**
MPN446	*rpsD*		AAACGGTGGGGAT**TTTAAT**CTTGATTATTAAGATAATT	TAAAAT	TGTCAC**C**
MPN444	lipoprotein	5	TCGAGCCGCATCCC**CGGTCA**AAGAAAGTGGTTAAATTT	TAGAAT	TATA**C**
MPN401	*greA*	1	TTTTTGGCAAATA**TTGGGA**AAAAAAAGTAGATTTAATT	TAACTT	TAAAC**A**
		1	AATTTTGCATTTTT**TTGGCA**AATATTGGGAAAAAAAG	TAGATT	TAATTT**A**
MPN396	*secD*	1	GTTCAATTTCAACTAAA**GTGGCC**ATTTTGCTAAATTTA	TATAAT	TTAAGT**C**
MPN393	*pdhA*	2	AAGCGCTTGGTTT**TAGTTA**CTATACCCCTAAAAATAAA	TAAGAT	TAAA**C**
MPN376	HP	2	CACGCGCAGAAGGA**TTGTCA**AAAATATTTTTAAAGTGC	TAGAAT	AAAG**C**
MPN309	p65		TAAATAGCATTTATTCA**TTTGCA**TTTTTTTAGATAAAT	TAAAAT	TAATGGT**A**
MPN304	*arcA* (Nter)	1	CACTTTGCCTTTTTTTC**TTGGGC**TTTTGTGTGGACGTG	TAACAA	TTAAA**G**
MPN288	lipoprotein	1	CACTGAATTTTTAGGA**ATTAAA**CTTGTTTAATCTTTAT	TAAGAT	TTC**T**
MPN281	lipoprotein	2	AAAATCAGATTCTT**TTTACT**TTCTTAAAAATATTTTTT	TAAGAT	TTTCC**A**
MPN271	lipoprotein	2	AATTTAACTCTTATGA**GTTAAA**AAGTTCATTTTTTAAA	TAACAT	TAAACC**G**
		2	TAAAAAGTTCATTTT**TTAAAT**AACATTAAACCGGAAAT	TTCAAT	TAGTTTCC**T**
MPN212	HP	1	CACCATCCAAATGG**CTTACA**ACGAACAACACCACAAAA	CACCAA	TG**A**

[a] For genes MPN531, MPN401 and MPN271 two different transcriptional start sites were determined.
[b] Current annotation. HP, hypothetical protein. All lipoproteins are defined as putative proteins.
[c] Number of alternative start points of transcription immediately 5' to the major start point.
[d] Putative -35 regions are highlighted in bold and underlined.
[e] Putative -10 regions.
[f] Major base for initiation of transcription is highlighted in bold.

Of note, transcription initiation has also been described in locations lacking an apparent Pribnow box, suggesting the existence of both non-canonical promoter elements and alternative sigma factors controlling transcription in mycoplasmas (Weber *et al.*, 2012; Lloréns-Rico *et al.*, 2015).

Sequence similarity and secondary structure prediction analyses suggest that the *M. pneumoniae* MPN626 protein is a putative sigma factor (Bornberg-Bauer and Weiner, 2002). Indeed, MPN626 shows distant similarity to sigma-70 factors of group IV, which accommodates the members of the extracytoplasmic function (ECF) subfamily of sigma-70 factors (Figure I.2) (Heimann, 2002). Typically, ECF sigma factors are members of cell-surface signaling systems that respond to signals arising from the extracytoplasmic environment. These signals are largely unknown, which limits the understanding of the physiological role of these transcriptional regulators. Of the four conserved regions characteristic of sigma-70 factors (Österberg *et al.*, 2011), only the region implicated in the recognition of the −35 promoter element (region 4) seems to be conserved in the MPN626 protein. Unfortunately, both proteomic and interactomic analyses of this putative sigma factor have been unsuccessful, possibly as a result of its implicit low cellular abundance (Maier *et al.*, 2011). Therefore, the nature of the MPN626 regulator and its associated regulatory DNA elements remain unknown.
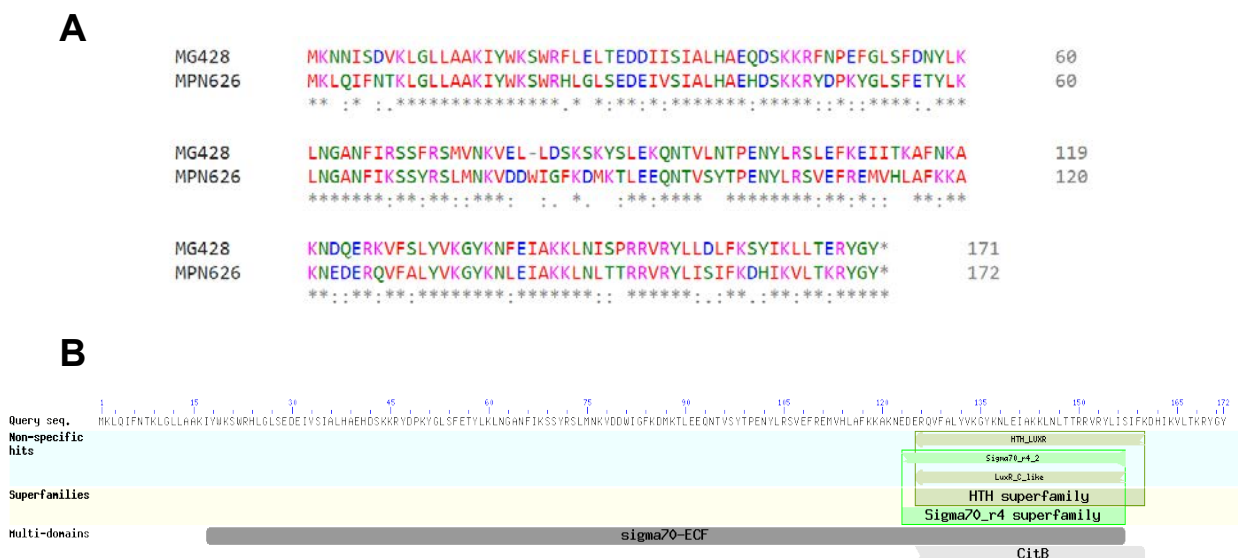
Figure I.2. (A) Sequence alignment between *M. genitalium* MG428 and *M. pneumoniae* MPN626 proteins, with more than 67% of identity among them. **(B)** Conserved domains of *M. pneumoniae* MPN626 protein. Most residues are encompassed in the sigma70-ECF multi-domain. At the C-terminal part of the protein, there is a well-defined helix-turn-helix domain (HTH), which is a classic DNA binding domain. This HTH domain has homology with sigma70 region 4.2, which can contact -35 consensus sequence of bacterial promoters (TTGACA) (see Figure I.1). Apparently, no conserved domains of regions 2 or 3 of sigma-70 are observed in MPN626.

Herein, we investigated the regulatory function of the MG428 protein, which is the *M. genitalium* orthologue of MPN626 (Figure I.2A). Our results demonstrate that MG428 is an alternative sigma factor that activates recombination in this human pathogen. Since

recombination plays a pivotal role in the generation of antigenic variation (Iverson-Cabral *et al.*, 2007; Ma *et al.*, 2007; Wood *et al.*, 2013), MG428 emerges as a novel factor contributing to *M. genitalium* virulence. Orthologues of the MG428 protein can be identified in other species such as *Mycoplasma gallisepticum* str. R(low) (MGA_0765), *Mycoplasma capricolum* subs. capricolum (MCAP_0855), *Mycoplasma suis* str. Illinois (MSU_0577) or *Acholeplasma* sp. CAG:878 (WP_021921817), suggesting that this protein and its associated regulatory functions might be widespread in the bacterial class Mollicutes.

I.2. RESULTS

I.2.1. Construction of a ΔMG_428 mutant and its complemented strain

To investigate the role of the MG428 protein, we constructed a MG_428 null mutant by allelic exchange (Figure I.3) using plasmid pΔMG_428. After electroporation with pΔMG_428, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_428 in five selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (SCRmg428-F and SCRmg428-R) were complementary to sequences of the chromosome located immediately outside of the MG_428 flanking regions cloned into the pΔMG_428 plasmid (Figure I.3A). This ensured that the screening method was specifically detecting the replacement of the MG_428 locus by the tetM438 marker, rather than the insertion of pΔMG_428 elsewhere in the genome. The amplification of a 4.2 kb band confirmed the replacement of the MG_428 gene (516 bp) by the tetM438 marker (1954 bp) in the five transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the MG_428 gene. As expected, a band of 2.7 kb was observed when the chromosomal DNA of the wild-type strain was analyzed (Figure I.3B).
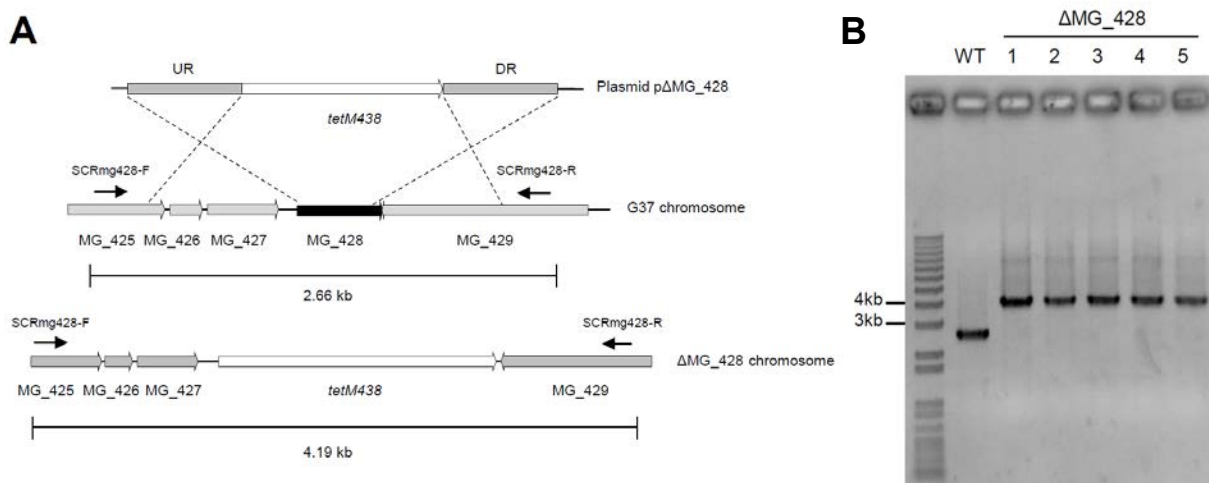
Figure I.3. Construction of a ΔMG_428 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_428 locus by the tetM438 marker in different ΔMG_428 mutant clones.
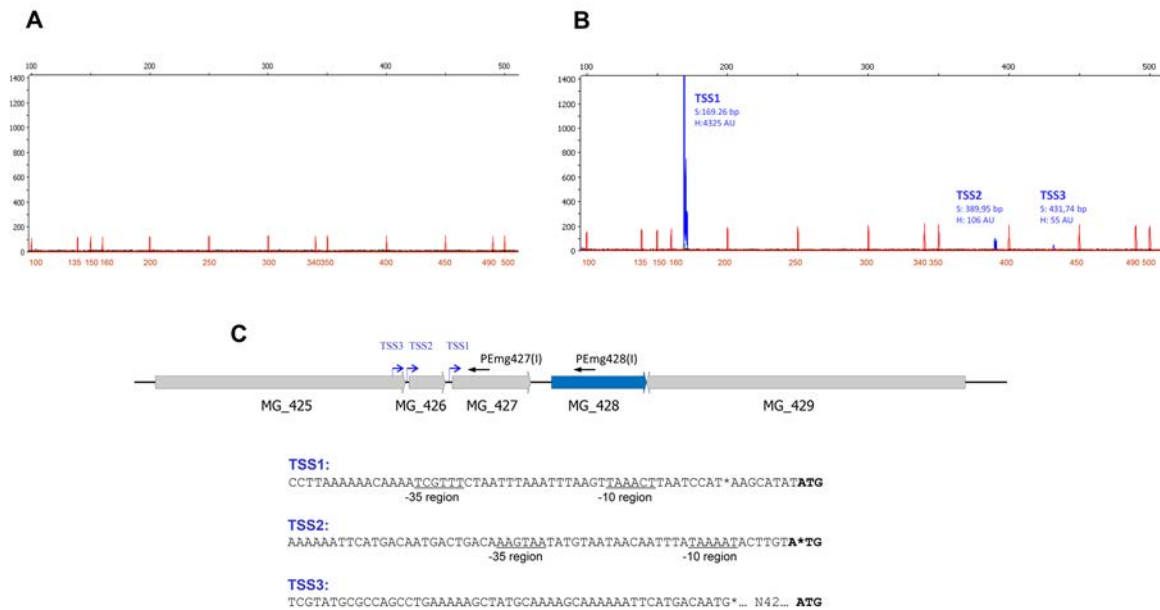
Figure I.4. Identification of the MG_428 TSS by primer extension. (A) Primer extension analysis of the MG_428 gene in the WT strain with primer PEmg428(I). No TSS could be detected within the five hundred nucleotide chromosomal segment located immediately upstream of the PEmg428(I) primer. These results were confirmed with primer PEmg428(II). **(B)** Primer extension analysis of the MG_427 gene in the wild-type with primer PEmg427(I). Three independent TSSs, designated TSS1, TSS2 and TSS3 were identified. These results were confirmed with primer PEmg427(II). All electropherograms were generated with Peak Scanner v1.0 (Applied Biosystems) analysis software. Red peaks represent ROX size standards while blue peaks correspond to the primer extension products. **(C)** Identification of putative promoter sequences upstream of the three identified TSSs. A schematic representation of the genome region analyzed is shown, and the presence of the identified promoters is indicated with blue arrows. Additionally, sequences flanking the identified TSSs are shown and the putative -35 and -10 regions underlined. Location of the identified TSSs is indicated with an asterisk and the translational start points of the corresponding genes are highlighted in bold letters.

For control purposes, the MG_428 wild-type (WT) allele was subsequently reintroduced into the ΔMG_428 mutant by transposon delivery to create its respective complemented strain. To this end we used the Tn*CatMG_428* minitransposon (MiniTnp), which carries the MG_428 gene under the control of its own promoter. Primer extension analysis revealed that this promoter is located upstream of the MG_427 gene (Figure I.4).

TSS-1 was located eight bases upstream of the translational start site of the MG_427 gene, which codes for a protein with homology to the osmotically inducible protein C (*osmC*). Consistently, a putative -10 promoter element (TAAACT) was found seven bases upstream of this TSS. This result is in agreement with data published in a recent study addressing the role of the MG427 protein in *M. genitalium* (Zhang and Baseman, 2014). TSS-2 overlapped with the translational start site of the MG_426 gene, which codes for the 50S ribosomal protein L28. A putative -10 promoter element (TAAAAT) was located six bases upstream of this TSS. TSS-3 was located forty-two bases upstream of the MG_426 translational start site. In this case, no apparent sigma-70 promoter elements could be recognized. In

agreement with previous reports, the three potential TSS identified were heterogeneous and devoid of well-conserved -35 promoter elements (III, Herrmann and Browning, 2000; Guell *et al.*, 2009).

Consistently, transformation efficiencies obtained in electroporation experiments of the ΔMG_428 mutant with the Tn*CatMG_428* MiniTnp were markedly low (~10^{-7} transformants per viable cell). On average, we only obtained seven chloramphenicol-resistant colonies per electroporation experiment. In contrast, electroporation of ΔMG_428 mutant with a Tn*Cat* MiniTnp without the MG_428 yielded ~10 thousand chloramphenicol-resistant colonies (~10^{-4} transformants per viable cell) (see Supp. Table S.6). The dramatic reduction in the number of viable transformants observed after electroporation with the Tn*CatMG_428* MiniTnp suggested that the expression of the MG_428 gene supplied *in trans* was deleterious to *M. genitalium*.

Twenty-five chloramphenicol-resistant colonies from four different electroporation experiments were picked up, propagated and the MiniTnp insertion point determined by sequencing with primers specific to the chloramphenicol acetyltransferase (*cat*) gene. We found the MiniTnp inserted within specific loci rather than randomly inserted throughout the chromosome (Figure I.5A, Supp. Table S.1). Of note, all the MiniTnp insertions were found in the antisense orientation with respect to the disrupted gene, indicating that this orientation is also selected. Furthermore, we determined that the two isolated mutants with a MiniTnp insertion within the MG_390 gene carried a truncated copy of the reintroduced MG_428 gene (Figure I.5B), reinforcing the hypothesis that the expression of the MG_428 gene supplied *in trans* was toxic. We also isolated a mutant bearing a truncated allele of MG_428 in a symmetric chromosomal inversion (Supp. Figure S.1). A possible mechanism to generate MG_428 truncated copies is described in Supp. Figure S.2.
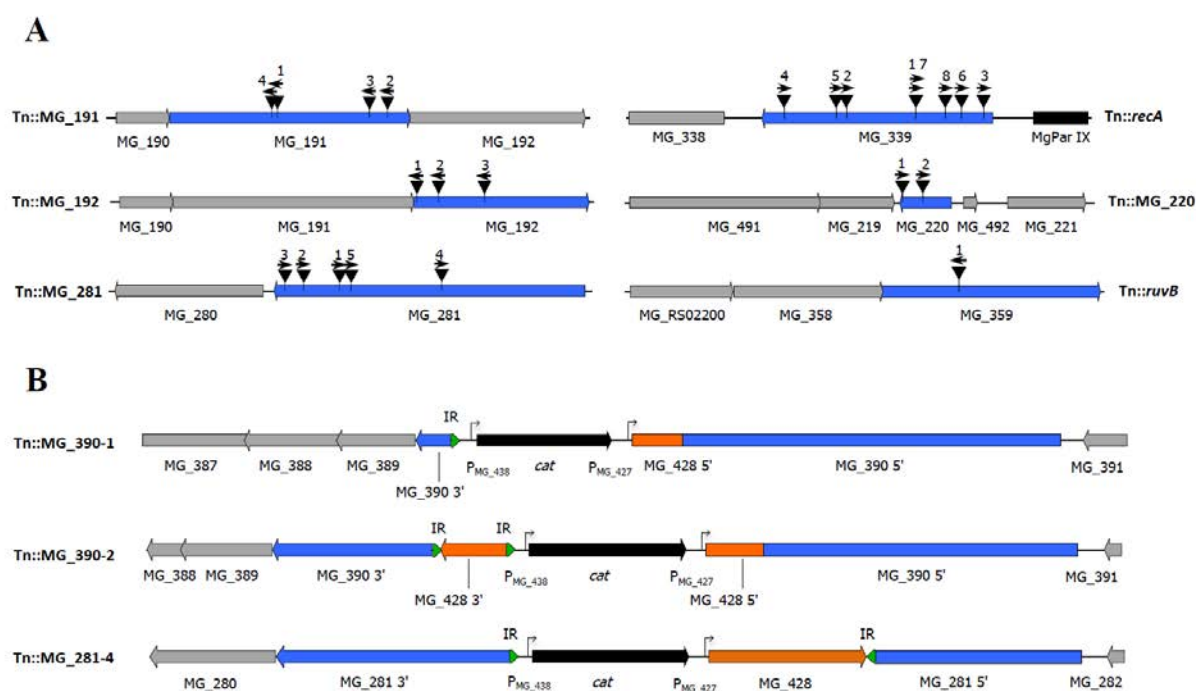
Figure I.5. Insertion points of the Tn*CatMG_428* MiniTnp in the *M. genitalium* genome. A) Schematic representation depicting the genes preferentially targeted by Tn*CatMG_428* in the genome of a ΔMG_428 mutant. Genes disrupted by Tn*CatMG_428* are shown in blue. Black filled triangles represent the transposon insertion points. Arrows above the triangles indicate the orientation of the transposon insertion. **B)** Schematic representation showing the presence of a truncated MG_428 ectopic copy in the genome of the Tn::MG_390-1 and -2 mutants. The MG_428 coding region is highlighted in orange and the inverted repeats (IR) of the Tn*CatMG_428* MiniTnp are shown in green. The Tn::MG_281-4 mutant, carrying a full copy of the MG_428 gene, is also shown for comparison.

I.2.2. Analysis of MG428 expression by Western blot

Expression of the MG428 protein was analyzed by Western blot using anti-MG428 polyclonal antibodies (Figure I.6). The MG428 protein was not detected in lysates of the WT strain, indicating that it was expressed at very low levels in *M. genitalium*. Likewise, MG428 expression was not detected in lysates of the ΔMG_428 mutant. In contrast, a band of the predicted molecular mass of the MG428 full-length protein (17kDa), was clearly detected in the complemented mutants.
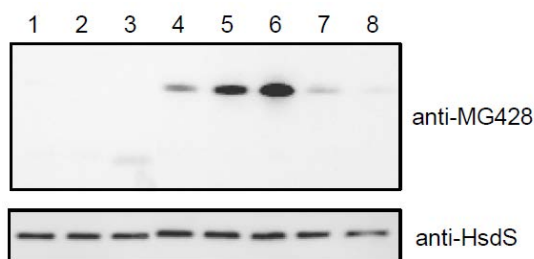
Figure I.6. Analysis of protein expression by western blot. Immunoblot analysis of MG428 expression in the WT strain and several representative mutants. Lane 1, WT; lane 2, ΔMG_428; lane 3, Tn::MG_390-1; lane 4, Tn::MG_281-1; lane 5, Tn::*recA*-1, lane 6, Tn::MG_220-1; lane 7, Tn::MG_191-2 and lane 8, Tn::MG_192-1. HsdS protein (MG438) was detected with a monoclonal antibody and used as a loading control.

This result indicates that transcriptional fusion of the MG_428 gene to its own promoter, located upstream of the MG_427 gene, leads to increased levels of MG428 expression as compared to the WT strain. In light of this finding, we analyzed the intergenic region between the MG_427 and MG_428 genes and we identified a possible Rho-independent terminator (Figure I.7). The presence of this putative terminator, which is supported by several transcription terminator prediction software (Gautheret and Lambert, 2001; Lesnik *et al.*, 2001; Kingsford *et al.*, 2007), could explain the reduced levels of MG428 expression observed in the WT strain.

533704 TAAAAAAGCGAAAAATACCTGTAATTTTGCAGGTATTTTTTTATTTAAAA 533753

Figure I.7. Predicted RNA secondary structure of the terminator sequence identified within MG_427 and MG_428 intergenic region. The RNA secondary structure prediction was generated with the RNAfold WebServer (Gruber *et al.*, 2008). The positional entropy box was chosen in the display options. The free energy of the thermodynamic ensemble is -16.34 kcal mol^{-1}. Intrinsic terminator is depicted at the illustration at the downstream region of MG_427. Exact sequence and base-pair location in the genome is also shown.

On the other hand, MG428 levels differed considerably among the complemented strains, indicating that the genetic context of the transposon insertion had a significant impact on the expression of the MG_428 ectopic copy. Finally, in agreement with the presence of a truncated copy of the MG_428 gene, a 9kDa band was detected in the Tn::MG_390-1 mutant (lane 3).

I.2.3. Analysis of gene expression by RNAseq

Transcriptional changes in a ΔMG_428 mutant were assessed by RNAseq analysis. As expected, no MG_428 transcript could be detected as compared to the WT strain (Figure I.8 and Table I.2). Excluding this, no other significant changes ($p < 0.05$) could be detected between the ΔMG_428 mutant and the WT strain (Table I.2). The lack of major transcriptional changes in the ΔMG_428 mutant suggests that the activation triggered by this putative sigma factor may not be constant in all the cell population at once.

Figure I.8. Plot of relative mRNA coverage at the MG_428 genomic region. The upper panels depict relative coverage of the WT strain in both DNA strands (Ref+ and Ref-). The lower panels represent relative coverage of the ΔMG_428 mutant in both DNA strands (Exp+ and Exp-). Differences in the positive strand of MG_429 gene in ΔMG_428 mutant can be explained as a polar effect of the expression of the tetracycline resistance gene.

Table I.2. Genes and regions up- or down-regulated in the ΔMG_428 mutant as compared to the WT strain.

Gene	Base mean	Log2 Fold Change	Standard Error	p-value
MG_284	30.57152	0.445578	0.465147	0.338099
MG_518	19.60915	0.404667	0.497669	0.416147
MG_184	208.6886	0.308362	0.26188	0.238998
MG_406	1.557241	-0.98313	1.852685	0.59566
MgPar5-G/+	13.43957	-1.0234	0.692316	0.139346
MgPar5-KL/+	64.55705	-1.21056	0.358486	0.000733
MG_381	3.857438	-1.30237	1.176973	0.268493
MG_428	133.4759	-4.10091	0.604659	1.18E-11

Additionally, we also analyzed one of the Tn::*recA* complemented mutants by RNAseq, which showed a marked MG428 overexpression by Western blot (Figure I.6). Consistently, we found increased levels of MG_428 transcript (~10.5-fold) as compared to the WT strain (Table I.3). No transcriptional changes of the MG_427 gene were observed, suggesting that MG428 overexpression has no effect on MG_427 transcript levels and, therefore, on MG_427 promoter.

On the other hand, several genes and regions were found to be upregulated upon MG428 overexpression. To simplify the analysis, we have divided them in different categories: genes targeted by the Tn*CatMG_428* minitransposon, rest of genes and non-coding RNAs.

All the RNAseq raw data is available in the Appendices section (Supp. Table S.2 and Supp. Table S.3).

Genes targeted by the Tn*CatMG_428* MiniTnp

- ***recA* or MG_339.** The gene coding for the recombinase A was found to be dramatically upregulated (~17.3-fold) in the Tn::*recA* mutant (Table I.3). To rule out possible polar effects due to the insertion of the Tn*CatMG_428* MiniTnp within the *recA* gene in this particular mutant, qRT-PCR analyses were also conducted with an intact *recA* gene (Supp. Figure S.3).

- ***ruvAB* operon or MG_358/MG_359.** The genes coding for the Holliday junction helicases A and B were also upregulated in the Tn::*recA* mutant (~16-fold). The *ruvAB* operon has recently been linked to the generation of antigenic variation in *M. genitalium* (Burgos and Totten, 2014a).

- **MG_220.** Transcription of the MG_220 gene was dramatically upregulated (~43-fold) in the Tn::*recA*-2 mutant. MG_220 is a small ORF of unknown function situated between the cytadherence regulatory locus and the division and cell wall gene cluster of *M. genitalium*.

Table I.3. MG_427, MG_428 and genes targeted by the Tn*CatMG_428* MiniTnp that are activated by MG428 in the complemented mutant as compared to the WT strain.

Gene	Base mean	Log2 Fold Change	Standard Error	p-value
MG_427	7617.683	-0.45138	0.187241	1.5E-2
MG_428	2954.158	3.351954	0.259285	3.14E-38
MG_220	1964.251	5.430537	0.152063	2.5E-279
MG_339	1864.701	4.115712	0.122973	1.4E-245
MG_358	891.9543	4.069886	0.155872	2.8E-150
MG_359	2293.894	3.96605	0.14897	3.7E-156
MG_389	33.85287	1.983034	0.297973	2.83E-11
MG_390	1122.28	1.168877	0.141805	1.68E-16

- ***sunT* or MG_390.** MG_390 was also activated at some extend (~2.5-fold) upon MG428 overexpression. Likewise, the MG_389 gene was also found to be upregulated, suggesting that both genes are co-transcribed. MG390 shows sequence similarity to SunT, a conserved uncharacterized protein among

mycoplasmas that resembles an ABC bacteriocin-processing exporter. Bacteriocins are antimicrobial peptides produced by bacteria to inhibit cell growth of other competitor bacteria (Reeves, 1965), but no bacteriocins have ever been isolated from mycoplasmas to date.

Rest of upregulated genes

- **MG_285 and MG_286.** Upregulation of both genes in the complemented mutant is ~13-fold as compared to the WT strain (Table I.4). Both genes code for hypothetical proteins of *M. genitalium*, with known homologs in some mycoplasma species. Looking at their genomic organization, MG_285 and MG_286 seem to comprise an operon along with MG_284, but the latest does not seem to be under the control of MG428.

- **MG_414.** MG_414 is activated ~12-fold upon MG428 overexpression. The gene situated immediately upstream (MG_525, previously annotated as mg415) is also activated at some extend, but this change is not statistically significant. However, qRT-PCR analyses of several complemented mutants confirm the existence of this slight upregulation of MG_525 (Supp. Figure S.3). MG_414 and MG_525 are paralogues, probably generated by a gene duplication event. MG_414 codes for an expressed protein of unknown function.

- **MG_RS02200.** This coding region is upregulated ~10.8-fold in the Tn::*recA* mutant. It was not previously annotated in the *M. genitalium* G37 strain (until 2015), but it was present in other *M. genitalium* strain annotations as well as in *M. pneumoniae* M129 (MPN534 gene). MG_RS02200 has no orthologues in other mycoplasmas, except for the aforementioned MPN534 gene of *M. pneumoniae*. In both species, it is situated immediately upstream of the *ruvA* gene and a possible role in recombination was proposed in a recent work (Burgos and Totten, 2014a).

- **MG_412.** MG_412 is also upregulated (~2.5-fold) in the complemented mutant. MG_412 codes for an uncharacterized lipoprotein of *M. genitalium* present in a phosphate transport operon. It has been demonstrated that MG412 protein activates NF-κB pathway through TLR1 and TLR2 in 293T cells *in vitro* (Shimizu *et al.*, 2007); thus it is a potential virulence factor of *M. genitalium*.

- **MG_010.** Last, MG_010 gene is slightly upregulated upon MG428 overexpression (~2-fold). MG010 is a small primase-like protein unique to mycoplasmas. Its function is still unknown, but in other bacteria similar small TOPRIM domain-containing proteins have been redefined as RNA maturase M5 or ribonuclease

M5. These are enzymes responsible for the maturation of the 5S ribosomal RNA (Allemand *et al.*, 2005). However, 5S RNA maturation is not necessary for proper ribosome function (Condon *et al.*, 2001), so its biological relevance is still on debate.

Table I.4. Genes not targeted by the Tn*CatMG_428* MiniTnp that are activated by MG428 in the complemented mutant as compared to the WT strain.

Gene	Base mean	Log2 Fold Change	Standard Error	p-value
MG_285	1120.683	3.728067	0.160431	1.9E-119
MG_286	379.9618	3.647855	0.198142	1.09E-75
MG_414	1298.149	3.610389	0.173115	1.36E-96
MG_RS02200	1279.847	3.434345	0.14219	6.9E-129
MG_412	983.7487	1.376309	0.15686	1.72E-18
MG_010	57.25729	0.982329	0.310018	1.5E-3

Non-coding RNAs upregulated by MG428

Aside from the activation of specific ORFs of *M. genitalium*, we found that transcription of three non-coding regions was markedly upregulated upon MG428 overexpression. Essentially, the upregulated regions comprise DNA repeats with sequence homology to the MG_191 and MG_192 genes, which code for P140 and P110, the main cytadhesins of *M. genitalium*.
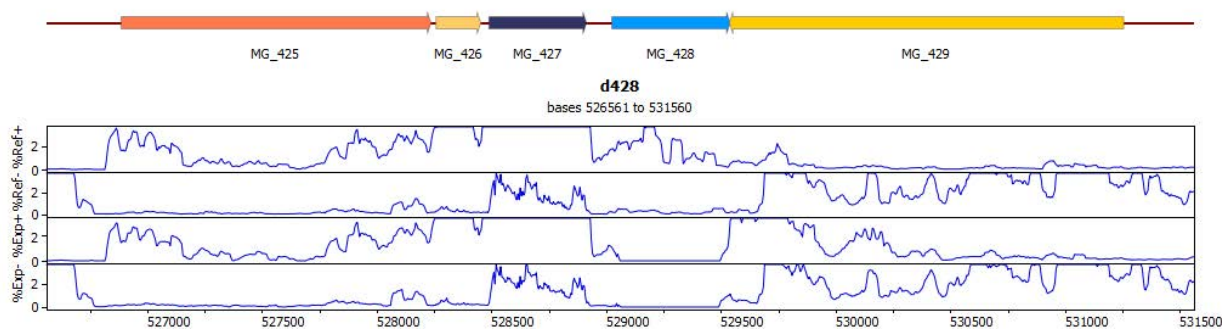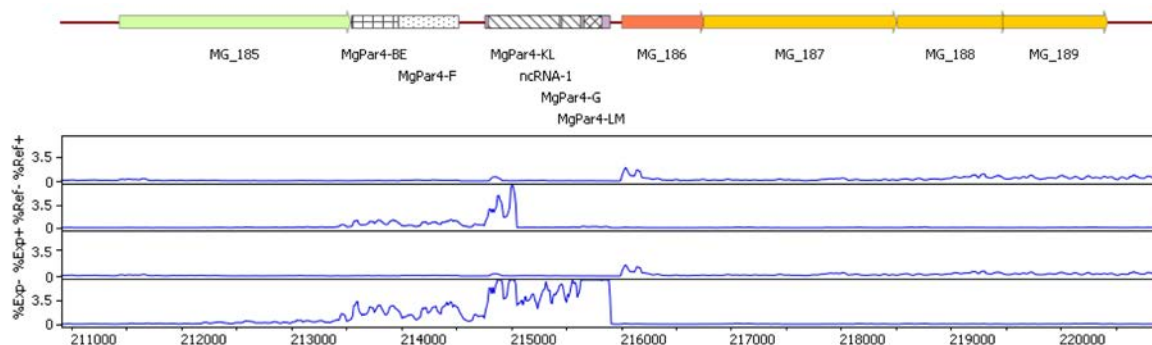

Figure I.9. Plot of relative RNA levels of the MgPar4. The upper panels depict relative coverage of the WT strain in both DNA strands (Ref+ and Ref-). The lower panels represent relative coverage of the Tn::*recA*-2 mutant in both DNA strands (Exp+ and Exp-). MgPar4 is depicted as texture-filled white boxes; while surrounding genes are represented by solid-colored arrows.

The non-coding region designated as ncRNA-1 lies within the MgPar-4 of *M. genitalium* chromosome (Iverson-Cabral *et al.*, 2007) in the reverse strand (Figure I.9). Upon MG428 overexpression, reverse transcript corresponding to the boxes KLM is up-regulated (Table I.5). Possibly, the other boxes of the MgPa region are also altered due to polar effects.

Table I.5. ncRNAs up-regulated in the Tn::*recA*-2 mutant as compared to the WT strain.

ncRNA	Base mean	Log2 Fold Change	Standard Error	p-value
ncRNA-1 (MgPar4/-)	3408.5	5.492	0.181	5.820E-203
ncRNA-2 (MgPar5/+)	850.8	3.761	0.232	2.780E-59
ncRNA-3/4 (MgPar/5-)	6638.7	2.524	0.424	2.672E-09

On the other hand, up-regulation of two different putative ncRNA in the MgPar5 (situated downstream of the cytadherence operon) was also observed upon MG428 overexpression. One of them (ncRNA-2) starts immediately after the MG_192 gene (coding for the p110 adhesin) in the positive strand and covers boxes B-EF of the MgPa region (Figure I.10). RNAseq results showed the presence of a strong intrinsic terminator at the end of the MG_192 gene, so transcription of the ncRNA must start as a consequence of an MG428-regulated promoter. Furthermore, *in vitro* translation of this RNA sequence exposed a putative small peptide present at the very beginning of this ncRNA (see Chapter II). At the same time, another small ncRNA (ncRNA-3/4) is present at the end of the MgPa region KL, in the reverse strand (Figure I.10). In this case, the activation of this ncRNA is not altering the expression of the nearby boxes of the MgPa region, most likely due to the presence of a strong rho-independent transcriptional terminator.
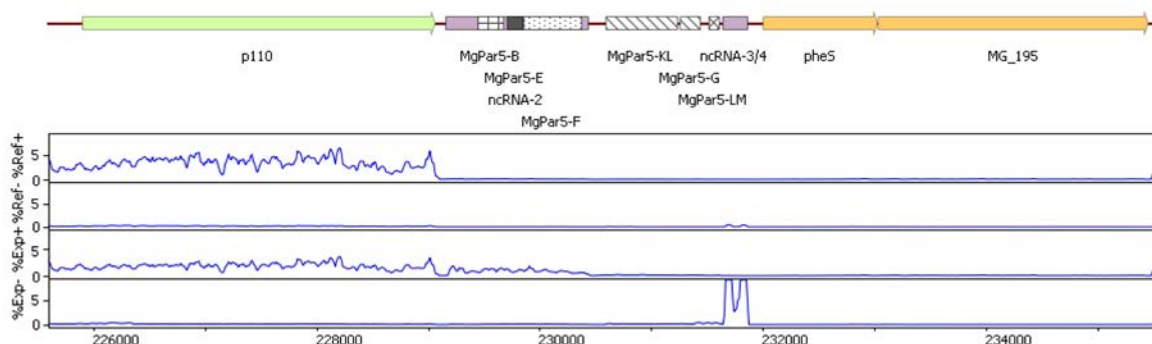
Figure I.10. Plot of relative RNA levels of the MgPar5. The upper panels depict relative coverage of the WT strain in both DNA strands (Ref+ and Ref-). The lower panels represent relative coverage of the Tn::*recA*-2 mutant in both DNA strands (Exp+ and Exp-). MgPar5 is depicted as texture-filled white boxes; while surrounding genes are represented by solid-colored arrows. ncRNA-2 and ncRNA-3/4 are indicated as purple boxes.

The transcription levels of this ncRNA are particularly high and divided in two distinct coverage peaks, suggesting the presence of multiple promoters driving its expression. The role of these two ncRNAs as part of the MG428 regulatory network remains unknown.

Activation of ncRNA-3 is particularly high in the Tn::*recA*-2 mutant (Table I.5). In fact, this activation upon MG428 expression can be detected in the WT strain as compared to the ΔMG_428 mutant (Figure I.11). Expression of ncRNA-3/4 in the ΔMG_428 mutant is almost null, while it can be detected in the WT strain at low levels. On the other hand, no other MG428-regulated ncRNAs nor genes show significant differences between the ΔMG_428 mutant and the WT strain. These results further suggest that MG428 is not active in most of the cells in the WT population at the same time. Since RNAseq allows quantitation of the transcriptome of the population as a whole, it is not possible to detect changes in RNA expression of a small subset of cells as they are masked by the RNA levels of the majority. Therefore, expression levels of the ncRNA-3 must be greatly increased in the MG428-activated cells of the WT population in order to show significant differences in the transcriptome.

Some of these results obtained by RNAseq were confirmed by qRT-PCR analysis of different MG_428 complemented mutants (see Appendices section S.1.3).
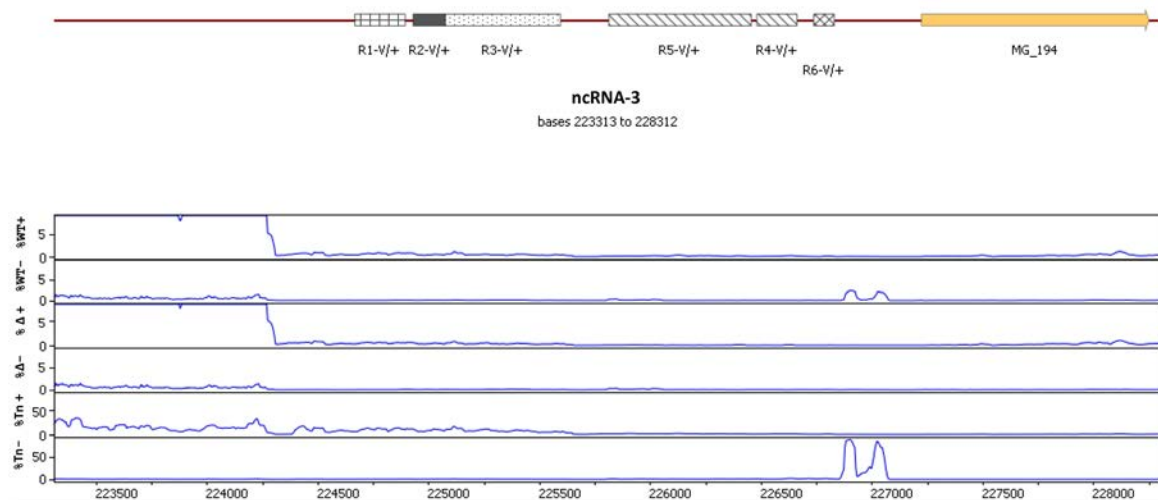
Figure I.11. Plot of relative RNA levels of the MgPar5 comparing WT, ΔMG_428 and Tn::*recA*-2 strains. The upper panels depict relative coverage of the WT strain in both DNA strands (WT+ and WT-). The middle panels represent relative coverage of the ΔMG_428 mutant in both DNA strands (Δ+ and Δ-). The lower panels represent relative coverage of the Tn::*recA*-2 mutant in both DNA strands (Tn+ and Tn-). Percentages of coverage in the WT and the ΔMG_428 mutant are zoomed 10-times with respect to the percentages of the coverage of the Tn::*recA*-2 mutant. MgPa region V is depicted as texture-filled white boxes; while surrounding genes are represented by solid-colored arrows.

I.2.4. MG428 interacts with the RNAP core enzyme

Interaction of the MG428 protein with the *E. coli* RNAP was analyzed by native gel electrophoresis (Figure I.12). A fixed amount of soluble recombinant MG428 protein was mixed with increasing concentrations of RNAP core enzyme, which is free of detectable sigma factors from *E. coli* (Table I.6).
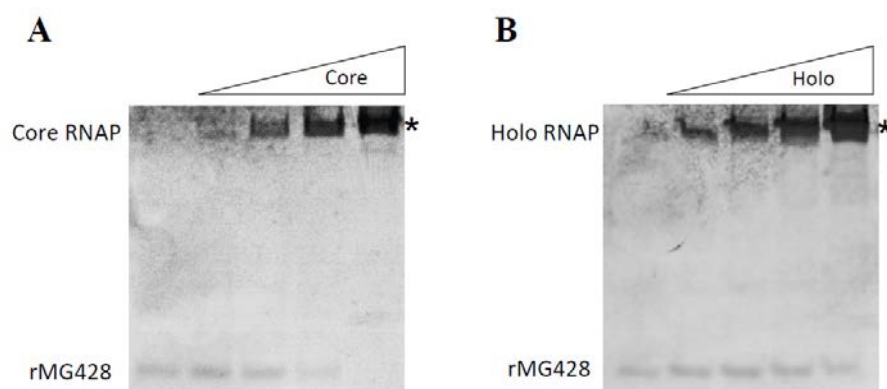
Figure I.12. Analysis of MG428-RNA polymerase interaction. A fixed amount of soluble recombinant MG428 protein (200ng) was incubated with increased concentrations (0, 2.5, 5, 10 and 20mM) of RNAP core enzyme **(A)** or RNAP holoenzyme **(B)**. Mixtures were separated on native discontinuous 4-15% polyacrylamide gels and stained with colloidal Coomassie. Bands indicated with an asterisk (*) were cut off the gel and analyzed by LC-MS (see Table I.6 and Table I.7).

Table I.6. List of identified proteins by LC-MS of core RNAP-MG428 complex (see Figure I.12A)

Accession	Protein	MW [KDa]	pI	Scores	Peptides	SC[a] [%]	RMS90[b] [ppm]	Spec. Counts
P0A8T7|RPOC_ECOLI	DNA-directed RNA polymerase subunit beta'	155.1	6.7	8632.6	93	51.1	1.32	296
P0A8V2|RPOB_ECOLI	DNA-directed RNA polymerase subunit beta	150.5	5	7515.3	90	50.4	1.2	270
P0A7Z4|RPOA_ECOLI	DNA-directed RNA polymerase subunit alpha	36.5	4.8	2372.2	22	50.5	1.79	88
P47667|Y428_MYCGE	Uncharacterized protein MG428	20.2	10.1	889.8	15	47.4	1.61	36
P0A800|RPOZ_ECOLI	DNA-directed RNA polymerase subunit omega	10.2	4.7	544.5	8	85.7	1.45	20
P30128|GREB_ECOLI	Transcription elongation factor GreB	18.5	8.8	304.7	11	42.4	2.7	17
P0ACJ8|CRP_ECOLI	cAMP-activated global transcriptional regulator CRP	23.6	9.2	59.7	5	27.6	2.63	5
P64588|YQJI_ECOLI	Transcriptional regulator YqjI	23.4	6.3	49.1	5	20.8	2.09	5

[a] % of sequence coverage
[b] root mean square in ppm

Table I.7. List of identified proteins by LC-MS of RNAP holoenzyme (see Figure I.12B)

Accession	Protein	MW [KDa]	pI	Scores	Peptides	SC[a] [%]	RMS90[b] [ppm]	Spec. Counts
P0A8T7|RPOC_ECOLI	DNA-directed RNA polymerase subunit beta'	155.1	6.7	20391.3	117	61.4	0.79	557
P0A8V2|RPOB_ECOLI	DNA-directed RNA polymerase subunit beta	150.5	5	18018.4	123	67.4	0.85	535
P0A6F5|CH60_ECOLI	60 kDa chaperonin	57.3	4.7	3719.7	34	57.7	0.93	91
P0A7Z4|RPOA_ECOLI	DNA-directed RNA polymerase subunit alpha	36.5	4.8	3372	27	67.2	0.8	108
P00579|RPOD_ECOLI	RNA polymerase sigma factor RpoD	70.2	4.5	2454.9	39	37	0.88	90
P0A800|RPOZ_ECOLI	DNA-directed RNA polymerase subunit omega	10.2	4.7	1864.3	11	92.3	1	43
P60240|RAPA_ECOLI	RNA polymerase-associated protein RapA	109.7	4.9	894.2	29	30.3	1.41	45
P47667|Y428_MYCGE	Uncharacterized protein MG428	20.2	10.1	417.1	11	42.1	1.28	21
P0AFF6|NUSA_ECOLI	Transcription termination/ antitermination protein NusA	54.8	4.4	378.5	11	23	1.2	15
P0A6J5|DADA_ECOLI	D-amino acid dehydrogenase	47.6	6.2	206.5	7	20.8	153.58	8
P0AFG0|NUSG_ECOLI	Transcription termination /antitermination protein NusG	20.5	6.4	117.2	4	23.8	1.42	5
P30128|GREB_ECOLI	Transcription elongation factor GreB	18.5	8.8	81.1	2	13.3	1	2
P0ADZ4|RS15_ECOLI	30S ribosomal protein S15	10.3	10.8	53.1	2	15.7	1.95	3
P0A9C5|GLNA_ECOLI	Glutamine synthetase	51.9	5.2	44.4	2	3.6	2.14	2
P60422|RL2_ECOLI	50S ribosomal protein L2	29.8	11.6	39.4	2	6.6	2.72	2

[a] % of sequence coverage
[b] root mean square in ppm

At the lowest concentrations of RNAP used, the recombinant MG428 protein and the core enzyme migrated as two discrete bands on the native gel. However, at the highest concentration of the core enzyme used, the band corresponding to the unbound recombinant MG428 protein was not detected, suggesting the formation of a MG428-RNAP complex with an electrophoretic mobility similar to that of the core enzyme alone.

Indeed, LC-MS/MS analysis of this band recovered from the native gel confirmed the presence of all expected RNAP subunits plus the recombinant MG428 protein (Table I.6). In contrast, this interaction was greatly reduced when the recombinant MG428 protein was mixed with increasing concentrations of the RNAP holoenzyme, which consists of the RNAP core enzyme saturated with sigma-70 factor (Figure I.12). Supporting this statement, the band corresponding to the unbound recombinant MG428 protein was clearly present in the native gel at all concentrations of the holoenzyme used.

I.2.5. Identification of a novel sigma-70 promoter sequence in the upstream region of the MG428-regulated genes

A putative promoter sequence with sigma-70 architecture was identified within the upstream region (UR) of the genes or operons up-regulated by the MG428 protein. This putative promoter was composed of two conserved elements of six residues separated by 18 or 19 nucleotides and showed the consensus 5'TTGTCA-N$_{18/19}$-ATTWAT-3' (Figure I.13A). Of note, a conserved sequence with sigma-70 promoter architecture was also recognized immediately upstream of all the *M. pneumoniae* genes homologous to the members of the *M. genitalium* MG428-regulon, except for the MG_412 homolog (MPN611) and the ncRNAs. The conserved sequence identified in *M. pneumoniae* showed the consensus 5'-TTGGCR-N$_{18/19}$-ATTTAT-3' (Figure I.13B). We conducted primer extension analyses of the *recA*, *ruvA*, MG_220, MG_414 and MG_285 genes and the ncRNAs-1, 2 and 3 (as defined in section I.2.3) in the WT strain and the Tn::*recA*-2 mutant (Figure I.14). In repeated experiments, we did not detect transcription initiation within the URs of these genes using RNA from the WT strain (data not shown). In contrast, single, unequivocal transcriptional start sites (TSS) were identified using RNA from the Tn::*recA*-2 mutant. As expected, all the identified TSSs were located immediately downstream of the anticipated promoter sequences (Figure I.13). For the *ruvA* gene, we found that transcription initiated in the promoter region of the MG_RS02200 ORF, which precedes the *ruvAB* genes.

A

```
MG_339        taataaaaaaATGTCAtttttatttttttgcacttATTTATaaggtgtAa
MG_RS02200    gtaaatttatTTGTCA-ctaatagtttttttagcacATTAATttgatagaA
MG_358        tagcagacttTTGTCTaaaggatgatactgaaaaaATTAATcaattctta
MG_220        ttataaaactTTGGCA-cttttagatttagatgtgATTAATagaCaaaag
MG_390        gcttacaaatTTGTCAtttttcacttttcattgctATTAATaagacaaaa
MG_010        aaaaattattTTGACAttaaaagttaaataaagcaATTTATttaacaaat
MG_414        gcaataaaaaTTGTCA-atatttcattaaatacacATTTATttaatgagA
MG_525        agtaaaaaaaCTGTCA-cttaataagtttttatttaATTTATggaatataa
MG_285        aatagtttttTTGACCattttgtttgagattcgtcAATATTagaCaaaag
MG_412        tttaaattaaTTGGCT-tgtgaaaaaactagactaAATAATaacCtgagt
ncRNA-1       gcaatttaatTGTTCA-aattgttaaaaaataacaATTTATagagtgggG
ncRNA-2       tttttgttttTTGTCA-cttttcactaaacttgcaATTTAGagagTggat
ncRNA-3 (I)   ggaacagaagTTGTATaaagtcatgttaatcttttATTTCTgttcccacT
ncRNA-3 (II)  aactatttttTTGTCA-aaaagctttttttgtgcaATTTAGtagatgggA
Consensus               TTGTCA                      ATTWAT
```


B

```
MPN490        atctaaaaaaATGGCGtttttccctttttgaatcaATTTAGatgatgtaa
MPN534        ataaaaaagtTTGGCGctttaaattctccttacacATTCATttagatagg
MPN535        aaaaataactTTGGCAacgaaaaactagtggatttAATTTTccacattgg
MPN313        ttataaaactTTGGCA-ttttactttagttatgtaATTCATttaaatgga
MPN571        cctttttttaTTGGCA-ccttagcctcattgggtcATTTATgtaagtgat
MPN014        tatttttattTTGGCGcttactagttcttgtaaccATTTATtagttaaat
MPN612        ttttaaaaaaTTGGCGtcaattcccttttttgagcaATTTAAtagatgaaa
MPN614        tttctaaaaaTTGGCAagttaaaaactttagttcaATTTATtgaatgtaa
MPN404        aatagtttttTTGACCattttgtctgggaggctccAATTTAtaagacaaa
Consensus               TTGGCR                      ATTTAT
```


Figure I.13. Identification of a conserved region with sigma-70 promoter architecture within the UR of the MG428-regulated genes. (A) Sequence logo generated with all the MG428-regulated genes, operons or small RNAs identified in this study. **(B)** Sequence logo generated with the UR of the respective *M. pneumoniae* homologues, except for the MgPa regions and MG_412 homolog. Nucleotides corresponding to the putative -35 and -10 promoter elements are boxed. Underlined and bolded nucleotides indicate experimentally determined TSSs (see Figure I.14). Due to the different length of the spacer region between the hexanucleotide promoter elements, one nucleotide gap was arbitrarily located after the -35 element when necessary.
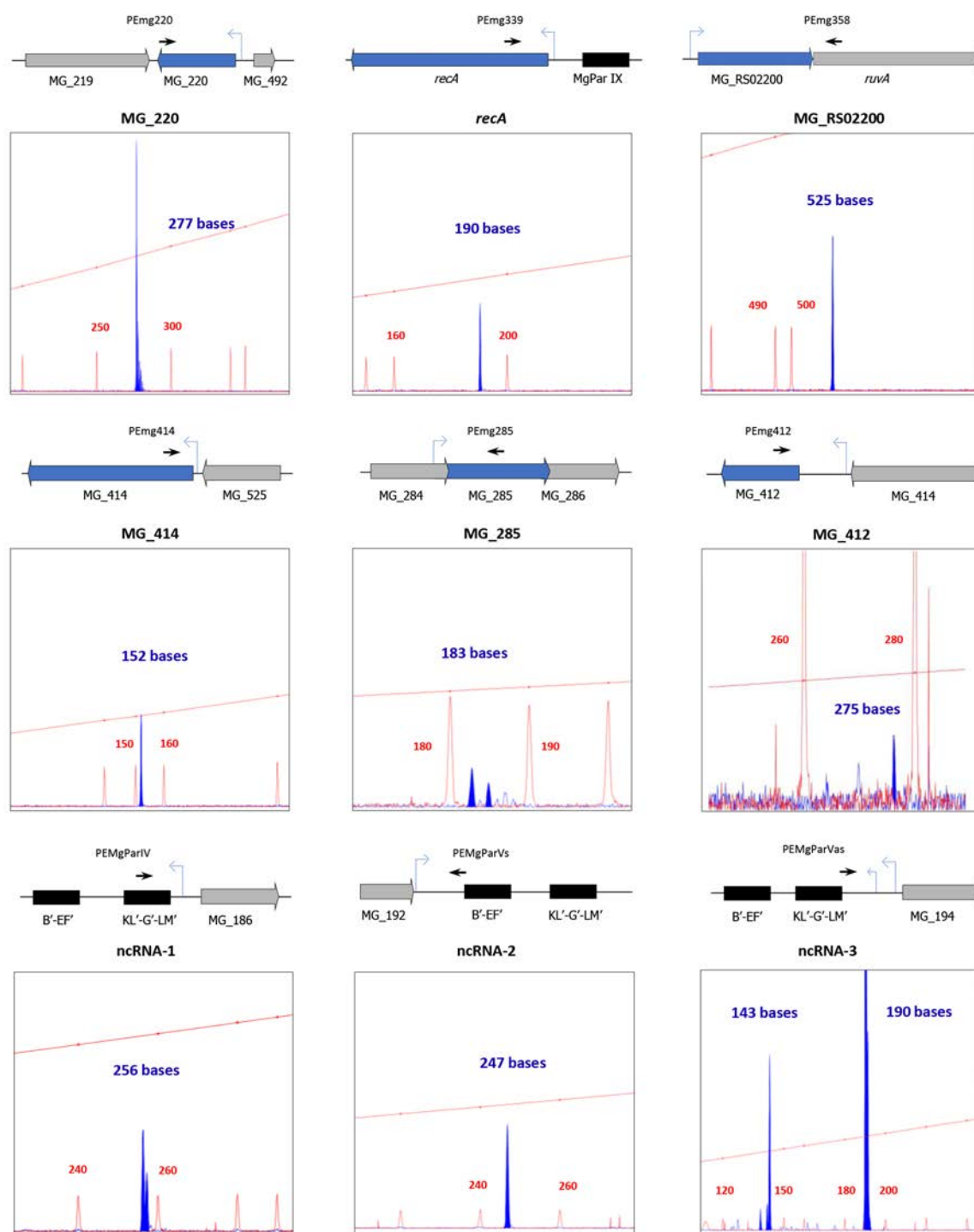
Figure I.14. Identification of the TSSs of several MG428-regulated genes. Primer extension analysis of the MG_220, *recA*, MG_RS02200, MG_285, MG_412 and MG_414 genes and MgPa regions 4 and 5 in the Tn::*recA*-2 mutant. All electropherograms were generated with Peak Scanner v1.0 (Applied Biosystems) analysis software. Red peaks represent ROX size standards while blue peaks correspond to the primer extension products. Schematic representations of the genome regions analyzed are shown and the presence of the identified promoters indicated with blue arrows. The approximate location of the primers used in these analyses is also indicated by arrows.

The specific TSS of MG_RS02200 was confirmed with a second primer complementary of the MG_RS02200 ORF region (Supp. Figure S.4). Therefore, MG428-dependent activation of the *ruvAB* genes is driven from the MG_RS02200 promoter. For the primer extension analysis of the ncRNA-3, two distinct transcriptional start sites were detected (Figure I.14). This result correlates with the RNA levels detected by RNAseq (Figure I.10).

For *M. pneumoniae*, a wide database with genomic, transcriptomic, proteomic and metabolomic databases are accessible online since early 2015 (Wodke *et al.*, 2015). However, TSS of the MG428-regulated orthologue genes are not defined in most cases, probably due to the low expression of these genes in the population. Although the conservation of the -35 element and the alternative Pribnow box is clear (Figure I.13B), TSS of all these genes should be determined in order to validate the activity of the putative alternative promoters in *M. pneumoniae*. A summary of all MG428-regulated genes, their orthologues in *M. pneumoniae* and in other mycoplasmas is listed in Table D.1.

I.2.6. Single cell analysis of MG428 and RecA expression by fluorescence microscopy

To monitor the expression of MG428 and RecA in the population of *M. genitalium*, mutants carrying MG428-, RecA- or a control Cat-mCherry fusion were obtained by allelic exchange (Figure I.15). Several clones of each mutant strain were analyzed by fluorescence microscopy (Figure I.16 and Table I.8).

Table I.8. Frequency of cells with detectable levels of different mCherry protein fusions

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
WT	1172	0	0
Cat:Ch	2249	2009	89.34
RecA:Ch	4529	30	0.66
MG428:Ch	2836	13	0.46
ΔMG_428-RecA:Ch	4551	0	0
RecA:Ch -10	3645	0	0
RecA:Ch -22	3296	19	0.58
RecA:Ch -35	2720	0	0
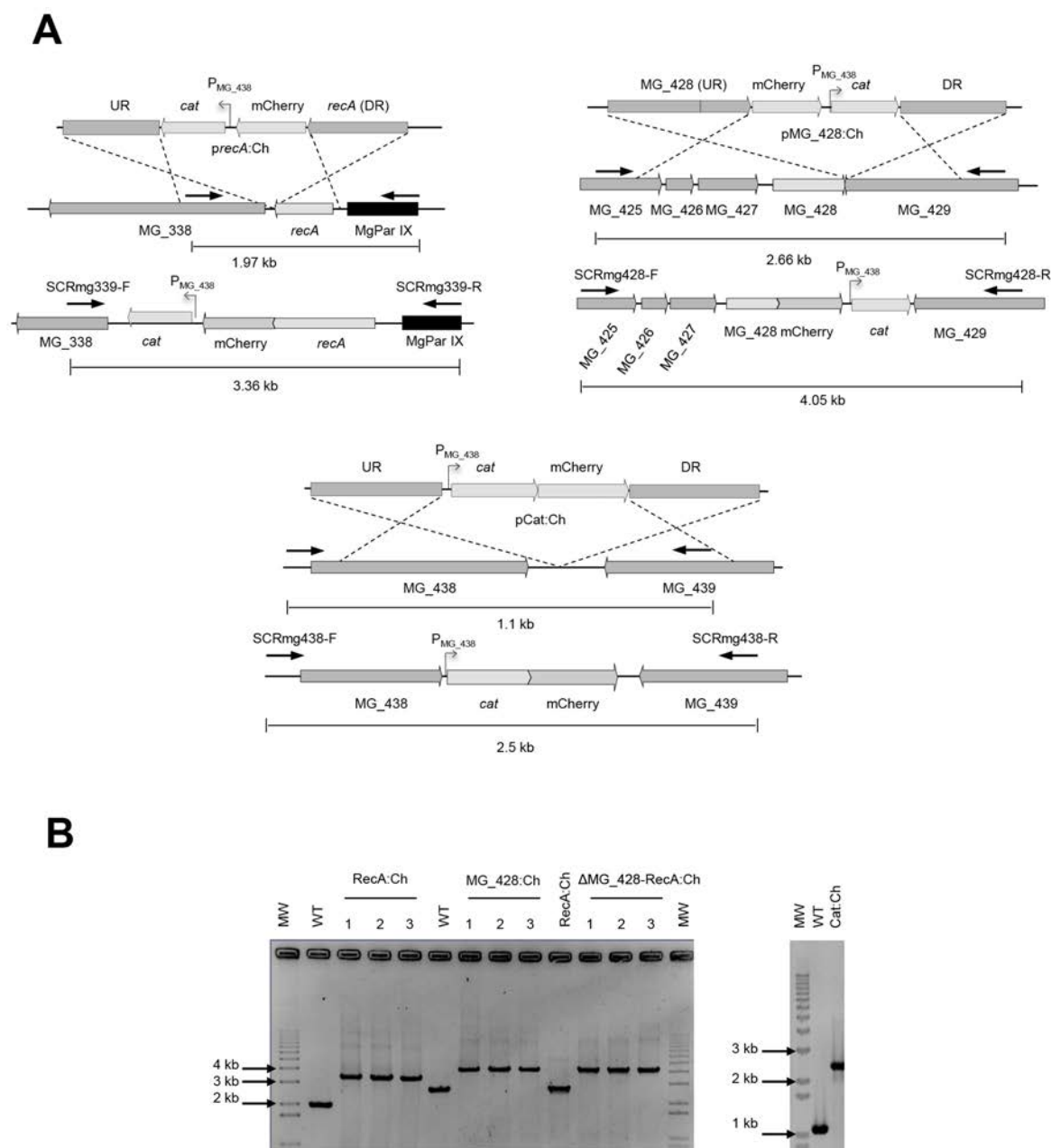
Figure I.15. Construction of *M. genitalium* mutants carrying Cat-, RecA- and MG428-mCherry fusions by allelic exchange. (A) Schematic representation of the construction of the RecA:Ch, MG428:Ch and Cat:Ch mutants by homologous recombination. Arrows indicate the location of the primers used for screening. **(B)** Electrophoresis gels demonstrating the generation of Cat-, RecA- and MG428-mCherry fusions in a wild-type strain background. The presence of a 2.5 kb band in the Cat:Ch mutant, a 3.4 kb band in the RecA:Ch mutant and a 4.05 kb band in the MG_428:Ch mutant was consistent with the intended genome modifications.
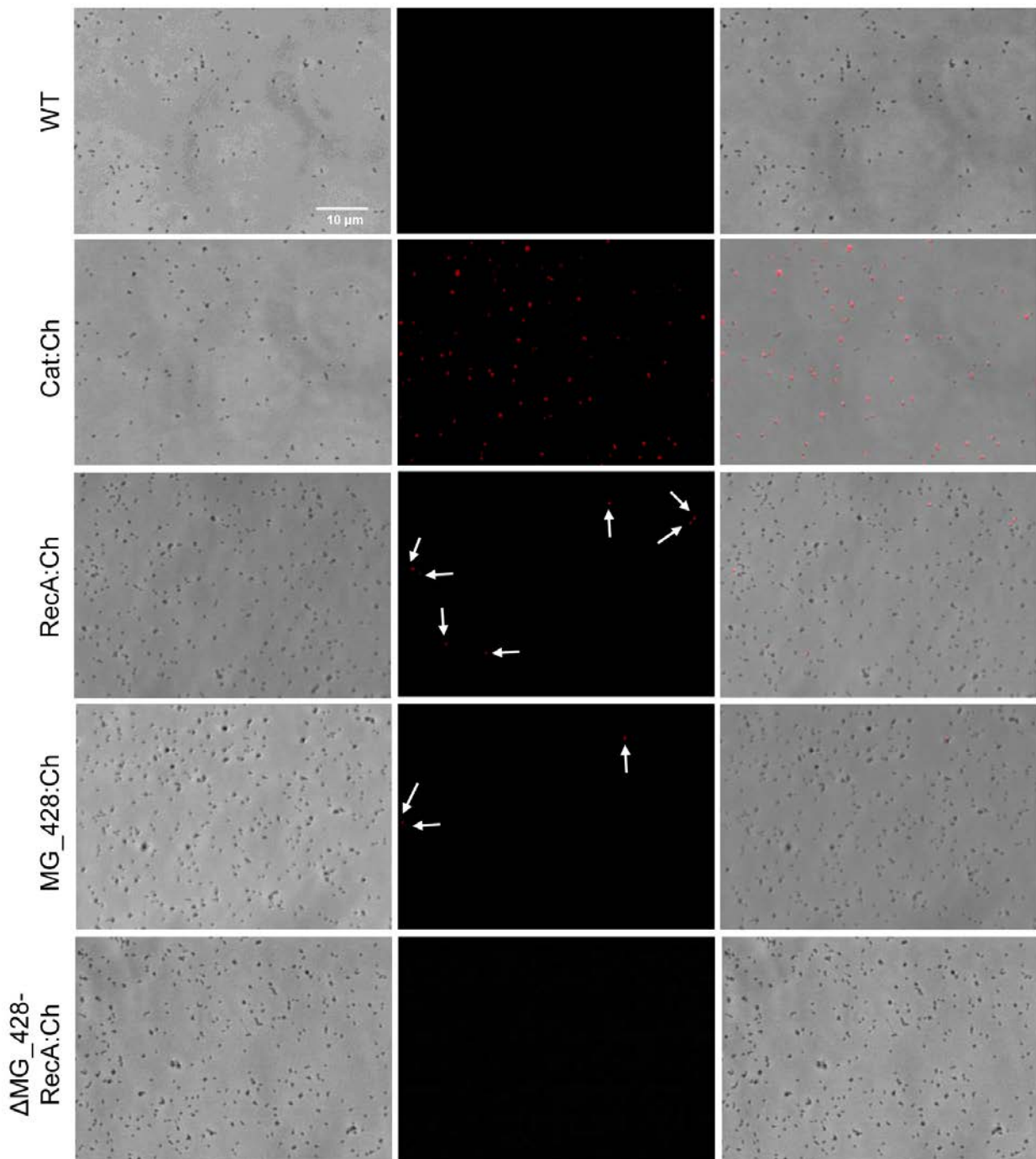
Figure I.16. Single cell analysis of Cat-, RecA- and MG428-mCherry expression by fluorescence microscopy. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. Arrows indicate the presence of fluorescent cells expressing either RecA- or MG428-mCherry fusions. All the pictures are shown at the same magnification.

As expected, expression of the Cat-mCherry fusion under the control of an endogenous *M. genitalium* promoter was detected in virtually the entire population. In contrast, only a 0.46% of the cells carrying the MG428-mChery fusion exhibited detectable fluorescence. Likewise, the RecA-mCherry fluorescence was detected in a 0.66% of the population. Interestingly, mCherry fluorescence was usually observed in cell pairs (Figure I.16), suggesting a spatial association between mycoplasmas with increased levels of MG428 or RecA. In addition, we explored RecA-mCherry expression in a ΔMG_428 mutant background. This strain was obtained by deleting the MG_428 gene from the mutant carrying the RecA-mCherry fusion (Figure I.15B). We could not detect any fluorescent cell in any of the clones analyzed, demonstrating that the presence of MG_428 gene was necessary for RecA expression at high levels.

On the other hand, we also created several RecA:Ch mutants with scrambled sequences centered at the -10, -22 and -35 regions of the *recA* promoter (Figure I.17). No fluorescent cells were observed in the RecA:Ch -10 or the RecA:Ch -35 mutants (Table I.8). This result demonstrates that the identified -10 and -35 promoter elements are essential for MG428-dependent activation of *recA*. In contrast, mutation of a non-conserved region located between the -10 and -35 *recA* promoter elements (mutant RecA:Ch -22) had very little or no impact on the percentage of fluorescent cells (0.58%).
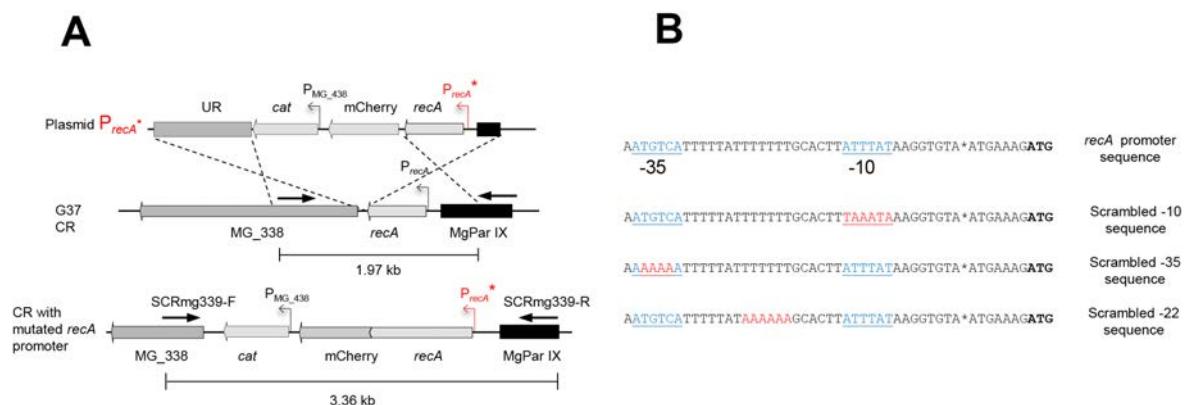
Figure I.17. Construction of RecA:Ch mutants with scrambled sequences within the *recA* promoter. (A) Scheme depicting the construction of RecA:Ch mutants carrying a modified *recA* promoter (highlighted in red) by allelic exchange. Transformants were screened for by PCR and sequencing of the resulting PCR products to confirm the presence of both the *recA*-mCherry fusion and the scrambled promoter sequences. **(B)** Sequence of *recA* promoter region in the WT strain and the RecA:Ch -10, RecA:Ch -22 and RecA:Ch -35 mutants. In blue, WT sequence; in red, scrambled sequence.

Of note, the RecA-mCherry fusion was detected by Western blot in all the tested mutants, including those mutants for which no fluorescent cells were observed (Figure I.18).

Therefore, basal levels of RecA expression can be detected by Western blot but not by fluorescence microscopy. In agreement with previous data, expression of the MG428-mCherry fusion was not detected by Western blot with anti-mCherry antibodies.
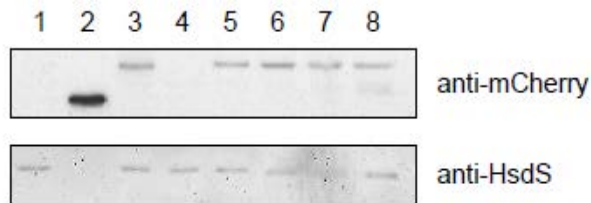
Figure I.18. Analysis of mCherry expression by Western blot. Lane 1, WT; lane 2, Cat:Ch; lane 3, RecA:Ch; lane 4, MG_428:Ch; lane 5, ΔMG_428-RecA:Ch; lane 6, RecA:Ch-10; lane 7, RecA:Ch-22 and lane 8, RecA:Ch-35. Because the Cat-mCherry fusion is expressed at very high levels compared to the RecA-mCherry fusion, the amount of total protein loaded for the Cat:Ch mutant was reduced 20 times. HsdS protein was detected with a monoclonal antibody and used as a loading control.

I.2.7. Quantitative assessment of the recombination capacity

As we described earlier, several genes up-regulated by the MG428 protein coded for proteins with a known role in homologous recombination (HR). This fact prompted us to investigate the capacity of the ΔMG_428 mutant and some complemented strains to be genetically modified by HR (Figure I.19 and Figure I.20). The transformation efficiency by homologous recombination of the WT strain with the suicide plasmid pΔMG_218-lacZ was ~1×10^{-6} transformants per viable cell. In contrast, repeated attempts to transform the ΔMG_428 strain by HR with the same plasmid failed.
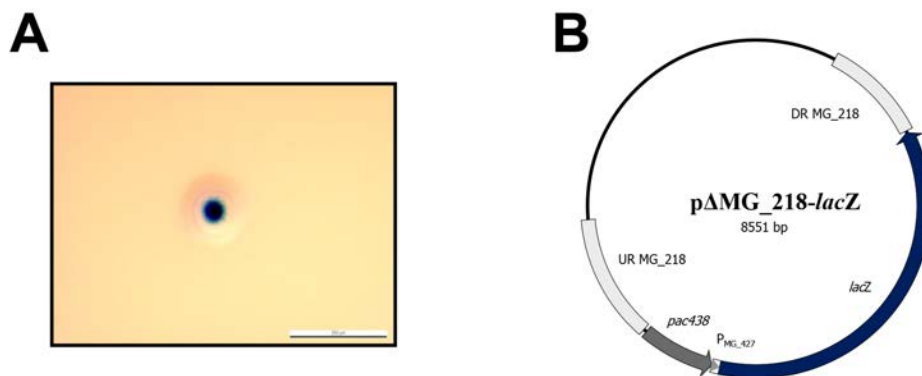
Figure I.19. β-galactosidase activity of *M. genitalium* colonies electroporated with the suicide plasmid pΔMG_218-lacZ. (A) Typical colony developed after transformation with the pΔMG_218-lacZ plasmid. The blue color in the center appears as a result of the β-galactosidase activity. **(B)** Schematic representation of the pΔMG_218-lacZ plasmid. Construction of this plasmid is described in the Appendices section.

Similarly, no transformant colonies could be observed after electroporation of the mutants carrying the Tn*CatMG_428* MiniTnp inserted within the *recA* or the *ruvB* genes (Figure I.20).

Instead, the presence of a WT copy of the MG_428 allele in the Tn::MG_281-1 mutant fully restored the capacity of this strain to be modified by HR (Figure I.20).
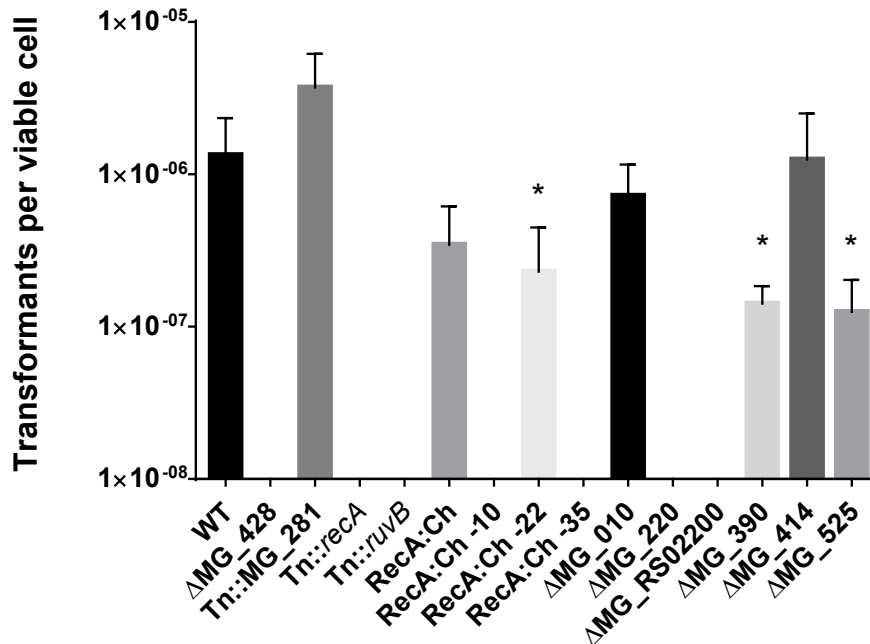
Figure I.20. Transformation efficiency by homologous recombination of different _M. genitalium_ strains. Bars show the averages and the standard deviations of at least three independent biological replicates for each strain. Transformation efficiency by HR of strains ΔMG_428, Tn::_recA_-2, Tn::_ruvB_, ΔMG_220, ΔMG_RS02200, RecA:Ch -10 and RecA:Ch -35 was below the limit of detection ($<10^{-8}$). Transformation efficiency by HR that were found to be statistically different than that of the WT strain are indicated with an asterisk. Statistical significance was assessed with Student's _t_ test ($P<0.05$).

Therefore, our data indicate that deletion of the MG_428 gene leads to severe recombination defects that can be complemented _in trans_ by the reintroduction of an ectopic copy of this gene.

On the other hand, we constructed null mutants of the MG_010, MG_220, MG_390, MG_414, MG_525 and MG_RS02200 genes by allelic exchange (see Appendices section 0) and assessed their capacity to be modified by HR (Figure I.20). Transformation efficiency by HR of the ΔMG_010 and ΔMG_414 mutants was similar to that of the WT strain. However, mutants lacking the MG_390 or MG_525 genes showed a ten-fold transformation efficiency by HR reduction as compared to the WT.

Furthermore, repeated attempts to transform the ΔMG_220 or ΔMG_RS022000 mutants by HR were unsuccessful. Of note, all these mutants were transformed with a Tn_Cat_ MiniTnp

with efficiencies similar to that of the WT strain, indicating that DNA entry was not compromised (data not shown).

The recombination capacity of the mutants expressing a RecA-mCherry fusion was also explored (Figure I.20). Transformation efficiency by HR of the RecA:Ch mutant was comparable to that of the WT strain. Moreover, transformation of the RecA:Ch -22 mutant with the pΔMG_218-lacZ plasmid was similar to that of the RecA:Ch mutant. In contrast, the RecA:Ch -10 and RecA:Ch -35 mutants could not be transformed by HR.

I.2.8. Analysis of genome variation by next-generation sequencing

HR plays an important role in the generation of genetic variants in the population (Vink *et al.*, 2012). Therefore, since MG428 functions as a novel regulator of HR in *M. genitalium*, we investigated whether high levels of MG428 expression correlated with increased genetic diversity. We found that DNA repeats of several MgPa regions exhibited high nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the WT strain or the Tn::MG_390-1 mutant (Figure I.21). Essentially, MgPa regions consist of repeated sequences that are homologous, but not identical, to sequences within the MG_191 and MG_192 genes. Similarly, DNA repeats within the MG_191 gene also exhibited high nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the two control strains (Figure I.21). Therefore, the relatively high levels of MG428 expression in the Tn::MG_192-1 mutant (Figure I.6) correlate with increased nucleotide sequence variation within DNA repeats. This result is in agreement with a recent report describing the possible function of the MG428 protein in the generation of antigenic diversity in *M. genitalium* (Burgos and Totten, 2014b).
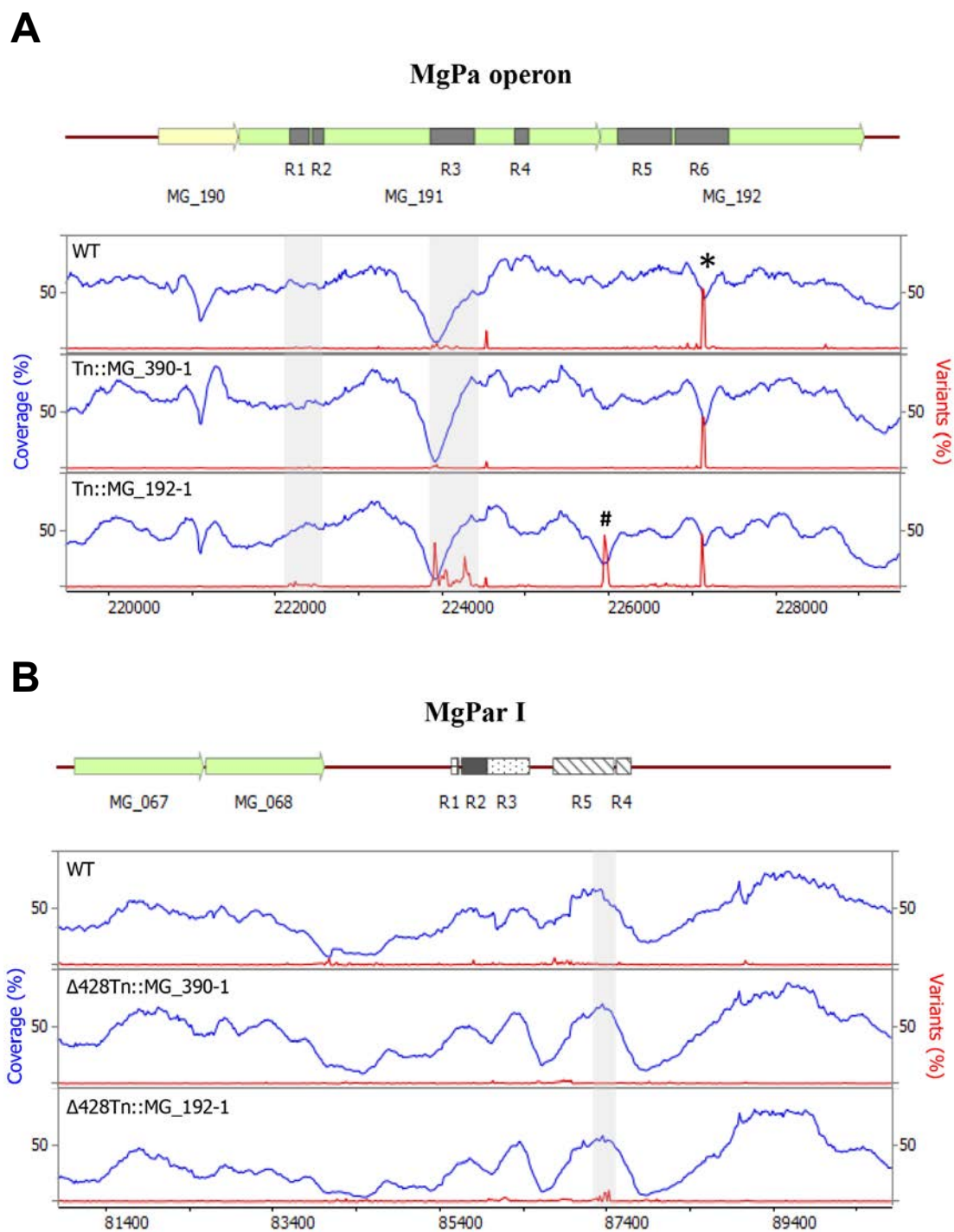
A

B

Figure I.21. Whole genome analysis of selected MG_428 complemented mutants by next generation sequencing. (*Continued on the next page*)
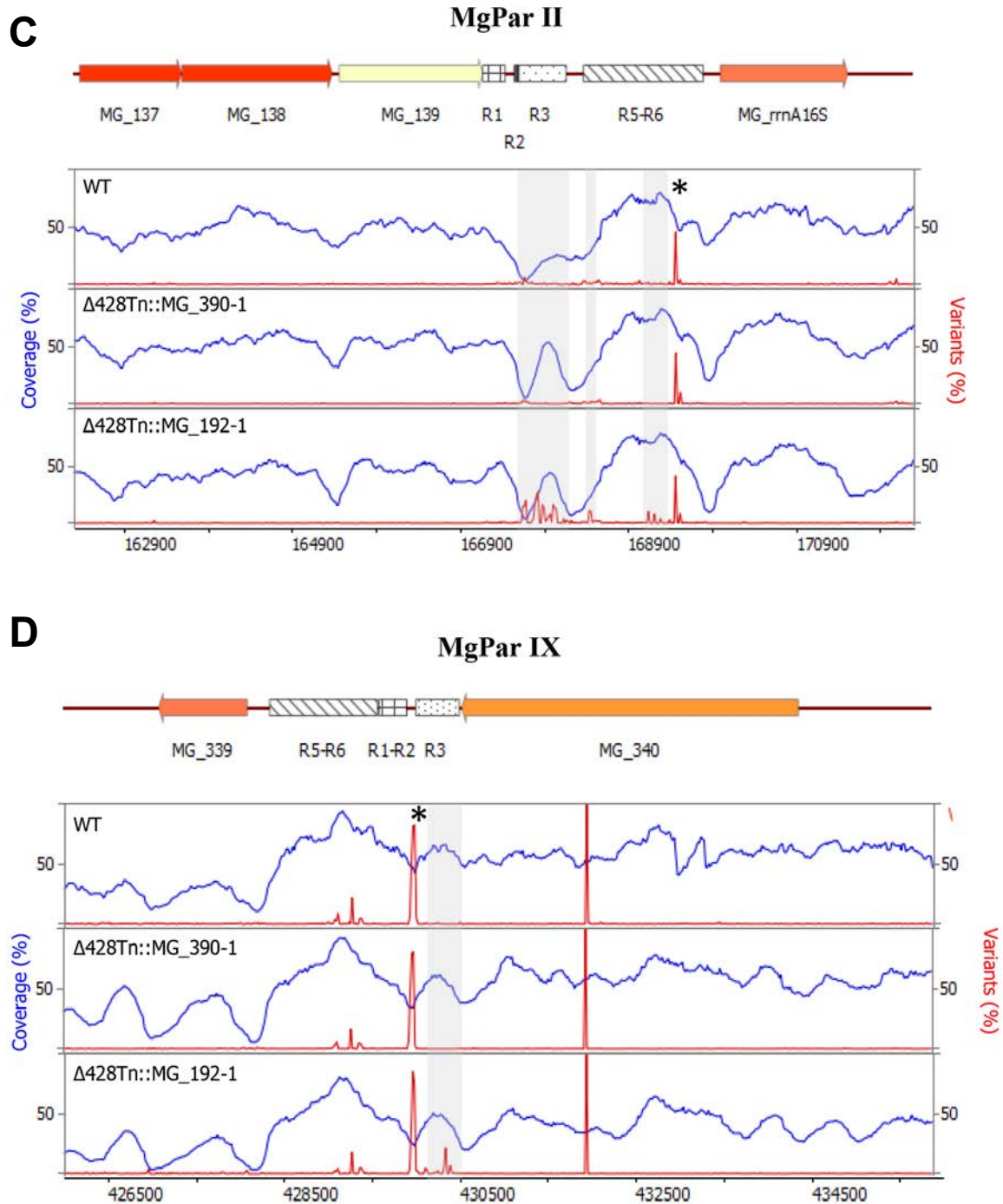
Figure I.21. Whole genome analysis of selected MG_428 complemented mutants by next generation sequencing. Schemes of the *M. genitalium* MgPa operon (A), MgPar1 **(B)**, MgPar2 **(C)** and MgPar9 **(D)** are shown along with its respective coverage data (in blue) and variant frequencies (in red). Discrete DNA repeats are boxed and numbered (R1 to R6). Genomic regions exhibiting a high rate of nucleotide variation in the Tn::MG_192-1 mutant as compared to the WT strain are shadowed. Variations labelled with an asterisk (*) correspond to length polymorphisms of a trinucleotide repeat. A variation labelled with a pound sign (#) corresponds to the insertion point of the Tn*CatMG_428* MiniTnp in the Tn::MG_192-1 mutant.

As showed in Figure I.21, frequency variants in the transposon insertion point (#) is ~50%.

A plausible explanation is that reads of the positive strand show only sequence variants

downstream of the transposon insertion point. Conversely, reads of the negative strand show only sequence variants upstream of the transposon insertion point. Since the graphic coverage shows combined reads from both the positive and the negative strand, this leads to a frequency of variation of about 50% in the regions flanking the transposon insertion site.

I.3. Discussion

Herein, we show that MG428 is a master regulator of recombination in *M. genitalium*. Transcription of the genes coding for RecA and the Holliday junction resolvases RuvA and RuvB was found to be extensively up-regulated by the MG428 protein. These three enzymes are nearly ubiquitous in bacteria, evidencing their fundamental role in HR (Rocha *et al.*, 2005). However, recombination requires the participation of other enzymes that are less conserved across bacterial species. For example, the RecBCD and RecFOR protein complexes provide two complementary pathways to load RecA onto ssDNA (Kowalczykowski *et al.*, 1994). Although several mycoplasma species possess homologs of the recombination genes *recD*, *recO* and *recR*, these genes appear to be absent in the chromosomes of *M. genitalium* and *M. pneumoniae* (Fraser *et al.*, 1995; Himmelreich *et al.*, 1996). Interestingly, we found that up to seven genes with unknown function were under the control of the MG428 protein (Table I.9). Deletion of two of these unknown genes (MG_220 and MG_RS02200) abrogated the capacity of *M. genitalium* to be genetically modified by HR. Moreover, our results indicate that the MG_390 and MG_525 genes may also participate in the recombination pathway of this bacterium.

Table I.9. Summary of experimentally validated MG428-regulated genes and function.

Gene/region	Annotation	Function	Mean transcript fold-increase[a]
MG_220	HP	Unknown, recombination	~43.1
ncRNA-1	-	Unknown	~42.2
MG_339	*recA*	Recombination	~17.3
MG_358	*ruvA*	Recombination and repair	~16.7
MG_359	*ruvB*	Recombination and repair	~15.6
ncRNA-2	-	Unknown	~13.6
MG_285	HP	Unknown	~13.2
MG_286	HP	Unknown	~12.5
MG_414	HP	Unknown	~12.2
MG_RS002200	HP	Unknown, recombination	~10.8
ncRNA-3/4	-	Unknown	~5.8
MG_389	HP	Unknown	~4
MG_412	*pstS*	Phosphate binding lipoprotein (putative)	~2.6
MG_390	*sunT*	Peptide secretion (putative)	~2.2
MG_010	HP	Replication (putative)	~2
MG_525	HP	Unknown	~1.7

[a] Data extracted from RNAseq analysis

This opens a new avenue in the understanding of recombination in this model organism, as some of these unknown genes may represent a novel set of recombination enzymes. Of note, inactivation of two MG428-regulated genes (MG_414 and MG_010) had very little or

no impact in the recombination capacity of *M. genitalium*, indicating that MG428 likely regulates other biological functions beyond recombination.

Unexpectedly, we found that insertion of a MiniTnp that leads to MG428 overexpression was highly selective for genes under the control of the MG428 protein. To our knowledge, this straightforward strategy to identify regulated genes is unprecedented and relies on the potential toxicity observed upon MG428 overexpression. In a previous report, Glass *et al*. described that 31% of the mutants isolated in a global transposon mutagenesis study carried transposons inserted within the MG_428, MG_339, MG_414 and MG_525 genes (Glass *et al.*, 2006). That is, three genes under the control of the MG428 protein and the MG_428 gene itself. In addition, the authors stated that mutants with a disrupted MG_414 or MG_525 genes grew significantly faster than the WT strain. These results were later confirmed in a comprehensive study comparing growth rates of different *M. genitalium* single-gene mutants (Sanghvi *et al.*, 2013). Altogether, these findings suggest that inactivation of MG_428 or specific MG428-regulated genes generate mutant cells with better fitness in vitro. Conservation of genes that are apparently detrimental for in vitro growth of this minimal bacterium, advocates for an important role of the MG_428 regulatory pathway in survival within the host. Work is in progress to identify other genes under the control of the MG428 protein using high throughput technologies.

In agreement with the low cellular abundance characteristic of transcriptional regulators, MG428 expression could not be detected by western blot in the WT strain. However, we found that MG428 was expressed at relatively high levels in a small subpopulation of cells. Similarly, only a few mycoplasmas exhibited detectable fluorescence upon the expression of a RecA-mCherry fusion protein from its native chromosomal locus. This is in contrast to observations in *E. coli*, where the majority of cells expressing a RecA-GFP fusion exhibited detectable fluorescence under standard laboratory conditions (Renzette *et al.*, 2005). Moreover, RecA-mCherry fluorescent cells were only observed in the presence of the MG_428 gene, indicating that *recA* activation is primarily under the control of the MG428 protein. Altogether, these data indicate that recombination is cell-specific in *M. genitalium* and strictly dependent on the MG428 protein. At the present time, the factors governing the activation of the MG428 regulatory pathway are unknown. Interestingly, it has been described that transcript levels of the *Saccharomyces cerevisiae* Rad51 gene, which codes for the functional homolog of RecA in eukaryotes, fluctuate throughout the cell cycle (Basile *et al.*, 1992). Activation of the MG428 regulatory pathway could also be coordinated with the cell division cycle in *M. genitalium* and accordingly, we observed a spatial

association between mycoplasmas with high levels of MG428 or RecA expression. Alternatively, MG428-activation could be linked to a yet unidentified mechanism of cell-to-cell DNA transfer in *M. genitalium*. Of note, conjugal transfer with the incorporation of the incoming DNA into the recipient chromosome by homologous recombination has been recently described in *Mycoplasma agalactiae* (Dordet-Frisoni *et al.*, 2014). The existence of a mechanism promoting horizontal gene transfer and the participation of the MG428-regulated proteins in this process would reinforce a role of MG428 in genome evolution and adaptation in *M. genitalium*.

Aside from the insight into the biological function of the MG428 protein, our data provide compelling evidence that this regulator is an alternative sigma factor of *M. genitalium*. Accordingly, we found that MG428 interacts with RNAP and that this binding is markedly reduced in the presence of the primary sigma-70 factor RpoD. Furthermore, we found that MG428 directs RNAP to specific promoter regions of *M. genitalium*. Of note, MG428-dependent promoters display a novel Pribnow box and a conserved −35 element. The identified −35 element resembles the canonical TTGACA box, which is rarely found in mycoplasma promoters (Weiner *et al.*, 2000). Unlike previous reports demonstrating that the -35 region is of minor importance for promoter function in *M. genitalium* and *M. pneumoniae* (Halbedel *et al.*, 2007; Zhang and Baseman, 2011a), herein we show that this element is essential for the activation of MG428-dependent promoters. Similarly, the divergent Pribnow box (ATTWAT) identified in this study is also necessary for MG428-regulation. Altogether, our findings demonstrate that transcription in *M. genitalium* is under the control of at least two sigma factors that direct RNAP to different promoter regions.

Overall, the results presented in this work challenge the current simplistic perception of gene regulation in mycoplasmas and demonstrate the existence of unique regulatory networks in these bacteria. Furthermore, since mycoplasmas are extremely refractory to genetic modification by homologous recombination, the identification of a protein stimulating recombination could greatly facilitate the study of these minimal cells.

CHAPTER II

RrlA and RrlB are two proteins that stabilize the alternative sigma factor σ^{20} of *M. genitalium*

In a recent work, we have demonstrated that antigenic variation in the human pathogen *Mycoplasma genitalium* is promoted via an alternative sigma factor (MG428 or σ^{20}) that is active in an intermittent fashion *in vitro*. σ^{20} activates transcription of MG_220 and MG_RS02200 (*rrlA* and *rrlB*), two genes coding for hypothetical proteins with a key role in homologous recombination. In this work, we have demonstrated that both proteins play an important role in σ^{20} stabilization and, therefore, allow a feed-forward loop that enhances σ^{20} activity. Despite this crucial role, recombination capacity of null mutants of *rrlA* and *rrlB* genes could be complemented *in trans* using great amounts of σ^{20} protein. Furthermore, we describe an unorthodox DNA transfer mechanism among *M. genitalium* cells that is strictly dependent on σ^{20} activation and the recombination machinery of this bacterium. DNA was horizontally transferred from a strain overexpressing the σ^{20} regulon to a recipient strain with the σ^{20} pathway unaltered. Altogether, we provide further knowledge regarding the circuitry controlling σ^{20} stabilization and also demonstrate a novel form of horizontal gene transfer in *M. genitalium*, which could direct genome plasticity and genetic variation in an unprecedented way in Mycoplasmas.

II.1. Introduction

In many bacteria, recombination is a tightly regulated process that involves a large genetic network and has several regulatory checkpoints (Cox, 2007). At the transcriptional level, the LexA repressor can activate genetic recombination in response to DNA insults in a RecA-directed autoproteolysis. Remarkably, Mycoplasmas are devoid of the classic transcriptional regulators controlling the main cellular processes, and LexA is no exception. In a recent report, we identified an alternative factor MG428 (herein designated as σ^{20}) controlling recombination in *Mycoplasma genitalium* (Torres-Puig *et al.*, 2015). σ^{20} regulates transcription of recombination genes such as *recA* (MG_339), *ruvA* (MG_358), *ruvB* (MG_359) and several genes of unknown function. The σ^{20} regulon is activated in a small subset of cells in the bacterial population and the factors triggering σ^{20} activation have not been identified. This feature complicates the characterization of σ^{20} because many of the phenotypic effects cannot be detected when the whole population is analyzed, imposing single cell analyses as the only suitable techniques to study σ^{20} function. In addition, despite that overexpression of σ^{20} using strong promoters overcomes the activation in a small subset of cells and allows studies with the whole bacterial population, it is markedly toxic since it induces a robust hyperrecombinogenic state in this bacterium.

The recombination machinery of *M. genitalium* is limited to RecA, the Holliday junction resolvases RuvAB/RecU and the ssDNA binding protein Ssb (Sluijter *et al.*, 2008, 2009, 2012; Estevao *et al.*, 2011). A recent study indicates that recombination has a minor role in DNA repair (Burgos *et al.*, 2012). Proteins related to the RecA-loading pathways RecBCD or RecFOR are absent, as well as other regulatory proteins of the RecA function (Carvalho *et al.*, 2005). Nonetheless, in our previous work, we identified two novel genes involved in recombination, designated as **r**ecombination **r**egulatory **l**oci *rrlA* (MG_220) and *rrlB* (MG_RS02200). Null mutants of these genes had serious recombination defects, similar to those observed in the σ^{20} defective mutant.

M. genitalium generates antigenic variation by a unique mechanism involving homologous recombination (Iverson-Cabral *et al.*, 2007). The chromosome of this bacterium comprises nine repeats, designated MgPar, that contain sequences with homology to the genes coding for the main cytadhesins. The activity of the alternative sigma factor σ^{20} is a major determinant for the generation of genetic variants of these antigens (Burgos and Totten, 2014b; Torres-Puig *et al.*, 2015), which play an important role in survival and persistence within the host (Wood *et al.*, 2013). However, genetic variation via homologous

recombination in Mycoplasmas is not only present in *M. genitalium*. A recent study on several strains of *M. agalactiae* described the generation of chromosomic chimeras using both the enzymatic machinery of integrative conjugative elements (ICE) and homologous recombination (Dordet-Frisoni *et al.,* 2014). This unconventional horizontal gene transfer mechanism has only been observed in mycoplasma species encoding ICE or by means of membrane fusion-inducing agents, like PEG.

Overall, in the current study we provide further knowledge regarding the function of σ^{20} in *M. genitalium*. We identified several factors controlling σ^{20} activation and modulating recombination in this bacterium. We suggest that RrlA and RrlB play a role in σ^{20} stabilization and control its activity towards the transcription of the σ^{20} regulon. Moreover, we linked the activation of the σ^{20} regulon to a gene transfer system mediated by the recombination machinery. This novel function of the σ^{20} pathway defies known genetic variation mechanisms and might be a valuable tool for genome evolution in this small pathogen.

II.2. RESULTS

II.2.1. Expression of RrlA, RrlB and ORF192.1 in single cells

As described earlier, the *rrlA* and *rrlB* loci code for two novel proteins necessary for recombination. However, so far, the expression of RrlA and RrlB in *M. genitalium* has not been demonstrated. Similarly, the expression of the ORF192.1 identified in this work has not been established either. To this end, we obtained mutants carrying transcriptional fusions of *rrlA*, *rrlB* or *orf192.1* to the *mcherry* fluorescent marker at their respective native loci (Figure II.1). As previously observed for other σ20-regulated genes (Torres-Puig *et al.*, 2015), only a small percentage of the mutant population displayed mCherry fluorescence (Figure II.2 and Table II.1). Supporting the σ20-controlled expression of *rrlA*, *rrlB* and *orf192.1*, we found that deletion of the MG_428 gene completely abrogated mCherry fluorescence in all cases (Figure II.2 and Table II.1).
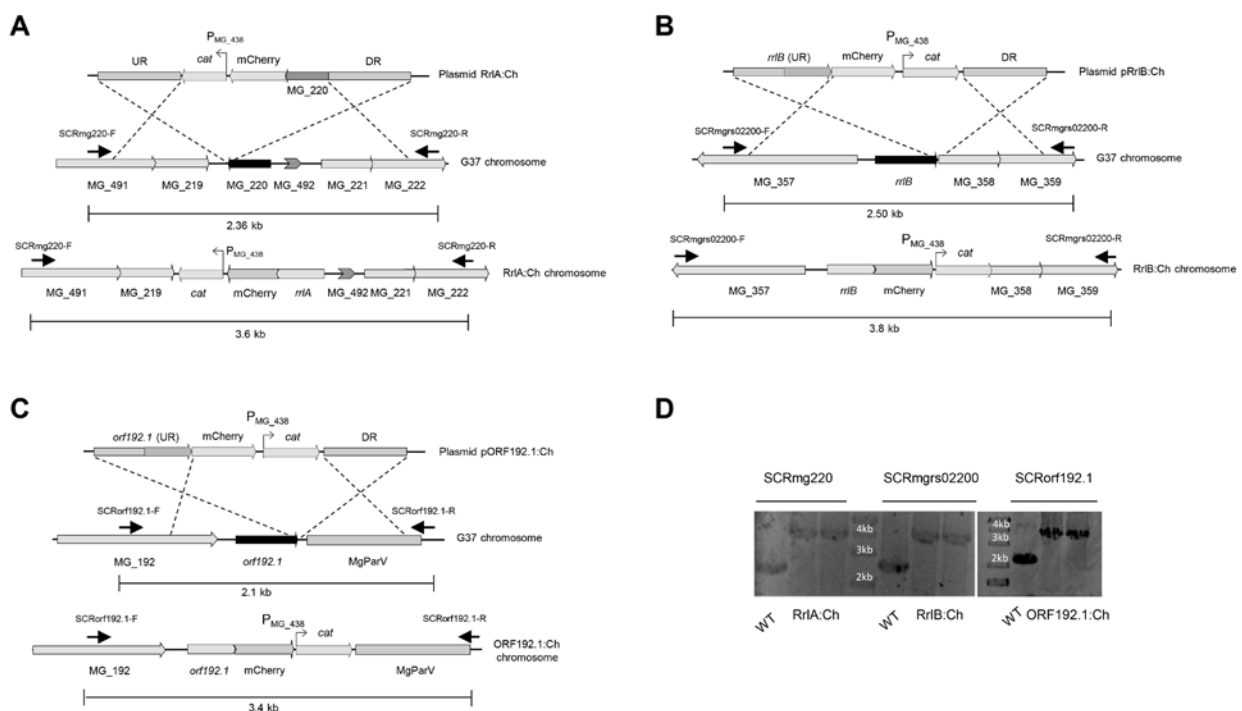
Figure II.1. Construction of *M. genitalium* mutants carrying RrlA-, RrlB- and ORF192.1-mCherry fusions by allelic exchange. Schematic representation of the construction of the RrlA:Ch **(A)**, RrlB:Ch **(B)** and ORF192.1:Ch **(C)** mutants by homologous recombination. Arrows indicate the location of the primers used for screening. **(D)** Electrophoresis gels demonstrating the generation of RrlA-, RrlB- and ORF192.1-mCherry fusions in a wild-type strain background. The presence of a 3.6kb band in the RrlA:Ch mutant, a 3.8kb band in the RrlB:Ch mutant and a 3.4kb band in the ORF192.1:Ch mutant was consistent with the intended genome modifications.

Next, we reintroduced the MG_428 gene to the mutants lacking σ20 described above by transposon delivery. Transcription of the transposon-encoded copy of the σ20 gene was driven by the MG_427 promoter. Our transcriptional data shows that the MG_427 promoter is not regulated by σ20 and strains carrying transcriptional fusions of the MG_427 promoter to the *mcherry* marker show a bright, sharp fluorescence that is observed in virtually the entire population (Figure II.2). Upon σ20 overexpression, mCherry fluorescence was restored to all mutants and the percentage of fluorescent cells observed was above that of the strains with wild-type levels of σ20 expression (Figure II.2 and Table II.1). The different percentages of fluorescent cells observed in the complemented strains likely correlate with the levels of σ20 expression, which in turn depend on the context of the transposon insertion within the chromosome. Of note, the majority of cells observed did not exhibit mCherry fluorescence. This was not anticipated because σ20 expression was expected to be strong and constitutive with the MG_427 promoter. This finding suggests that transcription of the target genes is still regulated in the complemented strains. Therefore, either the expression of *rrlA*, *rrlB* and orf192.1 is under the control of additional factors beyond σ20 or the activity of σ20 is regulated at the post-transcriptional level in *M. genitalium*.

Table II.1. Frequency of cells with detectable levels of different mCherry protein fusions (section II.2.1)

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
MG427:Ch	3468	3448	99.42%
RrlA:Ch	4571	16	0.35%
RrlB:Ch	2460	47	1.91%
ORF192.1:Ch	3526	19	0.54%
ΔMG_428-RrlA:Ch	3867	0	0
ΔMG_428-RrlB:Ch	4173	0	0
ΔMG_428-ORF192.1:Ch	2988	0	0
ΔMG_428-RrlA:Ch +Tn428	1563	127	8,13%
ΔMG_428-RrlB:Ch +Tn428	1238	423	34.17%
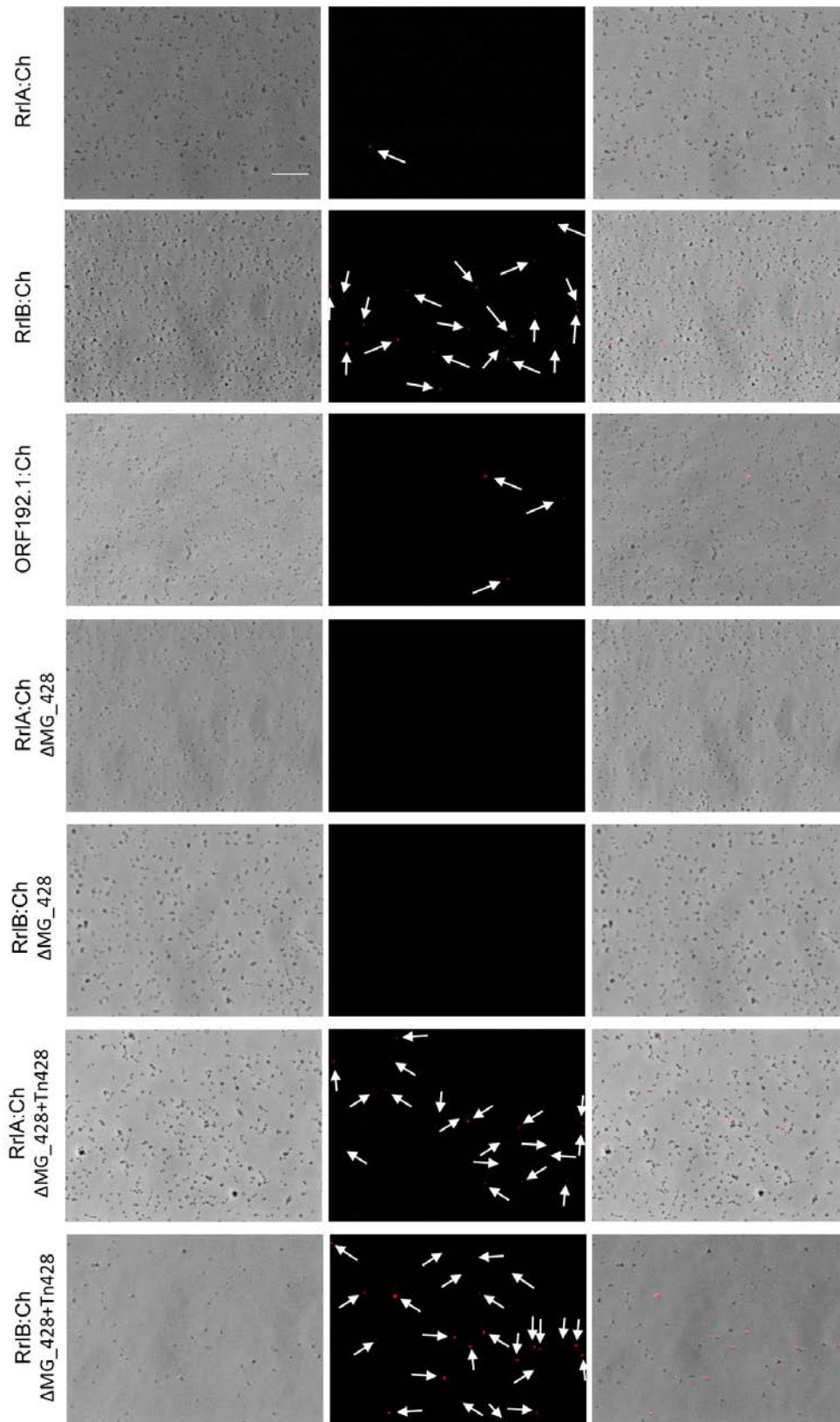
Figure II.2. Single cell analysis of RrlA-, RrlB- and ORF192.1-mCherry expression by fluorescence microscopy. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. Arrows indicate the presence of fluorescent cells expressing either RrlA-, RrlB or ORFMG192.1-mCherry fusions. All the pictures are shown at the same magnification, scale bar is 10μm.
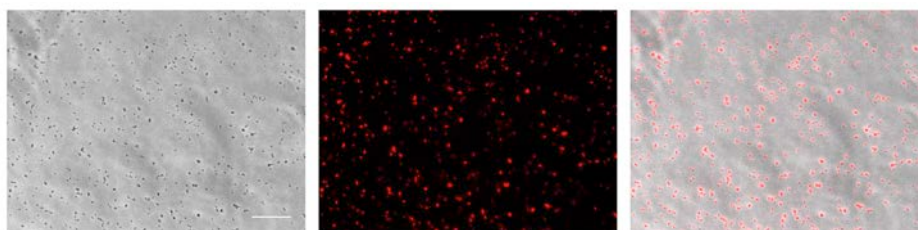
Figure II.3.Single cell analysis of MG427-mCherry expression by fluorescence microscopy. Series of three images correspond to the phase contrast, the Texas Red channel and the resulting overlay, respectively. All the pictures are shown at the same magnification, scale bar is 10μm.

II.2.2. RrlA and RrlB are necessary for the activation of the σ²⁰-regulon

In order to understand the role of RrlA or RrlB in recombination, we decided to monitor the expression of σ²⁰ and RecA in strains lacking any of these two recombination regulatory proteins. To this end, we deleted the *rrlA* or *rrlB* genes from strains carrying MG_428- or *recA*-mcherry fusions at their respective native loci. The absence of RrlA or RrlB prompted a dramatic decrease in the percentage of σ²⁰-mCherry and RecA-mCherry fluorescent cells (Figure II.4, Figure II.5 and Table II.2). In contrast, deletion of other genes under the control of σ²⁰ such as MG_390 or MG_414 had very little or no impact on mCherry fluorescence (Table II.2). Reintroduction of *rrlA* or *rrlB* restored the percentage of fluorescent cells to both mutant backgrounds (Figure II.4, Figure II.5 and Table II.2). Hence, our data indicate that RrlA and RrlB are necessary to activate recombination through the σ²⁰ pathway.

Table II.2. Frequency of cells with detectable levels of different mCherry protein fusions (section II.2.2)

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
RecA:Ch	4529	30	0.66%
RecA:Ch Δ*rrlA*	4239	0	0
RecA:Ch Δ*rrlB*	3765	0	0
RecA:Ch ΔMG_390	4880	25	0.51%
RecA:Ch ΔMG_414	3769	22	0.58%
RecA:Ch Δ*rrlA*+Tn*rrlA*	4587	10	0.22%
RecA:Ch Δ*rrlB*+Tn*rrlB*	3741	9	0.24%
σ²⁰:Ch	2836	13	0.46%
σ²⁰:Ch Δ*rrlA*	5234	0	0
σ²⁰:Ch Δ*rrlB*	4872	0	0
σ²⁰:Ch Δ*rrlA*+Tn*rrlA*	3769	5	0.13%
σ²⁰:Ch Δ*rrlB*+Tn*rrlB*	3564	7	0.20%
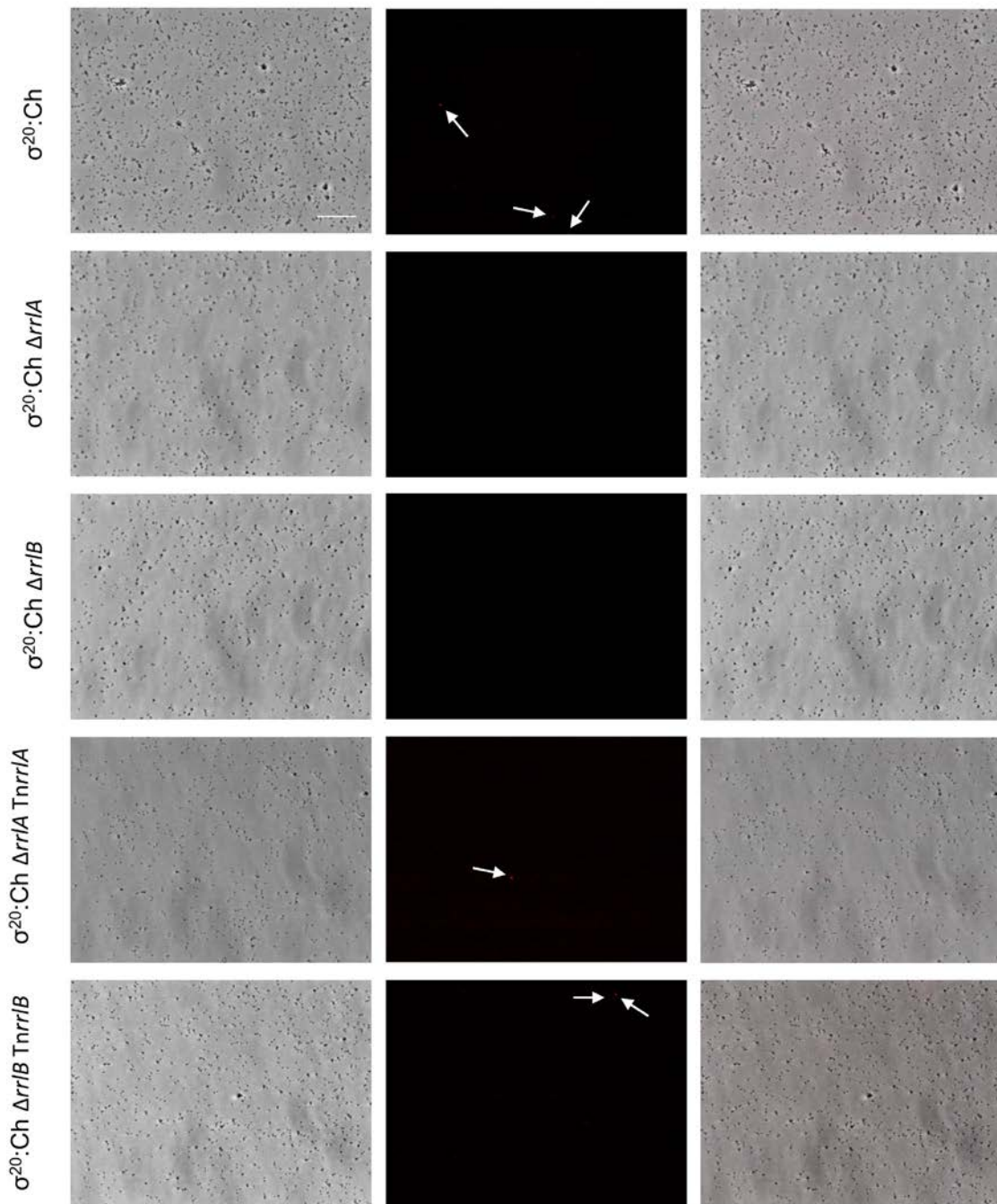
Figure II.4. Single cell analysis of σ^{20}-mCherry expression by fluorescence microscopy in different mutant backgrounds. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. Arrows indicate the presence of fluorescent cells expressing σ^{20}-mCherry fusion. All the pictures are shown at the same magnification, scale bar is 10μm.
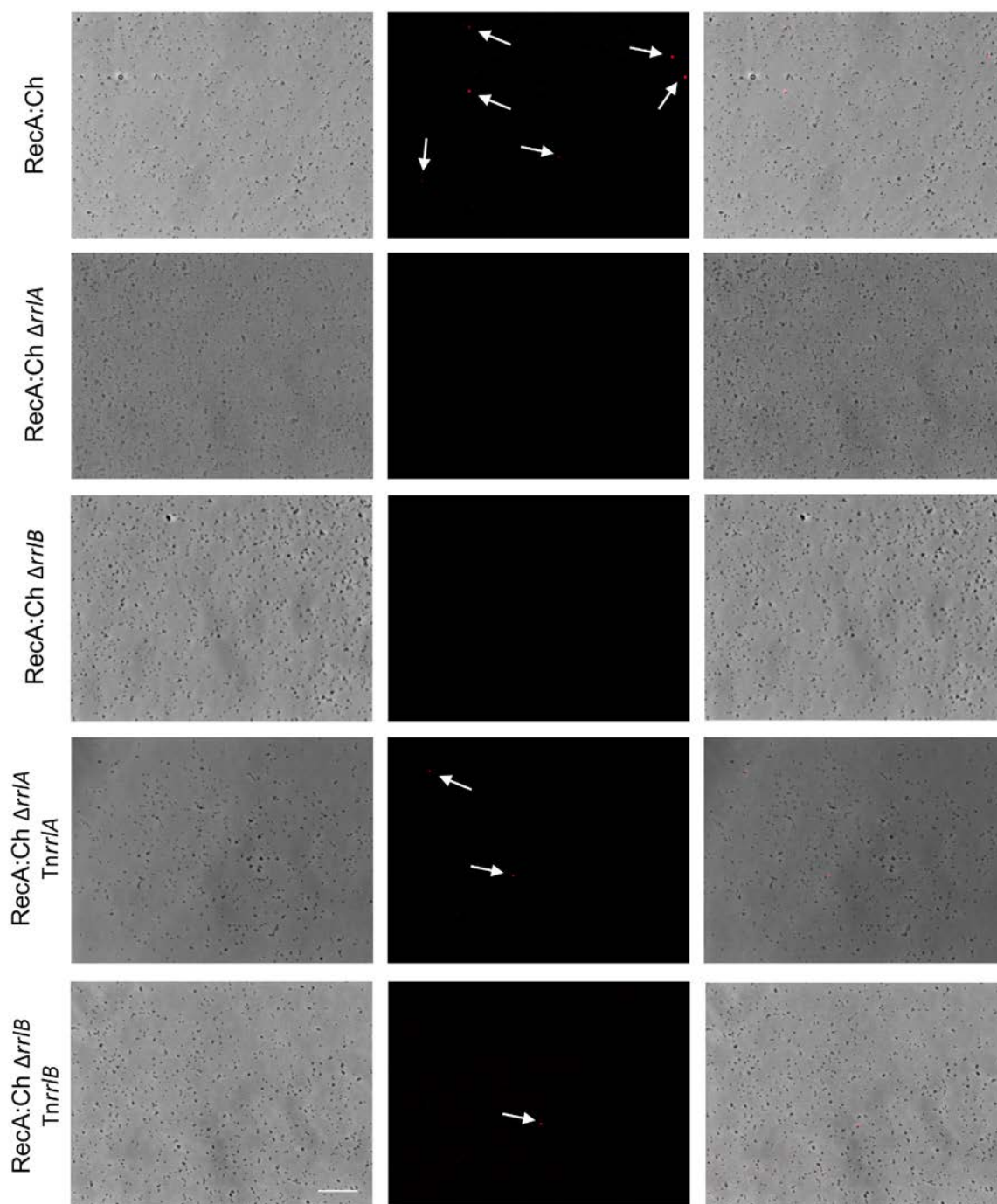
Figure II.5. Single cell analysis of RecA-mCherry expression by fluorescence microscopy in different mutant backgrounds. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. Arrows indicate the presence of fluorescent cells expressing RecA-mCherry fusion. All the pictures are shown at the same magnification, scale bar is 10μm.

II.2.3. Activation of the σ²⁰ regulon increases the stability of σ²⁰

We wanted to get further knowledge on the factors regulating σ²⁰. To this end, we introduced an ectopic copy of the MG_428 gene fused to the *eYFP* marker into the wild-type strain and mutants carrying MG_428- or *recA-mcherry* fusions at their respective native loci. In all cases, the majority of cells exhibited a diffuse eYFP fluorescence, although we also observed many cells without fluorescence (Figure II.6 and Table II.3). Additionally, we identified a few cells displaying a marked eYFP fluorescence. On the other hand, we observed a substantial increase in the percentage of cells showing σ²⁰-mCherry fluorescence, demonstrating the existence of a positive feedback loop underlying σ²⁰ activation (Figure II.6 and Table II.3). Similarly, the percentage of RecA-mCherry fluorescent cells was also higher, indicating that the constitutive expression of the σ²⁰-YFP protein prompted a more frequent and/or prolonged activation of the σ²⁰ regulon. Of note, in the RecA-mCherry background, all cells with intense eYFP fluorescence displayed also mCherry fluorescence. This finding suggests that the σ²⁰-YFP fusion is more stable during the periods of activation of the σ²⁰ regulon. In contrast, in the σ²⁰-mCherry mutant, the overlap within eYFP and mCherry fluorescence was not as apparent. We hypothesize that the two different σ²⁰-fusion proteins expressed in this mutant compete for the stabilizing factors and only one form of σ²⁰ prevails in the cells.
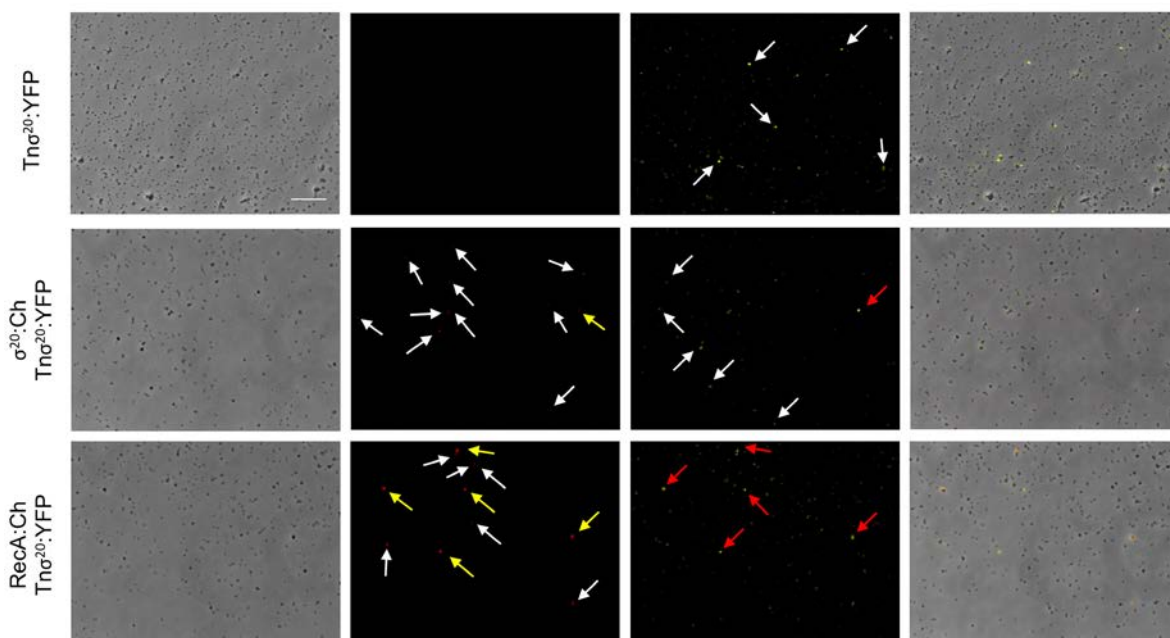
Figure II.6. Single cell analysis of σ²⁰-eYFP and RecA- and σ²⁰-mCherry expression by fluorescence microscopy. Each row contains a series of four images corresponding to the phase contrast, the Texas Red channel, the yellow fluorescent channel and the resulting overlay, respectively. White arrows indicate the presence of cells expressing sharp fluorescence, either mCherry- or eYFP-derived. Yellow arrows indicate mCherry fluorescent cells showing also sharp eYFP fluorescence. Red arrows indicate eYFP fluorescent cells showing mCherry-derived fluorescence. All the pictures are shown at the same magnification, scale bar is 10μm.

Table II.3. Frequency of cells with detectable levels of different mCherry protein fusions (section II.2.3)

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
σ^{20}:Ch	2836	13	0.46%
σ^{20}:Ch Tnσ^{20}:YFP	3643	110	3.02%
RecA:Ch	4529	30	0.66%
RecA:Ch Tnσ^{20}:YFP	3334	107	3.21%

II.2.4. RrlA and RrlB are necessary for σ^{20} stability

To further explore the role of RrlA and RrlB in the activation of the σ^{20} regulon, we deleted the *rrlA* or *rrlB* genes from the reporter strain carrying the RecA-mCherry and σ^{20}-YFP fusions described above. This strain allows us to monitor σ^{20} expression (via σ^{20}-YFP) and activity (through RecA-mCherry) at the same time. As expected, we found that the percentage of cells showing high levels of RecA expression decreased significantly (Figure II.7 and Figure II.9). Concurrently, the presence of cells with intense YFP fluorescence was virtually abrogated, indicating that the absence of RrlA or RrlB prevented the periods of σ^{20}-YFP increased stability described earlier. In contrast, the percentage of RecA-mCherry or σ^{20}-YFP fluorescent cells was not altered upon RrlA or RrlB overexpression (Figure II.8 and Figure II.9). Of note, the majority of mutants recovered in the latter experiments were non-adherent, suggesting that RrlA or RrlB overexpression enhances somehow the generation of P140 and P110 non-adherent variants (data now shown).
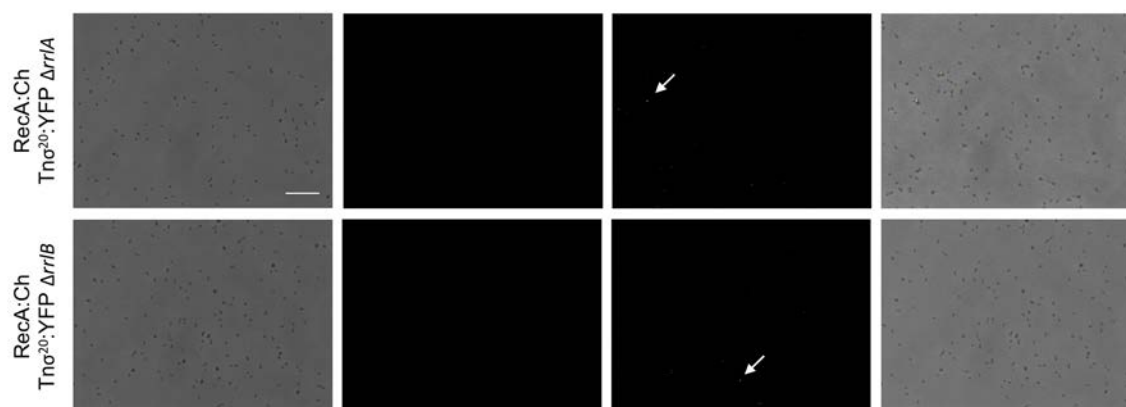
Figure II.7. Single cell analysis of σ^{20}-eYFP and RecA-mCherry expression by fluorescence microscopy in Δ*rrlA* and Δ*rrlB* mutants. Each row contains a series of four images corresponding to the phase contrast, the Texas Red channel, the yellow fluorescent channel and the resulting overlay, respectively. White arrows indicate the presence of cells expressing sharp eYFP fluorescence. Note the absence of mCherry fluorescent cells and reduced number of cells displaying diffuse and sharp eYFP-derived fluorescence. All the pictures are shown at the same magnification, scale bar is 10μm.
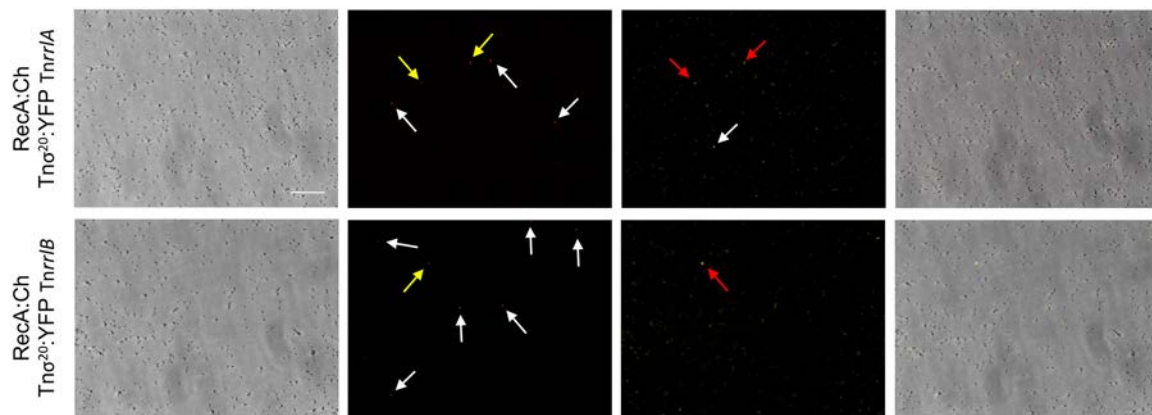
Figure II.8. Single cell analysis of σ²⁰-eYFP and RecA-mCherry expression by fluorescence microscopy in mutants overexpressing *rrlA* or *rrlB*. Each row contains a series of four images corresponding to the phase contrast, the Texas Red channel, the yellow fluorescent channel and the resulting overlay, respectively. White arrows indicate the presence of cells expressing sharp fluorescence, either mCherry- or eYFP-derived. Yellow arrows indicate mCherry fluorescent cells showing also sharp eYFP fluorescence. Red arrows indicate eYFP fluorescent cells showing mCherry-derived fluorescence. All the pictures are shown at the same magnification, scale bar is 10μm.
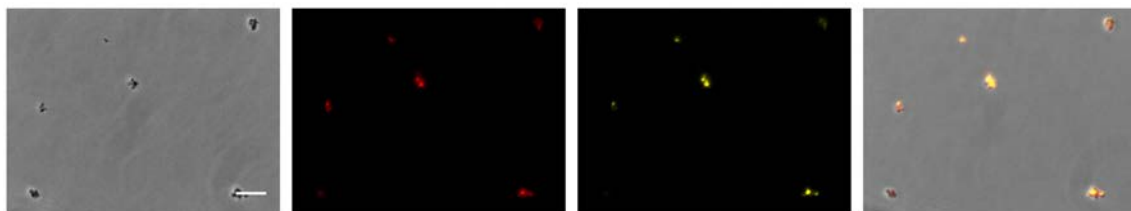
Figure II.9. Single cell analysis of σ²⁰-eYFP and RecA-mCherry expression by fluorescence microscopy in a mutant strain overexpressing the *rrlA* gene and bearing adhesin variants. Each row contains a series of four images corresponding to the phase contrast, the Texas Red channel, the yellow fluorescent channel and the resulting overlay, respectively. All the pictures are shown at the same magnification, scale bar is 10μm. Due to the high cell aggregation of this mutant, frequency of fluorescent cells could not be properly assessed and it is not displayed in Table II.4.

Table II.4. Frequency of cells with detectable levels of RecA-mCherry protein fusions (section II.2.4)

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
RecA:Ch Tnσ²⁰:YFP Δ*rrlA*	2836	0	0
RecA:Ch Tnσ²⁰:YFP Δ*rrlB*	3643	0	0
RecA:Ch Tnσ²⁰:YFP Tn*rrlA*	1190	31	2.60%
RecA:Ch Tnσ²⁰:YFP Tn*rrlB*	1476	42	2.85%

We could isolate one single mutant overexpressing *rrlA* gene that had activation of the σ[20] pathway in virtually all cells (Figure II.9). This particular mutant was mainly non-adherent, but some cells were still able to adhere to plastic surfaces, allowing single cell studies. These analyses showed that all cells presented both mCherry and eYFP sharp fluorescence, although cells were highly aggregated and displayed morphologies similar to those of non-adherent strains. Results derived from this strain suggest that overexpression of *rrlA* can trigger the σ[20] pathway and, as a result, originate phase variants. Therefore, unaltered fluorescence patterns displayed by adherent strains overexpressing *rrlA* or *rrlB* might be originated by gene silencing of the ectopic copies derived in the miniTnps. Altogether, our findings indicate that RrlA and RrlB are necessary but might not be sufficient for the stability of σ[20].

II.2.5. σ[20] overexpression restores recombination to *rrlA⁻* and *rrlB⁻* mutants

We previously demonstrated that overexpression of a σ[20] ectopic copy restores the recombination defects of a σ[20] mutant (Torres-Puig *et al.*, 2015). To confirm the accessory role of RrlA and RrlB in recombination, we overexpressed σ[20] on strains carrying deletions on *rrlA*, *rrlB* or *recA*. σ[20] overexpression restored recombination to mutants lacking either *rrlA* or *rrlB*, although to different extents (Figure II.10). In contrast, recombination could not be restored to a mutant lacking the *recA* gene. Therefore, unlike RecA, the RrlA and RrlB proteins play an accessory role in recombination.
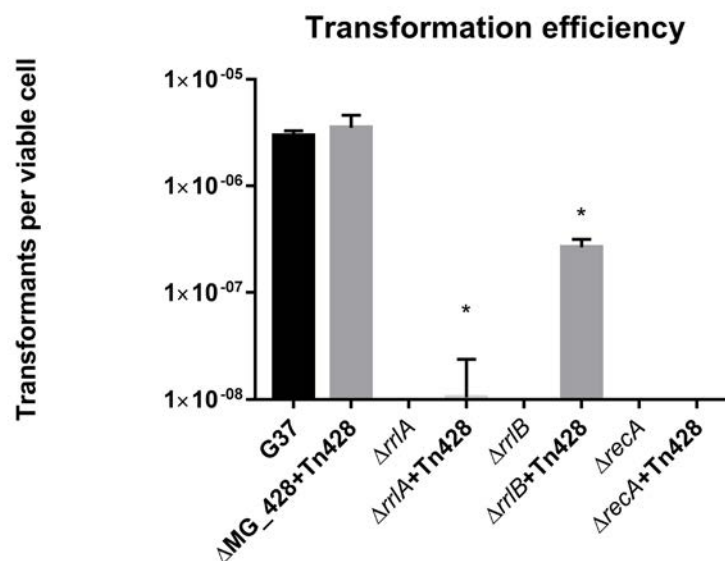
Figure II.10.Transformation efficiency by homologous recombination of different *M. genitalium* strains. Bars show the averages and the standard deviations of at least three independent biological replicates for each strain. Transformation efficiency by HR of strains Δ*rrlA*, Δ*rrlB* and Δ*recA* was below the limit of detection (<10⁻⁸). Transformation efficiency by HR that were found to be statistically different than that of the WT strain are indicated with an asterisk. Statistical significance was assessed with Student's *t* test (*P*<0.05).

II.2.6. σ²⁰ overexpression promotes horizontal gene transfer

In a previous report, we showed that activation of the σ²⁰ regulon was often observed in cell pairs (Torres-Puig *et al.*, 2015). Based on this observation, we wondered whether the activation of recombination mediated by σ²⁰ was linked to a mechanism promoting horizontal gene transfer. To ascertain this question, we mixed strains carrying different antibiotic resistance markers in the presence of high concentrations of DNase and assessed the generation of double resistant mutants. In the first attempts, we mixed strains with wild-type levels of σ²⁰ expression. We used combinations of different antibiotic resistance markers and strains, but we repeatedly failed to isolate double resistant mutants. In the next experiments, at least one of the mixed strains overexpressed σ²⁰ (Figure II.11).
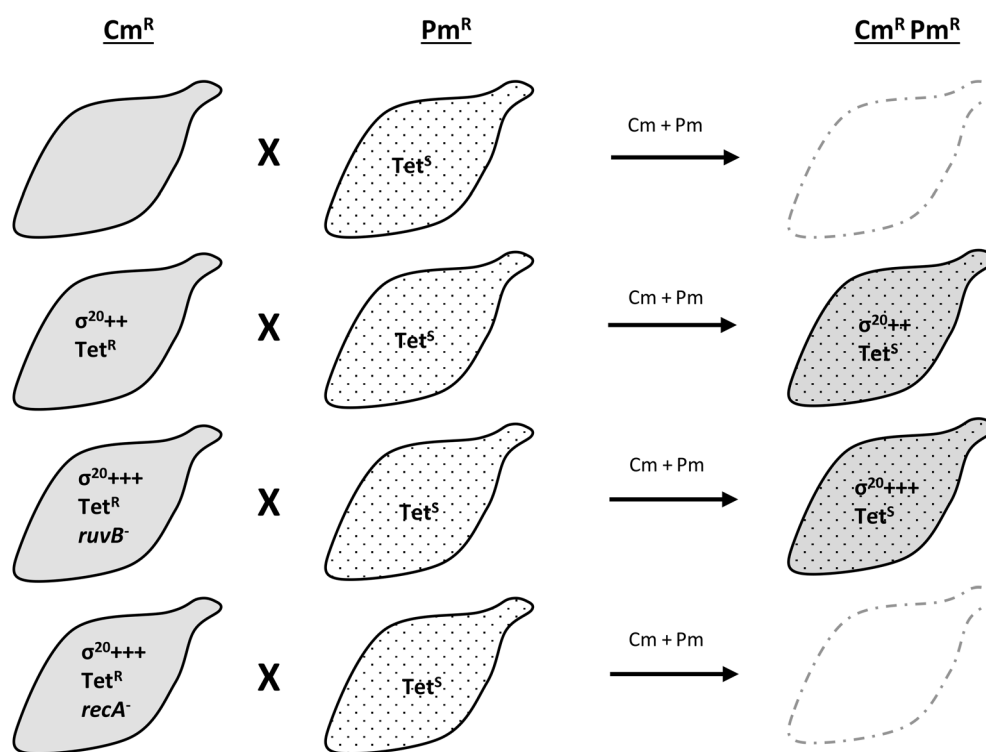
Figure II.11. Schematic representation of the different mating experiments performed in this work. Donor strains are depicted as grey-shaped mycoplasma cells, receiver strain is depicted as a dotted mycoplasma cell and viable transconjugant is drawn as a grey-shaped, dotted mycoplasma cell. Unsuccessful DNA transfers are indicated as empty, dash-lined mycoplasma cells. Protein levels of σ²⁰ are shown as + (constitutive levels of σ²⁰, higher than WT) or +++ (high overexpression of σ²⁰), as shown in Figure I.6.

This time, we successfully isolated mutants resistant to both chloramphenicol (Cm) and puromycin (Pm) upon incubation of a strain resistant to Cm and tetracycline (Tet) and different strains resistant to Pm. The presence of the *cat* and *pac* markers in the recovered

mutants was confirmed by PCR and sequencing (Figure II.12). Of note, the recovered mutants were susceptible to Tet, suggesting the transfer of the *cat* gene from the strain overexpressing σ^{20} to the Pm resistant strain.
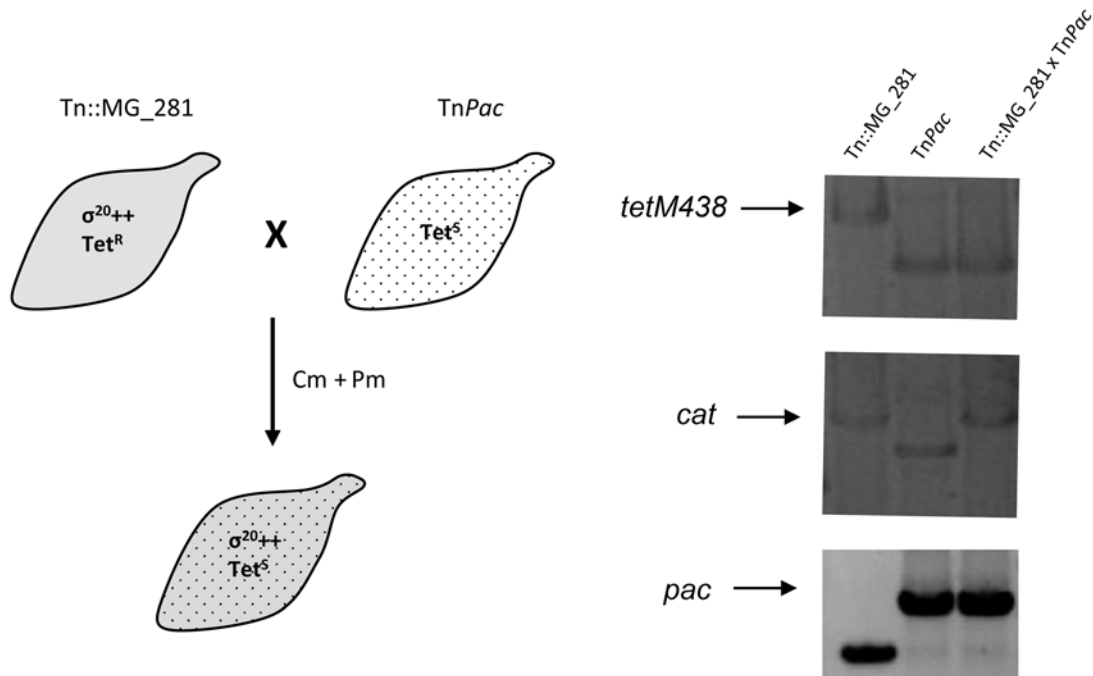
Figure II.12. Genetic analysis by PCR of the different strains used and obtained in the mating experiments. Arrows indicate expected molecular weigth for each antibiotic marker. Strain resulting from the HGT event was passaged three times before the analysis to avoid detection of the genomic regions of the parental strains.

The location of the *cat* gene in the receptor strain was the same as in the donor strain, suggesting that the *cat* marker was integrated by homologous recombination. Similarly, we also observed the acquisition of the *pac* gene by a receptor strain carrying the *cat* marker (data not shown). Additionally, we found that deletion of MG_428 or *recA* prevented the isolation of double resistant mutants (Figure II.11). Similarly, heat inactivation of the donor strain prevented the isolation of double resistant mutants. We could not detect the transfer of the *cat* marker when the Pm strains were incubated with naked DNA, either chromosomal or plasmid, containing a copy of the *cat* gene. Altogether, these data indicate the participation of σ^{20} in a system promoting horizontal gene transfer in *M. genitalium.*

II.3. DISCUSSION

The study of the factors regulating recombination in Mycoplasmas are scarce. In *M. genitalium*, homologous recombination is under the control of a complex regulatory network that we are just beginning to elucidate. Herein, we describe the regulation of σ^{20} function by two uncharacterized proteins. RrlA and RrlB are two proteins of *M. genitalium* that are under the control of σ^{20}. mCherry translational fusions to these proteins at their respective loci allowed the detection of both RrlA and RrlB only in a low percentage of cells, an expression pattern previously observed in another σ^{20}-regulated gene (*recA*) and σ^{20} itself. This intermittent expression pattern and the impairment in recombination capacity displayed by *rrlA* and *rrlB* mutants shown in our previous work (Torres-Puig *et al.*, 2015) suggested that these proteins could play a role in the activation of σ^{20}. Null mutants of *rrlA* and *rrlB* in σ^{20}-mCherry or RecA-mCherry mutant backgrounds abolished the presence of σ^{20}-activated cells in both cases. These results advocate a role for both proteins in the regulation of σ^{20}, that in turn have further consequences on the σ^{20} regulon. Activation of the σ^{20} pathway in $\Delta rrlA$ and $\Delta rrlB$ could be restored by the reintroduction of the respective wild-type alleles by transposon delivery. This fact suggests that protein levels of both RrlA and RrlB tightly stabilize or promote activity of σ^{20} and are highly dependent on the genetic context surrounding the transposon insertion.

On the other hand, introduction into a σ^{20}-mCherry strain of an ectopic σ^{20} copy under the control of a constitutive promoter resulted in a non-negligible increase in cells displaying mCherry fluorescence, suggesting the existence of a feed-forward loop activating σ^{20}. The effect of constitutive expression of σ^{20} had a similar effect on RecA expression, confirming that stabilization of the natural σ^{20} has an impact on the σ^{20} regulon. Remarkably, deletion of *rrlA* or *rrlB* reduced drastically the protein levels of the ectopic σ^{20} as well as RecA expression, suggesting that the feed-forward loop was strictly dependent on the expression of these two proteins.

We could not see an effect on σ^{20} activation when overexpressing either RrlA or RrlB. However, most of the isolated mutants were non-adherent and formed large aggregates, which makes very difficult the characterization of these strains. Phase variants in *M. genitalium* cannot adhere to plastic surfaces or to blood cells, and are generated by homologous recombination of the MgPa operon and scattered homologous regions in the chromosome (MgPar) (Burgos *et al.*, 2006). As homologous recombination is the major

mechanism to generate these mutants, it is tempting to speculate that overexpression of RrlA or RrlB could exacerbate σ^{20} activity and increase the recombination capacity. Then, adherent isolated mutants shown in this work might not exhibit enough levels of RrlA or RrlB to induce any detectable change.

Altogether, these results suggest that RrlA and RrlB are sigma auxiliary proteins that aid σ^{20} in the activation of the regulon (Figure II.13). This mechanism is not unprecedented in bacteria: cells use many different strategies to force transcription transition from the primary sigma factor to alternative sigma factors (Österberg *et al.*, 2011). *M. genitalium* reduced chromosome does not code for any known anti-sigma factors, although this kind of regulators have low sequence conservation (Paget, 2015) and could be hidden in the form of hypothetical proteins with unknown function. The presence of RrlA and RrlB as positive regulators of the σ^{20} use resembles the activity of the Crl protein in *Escherichia coli* or *Salmonella typhimurium*. Crl protein can bind to the alternative sigma factor σ^{S} and enhance the formation of a σ^{S}-RNA polymerase complex (Gaal *et al.*, 2006; Typas *et al.*, 2007). This factor is critical for the σ^{S} activity as it aids overcoming the low affinity of the alternative sigma factor for the RNA polymerase core enzyme as compared to the primary sigma particle. In the case of *M. genitalium*, RrlA and RrlB could help during σ^{20} activation to redirect transcription to the σ^{20} regulon and thus promote homologous recombination and other possible unknown cellular processes.
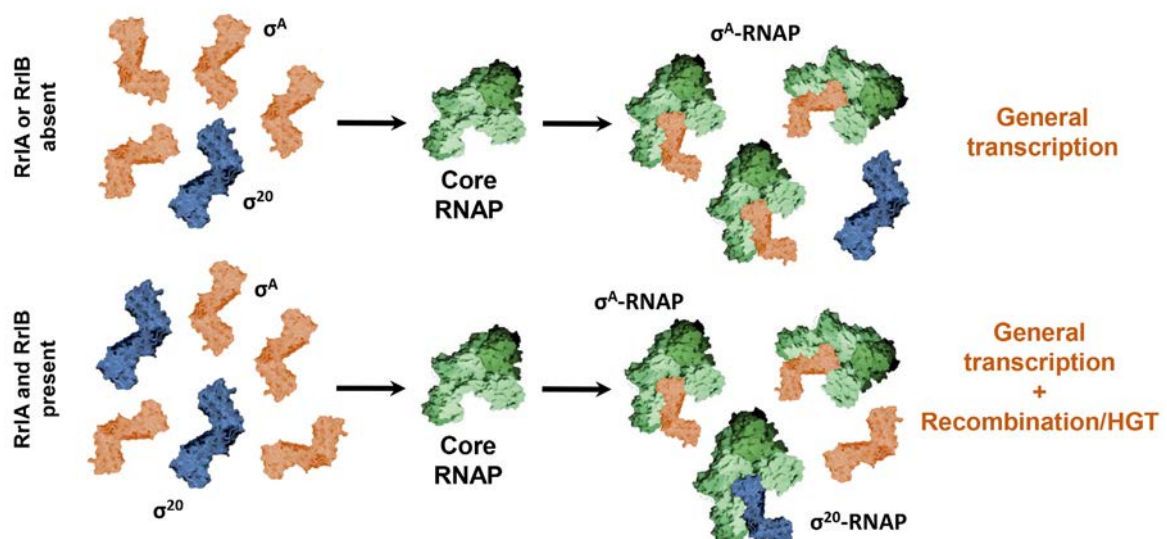
Figure II.13. Schematic representation depicting the possible role for RrlA and RrlB in the activation of the MG428 pathway. We hypothesize that RrlA and RrlB stabilize the σ^{20} protein, which is critical during the induction of the σ^{20}-dependent recombination pathway because the levels of σ^{20} are low. The RrlA-RrlB dependent stabilization of σ^{20} is necessary to allow binding of this alternative sigma factor to the RNA polymerase core enzyme. Otherwise, the presence of large amounts of σ^{A} (MG249) prevents σ^{20} binding to the RNAP.

On the other hand, herein we have described an unprecedented form of HGT in Mycoplasmas. We have observed that strains overexpressing the alternative sigma factor σ^{20} can act as donor cells in a novel DNA transfer system. HGT in Mycoplasmas is a trait only detected in certain ruminant species that encode integrative conjugative elements (ICE) or ICE-like sequences in their chromosomes (Dordet Frisoni *et al.*, 2013; Tardy *et al.*, 2015). In a recent work, large chromosomal transfers via homologous recombination between mycoplasma strains was documented, although it was dependent at some extent on the ICE enzymatic machinery for membrane fusions (Dordet-Frisoni *et al.*, 2014). *M. genitalium* apparently lacks any ICE or ICE-like elements, thus this gene transfer would be catalyzed by intrinsic non-mobilizable factors encoded in the chromosome. The strict dependence of this HGT system on σ^{20} overexpression prompts us to speculate that proteins under the control of σ^{20} are the main effectors of this unconventional gene variation mechanisms. Besides, several σ^{20}-regulated proteins are thought to be located at the cell membrane, which could enhance the formation of cell to cell contacts that promote DNA exchange. Work is in progress to determine which is the role of these hypothetical proteins in this and other cellular processes.

Overall, in this work we have defined two factors that regulate σ^{20} activity and are key on proper antigenic variation catalyzed by this alternative sigma factor. Moreover, the discovery of a form of horizontal gene transfer in *M. genitalium* widely increases the genome plasticity of this small pathogen and represents a novel form of antigenic variation that could play an important role in bacterial persistence and immune evasion.

GENERAL DISCUSSION

D.1. MG428, AN ALTERNATIVE SIGMA FACTOR OF *MYCOPLASMA GENITALIUM*

D.2. REGULATION OF **MG428** PATHWAY ACTIVATION. ROLE OF SIGMA ACCESSORY PROTEINS

D.3. IMPACT OF **MG428** REGULATORY PATHWAY IN *M. GENITALIUM* LIFESTYLE

D.1. MG428, AN ALTERNATIVE SIGMA FACTOR OF *MYCOPLASMA GENITALIUM*

D.1.1. MG428 recognizes a novel promoter with σ⁷⁰ architecture

In the recent years, rudimentary control mechanisms of transcription in Mycoplasmas have been challenged by the discovery of supraoperonic transcript organization and global gene regulation though terminators and terminator-like sequences (Mazin *et al.*, 2014; Junier *et al.*, 2016). However, regulation through classic transcription factors has been poorly described despite that many tools for the identification of these regulators are currently available. *M. genitalium*, the microorganism used in this study, has more than 10 proteins with DNA-binding domains that could serve as transcription factors, but to date only the function of the heat-shock regulator HrcA has been investigated (Musatovova *et al.*, 2006).

In this work, we have characterized the MG428 protein, a putative transcriptional regulator of the LuxR family. The HTH domain of this family of regulators resembles region r4 from sigma factors of the sigma-70 family, the region involved in recognition of -35 elements of bacterial promoters. Unlike most transcription factors, sigma factors cannot bind DNA if they are not associated to the RNA polymerase complex. We demonstrated that MG428 binds to the *E. coli* RNA polymerase when the latter is devoid of other sigma factors. Moreover, upon overexpression of MG428, several genes under the control of novel promoters were activated, suggesting that they were recognized by MG428. Noteworthy, MG428-regulated promoters contained a well-defined, classic -35 element (TTGACA), a rare feature among mycoplasma promoters.

The primary sigma factor (σ^A) of *M. genitalium*, encoded by the MG_249 gene, bears an r4 region in its C-terminal domain (Figure D.1). This domain is responsible for binding to the -35 element, which is often missing in housekeeping promoters recognized by σ^A.
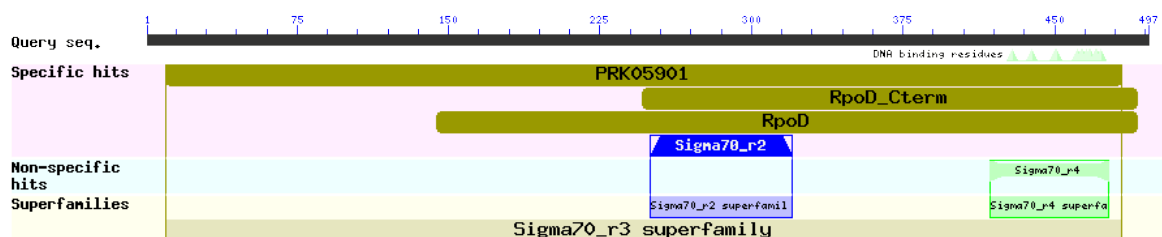
Figure D.1. Conserved domains on MG249 protein of *M. genitalium*, the primary sigma factor in this bacterium. Note the presence of the r4 region at the C-terminal domain as well as the r2 region, responsible for recognition of the Pribnow box of bacterial promoters.

In other bacteria, loss of the -35 promoter element results in poor recognition by the primary sigma factor, causing a deficient initiation of transcription (Reznikoff *et al.*, 1985). Therefore, efficient recognition of promoters lacking -35 elements by σ^A of *M. genitalium* remains unknown. However, the conservation of the r4 in the primary sigma factor of *M. genitalium* might indicate that σ^A is also able to recognize -35 elements in promoters. This fact could explain the low levels of mRNA expression observed for the majority of the MG428-regulated genes. That is, MG428-regulated promoters could be recognized by σ^A through its -35 element and basal transcription initiated despite the alternative Pribnow box present in these promoters. During activation of MG428, this alternative sigma factor probably binds to these promoters with higher affinity than the primary σ and drives transcription at higher levels. In fact, MG428 might recognize promoters more efficiently and with better specificity than the primary sigma factor, as it can recognize the -35 and the alternative -10 elements from the alternative promoters.

Despite all the evidence presented in this work, the role of MG428 as a sigma factor could be further characterized. Experiments showing binding of MG428 to the RNA polymerase of *M. genitalium* and *in vitro* transcription from MG428-regulated promoters using purified RNAP-MG428 complexes will reinforce the findings presented in this work. However, the slow growth and limited biomass production of *M. genitalium*, hampers the purification of large amounts of RNAp-MG428 complexes.

D.1.2. MG428 activates transcription of a unique regulon

Overexpression of the alternative sigma factor σ^{20} leads to the upregulation of several genes with distinct functions. The first genes under the control of the MG428 regulator were identified because they were preferentially targeted by transposons overexpressing MG428. This serendipitous procedure to identify MG428-regulated genes relies on the toxicity associated to MG428 overexpression. We hypothesize that only those mutants with transposon insertions that neutralize the toxic effect derived from MG428 overexpression survive and can be therefore isolated. In this sense, in the isolated clones, the transposon-encoded allele of the MG428 gene was always in the antisense orientation with respect to the disrupted gene. Insertion of the transposon in the antisense direction could provide some means to downregulate protein expression of the ectopic copy of MG428 and thus reduce toxicity. If this is the case, this experiment indicates that RNA-RNA hybrids could alter translation efficiency by the ribosomes, either blocking ribosome binding to RNA or recruiting specific RNases that could target and cleave these hybrids.

Interestingly, the aforementioned transposon overexpressing MG428 targeted *recA* and *ruvB*, which code for two of key enzymes of the recombination machinery. This fact prompted us to think that other genes coding for proteins involved in homologous recombination could be under the control of MG428. However, RNAseq analyses confirmed that only *recA* and the Holliday resolvases *ruvAB* were regulated by this alternative sigma factor. Other genes such as *recU* or *ssbA*, were not found to be upregulated upon MG428 overexpression. Single-stranded DNA binding protein SsbA is responsible for protecting ssDNA from nuclease activity during replication and homologous recombination. It has been described in other bacteria that SsbA and SsbB proteins inhibit RecA filamentation on ssDNA during recombination when it is bound to ATP (RecA-ATP), although they are required for efficient strand exchange when RecA is bound to dATP (Figure D.2) (Yadav *et al.*, 2012). Other DNA-binding proteins, such as RecO or DprA, can aid in this process, (Mortier-Barrière *et al.*, 2007; Yadav *et al.*, 2012, 2013), but they are missing in the reduced chromosome of *M. genitalium*. Therefore, we hypothesize that *ssbA* upregulation during a σ^{20} activation interval would impede both efficient RecA loading to ssDNA and promotion of strand exchange, obstructing the recombination process in this small pathogen.
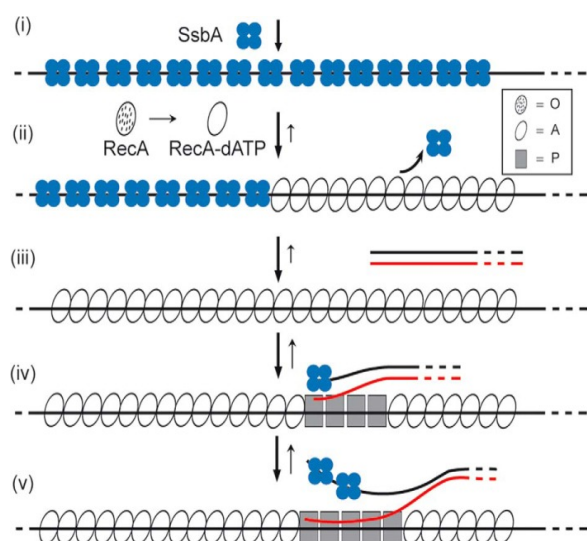
Figure D.2. Polymerization on DNA, pairing, and strand exchange mediated by RecA in the presence of dATP. The template for RecA-dATP-Mg^{2+} assembly *in vivo* is SsbB- or SsbA-coated ssDNA (*step i*). Based primarily on biochemical analyses, there are at least three significant functional stages of RecA. In the first state, RecA cannot bind ssDNA (apo RecA, *dotted ovals*, O state), but upon binding to dATP-Mg^{2+}, it forms the active state denoted as RecA-dATP (*empty oval*, A state) that can nucleate on the SsbB- or SsbA-coated ssDNA (*step ii*). RecA-dATP upon nucleation undergoes a second functional transition, which allows efficient displacement of SsbA and polymerization on the ssDNA (*step iii*). RecA-dATP by binding to dsDNA acquires a third structural transition state (*filled squares*, P state) that can catalyze DNA pairing (*step iv*) and DNA strand exchange (*step v*). Extracted from Yadav *et al.*, 2012.

On the other hand, in Gram-positive bacteria, *recU* codes for a Holliday endonuclease with a similar function to RuvC in Gram-negative microorganisms. Along with RuvA and RuvB, RecU is essential for proper Holliday junction resolution. Our transcriptional data suggest that *recU* (MG_352) and the *ppi* gene (MG_351) comprise an operon in *M. genitalium*. Interestingly, *in silico* analyses identify a putative MG428-regulated promoter at the upstream region of MG_351. However, none of these two genes were found to be upregulated upon MG428 overexpression. It has been shown that RecU can aid in strand

exchange by recruiting RecA onto ssDNA and promote D-loop formation in DNA in *B. subtilis*. However, the presence of high amounts of RecU also inhibits RecA-mediated strand exchange by preventing RecA disassembly (Carrasco *et al.*, 2005). This dual effect could explain why *recU* is not upregulated by σ[20] along with *recA* or *ruvAB*.

The function of other genes regulated by MG428 and targeted by the transposon is unclear. The biological role of the MG220 and the protein encoded by the MG_RS02200 gene (herein designated as RrlA and RrlB, respectively) will be discussed thoroughly in the section D.2 of this thesis. In the case of MG_390, some information regarding its putative function is available. MG390 resembles a C39 membrane peptidase which are usually fused to an ABC transporter domain. In other bacteria, these proteins are encoded by the *sunT* gene. C39 peptidases play a role in bacteriocin peptide export in *B. subtilis* (Paik *et al.*, 1998) and quorum sensing and competence in *Streptococcus pneumoniae* (Hui *et al.*, 1995). Despite most mycoplasma species code for a *sunT* homolog (see Table D.1), these genes have not been studied so far. C39 peptidases of the SunT family recognize proteins with leader peptide sequences bearing a double glycine motif (GG), and cleave and export the processed peptides (Gebhard, 2012). It was surprising to notice that the RrlA protein of *M. genitalium* has a putative leader peptide sequence similar to those recognized by SunT exporters (Supp. Figure S.5). However, we have no evidence suggesting the secretion of RrlA to the extracellular medium, probably because detection of bacterial hormones requires high sensitive methods. Nevertheless, a null mutant for MG_390 showed lower transformation efficiency by HR compared to the WT strain, reinforcing the possible connection between MG390 and the RrlA protein. Interestingly, we have observed the presence of small proteins bearing GG leader peptides in other Mycoplasmas (Table D.1), which should motivate the further study of these systems and their role in mycoplasma populations.

Other genes under the control of σ[20] code for hypothetical proteins. MG_414 and MG_525 are paralogues, probably generated by duplication of a common ancestor gene. Bioinformatic predictions indicate that they are both surface proteins that share a unique transmembrane helix domain at the C-terminal end that probably anchors them at the cell membrane. They are conserved among most Mycoplasmas of the pneumoniae cluster, but their function is still unknown.

Table D.1. Putative sigma factors of the ECF subfamily and homologues of σ^{20}-regulated genes in other mycoplasma species

Mycoplasma strain	ECF-sigma 70	*recA*	*ruvA*	*ruvB*	*sunT*	*rrlA*[1]	*rrlB*	M5 RNase[2]/ *recR*/primase	HP	HP
M. genitalium G37	MG_428	MG_339	MG_358	MG_359	MG_390	MG_220	MG_RS02200	MG_010	**MG_285**	**MG_414**
M. pneumoniae M129	MPN626	MPN490	MPN535	MPN536	MPN571	MPN313	MPN534	MPN014	MPN404	MPN612
M. pneumoniae FH	F539_03525	F539_02755	F539_03040	F539_03045	F539_03230	F539_01765	F539_03035	F539_00075	F539_02285	F539_03455
M. gallisepticum R(low)	MGA_0765	MGA_0143	MGA_0407	MGA_0404	MGA_0022	MGA_1216	-	MGA_0346[4]	MGA_1029	MGA_0754
U. parvum ser3 ATCC 27815	-	UPA3_0083	UPA3_0468	UPA3_0467	UPA3_0540	UPA3_0412	-	UPA3_0200[4]	-	-
M. penetrans HF-2	-	MYPE6390	MYPE2790	MYPE2780	MYPE7470	-	-	MYPE9440[4]	-	-
M. hyopneumoniae 232	mhp659	mhp041	mhp421	mhp422	mhp222	mhp253	-	mhp121[3]	-	-
M. agalactiae PG2	MAG_0420	MAG_5500	MAG_2230	MAG_2240	MAG_1130[P]	MAG_5560	-	MAG_6860[3]	-	-
M. agalactiae 5632	MAGa0420	MAGa6140	MAGa2400	MAGa2410	MAGa1190	MAGa6200	-	MAGa7820[3]	-	-
M. bovis PG45	MBOVPG45_0043	MBOVPG45_0258	MBOVPG45_0596	MBOVPG45_0597	MBOVPG45_0122	MBOVPG45_0251	-	MBOVPG45_0783[3]	-	-
M. fermentans PG18	MBIO_0783	MBIO_0827	MBIO_0357 (?)	MBIO_0358 (?)	MBIO_0130	MBIO_0388	-	MBIO_0768[3]	-	-
M. pulmonis UAB CTIP	MYPU_1790	MYPU_2520	MYPU_6570	MYPU_6580	MYPU_3760	MYPU_3310	-	MYPU_0510[3]	-	-
M. mycoides mycoides SC. PG1	-	MSC_0421	MSC_0542	MSC_0543	MSC_0459[T]	-	-	MSC_0003[4]	-	-
M. capricolum ATCC 27343	-	MCAP_0551	MCAP_0429	MCAP_0428	MCAP_0509	-	-	MCAP_0003	-	-
M. synoviae ATCC 25204	-	VY93_03100	VY93_00890	VY93_00895	VY93_02115	-	-	VY93_00285[3]	-	-
M. mobile 163K	-	MMOB5150	MMOB2700	MMOB2710	MMOB3500	MMOB3140	-	MMOB0750[3]	-	-
M. hyorhinis SK76	MOS_271	MOS_234	MOS_295	MOS_294	MOS_347	MOS_RS02890	-	MOS_115	-	-
M. hominis ATCC 23114	-	MHO_1190	MHO_1840	MHO_1850	MHO_4410	MHO_3030	-	MHO_0970[3]	-	-
M. mycoides JCVI-Syn3.0	-	-	-	-	JCVISYN3_0401	-	-	JCVISYN3_0003[4]	-	-

[1] Homology to double-glycine (GG) leader peptide-containing proteins of ~10 kDa
[2] Annotated as suggested by Allemand *et al.*, 2005
[3] Annotated as RecR
[4] These organisms contain both RecR and M5 ribonuclease proteins. ORF listed in this table codes for the putative M5 ribonuclease

(?) *M. fermentans* PG18 strain encodes several *ruvAB*-like ORFs in mobile elements
[T] This gene only codes for the N-ter (C39 domain) of SunT, the ABC transporter domain is missing
[P] Pseudogene

As stated by Glass and colleagues (Glass *et al.*, 2006), null mutants of MG_414 grow significantly faster (data not shown), a treat not observed in MG_525 null mutants. Moreover, our data suggests that MG525 plays a certain role in recombination while ΔMG_414 mutant shows no recombination defects. However, more experimental data is needed to better understand the role of these proteins in the σ^{20} pathway.

RNAseq analyses showed that MG_285 and MG_286, two genes coding for hypothetical proteins, are also under the control of MG428. For its chromosomal localization, one could predict a possible operonic structure along with the MG_284 gene. However, primer extension analysis revealed that σ^{20}-mediated transcription initiates a few bases upstream of the translational start of the MG_285 gene. Indeed, a putative σ^{20}-regulated promoter can be identified in this region (inside the MG_284 gene). From their secondary structure, we can hypothesize that MG285 and MG286 are membrane proteins, but their specific role remains unknown.

MG010 is a putative DNA primase-like protein of *M. genitalium* and it is also regulated by σ^{20}. *M. genitalium* has another gene (*dnaG* or MG_250) coding for the primary DNA primase, a protein essential for chromosome replication as it synthesizes short RNA primers in the lagging strand. Although MG010 is smaller than housekeeping DNA primases, the presence of two proteins with TOPRIM domains in the small chromosome of *M. genitalium* is puzzling. In other bacteria, small TOPRIM domain-containing proteins code for a specific RNase called M5 ribonuclease, involved in 5S rRNA maturation (Allemand *et al.*, 2005). Moreover, TOPRIM domain is also present in the RecR recombination protein, a member of the RecFOR pathway that aids RecA in homologous recombination. RecR function is highly dependent on the presence of RecO, as they form clusters to favor RecA loading to ssDNA. Some mycoplasma species carry one copy of RecO in their small genomes, so it is not surprising the presence of genes coding for small proteins with TOPRIM domains annotated as RecR. However, *M. genitalium* and *M. pneumoniae* do not bear RecF or RecO, therefore it is unlikely that MG_010 codes for a protein related to recombination. Supporting this hypothesis, we show that a null mutant of MG_010 have similar recombination efficiencies than the WT strain. In contrast, *recR* mutants usually result in recombination defects in other bacteria (Alonso *et al.*, 1991). Alternatively, this regulated alternative primase could be involved in horizontal gene transfer. Work is in progress to test this hypothesis.

Finally, RNAseq analysis revealed that four putative non-coding RNAs were upregulated upon σ^{20} overexpression. One of them, ncRNA-2, encompasses a novel ORF coding for a small protein with unknown function. The role of these non-coding RNAs is unknwon, but they could be involved in regulation of n of adhesion expression. Suporting this idea, part of the ncRNA-1 and 2 comprise MgPa repeat regions in the antisense orientation and could generate dsRNA hybrids with the *mgpB* or *mgpC*.

The presence of putative alternative sigma factors of the ECF subfamily in other mycoplasma species (Table D.1), suggests that these bacteria could use these proteins to regulate recombination and DNA transfer. However, the uniqueness of antigenic variation in *M. genitalium,* which is stimulated by σ^{20}, suggests that alternative sigma factors of other Mycoplasmas will likely be involved in other cellular processes. Indeed, some mycoplasmas of the *hominis* cluster code for a putative alternative sigma factor situated immediately after the 23S rRNA methyltransferase RlmB, suggesting a possible role of this transcription factor linked to rRNA methylation and large translational control.

D.2. REGULATION OF MG428-PATHWAY ACTIVATION. ROLE OF SIGMA AUXILIARY PROTEINS

Homologous recombination plays a pivotal role for survival mechanism for *M. genitalium* within the host. In this sense, it is the prime way to generate antigenic diversity and elude the immune system of the host. However, the presence of active recombination systems and several homologous sequences scattered all around the genome (MgPa repeats) is a double-edged sword. Recombination in single crossovers can lead to major chromosome rearrangements and deletions, which may result in reduced cellular fitness or cell death. For this reason, it is reasonable to think that the σ^{20} pathway, responsible for the activation of recombination in *M. genitalium*, will have several regulatory layers to avoid undesired hyperrecombination states. We have determined three different levels of regulation of the expression, stability or activation of the MG428 protein in *M. genitalium*.

D.2.1. Regulation of σ^{20} expression

First, MG428 expression is regulated at the transcriptional level. In this work, we show that the MG_427 promoter (P_{MG_427}) drives MG_428 expression. However, RNAseq analyses confirmed that P_{MG_427} is a strong promoter and transcriptional fusion to the MG_428 gene

leads to a pronounced σ^{20} overexpression. Therefore, a genetic element between the MG_427 and MG_428 genes must act *in cis* to limit MG_428 mRNA levels in the cell. Bioinformatics analyses using intrinsic terminator predictors strongly suggested the presence of a terminator sequence at the end of MG_427 gene. In fact, the presence of this terminator sequence is in agreement with our RNAseq coverage analyses of the WT strain (Figure D.3A). RNA levels in the positive strand of the MG_427 drop abruptly at the 3' end of this gene, resulting in a dramatic reduction of MG_428 expression.
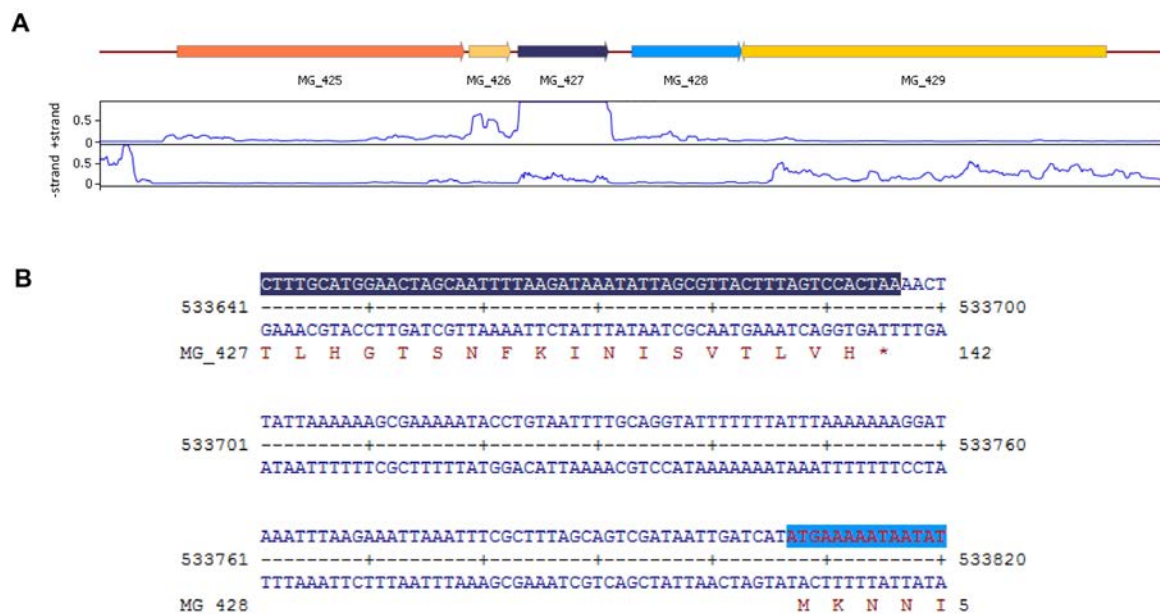
Figure D.3. Chromosome organization of the MG_428 locus. (A) Relative RNA levels of the MG_428 locus in the WT strain. The upper and lower panels depict coverage of total RNA of the positive and negative strand, respectively. Note the coverage drop at the 3' end of the MG_427 gene. **(B)** DNA sequence at the MG_427-MG_428 intergenic region. Sequence highlighted in dark blue corresponds to the 3' end of MG_427. Sequence highlighted in light blue covers the 5' end of the MG_428 gene. Letters in red show the amino acid sequence of each translatable region.

In addition, the intergenic region between the MG_427 and MG_428 genes is 110 nucleotides long and no apparent Shine-Delgarno sequence is present (Figure D.3B). This fact could limit translation efficiency of the MG_428 gene and account for the low levels of σ^{20} present in the WT strain. Translational coupling between genes separated with more than 20 nucleotides is highly inefficient (Schümperli *et al.*, 1982; Levin-Karp *et al.*, 2013). Therefore, a terminator sequence between these two genes can have an effect both in transcription and translation of the MG_428 product.

ECF-sigma factor expression is normally triggered by external signals or certain stresses. This is the case of well-studied alternative sigma factors such as σ^W, regulating the envelope stress response in *B. subtilis*; σ^{FecI}, responsible for the iron limitation response in *E. coli*; or σ^R, involved in the regulation of oxidative stress in *S. coelicolor* (Paget and Helmann, 2003). We have subjected *M. genitalium* to various stresses and analyzed σ^{20} expression by single cell analysis and qRT-PCR. So far, we have not found any condition stimulating σ^{20} production (see Appendices section S.2.2). Actually, the fact that the σ^{20} pathway is only activated in a small subset of cells at a time, suggests that the trigger factor of σ^{20} activation could be intrinsic and likely cyclical, as for example DNA replication or cell division.

D.2.2. Regulation of σ^{20} stability

The nature of the σ^{20} protein as an alternative sigma factor of the ECF subfamily has also some implications regarding its regulation. The primary sigma factor, σ^A, controls the initiation of transcription of the housekeeping genes of *M. genitalium*, and it is probably responsible for the basal levels of transcription of σ^{20}-regulated genes. For this reason, as it is well described in other bacteria, it is assumed that the affinity of σ^A for the core RNA polymerase enzyme is higher than the affinity of alternative sigma factors (Maeda *et al.*, 2000; Grigorova *et al.*, 2006; Mauri and Klumpp, 2014). Moreover, σ^A is typically more abundant in the cell than alternative sigma factors. As a result, during the sigma cycle (see Figure D.4), σ^A-binding to core RNA polymerase is favored, which limits the function of alternative sigma factors.
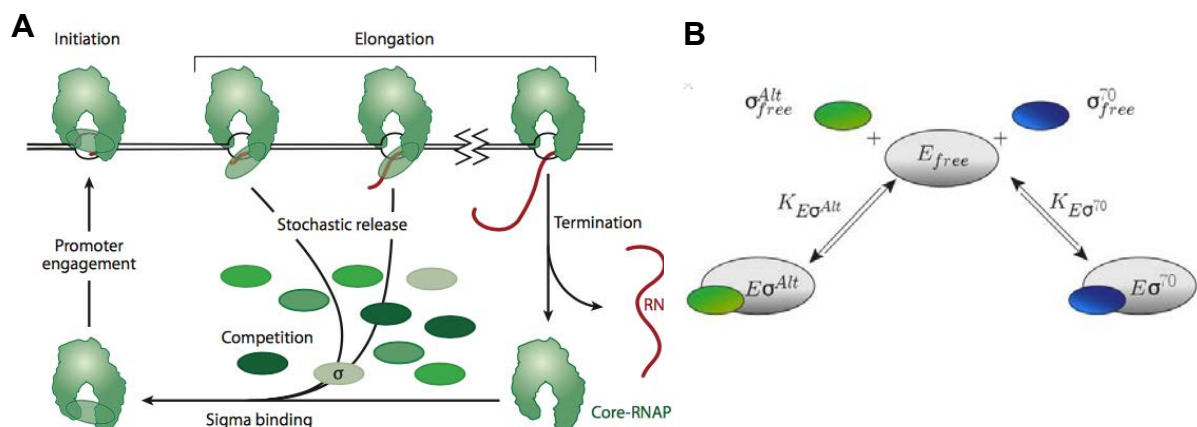
Figure D.4.The sigma cycle allows reprogramming of the core RNA polymerase (core-RNAP). (A) Schematic illustration of the transcription cycle in which sigma factors compete for association with the core-RNA polymerase to direct the holoenzyme to engage promoters. **(B)** Simplified core model for RNAP holoenzyme formation in the presence of two different sigma factors. Adapted from Österberg, Peso-Santos and Shingler, 2011 and Mauri and Klumpp, 2014.

Therefore, bacterial cells have evolved a large repertoire of strategies to increase competitiveness of alternative sigma factors, including alarmones, such as ppGpp, or proteins like anti-sigma factors, anti-antisigma factors, RNA polymerase-binding proteins or σ-binding proteins (Österberg *et al.*, 2011).

In this work, we have identified two proteins that impact σ^{20} stability: the RrlA and RrlB (Rrl stands for <u>r</u>ecombination <u>r</u>egulatory <u>l</u>oci). Mutants of MG_220 (*rrlA*) or MG_RS02200 *(rrlB)* showed severe recombination defects. Therefore, we wondered whether they were a novel subset of recombination enzymes or they rather played a role in σ^{20} activation. Bioinformatics analyses indicated that neither RrlA nor RrlB showed DNA-binding or other conserved domains. However, null mutants of both genes did not allow the activation of the MG428 pathway and the peaks of σ^{20} expression were no longer observed. Moreover, our results showed that activation of the σ^{20} pathway increases σ^{20} stability in a feed-forward loop involving RrlA and RrlB. Overall, these results demonstrate a role for RrlA and RrlB in σ^{20} stability and, eventually, σ^{20} pathway activation.

In contrast to the observed with σ^{20}, overexpression of RrlA and RrlB in *M. genitalium* was not toxic, but most of the mutants isolated were non adherent and therefore presented phase variants. Despite this, the σ^{20} regulon was not upregulated in these mutants indicating that the σ^{20} pathway was not activated in the population. In this work, we have demonstrated that upregulation of the σ^{20} pathway enhances the generation of phase variants in the adhesin genes. Therefore, we can speculate that overexpression of RrlA and RrlB could also trigger σ^{20} activation.

Proteins stabilizing alternative sigma factors are not exclusive of *M. genitalium*. Three examples are Rsd (<u>r</u>egulator of <u>s</u>igma <u>D</u>), RssB (<u>r</u>egulator of <u>σ</u><u>S</u>) and Crl (<u>c</u>urli <u>r</u>egulatory <u>l</u>oci). Each of them represents a different model to stabilize/destabilize alternative sigma factors and to enhance the formation of active complexes with the RNA polymerase core enzyme.

- Rsd protein was isolated as a σ^{70}-sequestering protein in *E. coli*. It can bind to and remove the primary sigma factor in the RNAP holoenzyme, allowing other sigma factors to form complexes with the released RNAP core enzyme (Jishage and Ishihama, 1998; Mitchell *et al.*, 2007). However, unlike anti-sigma factors, overexpression of Rsd does not lead to lethal phenotypes nor Rsd null mutants have any impact on alternative sigma factor-dependent regulation (Mitchell *et al.*, 2007).

- RssB protein is a response regulator from *E. coli* and regulates σ^S proteolysis (Muffler *et al.*, 1996). Phosphorylated RssB can bind to σ^S and stimulate its recognition by the specialized protease ClpXP, promoting σ^S protein turnover. Dephosphorylation of RssB reduces its affinity for σ^S and therefore enhances its stability (Becker *et al.*, 1999).

- Crl protein is a σ^S-binding protein of *E. coli*. It interacts directly with the alternative sigma factor σ^S (Bougdour *et al.*, 2004) and facilitates RNA polymerase holoenzyme formation (Gaal *et al.*, 2006). Moreover, Crl protein can modulate σ^S susceptibility to other negative regulator factors such as RssB and reorganize transcription machinery towards σ^S-dependent promoters (Typas *et al.*, 2007).
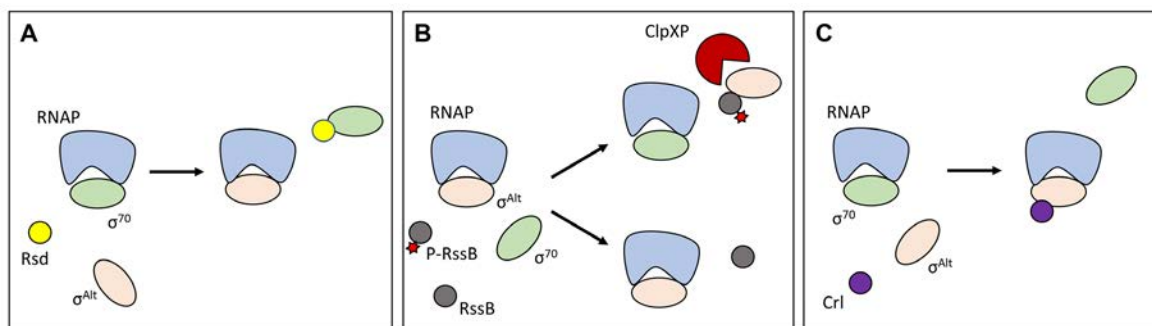
Figure D.5. Schematic representation of different mechanisms to stabilize alternative sigma factors. Primary sigma sequestration by Rsd **(A),** alternative sigma factor turnover promotion by phosphorylated RssB **(B)** and promotion of alternative holoenzyme formation by Crl **(C)** in *E. coli* are depicted.

Whether RrlA or RrlB fit in one or more of these models is yet unclear. Based on the behavior of the corresponding null mutants in single cell analyses, we can speculate a similar role to Rsd or Crl proteins, as null mutants for RssB in *E. coli* result in increased stabilization of σ^S (Muffler *et al.*, 1996). Furthermore, no specialized protease like ClpXP is present in the *M. genitalium* chromosome apparently, although moonlighting proteins or classic proteases could perform this function instead. Nonetheless, more experimental data is needed to elucidate the mechanism of action of RrlA and RrlB. Isolation of RNAP holoenzyme with σ^{20} might be useful to better characterize these σ^{20}-stabilizer proteins. Tandem Affinity Purification attempts with both the σ^{20} and RrlA and RrlB proteins were unsuccessful (see Appendices section S.2.4), probably due to the low intracellular amounts of σ^{20}-RNAP complexes in most of the cell population. Therefore, other strategies should be taken into consideration to determine the mechanism of action of these putative sigma auxiliary proteins.

Finally, the chromosome of *M. genitalium* codes for an anti-RpoA factor (MG_127, SpxA) and a putative sigma modulator protein (MG_022, RpoE). ECF-sigma factors are often regulated by anti-sigma factors that bind to them and prevent their interaction with the RNA polymerase (Helmann, 1999). At the same time, anti-sigma factors undergo similar complex regulatory mechanisms to release alternative sigma factors during a certain stress or in response to an extracellular signal (Paget, 2015). The Spx anti-RpoA factor is a global transcriptional regulator that interacts with the C-terminal region of the α subunit of the RNA polymerase and induces transcription of discrete genes involved in thiol homeostasis. At the same time, it represses growth-related cell functions in *B. subtilis* (Nakano *et al.*, 2003). In this way, Spx protein can serve both as a transcription factor and an anti-sigma factor. However, Spx protein of *M. genitalium* was found not to be related with the σ^{20}-pathway activation (see Appendices section S.2.3) and its exact role in transcriptional control remains obscure. On the other hand, *M. genitalium* also encodes for the cryptic RNA polymerase delta subunit RpoE. This subunit has been linked with promoter recognition and RNA polymerase turnover, but its exact role in model bacteria is still a enigmatic (Weiss and Shaw, 2015). However, the presence of the putative DNA-binding domain HARE-HTH in the RpoE subunit has raised the hypothesis that it could act as a sigma-like particle (Aravind and Iyer, 2012; Weiss and Shaw, 2015). Work is in progress to determine whether RpoE subunit can alter RNA polymerase recognition by the MG428 protein or redirect transcription to classic sigma-70 promoters. Anyhow, the presence of both *spxA* and *rpoE* in the minimal gene complement *M. genitalium* (Glass *et al.*, 2006) suggests a crucial role for these proteins in bacterial homeostasis.

D.2.3. Regulation of σ^{20} activity

In a wild-type genetic context, we have shown that the σ^{20}-regulon is activated in a cell-specific manner. Initially, we hypothesized that this pattern could be induced by the intermittent expression of σ^{20} or the regulated genes, probably by means of a subjacent transcriptional control. Nonetheless, overexpression of σ^{20} using two different constitutive promoters (MG_427 promoter, P_{MG_427} and MG_438 promoter, P_{MG_438}) did not result in a generalized expression of σ^{20} or the regulated genes. On the contrary, the characteristic intermittent activation pattern was still observed, although more cells showed activation at the same time. Altogether, we speculate that an unknown factor is responsible for σ^{20} activation through a yet unidentified mechanism. Once activated, σ^{20} will drive the expression of RrlA and RrlB, which in turn will stabilize the σ^{20} protein and induce transcription of the rest of the regulon.

Strikingly, very few examples of single cell activation of a sigma-mediated regulon are mentioned in the scientific literature. In *B. subtilis*, a similar "pulse regulation" of the alternative sigma σ^B has been described in individual cells upon stress (Locke *et al.*, 2011). However, eventually, the entire population becomes active upon continued exposure to the stressing condition. Miyazaki and colleagues described another example of subpopulation activation years ago. In their work, they found that integrative conjugative element (ICE) expression in *Pseudomonas* was activated by σ^S in ~5% of the population (Miyazaki *et al.*, 2012), opening a new field in transcriptional regulation of ICE activation.

Another interesting topic is how cells manage to turn off the hyperrecombinative status triggered by a fully active σ^{20}. We anticipate that the cell could target the RrlA and RrlB proteins for degradation and thus destabilize σ^{20}, or directly remove the alternative sigma factor from the RNA polymerase complex. Another hypothesis is that the σ^{20} protein could be outcompeted by the primary σ^A for RNAP binding. During the peaks of σ^{20} activation, if the σ^A function is impaired, transcription from the P_{MG_427} promoter could no longer support enough expression of σ^{20} and this would result in a progressive end of the σ^{20} peak. However, this passive strategy would have a huge impact on global transcription, which it appears not to be the case in mutants overexpressing σ^{20}.

Overall, taking together all the regulatory levels described in this section, we can conclude that regulation of alternative sigma factor use is astonishingly complex in bacteria, even in a genome-reduced pathogen like *M. genitalium*.

D.3. IMPACT OF THE MG428 REGULATORY PATHWAY IN *M. GENITALIUM* LIFESTYLE

D.3.1. Activation of the σ^{20} regulon *M. genitalium* triggers antigenic variation

Antigenic variation in *M. genitalium* has been characterized in depth throughout the recent years. Basically, it is based on the interchange between certain sequences of the cytadherence *mgpB* or *mgpC* genes and homologous regions present in the genome by means of double crossover events (Iverson-Cabral *et al.*, 2007). This sequence swap is highly dependent on the activity of the recombination machinery of *M. genitalium*, led by the RecA recombinase (Burgos *et al.*, 2012) and the Holliday junction resolvases RuvAB (Burgos and Totten, 2014a). In this work, we demonstrated that activation of the alternative

sigma factor σ^{20} stimulates sequence variation of the MgPa adhesins. The intrinsic activation of σ^{20} *in vitro* suggests that *M. genitalium* population is highly variable and cells display different MgPa variants despite the absence of the selective pressure exerted by an active immune system.

This inherent heterogeneity of the population seems critical to allow immune evasion of a few cells within the host and persistence of mycoplasma infections. Studies of *M. genitalium* infection in macaques show that antibodies produced during infection preferentially recognize peptides of the infecting strain instead of the predominant variants raised weeks after of the initial infection (Wood *et al.*, 2013). Therefore, the immune system displayed by the host is responsible for the selection of the emerging variant strains, which at the same time display other variations to face new adaptive immune responses. We hypothesize that this cycle is strictly dependent on the activation of the σ^{20} pathway, as it coordinates antigenic variation *in vitro*. Animal infection studies with *M. genitalium* ∆MG_428 mutant should be performed to elucidate the potential role of this alternative sigma factor in bacterial persistence and immune evasion.

D.3.2. Activation of the σ^{20} pathway promotes horizontal gene transfer

Horizontal gene transfer (HGT) is a source of genome innovation in bacterial genomes and can alter antibiotic susceptibility, metabolic pathways, virulence or cellular structures (Ochman *et al.*, 2000). HGT in Mollicutes has been identified in ruminant mycoplasma species, such as *M. agalactiae* or *M. bovis* (Tardy *et al.*, 2015), and in *M. pulmonis* (Teachman *et al.*, 2002). Most ruminant mycoplasmas encode integrative conjugative elements (ICE) in their small chromosomes that favor transfer of DNA to a recipient cell (Dordet Frisoni *et al.*, 2013). In the case of *M. pulmonis*, HGT was achieved using the non-conjugative Tn*4001* transposon in an unidentified transfer mechanism that was not related to homologous recombination but to a transposition event (Teachman *et al.*, 2002). Despite HGT has only been stablished in a few Mycoplasma species, high-throughput analyses of several genomes suggest that HGT might not be marginal, but play a significant role in genome evolution of these small bacteria (Sirand-Pugnet *et al.*, 2007).

In this work, we have observed transfer of antibiotic markers by homologous recombination between two distinct *M. genitalium* mutant strains. DNA transfer was unidirectional and DNaseI-resistant. Moreover, we could only detect DNA transfer when the donor strain overexpressed σ^{20}; previous attempts with strains with normal levels of σ^{20} were

unsuccessful. Unlike HGT in *M. pulmonis*, DNA transfer in *M. genitalium* relied on the homologous recombination system.

This means of HGT in bacteria lacking conjugative plasmids or ICE is rare in prokaryotes. Typically, specific machinery encoded in these mobile elements is essential for DNA transfer, and very few chromosomal genes are involved in this process. For instance, most ICEs encode specific integrases for ICE-excision and integration, proteins to stabilize extrachromosomal ICE forms, relaxases for rolling-circle replication, secretion systems and even regulatory proteins (Wozniak and Waldor, 2010). However, *M. genitalium* does not encode ICE-like machinery or vestigial proteins that could mimic integrative conjugative element transmission; instead, it seems that HGT is catalyzed by specific cellular machinery. Recently, horizontal transfer of large chromosomic regions not-subjected to any previously known conjugative system has been described in Mycobacteria (Wang *et al.*, 2005; Gray *et al.*, 2013).
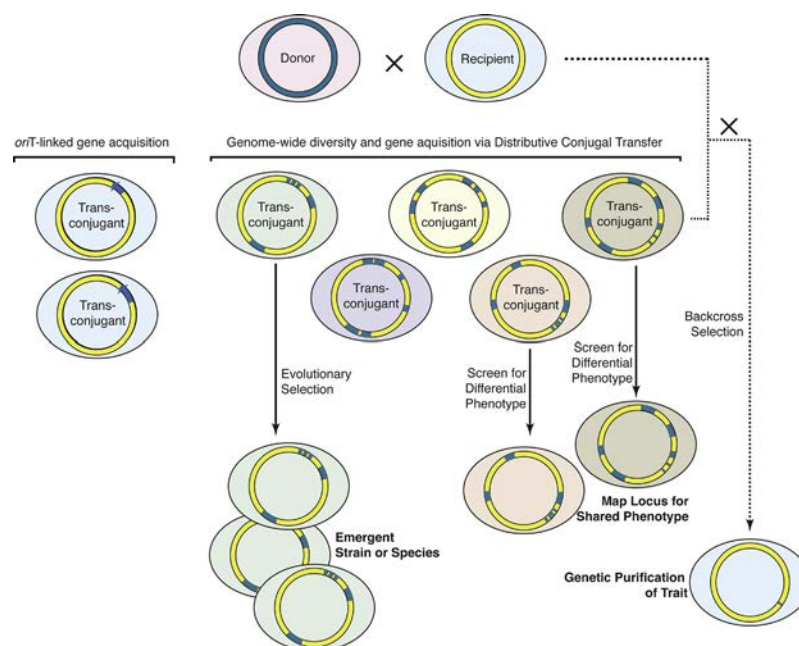
Figure D.6. Graphic summary of the evolutionary and gene mapping potential of distributive conjugal transfer in comparison to oriT-mediated transfer. The parental donor and recipient strains are schematically shown at the top, with their native chromosomes (blue and yellow circles, respectively) that confer different phenotypes (pink and blue backgrounds, respectively). Co-incubation of the donor and recipient strains on solid media (agar plates) or in a biofilm, permits conjugation. For oriT-mediated transfer (left), all transferred segments of DNA are linked to oriT, which limits the extent of genetic diversity among the transconjugants. This contrasts with distributive conjugal transfer (DCT), wherein random segments of the donor chromosome are transferred to the recipient, generating unique transconjugants. Each transconjugant has a novel genotype that confers a unique phenotypic profile (different colored background). Extracted from Gray *et al.*, 2013.

Chromosomal DNA transfer (CT) in *Mycobacterium smegmatis* is not dependent on any *oriT* sequence and is apparently locus-independent, although it is dependent on the

presence of the mating identity locus (*mid*), and the *esx1* gene, that codes for a Type VII secretion system (Gray *et al.*, 2013, 2016). As this chromosomal transfer is not limited by the *oriT* sequence, the authors named it as distributive conjugal transfer (DCT) (see Figure D.6). The existence of this alternative system in *M. smegmatis* suggests that other mechanisms rather than classic conjugation machinery may be present in bacteria.

Despite the uniqueness of HGT in *M. genitalium*, a similar recombination-based CT has been identified in *M. agalactiae* and *M. bovis*. ICE-dependent-recombination of homologous sequences from different strains of *M. agalactiae* has been observed (Dordet-Frisoni *et al.*, 2014). However, instead of transferring chromosomic DNA from ICE-positive to ICE-negative cells, chromosomal transfers occur in the opposite direction, suggesting that ICE-encoded machinery might be used merely as a mean to allow cell-to-cell communication (Dordet-Frisoni *et al.*, 2014). In the absence of ICE, these chromosomal transfers could only be observed in the presence of a membrane fusion agent (PEG). This system contrasts with the CT observed in *M. genitalium*, as no membrane fusion agent or ICE-mediated transfer was necessary, suggesting that any required component for cell-to-cell communication might be encoded in the genome.

Apparently, no differences in cell morphology were observed in strains overexpressing the σ^{20} pathway. However, the inherent pleomorphism and propensity to form aggregates of *M. genitalium* cells could be in part derived from differential membrane traits displayed upon σ^{20} activation. In this sense, major Gram-positive bacteria display specific cell-to-cell contact mechanisms to promote bacterial communication, such as the recently characterized bacterial nanotubes (Dubey and Ben-Yehuda, 2011; Dubey *et al.*, 2016). *B. subtilis* cells could transfer proteins, intracellular molecules and even non-conjugative plasmids through these small capillaries (Dubey and Ben-Yehuda, 2011). Whether a conduct of a similar nature might be used by *M. genitalium* cells for CT will be addressed in the future.

In summary, the presence of a HGT mechanism to promote genome evolution should be considered as a new means for rapid adaptation and persistence during infection by *M. genitalium*. Together with reciprocal recombination of the MgPa operon sequences and the putative protective activity of protein M, this system could provide a new molecular pathway to propagate variants to better face specialized immune response triggered by the host.

Overall, in this work we have shown that the alternative sigma factor σ^{20} is a major regulator of recombination activation, a trigger factor for antigenic variation and an inducer of

horizontal gene transfer in *M. genitalium*, but might also be controlling other unknown cellular processes. Figure D.7 summarizes the highlights of the σ^{20} regulatory pathway described in this work.
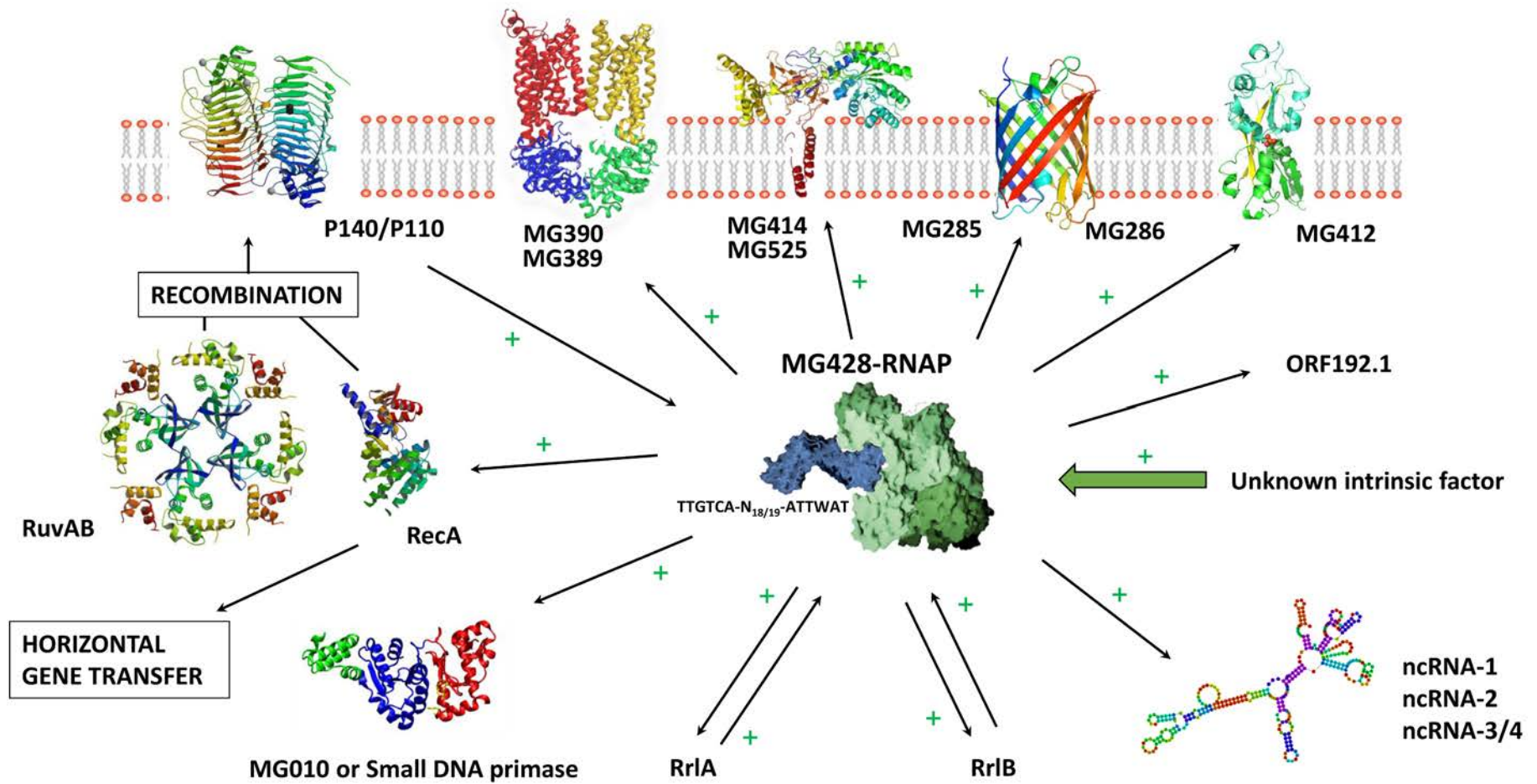
Figure D.7. Graphic summary of the σ²⁰ pathway and the regulatory networks stablished in this work. Protein structures depicted in this figure are figurative and are based on real structures of the related proteins from other bacteria. Secondary structure of the ncRNAs does not correspond to the real conformation of the regulated ncRNAs stated in this work.

CONCLUSIONS

Chapter I. MG428 is an alternative sigma factor that promotes antigenic variation in *M. genitalium*

1. Overexpression of MG_428 in *M. genitalium* via transposon delivery leads to a toxic phenotype, only circumvented by transposon insertion in the antisense direction in specific chromosomal loci.

2. Protein levels of σ^{20} are low in WT cells and are only detectable in a small subset of the bacterial population.

3. Overexpression of σ^{20} activates transcription of a unique regulon composed by known recombination genes (*recA*, *ruvAB*) as well as genes with unknown function (MG_010, MG_220, MG_RS02200, MG_285, MG_286, MG_389, MG_390, MG_412, MG_414 and MG_525).

4. The σ^{20} regulon also comprises three distinct non-coding RNAs, named ncRNA-1, ncRNA-2 and ncRNA-3/4, with an unknown function.

5. A unique conserved promoter with σ^{70} architecture can be found in the upstream region of all the regulated genes and ncRNAs. This conserved sequence is formed by a classic -35 and a novel -10 element, which are also conserved in *M. pneumoniae*.

6. σ^{20} binds to the RNA polymerase core enzyme, which is devoid of any sigma factor.

7. Activation of the σ^{20} pathway only takes place in a small subset of cells at the same time.

8. Null mutants of MG_428 and scrambled mutations at the promoter regions of MG428-regulated genes abrogate single cell activation of the σ^{20} pathway.

9. Null mutants of MG_428 have highly reduced capacity to incorporate exogenous DNA by homologous recombination.

10. Strains with scrambled mutations at the σ^{20}-regulated -10 and -35 element of *recA* show severe recombination defects.

11. Null mutant of MG_428 have reduced capacities to generate MgPa variants, while mutants overexpressing this regulator have increased nucleotide variation in the MgPa operon as well as in the MgPa regions.

Chapter II. RrlA and RrlB are two proteins that stabilize the alternative sigma factor σ²⁰ of *M. genitalium*

12. RrlA and RrlB are expressed proteins of *M. genitalium* under the control of the alternative sigma factor σ^{20}.

13. orf192.1 is a non-annotated protein coding sequence of *M. genitalium* that is encoded in the σ^{20}-regulated ncRNA-2.

14. *rrlA* and *rrlB* null mutants prevent σ^{20} activation in the bacterial population, as well as single cell activation of σ^{20}-regulated genes.

15. σ^{20} overexpression induces its own stabilization via RrlA and RrlB proteins.

16. RrlA and RrlB overexpression does not increase the number of percentage of σ^{20}-activated cells in adherent strains, but promotes the generation of non-adherent variants.

17. σ^{20} overexpression restores recombination capacity to *rrlA* or *rrlB* mutants, but is unable to restore recombination of a *recA* mutant.

18. *M. genitalium* can transfer DNA from one strain to another via an unknown system that is dependent on high levels of σ^{20} and a proficient recombination pathway in the donor strain.

EXPERIMENTAL PROCEDURES

E.1. BIOLOGIC MATERIAL

E.2. DNA MANIPULATION

E.3. RNA MANIPULATION AND TRANSCRIPTOMIC METHODS

E.4. PROTEIN ANALYSIS METHODS

E.5. EPIFLUORESCENCE MICROSCOPY

E.1. BIOLOGIC MATERIAL

E.1.1. Bacterial strains

All bacterial strains used in this work are listed below. *E. coli* strains were preserved in LB medium with 15% glycerol (v/v) at -80ºC. *M. genitalium* strains were kept in SP4 medium without adding glycerol directly at -80ºC, as the high concentration of fetal bovine serum present in the culture medium acts as a cryoprotector.

Table E.1. Escherichia coli strains used in this work

XL1-Blue	Strain used for cloning purposes. It is deficient in recombination and has a high transformation efficiency. Complete genotype: *endA*1 *gyrA*96(nal^R) *thi*-1 *recA*1 *relA*1 *lac glnV*44 *hsdR*17(r$_K^-$ m$_K^+$) F' [::Tn10 *pro*AB$^+$ *lacI*^{q} Δ(*lacZ*)M15]. It is resistant to nalidixic acid and tetracycline.
BL21(DE3)	Strain used for recombinant protein expression. It contains the RNA polymerase from phage T7 under the control of the *lac*UV5 promoter sequence, which is inducible with IPTG. This feature allows high expression levels of genes cloned under the control of the T7 promoter, present for example in pET protein expression plasmids. Complete genotype: F⁻ *ompT gal dcm lon hsdS$_B$(r$_B$⁻m$_B$⁻)* λ(DE3[*lacI lacUV5-T7 p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S).

Table E.2. Mycoplasma genitalium strains used in Chapter I

WT or G37	*Mycoplasma genitalium* G37 is the reference strain. It was isolated from a male patient with non-gonococcal urethritis (Tully *et al.*, 1983). Source: ATCC 33530.
ΔMG_428	This strain lacks the MG_428 gene and it is resistant to tetracycline. Genotype: ΔMG_428::*tetM*. Source: this work.
Tn::MG_281	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the MG_281 gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_281::Tn*Cat*MG_428. Source: this work.
Tn::*recA-2*	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the *recA* gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_339::Tn*Cat*MG_428. Source: this work.
Tn::*ruvB*	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the *ruvB* gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_359::Tn*Cat*MG_428. Source: this work.
Tn::MG_220	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the MG_220 gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_220::Tn*Cat*MG_428. Source: this work.
Tn::MG_191	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the *mgpB* gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_191::Tn*Cat*MG_428. Source: this work.

Tn::MG_192	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the *mgpC* gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_191::Tn*Cat*MG_428. Source: this work.
Tn::MG_390	This strain is a defective mutant for MG_428 gene complemented with the Tn*CatMG_428* MiniTnp inserted in the *sunT* gene. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_428::*tetM*, MG_390::Tn*Cat*MG_428*. Source: this work.
Cat:Ch	This strain bears a transcriptional fusion between *cat* gene and mCherry fluorescent protein gene under the control of MG_438 promoter. This fusion was inserted between MG_438 and MG_439 genes by allelic exchange. It is resistant to chloramphenicol. Genotype: *cat*:mCherry. Source: Broto A., unpublished.
RecA:Ch	This strain bears a transcriptional fusion between *recA* gene and mCherry fluorescent protein gene at the chromosomal locus. It is resistant to chloramphenicol. Genotype: *cat*, *recA*:mCherry. Source: this work.
MG_428:Ch	This strain bears a transcriptional fusion between MG_428 gene and mCherry fluorescent protein gene at the chromosomal locus. It is resistant to chloramphenicol. Genotype: *cat*, MG_428:mCherry. Source: this work.
RecA:Ch ΔMG_428	This strain was obtained modifying the RecA:Ch strain. It lacks the MG_428 gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_428::*tetM*, *cat*, *recA*:mCherry. Source: this work.
RecA:Ch -10	This strain bears a transcriptional fusion between *recA* gene and mCherry fluorescent protein gene at the chromosomal locus. Moreover, Pribnow box sequence is scrambled. It is resistant to chloramphenicol. Genotype: *cat*, *recA*:mCherry, scrambled -10 sequence. Source: this work.
RecA:Ch -22	This strain bears a transcriptional fusion between *recA* gene and mCherry fluorescent protein gene at the chromosomal locus. Moreover, sequence between the -35 and -10 elements is scrambled. It is resistant to chloramphenicol. Genotype: *cat*, *recA*:mCherry, scrambled -22 sequence. Source: this work.
RecA:Ch -35	This strain bears a transcriptional fusion between *recA* gene and mCherry fluorescent protein gene at the chromosomal locus. Moreover, -35 element is scrambled. It is resistant to chloramphenicol. Genotype: *cat*, *recA*:mCherry, scrambled -35 sequence. Source: this work.
ΔMG_010	This strain lacks the MG_010 gene and it is resistant to tetracycline. Genotype: ΔMG_010::*tetM*. Source: this work.
ΔMG_220	This strain lacks the MG_220 gene and it is resistant to tetracycline. Genotype: ΔMG_220::*tetM*. Source: this work.
ΔMG_390	This strain lacks the MG_390 gene and it is resistant to tetracycline. Genotype: ΔMG_390::*tetM*. Source: this work.
ΔMG_414	This strain lacks the MG_414 gene and it is resistant to tetracycline. Genotype: ΔMG_414::*tetM*. Source: this work.
ΔMG_525	This strain lacks the MG_525 gene and it is resistant to tetracycline. Genotype: ΔMG_525::*tetM*. Source: this work.
ΔMG_RS02200	This strain lacks the MG_RS02200 gene and it is resistant to tetracycline. Genotype: ΔMG_RS02200::*tetM*. Source: this work.

*This mutant carries a truncated copy of the MG_428 gene coded in the Tn*CatMG_428* MiniTnp.

Table E.3. *Mycoplasma genitalium* strains used in Chapter II

WT or G37	See description in Table E.2
ΔMG_428	See description in Table E.2
Tn::*recA*-2	See description in Table E.2
MG_427:Ch	This strain bears a transcriptional fusion between MG_427 gene and mCherry fluorescent protein gene at the chromosomal locus. It is resistant to chloramphenicol. Genotype: *cat*, MG_427:mCherry. Source: this work.
MG_428:Ch Δ*recA*	This strain was obtained modifying the MG_428:Ch strain. It lacks the *recA* gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_339::*tetM*, *cat*, MG_428:mCherry. Source: this work.
MG_428:Ch Δ*recA* Tn*Pac*MG_428	This strain was obtained modifying the MG_428:Ch Δ*recA* strain introducing a MiniTnp bearing an ectopic copy of MG_428 under the control of its own promoter. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_339::*tetM*, *cat*, MG_428:mCherry Tn*Pac*MG_428 (pool). Source: this work.
RrlA:Ch	This strain bears a transcriptional fusion between *rrlA* gene (MG_220) and mCherry fluorescent protein gene at the chromosomal locus. It is resistant to chloramphenicol. Genotype: *cat*, *rrlA*:mCherry. Source: this work.
RrlB:Ch	This strain bears a transcriptional fusion between *rrlB* gene (MG_RS022000) and mCherry fluorescent protein gene at the chromosomal locus. It is resistant to chloramphenicol. Genotype: *cat*, *rrlB*:mCherry. Source: this work.
RrlA:Ch ΔMG_428	This strain was obtained modifying the RrlA:Ch strain. It lacks the MG_428 gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_428::*tetM*, *cat*, *rrlA*:mCherry. Source: this work.
RrlB:Ch ΔMG_428	This strain was obtained modifying the RrlB:Ch strain. It lacks the MG_428 gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_428::*tetM*, *cat*, *rrlB*:mCherry. Source: this work.
RrlA:Ch ΔMG_428 Tn*Pac*MG_428	This strain was obtained modifying the RrlA:Ch ΔMG_428 strain. It carries a MiniTnp with the MG_428 gene under the control of its own promoter inserted in the MG_281 gene. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_428::*tetM*, *cat*, *rrlA*:mCherry MG_281::Tn*Pac*MG_428. Source: this work.
RrlB:Ch ΔMG_428 Tn*Pac*MG_428	This strain was obtained modifying the RrlB:Ch ΔMG_428 strain. It carries a MiniTnp with the MG_428 gene under the control of its own promoter inserted in the *recA* gene. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_428::*tetM*, *cat*, *rrlA*:mCherry MG_339::Tn*Pac*MG_428. Source: this work.
MG_428:Ch Δ*rrlA*	This strain was obtained modifying the MG_428:Ch strain. It lacks the *rrlA* gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_220::*tetM*, *cat*, MG_428:mCherry. Source: this work.
MG_428:Ch Δ*rrlB*	This strain was obtained modifying the MG_428:Ch strain. It lacks the *rrlB* gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_RS02200::*tetM*, *cat*, MG_428:mCherry. Source: this work.
MG_428:Ch Δ*rrlA* Tn*Pac-rrlA*	This strain was obtained modifying the MG_428:Ch Δ*rrlA* strain. It carries the *rrlA* gene in a Tn*Pac* MiniTnp under the control of its own promoter. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_220::*tetM*, *cat*, MG_428:mCherry Tn*Pac-rrlA*. Source: this work.
MG_428:Ch Δ*rrlB* Tn*Pac-rrlB*	This strain was obtained modifying the MG_428:Ch Δ*rrlB* strain. It carries the *rrlB* gene in a Tn*Pac* MiniTnp under the control of its own promoter. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_RS02200::*tetM*, *cat*, MG_428:mCherry Tn*Pac-rrlB*. Source: this work.
RecA:Ch Δ*rrlA*	This strain was obtained modifying the RecA:Ch strain. It lacks the *rrlA* gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_220::*tetM*, *cat*, *recA*:mCherry. Source: this work.
RecA:Ch Δ*rrlB*	This strain was obtained modifying the RecA:Ch strain. It lacks the *rrlB* gene and it is resistant to chloramphenicol and tetracycline. Genotype: ΔMG_RS02200::*tetM*, *cat*, *recA*:mCherry. Source: this work.
RecA:Ch Δ*rrlA* Tn*Pac-rrlA*	This strain was obtained modifying the RecA:Ch Δ*rrlA* strain. It carries the *rrlA* gene in a Tn*Pac* MiniTnp under the control of its own promoter. It is resistant to

	chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_220::*tetM*, *cat*, *recA*:mCherry Tn*Pac-rrlA*. Source: this work.
RecA:Ch Δ*rrlB* Tn*Pac-rrlB*	This strain was obtained modifying the RecA:Ch Δ*rrlB* strain. It carries the *rrlB* gene in a Tn*Pac* MiniTnp under the control of its own promoter. It is resistant to chloramphenicol, tetracycline and puromycin. Genotype: ΔMG_220::*tetM*, *cat*, *recA*:mCherry Tn*Pac-rrlB* Source: this work.
TnMG_428:YFP	This strain bears a transcriptional fusion between MG_428 gene and eYFP fluorescent protein gene in a Tn*Pac* MiniTnp under the control of the MG_438 promoter. It is resistant to puromycin. Genotype: MG_343::Tn*Pac*-P438MG_428:YFP. Source: this work.
MG_428:Ch TnMG_428:YFP	This strain was obtained modifying the MG_428:Ch strain. It carries the MG_428 gene fused with the eYFP gene in a Tn*Pac* MiniTnp under the control of the MG_438 promoter. It is resistant to chloramphenicol and puromycin. Genotype: *cat*, MG_428:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP. Source: this work.
RecA:Ch TnMG_428:YFP	This strain was obtained modifying the RecA:Ch strain. It carries the MG_428 gene fused with the eYFP gene in a Tn*Pac* MiniTnp under the control of the MG_438 promoter. It is resistant to chloramphenicol and puromycin. Genotype: *cat*, *recA*:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP. Source: this work.
RecA:Ch TnMG_428:YFP Δ*rrlA*	This strain was obtained modifying the RecA:Ch TnMG_428:YFP strain. It lacks the *rrlA* gene and it is resistant to chloramphenicol, puromycin and tetracyciline. Genotype: *cat*, *recA*:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP, ΔMG_220::*tetM*. Source: this work.
RecA:Ch TnMG_428:YFP Δ*rrlB*	This strain was obtained modifying the RecA:Ch TnMG_428:YFP strain. It lacks the *rrlB* gene and it is resistant to chloramphenicol, puromycin and tetracyciline. Genotype: *cat*, *recA*:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP, ΔMG_RS02200::*tetM*. Source: this work.
RecA:Ch TnMG_428:YFP TnTet-*rrlA*TAP	This strain was obtained modifying the RecA:Ch TnMG_428:YFP strain. It bears a MiniTnp carrying a transcriptional fusion between *rrlA* and a TAP tag under the control of MG_427 promoter. It is resistant to chloramphenicol, puromycin and tetracyciline. Genotype: *cat*, *recA*:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP, Tn*TetM*-P427*rrlA*:TAP. Source: this work.
RecA:Ch TnMG_428:YFP TnTet-*rrlB*TAP	This strain was obtained modifying the RecA:Ch TnMG_428:YFP strain. It bears a MiniTnp carrying a transcriptional fusion between *rrlB* and a TAP tag under the control of MG_427 promoter. It is resistant to chloramphenicol, puromycin and tetracyciline. Genotype: *cat*, *recA*:mCherry, MG_343::Tn*Pac*-P438MG_428:YFP, Tn*TetM*-P427*rrlB*:TAP. Source: this work.
Δ*rrlA*	Described in Table E.2 as ΔMG_220
Δ*rrlB*	Described in Table E.2 as ΔMG_RS02200
Δ*recA*	This strain lacks the *recA* gene and it is resistant to tetracycline. Genotype: ΔMG_339::*tetM*. Source: this work.
Δ*rrlA* Tn*Cat*MG_428	This strain was obtained modifying the Δ*rrlA* strain. It carries the Tn*Cat*MG_428 MiniTnp, used in complementation of the ΔMG_428 strain in Chapter I. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_220::*tetM*, MG_281::Tn*Cat*MG_428. Source: This work.
Δ*rrlB* Tn*Cat*MG_428	This strain was obtained modifying the Δ*rrlB* strain. It carries the Tn*Cat*MG_428 MiniTnp, used in complementation of the ΔMG_428 strain in Chapter I. It is resistant to tetracycline and chloramphenicol. Genotype: ΔMG_RS02200::*tetM*, MG_281::Tn*Cat*MG_428. Source: This work.
Δ*recA* Tn*Cat*MG_428	This strain was obtained modifying the Δ*recA* strain. It carries the Tn*Cat*MG_428 MiniTnp, used in complementation of the ΔMG_428 strain in Chapter I. It is resistant to tetracycline and chloramphenicol. Genotype: Δ*recA*::*tetM*, Tn*Cat*MG_428 (pool). Source: This work.

E.1.2. Culture mediums and bacterial growth conditions

E.1.2.1. Culture of *E. coli* strains

LB (Lysogeny Broth) medium was used for most applications. It is a nutritionally rich broth ideal to cultivate recombinant strains of *E. coli* for molecular microbiology purposes (Sambrook *et al.*, 1989). The composition of LB broth is 10g/L tryptone (Scharlau Microbiology), yeast extract 5g/L (Scharlau Microbiology) and NaCl 10g/L (Sigma-Aldrich). For LB agar plates, 1.5% bacto-agar (w/v) (Scharlau Microbiology) was added before sterilization. LB medium was autoclaved at 121ºC for 15 minutes. Plates were kept at 4ºC for long term storage.

SOB (Super Optimal Broth) medium was used to grow XL1-Blue during competent cell preparation. The composition of SOB is 20g/L tryptone (Scharlau Microbiology), 5g/L yeast extract (Scharlau Microbiology), NaCl 0.5g/L (Sigma-Aldrich), 2.5mM KCl and pH 7 adjusted with 2M KOH. SOB was autoclaved at 121ºC for 15 minutes. Sterile $MgCl_2$ at a final concentration of 2mM was added fresh prior growing any culture.

Supplements and antibiotics were added to sterile LB when necessary. In the case of LB plates, all components were added to warm LB-agar solution (50-55ºC) to prevent antibiotic inactivation and the mix poured onto sterile plates.

The following LB supplements were used in this work:
- Ampicillin (Sigma-Aldrich). Stock solution was prepared at 200mg/mL in sterile MilliQ H_2O and filtered through a 0.22μm membrane filter (Merck-Millipore) and stored at -20ºC. Ampicillin was added to a final working concentration of 100μg/mL. Resistance to this antibiotic is conferred by the *bla* gene, coding for a β-lactamase, present in all the plasmids used in this work.
- Tetracycline (Sigma-Aldrich). Stock solution was prepared at 5mg/mL in absolute ethanol in a sterile environment and kept in the dark at -20ºC. Working concentration of tetracycline for XL1-Blue selection was 10μg/mL. Resistance to tetracycline is derived from the activity of *tetA* and *tetR* genes.
- Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich). IPTG was used to induce transcription of genes under the control of *lac*-derived promoters. Stock solution was prepared at 1M using sterile MilliQ H_2O and stored at -20ºC. IPTG was used at working concentrations of 0.5-1mM.

- 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (VWR Chemicals). This compound is an analogue of lactose that derives in a blue insoluble chemical upon hydrolyzation. This feature allows simple detection of β-galactosidase activity from *lacZ*-coding plasmids in XL1-Blue colonies. Stock solution was prepared fresh dissolving 40mg of X-Gal in 1mL N,N-Dimethylformamide. It was only used in LB plates at a final concentration of 40μg/mL. X-Gal-containing plates were stored in the dark as this component is photosensitive.

To grow all *E. coli* strains in liquid cultures, tubes were shaken at 250rpm in an orbital shaking incubator at 37ºC O/N, keeping 1:5 volume proportion in the tubes for proper aeration. 10μL of working stocks were used to inoculate 3-50mL of LB. For larger volumes, a starter culture was set until OD_{600}~1 and used for inoculation (typically 1% of the final culture volume).

LB agar plates were used for colony isolation upon transformation. Colonies were picked up using sterile wooden inoculation sticks and propagated O/N in 3mL LB (with ampicillin 100μg/mL) in sterile culture tubes.

E.1.2.2. Culture of *M. genitalium* strains

Spiroplasma medium 4 (SP4) broth was used to grow all *M. genitalium* strains of this work. It is a very rich medium recommended for culturing mycoplasmas and spiroplasmas (Tully *et al.*, 1979). SP4 must be prepared in two sequential steps:

First; a medium base is prepared, autoclaved and finally supplemented to obtain the fully complemented broth. Base is composed of PPLO 3.5g/L (Pleuro-pneumoniae like organism) broth (Becton Dickinson), Tryptone 10g/L (Becton Dickinson), Bacto-peptone 5.3g/L (Becton Dickinson) and Glucose 5g/L (Sigma-Aldrich) dissolved in MilliQ H_2O. pH is adjusted to 7.6-7.8 using 2M NaOH solution. To prepare SP4 agar plates, 0.8% bactoagar (Becton Dickinson) was added to the mixture prior autoclaving. The SP4 base was autoclaved at 121ºC for 15 minutes.

Before supplementation, autoclaved SP4 base was cooled down (to 55ºC in case of SP4-agar solution). For 1L SP4 broth preparation 12mL 0.1% Phenol red (w/v) pH 7 (Sigma-Aldrich), 100mL 2% Yeastolate (w/v) (Becton Dickinson), 50mL 10x CMRL (Life Technologies), 170mL Fetal Bovine Serum (Life Technologies), 34mL yeast extract

(handmade), 3.42mL Glutamine 29.2mg/mL (Sigma-Aldrich) and 500μL ampicillin 200mg/mL (Sigma-Aldrich) were added to the base. SP4 pH was readjusted to 7.8 using NaOH 2M solution. To avoid the presence of killed yeast cells in the final broth, it was filtered through a 0.22μm Stericup® filtering system (Merck-Millipore).

To prepare SP4 agar plates, all supplements were mixed up together and warmed up to 55ºC. Then, SP4 base was supplemented and poured onto 55mm sterile plates (typically, 6mL per plate). Antibiotics and other additives were added right before plating.

Fetal Bovine Serum (FBS) was heated up to 56ºC for 30 minutes to inactivate the complement system. Then, FBS was aliquoted and stored at -20ºC. For the yeast extract preparation, 250g of fresh yeast was dissolved in 1L of MilliQ H_2O and autoclaved at 115ºC for 10 minutes. Then, it was centrifuged at low speed (400g) for 10 minutes and pellet was discarded. Supernatants were autoclaved again at 115ºC for 10 minutes, aliquoted and stored at -20ºC. Phenol Red and Yeastolate solutions were autoclaved at 121ºC for 15 minutes before use.

The following additives were added to SP4 broth when necessary:
- Tetracycline (Sigma-Aldrich). Stock solution was prepared as described before. Tetracycline was used at 3-5μg/mL for mutant selection and kept in the dark. Tetracycline resistance in *M. genitalium* is derived from the presence of *tetM438* marker (Pich *et al.*, 2006b).
- Chloramphenicol (Sigma-Aldrich). Stock solution was prepared at 34mg/mL in 70% ethanol and kept at -20ºC. For mutant selection, chloramphenicol was used at concentrations ranging 17-68μg/mL. *cat* marker provides antibiotic resistance to chloramphenicol in *M. genitalium* (Calisto *et al.*, 2012).
- Puromycin (Life Technologies). Puromycin stock solution (10mg/mL) was stored at -20ºC and kept away from the light, following manufacturer's instructions. Puromycin was used at 3μg/mL to select *M. genitalium* mutants. Puromycin resistance arises from the *pac* marker under the control of MG_438 promoter region (this work).
- X-Gal (VWR). X-Gal was dissolved in 1mL of N,N-Dimethylformafide as described before. For β-galactosidase activity detection in SP4 plates, X-Gal solution was added at a final concentration of 150μg/mL as in a previous work (Lluch-Senar *et al.*, 2007).

M. genitalium strains were grown in 5, 20 or 35mL of SP4 in tissue culture flasks of 25, 75 or 175 cm^2 (SPL Life Sciences) respectively at 37ºC and 5% CO_2. Cultures were kept in these conditions until mid-log phase (normally 3-7 days); then they were scrapped off or centrifuged depending on the subsequent application and the adherence to the flask.

For *M. genitalium* solid cultures, 100-200μL were plated onto SP4 agar plates and kept at 37ºC 5% CO_2 for 12-18 days. Colonies were spotted using a binocular stereomicroscope, picked up using sterile cut microtips and propagated in 5mL of SP4 with the selection marker.

All the manipulations of living mycoplasmas were performed in a BS-2 laminar flow hood.

E.1.3. Bacterial transformation

E.1.3.1. *E. coli* transformation

E. coli competent cells

E. coli XL1-Blue cells are not naturally competent. For competence induction in this strain, the inoue method was followed as described in Sambrook, Fritsch and Maniatis, 1989. Briefly, two XL1-Blue colonies were propagated in 10mL LB with tetracycline (10μg/mL) at 37ºC and 250rpm for 6-8 hours as a starter culture. Then, 3x10mL of SOB broth were inoculated with 100μL each and incubated O/N at 37ºC, 250rpm. Then, 3x250mL SOB were inoculated with 2, 4 and 10mL of the 3x10mL culture, respectively; and incubated at 20ºC, 250rpm for ~20h.

Cultures were grown until OD_{600} ~0.5-0.6. Afterwards, cultures were kept on ice for 10 minutes and centrifuged at 3000*g* for 15 minutes at 4ºC. Pellet was vacuum dried, swirled in 80mL of ice-cold inoue buffer and centrifuged again at 3000*g* for 10 minutes at 4ºC. Pellets were swirled in 20mL ice-cold inoue buffer and 1.5mL of DMSO. Then, 100μL aliquots were flash-frozen in liquid nitrogen and stored at -80ºC.

Inoue transformation buffer is composed by 55mM $MnCl_2$, 15mM $CaCl_2$, 250mM KCl and 10mM PIPES pH 6.7. It was filter-sterilized, aliquoted and stored at -20ºC.

Competence of *E. coli* BL21(DE3) strain was induced with the $CaCl_2$ method, also described in Sambrook, Fritsch and Maniatis, 1989. Briefly, pelleted BL21(DE3) cells were

subsequently washed in ice-cold $CaCl_2$-$MgCl_2$ solution (80mM $MgCl_2$, 20mM $CaCl_2$) and resuspended in ice-cold 0.1M $CaCl_2$. Cells were aliquoted and stored at -80ºC.

Heat-shock transformation protocol

Transformation of *E. coli* strains was performed following standard heat-shock protocol, described in Sambrook, Fritsch and Maniatis, 1989. For *E. coli* XL1-Blue strains, cells were heated at 42ºC for 90 seconds, while BL21(DE3) strains exposure time was reduced to 45 seconds.

E.1.3.2. *M. genitalium* transformation and screening for mutants

A protocol to transform *M. genitalium* by electroporation was first described in 1996 (Reddy *et al.*, 1996). In this work, a modified protocol has been used to create all *M. genitalium* mutant strains and it is detailed here:

- An *M. genitalium* culture was set in a 75cm^2 flask with 20mL fresh SP4 medium at 37ºC 5% CO_2.
- Typically, the culture reached the mid-log phase after 3 days. 10mL of SP4 were discarded and cells were scrapped off in the remaining volume of SP4. Then, they were filtered through a 0.45μm low protein binding filter to disaggregate cell clusters. Afterwards:
    a) To transform with one or two plasmids, 3-5mL of the disaggregated cells were seeded in a new 75cm^2 flask with 15-17mL fresh SP4 medium.
    b) To transform several plasmids, all the filtered cells were seeded in a 175cm^2 flask with 25mL fresh SP4 medium for a final volume of ~35mL.
- Cell passage was left O/N at 37ºC 5%CO_2. Next, cells were washed three times with electroporation buffer (0.272M Sucrose, 8mM HEPES pH 7.2) and scrapped off in 300μL or 1.5mL if cells were seeded in a 75cm^2 or 175cm^2 flask, respectively.
- 100μL of cells were mixed with 5 or 30μg of plasmid DNA, depending if a minitransposon or a suicide plasmid for gene replacement mutagenesis was used, respectively.
- Immediately, the mixture was placed in a 0.2mm gapped electroporation cuvette and electroporated at 2500V, 250Ω and 25mF. Normally, electroporation time constants in these conditions were 5.0-6.8 milliseconds.
- Then, cuvettes were kept on ice for 15 minutes. Afterwards, cells were resuspended in 950μL of fresh SP4 and transferred in a microtube. Tubes were left at 37ºC 5% CO_2 for at least 2 hours to let phenotypic expression of the resistance marker. Next,

a) Cells could be plated on SP4 agar to isolate colonies. Normally, serial dilutions were plated in the case of transformation by transposon delivery; cells transformed by homologous recombination were not diluted.

b) A mutant pool could be obtained seeding cells in a 75cm^2 flask with 20mL of fresh SP4 with the corresponding antibiotic.

- Colonies could be spotted 12-16 days after electroporation in agar plates.

To transform non-adherent mutants of *M. genitalium*, a similar procedure was performed. The starter culture was not filtered through a 0.45μm filter, instead it was forced through a 25G needle syringe several times. Furthermore, as cells were not adhered to the plastic surface, washes with electroporation buffer were performed in three consecutive centrifuge rounds of 15 minutes at 16000g.

Transposon mutants were screened for by sequencing with the corresponding primers (see Supp. Table S.12). Mutants obtained by allelic exchange were screened for by PCR and sequencing of the resulting amplicons.

E.1.4. Quantitative assessment of the recombination capacity

Mycoplasma genitalium cultures were grown and prepared for electroporation as described earlier (E.1.3.2). Mycoplasma cell suspensions were adjusted to ~10^9 cells/mL, which was monitored by diluting the resuspended cells to an OD600 ~2. Next, 100μL cell aliquots were mixed with 30μg of the suicide plasmid pΔMG_218-lacZ and electroporated. After the addition of 1mL of SP4, electroporated cells were incubated at 37°C for 4h, and 200μL aliquots seeded onto SP4 plates containing puromycin and X-Gal (see E.1.2.2). After 21 days of incubation at 37°C, plates were screened for the presence of puromycin resistant, blue colonies. A double recombination event between homologous sequences from the pΔMG_218-lacZ plasmid and the chromosome of *M. genitalium* promotes the replacement of the MG_218 gene by the puromycin resistance marker. The MG_218 gene, which codes for the cytadherence-associated protein HMW2, was shown to be dispensable for growth under laboratory culture conditions in previous studies (Dhandayuthapani *et al.*, 1999; Pich *et al.*, 2008). Deletion of the MG_218 gene was chosen because previous work from our laboratory showed that transformants were obtained at a high frequency (Pich *et al.*, 2008).The *lacZ* gene present in the pΔMG_218-lacZ plasmid confers β-galactosidase activity and facilitates the detection of the transformant colonies. The transformation efficiency (number of transformants per viable cell) was calculated as an indicator of the

recombination capacity of selected *M. genitalium* strains. Results presented in this work correspond to at least three independent biological repeats.

E.1.5. Mating experiments in *M. genitalium*

Mating experiments were carried out following this protocol.

- Two *M. genitalium* cultures were set in 75cm^2 flasks and 20mL of SP4 at 37ºC 5% CO_2 until mid-log phase. One of them (donor strain) was resistant to a certain antibiotic and overexpressed the MG428 regulatory pathway, while the other strain (recipient strain) carried a different selectable marker than the donor strain.
- 10 mL of SP4 medium were discarded and cells were scrapped off the flasks in the remaining medium in the flask.
- Cells were individualized through a 0.45µm filter.
- 4mL of filtrated cells of the donor strain (strain that overexpresses MG428) were mixed with 2mL of filtrated cells of the recipient strain and inoculated in 14mL of fresh SP4 in 75cm^2 flasks. Cells were left 24h at 37ºC 5%CO_2 without antibiotic selection.
- SP4 medium is aspirated and cells are scrapped off in 1mL of fresh SP4.
- 100-250µL of the scrapped cells are inoculated in a 75cm^2 flask with SP4 and dual antibiotic selection.
- Typically, 14-16 days were needed for the cells to grow all over the flask. Medium was changed at day 9-12 to discard dead cells. Dual antibiotic selection was kept during all the process.
- To avoid the detection of gDNA of the parental strains, 2-3 cell passages were performed.

E.2. DNA MANIPULATION

E.2.1. Plasmid DNA extraction

<u>Minipreparations</u>

Plasmid minipreparations were carried out using 2mL of an O/N *E. coli* culture using the *FastPlasmid*^{TM} *Minikit* (5prime) or the *GeneJET Miniprep Kit* (Fermentas), following manufacturer's instructions. Typically, up to 12µg of plasmid DNA could be obtained from each extraction. Minipreparations were used for cloning purposes.

<u>Midipreparations</u>

Midipreparations were performed using 50mL of an O/N *E. coli* culture using *GenElute HP Midiprep Kit* (Sigma-Aldrich), following manufacturer's instructions. Normally, up to 250µg of plasmid DNA could be obtained. DNA midipreparations were used to transform *M. genitalium*.

<u>Maxipreparations</u>

Plasmid Maxipreparations were performed with 250-400mL of an O/N *E. coli* culture using *GeneJET Maxiprep Kit* (Fermentas), following manufacturer's instructions. Typically, up to 500µg of plasmid DNA was extracted in each maxiprep. Maxipreparations were used to transform *M. genitalium* strains.

E.2.2. Genomic DNA extraction of *M. genitalium* strains

To prepare genomic DNA (gDNA) for Next Generation sequencing analysis, this protocol was followed:

- 20mL of fresh SP4 were inoculated with 50-70µL from a stock in a 75cm^2 flask and grown until mid-log phase.
- Cells were washed twice in 1xPBS (Sigma-Aldrich) and scrapped off in 0.5mL of solution I (0.1M Tris-HCl pH 8.0, 0.5M NaCl, 10mM EDTA). Non-adherent strains were washed and recovered by centrifugation at 18000*g* for 15 minutes.
- Cells were lysed with 12.5µL of SDS 20% (w/v) and treated with 50µg/mL proteinase K (Life Technologies) for 1h at 55ºC.
- Then, two consecutive phenol-chloroform extractions were performed. Phenol solution contained phenol-chloroform-isoamyl alcohol at a 25:24:1 ratio. Aqueous phase was transferred to a new tube

- DNA was precipitated adding 2 volumes of absolute ethanol and mixed gently. gDNA was recovered by centrifugation at 16000*g* for 10 minutes. Pellet was washed with 70% ethanol.

- Pellet was vacuum dried and resuspended in 0.2x TE buffer (Tris-HCl 10mM pH 8, EDTA 1mM) with 15µg/mL of RNase A (Life Technologies). Suspension was left O/N at 4ºC, vortexed and stored at -20ºC.

E.2.3. Obtention of *M. genitalium* lysates

Generally, to screen for *M. genitalium* mutants, high quality, intact gDNA was not necessary. Lysates of *M. genitalium* usually contained enough DNA to perform a screening by PCR or by DNA sequencing. To obtain these lysates, this protocol was followed:

- An *M. genitalium* culture was set in a small flask (25cm^2) with 5mL of SP4 broth supplemented with any necessary antibiotics and grown until mid-log phase.

- SP4 was aspirated and cells were scrapped off in 1mL of 1xPBS and transferred into a microtube. Tubes were spun at 16000*g* for 10 minutes. For non-adherent cultures, cells were centrifuged in SP4 and the pellet was washed once in 1xPBS before lysing.

- Pellets were lysed using 20-50µL of lysis buffer (0.1M Tris-HCl pH 8.5, 0.05% Tween-20, 250 µg/mL Proteinase K) and the mix was incubated at 37ºC for 1h.

- Proteinase K was inactivated at 95ºC for 10 minutes in a thermoblock.

- Lysates were stored at -20ºC.

Normally, 2-5µL of the lysate were sufficient for PCR amplification or DNA sequencing.

E.2.4. DNA amplification

DNA amplification was used to screen for mutants in *M. genitalium* as well as for cloning purposes. DNA was amplified by PCR in a *MJ Mini Thermal Cycler* (Bio-Rad). *Phusion*® *High Fidelity DNA polymerase* (Thermo Fisher) was used for molecular cloning applications, while Taq DNA polymerase (Sigma-Aldrich) was used in screenings. The usual PCR protocol was:

Reaction Mix

- 10-50ng of DNA template
- 0.5µM of each oligonucleotide
- 200µM dNTP mix (Sigma-Aldrich)
- 1x HF Green Buffer (Phusion) / 1x PCR reaction buffer (Taq)
- 0.6U Phusion polymerase / 1.25U Taq polymerase

- Sterile MilliQ H$_2$O to 50µL (Phusion) or 25µL (Taq)

<u>Thermal cycler protocol</u>

1. Initial denaturation: 2 minutes at 98ºC (Phusion) or at 94ºC (Taq)

2. Denaturation step: 20s at 98ºC (Phusion) or at 94ºC (Taq)

3. Primer annealing step: 20s at (Tm+3)ºC (Phusion) or 20s at (Tm-5)ºC

4. Elongation step: 72ºC, 15-30s/kb (Phusion); 72ºC, 1 minute/kb (Taq)

5. Repeat steps 2-4 for 34 times

6. Final elongation step: 5 minutes at 72ºC (both polymerases)

7. END

Most plasmid inserts were created by DNA fragments joined together by Splicing by Overlap Extension PCR (SOE-PCR). This technique is based in the amplification of a DNA fragment using oligonucleotides with a 5' overhang complementary to the end of the molecule to be fused with. Mixing both DNA fragments in the same PCR reaction, they will be annealed together and extended generating a new product formed by the unification of both fragments. This technique allows the ligation of several DNA pieces without the use of restriction enzymes nor DNA ligases.

E.2.5. Gel electrophoresis and DNA quantification

DNA fragments and PCR products were separated by agarose gel electrophoresis following standard procedures (Sambrook *et al.*, 1989). Low electroendosmosis agarose (Lonza) was used at 0.7-2% (w/v) diluted in 1xTAE buffer (40mM Tris, 20mM acetic acid and 1mM EDTA). DNA samples were diluted in loading buffer (40mM Tris-acetate pH 8, 1mM EDTA, Bromophenol blue 250µg/mL, xylene cyanol 250µg/mL and 30% glycerol) when necessary and loaded into the agarose gel. Electrophoresis was run at 80-90V for approximately 1h in 1x TAE buffer. Molecular weight was determined using *GeneRuler^{TM} 1kb Plus DNA ladder* (Thermo Scientific). Once the electrophoresis finished, gel was soaked in staining buffer (0.1M NaCl, 0.02% Midori Green Advanced (v/v)) for 20-30 minutes, and visualized under UV light or black light in a *GelDoc^{TM} XR+ Imaging System* (Bio-Rad). When necessary, gel bands were recovered from the gel using a clean scalpel and DNA was extracted using the *NucleoSpin Gel and PCR Clean-up* kit (Macherey-Nagel), following manufacturer's instructions.

DNA concentration was determined using a *NanoDrop 1000 Spectophotometer* (Thermo Scientific) following the standard procedures as described by the manufacturer.

E.2.6. DNA digestion

DNA fragments were cleaved with restriction enzymes (Fermentas) following the manufacturer's guidelines. Restriction of DNA was used for plasmid screenings and DNA cloning.

E.2.7. Ligation of DNA fragments and plasmid dephosphorylation

DNA fragments were ligated into digested plasmids with T4 DNA ligase (Sigma-Aldrich). Normally, ligation reactions were left at RT(ºC) for 1h or O/N at 16ºC, following manufacturer's instructions.

Blunt-end digested plasmids or cleaved plasmids bearing compatible ends were treated with the *rAPid alkaline phosphatase* Kit (Roche) before ligation to decrease the probability of re-circularization of the plasmid. This procedure was performed under the manufacturer's guidelines.

E.2.8. DNA sequencing

Sanger sequencing

Genomic DNA, plasmids and PCR products were Sanger sequenced at Servei de Genòmica i Bioinformàtica (UAB). Reactions were performed with the *BigDye® v3.1 Cycle Sequencing Kit* (Thermo Fisher) and analyzed in an ABI PRISM 3130xl Genetic Analyser.

Next-Generation Sequencing

Genomic DNA was prepared with the *Nextera XT DNA Sample Preparation Kit* (Illumina) and analyzed using a MiSeq Desktop Sequencer (Illumina) at Servei de Genòmica i Bioinformàtica (UAB). Illumina reads were 251 nucleotides long. Bowtie 2 (Langmead and Salzberg, 2012) and BWA-MEM (Li, 2012) were used to align sequences, which were then piled up with SAMtools (Li *et al.*, 2009). Finally, variants were called using VarScan (Koboldt *et al.*, 2009). Variants were only selected when the corresponding nucleotide had a phred quality score >20 (Ewing and Green, 1998; Ewing *et al.*, 1998a).

E.2.9. Plasmid construction

A total of 29 plasmids were constructed and used in this work (Table E.4). Strategies on the construction of the different vectors were similar and can be classified in three approaches:

construction of plasmids used to generate KOs in *M. genitalium*, construction of plasmids used to introduce fluorescent fusion proteins in *M. genitalium* and construction of minitransposons to re-introduce or overexpress genes in *M. genitalium*.

Table E.4.List of plasmids used in this work

	Plasmid name	Aim
	pΔMG_428	To create an MG_428 defective mutant
	pMTnCatMG_428	To complement MG_428 gene/To overexpress MG_428
	pET21aMG428	To produce MG428 recombinant protein in *E. coli*
	pΔMG_218-lacZ	To quantify transformation efficiency by homologous recombination in *M. genitalium*
	pMG_428:Ch	To introduce an mCherry fluorescent protein fusion to MG_428 gene
	pRecA:Ch	To introduce an mCherry fluorescent protein fusion to *recA* gene
	pRecA:Ch-10	To introduce an mCherry fluorescent protein fusion to *recA* gene and a mutation at the Pribnow box
Chapter I	pRecA:Ch-22	To introduce an mCherry fluorescent protein fusion to *recA* gene and a mutation at -22 bases upstream of the TSS
	pRecA:Ch-35	To introduce an mCherry fluorescent protein fusion to *recA* gene and a mutation at the -35 element
	pΔMG_010	To create an MG_010 defective mutant
	pΔMG_220	To create an MG_220 defective mutant
	pΔMG_RS02200	To create an MG_RS02200 defective mutant
	pΔMG_285	To create an MG_285 defective mutant
	pΔMG_390	To create an MG_390 defective mutant
	pΔMG_414	To create an MG_414 defective mutant
	pΔMG_525	To create an MG_525 defective mutant
	pΔ*recA*	To create *recA* defective mutant
	pMG_427:Ch	To introduce an mCherry fluorescent protein fusion to MG_427 gene
	pMTnPacMG_428	To complement MG_428/To overexpress MG_428
Chapter II	pRrlA:Ch	To introduce an mCherry fluorescent fusion to *rrlA* gene
	pRrlB:Ch	To introduce an mCherry fluorescent fusion to *rrlB* gene
	pMTnPacrrlA	To complement *rrlA* gene
	pMTnPacrrlB	To complement *rrlB* gene
	pMTnPacMG_428:eYFP	To overexpress MG_428 with an eYFP fusion protein
	pMTnTetrrlA:TAP	To overexpress *rrlA* with an TAP fusion tag
	pMTnTetrrlB:TAP	To overexpress *rrlB* with an TAP fusion tag
	pMTnPac	To create a minitransposon bearing the puromycin resistance gene
Appendices	pMTnPacSpx	To create a minitransposon to overexpress *spx* gene
	pΔMG_281	To create an MG_281 defective mutant

Plasmids to generate knock-out mutants in *M. genitalium*

KO mutants in *M. genitalium* were generated replacing the gene of interest with an antibiotic resistance marker by a double crossover event. This gene replacement is achieved via transformation with a suicide plasmid bearing the antibiotic cassette neighbored by 1kb of the upstream and downstream regions of the gene to be replaced.

Upstream and downstream regions as well as the antibiotic marker were amplified by PCR. Then, upstream region and antibiotic marker amplicons were mixed and used as a template to ligate and amplify them by SOE-PCR (see E.2.4). Finally, the amplicon containing the upstream region and the cassette was mixed with the downstream region amplicon and amplified by SOE-PCR to generate the final DNA insert (upstream region + marker + downstream region). Then, this insert was cloned in an *Eco*RV-digested pBE (Pich *et al.*, 2006b) to generate the suicide plasmid used to transform *M. genitalium*.

<u>Plasmids used to introduce fluorescent fusion proteins in *M. genitalium*</u>
Transcriptional fusions with fluorescent proteins in the chromosomal loci were generated using a similar strategy as described before. Fluorescent fusion coding gene and antibiotic marker were knocked-in at the 3' end of the gene of interest, deleting its stop codon to allow full-length translation of the fusion. Upstream and downstream regions, fluorescent protein and antibiotic marker were amplified individually by PCR and subsequently ligated using consecutive SOE-PCR rounds. The final DNA insert was cloned in an *Eco*RV-digested pBE to generate the suicide plasmid used to transform *M. genitalium*.

<u>Plasmids bearing minitransposons to re-introduce or overexpress genes in *M. genitalium*</u>
Plasmids carrying the miniTn4001-derived transposon were used to re-introduce or overexpress genes in the *M. genitalium* chromosome. Plasmids pMTnTetM438 (Pich *et al.*, 2006b), pMTnCat (Calisto *et al.*, 2012) or pMTnPac (this work) were digested with *Apa*I-*Xho*I or *Xba*I-*Not*I and ligated with similarly digested inserts. Transcription of genes cloned between *Xba*I-*Not*I sites was significantly increased due to the effect of the MG_438 promoter that allows antibiotic cassette expression (Figure E.1).

Figure E.1. Schematic representation of pMTnPac plasmid backbone. Restriction sites used for cloning are highlighted in red. Puromycin acetyl transferase (*pac*) gene is depicted as a green arrow.
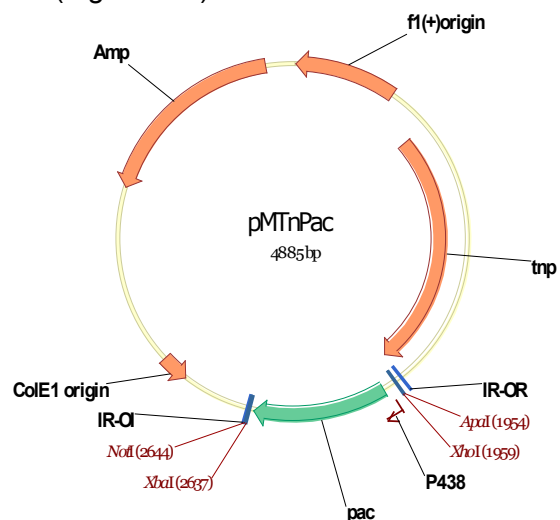
A detailed description of the construction of each plasmid listed in Table E.4 can be found in the Appendices section S.3.2.

E.2.10. Oligonucleotides used in this work

A complete list of all the oligonucleotides used in this work can be found in the Appendices section S.3.1.

E.3. RNA MANIPULATION AND TRANSCRIPTOMIC METHODS

E.3.1. RNA extraction and gDNA depletion

Prior working with RNA, lab surfaces and equipment were treated with *RNase ZAP® RNase decontamination solution* (Life Technologies) to avoid RNA degradation by RNases. To obtain total RNA of *M. genitalium* strains for qRT-PCR or primer extension analyses, the *RNAqueous Total RNA Isolation Kit* (Life Technologies) was used. Mid-log phase cultures were washed twice with 1x PBS and lysed with 500µL of lysis buffer provided in the kit. Then, instructions provided by the manufacturer were followed. Eluted RNA was treated with *Turbo DNA-free Kit* (Life Technologies) to eliminate any traces of gDNA. RNAs were stored at -80ºC until used.

To obtain total RNA of *M. genitalium* for RNAseq analysis the *miRNeasy Mini Kit* (QIAgen) was used. This kit allowed the purification of total mRNA, including miRNAs and the 5S rRNA that could not be purified using *RNAqueous Total RNA Isolation Kit*. Late phase cultures (25cm^2 flasks) were subcultured in two 25cm^2 flasks with fresh SP4 and incubated at 37ºC for 6 hours. Then, flasks were washed three times with pre-warmed (37ºC) 1x PBS and cells were lysed with 500µL of QIAzol Lysis Reagent. RNA extraction was then performed as described in the manufacturer's protocol. gDNA depletion was ensured using the *RNase*-free DNase Set (QIAgen) following manufacturer's instructions. RNA extractions were stored immediately at -80ºC.

RNA quantity was assessed using a NanoDrop 1000 Spectophotometer (Thermo Scientific) following the standard procedures as described by the manufacturer. RNA quality was determined using the *Agilent RNA 6000 Nano Kit* (Agilent Technologies) in an Agilent 2100 Bioanalyzer (Agilent Technologies).

E.3.2. Primer Extension analysis

Primer extension analyses were performed with 20µg of total RNA. Reverse transcription reaction was carried out using the SuperScript® III First-Strand Synthesis System (Life Technologies) following manufacturer's instructions. Briefly, the following mix was heated at 65ºC for 5 minutes in a thermal cycler:

- 20µg of total RNA
- 0.5µL of a 50µM 6-Fam-labelled primer
- 1.5µL of dNTPs

This mix was immediately cooled down in ice and the following components were added:

- 3µL of 10x RT Buffer
- 3µL of 25mM MgCl$_2$
- 2µL of 0.1M DTT
- 1.5µL RNaseOUT™ (40U/µL)
- 1.5µL SuperScript® III RT (200/µL)

Then, the retrotranscription reaction was performed at 50ºC for 50 minutes, followed by a step at 85ºC for 5 minutes to inactivate the enzyme. Finally, cDNA was treated with 2U RNase H for 20 minutes at 37ºC.

6-FAM-labelled cDNA was precipitated using 0.1 volumes of 3M sodium acetate and 2.5 volumes of absolute ethanol. Pellet was dissolved in 10µL of highly deionized formamide (Hi-Di Formamide) (Thermo Fisher) and mixed with 0.5µL of ROX400HD or ROX500 markers (Thermo Fisher). Fragments were separated in an ABI 3130xl Genetic Analyzer (Applied Biosystems) and analyzed using PeakScanner v1.0 software (Thermo Fisher). Fragment separation was executed at Servei de Genòmica i Bioinformàtica (UAB). At least two independent primer extension experiments were performed with each primer.

E.3.3. cDNA synthesis and qRT-PCR

cDNA for quantitative PCR analysis was synthesized using *iScript cDNA Synthesis Kit* (Bio-Rad) with 1µg total RNA, following manufacturer's guidelines. cDNA was then diluted 1:15 and stored at -20ºC. Negative controls without reverse transcriptase were also synthesized.

Quantitative Real-Time PCRs were carried out in 96-well or 384-well plates (Bio-Rad) using *iTaq™ Universal SYBR® Green Supermix* and CFX96 or CFX384 Real-Time PCR Detection Systems (Bio-Rad) from Laboratori de Luminiscència i Espectroscòpia de Biomolècules (UAB). PCR reactions were set as follows:

Per well, to a final volume of 10µL:

- 3µL diluted cDNA
- 1µL of each oligonucleotide
- 5µL of iTaq Universal SYBR Green Supermix

Thermal cycler setup was designed as a 3-step PCR:

1. Initial denaturation at 95ºC for 3 minutes
2. Denaturation step at 95ºC for 10s
3. Annealing step at 56ºC for 20s
4. Elongation step at 72ºC for 20s
5. Plate read
6. Repeat steps 2-5 for 39 more times
7. Denaturation step at 95ºC for 10s
8. Melting curve from 65ºC to 95ºC with temperature increments of 0.5ºC. Plate read every 5s.
9. END

The standard non-template controls (NTC) and non-reverse transcribed controls (NRT) were considered. For NRT controls, negative controls of cDNA synthesis were taken as template for the PCR reaction. NTC consisted on using DNase, RNase-free water as template for amplification. NRT allowed the detection of gDNA presence in cDNA samples, while NTC allowed the detection of possible contaminants in the primer set or reaction buffer. Furthermore, melting curves programmed at the end of the qRT-PCR served as another control to detect the amplification of non-specific amplicons.

All reactions were set as technical duplicates and three biological repeats were performed per experiment.

E.3.4. qRT-PCR data analysis

Cycle thresholds (Ct) were analyzed using CFX Manager software (Bio-Rad). Relative gene expression was calculated using the Pfaffl method (Pfaffl, 2001), which considers the amplification efficiencies (see E.3.4.2) of the target and the reference genes. Accurate normalization was achieved by geometric averaging of multiple reference genes (Vandesompele *et al.*, 2002). Differential gene expression was judged based on the common arbitrary 2-fold cutoff.

E.3.4.1. Primer design and selection of reference genes

Primers were designed using Primer3Plus software (Rozen and Skaletsky, 2000). Primers were usually 18-22nt long and 40-50% G+C, and annealed near the 5' end of the target gene.

Several reference genes were used in this study to normalize expression between samples (Table E.5). 16S rRNA is often selected as a housekeeping gene for normalization, but some studies have shown that rRNA can vary greatly between the exponential growth and stationary phases (Piir *et al.*, 2011) or during stringent response (Kaplan and Apirion, 1975). Moreover, ribosomal RNA represents 90-95% of the total RNA of a bacterial cell and their expression levels are far greater than most mRNAs. Therefore, ribosomal RNAs were not considered for reference gene selection.

Table E.5. List of reference genes used in qRT-PCR analysis

Locus tag	Gene annotation	Function
MG_177	*rpoA*	DNA-directed RNA polymerase, α subunit
MG_418	*rplM*	Ribosomal protein L13
MG_430	*gpmI*	2,3-biphosphoglycerate-independent phosphoglycerate mutase
MG_341	*rpoB*	DNA-directed RNA polymerase, β subunit
MG_238	*tig*	Trigger factor

E.3.4.2. Amplification efficiency calculation

Amplification efficiency for each oligonucleotide pair was calculated using a calibration curve. Genomic DNA was diluted serially (1:4) 6 times and the dilutions were used as a template for amplification. Considering a theoretical efficiency of 100%, each dilution should be amplified in Ct increments of 2 units. Plotting the logarithm of the template quantity (arbitrary units) vs. Ct, a linear regression could be drawn and the slope of the curve allowed the calculation of the amplification efficiency (Figure E.2). Amplification efficiency was calculated following the equation:

$$\mathrm{E}(\%) = \left(10^{-1/\mathrm{slope}} - 1\right) \cdot 100$$

Oligonucleotides with amplification efficiencies within 85-110% were used for qRT-PCR analysis. A table listing all oligonucleotides used in qRT-PCR and their calculated efficiency can be found at Supp. Table S.13.

MG_177   Amplicon size: 184 bp   Amplification efficiency: 92.4%   Observed Tm: 76.4ºC
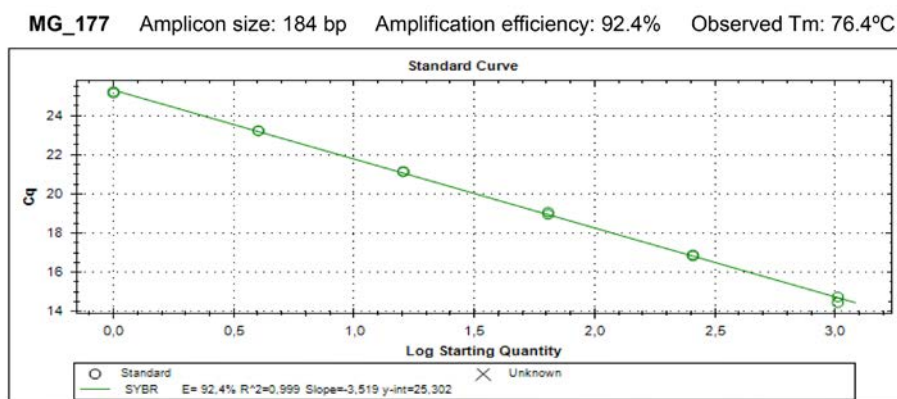
Figure E.2. Example of the amplification efficiency calibration curve of MG_177 primers. Linear regression had the equation y=-3.519x + 25.302 with an r^2=0.999. Applying the aforementioned formula, efficiency of MG_177 primers was E=92.4%. Tm, melting temperature.

E.3.5. RNAseq data analysis

To conduct the RNAseq study, three independent biological repeats of each strain were submitted to analysis. RNA libraries were prepared with *TruSeq Stranded Total RNA Library Prep Kit* (Illumina) and analyzed using a HiSeq 3000 System (Illumina) at the Genomics Unit from Center for Genomic Regulation (CRG), Barcelona. cDNA clusters were immobilized in sequencing lanes of 2x50 reads. Prior to any data analysis, reverse and complementary was computed for sequences coming from Read2 primer. Data analysis and sequence alignment was performed using *Map v1.7* software (Pinyol, unpublished) with Bowtie2 tool (Langmead and Salzberg, 2012) in the End-to-End mode and Forward-Forward paired-ends. Sequences were piled up using SAMtools (Li *et al.*, 2009) with no limit set to the number of sequences in the alignment. Counts in the different ORFs were performed with a standalone version of *featureCounts* program (Liao *et al.*, 2014) without counting the multi-mapping reads and disabling multi-overlaping reads.

Counted features were then submitted to the R/Bioconductor package DESeq2 (Anders and Huber, 2010; Love *et al.*, 2014) for statistical analysis. DESeq2 analysis used a parametric fitType and a zero-mean normal prior on the non-intercept coefficients. Data was sorted by log2 fold change and statistical significance was set at p-value<0.05. DESeq2 was chosen as the RNAseq normalization method over other widely used procedures, such as RPKM (Reads Per Kilobase per Million mapped reads) or TC (Total Count), since a recent study has shown that DESeq2 normalization can maintain a reasonable false-positive rate in different library sizes and widely different library compositions (Dillies *et al.*, 2013).

E.4. PROTEIN ANALYSIS METHODS

E.4.1. Protein extraction and quantification

To obtain *M. genitalium* protein extracts for analysis, mid -log phase cultures grown in 75cm^2 flasks were washed 4 times with 1x PBS (Sigma-Aldrich) and scrapped off in 125µL 1x PBS. Then, cells were lysed with 25µL 6x Laemmli sample buffer (0.375M Tris-HCl pH 6.8, 12% SDS, 60% glycerol, 0.6% β-mercaptoethanol and 0.003% bromophenol blue) and boiled for 10min to fully denature proteins. Protein extractions were cooled down and stored at -20ºC.

Protein extracts were quantified using *Pierce^{TM} BCA Protein Assay Kit* (Thermo Fisher) following manufacturer's instructions, or using NanoDrop 1000 Spectophotometer (Thermo Fisher).

E.4.2. Protein electrophoresis and staining

SDS-PAGE or denaturing protein electrophoresis
Proteins were separated in denaturing polyacrylamide gels following standard procedures (Shapiro *et al.*, 1967). SDS-PAGE gels were polymerized using the Mini-PROTEAN® handcast system (Bio-Rad) and 40% acrylamide:bisacrylamide (37.5:1) solution (Bio-Rad). Gels were run at 0.02-0.03A for 1h 30min in 1x Running buffer (25mM Tris, 192mM Glycine, 0.1% SDS). Molecular weight was determined using *PageRuler^{TM} Unstained Protein Ladder* (Thermo Fisher). For SDS-PAGE gels to be transferred to membranes, *PageRuler^{TM} Plus Prestained Protein Ladder* (Thermo Fisher) was used as a molecular weight ladder.

Native gel electrophoresis
Native gel electrophoresis was carried out in AnykD^{TM} Mini-PROTEAN® TGX^{TM} Precast gels (Bio-Rad) using 1x Running buffer without SDS. 10% glycerol was added to the samples and they were not boiled before running native electrophoresis. Gels were run in the standard conditions used in SDS-PAGE. No molecular weight markers were used in native gel electrophoresis. Gels were run at 80V for 90 minutes.

When the electrophoresis run was finished, gels were rinsed twice with distilled water to wash out SDS and soaked in Coomassie staining solution (0.1% Coomassie R250 (w/v), 10% acetic acid (v/v), 40% methanol (v/v)) for 20 minutes on a rocking table. Afterwards,

gels were distained in 10% acetic acid solution for several hours. Sponges were used to shorten the distaining process. Standard Coomassie staining allowed the detection of bands containing at least ~0.3µg of protein.

When higher sensitivity was required, gels were stained in colloidal Coomassie staining solution as described elsewhere (Dyballa and Metzger, 2009). This protocol allowed detection of at least ~5ng of protein/band.

E.4.3. Western blot method

For specific detection of *M. genitalium* proteins, SDS-PAGE gels were electrotransferred to PVDF membranes using cold Towbin buffer (25mM Tris, 192mM Glycine, 20% Methanol (v/v)) following the standard protocols described in Towbin *et al.*, 1979 and Mahmood and Yang, 2012.

Table E.6. Antibodies used in this study

Antibody	Host	Dilution	Conjugated enzyme	Source
Polyclonal anti-MG428	Mouse	1:100	-	This work
Polyclonal anti-mCherry	Rabbit	1:1000	-	BioVision, SKU: 5993
Monoclonal anti-MG438	Mouse	1:2000	-	Pich OQ, unpublished
Polyclonal anti-peroxidaseΨ		1:2000	HRP	Thermo Scientific
Polyclonal anti-Rabbit IgG (H+L)	Goat	1:5000	HRP	Thermo Scientific
Polyclonal anti-Mouse IgG (H+L)	Goat	1:5000	HRP	Bio-rad

* Kindly provided by Dr. Joel Baseman (UTHSCSA, USA)
Ψ Antibody used to screen for mutants in TAP analyses (see Appendices section S.2.4)

Membranes were probed with the appropriate primary antibody in blocking solution, which was then detected using a secondary antibody conjugated with horseradish peroxidase (Bio-Rad). Bioluminiscence reaction was catalyzed with *Luminata Forte*TM *Western HRP substrate* (Merck Millipore). Visualization and image optimization was performed in a VersaDoc Imaging System (Bio-Rad) using QuantityOne® software (Bio-Rad).

E.4.4. Recombinant expression and purification of MG428

Most recombinant MG428 protein expressed in *E. coli* in standard conditions was insoluble. Therefore, purification under denaturing conditions was carried out following *His-Bind Buffer Kit* protocols from Novagen® (Merck-Millipore).

The *E. coli* BL21 (DE3) strain carrying the pET21aMG428 plasmid was grown in 2L of LB medium at 37ºC to an OD_{600} of ~0.6 and MG428 expression was induced with 1mM IPTG. Protein expression was carried out at 37ºC for 3h. Then, cells were harvested by centrifugation at 5000x*g* for 5 minutes and pellet was resuspended in 40mL of binding buffer (20mM Tris-HCl pH 7.9, 0.5M NaCl and 60mM imidazole) with 0.1M PMSF protease inhibitor. After, cells were lysed by sonication and centrifuged again at 20000x*g* for 15 minutes to collect the inclusion bodies and cellular debris. Pellet was then resuspended in 20mL of binding buffer, sonicated and centrifuged again at 20000x*g* for 15 minutes. Finally, pellet was suspended in 20mL of binding buffer containing 6M urea and it was left on ice for 1h to completely dissolve the proteins. Insoluble material was discarded by centrifugation at 39000x*g* for 20 minutes, and supernatant was filtered through a 0.45µm low protein binding filter (Merck-Millipore). Purification was achieved using a HisTrap Fast Flow affinity purification column (GE Healthcare) and monitored in an Âkta Prime chromatography equipment (GE Healthcare). Eluate was analyzed by SDS-PAGE and the purest fractions were selected for dialysis in 1xPBS.

E.4.5. Production of MG428 antiserum

Polyclonal antibodies to the MG428 protein were obtained by repeated immunization of BALB/c mice with 0.5mg/mL purified recombinant MG428. Serum was tested against purified MG428 recombinant protein (data not shown).

E.4.6. Analysis of MG428-RNA polymerase interaction

200ng of soluble recombinant MG428 protein were mixed with 1, 2, 4 and 8µg of *E. coli* RNA polymerase (RNAP) core- or holoenzymes (Epicentre) in a buffer containing 25mM HEPES pH 8, 10mM $MgCl_2$, 100mM potassium glutamate and 2mM Dithiothreitol (DTT). The mix was incubated at 37ºC for 30min. Then, samples were analyzed in a native gel electrophoresis (see E.4.2).

E.4.7. Liquid chromatography-mass spectrometry (LC-MS) analysis

Samples were analyzed on a Maxis Impact Q-TOF spectrometer (Bruker), coupled to a nano-HPLC system (Proxeon). Samples were dissolved in 5% acetonitrile, 0.1% formic acid in water, concentrated on a 100μm ID-2cm Proxeon nanotrapping column and then loaded onto a 75μm ID-25 cm Acclaim PepMap nanoseparation column (Dionex). Chromatography was run using a 0.1% formic acid-acetonitrile gradient (5-35% in 45min; flow rate 300nL/min). The column was coupled to the mass spectrometer inlet through a Captive Spray (Bruker) ionization source. MS acquisition was set to 3 second cycles of MS (0.5Hz), followed by MS/MS (4-16Hz, intensity depending) of a variable number of the most intense precursor ions, with an intensity threshold for fragmentation of 2000 counts, and using a dynamic exclusion time of 2min, with an automated precursor re-selection when a 3-fold increase in intensity was observed. All spectra were acquired on the range 150-2200Da. LC-MSMS data was analyzed using the *Data Analysis 4.0* software (Bruker).

Proteins were identified using Mascot (Matrix Science) by search on a database constructed with *E. coli* K-12 substrain MG1655 and *M. genitalium* G37 sequences in the SwissProt Database (4738 sequences). MS/MS spectra were searched with a precursor mass tolerance of 10ppm, fragment tolerance of 0.05Da, trypsin specificity with a maximum of two missed cleavages, cysteine carbamidomethylation set as fixed modification and methionine oxidation as variable modification. Significance threshold for the identifications was set to give a False Discovery Rate <1% at the peptide level measured by searching a Decoy database. LC-MS analysis and identification was performed by the Proteomics Laboratory at Vall d'Hebron (VHIO), Barcelona.

E.4.8. Tandem Affinity Purification (TAP) protocol

This procedure was adapted from Rigaut *et al.*, 1999 and Kuhner *et al.*, 2009. Before starting this protocol, expression of the tagged protein was assessed by Western blot using anti-Peroxidase antibody (see Table E.6).

<u>Cell growth and lysis</u>

Ten 175cm^2 flasks were inoculated with the *M. genitalium* strain bearing the TAP tag and incubated at 37ºC 5%CO$_2$ until mid-log phase. Cells were scrapped off the flasks and washed three times with 1x PBS. Pelleted cells were resuspended in 10-20mL of NP-40 Buffer (6mM Na$_2$HPO$_4$, 4mM NaH$_2$PO$_4$·H$_2$O, 1% Nonidet P-40 (v/v), 300mM NaCl, 2mM EDTA, 50mM NaF, leupeptin 4μg/mL, benzamidine 16μg/mL and 0.1mM Na$_3$VO$_4$).

Protease inhibitor cocktail (cOmplete tablet, Roche) was added fresh right before use. Resuspended pellet was lysed in an homogenizer chamber using an ice-water jacked. Lysed cells were centrifuged to discard cell debris and supernatant was retrieved for purification.

Protein complex purification

Supernatant was mixed with 800µL of IgG/Sepharose bead (Pharmacia) 1:1 slurry in NP-40 Buffer and incubated on a rotating platform for 1h at 4ºC. Then, lysate was poured into a Poly-Prep Chromatography column (0.8x4cm) (Bio-Rad) and packed by gravity. Afterwards, beads were washed with 30mL IPP150 Buffer (10mM Tris-HCl pH 8.0, 300mM NaCl, 0.1% Nonidet P-40 (v/v)) and with 10mL TEV Cleavage Buffer (10mM Tris-HCl pH 8.0, 300mM NaCl, 0.1% Nonidet P-40 (v/v), 0.5mM EDTA and 0.5mM DTT; prepared fresh). Once washed, beads were treated with 300-500U of TEV protease in 1mL TEV cleavage buffer incubating the column on a rotating platform for 1.5h at 16ºC. Eluate was drained into a new bottom sealed column and 6mL of CB Buffer (10mM Tris-HCl pH 8.0, 300mM NaCl, 1mM Imidazole, 1mM magnesium acetate, 2mM $CaCl_2$ and 10mM β-mercaptoethanol) were added plus 6µL 1M $CaCl_2$ and 300µL of Calmodulin resin (Stratagene) 1:1 slurry in CB Buffer. Mix was incubated on a rotating platform for 1h at 4ºC. Then, resin was washed with CB Buffer with 0.1% NP-40 (v/v) twice and once again with CB Buffer with 0.02% NP-40 (v/v). Protein complexes were eluted using 1mL CE Buffer (10mM Tris-HCl pH 8.0, 300mM NaCl, 1mM Imidazole, 1mM magnesium acetate, 20mM EGTA, 0.02% Nonidet P-40 (v/v) and 10mM β-mercaptoethanol).

Protein precipitation

Eluate was adjusted to 25% trichloroacetic acid (TCA) with 100% TCA and tubes were placed on ice for 30min with brief periodic vortexing or O/N at -20ºC. Then, samples were spun at top speed at 4ºC for 10min and pellet was washed twice with ice-cold (-20ºC) acetone containing 0.05N HCl or ice-cold (-20ºC) pure acetone. Pellet was vacuum dried in a centrifugal evaporator.

Sample analysis

Typically, pellet can be resuspended in 1x Laemmli Sample Buffer and resolved in an SDS-PAGE with posterior silver staining or colloidal Coomassie staining. Otherwise, pellet can also be analyzed by MS.

E.5. EPIFLUORESCENCE MICROSCOPY

M. genitalium cells were grown in filtered (0.22µm) SP-4 medium on µ-Slide 8 well plates (Ibidi) for 16h at 37ºC 5%CO_2. Then, cells were washed once with 1x PBS with $CaCl_2$ and $MgCl_2$ (Sigma-Aldrich) and visualized on a Nikon Eclipse TE 2000-E microscope. All strains were grown and visualized under homogenous conditions. Phase contrast and epifluorescence images were captured with a Digital Sight DS-SMC Nikon camera controlled by NIS-Elements BR software. To detect eYFP fluorescence, samples were exposed with 500nm wavelength for 1s and images were captured at 542nm. mCherry fluorescence was detected exposing samples with 560nm wavelength for 2s and emitted light was captured at 630nm.

Images were analyzed using ImageJ software (NIH) and GDSC plug-in was used for cell counts.

APPENDICES

S.1. APPENDIX OF CHAPTER I

S.2. APPENDIX OF CHAPTER II

S.3. APPENDIX OF EXPERIMENTAL PROCEDURES

S.1. APPENDIX OF CHAPTER I

S.1.1. ΔMG_428 complemented mutants and TnCatMG_428 MiniTnp truncation

In this section, more information is shown on the different complemented ΔMG_428 mutant derivatives using Tn*CatMG_428* MiniTnp. A list with all the isolated mutants is presented in the Supp. Table S.1.

Supp. Table S.1. Insertion point of the Tn*CatMG_428* MiniTnp in the chromosome of the *M. genitalium* ΔMG_428 mutant derivatives.

Mutant designation	Disrupted gene	Insertion point[1]	Orientation
Tn::*ruvB*	MG_359	457688	as[2]
Tn::*recA*-1	MG_339	427770	as
Tn::*recA*-2	MG_339	427465	as
Tn::*recA*-3	MG_339	428074	as
Tn::*recA*-4	MG_339	427187	as
Tn::*recA*-5	MG_339	427419	as
Tn::*recA*-6	MG_339	427974	as
Tn::*recA*-7	MG_339	427773	as
Tn::*recA*-8	MG_339	427903	as
Tn::MG_220-1	MG_220	266090	as
Tn::MG_220-2	MG_220	266200	as
Tn::MG_281-1	MG_281	342383	as
Tn::MG_281-2	MG_281	342190	as
Tn::MG_281-3	MG_281	342091	as
Tn::MG_281-4	MG_281	342932	as
Tn::MG_281-5	MG_281	342444	as
Tn::MG_191-1[3]	MG_191	223515	as
Tn::MG_191-2	MG_191	225497	as
Tn::MG_191-3	MG_191	225180	as
Tn::MG_191-4	MG_191	223413	as
Tn::MG_192-1	MG_192	225962	as
Tn::MG_192-2	MG_192	226352	as
Tn::MG_192-3	MG_192	227185	as
Tn::MG_390-1[4]	MG_390	491702	as
Tn::MG_390-2[4]	MG_390	492205	as
Tn::MG_438/MG_024[4]	-	544221-28501	-

[1] Coordinates of the G37 genome corresponding to the transposon insertion point determined by sequencing with the CmUp and CmDown primers.
[2] as, transposon insertion in antisense orientation with respect to the disrupted gene.
[3] This transformant carries a copy of the transposase gene upstream of the Cm marker.
[4] These transformants carry a truncated copy of the MG_428 gene.
Mutants highlighted in bold were taken as reference in the Western blot analysis.

One of the complemented mutants, Tn::MG_438/MG_024, was found to bear a symmetric chromosomal inversion near the insertion point of the Tn*CatMG_428* MiniTnp. Large chromosomal inversions centered around the replication origin are rare but have been described in many bacteria (Eisen *et al.*, 2000). To our knowledge, this is the first report of

a viable chromosomal inversion in mycoplasmas probably generated during the transposition event. Symmetric chromosomal inversions centered at the oriC region are not deleterious due to the maintenance of the gene location with respect to the leading and lagging DNA strands (Mackiewicz *et al.*, 2001). Moreover, mutant Tn::MG_438/MG_024 also bears a truncated copy of the MG_428 gene, so its derived toxicity is highly reduced (Supp. Figure S.1).
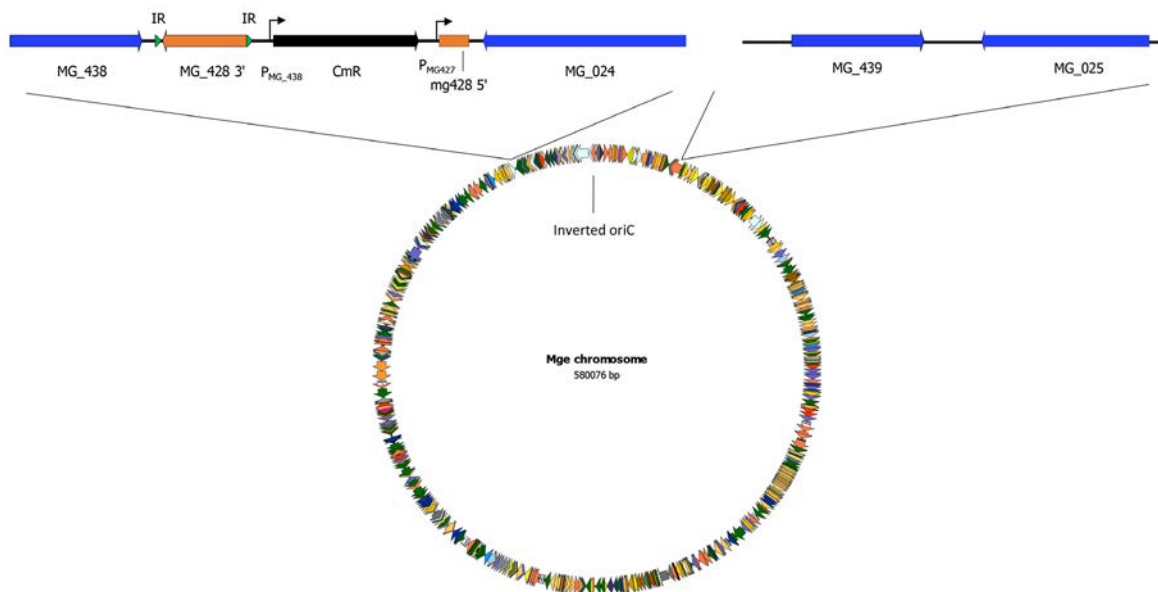
Supp. Figure S.1. Schematic representation of the symmetric chromosomal inversion present in the Tn::MG_438/MG_024 mutant. Genes surrounding the chromosomal inversion are highlighted as blue-filled arrows. MG_428 fragments of the ectopic copy introduced in the Tn*CatMG_428* MiniTnp are marked in orange and inverted repeats are drawn as green triangles. Whole genome analysis confirmed that the inversion was the only major modification in the genome of this mutant.

We found that MG_428 ectopic copy in this mutant (Tn::MG_438/MG_024) and in Tn::MG_390-1 and 2 mutants was truncated. The MG_428 gene is a major hotspot for Tn*4001*-derived transposons (Glass *et al.*, 2006) and this fact could provide a practical way to generate MG_428 truncated copies. In particular, insertion of a MiniTnp within the MG_428 copy of another MiniTnp would generate a sequence with four IRs, which will allow the mobilization of different DNA segments of the MG_428 gene (Supp. Figure S.2).
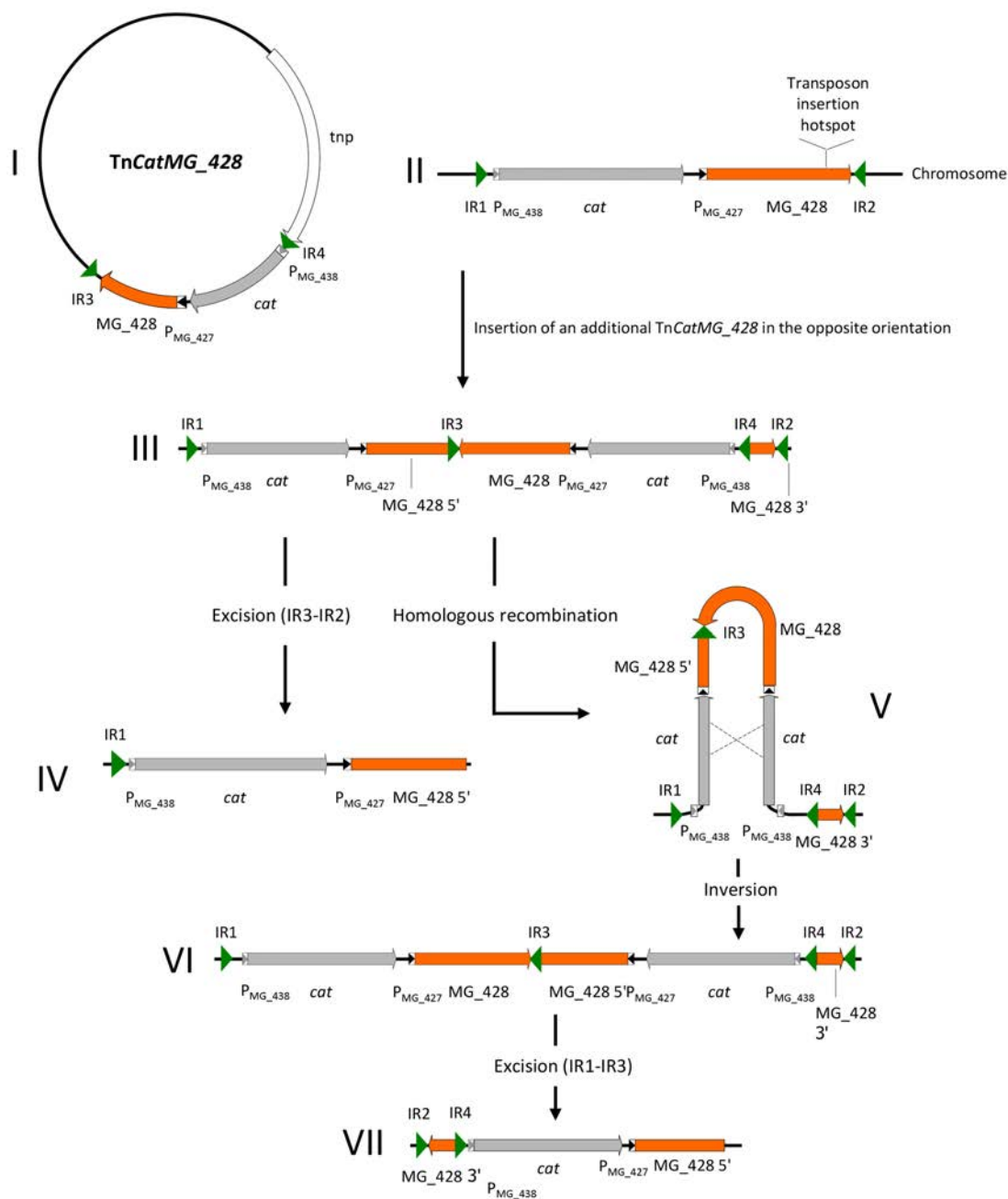
Supp. Figure S.2. Schematic depiction of two possible ways to generate MG_428 truncated copies derived from the TnCat*MG_428* MiniTnp. The MG_428 coding region is highlighted in orange and the IRs of the Tn*CatMG_428* MiniTnp are shown in green. Tnp, transposase.

S.1.2. Analysis of gene expression by RNAseq

Whole RNAseq data of the ΔMG_428 and the Tn::*recA*-2 mutants vs. the WT strain is detailed below in Supp. Table S.2 and Supp. Table S.3. Only statistically significant (p<0.05) transcription changes are listed.

Supp. Table S.2. Whole RNAseq data of the ΔMG_428 strain

Gene	Base mean	Log2 Fold Change	Standard Error	p-value
MG_428	133.6	-4.065	0.570	1.017E-12
ncRNA-3	73.9	-5.661	0.835	1.225E-11
ncRNA-1	88.9	-1.077	0.310	5.202E-04
MgPa-V, R5	64.6	-1.210	0.362	8.454E-04
MG_032	2533.2	-0.431	0.158	6.299E-03
MG_149	2633.3	-0.400	0.181	0.02709

Supp. Table S.3. Whole RNAseq data of the Tn::*recA* mutant

Gene	Base mean	Log2 Fold Change	Standard Error	p-value
MG_220	1963.165	5.602105	0.158967	4.8E-272
MG_339	1863.886	4.192576	0.125828	2E-243
ncRNA-1	3408.491	5.492319	0.18068	5.8E-203
MG_359	2292.914	4.079217	0.1543	5.2E-154
MG_358	891.5288	4.201784	0.16196	2.2E-148
MG_RS02200	1279.384	3.522364	0.146423	7.2E-128
MG_285	1120.22	3.855721	0.166988	5.9E-118
MgPa-V, R3	1894.587	3.899757	0.169706	7.5E-117
MgPa-V, R1	1105.688	3.784697	0.178866	2.3E-99
MG_414	1297.612	3.757622	0.180797	6.09E-96
MG_286	379.784	3.859755	0.211849	3.62E-74
ncRNA-2	850.7733	3.760953	0.231635	2.78E-59
MgPa-V, R5	446.9585	2.936213	0.20012	9.7E-49
MG_428	2952.72	3.770518	0.2897	1E-38
MgPa-V, R2	303.6127	3.553966	0.283216	4.05E-36
MG_412	983.9863	1.417722	0.161133	1.39E-18
MG_390	1122.805	1.194644	0.146069	2.87E-16
MgPa-V, R4	74.28684	2.302253	0.331623	3.86E-12
MG_389	33.84102	2.420295	0.364086	2.98E-11
ncRNA-3	6638.749	2.523997	0.424158	2.67E-09
MgPa-V, R6	25.91034	2.063547	0.36615	1.74E-08
MG_223	542.7662	0.827592	0.151907	5.09E-08
MgPa-IX, R1-R2	94.00763	1.557629	0.286036	5.16E-08
MG_084	853.1529	0.845246	0.157819	8.52E-08
MG_352	256.8799	0.817508	0.170742	1.68E-06
MG_221	461.7668	0.719275	0.158337	5.55E-06
MG_150	1872.933	0.455106	0.106899	2.07E-05
MG_195	981.6659	0.410104	0.097223	2.46E-05
MgPa-VIII, R3	423.3231	0.725491	0.173164	2.79E-05
MG_355	586.562	0.728823	0.178185	4.31E-05
MG_224	630.6819	0.605324	0.148712	4.69E-05
MG_096	11799.93	-0.44382	0.110067	5.52E-05
MG_085	730.0809	0.424675	0.112558	1.61E-04
MG_222	340.3917	0.642737	0.170994	1.71E-04
MG_388	261.6538	0.781461	0.224322	4.95E-04
MG_460	964.9949	-0.45624	0.131183	5.05E-04
MG_004	3665.239	-0.25223	0.074367	6.95E-04
MG_214	572.3651	-0.48165	0.14678	1.033E-03
MG_399	2644.851	-0.3508	0.107626	1.116E-03
MG_171	1224.802	-0.38714	0.120312	1.292E-03
MG_010	57.24838	1.276929	0.401141	1.456E-03
MG_454	261.0849	-0.57219	0.180222	1.499E-03
MG_288	229.6384	-0.7129	0.227429	1.721E-03
MgPa-II, R3	516.5142	-0.57733	0.184386	1.742E-03
MG_032	2679.44	-0.37904	0.121627	1.83E-03
MG_192	66346.8	-0.70753	0.227375	1.86E-03
MG_349	195.8713	0.539613	0.176076	2.179E-03
MgPa-II, R5-R6	2307.252	-0.73343	0.242789	2.521E-03
MG_194	271.9823	0.447879	0.151581	3.13E-03

MG_497	59.38309	0.769006	0.261372	3.259E-03
MG_218	10165.98	-0.30676	0.104783	3.416E-03
MG_190	1166.604	-0.30696	0.106212	3.852E-03
MG_189	1039.437	-0.4306	0.151368	4.445E-03
MG_149	2674.771	-0.47464	0.167287	4.55E-03
MG_384	249.5281	-0.53163	0.189415	5.006E-03
MG_072	4033.26	-0.33397	0.120297	5.499E-03
MG_206	237.7667	0.700885	0.255106	6.007E-03
MG_142	2403.192	-0.28028	0.105339	7.796E-03
MG_241	305.9248	0.383417	0.144834	8.114E-03
MG_479	1566.347	0.919436	0.349805	8.578E-03
MG_151	2885.984	0.265094	0.103132	0.010157
MG_213	878.4989	-0.47283	0.184896	0.01055
MG_427	7630.164	-0.47868	0.197404	0.015312
MG_083	588.4194	0.700727	0.288992	0.01532
MgPa-VIII, R5-R6	1911.212	-0.61647	0.256151	0.016098
MG_141	1964.586	-0.36528	0.153689	0.017464
MG_071	5165.396	-0.35432	0.154658	0.021964
MG_191	98525.89	-0.52772	0.231789	0.022804
MG_340	12535.91	-0.20578	0.090587	0.02311
MG_504	351.453	0.870787	0.383592	0.023202
MG_284	39.53608	0.754464	0.333221	0.023565
MG_276	622.8263	-0.26106	0.116795	0.025402
MG_367	352.5186	-0.39881	0.178968	0.025854
MG_461	344.4994	-0.31385	0.146593	0.032277
MG_287	65.36049	0.818662	0.383876	0.032956
MG_351	582.0411	0.436086	0.20576	0.034057
MG_051	772.1972	0.318748	0.151895	0.035862
MG_212	1025.458	-0.32858	0.15773	0.037234
MG_341	15303.29	-0.23526	0.113251	0.03777
MG_307	2443.415	-0.25661	0.12626	0.042116
MG_035	321.5621	0.454335	0.225249	0.043692
MG_217	3801.274	-0.23422	0.116406	0.044215
MG_278	338.6737	0.339794	0.168913	0.044257
MG_037	579.974	-0.28533	0.143448	0.046694
MG_129	92.97199	0.47322	0.238207	0.046967
MG_350	129.5753	0.527997	0.269283	0.049908

S.1.3. Analysis of gene expression by qRT-PCR

Results obtained by RNAseq of the different transcriptional changes due to the overexpression of MG428 were confirmed by qRT-PCR analysis (Supp. Figure S.3).
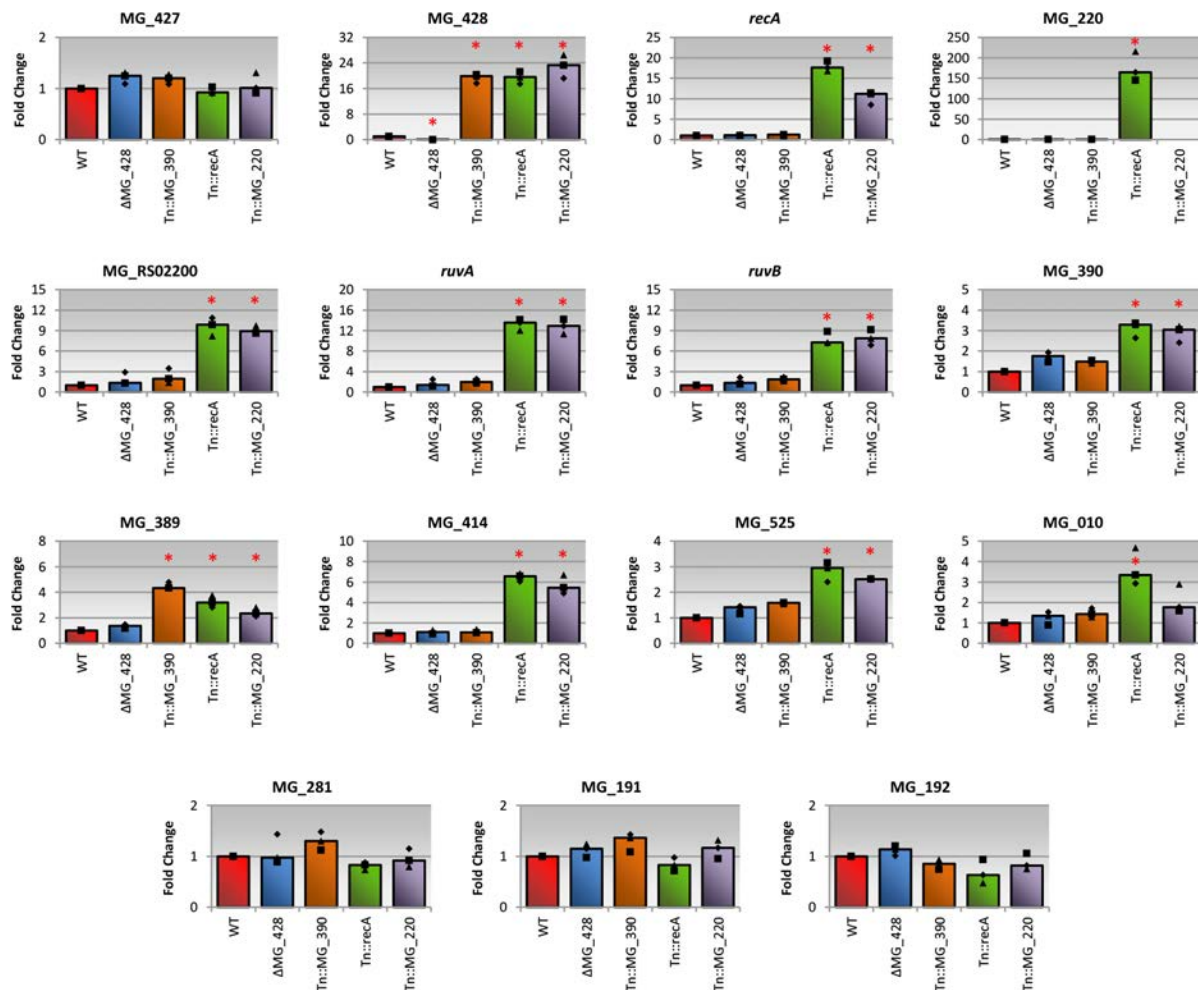
Supp. Figure S.3. Analysis of gene expression by qRT-PCR. Transcriptional analysis of selected *M. genitalium* genes in the WT and several mutant strains. Three independent biological repeats were performed and the respective fold-changes in gene expression are indicated with diamonds, squares and triangles. Mean fold-changes for each target gene are represented by color bars. Statistical significance of mean fold-changes above the cutoff (>2) was assessed with Student's *t* test. Statistically significant values (*P* < 0.05) are indicated with a red asterisk. Transcription of MG_220 (309 bp) in the Tn::MG_220 mutant could not be assessed by qRT-PCR due to the presence of a Tn*CatMG_428* MiniTnp insertion in this gene.

The mRNA levels of the MG428-regulated genes in the ΔMG_428 and the Tn::MG_390-1 mutants were similar to those of the WT strain. We observed that MG_389 transcription was elevated in the Tn::MG_390-1 mutant, which carried the Tn*CatMG_428* MiniTnp inserted within the MG_390 locus. The increased MG_389 transcript levels in this mutant are likely the result of the activity of cryptic promoters located within the MiniTnp. In contrast, transcription of the MG_281, MG_191 and MG_192 genes was not found to be altered in the ΔMG_428 or the complemented mutants, indicating that expression of these genes is not under the control of the MG428 protein.

S.1.4. Identification of the TSS of the MG_RS02200 gene

The MG_RS02200 gene is situated immediately upstream of the *ruvAB* operon of *M. genitalium*. In a primer extension analysis, we found that there was no TSS in the upstream region of the *ruvA* gene using a specific primer, but we could detect the TSS of MG_RS02200 out of the linear range (Figure I.14). For this reason, we confirmed the initiation of transcription using a new primer from within the MG_RS02200 gene (Supp. Figure S.4).

Supp. Figure S.4. Identification of the MG_RS02200 TSS by primer extension. Analysis was performed in the WT strain and the Tn::*recA*-2 mutant using the PEmg_rs02200 primer. TSS was only detected in the Tn::*recA*-2 mutant. Electropherograms were generated with Peak Scanner v1.0 (Applied Biosystems) analysis software. Red peaks represent ROX size standards while blue peaks correspond to the primer extension products. A schematic representation of the genome region is also shown; the identified promoter is indicated with a blue arrow and the approximate location of the PEmg_rs02200 primer is depicted with a black arrow. Additionally, the sequence neighboring the identified TSS is shown and the predicted -10 and -35 promoter elements are underlined. Location of the identified TSS is indicated with a +1 and the translational start point of the MG_RS02200 gene is highlighted in bold letters. S, peak size (in bp); H, peak high (in arbitrary units).

S.2. APPENDIX OF CHAPTER II

S.2.1. RrlA resembles bacterial hormones with a double-glycine motif

Many Gram-positive bacteria can produce peptidic hormones to interact with other individuals of the population or to respond to environment variations. Among them, the most studied are autoinducers like the competence-inducing peptides, which stimulate competence in selected bacterial species (Dunny and Leonard, 1997); or bacteriocins, peptides with anti-microbial activity that are used to kill or inhibit growth of bacterial strains living in similar biological niches (Gratia, 1925).

Many bacteriocins of Gram-positive bacteria are small peptides (<10kDa) that bear a leader sequence with a well-conserved double-glycine motif (GG). The GG leader sequence is cleaved by a specific ABC transporter with a C39 peptidase domain, and then, active bacteriocin is secreted to the extracellular medium (Havarstein *et al.*, 1995). Beside this leader sequence, bacteriocins often lack any other common sequence similarities. *rrlA* of *M. genitalium* codes for a small protein of ~10kDa (93 amino acids) with a putative double-glycine-type leader sequence that is conserved in other Mycoplasmas (Supp. Figure S.5 and Table D.1). The putative leader peptide is comprised by 15 residues, which are similar to those described as the consensus sequence for analogous proteins.

Supp. Figure S.5. Analysis of the N-terminal end sequence of RrlA. (A) Multiple sequence alignment of the first residues of the RrlA polypeptide (Mge), its homologues in other Mycoplasmas: *M. pneumoniae* (Mpn), *M. testudinis* (Mte), *M. pirum* (Mpi) and *M. iowae* (Mio); and known bacteriocins of other bacteria: *Streptococcus pneumoniae* (Spn), *Bacillus sp.* (Bac), *Enterococcus faecalis* (Efa), *Streptomyces coelicolor* (Sco) and *Clostridium acetobutylicum* (Cac) (Dirix *et al.*, 2004). **(B)** Sequence logo generated using the aligned sequences. Below, comparison between the residues immediately upstream of the double-glycine motif (GG) of the RrlA sequence and the GG-type leader sequence consensus (Havarstein *et al.*, 1994). Identical residues are highlighted in green, while conservative mutations are marked in yellow.

Most mycoplasma species encode at least one protein with a C39 protease domain in their minimal chromosomes (see Table D.1). In the case of *M. genitalium*, σ^{20}-regulated MG_390 gene (*sunT*) codes for a putative ABC transporter with a C39 peptidase domain at the N-terminal end. Despite there is no experimental data linking RrlA and SunT, it is tempting to speculate that both proteins might be related as they are under the control of the σ^{20} factor.

Work is in progress to study this possible link between these two proteins.

S.2.2. σ^{20} pathway is not activated by external factors, but can be disabled by oxidative stress

We have submitted RecA:Ch strain to several stress conditions to determine whether the σ^{20} pathway could be activated by an external factor. Apparently, DNA damaging agents such as mitomycin C or ciprofloxacin did not alter substantially the number of activated cells, discarding DNA insults as the possible trigger factor (Supp. Figure S.6 and Supp. Table S.4). Comparable results were obtained when cells were subjected to UV radiation. This fact suggests that σ^{20} pathway is not a form of SOS response of *M. genitalium*, regulated in other bacteria by the transcriptional repressor LexA. However, activation of recombination enzymes in a small percentage of the cell population might be useful to face and overcome DNA damage and, eventually, increase survival rates. Interestingly, the number of activated cells during ciprofloxacin treatment was slightly above the observed in untreated cells (Supp. Table S.4). Ciprofloxacin is a second-generation fluoroquinolone that inhibits DNA gyrase and topoisomerase activities, causing cell division arrest and, eventually, cell death. This mechanism of action can explain the high percentage of cells in division after 16h of treatment (Supp. Figure S.6). In this case, most activated cells are forming cell pairs united in membrane filaments, suggesting that activation of one cell can trigger activation of its partner via the membrane union. However, despite all cells are under ciprofloxacin treatment, σ^{20} pathway activation is still only present in a small subset of the population, suggesting that ciprofloxacin is not the trigger, but a possible enhancer.
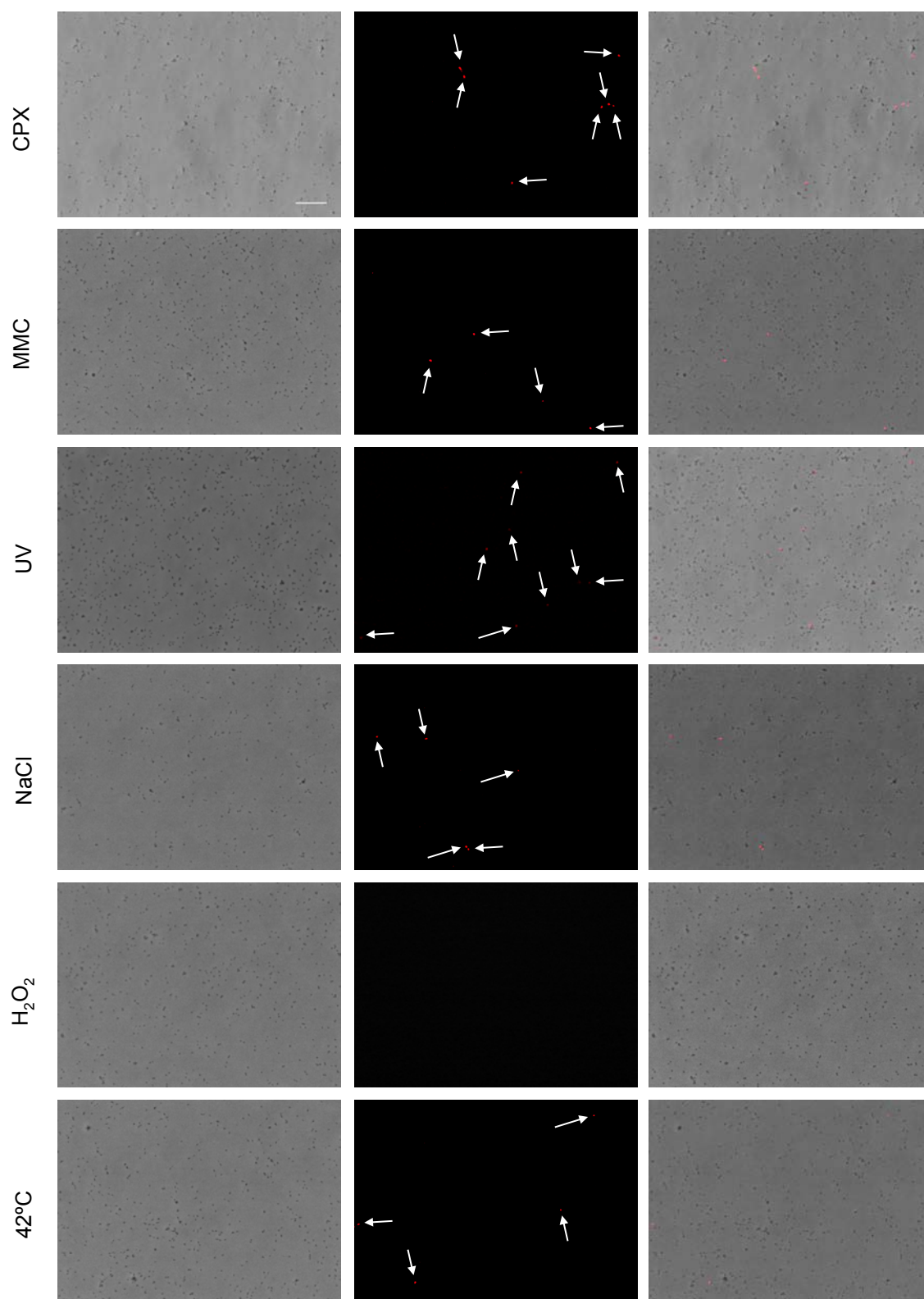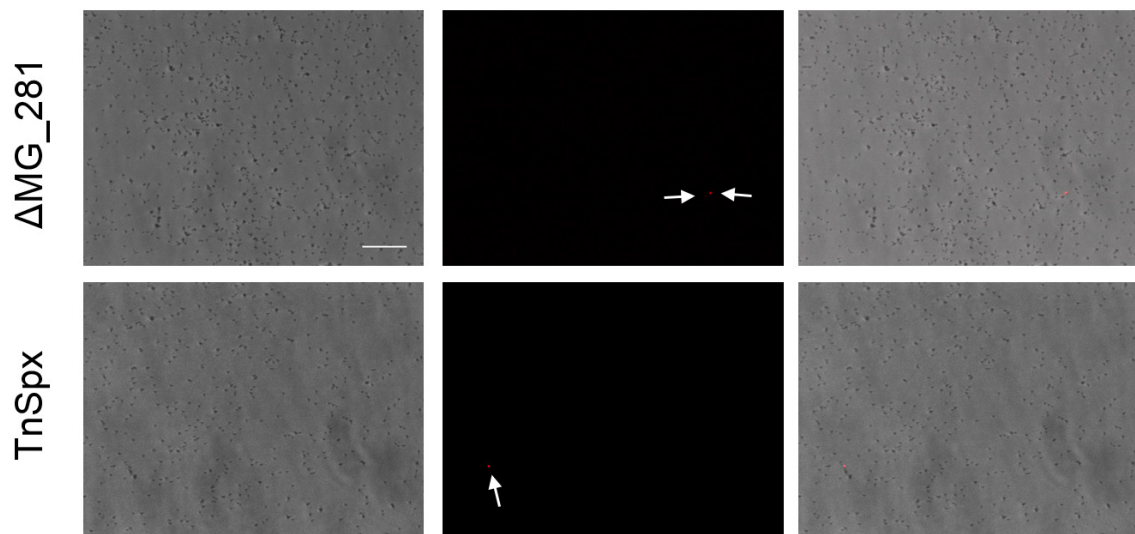
Supp. Figure S.6. Single cell analysis of RecA-mCherry expression by fluorescence microscopy upon inducing several stresses. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. White arrows point to mCherry fluorescent cells. All the pictures are shown at the same magnification, scale bar is 10µm.

At the same time, physical agents such as heat stress, continued osmotic shock or oxidative stress did not trigger the σ^{20} pathway (Supp. Figure S.6 and Supp. Table S.4). Nonetheless, treatment with 2% H_2O_2 during 1h could completely disable σ^{20} activation, suggesting that cells cannot activate this pathway during oxidative stress. This result is in agreement with a previous study regarding MG_427 induction upon different stresses (Zhang and Baseman, 2014). In that work, it is described that osmotically inducible protein C (MG_427) expression is greatly downregulated at high concentrations of H_2O_2. We have demonstrated that transcription of MG_428 is driven by MG_427 promoter; as a consequence, any disturbance regarding this promoter would have a consequence in both MG_427 and MG_428 transcription. However, the implications of the inactivation of the σ^{20} pathway in correlation with oxidative stress are intriguing and they should be addressed in future experiments.

Despite that none of these external factors had any triggering effect on σ^{20} activation, the nature of MG428 as an alternative sigma factor of the ECF subfamily prompts us to speculate that it might be subjected to the control of external agents present within the host. Unfortunately, we could not identify such factors or we failed to imitate the tested conditions as they take place *in vivo*.

S.2.3. σ^{20} pathway is not controlled by protein M nor RNAP regulator SpxA

Transformation of a MiniTnp bearing MG_428 gene under the control of MG_427 promoter leads to a toxic phenotype unless the transposon is inserted in the transcription antisense direction in specific loci along the chromosome. Some of the insertion points corresponded to σ^{20}-regulated genes, but the presence of an insertion hotspot at the MG_281 gene was intriguing. MG_281 is not under the control of σ^{20} (see Supp. Figure S.3), but it could play a role in σ^{20} regulation. For this reason we constructed a ΔMG_281 mutant in a σ^{20}:Ch and RecA:Ch genetic background. Single cell analyses revealed that cells lacking protein M have alterations in σ^{20} activation pattern (Supp. Figure S.7 and Supp. Table S.4). Precisely, activation of the σ^{20} pathway was reduced ~2.5-fold. Taken together, these results and Western blot analyses (Figure I.6) suggest that insertions within the MG_281 gene were selected because they led to lesser amounts of σ^{20} due to antisense translation blockage, but also because σ^{20} pathway activation is reduced in this background.

Supp. Figure S.7. Single cell analysis of RecA-mCherry expression by fluorescence microscopy in two different mutant backgrounds, a null mutant of protein M (ΔMG_281) and a mutant overexpressing Spx (TnSpx). Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. White arrows point to mCherry fluorescent cells. All the pictures are shown at the same magnification, scale bar is 10μm.

SpxA protein is a RNAP alpha subunit-binding protein that can modulate transcription in *B. subtilis* acting in a similar manner of the anti-sigma factors (Nakano *et al.*, 2003; Newberry *et al.*, 2005). Moreover, SpxA can sense oxidative stress and redirect transcription towards genes coding for protective enzymes. The presence of this protein coded in the *M. genitalium* chromosome (MG_127) and the fact that σ^{20} activation was inhibited upon oxidative stress prompted us to speculate that it could modulate σ^{20} binding to RNAP and control the expression of σ^{20}-regulated genes. Several attempts on constructing a ΔMG_127 mutant were unsuccessful (data not shown) suggesting that it is essential for growth in *M. genitalium*. Therefore, to study the impact of Spx on the σ^{20} pathway we decided to upregulate MG_127 in a RecA:Ch mutant using a MiniTnp. Activation of the σ^{20} pathway was reduced in all mutants analyzed but not completely abrogated (Supp. Figure S.7 and Supp. Table S.4), suggesting that Spx protein might modulate σ^{20} activation. Spx protein has two different conformations depending on the oxidation state of the cell. Oxidized form of Spx can bind the RNA polymerase complex, while Spx in its reduced state is inactive (Zuber, 2013). Therefore, we cannot rule out that this switch between active and inactive states of Spx can affect the results presented in this section.

In this sense, we are currently evaluating the role of RpoE (MG_022), another RNAP-binding protein that can modulate transcription in other bacteria (Weiss and Shaw, 2015).

As in the case of *spx*, mutants of MG_022 are not viable suggesting a crucial role of *rpoE* in *M. genitalium* homeostasis. Experiments overexpressing this component of the RNAP complex will allow us to determine if it plays a significant role in the σ^{20} pathway.

Supp. Table S.4. Summary of the different conditions tested on a RecA:Ch mutant strain

Condition	Strain	% of fluorescent cells
Untreated	RecA:Ch	0.66%
Ciprofloxacin (2µg/mL, 16h)	RecA:Ch	0.89%
Mitomycin C (40ng/mL, 16h)	RecA:Ch	0.55%
UV light (40J/m^2)	RecA:Ch	0.62%
Heat shock (42ºC, 15min)	RecA:Ch	0.70%
Osmotic shock (0.2M NaCl, 1h)	RecA:Ch	0.52%
Oxidative stress (2% H_2O_2, 1h)	RecA:Ch	0%
MG_281 null mutant (permanent)	RecA:Ch	0.24%
Spx overexpression (permanent)*	RecA:Ch	0.42%

*Average of four clones analized

S.2.4. Interactomic analyses of σ^{20}, RrlA, RrlB and RecA

To determine the molecular mechanisms driving σ^{20} activation and stabilization, we decided to conduct a Tandem Affinity Purification (TAP) using σ^{20}, RrlA, RrlB and RecA as baits. These interactomic analyses were performed successfully in *M. pneumoniae* in a global study (Kuhner *et al.*, 2009), but no information related to the orthologues of these proteins could be obtained.

To conduct a TAP analysis, overexpression of a TAP-tagged bait is indispensable. We transformed a WT strain with MiniTnp carrying σ^{20}-, RrlA-, RrlB- and RecA-TAP fusions under the control of the MG_427 promoter, with different outcomes. As expected, overexpression of σ^{20} resulted in few mutant colony recovery, with insertions at the same spots as defined in Figure I.5 and Supp. Table S.1. We selected a mutant carrying the transposon inserted at the MG_281 gene for the TAP experiment, but we could not isolate any protein counterpart nor the σ^{20} itself (data not shown), suggesting that there was not enough expressed bait for the experiment.

Overexpression of RecA did not render any mutant colonies, probably due to the high hyperrecombinogenic state induced in these cells. We could only isolate mutants bearing truncated copies of RecA-TAP that were useless for further analyses (data not shown).

In the case of RrlA and RrlB, we could isolate several colonies carrying RrlA- and RrlB-TAP fusions. Nevertheless, most isolated colonies presented mutations in the adhesins and showed no-adherent phenotypes. We decided to use adherent mutant pools for the analysis and we proceeded to perform a TAP assay. However, we could only identify proteins unrelated to the transcription machinery or the σ^{20} pathway in the case of RrlA (Supp. Table S.5), and no significant results were obtained in the case of RrlB (data not shown).

Supp. Table S.5. Identified proteins that interact with RrlA in a TAP assay.

Protein	Function	MW [kDa]	pI	Score	#Peptides	SC [%]	RMS90 (ppm)	#Spec. counts
MG312	Gliding motility	130.5	3.9	848.8	25	33.2	1.32	32
MG377	HP	22.6	6.3	218	8	46.6	1.64	13
MG183	Oligoendopeptidase F homolog	70.8	9.8	357.5	9	23.6	1.76	9
MG269	HP	39.6	9.2	258.9	8	25.9	0.78	8
MG318	Gliding motility	32.1	11.2	168.9	4	29.6	0.97	7
MG220	σ^{20} stability	11.3	9.7	81	2	9.8	1.94	3

Overall, these results indicate that TAP assays are not a suitable technique to identify interactant proteins of σ^{20} or σ^{20}-regulated proteins. The necessary overexpression of tagged proteins as baits and the intermittent nature of the σ^{20} pathway might be a considerable impediment for the realization of these studies.

Molecular cloning, colony isolation and TAP assays were conducted with the assistance of Ignasi Granero Moya, a former undergraduate student in the lab.

S.2.5. Transformation efficiency using miniTn*CmR*MG_428

Overexpression of MG_428 in the ΔMG_428 strain using the miniTn*CmR*MG_428 plasmid resulted in high toxicity, probably due to the induction of a hyperrecombinogenic state. Transformation efficiency in other genetic backgrounds is similar to those observed in the ΔMG_428 strain, with the notable exception of the Δ*recA* strain, in which the efficiency is 100 times higher than the WT strain (Supp. Table S.6).

Supp. Table S.6. Transformation efficiency of different strains using miniTn*CmR*MG_428 plasmid

Strain	Transformation efficiency using an empty miniTn*CmR*	Transformation efficiency using miniTn*CmR*MG_428
G37	3.4×10^{-3}	5×10^{-8}
ΔMG_428	2×10^{-3}	10^{-7}
Δ*rrlA*	1.1×10^{-3}	3×10^{-7}
Δ*rrlB*	1.3×10^{-3}	4×10^{-7}
Δ*recA*	1.1×10^{-3}	1×10^{-5}

S.3. APPENDIX OF EXPERIMENTAL PROCEDURES

S.3.1. List of oligonucleotides used in this work

Supp. Table S.7. Oligonucleotides used for mutant construction

Construction of mutants of *M. genitalium*

Name	Sequence 5'→ 3'	Use
Tc-F	CTCGAGTAGTATTTAGAATTAATAAAG	Amplify TetM438 marker.
Tc-R	GGATCCCTAAGTTATTTTATTGAAC	Used to build all pΔ plasmids
mg428 (Up-F)	GGTGGTATTTCACTTGATTG	Build pΔMG_428
mg428 (Up-R)	TAATTCTAAATACTAGAATTCGTCCCAACTTTACATCACTA	
mg428 (Down-F)	CAATAAAATAACTTAGGGATCCGCTACATCAAATTGCTAACAG	
mg428 (Down-R)	GGTATCAATGGTAGAAAAGG	
mg220 (Up-F)	GTGATCCTGATCCAATCCAA	Build pΔMG_220
mg220 (Up-R)	GTTCAATAAAATAACTTAGGGATCCGTTAAAACTACTGTTGTTAACAC	
mg220 (Down-F)	TTTATTAATTCTAAATACTACTCGAGCTCCTTTTGTCTATTAATCACATC	
mg220 (Down-R)	GTTGTTTAAGATGAACTGGG	
mpn534h (Up-F)	ATCGCACAAACACTTGCACC	Build pΔMG_RS02200 (pΔrrlB)
mpn534h (Up-R)	TAATTCTAAATACTACTCGAGGTGTTTCTCTGATAGTTTCC	
mpn534h (Down-F)	CAATAAAATAACTTAGGGATCCTGAATGCAAGGGATCAAATC	
mpn534h (Down-R)	CTTGGAATTAAAATCCTTGCC	
mg010 Up-F	TCTGATTGAAAATTGACAGC	Build pΔMG_010
mg010 Up-R	TTTATTAATTCTAAATACTACTCGAGCTTTCAGTTCTAACATCCAT	
mg010 Down-F	GTTCAATAAAATAACTTAGGGATCCAACTCAAATTGATCAGCG	
mg010 Down-R	ATCAAGACAGATCATTGATC	
mg414 Up-F	GATTAGAAGGCAAAACTTCC	Build pΔMG_414
mg414 Up-R	GTTCAATAAAATAACTTAGGGATCCTGATAAGTTCTGCTGTTTTG	
mg414 Down-F	TTTATTAATTCTAAATACTACTCGAGAAGCATTAGCTAAAGCAACG	
mg414 Down-R	GATTTAACTAAATGGGGTTC	
mg525 Up-F	TGTCATTGCTAAACCTCATC	Build pΔMG_525
mg525 Up-R	GTTCAATAAAATAACTTAGGGATCCCACTGTTTTTATTTACTTTC	
mg525 Down-F	TTTATTAATTCTAAATACTACTCGAGGAAACTCACTTTAAATAGCG	
mg525 Down-R	AAAGAAACGCTCTTTATCAC	
mg390 Up-F	TCAGTAAAACCAGGGACATC	Build pΔMG_390
mg390 Up-R	GTTCAATAAAATAACTTAGGGATCCATGAAGCAGAAAATTTTAGGAA	
mg390 Down-F	TTTATTAATTCTAAATACTACTCGAGCACTCATTCTGTTGTTCTTG	
mg390 Down-R	AGACACTCCTTCAGATCTAC	
recA (Up-F)	CCACTTTCACTAGGATCTTG	Build pΔrecA
recA (Up-R)	GTTCAATAAAATAACTTAGGATCCTAACTGTTCAACAATTGGC	
recA (Down-F)	TTTATTAATTCTAAATACTACTCGAGGAGCCATCTTTCATTACACC	
recA (Down-R)	CAGTCTAAATTGGTGTTAGC	

mg192.1 Up-F	GGTTGTTTGAAATTGATACC	
mg192.1 Up-R	TTTATTAATTCTAAATACTACTCGAGCATATCCACTCTCTAAATTG	Build pΔMG_192.1ter
mg192.1 Down-F	TCGGATTCCTCGAGGGATCCGACAATATAAACCATTGAAAAG	
mg192.1 Down-R	CTCCTAATTTATCTGTTTCG	
mg281 Up-F	GGAAGATATTTTGGGAGATC	
mg281 Up-R	GTTCAATAAAATAACTTAGGGATCCAACTTAAAACCAACGCATTG	Build pΔMG_281
mg281 Down-F	TTTATTAATTCTAAATACTACTCGAGTTTACGCTTAAATTTCACAC	
mg281 Down-R	GAAAATGCTGATTATGATGC	
COMmg428-F(*Xba*I)	AGTTCTAGACCTTAAAAAACAAATCGTTTCTAATTTAAATTTAAGTTAAACTTAATCCATAAGCATATATGAAAAATAATATTAGTGATGTAAAGTTGGG	Build pMTnCatMG_428
COMmg428-R(*Not*I)	AAAGCGGCCGCTTAATATCCATATCTTTCTGTTAGC	
COMmg428-F (*Apa*I)	AGTGGGCCCTAGTATTTAGAATTAATAAAGTATGAAAAATAATATTAGTGATGTAAAGTTGGG	Build pMTnPacMG_428:YFP
COMmg220-F	AGTGGGCCCTAGTATTTAGAATTAATAAAGTATGTATAAATTAGAAAAAGCAC	Build pMTnPac-rrlA
COMmg220-R	AGTCTCGAGTTAACTAAAAACTGGAAAGG	
COMmgrs02200-F	AGTGGGCCCTAGTATTTAGAATTAATAAAGTATGGAAACTATCAGAGAAAC	Build pMTnPac-rrlB
COMmgrs02200-R	AGTCTCGAGTTATTTGCCCTCCATTATTT	
TER305-F	GCGGGATCCCACAAGCAAAATAACCTGTTC	Amplify *dnaK* terminator sequence
TER305-R	GCGGGATCCCTCGAGCTAAAAATCTGTTTTTTGGT	
COMmg127-F	AGTTCTAGACCTTAAAAAACAAATCGTTTCTAATTTAAATTTAAGTTAAACTTAATCCATAAGCATATATGAAAAAGCGGGTAAAAAAATAGTG	Build pMTnPacSpx
COMmg127-R	AAAGCGGCCGCTTATTGTTGTATTATAGGCTTTGG	

Construction of mutants used in single cell analysis

Name	Sequence 5'→ 3'	Use
Ch-F	ATGGTGAGCAAGGGCGAGGA	To amplify mCherry/eYFP gene
Ch-R	CTACTTGTACAGCTCGTCCA	
YFP-R (XhoI)	AGTCTCGAGTTACTTGTACAGCTCGTCC	
ch:cat-F	TGGACGAGCTGTACAAGTAGAATTGTAGTATTTAGAATTAATAAAG	To amplify *cat* marker
cat-R	TTACGCCCCGCCCTGCCAC	
mg428:Ch Up-R	TCCTCGCCCTTGCTCACCATATATCCATATCTTTCTGTTAGC	To generate pMG428:Ch
mg428:Ch Down-F	GTGGCAGGGCGGGGCGTAACAGAAAGATATGGATATTAAATAAAG	
RecA:Ch Up-F	GGTGTAATGAAAGATGGCTC	To generate pRecA:Ch
RecA:Ch Up-R	TCCTCGCCCTTGCTCACCATGCTAGCTGTTTGTTGAAATG	
RecA:Ch Down-F	GTGGCAGGGCGGGGCGTAAAAACTAATTTTTTGATTTTCTGG	
RecA:Ch Down-R	CCACCACTTTCACTAGGATC	
mg220:Ch Up-R	GTGGCAGGGCGGGGCGTAAGTTAAAACTACTGTTGTTAACAC	To generate pRrlA:Ch
mg220:Ch Down-F	TCCTCGCCCTTGCTCACCATACTAAAAACTGGAAAGGAATAG	
mgrs02200:Ch Up-R	TCCTCGCCCTTGCTCACCATTTTGCCCTCCATTATTTTTTTG	To generate pRrlB:Ch
mgrs02200:Ch Down-F	GTGGCAGGGCGGGGCGTAAAAATGATTACATCTATCTTTGG	
MutRecA-10-F	GAGCCATCTTTCATTACACCTTTATTTAAAGTGCAAAAAAATAAAAATGACATTTTTTTATTAGAAATTTTG	To generate RecA:Ch mutants with scrambled promoter sequences
MutRecA-35-F	GAGCCATCTTTCATTACACCTTATAAATAAGTGCAAAAAAATAAAAATTTTTTTTTTTTTTATTAGAAATTTTG	
MutRecA-Int-F	GAGCCATCTTTCATTACACCTTATAAATAAGTGCTTTTTTATAAAAATGACATTTTTTTATTAGAAATTTTG	
MutRecA-R	ACCATCGTTCACTTGATGCC	

mg427:Ch Up-R	TCCTCGCCCTTGCTCACCATGTGGACTAAAGTAACGCTAA	To generate pMG427:Ch
mg427:Ch Down-F	GTGGCAGGGCGGGGCGTAAAACTTATTAAAAAAGCGAAAAATACCTG	
mg427:Ch Down-R	TTTCTCAATAGGTAGCAATG	
mg192.1Ch Up-F	GGTTGTTTGAAATTGATACC	To generate pMG192.1:Ch
mg192.1Ch Up-R	CTCGCCCTTGCTCACCATATTTTTAAAATAACGGCGTT	
mg192.1Ch Down-F	GTGGCAGGGCGGGGCGTAACTATGTATTCTACTTTAAGAC	
mg192.1Ch Down-R	CTCCTAATTTATCTGTTTCG	
COMmg220-F P427	AGTGGGCCCTTAAAAAACAAAATCGTTTCTAATTTAAATTTAAGTTAAACTTAATCCATAAGCATATATGTATAAATTAGAAAAAGCAC	To generate pTnTet-TAP plasmids
COMmgrs02200-F P427	AGTGGGCCCTTAAAAAACAAAATCGTTTCTAATTTAAATTTAAGTTAAACTTAATCCATAAGCATATATGGAAACTATCAGAGAAAC	
TAPmg220-F	ATTCCTTTCCAGTTTTTAGTGGCCTTGCTGGCCTAATGG	
TAPrrlB-F	AAAAAATAATGGAGGGCAAAGGCCTTGCTGGCCTAATGG	
TAP-R	AGTCTCGAGTTAGGTTGACTTCCCCGC	

Supp. Table S.8. Oligonucleotides used in mutant screenings

Name	Sequence 5'→ 3'	Use
SCRmg428-F	CACTGATAATCGAATAGCTG	To screen for MG_428 mutants
SCRmg428-R	AGTTGAAGATGGCAAAACAG	
SCRmg220-F	TGATGTTATTGGCGGTAGTG	To screen for *rrlA* mutants
SCRmg220-R	AGGTCCATTAAGATCCCATC	
SCRmpn534h-F	CGATGTGTTTAGTGATTTGC	To screen for *rrlB* mutants
SCRmpn534h-R	CAAGTTGTGTAGTTGCACCA	
SCRmg010-F	GTATTTCTGCAAAGCAAAAG	To screen for MG_010 mutants
SCRmg010-R	ATGCAACTTTTTTATTGCGG	
SCRmg390-F	ATTCCACTTAATGCTTTGCG	To screen for *sunT* mutants
SCRmg390-R	GTGCTGTTGAATATCACAAC	
SCRmg414-F	TTCTCTTCAGAAAGTGAAAC	To screen for MG_414 mutants
SCRmg414-R	AAGCCCAGTGTTAAAAATGG	
SCRmg525-F	GGAGGAACTTTCACTAGTTC	To screen for MG_525 mutants
SCRmg525-R	GTTGAACTTAACCCAATTGG	
SCRrecA-F	TGAACTTTGTGATGAACCAC	To screen for *recA* mutants
SCRrecA-R	AACTGGTTCTCTACAAATTG	
SCRmg427-F	CAATCATAAAGGGACATGTG	To screen for MG_427 mutants
SCRmg281-F	GATTTTTCAACTTCTGAAAG	To screen for MG_281 mutants
SCRmg281-R	TTAAACGTCCTGAGATTATC	
SCRmg192.1-F	CAGTGCAGTTATTGCTAAAG	To screen for MG_192.1 mutants
SCRmg192.1-R	ATTTAGTTGATAAAAACCGC	
SCRmg438-F	GTTCTTTAACAGTGATTCGTG	To screen for Cat:Ch mutants
SCRmg438-R	GTGCTTTATTCCATTTTGTAATG	

Supp. Table S.9. Oligonucleotides used for recombinant protein expression

Name	Sequence 5'→ 3'	Use
EXPmg428-F	CATATGAAAAATAATATTAGTGATGTAAAGTTGGGACTGTTAG CAGCAAAAATTTATTGGAAATCTTGGCGCTTTTTAGAG	Build pET21aMG428
EXPmg428-R	CTCGAGACCCTGAAAATACAGATTTTCATATCCATATCTTTCT GTTAG	

Supp. Table S.10. Oligonucleotides used in Primer Extension analysis

Name	Sequence 5'→ 3'	Use
PEmg428 (I)	[6-FAM] GCTCCATTTAATTTGAGATAG	To rule out the presence of a TSS immediately
PEmg428 (II)	[6-FAM] TTCTTAGAATCTTGCTCTGC	upstream of the MG_428 gene.
PEmg427 (I)	[6-FAM] GTTCACATGCAGCTAAACCA	To determine the TSS of MG_427 and MG_428
PEmg427 (II)	[6-FAM] CCTTTGATGTTAATCAGTGC	genes
PEmg339	[6-FAM] CTACCTAAAGGTAGACCACC	To determine TSS of *recA* gene
PEmg220	[6-FAM] GGAATAGGAAGAACTACCTG	To determine TSS of *rrlA* gene
PEmgRS02200	[6-FAM] GAATTGCGCAGCTTATCGTC	To determine TSS of *rrlB* gene
PEmg414	[6-FAM] ATCCAGCTGTATAGCTTCTG	To determine TSS of MG_414 gene
PEmg358	[6-FAM] GAAATCCAGTTGTGCTCAAC	To determine TSS of *ruvA* gene (*rrlB* operon)
PEmg192.1	[6-FAM] CTTCCGTTTTGAGATCTTGTTT	To determine TSS of MG_192.1 gene
PEncRNA-3	[6-FAM] GCCTATGGGTTAACAAAGAGATG	To determine TSS of the ncRNA-3/4
PEncRNA-1	[6-FAM] AGGGAAGTTTCAGTGTCC	To determine TSS of the ncRNA-1
PEmg285	[6-FAM] TAGTAAGGTGAATTTGAGCC	To determine TSS of MG_285 gene
PEmg412	[6-FAM] CTGATTCCAGCATTTGA	To determine TSS of MG_412 gene

Supp. Table S.11. Oligonucleotides used in qRT-PCR analysis

Name	Sequence 5'→ 3'	Use
RTPCRmg418-F	TGTTGACGCTAGTGGTTTGG	Housekeepìng gene (*rplM*)
RTPCRmg418-R	TTCCACCCATGTATTGAGAGTG	
RTPCRmg177-F	TGAGTGTCCAGCTGGTTTTG	Housekeepìng gene (*rpoA*)
RTPCRmg177-R	AACCGGGGAAAAGTTAGCAT	
RTPCRmg341-F	TGTCCTCTCCTTGCTCAGGT	Housekeepìng gene (*rpoB*)
RTPCRmg341-R	CCCTTTGGCCTATTTCAACA	
RTPCRmg238-F	TCCCTGAACAATTCCAAAGC	Housekeepìng gene (*tuf*)
RTPCRmg238-R	CGCTTCTGGAAGAAAAGGAA	
RTPCRmg430-F	GGAAGCAGTTGGATTGCCTA	Housekeepìng gene (*gpmI*)
RTPCRmg430-R	ATGCACTCCTCCATTGGAAA	
RTPCRmg427-F	GCAGCTGCTAAGATGATAAC	mRNA quantification of *osmC* gene
RTPCRmg427-R	GTGGACTAAAGTAACGC	
RTPCRmg428-F	GTGATGTAAAGTTGGGACTG	mRNA quantification of MG_428 gene
RTPCRmg428-R	CAAGCAATTCAACTTTATTCACC	
RTPCRmg281-F	ACGGAAGTTAGCACCACCAG	mRNA quantification of protein M gene
RTPCRmg281-R	GCTGGAGCGTTAACCACATT	
RTPCRmg339-F	CCATGCCTAGGATGAGCAAT	mRNA quantification of *recA* gene
RTPCRmg339-R	GCATTAGGGTCTGGTGGTCT	
RTPCRmgrs02200-F	GATAAGCTGCGCAATTCGAT	mRNA quantification of *rrlB* gene
RTPCRmgrs02200-R	GCATCACCAATTTGAAAAATCA	
RTPCRmg358-F	TGCTCAATAATGACCTTGAGGA	mRNA quantification of *ruvA* gene
RTPCRmg358-R	CCATCAGCAACTTGAACTCC	
RTPCRmg359-F	CCACCTGGTGTGGGTAAAAC	mRNA quantification of *ruvB* gene
RTPCRmg359-R	ATTAGGTGCTACGGCATGGA	
RTPCRmg191-F	CTGGAGAGAACCCAGGATCA	mRNA quantification of *mgpB* gene
RTPCRmg191-R	AGGGGTTTTCCATTTTTGCT	
RTPCRmg192-F	TCCCCTAATGAATTGCGAAG	mRNA quantification of *mgpC* gene
RTPCRmg192-R	CAGGGGCAATTGATTTAAGC	
RTPCRmg220-F	AAACCGTGCGGTGTTAAAGT	mRNA quantification of *rrlA* gene
RTPCRmg220-R	GCTAAGCAAGTTGTTGGTGGA	

RTPCRmg390-F	CTCTGAACGGGTTATGCAGC	mRNA quantification of *sunT* gene
RTPCRmg390-R	TATGGTACGCGATGCTGTTG	
RTPCRmg389-F	TCTTTGAACTACGCAACAACAA	mRNA quantification of MG_389 gene
RTPCRmg389-R	TCAGCCCTGTTCTAGAGTAACA	
RTPCRmg414-F	TGGCCAAGAATCAAGAGGAT	mRNA quantification of MG_414 gene
RTPCRmg414-R	CCAGGAAAAGGTTCGTTTCA	
RTPCRmg525-F	GCTTTGGTGTTATTTTTGATCC	mRNA quantification of MG_525 gene
RTPCRmg525-R	TTCCTTTGCGATAATTCCCTTA	
RTPCRmg010-F	AAAAAGGCGTGGATTCAATTT	mRNA quantification of MG_010 gene
RTPCRmg010-R	AAAAGCTTCGCTCTTTGAAGG	
RTPCRmCherry-F	GCCCCTAATGCAGAAGAAG	mRNA quantification of *mCherry* gene
RTPCRmCherry-R	GTGTAGTCCTCGTTGTGGGA	
RTPCRmg285-F	GATGTACCCCGTTCAGCACT	mRNA quantification of MG_285 gene
RTPCRmg285-R	TCATGTTGCCAAAGTTGATTG	

Supp. Table S.12. Oligonucleotides used in sequencing reactions

Name	Sequence 5'→ 3'	Description
Fup17	GTAAAACGACGGCCAGT	Universal Forward primer
Rup17	GGAAACAGCTATGACCATG	Universal Reverse primer
TetUp	TTCCTGCATCAACATGAG	To determine miniTnTet Tnp insertion point
TetDown	GTCGTCCAAATAGTCGGA	To determine miniTnTet Tnp insertion point
CmUp	CAACGGTGGTATATCCAG	To determine miniTnCat Tnp insertion point
CmDown	CAGTACTGCGATGAGTGGCA	To determine miniTnCat Tnp insertion point
PacUp	GTAGCTAATCTAACAGTAGG	To determine miniTnPac Tnp insertion point
PacDown	GTCCTAGAACTTGGTGTATG	To determine miniTnPac Tnp insertion point
T7	TAATACGACTCACTATAGGG	To sequence pET21 plasmid
T7prom	GCTAGTTATTGCTCAGCGG	To sequence pET21 plasmid
Tnp3	CATGATGAATGGATTTATTC	To sequence DNA inserts cloned in a miniTnp Tnp

S.3.2. Plasmid construction and generation of mutants

S.3.2.1. Construction of pΔMG_428 plasmid and generation of ΔMG_428 mutant strain

The suicide plasmid pΔMG_428 was created to generate a *M. genitalium* MG_428 null mutant by homologous recombination (HR). First, the MG_428 upstream region (UR) was PCR-amplified with the *mg428 Up-F* and *mg428 Up-R* primers (see Supp. Table S.7). Similarly, the MG_428 downstream region (DR) was PCR-amplified with the *mg428 Down-F* and *mg428 Down-R* primers. In parallel, the *tetM438* selectable marker (Pich *et al.*, 2006b) was PCR-amplified with the *Tc-F* and *Tc-R* primers. Then, the UR and *tetM438* PCR products were joined by Splicing by Overlap Extension (SOE) PCR with the *mg428 Up-F* and *Tc-R* primers. Similarly, the obtained recombinant amplicon and the DR PCR product were also joined by SOE-PCR with the *mg428 Up-F* and *mg428 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid (Pich *et al.*, 2006b).
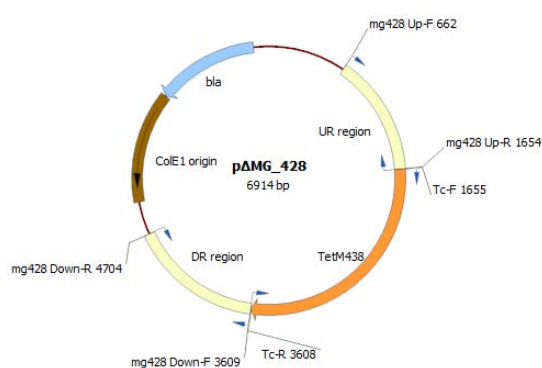
**Supp. Figure S.8. pΔMG_428 plasmid map.** Oligonucleotides used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

Generation of ΔMG_428 strain is described in Chapter I, Figure I.3.

S.3.2.2. Construction of pMTn*Cat*MG_428 plasmid

This plasmid carries the TnCatMG_428 MiniTnp, which was used to reintroduce a wild-type copy of the MG_428 allele under the control of its own promoter into a ΔMG_428 mutant. The MG_428 allele was PCR-amplified with the *COMmg428-F*(*Xba*I) and *COMmg428-R* (*Not*I) primers, digested with *Not*I and *Xba*I and ligated into a similarly digested pMTnCat plasmid (Calisto *et al.*, 2012).
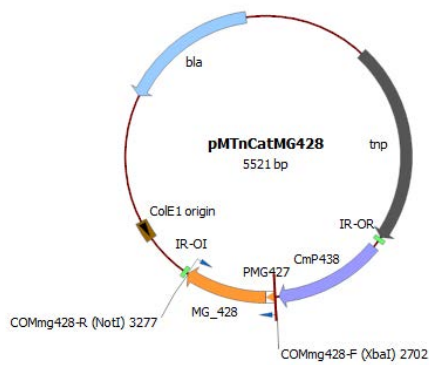
Supp. Figure S.9. pMTn*Cat*MG_428 plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; PMG427, MG_427 promoter; CmP438, *cat* marker with MG_438 promoter; tnp, transposase; *bla*, β-lactamase.

S.3.2.3. Construction of pET21aMG428 plasmid

This plasmid was used to express a C-terminal His-tagged MG428 recombinant protein. The MG_428 allele was PCR-amplified with the *EXPmg428-F* and *EXPmg428-R* primers. In this step, a TGA codon at the 5' end of the MG_428 gene was changed to TGG to prevent premature translation termination in *E. coli*. The PCR fragment was cloned into an *Eco*RV-digested pBE plasmid, excised with *Nde*I-*Xho*I and ligated into a similarly digested pET21a vector (Novagen).
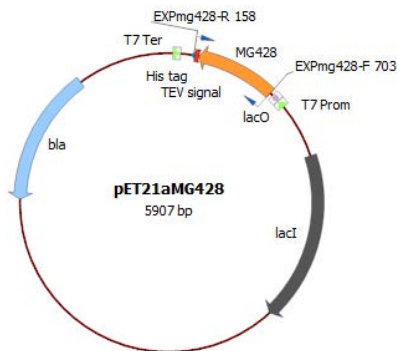
Supp. Figure S.10. pET21MG428 plasmid map. Primers used to clone MG428 in this plasmid are indicated as blue arrows. T7Prom, T7 promoter region; lacO, lac operator region; TEV signal, TEV protease cleavage site; T7Ter, T7 terminator sequence; *bla*, β-lactamase.

S.3.2.4. Construction of pΔMG_218-lacZ plasmid and generation of ΔMG_218 mutants

The suicide plasmid pΔMG_218-lacZ was created to assess recombination in different *M. genitalium* strains. The puromycin resistance cassette (Algire *et al.*, 2009) (*pac*) was PCR-amplified from plasmid pMiniHimar1BSC1 (Maglennon *et al.*, 2013a) with the *EcoRIpac-F* and *BamHI pac-R* primers, cloned into an *Eco*RV-digested pBE and digested with *Eco*RI and *Bam*HI. The *E. coli lacZ* gene was PCR-amplified from plasmid pMC1871 (Shapira SK *et al.*, 1983) with the *SalI lacZ-F* and *XhoI lacZ-R* primers, cloned into an *Eco*RV-digested pBE and excised with *Sal*I-*Xho*I. Finally, the pΔMG_218 plasmid (Pich *et al.,* 2008) was digested with *Eco*RI and *Sal*I to excise the *tetM438* marker and replace it by the *pac438* marker plus the *lacZ* gene to obtain pΔMG_218-lacZ.
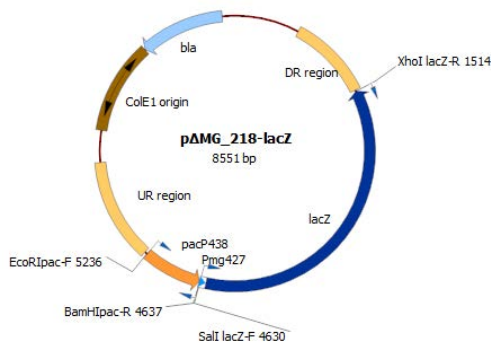
Supp. Figure S.11. pΔMG_218-lacZ plasmid map. Oligonucleotides used to create this plasmid are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; pacP438, puromycin resistance cassette with MG_438 promoter; *bla*, β-lactamase.

ΔMG_218-lacZ mutants were generated by allelic exchange in a double crossover when transforming with pΔMG_218-lacZ suicide plasmid. Screening for mutants was performed visualizing the *lacZ*-expressing colonies (see Figure I.19).

S.3.2.5. Construction of pMG_428:Ch plasmid and generation of MG428:Ch strains

This suicide plasmid was created to generate a C-terminal fusion of the MG428 polypeptide to the fluorescent protein mCherry. A 1.4 kb fragment containing the MG_428 ORF without the stop codon was PCR-amplified with the *mg428 Up-F* and *mg428:Ch Up-R* primers. In parallel, a 1 kb PCR fragment encompassing the MG_428 DR was PCR-amplified with the *mg428:Ch Down-F* and *mg428 Down-R* primers. The *cat438* marker (687 bp) was PCR-amplified from plasmid pMTn*Cat* with the *ch:cat-F* and *cat-R* primers, and the mCherry coding region (711 bp) was PCR-amplified from plasmid pmCherry (Clontech-Takara) with the *Ch-F* and *Ch-R* primers. Then, the *cat438* marker and the mCherry PCR product were joined by SOE-PCR with the *Ch-F* and *cat-R* primers. The resulting recombinant amplicon (1398 bp) and the 1.4 kb PCR fragment containing the MG_428 ORF were also joined by SOE-PCR with the *mg428 Up-F* and *cat-R* primers. Finally, the obtained 2.8 kb PCR product and the 1 kb PCR fragment encompassing the MG_428 DR were joined by SOE-PCR with the *mg428 Up-F* and *mg428 Down-R* primers. The final PCR product (3.8 kb) was cloned into a *Eco*RV-digested pBE to create pMG_428:Ch.

Supp. Figure S.12. pMG_428:Ch plasmid map. Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase.
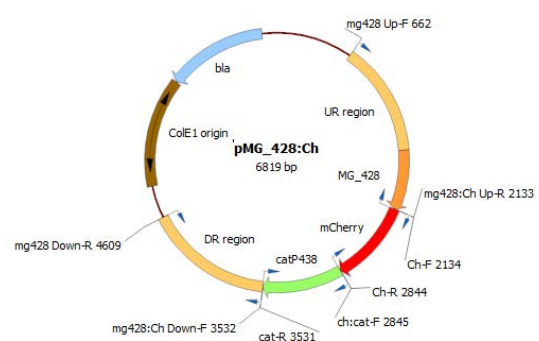
Generation of MG428:Ch strains is described in Figure I.15.

S.3.2.6. Construction of pRecA:Ch plasmid and generation of RecA:Ch strains

This suicide plasmid was created to generate a C-terminal fusion of the MG339 (RecA) polypeptide to the fluorescent protein mCherry. A 1 kb fragment containing the MG_339 ORF without the stop codon was PCR-amplied with the *recA:Ch Up-F* and *recA:Ch Up-R* primers. This PCR product and the 1398 bp recombinant amplicon containing the *cat438* marker plus the mCherry coding region (see pMG428:Ch) were joined by SOE-PCR with the *recA:Ch Up-F* and *cat-R* primers. Then, a 1 kb amplicon encompassing the MG_339 DR was PCR-amplified with the *recA:Ch Down-F* and *recA:Ch Down-R* primers. The final PCR product (3.4 kb) was obtained by SOE-PCR with the *recA:Ch Up-F* and *recA:Ch Down-R* primers, and cloned into a *Eco*RV-digested pBE to create pRecA:Ch.
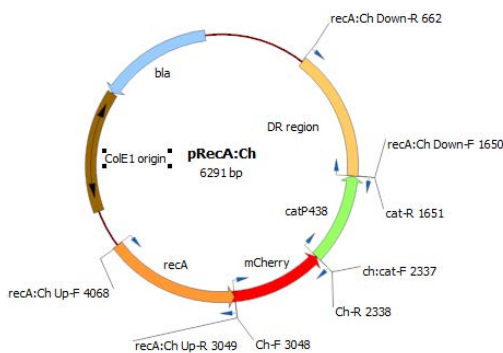
Supp. Figure S.13. pRecA:Ch plasmid map. Primers are mapped as blue arrows. DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase. *recA* gene sequence assists as homologous upstream region for allelic exchange.

Generation of RecA:Ch strains is described in in Figure I.15.

S.3.2.7. Construction of pCat:Ch plasmid and generation of Cat:Ch strain

This plasmid was created in a previous work (Broto, unpublished). Detailed description of the construction of this plasmid can be found in the supplementary material of Torres-Puig *et al.*, 2015.
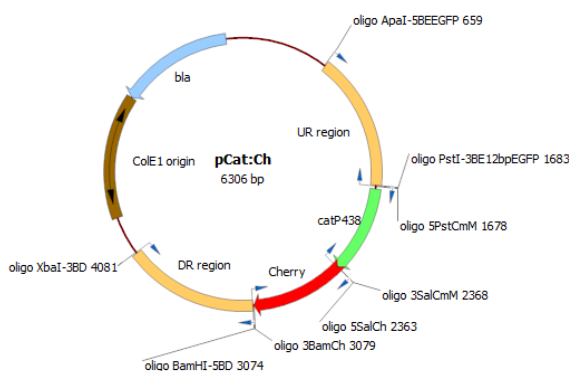
Supp. Figure S.14. pCat:Ch plasmid map. Primers and restriction sites used to clone each DNA fragment are indicated with blue arrows.

Generation of Cat:Ch strain is described in in Figure I.15 as well as in Torres-Puig *et al.*, 2015.

S.3.2.8. Construction of pRecA:Ch-10, pRecA:Ch-22 and pRecA:Ch-35 plasmids and generation of the corresponding mutant strains

pRecA:Ch-10. This suicide plasmid was created to generate a C-terminal fusion of the MG339 (RecA) polypeptide to the fluorescent protein mCherry under the control of a *recA* promoter with a scrambled -10 region (see Figure I.17). First, the UR of the *recA* gene (350 bp) carrying the scrambled -10 promoter region was PCR-amplified with the *MutRecA-10-F* and *MutRecA-R* primers. This amplicon was then joined with the PCR fragment encompassing the MG_339 ORF without the stop codon (see pRecA:Ch plasmid, S.3.2.6) by SOE-PCR with the *MutRecA-R* and *recA:Ch Up-R* primers. This PCR product and the recombinant amplicon containing the *cat438* marker plus the mCherry coding region (see pMG_428:Ch plasmid) were joined by SOE-PCR with the *MutRecA-R* and *cat-R* primers. Next, a 1 kb amplicon encompassing the *recA* DR was PCR-amplified with the *recA:Ch Down-F* and *recA:Ch Down-R* primers. The final PCR product (3.8 kb) was obtained by SOE-PCR with the *MutRecA-R* and *recA:Ch Down-R* primers, and cloned into a *Eco*RV-digested pBE to create pRecA:Ch-10.

pRecA:Ch-22. This suicide plasmid was created to generate a C-terminal fusion of the MG339 (RecA) polypeptide to the fluorescent protein mCherry under the control of a *recA* promoter with a scrambled -22 region (see Figure I.17). This plasmid was constructed using the same steps as for the construction of the *pRecA:Ch-10* plasmid, except that the UR of the *recA* gene (350 bp) carrying the scrambled -22 promoter region was PCR-amplified with the *MutRecA-22-F* and *MutRecA-R* primers.

pRecA:Ch-35. This suicide plasmid was created to generate a C-terminal fusion of the MG339 (RecA) polypeptide to the fluorescent protein mCherry under the control of a *recA* promoter with a scrambled -35 region (see Figure I.17). This plasmid was constructed using the same steps as for the construction of the *pRecA:Ch-10* plasmid, except that the UR of the *recA* gene (350 bp) carrying the scrambled -35 promoter region was PCR-amplified with the *MutRecA-35-F* and *MutRecA-R* primers.
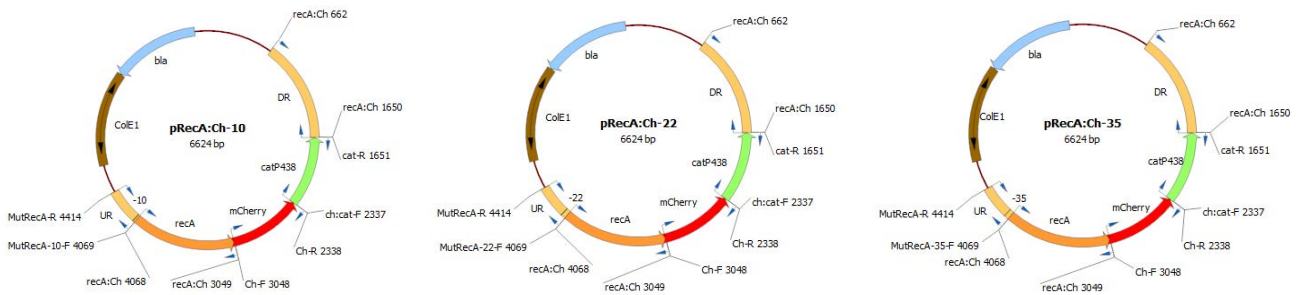
Supp. Figure S.15. pRecA:Ch-10, pRecA:Ch-22 and pRecA:Ch-35 plasmid maps. Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase.

Generation of RecA:Ch-10, RecA:Ch-35 and RecA:Ch-22 strains is described in see Figure I.17.

S.3.2.9. Construction of pΔMG_010 plasmid and generation of ΔMG_010 mutant strain.

This suicide plasmid was created to generate an *M. genitalium* MG_010 null mutant by HR. First, the MG_010 UR was PCR-amplified with the *mg010 Up-F* and *mg010 Up-R* primers. Similarly, the MG_010 DR was PCR-amplified with the *mg010 Down-F* and *mg010 Down-R* primers. Then, the MG_010 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg010 Up-F* and *Tc-R* primers. Next, the obtained recombinant amplicon and the MG_010 DR were joined by SOE-PCR with the *mg010 Up-F* and *mg010 Down-R* primers. Finally, the resulting PCR product (~3.8 kb) was cloned into an *Eco*RV-digested pBE plasmid.
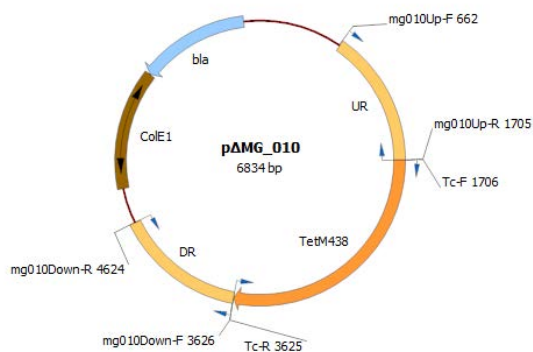
Supp. Figure S.16. pΔMG_010 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.
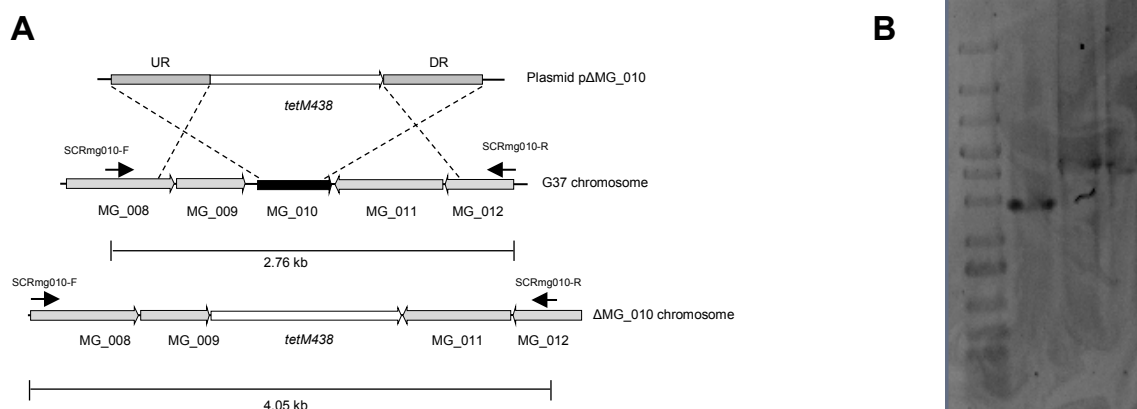
Supp. Figure S.17. Construction of a ΔMG_010 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_010 locus by the tetM438 marker in different ΔMG_010 mutant clones.

After electroporation with pΔMG_010, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_428 in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg010-F* and *SCRmg010-R*) were complementary to sequences of the chromosome located immediately outside of the MG_010 flanking regions cloned into the pΔMG_010 plasmid (Supp. Figure S.17). The amplification of a 4.0 kb band confirmed the replacement of the MG_010 gene (657 bp) by the tetM438 marker (1954 bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the MG_428 gene. As expected, a band of 2.7 kb was observed when the chromosomal DNA of the wild-type strain was analyzed.
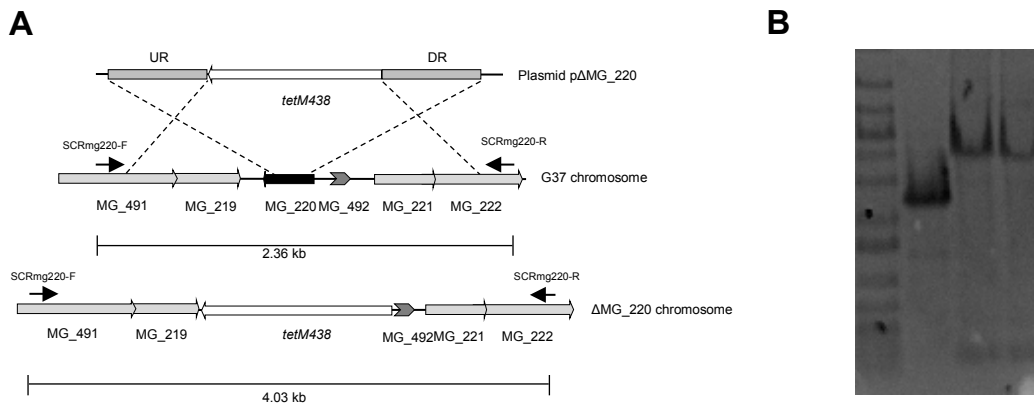
S.3.2.10. Construction of pΔMG_220 plasmid and generation of ΔMG_220 mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_220 null mutant by HR. First, the MG_220 UR was PCR-amplified with the *mg220 Up-F* and *mg220 Up-R* primers. Similarly, the MG_220 DR was PCR-amplified with the *mg220 Down-F* and *mg220 Down-R* primers. Then, the MG_220 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg220 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_220 DR were joined by SOE-PCR with the *mg220 Up-F* and *mg220 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.
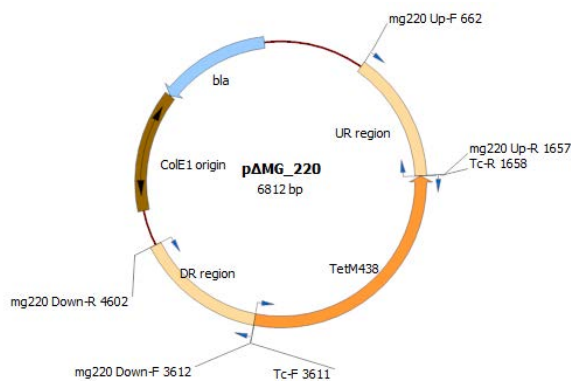
Supp. Figure S.18. pΔMG_220 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

After electroporation with pΔMG_220, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_220 (*rrlA*) in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg220-F* and *SCRmg220-R*) were complementary to sequences of the chromosome located immediately outside of the MG_220 flanking regions cloned into the pΔMG_220 plasmid (Supp. Figure S.19). The amplification of a 4.0 kb band confirmed the replacement of the MG_220 gene (282 bp) by the tetM438 marker (1954 bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the MG_220 gene. As expected, a band of 2.3 kb was observed when the chromosomal DNA of the wild-type strain was analyzed.
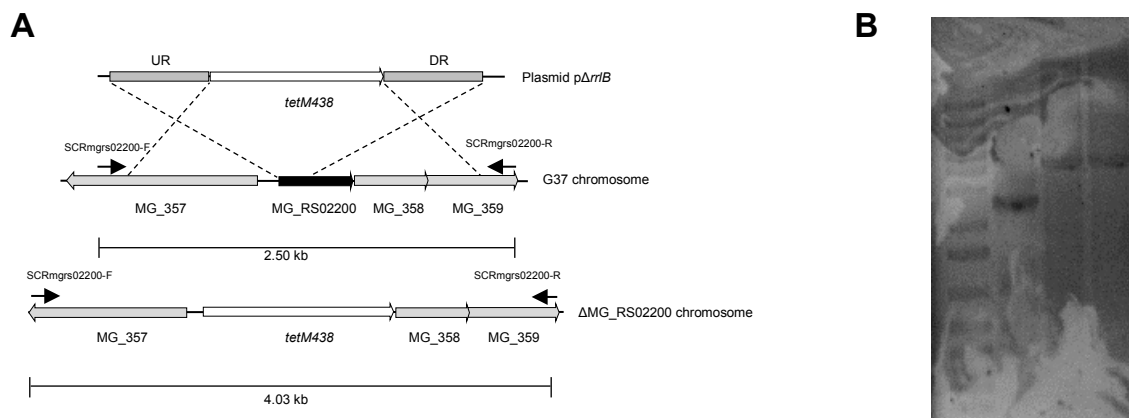

Supp. Figure S.19. Construction of a ΔMG_220 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_220 locus by the tetM438 marker in different ΔMG_220 mutant clones.

S.3.2.11. Construction of pΔMG_RS02200 (pΔ*rrlB*) plasmid and generation of ΔMG_RS02200 (Δ*rrlB*) mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_RS02200 (*rrlB*) null mutant by HR. First, the *rrlB* UR was PCR-amplified with the *mg_rs02200 Up-F* and *mg_rs02200 Up-R* primers. Similarly, the *rrlB* DR was PCR-amplified with the *mg_rs02200 Down-F* and *mg_rs02200 Down-R* primers. Then, the *rrlB* UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg_rs02200 Up-F* and *Tc-R* primers. Next, the obtained recombinant amplicon and the *rrlB* DR were joined by SOE-PCR with the *mg_rs02200 Up-F* and *mg_rs02200 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.
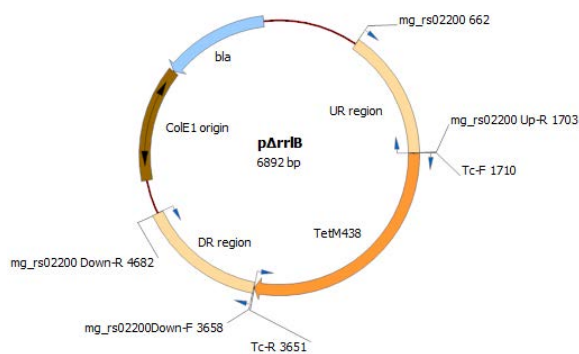
Supp. Figure S.20. pΔMG_RS02200 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

Supp. Figure S.21. Construction of a ΔMG_RS02200 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the *rrlB* locus by the tetM438 marker in different ΔMG_RS02200 mutant clones.
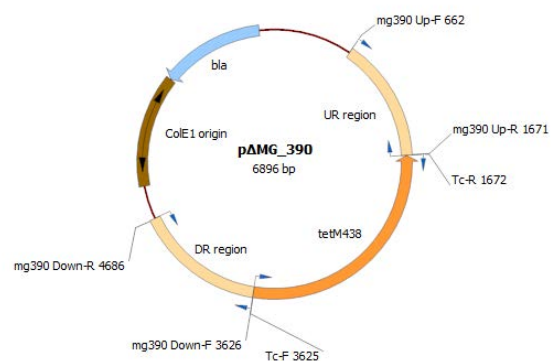
After electroporation with pΔMG_RS02200, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_RS02200 (*rrlB*) in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmgrs022000-F* and *SCRmgrs02200-R*) were complementary to sequences of

the chromosome located immediately outside of the *rrlB* flanking regions cloned into the pΔMG_RS02200 plasmid (Supp. Figure S.21). The amplification of a 4.0 kb band confirmed the replacement of the *rrlB* gene (426 bp) by the tetM438 marker (1954 bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the *rrlB* gene. As expected, a band of 2.3 kb was observed when the chromosomal DNA of the wild-type strain was analyzed.
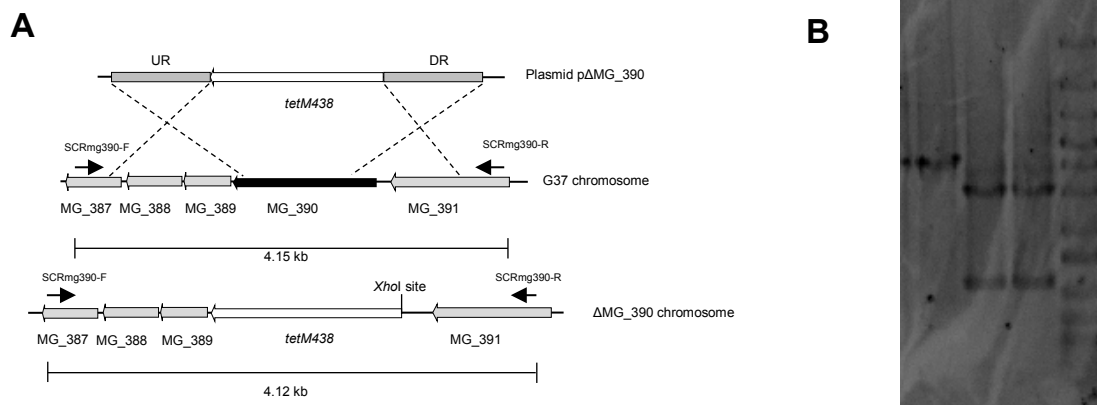
S.3.2.12. Construction of pΔMG_390 plasmid and generation of ΔMG_390 mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_390 null mutant by HR. First, the MG_390 UR was PCR-amplified with the *mg390 Up-F* and *mg390 Up-R* primers. Similarly, the MG_390 DR was PCR-amplified with the *mg390 Down-F* and *mg390 Down-F* primers. Then, the MG_390 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg390 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_390 DR were joined by SOE-PCR with the *mg390 Up-F* and *mg390 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.

Supp. Figure S.22. pΔMG_390 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

After electroporation with pΔMG_390, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_390 in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg390-F* and *SCRmg390-R*) were complementary to sequences of the chromosome located immediately outside of the MG_390 flanking regions cloned into the pΔMG_390 plasmid (Supp. Figure S.23). Then, the PCR product was digested with *Xho*I. The presence of two bands of ~1kb and ~3kb confirmed the replacement of the MG_390 gene (1983 bp) by the tetM438 marker (1954 bp, it bears a *Xho*I site) in the two transformants analyzed. Sequencing of the digested PCR products corroborated the intended deletion of the MG_390 gene. As expected, a single band of 4.15 kb was observed when the chromosomal DNA of the wild-type strain was analyzed and digested with *Xho*I.
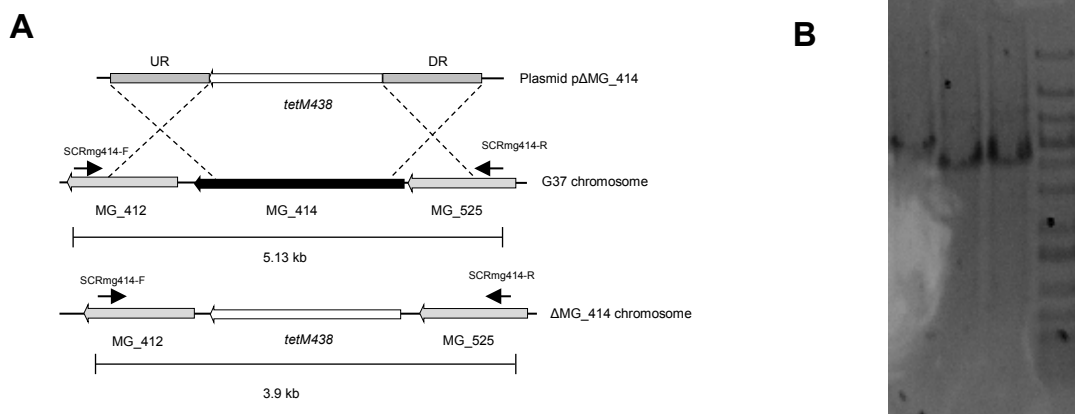
Supp. Figure S.23. Construction of a ΔMG_390 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_390 locus by the tetM438 marker in different ΔMG_390 mutant clones.

S.3.2.13. Construction of pΔMG_414 plasmid and generation of ΔMG_414 mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_414 null mutant by HR. First, the MG_414 UR was PCR-amplified with the *mg414 Up-F* and *mg414 Up-R* primers. Similarly, the MG_414 DR was PCR-amplified with the *mg414 Down-F* and *mg414 Down-R* primers. Then, the MG_414 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg414 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_414 DR were joined by SOE-PCR with the *mg414 Up-F* and *mg414 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.
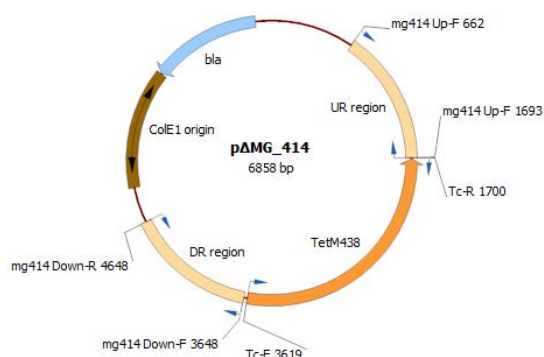
Supp. Figure S.24. pΔMG_414 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

A

B

Supp. Figure S.25. Construction of a ΔMG_414 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_414 locus by the tetM438 marker in different ΔMG_414 mutant clones.

After electroporation with pΔMG_414, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_414 in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg414-F* and *SCRmg414-R*) were complementary to sequences of the chromosome located immediately outside of the MG_414 flanking regions cloned into the pΔMG_414 plasmid (Supp. Figure S.25). The amplification of a 3.9 kb band confirmed the replacement of the MG_414 gene (3111 bp) by the tetM438 marker (1954 bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the MG_414 gene. As expected, a band of 5.1 kb was observed when the chromosomal DNA of the wild-type strain was analyzed.
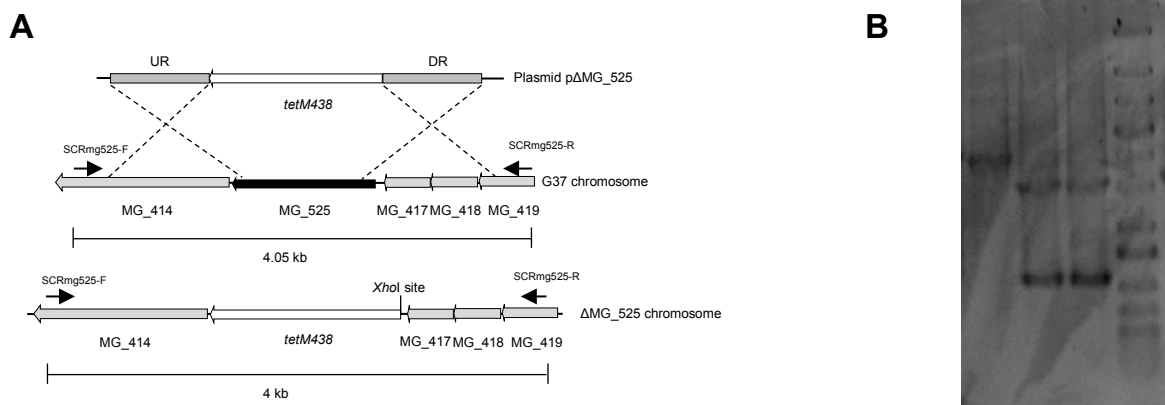
S.3.2.14. Construction of pΔMG_525 plasmid and generation of ΔMG_525 mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_525 null mutant by HR. First, the MG_525 UR was PCR-amplified with the *mg525 Up-F* and *mg525 Up-R* primers. Similarly, the MG_525 DR was PCR-amplified with the *mg525 Down-F* and *mg525 Down-R* primers. Then, the MG_525 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg525 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_525 DR were joined by SOE-PCR with the *mg525 Up-F* and *mg525 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.
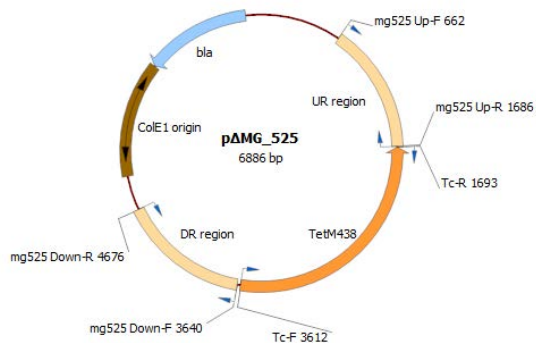
Supp. Figure S.26. pΔMG_525 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

After electroporation with pΔMG_525, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_525 in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg525-F* and *SCRmg525-R*) were complementary to sequences of the chromosome located immediately outside of the MG_525 flanking regions cloned into the pΔMG_525 plasmid (Supp. Figure S.27). Then, the PCR product was digested with *Xho*I. The presence of two bands of ~1kb and ~3kb confirmed the replacement of the MG_525 gene (1995 bp) by the tetM438 marker (1954 bp, it bears a *Xho*I site) in the two transformants analyzed. Sequencing of the digested PCR products corroborated the intended deletion of the MG_525 gene. As expected, a single band of 4.05 kb was observed when the chromosomal DNA of the wild-type strain was analyzed and digested with *Xho*I.
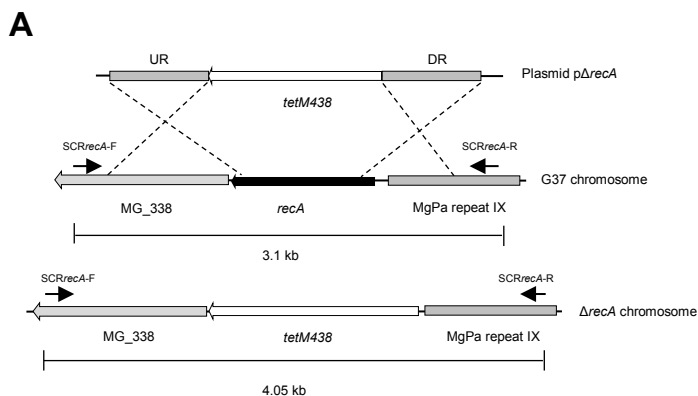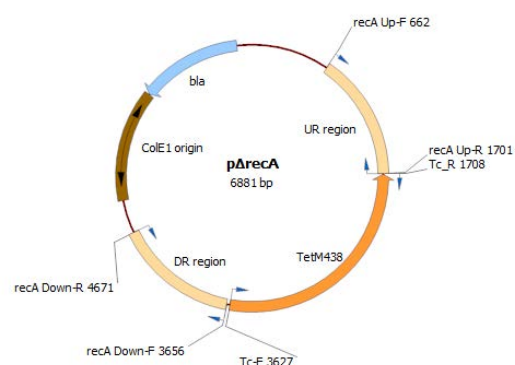

Supp. Figure S.27. Construction of a ΔMG_525 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_525 locus by the tetM438 marker in different ΔMG_525 mutant clones.

S.3.2.15. Construction of pΔ*recA* plasmid and generation of Δ*recA* mutant strain

This suicide plasmid was created to generate a *M. genitalium recA* null mutant by HR. First, the MG_339 UR was PCR-amplified with the *mg339 Up-F* and *mg339 Up-R* primers. Similarly, the MG_339 DR was PCR-amplified with the *mg339 Down-F* and *mg339 Down-R* primers. Then, the MG_339 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg339 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_339 DR were joined by SOE-PCR with the *mg339 Up-F* and *mg339 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.

Supp. Figure S.28. pΔ*recA* plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.


Supp. Figure S.29. Construction of a Δ*recA* deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_339 locus by the tetM438 marker in different Δ*recA* mutant clones.

After electroporation with pΔ*recA*, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of *recA* in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRrecA-F* and *SCRrecA-R*) were complementary to sequences of the chromosome located immediately outside of the MG_339 flanking regions cloned into the pΔ*recA* plasmid (Supp. Figure S.29). The amplification of a 3.9 kb band confirmed the replacement of the MG_339 gene (1023

bp) by the tetM438 marker (1954 bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the *recA* gene. As expected, a band of 3.1 kb was observed when the chromosomal DNA of the wild-type strain was analyzed.

S.3.2.16. Construction of pMG427:Ch and generation of MG_427:Ch mutant strain

This suicide plasmid was created to generate a C-terminal fusion of the MG427 polypeptide to the fluorescent protein mCherry. A 1kb fragment containing the MG_427 ORF without the stop codon was PCR-amplified with the *SCRmg428-F* and *mg427:Ch Up-R* primers. In parallel, a 1 kb PCR fragment encompassing the MG_427 DR was PCR-amplified with the *mg427:Ch Down-F* and *mg427 Down-R* primers. The *cat438* marker (687 bp) and the mCherry coding region (711bp) was PCR-amplified and joined together using SOE-PCR as described in S.3.2.5. The resulting recombinant amplicon and the 1kb PCR fragment containing the MG_427 ORF were also joined by SOE-PCR with the *SCRmg428-F* and *cat-R* primers. Finally, the obtained 2.4kb PCR product and the 1kb PCR fragment encompassing the MG_427 DR were joined by SOE-PCR with the *SCRmg428-F* and *mg427:Ch Down-R* primers. The final PCR product (3.4kb) was cloned into a *Eco*RV-digested pBE to create pMG427:Ch.
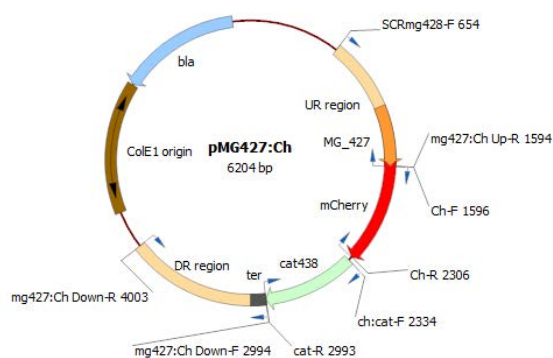
Supp. Figure S.30. pMG_427:Ch plasmid map. Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; ter, MG_427 terminator region; *bla*, β-lactamase.

Upon electroporation with pMG427:Ch plasmid, several chloramphenicol-resistant colonies were picked up and propagated. The intended transcriptional fusion of MG_427 with mCherry gene in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg427-F* and *SCRmg428-R*) were complementary to sequences of the chromosome located immediately outside of the MG_427 flanking regions cloned into the pMG427:Ch plasmid (Supp. Figure S.31).

A

B
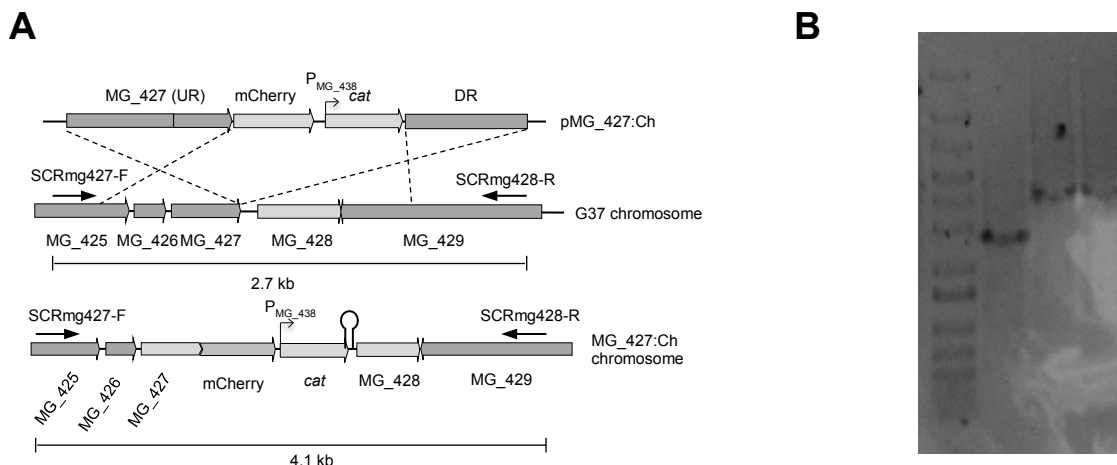
Supp. Figure S.31. Construction of a transcriptional fusion of MG_427 with mCherry by allelic exchange. **(A)** Schematic representation illustrating the construction of a transcriptional fusion in *M. genitalium* by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the insertion of mCherry and *cat* marker in different mutant clones.

The amplification of a 4.1 kb band confirmed the insertion of the intended fusion in the two transformants analyzed. Sequencing of the PCR product corroborated that the fusion was correct. As expected, a band of 2.7 kb was observed when the chromosomal DNA of the wild-type strain was analyzed. Special care to maintain the MG_427 independent terminator sequence was taken in order not to cause disturbances in MG_428 expression.

S.3.2.17. Construction of pΔMG_281 plasmid and generation of ΔMG_281 mutant strain

This suicide plasmid was created to generate a *M. genitalium* MG_281 null mutant by HR. First, the MG_281 UR was PCR-amplified with the *mg281 Up-F* and *mg281 Up-R* primers. Similarly, the MG_281 DR was PCR-amplified with the *mg281 Down-F* and *mg281 Down-R* primers. Then, the MG_281 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg281 Up-F* and *Tc-F* primers. Next, the obtained recombinant amplicon and the MG_281 DR were joined by SOE-PCR with the *mg281 Up-F* and *mg281 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid.

Supp. Figure S.32. pΔMG_281 plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.

The construction of this plasmid was made with the inestimable assist of Ignasi Granero Moya, a former undergraduate student in the lab.

A

B
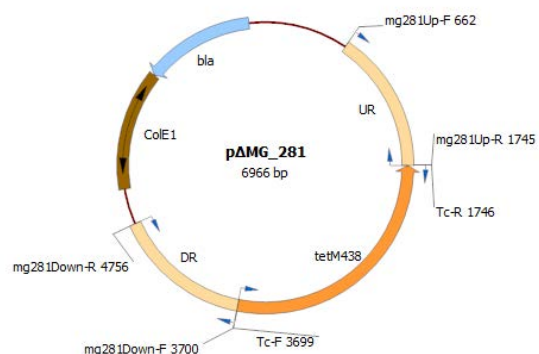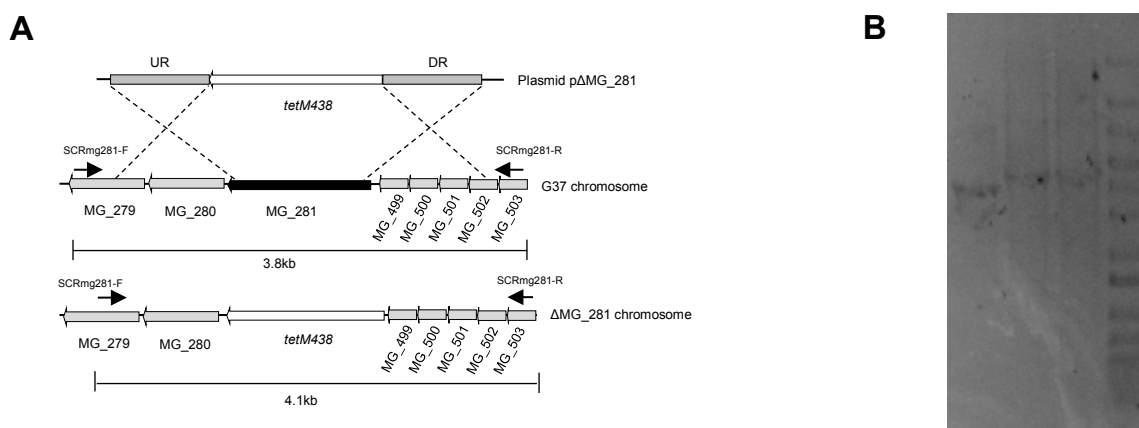
Supp. Figure S.33. Construction of a ΔMG_281 deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the MG_281 locus by the tetM438 marker in different ΔMG_281 mutant clones.

After electroporation with pΔMG_281, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of MG_281 in two selected transformants was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg281-F* and *SCRmg281-R*) were complementary to sequences of the chromosome located immediately outside of the MG_281 flanking regions cloned into the pΔMG_281 plasmid (Supp. Figure S.33). The amplification of a 4.1kb band confirmed the replacement of the MG_281 gene (1671bp) by the tetM438 marker (1954bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the MG_281 gene. As expected, a band of 3.8kb was observed when the chromosomal DNA of the wild-type strain was analyzed.

S.3.2.18. Construction of pMTnPac plasmid

pMTnPac was designed to introduce genes by transposon delivery in a *M. genitalium* chromosome using the puromycin acetyl transferase (*pac*) gene, which confers resistance to puromycin. This plasmid was created excising the TetM438 cassette from pMTn*TetM*438 plasmid (Pich *et al.*, 2006b) and cloning *pac438* marker in its place using restriction sites *Eco*RI/*Bam*HI. Puromycin acetyl transferase gene was obtained excising *pac438* marker from pΔMG_218-lacZ (this work) using *Eco*RI and *Bam*HI.

Supp. Figure S.34. pMTnPac plasmid map. IR, inverted repeat; P438, MG_438 promoter; *pac*, puromycin acetyl transferase; tnp, transposase; *bla*, β-lactamase.
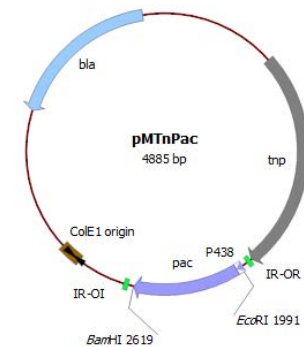
S.3.2.19. Construction of pMTnPacMG_428 plasmid

This plasmid carries the TnPacMG_428 MiniTnp, which was used to reintroduce a wild-type copy of the MG_428 allele under the control of its own promoter in the *M. genitalium* chromosome. The MG_428 allele was PCR-amplified with the *COMmg428-F(Xba*I) and *COMmg428-R* (*Not*I) primers, digested with *Not*I and *Xba*I and ligated into a similarly digested pMTnPac plasmid.

Supp. Figure S.35. pMTnPacMG_428 plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; PMG427, MG_427 promoter; *pac*, puromycin acetyl transferase marker; P438, MG_438 promoter; tnp, transposase; *bla*, β-lactamase.
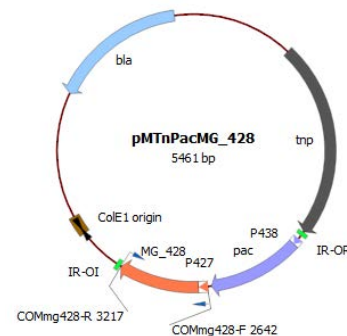
S.3.2.20. Construction of pRrlA:Ch plasmid and generation of RrlA:Ch mutant strain

This suicide plasmid was created to generate a C-terminal fusion of the MG220 polypeptide to the fluorescent protein mCherry. A 1kb fragment containing the MG_220 ORF without the stop codon was PCR-amplified with the *mg220 Up-F* and *mg220:Ch Up-R* primers. In parallel, a 1 kb PCR fragment encompassing the MG_220 DR was PCR-amplified with the *mg220:Ch Down-F* and *mg220 Down-R* primers. The *cat438* marker (687bp) and the mCherry coding region (711bp) was PCR-amplified and joined together using SOE-PCR as described in S.3.2.5. The resulting recombinant amplicon and the 1kb PCR fragment containing the MG_220 ORF were also joined by SOE-PCR with the *mg220 Up-F* and *cat-R* primers. Finally, the obtained 2.4kb PCR product and the 1kb PCR fragment encompassing the MG_220 DR were joined by SOE-PCR with the *mg220 Up-F* and *mg220 Down-R* primers. The final PCR product (3.4kb) was cloned into a *Eco*RV-digested pBE to create pRrlA:Ch.

Supp. Figure S.36. pRrlA:Ch plasmid map Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase.

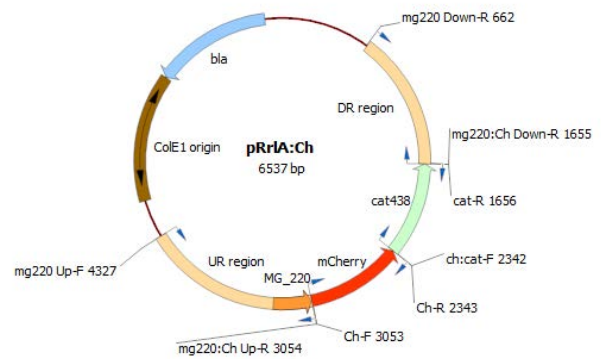Generation of RrlA:Ch strain is described in Figure II.1.

S.3.2.21. Construction of pRrlB:Ch plasmid and generation of RrlB:Ch mutant strain

This suicide plasmid was created to generate a C-terminal fusion of the MGRS02200 polypeptide (*rrlB*) to the fluorescent protein mCherry. A 1kb fragment containing the *rrlB* ORF without the stop codon was PCR-amplified with the *mgrs02200 Up-F* and *mgrs022000:Ch Up-R* primers. In parallel, a 1 kb PCR fragment encompassing the *rrlB* DR was PCR-amplified with the *mgrs02200:Ch Down-F* and *mgrs02200 Down-R* primers. The *cat438* marker (687bp) and the mCherry coding region (711bp) was PCR-amplified and joined together using SOE-PCR as described in S.3.2.5. The resulting recombinant amplicon and the 1kb PCR fragment containing the *rrlB* ORF were also joined by SOE-PCR with the *mgrs02200 Up-F* and *cat-R* primers. Finally, the obtained 2.4kb PCR product and the 1kb PCR fragment encompassing the *rrlB* DR were joined by SOE-PCR with the *mgrs02200 Up-F* and *mgrs02200 Down-R* primers. The final PCR product (3.4kb) was cloned into a *Eco*RV-digested pBE to create pRrlB:Ch
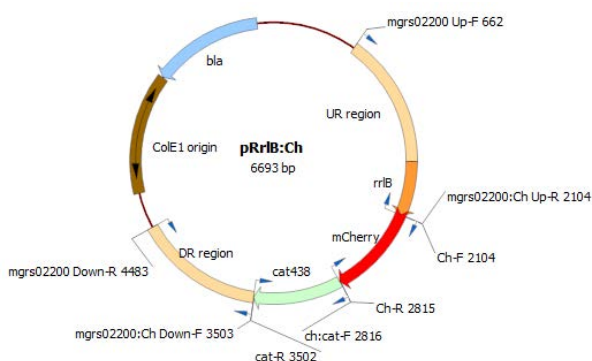
Supp. Figure S.37. RrlB:Ch plasmid map Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase.

Generation of RrlB:Ch strain is described in Figure II.1.

S.3.2.22. Construction of pMTnPac-*rrlA*

This plasmid carries the TnPac*rrlA* MiniTnp, which was used to reintroduce a wild-type copy of the MG_220 allele under the control of the MG_438 promoter in the *M. genitalium* chromosome. The MG_220 allele was PCR-amplified with the *COMmg220-F* and

COMmg220-R primers, digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnPac plasmid.
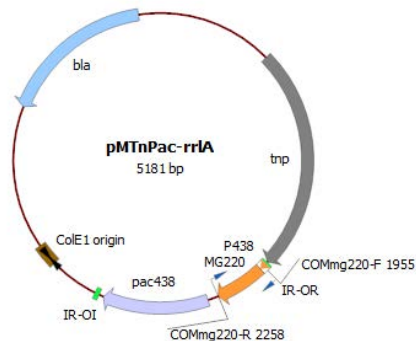
Supp. Figure S.38. pMTnPac*rrlA* plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P438, MG_438 promoter; *pac*, puromycin acetyl transferase marker; tnp, transposase; *bla*, β-lactamase.

S.3.2.23. Construction of pMTnPac-*rrlB*

This plasmid carries the TnPac*rrlB* MiniTnp, which was used to reintroduce a wild-type copy of the MG_RS02200 allele under the control of the MG_438 promoter in the *M. genitalium* chromosome. The MG_RS02200 allele was PCR-amplified with the *COMmgrs02200-F* and *COMmgrs02200-R* primers, digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnPac plasmid.

Supp. Figure S.39. pMTnPac*rrlB* plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P438, MG_438 promoter; *pac*, puromycin acetyl transferase marker; tnp, transposase; *bla*, β-lactamase.
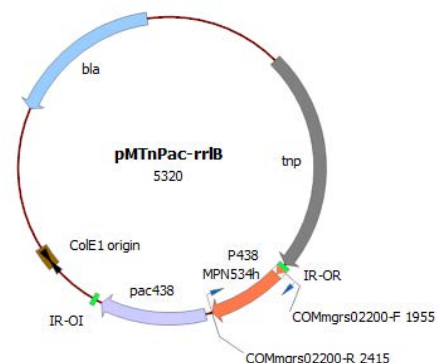
S.3.2.24. Construction of pTnMG_428:YFP

This plasmid carries the TnPacMG_428:YFP MiniTnp, which was used to reintroduce a wild-type copy of the MG_428 allele fused with the eYFP fluorescent protein under the control of the MG_438 promoter in the *M. genitalium* chromosome. The MG_428 allele was PCR-amplified with the *COMmg428-F*(*Apa*I) and *mg428:Ch Up-R** primers, while eYFP was amplified using *Ch-F** and *YFP-R* (*Xho*I) primers from pCAG-eYFP plasmid (Addgene). Then, both fragments were joined in a SOE-PCR using *COMmg428-F*(*Apa*I) and *YFP-R* (*Xho*I), digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnPac plasmid.
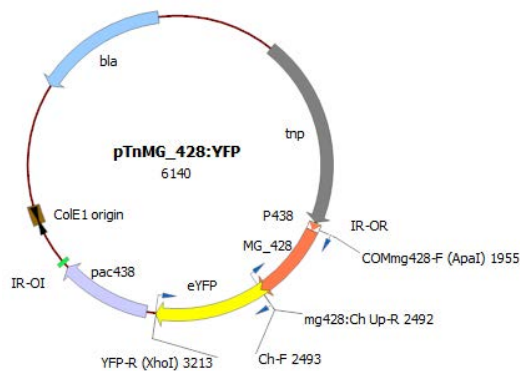
Supp. Figure S.40. pMTnMG_428:YFP plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P438, MG_438 promoter; *pac438*, puromycin acetyl transferase marker under the control of MG_438 promoter; tnp, transposase; *bla*, β-lactamase.

*5' and 3' ends of the mCherry and eYFP coding genes have identical sequences and can be amplified by PCR using the same oligos.
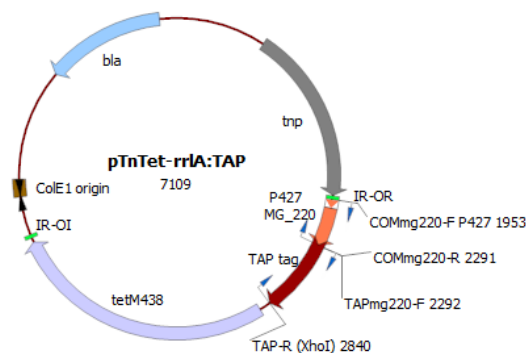
S.3.2.25. Construction of pTnTet-rrlA:TAP

This plasmid carries the TnTet-rrlA:TAP MiniTnp, which was used to overexpress a wild-type copy of the MG_220 allele fused with the TAP tag under the control of the MG_427 promoter in the *M. genitalium* chromosome. The MG_220 allele was PCR-amplified with the *COMmg220-F P427* and *COMmg220-R* primers, while TAP tag was amplified using *TAPmg220-F* and *TAP-R* (*Xho*I) primers from pTnMG328TAP (González-González, unpublished). Then, both fragments were joined in a SOE-PCR using *COMmg220-F P427* and *TAP-R* (*Xho*I), digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnTetM438 plasmid.

Supp. Figure S.41. pTnTet-rrlA:TAP plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P438, MG_438 promoter; *tetM438*, tetracycline resistance marker; tnp, transposase; *bla*, β-lactamase.

The construction of this plasmid was made with the inestimable assist of Ignasi Granero Moya, a former undergraduate student in the lab.

S.3.2.26. Construction of pTnTet-rrlB:TAP

This plasmid carries the TnTet-rrlB:TAP MiniTnp, which was used to overexpress a wild-type copy of the MG_RS02200 allele fused with the TAP tag under the control of the MG_427 promoter in the *M. genitalium* chromosome. The *rrlB* allele was PCR-amplified with the *COMrrlB-F P427* and *COMmgrs02200-R* primers, while TAP tag was amplified using *TAPrrlB-F* and *TAP-R* (*Xho*I) primers from pTnMG328TAP (González-González, unpublished). Then, both fragments were joined in a SOE-PCR using *COMrrlB-F P427* and

TAP-R (*Xho*I), digested with *Apa*I and *Xho*I and ligated into a similarly digested pMTnTetM438 plasmid.
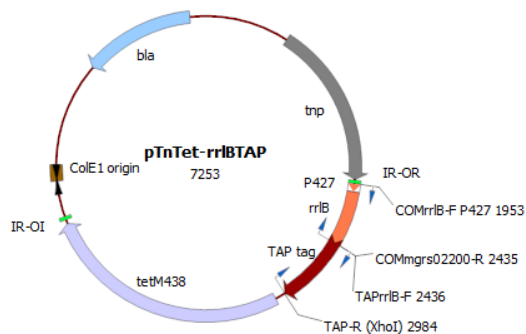

Supp. Figure S.42. pTnTet-rrlB:TAP plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P438, MG_438 promoter; *tetM438*, tetracycline resistance marker; tnp, transposase; *bla*, β-lactamase.

The construction of this plasmid was made with the inestimable assist of Ignasi Granero Moya, a former undergraduate student in the lab.

S.3.2.27. Construction of pMTnPacSpx

This plasmid carries the TnPacSpx MiniTnp, which was used to overexpress a wild-type copy of the MG_127 allele (*spx* gene) under the control of the MG_427 promoter in the *M. genitalium* chromosome. The *spx* allele was PCR-amplified with the *COMmg127-F* and *COMmg127-R* primers. Then, amplicon was digested with *Xba*I and *Not*I and ligated into a similarly digested pMTnPac plasmid.
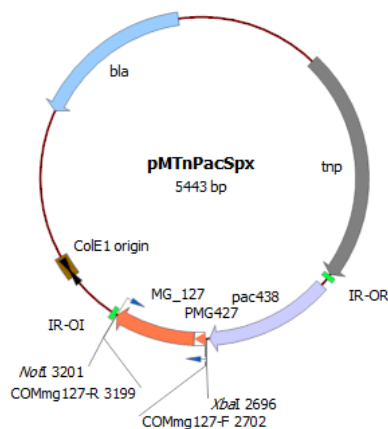
Supp. Figure S.43. pMTnPacSpx plasmid map. Primers used to build this plasmid are highlighted as blue arrows. IR, inverted repeat; P427, MG_427 promoter; *pac4388*, puromycin resistance marker; tnp, transposase; *bla*, β-lactamase.

S.3.2.28. Construction of pMG192.1:Ch and generation of MG192.1:Ch mutant strain

This suicide plasmid was created to generate a C-terminal fusion of the putative MG192.1 peptide to the fluorescent protein mCherry. A 1kb fragment containing the MG_192.1 ORF without the stop codon was PCR-amplified with the *mg192.1 Up-F* and *mg192.1:Ch Up-R* primers. In parallel, a 1 kb PCR fragment encompassing the MG_192.1 DR was PCR-amplified with the *mg192.1:Ch Down-F* and *mg192.1 Down-R* primers. The *cat438* marker (687 bp) and the mCherry coding region (711bp) was PCR-amplified and joined together using SOE-PCR as described in S.3.2.5. The resulting recombinant amplicon and the 1kb PCR fragment containing the MG_192.1 ORF were also joined by SOE-PCR with the

mg192.1 Up-F and *cat-R* primers. Finally, the obtained 2.4kb PCR product and the 1kb PCR fragment encompassing the MG_192.1 DR were joined by SOE-PCR with the *mg192.1 Up-F* and *mg192.1:Ch Down-R* primers. The final PCR product (3.4kb) was cloned into a *Eco*RV-digested pBE to create pMG192.1:Ch.
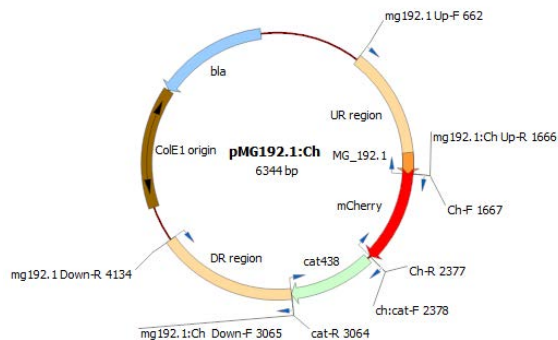
Supp. Figure S.44. pMG192.1:Ch plasmid map. Primers are mapped as blue arrows. UR, homologous upstream region; DR, homologous downstream region; catP438, chloramphenicol resistance cassette with MG_438 promoter; *bla*, β-lactamase.

Generation of MG192.1:Ch strain is described in Figure II.1.

S.3.2.29. Construction of pΔMG_192.1ter

This suicide plasmid was created to generate an *M. genitalium* MG_192.1 null mutant by HR. First, the MG_192.1 UR was PCR-amplified with the *mg192.1 Up-F* and *mg192.1 Up-R* primers. Similarly, the MG_192.1 DR was PCR-amplified with the *mg192.1 Down-F* and *mg192.1 Down-R* primers. Then, the MG_192.1 UR and *tetM438* PCR product (see pΔMG_428, S.3.2.1) were joined by SOE-PCR with the *mg192.1 Up-F* and *Tc-R* primers. Next, the obtained recombinant amplicon and the MG_192.1 DR were joined by SOE-PCR with the *mg192.1 Up-F* and *mg192.1 Down-R* primers. Finally, the resulting PCR product (~4 kb) was cloned into an *Eco*RV-digested pBE plasmid, forming pΔMG_192.1 plasmid. Afterwards, intrinsic terminator sequence of the MG_305 gene of *M. genitalium* (*dnaK*) (Torres-Puig *et al*., unpublished) was amplified with TER305-F and TER305-R, digested with *Bam*HI and cloned in a dephosphorylated, *Bam*HI-digested pΔMG_192.1 plasmid to finally create pΔMG_192.1ter.

Supp. Figure S.45. pΔMG_192.1ter plasmid map. Primers used for cloning are drawn as blue arrows. UR, homologous upstream region; DR, homologous downstream region; *bla*, β-lactamase.
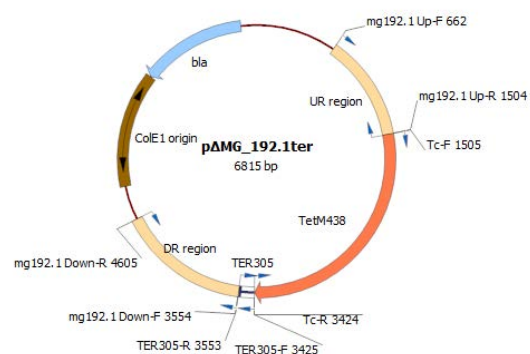
After electroporation with pΔMG_192.1ter, several tetracycline-resistant colonies were picked up and propagated. The intended deletion of *orf192.1* in two selected transformants

was screened for by PCR using chromosomal DNA as template. The primers used (*SCRmg192.1-F* and *SCRmg192.1-R*) were complementary to sequences of the chromosome located immediately outside of the *orf192.1* flanking regions cloned into the pΔMG_192.1ter plasmid (Supp. Figure S.46). The amplification of a 4.1kb band confirmed the replacement of the *orf192.1* gene (168bp) by the tetM438 marker (1954bp) in the two transformants analyzed. Sequencing of the PCR product corroborated the intended deletion of the *orf192.1* gene. As expected, a band of 3.9kb was observed when the chromosomal DNA of the wild-type strain was analyzed.
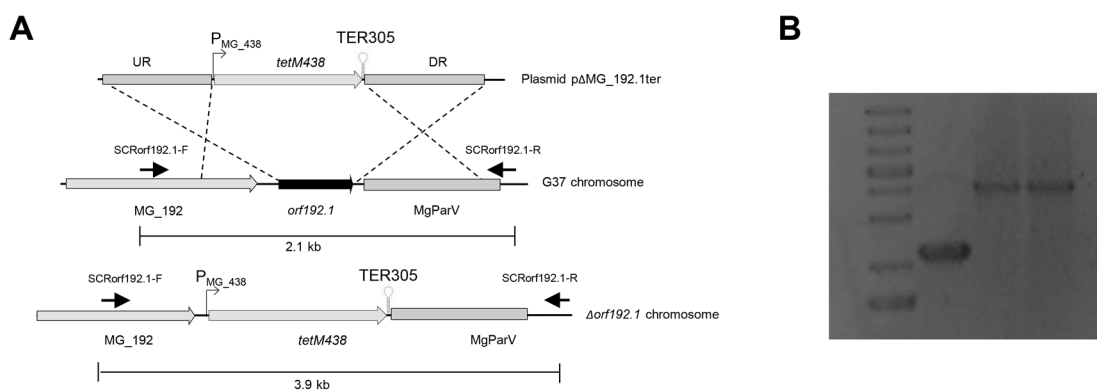
Supp. Figure S.46.Construction of a Δ*orf192.1* deletion mutant by allelic exchange. (A) Schematic representation illustrating the construction of a *M. genitalium* null mutant by homologous recombination. Arrows indicate the approximate location of the primers used for screening. **(B)** Electrophoresis gel demonstrating the replacement of the *orf192.1* locus by the tetM438 marker in different Δ*orf192.1* mutant clones.

S.3.3. Amplification efficiencies of oligonucleotides used in qRT-PCR analysis

To perform this analysis, amplification efficiencies of each primer pair was calculated using standard curves and serial dilutions of *M. genitalium* G37 gDNA.

Supp. Table S.13. Standard curves and amplification efficiencies of the primers used in the qRT-PCR analysis.

Gene	Primers	Amplicon Size	Amplification efficiency	Observed Tm	Trendline	R^2
MG_418 (*rplM*)	RTPCRmg418-F RTPCRmg418-R	200 bp	93,7%	74,5ºC	y=3,482x-25,104	0,998
MG_177 (*rpoA*)	RTPCRmg177-F RTPCRmg177-R	184 bp	92,4%	76,4ºC	y=3,519x-25,302	0,999
MG_430 (*rpmI*)	RTPCRmg430-F RTPCRmg430-R	223 bp	90,6%	76,1ºC	y=3,571x-24,986	0,998
MG_427	RTPCRmg427-F RTPCRmg427-R	243 bp	89,2%	76,1ºC	y=3,612x-25,484	0,999
MG_428	RTPCRmg428-F RTPCRmg428-R	228 bp	86,9%	75,3ºC	y=3,683x-25,591	1,000
MG_RS02200	RTPCRmg_rs02200-F RTPCRmg_rs02200-R	221 bp	87,8%	75,1ºC	y=3,653x-25,436	1,000
MG_358 (*ruvA*)	RTPCRmg358-F RTPCRmg358-R	172 bp	92,4%	73,5ºC	y=3,519x-25,152	0,997
MG_359 (*ruvB*)	RTPCRmg359-F RTPCRmg359-R	177 bp	86,9%	75,9ºC	y=3,683x-25,593	1,000
MG_339 (*recA*)	RTPCRmg339-F RTPCRmg339-R	218 bp	88,0%	76,2ºC	y=3,647x-24,920	0,999
MG_220	RTPCRmg220-F RTPCRmg220-R	204 bp	93,8%	76,0ºC	y=3,480x-24,503	0,999
MG_281	RTPCRmg281-F RTPCRmg281-R	213 bp	93,1%	77,2ºC	y=3,500x-24,551	0,999
MG_390	RTPCRmg390-F RTPCRmg390-R	174 bp	92,5%	74ºC	y=3,516x-24,648	0,999
MG_389	RTPCRmg389-F RTPCRmg389-R	226 bp	90,1%	74,1ºC	y=3,584x-24,583	0,999
MG_414	RTPCRmg414-F RTPCRmg414-R	183 bp	99,1%	76,9ºC	y=3,343x-24,020	0,997
MG_525	RTPCRmg525-F RTPCRmg525-R	236 pb	100,4%	74,3ºC	y=3,312x-23,635	0,995
MG_010	RTPCRmg010-F RTPCRmg010-R	199 bp	86,0%	73,2ºC	y=3,711x-23,337	0,999
MG_191 (*mgpB*)	RTPCRmg191-F RTPCRmg191-R	213 bp	88,1%	75,5ºC	y=3,646x-25,478	0,999
MG_192 (*mgpC*)	RTPCRmg192-F RTPCRmg192-R	205 bp	87,7%	78,1ºC	y=3,658x-25,175	0,999

S.3.4. Genome annotation and genetic nomenclature

In the recent years, *de novo* sequencing projects and gene reannotations are continuously updating bacterial genomes. However, most genes are still referred using the classic nomenclature to simplify and for better understanding among microbiologists. In the case of this doctoral thesis, the old locus tags of genes of *M. genitalium* have been used. The following table (Supp. Table S.14) contains the full list of genes and proteins mentioned in this work, with the corresponding actual locus tag, protein tag and protein annotation.

Supp. Table S.14. List of genes and proteins mentioned in this work using old and current annotations.

Old locus tag[1]	Locus tag[2]	Gene annotation	Protein ID	Protein annotation
-	MG_RS02200	*rrlB*[3]	WP_009885815.1	RrlB[3]
-	-	*orf192.1*[3]	-	ORF192.1[3]
MG_010	MG_RS00050	-	WP_010869288.1	HP
MG_019	MG_RS00095	*dnaJ*	WP_009885921.1	DnaJ
MG_022	MG_RS00110	*rpoE*	WP_009885918.1	RpoE
MG_073	MG_RS00415	*uvrB*	WP_010869318.1	UvrB
MG_091	MG_RS00505	*ssb*	WP_010869327.1	Ssb
MG_097	MG_RS00535	*ung*	WP_014894390.1	UDG
MG_103	MG_RS00565	*whiA*	WP_010869332.1	WhiA
MG_105	MG_RS00580	*disA*	WP_009885663.1	DisA
MG_127	MG_RS00690	*spxA*	WP_010869345.1	SpxA
MG_134	MG_RS00730	*ybaB*	WP_010869347.1	YbaB/EbfC
MG_177	MG_RS00990	*rpoA*	WP_009885862.1	RpoA
MG_191	MG_RS01075	*mgpB*	WP_010869366.1	P140
MG_192	MG_RS01080	*mgpC*	WP_041593683.1	P110
MG_200[4]	MG_RS02825	-	-[4]	TopJ
MG_205	MG_RS01160	*hrcA*	WP_009885743.1	HrcA
MG_206	MG_RS01165	*uvrC*	WP_010869373.1	UvrC
MG_217	MG_RS01275	*p65*	WP_010869376.1	P65
MG_218	MG_RS01280	*hmw2*	WP_010869377.1	HMW2
MG_219	MG_RS01290	*p24*	WP_010869379.1	P24
MG_220	MG_RS01295	*rrlA*[3]	WP_009885758.1	RrlA[3]
MG_221	MG_RS01305	*mraZ*	WP_010869380.1	MraZ
MG_224	MG_RS01320	*ftsZ*	WP_010869382.1	FtsZ
MG_235	MG_RS01380	*nfo*	WP_009885772.1	Nfo
MG_236	MG_RS01385	*fur*	WP_009885773.1	HP
MG_239	MG_RS01400	*lon*	WP_009885776.1	Lon
MG_244	MG_RS01425	*pcrA*	WP_010869390.1	PcrA/UvrD
MG_249	MG_RS01450	*sigA*	WP_009885787.1	σ^A
MG_262	MG_RS01555	-	WP_010869399.1	Pol I-like
MG_281	MG_RS01660	-	WP_010869405.1	Protein M
MG_285	MG_RS01705	-	WP_010869407.1	HP
MG_286	MG_RS01710	-	WP_010869408.1	HP
MG_301	MG_RS01810	*gpdH*	WP_009885879.1	GPDH
MG_305	MG_RS01830	*dnaK*	WP_009885883.1	DnaK
MG_312	MG_RS01865	*hmw1*	WP_010869424.1	HMW1
MG_317	MG_RS01890	*hmw3*	WP_010869426.1	HMW3
MG_318	MG_RS01895	*p32*	WP_010869427.1	P32
MG_332	MG_RS02020	*yebC*	WP_009885985.1	YebC
MG_339	MG_RS02065	*recA*	WP_009885970.1	RecA
MG_343	MG_RS02100	-	WP_010869442.1	HP
MG_352	MG_RS02165	*recU*	WP_009885810.1	RecU
MG_353	MG_RS02170	*himA*	WP_009885811.1	HU protein
MG_355	MG_RS02180	*clpB*	WP_010869446.1	ClpB
MG_358	MG_RS02205	*ruvA*	WP_010869447.1	RuvA
MG_359	MG_RS02210	*ruvB*	WP_009885816.1	RuvB
MG_360	MG_RS02215	*dinB*	WP_010869448.1	Pol IV
MG_389	MG_RS02370	-	WP_009885633.1	HP
MG_390	MG_RS02375	*sunT*	WP_010869463.1	SunT
MG_412	MG_RS02490	*pstS*	WP_010869472.1	PstS
MG_414	MG_RS02495	-	WP_010869473.1	HP
MG_418	MG_RS02510	*rplM*	WP_010869476.1	L13
MG_421	MG_RS02520	*uvrA*	WP_014894509.1	UvrA
MG_427	MG_RS02545	*osmC*	WP_009885605.1	OsmC
MG_428	MG_RS02550	*sig20*[3]	WP_009885604.1	σ^{20} [3]
MG_430	MG_RS02560	*rpmI*	WP_009885603.1	RpmI
MG_438	MG_RS02595	*hsdS*	WP_009885596.1	HsdS
MG_491	MG_RS01285	*p41*	WP_010869378.1	P41

MG_498	MG_RS01560	*mutM*	WP_010869400.1	FPG
MG_525	MG_RS02500	-	WP_010869474.1	HP

[1] Old locus tag used in chromosome annotation L43967 (2014)

[2] Locus tag used in chromosome annotation NC_000908.2 (2017)

[3] Annotated in this work

[4] MG_200 is annotated as a pseudogene in NC_000908.2, probably due to an error in contig assembly.

BIBLIOGRAPHY

Achberger, E. C. and Whiteley, H. R. (1981). The role of the delta peptide of the *Bacillus subtilis* RNA polymerase in promoter selection. *J. Biol. Chem.*, 256(14), pp. 7424–32.

Aldecoa-Otalora, E., Langdon, W. B., Cunningham, P. and Arno, M. J. (2009). Unexpected presence of mycoplasma probes on human microarrays. *Biotechniques*, 47(6), pp. 1013–1016.

Algire, M. A., Lartigue, C., Thomas, D. W., Assad-Garcia, N., Glass, J. I. and Merryman, C. (2009). New selectable marker for manipulating the simple genomes of Mycoplasma species. *Antimicrob. Agents Chemother.*, 53(10), pp. 4429–4432. doi: 10.1128/AAC.00388-09.

Allam, A. B., Reyes, L., Assad-Garcia, N., Glass, J. I. and Brown, M. B. (2010). Enhancement of targeted homologous recombination in *Mycoplasma mycoides* subsp. *capri* by inclusion of heterologous *recA*. *Appl. Environ. Microbiol.*, 76(20), pp. 6951–6954. doi: 10.1128/AEM.00056-10.

Allemand, F., Mathy, N., Brechemier-Baey, D. and Condon, C. (2005). The 5S rRNA maturase, ribonuclease M5, is a Toprim domain family member. *Nucleic Acids Res.*, 33(13), pp. 4368–4376. doi: 10.1093/nar/gki752.

Alonso, J. C., Luder, G. and Tailort, R. H. (1991). Characterization of *Bacillus subtilis* recombinational pathways. *J. Bacteriol.*, 173(13), pp. 3977–3980.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.*, 11, p. R106. doi: 10.1186/gb-2010-11-10-r106.

Andreev, J., Borovsky, Z., Rosenshine, I. and Rottem, S. (1995). Invasion of HeLa cells by *Mycoplasma penetrans* and the induction of tyrosine phosphorylation of a 145-kDa host cell protein. *FEMS Microbiol. Lett.*, 132(3), pp. 189–194. doi: 10.1016/0378-1097(95)00309-S.

Antunes, N. T., Assunção, P., Poveda, J. B. and Tavío, M. M. (2015). Mechanisms involved in quinolone resistance in *Mycoplasma mycoides* subsp. *capri*. *Vet. J.*, 204(3), pp. 327–332. doi: 10.1016/j.tvjl.2015.04.018.

Aravind, L. and Iyer, L. M. (2012). The HARE-HTH and associated domains. *Cell Cycle*, 11(1), pp. 119–131. doi: 10.4161/cc.11.1.18475.

Arfi, Y., Minder, L., Di Primo, C., Le Roy, A., Ebel, C., Coquet, L., Claverol, S., Vashee, S., Jores, J., Blanchard, A. and Sirand-Pugnet, P. (2016). MIB–MIP is a mycoplasma system that captures and cleaves immunoglobulin G. *Proc. Natl. Acad. Sci.*, 113(19), pp. 5406–5411. doi: 10.1073/pnas.1600546113.

Ayora, S., Carrasco, B., Cárdenas, P. P., César, C. E., Cañas, C., Yadav, T., Marchisone, C. and Alonso, J. C. (2011). Double-strand break repair in bacteria: a view from *Bacillus subtilis*. *FEMS Microbiol. Rev.*, 35(6), pp. 1055–1081. doi: 10.1111/j.1574-6976.2011.00272.x.

Ayora, S., Carrasco, B., Doncel-Perez, E., Doncel, E., Lurz, R. and Alonso, J. C. (2004). *Bacillus subtilis* RecU protein cleaves Holliday junctions and anneals single-stranded DNA. *Proc. Natl. Acad. Sci. U. S. A.*, 101(2), pp. 452–7. doi: 10.1073/pnas.2533829100.

Balish, M. F. (2014). *Mycoplasma pneumoniae*, an underutilized model for bacterial cell biology. *J. Bacteriol.*, 196(21), pp. 3675–3682. doi: 10.1128/JB.01865-14.

Balish, M. F., Santurri, R. T., Ricci, A. M., Lee, K. K. and Krause, D. C. (2003). Localization of *Mycoplasma pneumoniae* cytadherence-associated protein HMW2 by fusion with green fluorescent protein: implications for attachment organelle structure. *Mol. Microbiol.*, 47(1), pp. 49–60.

Barile, M. F., Chandler, D. K. F., Yoshida, H., Grabowski, M. W. and Razin, S. (1988). Hamster challenge potency asay for evaluation of *Mycoplasma pneumoniae* vaccines. *Infect. Immun.*, 56(9), pp. 2450–2457.

Baseman, J. B., Banai, M. and Kahane, I. (1982). Sialic acid residues mediate *Mycoplasma pneumoniae* attachment to human and sheep erythrocytes. *Infect. Immun.*, 38(1), pp. 389–91.

Baseman, J. and Tully, J. G. (1997). Mycoplasmas: sophisticated, reemerging, and burdened by their notoriety. *Emerg. Infect. Dis.*, 3(1), pp. 21–32. doi: 10.3201/eid0301.970103.

Basile, G., Aker, M. and Mortimer, R. K. (1992). Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Mol. Cell. Biol.*, 12(7), pp. 3235–46. doi: 10.1128/MCB.12.7.3235.

Becker, A., Kannan, T. R., Taylor, A. B., Pakhomova, O. N., Zhang, Y., Somarajan, S. R., Galaleldeen, A., Holloway, S. P., Baseman, J. B. and Hart, P. J. (2015). Structure of CARDS toxin, a unique ADP-ribosylating and vacuolating cytotoxin from *Mycoplasma pneumoniae*. *Proc. Natl. Acad. Sci.*, 112(16), pp. 5165–5170. doi: 10.1073/pnas.1420308112.

Becker, G., Klauck, E. and Hengge-Aronis, R. (1999). Regulation of RpoS proteolysis in *Escherichia coli*: the response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. *Proc. Natl. Acad. Sci. U. S. A.*, 96(11), pp. 6439–44.

Beeton, M. L. and Spiller, O. B. (2017). Antibiotic resistance among *Ureaplasma* spp. isolates: cause for concern? *J. Antimicrob. Chemother.*, 72(2), pp. 330–337. doi: 10.1093/jac/dkw425.

Bergemann, A. D., Whitley, J. C. and Finch, L. R. (1990). Taxonomic significance of differences in DNA methylation within the '*Mycoplasma mycoides* cluster' detected with restriction endonucleases MboI and DpnI. *Lett. Appl. Microbiol.*, 11(1), pp. 48–51. doi: 10.1111/j.1472-765X.1990.tb00134.x.

Bhugra, B. and Dybvig, K. (1993). Identification and characterization of IS1138, a transposable element from *Mycoplasma pulmonis* that belongs to the IS3 family. *Mol. Microbiol.*, 7(4), pp. 577–84.

Biernat-Sudolska, M., Szostek, S., Rojek-Zakrzewska, D., Klimek, M. and Kosz-Vnenchak, M. (2011). Concomitant infections with human papillomavirus and various mycoplasma and ureaplsasma species in women with abnormal cervical cytology. *Adv. Med. Sci.*, 56(2), pp. 299–303. doi: 10.2478/v10039-011-0028-9.

Biondo, A. W., Dos Santos, A. P., Guimarães, A. M. S., Vieira, R. F. da C., Vidotto, O., Macieira, D. de B., Almosny, N. R. P., Molento, M. B., Timenetsky, J., de Morais, H. A., González, F. H. D. and Messick, J. B. (2009). A review of the occurrence of Hemoplasmas (hemotrophic Mycoplasmas) in Brazil. *Rev. Bras. Parasitol. Vet.*, 18(3), pp. 1–7.

Blanchard, A. and Montagnier, L. (1994). Aids-Associated Mycoplasmas. *Annu. Rev. Microbiol.*, 48(1), pp. 687–712. doi: 10.1146/annurev.mi.48.100194.003351.

Blaylock, M. W., Musatovova, O., Baseman, J. G. and Baseman, J. B. (2004). Determination of infectious load of *Mycoplasma genitalium* in clinical samples of human vaginal cells. *J. Clin. Microbiol.*, 42(2), pp. 746–52.

Boonmee, A., Ruppert, T. and Herrmann, R. (2009). The gene mpn310 (*hmw2*) from *Mycoplasma pneumoniae* encodes two proteins, HMW2 and HMW2-s, which differ in size but use the same reading frame. *FEMS Microbiol. Lett.*, 290(2), pp. 174–81. doi: 10.1111/j.1574-6968.2008.01422.x.

Bornberg-Bauer, E. and Weiner, J. (2002). A putative transcription factor inducing mobility in *Mycoplasma pneumoniae*. *Microbiology*, 148(Pt 12), pp. 3764–5. doi: 10.1099/00221287-148-12-3764.

Bose, S., Segovia, J. A., Somarajan, S. R., Chang, T.-H., Kannan, T. R. and Baseman, J. B. (2014). ADP-ribosylation of NLRP3 by Mycoplasma pneumoniae CARDS toxin regulates inflammasome activity. *MBio*, 5(6), pp. e02186-14. doi: 10.1128/mBio.02186-14.

Bougdour, A., Lelong, C. and Geiselmann, J. (2004). Crl, a low temperature-induced protein in *Escherichia coli* that binds directly to the stationary phase sigma subunit of RNA polymerase. *J. Biol. Chem.*, 279(19), pp. 19540–50. doi: 10.1074/jbc.M314145200.

Bradbury, J. M. (1998). Recovery of Mycoplasmas from Birds. in *Mycoplasma Protoc.*, pp. 45–51. doi: 10.1385/0-89603-525-5:45.

Bradshaw, C. S., Jensen, J. S., Tabrizi, S. N., Read, T. R. H., Garland, S. M., Hopkins, C. A., Moss, L. M. and Fairley, C. K. (2006). Azithromycin failure in *Mycoplasma genitalium* urethritis. *Emerg. Infect. Dis.*, 12(7), pp. 1149–1152. doi: 10.3201/eid1207.051558.

Breton, M., Sagne, E., Duret, S., Beven, L., Citti, C. and Renaudin, J. (2010). First report of a tetracycline-inducible gene expression system for mollicutes. *Microbiology*, 156(1), pp. 198–205. doi: 10.1099/mic.0.034074-0.

Brocchi, M., Vasconcelos, A. T. R. de and Zaha, A. (2007). Restriction-modification systems in *Mycoplasma spp*. *Genet. Mol. Biol.*, 30(1), pp. 236–244. doi: 10.1590/S1415-47572007000200011.

Burgos, R., Pich, O. Q., Ferrer-Navarro, M., Baseman, J. B., Querol, E. and Pinol, J. (2006). *Mycoplasma genitalium* P140 and P110 cytadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development. *J. Bacteriol.*, 188(24), pp. 8627–8637. doi: 10.1128/JB.00978-06.

Burgos, R., Pich, O. Q., Querol, E. and Piñol, J. (2007). Functional analysis of the *Mycoplasma genitalium* MG312 protein reveals a specific requirement of the MG312 N-terminal domain for gliding motility. *J. Bacteriol.*, 189(19), pp. 7014–23. doi: 10.1128/JB.00975-07.

Burgos, R., Pich, O. Q., Querol, E. and Piñol, J. (2008). Deletion of the *Mycoplasma genitalium* MG_217 gene modifies cell gliding behaviour by altering terminal organelle curvature. *Mol. Microbiol.*, 69(4), pp. 1029–1040. doi: 10.1111/j.1365-2958.2008.06343.x.

Burgos, R. and Totten, P. A. (2014a). Characterization of the operon encoding the holliday junction helicase RuvAB from *Mycoplasma genitalium* and its role in *mgpB* and *mgpC* gene variation. *J. Bacteriol.*, 196(8), pp. 1608–1618. doi: 10.1128/JB.01385-13.

Burgos, R. and Totten, P. A. (2014b). MG428 is a novel positive regulator of recombination that triggers mgpB and mgpC gene variation in *Mycoplasma genitalium*. *Mol. Microbiol.*, 94(2), pp. 290–306. doi: 10.1111/mmi.12760.

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. *Mol. Microbiol.*, 85(4), pp. 669–683. doi: 10.1111/j.1365-2958.2012.08130.x.

Calcutt, M. J., Lewis, M. S. and Wise, K. S. (2002). Molecular genetic analysis of ICEF, an integrative conjugal element that is present as a repetitive sequence in the chromosome of *Mycoplasma fermentans* PG18. *J. Bacteriol.*, 184(24), pp. 6929–41. doi: 10.1128/jb.184.24.6929-6941.2002.

Calisto, B. M., Broto, A., Martinelli, L., Querol, E., Piñol, J. and Fita, I. (2012). The EAGR box structure: a motif involved in Mycoplasma motility. *Mol. Microbiol.*, 86(2), pp. 382–393. doi: 10.1111/j.1365-2958.2012.08200.x.

Calisto, B. M., Pich, O. Q., Piñol, J., Fita, I., Querol, E. and Carpena, X. (2005). Crystal structure of a putative type I restriction-modification S subunit from *Mycoplasma genitalium*. *J. Mol. Biol.*, 351(4), pp. 749–762. doi: 10.1016/j.jmb.2005.06.050.

Carrasco, B., Ayora, S., Lurz, R. and Alonso, J. C. (2005). *Bacillus subtilis* RecU Holliday-junction resolvase modulates RecA activities. *Nucleic Acids Res.*, 33(12), pp. 3942–52. doi: 10.1093/nar/gki713.

Carvalho, F. M., Fonseca, M. M., Batistuzzo De Medeiros, S., Scortecci, K. C., Blaha, C. A. and Agnez-Lima, L. F. (2005). DNA repair in reduced genome: the Mycoplasma model. doi: 10.1016/j.gene.2005.06.012.

Cecchini, K. R., Gorton, T. S. and Geary, S. J. (2007). Transcriptional responses of *Mycoplasma gallisepticum* strain R in association with eukaryotic cells. *J. Bacteriol.*, 189(16), pp. 5803–5807. doi: 10.1128/JB.00667-07.

Cherry, J. D. and Taylor-Robinson, D. (1970). Peroxide production by Mycoplasmas in chicken tracheal organ cultures. *Nature*, 228(5276), pp. 1099–100.

Chopra-Dewasthaly, R., Baumgartner, M., Gamper, E., Innerebner, C., Zimmermann, M., Schilcher, F., Tichy, A., Winter, P., Jechlinger, W., Rosengarten, R. and Spergser, J. (2012). Role of Vpma phase variation in Mycoplasma agalactiae pathogenesis. *FEMS Immunol. Med. Microbiol.*, 66(3), pp. 307–322. doi: 10.1111/j.1574-695X.2012.01010.x.

Chopra-Dewasthaly, R., Citti, C., Glew, M. D., Zimmermann, M., Rosengarten, R. and Jechlinger, W. (2008). Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation. *Mol. Microbiol.*, 67(6), pp. 1196–1210. doi: 10.1111/j.1365-2958.2007.06103.x.

Chopra-Dewasthaly, R., Marenda, M., Rosengarten, R., Jechlinger, W. and Citti, C. (2005). Construction of the first shuttle vectors for gene cloning and homologous recombination in *Mycoplasma agalactiae*. *FEMS Microbiol. Lett.*, 253(1), pp. 89–94. doi: 10.1016/j.femsle.2005.09.021.

Chopra, I. and Roberts, M. (2001). Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol. Mol. Biol. Rev.*, 65(2), pp. 232–60. doi: 10.1128/MMBR.65.2.232-260.2001.

Chowdhury, M. I., Munakata, T., Koyanagi, Y., Arai, S. and Yamamoto, N. (1994). Mycoplasma stimulates HIV-1 expression from acutely- and dormantly-infected promonocyte/monoblastoid cell lines. *Arch. Virol.*, 139(3–4), pp. 431–8.

Citti, C., Kim, M. F. and Wise, K. S. (1997). Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth- inhibiting host antibodies. *Infect. Immun.*, 65(5), pp. 1773–1785.

Citti, C., Nouvel, L.-X. and Baranowski, E. (2010). Phase and antigenic variation in Mycoplasmas. *Future Microbiol.*, 5(7), pp. 1073–1085. doi: 10.2217/fmb.10.71.

Citti, C. and Wise, K. S. (1995). *Mycoplasma hyorhinis vlp* gene transcription: critical role in phase variation and expression of surface lipoproteins. *Mol. Microbiol.*, 18(4), pp. 649–60.

Clarivet, B., Picot, E., Marchandin, H., Tribout, V., Rachedi, N., Schwartzentruber, E., Ledésert, B., Dereure, O., Guillot, B. and Picot, M.-C. (2014). Prevalence of *Chlamydia trachomatis*, *Neisseria gonorrhoeae* and *Mycoplasma genitalium* in asymptomatic patients under 30 years of age screened in a French sexually transmitted infections clinic. *Eur. J. Dermatol.*, 24(5), pp. 611–6. doi: 10.1684/ejd.2014.2413.

Cloward, J. M. and Krause, D. C. (2011). Loss of co-chaperone TopJ impacts adhesin P1 presentation and terminal organelle maturation in *Mycoplasma pneumoniae*. *Mol. Microbiol.*, 81(2), pp. 528–539. doi: 10.1111/j.1365-2958.2011.07712.x.

Cole, B. C., Ward, J. R. and Martin, C. H. (1968). Hemolysin and peroxide activity of Mycoplasma species. *J. Bacteriol.*, 95(6), pp. 2022–30.

Condon, C., Brechemier-Baey, D., Beltchev, B., Grunberg-Manago, M. and Putzer, H. (2001). Identification of the gene encoding the 5S ribosomal RNA maturase in *Bacillus subtilis*: mature 5S rRNA is dispensable for ribosome function. *RNA*, 7(2), pp. 242–53.

Cordova, C. M. M., Lartigue, C., Sirand-Pugnet, P., Renaudin, J., 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. *J. Bacteriol.*, 184(19), pp. 5426–35.

Cox, M. M. (1991). The RecA protein as a recombinational repair system. *Mol. Microbiol.*, 5(6), pp. 1295–1299. doi: 10.1111/j.1365-2958.1991.tb00775.x.

Cox, M. M. (2007). Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.*, 42(1), pp. 41–63. doi: 10.1080/10409230701260258.

Coyle, M., Hu, J. and Gartner, Z. (2016). Mysteries in a minimal genome. *ACS Cent. Sci.*, 2(5), pp. 274–7. doi: 10.1021/acscentsci.6b00110.

Dahl, J. (1993). The role of cholesterol in Mycoplasma membranes. in, pp. 167–188. doi: 10.1007/978-1-4615-2924-8_5.

Daley, G., Russell, D., Tabrizi, S. and McBride, J. (2014). *Mycoplasma genitalium*: a review. *Int. J. STD AIDS*, 25(7), pp. 475–487. doi: 10.1177/0956462413515196.

DaMassa, A. J., Wakenell, P. S. and Brooks, D. L. (1992). Mycoplasmas of goats and sheep. *J. Vet. Diagnostic Investig.*, 4(1), pp. 101–113. doi: 10.1177/104063879200400126.

Das, K., De la Garza, G., Siwak, E. B., Scofield, V. L. and Dhandayuthapani, S. (2014). Mycoplasma genitalium promotes epithelial crossing and peripheral blood mononuclear cell infection by HIV-1. *Int. J. Infect. Dis.*, 23, pp. 31–38. doi: 10.1016/j.ijid.2013.11.022.

Dhandayuthapani, S., Blaylock, M. W., Bebear, C. M., Rasmussen, W. G. and Baseman, J. B. (2001). Peptide methionine sulfoxide reductase (MsrA) is a virulence determinant in *Mycoplasma genitalium*. *J. Bacteriol.*, 183(19), pp. 5645–5650. doi: 10.1128/JB.183.19.5645-5650.2001.

Dhandayuthapani, S., Rasmussen, W. G. and Baseman, J. B. (1998). Identification of mycoplasmal promoters in *Escherichia coli* using a promoter probe vector with Green Fluorescent Protein as reporter system. *Gene*, 215(1), pp. 213–222. doi: 10.1016/S0378-1119(98)00260-1.

Dhandayuthapani, S., Rasmussen, W. G. and Baseman, J. B. (1999). Disruption of gene mg218 of *Mycoplasma genitalium* through homologous recombination leads to an adherence-deficient phenotype. *Proc. Natl. Acad. Sci. U. S. A.*, 96(9), pp. 5227–32. doi: 10.1073/pnas.96.9.5227.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloe, D., Le Gall, C., Schaeffer, B., Le Crom, S., Guedj, M. and Jaffrezic, F. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief. Bioinform.*, 14(6), pp. 671–683. doi: 10.1093/bib/bbs046.

Dillingham, M. S., Soultanas, P., Wiley, P., Webb, M. R. and Wigley, D. B. (2001). Defining the roles of individual residues in the single-stranded DNA binding site of PcrA helicase. *Proc. Natl. Acad. Sci. U. S. A.*, 98(15), pp. 8381–7. doi: 10.1073/pnas.131009598.

Dirix, G., Monsieurs, P., Marchal, K., Vanderleyden, J. and Michiels, J. (2004). Screening genomes of Gram-positive bacteria for double-glycine-motif-containing peptides. *Microbiology*, 150(5), pp. 1121–1126. doi: 10.1099/mic.0.27040-0.

Djordjevic, S. R., Forbes, W. A., Forbes-Faulkner, J., Kuhnert, P., Hum, S., Hornitzky, M. A., Vilei, E. M. and Frey, J. (2001). Genetic diversity among Mycoplasma species bovine group 7: clonal isolates from an outbreak of polyarthritis, mastitis, and abortion in dairy cattle. *Electrophoresis*, 22(16), pp. 3551–61.

Dordet-Frisoni, E., Sagné, E., Baranowski, E., Breton, M., Nouvel, L. X., Blanchard, A., Marenda, M. S., Tardy, F., Sirand-Pugnet, P. and Citti, C. (2014). Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *MBio*, 5(6). doi: 10.1128/mBio.01958-14.

Dordet Frisoni, E., Marenda, M. S., Sagné, E., Nouvel, L. X., Guérillot, R., Glaser, P., Blanchard, A., Tardy, F., Sirand-Pugnet, P., Baranowski, E. and Citti, C. (2013). ICEA of *Mycoplasma agalactiae* : a new family of self-transmissible integrative elements that confers conjugative properties to the recipient strain. *Mol. Microbiol.*, 89(6), pp. 1226–1239. doi: 10.1111/mmi.12341.

Dowding, J. E. (1977). Mechanisms of gentamicin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.*, 11(1), pp. 47–50.

Drexler, H. G. and Uphoff, C. C. (2002). Mycoplasma contamination of cell cultures: incidence, sources, effects, detection, elimination, prevention. *Cytotechnology*, 39(2), pp. 75–90. doi: 10.1023/A:1022913015916.

Dronkert, M. L. G. and Kanaar, R. (2001). Repair of DNA interstrand cross-links. *Mutat. Res.*, 486, pp. 217–247.

Dubey, G. P. and Ben-Yehuda, S. (2011). Intercellular nanotubes mediate bacterial communication. *Cell*, 144(4), pp. 590–600. doi: 10.1016/j.cell.2011.01.015.

Dubey, G. P., Malli Mohan, G. B., Dubrovsky, A., Amen, T., Tsipshtein, S., Rouvinski, A., Rosenberg, A., Kaganovich, D., Sherman, E., Medalia, O. and Ben-Yehuda, S. (2016). Architecture and characteristics of bacterial nanotubes. *Dev. Cell*, 36(4), pp. 453–461. doi: 10.1016/j.devcel.2016.01.013.

Dunny, G. M. and Leonard, B. A. B. (1997). Cell-cell communication in Gram-positive bacteria. *Annu. Rev. Microbiol.*, 51(1), pp. 527–564. doi: 10.1146/annurev.micro.51.1.527.

Dyballa, N. and Metzger, S. (2009). Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *J. Vis. Exp.*, (30), pp. 2–5. doi: 10.3791/1431.

Dybvig, K., Cao, Z., French, C. T. and Yu, H. (2007). Evidence for type III restriction and modification systems in *Mycoplasma pulmonis*. *J. Bacteriol.*, 189(6), pp. 2197–202. doi: 10.1128/JB.01669-06.

Dybvig, K. and Cassell, G. H. (1987). Transposition of gram-positive transposon Tn916 in *Acholeplasma laidlawii* and *Mycoplasma pulmonis*. *Science*, 235(4794), pp. 1392–4.

Dybvig, K., Sitaraman, R. and French, C. T. (1998). A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. *Proc. Natl. Acad. Sci. U. S. A.*, 95(23), pp. 13923–8.

Dybvig, K. and Voelker, L. L. (1996). Molecular biology of Mycoplasmas. *Annu. Rev. Microbiol.*, 50(1), pp. 25–57. doi: 10.1146/annurev.micro.50.1.25.

Edlund, M., Blaxhult, A. and Bratt, G. (2012). The spread of *Mycoplasma genitalium* among men who have sex with men. *Int. J. STD AIDS*, 23(6), pp. 455–456. doi: 10.1258/ijsa.2009.009411.

Edward, D. G. and Freundt, E. A. (1956). The classification and nomenclature of organisms of the pleuropneumonia group. *J. Gen. Microbiol.*, 14(1), pp. 197–207. doi: 10.1099/00221287-14-1-197.

Edward, D. G. and Freundt, E. A. (1967). Proposal for Mollicutes as name of the class established for the order Mycoplasmatales. *Int. J. Syst. Bacteriol.*, 17, pp. 267–268.

Eisen, J. A., Heidelberg, J. F., White, O. and Salzberg, S. L. (2000). Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol.*, 1(6), p. research0011.1. doi: 10.1186/gb-2000-1-6-research0011.

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*. *J. Bacteriol.*, 196(11), pp. 2053–66. doi: 10.1128/JB.01370-13.

Estevao, S., Sluijter, M., Hartwig, N. G., van Rossum, A. M. C. and Vink, C. (2011). Functional characterization of the RuvB hHomologs from Mycoplasma pneumoniae and Mycoplasma genitalium. *J. Bacteriol.*, 193(23), pp. 6425–6435. doi: 10.1128/JB.06003-11.

van Etten, W. J. and Janssen, G. R. (1998). An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in Escherichia coli. *Mol. Microbiol.*, 27(5), pp. 987–1001.

Ewing, B. and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.*, 8(3), pp. 186–194. doi: 10.1101/gr.8.3.186.

Ewing, B., Hillier, L., Wendl, M. C. and Green, P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.*, 8(3), pp. 175–185. doi: 10.1101/gr.8.3.175.

Feng, S.-H., Tsai, S., Rodriguez, J. and Lo, S.-C. (1999). Mycoplasmal infections prevent apoptosis and induce malignant transformation of interleukin-3-dependent 32D hematopoietic cells. *Mol. Cell. Biol.*, 19(12), pp. 7995–8002.

Fisseha, M., Göhlmann, H. W., Herrmann, R. and Krause, D. C. (1999). Identification and complementation of frameshift mutations associated with loss of cytadherence in *Mycoplasma pneumoniae*. *J. Bacteriol.*, 181(14), pp. 4404–10.

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, pp. 59–65. doi: 10.1016/j.biochi.2016.02.016.

Flitman-Tene, R., Mudahi-Orenstein, S., Levisohn, S. and Yogev, D. (2003). Variable lipoprotein genes of *Mycoplasma agalactiae* are activated in vivo by promoter addition via site-specific DNA inversions. *Infect. Immun.*, 71(7), pp. 3821–30.

Franke, A. E. and Clewell, D. B. (1981). Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of 'conjugal' transfer in the absence of a conjugative plasmid. *J. Bacteriol.*, 145(1), pp. 494–502.

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. *et al.* (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science*, 270(5235), pp. 397–403. doi: 10.1126/science.270.5235.397.

Friedman, D. I., Schauer, A. T., Baumann, M. R., Baron, L. S. and Adhya, S. L. (1981). Evidence that ribosomal protein S10 participates in control of transcription termination. *Proc. Natl. Acad. Sci. U. S. A.*, 78(2), pp. 1115–8.

Fritsch, T. E., Siqueira, F. M. and Schrank, I. S. (2015). Intrinsic terminators in *Mycoplasma hyopneumoniae* transcription. *BMC Genomics*, 16(1), p. 273. doi: 10.1186/s12864-015-1468-6.

Gaal, T., Mandel, M. J., Silhavy, T. J. and Gourse, R. L. (2006). Crl facilitates RNA polymerase holoenzyme

formation. *J. Bacteriol.*, 188(22), pp. 7966–7970. doi: 10.1128/JB.01266-06.

García-Morales, L., González-González, L., Costa, M., Querol, E. and Piñol, J. (2014). Quantitative assessment of Mycoplasma hemadsorption activity by flow cytometry. *PLoS One*. Edited by L. G. Filion, 9(1), p. e87500. doi: 10.1371/journal.pone.0087500.

García-Morales, L., González-González, L., Querol, E. and Piñol, J. (2016). A minimized motile machinery for *Mycoplasma genitalium*. *Mol. Microbiol.*, 100(1), pp. 125–138. doi: 10.1111/mmi.13305.

Gautheret, D. and Lambert, A. (2001). Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J. Mol. Biol.*, 313(5), pp. 1003–11. doi: 10.1006/jmbi.2001.5102.

Gebhard, S. (2012). ABC transporters of antimicrobial peptides in Firmicutes bacteria - phylogeny, function and regulation. *Mol. Microbiol.*, 86(6), pp. 1295–1317. doi: 10.1111/mmi.12078.

Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison, C. A. and Smith, H. O. (2008). Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, 319(5867), pp. 1215–1220. doi: 10.1126/science.1151721.

Gibson, D., Glass, J., Lartigue, C., Noskov, V., Chuang, R.-Y., Algire, M., Benders, G., Montague, M., Ma, L., Moodie, M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E., Young, L., Qi, Z., Segall-Shapiro, T. *et al.* (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987), pp. 52–56. doi: 10.1126/science.1190719.

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. *Proc. Natl. Acad. Sci. U. S. A.*, 103(2), pp. 425–430. doi: 10.1073/pnas.0510013103.

Glew, M. D., Baseggio, N., Markham, P. F., Browning, G. F. and Walker, I. D. (1998). Expression of the pMGA genes of *Mycoplasma gallisepticum* is controlled by variation in the GAA trinucleotide repeat lengths within the 5' noncoding regions. *Infect. Immun.*, 66(12), pp. 5833–5841.

Glew, M. D., Marenda, M., Rosengarten, R. and Citti, C. (2002). Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the vpma multigene locus. *J. Bacteriol.*, 184(21), pp. 5987–98.

Gorbachev, A. Y., Fisunov, G. Y., Izraelson, M., Evsyutina, D. V, Mazin, P. V, Alexeev, D. G., Pobeguts, O. V, Gorshkova, T. N., Kovalchuk, S. I., Kamashev, D. E. and Govorun, V. M. (2013). DNA repair in *Mycoplasma gallisepticum*. *BMC Genomics*, 14. doi: 10.1186/1471-2164-14-726.

Gottesman, T., Yossepowitch, O., Samra, Z., Rosenberg, S. and Dan, M. (2017). Prevalence of *Mycoplasma genitalium* in men with urethritis and in high risk asymptomatic males in Tel Aviv: a prospective study. *Int. J. STD AIDS*, 28(2), pp. 127–132. doi: 10.1177/0956462416630675.

Gratia, A. (1925). Sur un rémarquable example d'antagonisme entre deux souches de colibacille. *Biochimie*, pp. 1040–1041. doi: 10.1128/JB.185.3.897.

Grau, O., Slizewicz, B., Tuppin, P., Launay, V., Bourgeois, E., Sagot, N., Moynier, M., Lafeuillade, A., Bachelez, H. and Clauvel, J. P. (1995). Association of *Mycoplasma penetrans* with human immunodeficiency virus infection. *J. Infect. Dis.*, 172(3), pp. 672–81.

Gray, T. A., Clark, R. R., Boucher, N., Lapierre, P., Smith, C. and Derbyshire, K. M. (2016). Intercellular communication and conjugation are mediated by ESX secretion systems in mycobacteria. *Science*, 354(6310).

Gray, T. A., Krywy, J. A., Harold, J., Palumbo, M. J. and Derbyshire, K. M. (2013). Distributive conjugal transfer in Mycobacteria generates progeny with meiotic-like genome-wide mosaicism, allowing mapping of a mating identity locus. *PLoS Biol.* Edited by J. A. Eisen, 11(7), p. e1001602. doi: 10.1371/journal.pbio.1001602.

Grigorova, I. L., Phleger, N. J., Mutalik, V. K. and Gross, C. A. (2006). Insights into transcriptional regulation and competition from an equilibrium model of RNA polymerase binding to DNA. *Proc. Natl. Acad. Sci.*, 103(14), pp. 5332–5337. doi: 10.1073/pnas.0600828103.

Grossman, A. D., Erickson, J. W. and Gross, C. A. (1984). The *htpR* gene product of *E. coli* is a sigma factor

for heat-shock promoters. *Cell*, 38(2), pp. 383–90.

Grover, R. K., Zhu, X., Nieusma, T., Jones, T., Boero, I., MacLeod, A. S., Mark, A., Niessen, S., Kim, H. J., Kong, L., Assad-Garcia, N., Kwon, K., Chesi, M., Smider, V. V, Salomon, D. R., Jelinek, D. F., Kyle, R. A., Pyles, R. B., Glass, J. I. *et al.* (2014). A structurally distinct human mycoplasma protein that generically blocks antigen-antibody union. *Science*, 343(6171), pp. 656–61. doi: 10.1126/science.1246135.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. and Hofacker, I. L. (2008). The Vienna RNA websuite. *Nucleic Acids Res.*, 36. doi: 10.1093/nar/gkn188.

Güell, M., van Noort, V., Yus, E., Chen, W. H., Leigh-Bell, J., Michalodimitrakis, K., Yamada, T., Arumugam, M., Doerks, T., Kuhner, S., Rode, M., Suyama, M., Schmidt, S., Gavin, A. C., Bork, P. and Serrano, L. (2009). Transcriptome complexity in a genome-reduced bacterium. *Science*, 326(5957), pp. 1268–1271. doi: 10.1126/science.1176951.

Hahn, T. W., Mothershed, E. A., Waldo, R. H. and Krause, D. C. (1999). Construction and analysis of a modified Tn4001 conferring chloramphenicol resistance in *Mycoplasma pneumoniae. Plasmid*, 41(2), pp. 120–4. doi: 10.1006/plas.1998.1387.

Halbedel, S., Eilers, H., Jonas, B., Busse, J., Hecker, M., Engelmann, S. and Stülke, J. (2007). Transcription in *Mycoplasma pneumoniae*: analysis of the promoters of the *ackA* and *ldh* genes. *J. Mol. Biol.*, 371(3), pp. 596–607. doi: 10.1016/j.jmb.2007.05.098.

Halden, N. F., Wolf, J. B. and Leonard, W. J. (1989). Identification of a novel site specific endonuclease produced by *Mycoplasma fermentans*: discovery while characterizing DNA binding proteins in T lymphocyte cell lines. *Nucleic Acids Res.*, 17(9), pp. 3491–9.

Hamasuna, R., Osada, Y. and Jensen, J. S. (2007). Isolation of *Mycoplasma genitalium* from first-void urine specimens by coculture with Vero cells. *J. Clin. Microbiol.*, 45(3), pp. 847–850. doi: 10.1128/JCM.02056-06.

Hames, C., Halbedel, S., Hoppert, M., Frey, J. and Stülke, J. (2009). Glycerol metabolism is important for cytotoxicity of *Mycoplasma pneumoniae. J. Bacteriol.*, 191(3), pp. 747–53. doi: 10.1128/JB.01103-08.

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*. *Proc. Natl. Acad. Sci.*, 103(44), pp. 16478–16483. doi: 10.1073/pnas.0608051103.

Hasselbring, B. M. and Krause, D. C. (2007). Cytoskeletal protein P41 is required to anchor the terminal organelle of the wall-less prokaryote *Mycoplasma pneumoniae*. *Mol. Microbiol.*, 63(1), pp. 44–53. doi: 10.1111/j.1365-2958.2006.05507.x.

Havarstein, L. S., Diep, D. B. and Nes, I. F. (1995). A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol. Microbiol.*, 16(2), pp. 229–240. doi: 10.1111/j.1365-2958.1995.tb02295.x.

Havarstein, L. S., Holo, H. and Nes, I. F. (1994). The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by Gram-positive bacteria. *Microbiology*, 140(9), pp. 2383–2389. doi: 10.1099/13500872-140-9-2383.

He, Y., Racz, R., Sayers, S., Lin, Y., Todd, T., Hur, J., Li, X., Patel, M., Zhao, B., Chung, M., Ostrow, J., Sylora, A., Dungarani, P., Ulysse, G., Kochhar, K., Vidri, B., Strait, K., Jourdian, G. W. and Xiang, Z. (2014). Updates on the web-based VIOLIN vaccine database and analysis system. *Nucleic Acids Res.*, 42(D1), pp. D1124–D1132. doi: 10.1093/nar/gkt1133.

Hedreyda, C. T., Lee, K. K. and Krause, D. C. (1993). Transformation of *Mycoplasma pneumoniae* with Tn4001 by electroporation. *Plasmid*, 30(2), pp. 170–175. doi: 10.1006/plas.1993.1047.

Le Hégarat, F., Salti-Montesanto, V., Hauck, Y. and Hirschbein, L. (1993). Purification and characterization of the HU-like protein HPB9 from the *Bacillus subtilis* nucleoid. *Biochim. Biophys. Acta*, 1172(1–2), pp. 101–7.

Heimann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.*, pp. 47–110. doi: 10.1016/S0065-2911(02)46002-X.

Helmann, J. D. (1999). Anti-sigma factors. *Curr. Opin. Microbiol.*, 2(2), pp. 135–141. doi: 10.1016/S1369-5274(99)80024-1.

Henderson, G. P. and Jensen, G. J. (2006). Three-dimensional structure of *Mycoplasma pneumoniae*'s attachment organelle and a model for its role in gliding motility. *Mol. Microbiol.*, 60(2), pp. 376–385. doi: 10.1111/j.1365-2958.2006.05113.x.

Henning, D., Eade, D., Langstone, A., Bean-Hodges, A., Marceglia, A. and Azzopardi, P. (2014). Asymptomatic *Mycoplasma genitalium* infection amongst marginalised young people accessing a youth health service in Melbourne. *Int. J. STD AIDS*, 25(4), pp. 299–302. doi: 10.1177/0956462413502317.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B. C. and Herrmann, R. (1996). Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.*, 24(22), pp. 4420–49.

Hjorth, S. V., Bjornelius, E., Lidbrink, P., Falk, L., Dohn, B., Berthelsen, L., Ma, L., Martin, D. H. and Jensen, J. S. (2006). Sequence-based typing of *Mycoplasma genitalium* reveals sexual transmission. *J. Clin. Microbiol.*, 44(6), pp. 2078–2083. doi: 10.1128/JCM.00003-06.

Horino, A., Kenri, T., Sasaki, Y., Okamura, N. and Sasaki, T. (2009). Identification of a site-specific tyrosine recombinase that mediates promoter inversions of phase-variable *mpl* lipoprotein genes in Mycoplasma penetrans. *Microbiology*, 155(4), pp. 1241–1249. doi: 10.1099/mic.0.025437-0.

Horino, A., Sasaki, Y., Sasaki, T. and Kenri, T. (2003). Multiple promoter inversions generate surface antigenic variation in *Mycoplasma penetrans*. *J. Bacteriol.*, 185(1), pp. 231–42.

Horinouchi, S. and Weisblum, B. (1982). Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. *J. Bacteriol.*, 150(2), pp. 815–825.

Hui, F. M., Zhou, L. and Morrison, D. A. (1995). Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes. *Gene*, 153(1), pp. 25–31. doi: 10.1016/0378-1119(94)00841-F.

Hutchison, C. A., Chuang, 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., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O. *et al.* (2016). Design and synthesis of a minimal bacterial genome. *Science*, 351(6280), p. 6253. doi: 10.1126/science.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*, 286(5447), pp. 2165–9.

Hyde, E. I., Hilton, M. D. and Whiteley, H. R. (1986). Interactions of *Bacillus subtilis* RNA polymerase with subunits determining the specificity of initiation. Sigma and delta peptides can bind simultaneously to core. *J. Biol. Chem.*, 261(35), pp. 16565–70.

Idahl, A., Lundin, E., Jurstrand, M., Kumlin, U., Elgh, F., Ohlson, N. and Ottander, U. (2011). *Chlamydia trachomatis* and *Mycoplasma genitalium* plasma antibodies in relation to epithelial ovarian tumors. *Infect. Dis. Obstet. Gynecol.*, 2011, p. 824627. doi: 10.1155/2011/824627.

Inagaki, Y., Bessho, Y., Hori, H. and Osawa, S. (1996). Cloning of the *Mycoplasma capricolum* gene encoding peptide-chain release factor. *Gene*, 169(1), pp. 101–3.

Inagaki, Y., Bessho, Y. and Osawa, S. (1993). Lack of peptide-release activity responding to codon UGA in *Mycoplasma capricolum*. *Nucleic Acids Res.*, 21(6), pp. 1335–8.

Indikova, I., Much, P., Stipkovits, L., Siebert-Gulle, K., Szostak, M. P., Rosengarten, R. and Citti, C. (2013). Role of the GapA and CrmA cytadhesins of *Mycoplasma gallisepticum* in promoting virulence and host colonization. *Infect. Immun.*, 81(5), pp. 1618–1624. doi: 10.1128/IAI.00112-13.

Into, T., Kiura, K., Yasuda, M., Kataoka, H., Inoue, N., Hasebe, A., Takeda, K., Akira, S. and Shibata, K. (2004). Stimulation of human Toll-like receptor (TLR) 2 and TLR6 with membrane lipoproteins of *Mycoplasma fermentans* induces apoptotic cell death after NF-kappa B activation. *Cell. Microbiol.*, 6(2), pp. 187–99.

Ireland, D. J. and Keelan, J. A. (2014). The maternal serological response to intrauterine *Ureaplasma sp.* infection and prediction of risk of pre-term birth. *Front. Immunol.*, 5, p. 624. doi: 10.3389/fimmu.2014.00624.

Ishag, H. Z. A., Liu, M. J., Yang, R. S., Xiong, Q. Y., Feng, Z. X. and Shao, G. Q. (2016). A replicating plasmid-based vector for GFP expression in *Mycoplasma hyopneumoniae*. *Genet. Mol. Res.*, 15(2). doi: 10.4238/gmr.15027832.

Ishag, H. Z. A., Xiong, Q., Liu, M., Feng, Z. and Shao, G. (2017). E. coli recA gene improves gene targeted homologous recombination in *Mycoplasma hyorhinis*. *J. Microbiol. Methods*, 136, pp. 49–56. doi: 10.1016/j.mimet.2017.03.004.

Iverson-Cabral, S. L., Astete, S. G., Cohen, C. R., Rocha, E. P. C. and Totten, P. A. (2006). Intrastrain heterogeneity of the *mgpB* gene in *Mycoplasma genitalium* is extensive *in vitro* and *in vivo* and suggests that variation is generated via recombination with repetitive chromosomal sequences. *Infect. Immun.*, 74(7), pp. 3715–3726. doi: 10.1128/IAI.00239-06.

Iverson-Cabral, S. L., Astete, S. G., Cohen, C. R. and Totten, P. A. (2007). *mgpB* and *mgpC* sequence diversity in *Mycoplasma genitalium* is generated by segmental reciprocal recombination with repetitive chromosomal sequences. *Mol. Microbiol.*, 66(1), pp. 55–73. doi: 10.1111/j.1365-2958.2007.05898.x.

Jacobs, E., Stuhlert, A., Drews, M., Pumpe, K., Schaefer, H. E., Kist, M. and Bredt, W. (1988). Host reactions to *Mycoplasma pneumoniae* infections in guinea-pigs preimmunized systemically with the adhesin of this pathogen. *Microb. Pathog.*, 5(4), pp. 259–265. doi: 10.1016/0882-4010(88)90098-8.

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*. *Appl. Environ. Microbiol.*, 71(6), pp. 2888–2893. doi: 10.1128/AEM.71.6.2888-2893.2005.

Jensen, J. S., Blom, J. and Lind, K. (1994). Intracellular location of *Mycoplasma genitalium* in cultured Vero cells as demonstrated by electron microscopy. *Int. J. Exp. Pathol.*, 75(2), pp. 91–8.

Jensen, J. S. and Bradshaw, C. (2015). Management of *Mycoplasma genitalium* infections - can we hit a moving target? *BMC Infect. Dis.*, 15, p. 343. doi: 10.1186/s12879-015-1041-6.

Jensen, J. S., Hansen, H. T. and Lind, K. (1996). Isolation of *Mycoplasma genitalium* strains from the male urethra. *J. Clin. Microbiol.*, 34(2), pp. 286–291.

Jiang, X., Lv, Y.-Q., Zhang, J.-N., Shi, Y.-L. and Xu, F.-F. (2013). *Mycoplasma penetrans* infection is a potential cause of immunoglobulin A nephropathy: a new animal model. *J. Nephrol.*, 26(3), pp. 470–475. doi: 10.5301/jn.5000180.

Jishage, M. and Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.*, 95(9), pp. 4953–8.

Juang, Y.-L. and Helmann, J. D. (1994). The δ Subunit of *Bacillus subtilis* RNA Polymerase. *J. Mol. Biol.*, 239(1), pp. 1–14. doi: 10.1006/jmbi.1994.1346.

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 Syst.*, 2(6), pp. 391–401. doi: 10.1016/j.cels.2016.04.015.

Jutras, B. L., Bowman, A., Brissette, C. A., Adams, C. A., Verma, A., Chenail, A. M. and Stevenson, B. (2012). EbfC (YbaB) is a new type of bacterial nucleoid-associated protein and a global regulator of gene expression in the Lyme disease spirochete. *J. Bacteriol.*, 194(13), pp. 3395–406. doi: 10.1128/JB.00252-12.

Kaiser, B. K. and Stoddard, B. L. (2011). DNA recognition and transcriptional regulation by the WhiA sporulation factor. *Sci. Rep.*, 1, p. 156. doi: 10.1038/srep00156.

Kannan, T. R. and Baseman, J. B. (2006). ADP-ribosylating and vacuolating cytotoxin of *Mycoplasma pneumoniae* represents unique virulence determinant among bacterial pathogens. *Proc. Natl. Acad. Sci.*, 103(17), pp. 6724–6729. doi: 10.1073/pnas.0510644103.

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. *Mol. Microbiol.*, 76(5), pp. 1127–1141. doi: 10.1111/j.1365-2958.2010.07092.x.

Kaplan, R. and Apirion, D. (1975). Decay of ribosomal ribonucleic acid in *Escherichia coli* cells starved for various nutrients. *J. Biol. Chem.*, 250(8), pp. 3174–8.

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), pp. 389–401. doi: 10.1016/j.cell.2012.05.044.

Keane, F. E. A., Thomas, B. J., Gilroy, C. B., Renton, A. and Taylor-Robinson, D. (2000). The association of *Chlamydia trachomatis* and *Mycoplasma genitalium* with non-gonococcal urethritis: observations on heterosexual men and their female partners. *Int. J. STD AIDS*, 11(7), pp. 435–439. doi: 10.1258/0956462001916209.

Kenri, T., Taniguchi, R., Sasaki, Y., Okazaki, N., Narita, M., Izumikawa, K., Umetsu, M. and Sasaki, T. (1999). Identification of a new variable sequence in the P1 cytadhesin gene of *Mycoplasma pneumoniae*: evidence for the generation of antigenic variation by DNA recombination between repetitive sequences. *Infect. Immun.*, 67(9), pp. 4557–62.

Kent, B. N., Foecking, M. F. and Calcutt, M. J. (2012). Development of a novel plasmid as a shuttle vector for heterologous gene expression in *Mycoplasma yeatsii*. *J. Microbiol. Methods*, 91(1), pp. 121–127. doi: 10.1016/j.mimet.2012.07.018.

King, K. W. and Dybvig, K. (1991). Plasmid transformation of *Mycoplasma mycoides* subspecies *mycoides* is promoted by high concentrations of polyethylene glycol. *Plasmid*, 26(2), pp. 108–15.

King, K. W. and Dybvig, K. (1992). Nucleotide sequence of *Mycoplasma mycoides* subspecies *mycoides* plasmid pKMK1. *Plasmid*, 28(1), pp. 86–91.

Kingsford, C. L., Ayanbule, K. and Salzberg, S. L. (2007). Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.*, 8(2), p. R22. doi: 10.1186/gb-2007-8-2-r22.

Klieneberger, E. (1935). The natural occurrence of pleuropneumonia-like organism in apparent symbiosis with *Streptobacillus moniliformis* and other bacteria. *J. Pathol. Bacteriol.*, 40(1), pp. 93–105. doi: 10.1002/path.1700400108.

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., Weinstock, G. M., Wilson, R. K. and Ding, L. (2009). VarScan: Variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, 25(17), pp. 2283–2285. doi: 10.1093/bioinformatics/btp373.

Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. and Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, 58(3), pp. 401–65.

Krause, D. C. and Balish, M. F. (2001). Structure, function, and assembly of the terminal organelle of *Mycoplasma pneumoniae*. *FEMS Microbiol. Lett.*, 198(1), pp. 1–7. doi: 10.1111/j.1574-6968.2001.tb10610.x.

Krause, D. C. and Balish, M. F. (2004). Cellular engineering in a minimal microbe: structure and assembly of the terminal organelle of *Mycoplasma pneumoniae*. *Mol. Microbiol.*, 51(4), pp. 917–24.

Krause, D. C. and Baseman, J. B. (1982). *Mycoplasma pneumoniae* proteins that selectively bind to host cells. *Infect. Immun.*, 37(1), pp. 382–6.

Krishnakumar, R., Assad-Garcia, N., Benders, G. A., Phan, Q., Montague, M. G. and Glass, J. I. (2010). Targeted chromosomal knockouts in Mycoplasma pneumoniae. *Appl. Environ. Microbiol.*, 76(15), pp. 5297–5299. doi: 10.1128/AEM.00024-10.

Kuhner, S., van Noort, V., Betts, M. J., Leo-Macias, A., Batisse, C., Rode, M., Yamada, T., Maier, T., Bader, S., Beltran-Alvarez, P., Castano-Diez, D., Chen, W.-H., Devos, D., Guell, M., Norambuena, T., Racke, I., Rybin, V., Schmidt, A., Yus, E. *et al.* (2009). Proteome organization in a genome-reduced bacterium. *Science*, 326(5957), pp. 1235–1240. doi: 10.1126/science.1176343.

Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9(4), pp. 357–359. doi: 10.1038/nmeth.1923.

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 Res.*, 31(22), pp. 6610–8.

Ledford, H. (2015). CRISPR, the disruptor. *Nature*, 522(7554), pp. 20–24. doi: 10.1038/522020a.

Lee, S.-W., Browning, G. F. and Markham, P. F. (2008). Development of a replicable oriC plasmid for *Mycoplasma gallisepticum* and *Mycoplasma imitans*, and gene disruption through homologous recombination in *M. gallisepticum*. *Microbiology*, 154(9), pp. 2571–2580. doi: 10.1099/mic.0.2008/019208-0.

Lesnik, E. A., Sampath, R., Levene, H. B., Henderson, T. J., McNeil, J. A. and Ecker, D. J. (2001). Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res.*, 29(17), pp. 3583–94. doi: 10.1093/nar/29.17.3583.

Levin-Karp, A., Barenholz, U., Bareia, T., Dayagi, M., Zelcbuch, L., Antonovsky, N., Noor, E. and Milo, R. (2013). Quantifying translational coupling in E. coli synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth. Biol.*, 2(6), pp. 327–336. doi: 10.1021/sb400002n.

Li, H. (2012). Exploring single-sample SNP and indel calling with whole-genome *de novo* assembly. *Bioinformatics*, 28(14), pp. 1838–1844. doi: 10.1093/bioinformatics/bts280.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Li, Q., Feng, J., Hu, H.-L., Chen, X.-C., Li, F.-Q. and Hong, G.-F. (2004). A HU-like gene mutation in *Rhizobium leguminosarum* bv. *viciae* affects the expression of nodulation genes. *Mol. Microbiol.*, 51(3), pp. 861–71.

Liang, H., Li, L., Dong, Z., Surette, M. G. and Duan, K. (2008). The YebC family protein PA0964 negatively regulates the *Pseudomonas aeruginosa* quinolone signal system and pyocyanin production. *J. Bacteriol.*, 190(18), pp. 6217–6227. doi: 10.1128/JB.00428-08.

Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), pp. 923–930. doi: 10.1093/bioinformatics/btt656.

Lim, K., Tempczyk, A., Parsons, J. F., Bonander, N., Toedt, J., Kelman, Z., Howard, A., Eisenstein, E. and Herzberg, O. (2002). Crystal structure of YbaB from *Haemophilus influenzae* (HI0442), a protein of unknown function coexpressed with the recombinational DNA repair protein RecR. *Proteins Struct. Funct. Bioinforma.*, 50(2), pp. 375–379. doi: 10.1002/prot.10297.

Linchevski, I., Klmenet, E. and Nir-Paz, R. (2009). *Mycoplasma pneumoniae* vaccine protective efficacy and adverse reactions—Systematic review and meta-analysis. *Vaccine*, 27(18), pp. 2437–2446. doi: 10.1016/j.vaccine.2009.01.135.

Lloréns-Rico, V., Lluch-Senar, M., Serrano, L., Llorens-Rico, V., Lluch-Senar, M., Serrano, L., Lloréns-Rico, V., Lluch-Senar, M., Serrano, L., Llorens-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 Res.*, 43(7), pp. 3442–3453. doi: 10.1093/nar/gkv170.

Lluch-Senar, M., Delgado, J., Chen, W.-H., Lloréns-Rico, V., O'Reilly, F. J., Wodke, J. A., 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. (2015). Defining a minimal cell: essentiality of small ORFs and ncRNAs in a genome-reduced bacterium. *Mol. Syst. Biol.*, 11(1), p. 780. doi: 10.15252/msb.20145558.

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 Genet.*, 9(1). doi: 10.1371/journal.pgen.1003191.

Lluch-Senar, M., Querol, E. and Piñol, J. (2010). Cell division in a minimal bacterium in the absence of *ftsZ*. *Mol. Microbiol.*, 78(2), pp. 278–289. doi: 10.1111/j.1365-2958.2010.07306.x.

Lluch-Senar, M., Vallmitjana, M., Querol, E. and Piñol, J. (2007). A new promoterless reporter vector reveals antisense transcription in *Mycoplasma genitalium*. *Microbiology*, 153(8), pp. 2743–2752. doi: 10.1099/mic.0.2006/007559-0.

Lo, S. C., Hayes, M. M., Kotani, H., Pierce, P. F., Wear, D. J., Newton, P. B., Tully, J. G. and Shih, J. W. (1993). Adhesion onto and invasion into mammalian cells by Mycoplasma penetrans: a newly isolated Mycoplasma from patients with AIDS. *Mod. Pathol.*, 6(3), pp. 276–80.

Locke, J. C. W., Young, J. W., Fontes, M., Jimenez, M. J. H. and Elowitz, M. B. (2011). Stochastic pulse regulation in bacterial stress response. *Science*, 334(6054), pp. 366–369. doi: 10.1126/science.1208144.

Loechel, S., Inamine, J. M. and Hu, P. C. (1991). A novel translation initiation region from *Mycoplasma genitalium* that functions in *Escherichia coli*. *Nucleic Acids Res.*, 19(24), pp. 6905–11.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.

Lusetti, S. L. and Cox, M. M. (2002). The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu. Rev. Biochem.*, 71(1), pp. 71–100. doi: 10.1146/annurev.biochem.71.083101.133940.

Lyon, B. R., May, J. W. and Skurray, R. A. (1984). Tn4001: a gentamicin and kanamycin resistance transposon in *Staphylococcus aureus*. *Mol. Gen. Genet.*, 193(3), pp. 554–6.

Ma, L., Jensen, J. S., Myers, L., Burnett, J., Welch, M., Jia, Q. and Martin, D. H. (2007). *Mycoplasma genitalium*: an efficient strategy to generate genetic variation from a minimal genome. *Mol. Microbiol.*, 66(1), pp. 220–236. doi: 10.1111/j.1365-2958.2007.05911.x.

Mackiewicz, P., Mackiewicz, D., Kowalczuk, M. and Cebrat, S. (2001). Flip-flop around the origin and terminus of replication in prokaryotic genomes. *Genome Biol.*, 2(12), p. 1004.1-1004.4.

Madsen, M. L., Nettleton, D., Thacker, E. L., Edwards, R. and Minion, F. C. (2006a). Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays. *Infect. Immun.*, 74(1), pp. 160–166. doi: 10.1128/IAI.74.1.160-166.2006.

Madsen, M. L., Nettleton, D., Thacker, E. L. and Minion, F. C. (2006b). Transcriptional profiling of *Mycoplasma hyopneumoniae* during iron depletion using microarrays. *Microbiology*, 152(4), pp. 937–944. doi: 10.1099/mic.0.28674-0.

Madsen, M. L., Puttamreddy, S., Thacker, E. L., Carruthers, M. D. and Minion, F. C. (2008). Transcriptome changes in *Mycoplasma hyopneumoniae* during infection. *Infect. Immun.*, 76(2), pp. 658–663. doi: 10.1128/IAI.01291-07.

Maeda, H., Fujita, N. and Ishihama, A. (2000). Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res.*, 28(18), pp. 3497–503.

Maglennon, G. A., Cook, B. S., Deeney, A. S., Bossé, J. T., Peters, S. E., Langford, P. R., Maskell, D. J., Tucker, A. W., Wren, B. W., Rycroft, A. N. and BRaDP1T consortium (2013a). Transposon mutagenesis in *Mycoplasma hyopneumoniae* using a novel mariner-based system for generating random mutations. *Vet. Res.*, 44(1), p. 124. doi: 10.1186/1297-9716-44-124.

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 BRaDP1T consortium (2013b). Development of a self-replicating plasmid system for *Mycoplasma hyopneumoniae*. *Vet. Res.*, 44(1), p. 63. doi: 10.1186/1297-9716-44-63.

Mahairas, G. G. and Minion, F. C. (1989). Random insertion of the gentamicin resistance transposon Tn4001 in *Mycoplasma pulmonis*. *Plasmid*, 21(1), pp. 43–7.

Mahmood, T. and Yang, P.-C. (2012). Western blot: technique, theory, and trouble shooting. *N. Am. J. Med. Sci.*, 4(9), pp. 429–34. doi: 10.4103/1947-2714.100998.

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*. *Mol. Biosyst.*, 9(7), pp. 1743–55. doi: 10.1039/c3mb70113a.

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. *Mol. Syst. Biol.*, 7(1), p. 511. doi: 10.1038/msb.2011.38.

Manhart, L. E. (2012). Another STI associated with HIV-1 acquisition: now what? *AIDS*, 26(5), pp. 635–637. doi: 10.1097/QAD.0b013e32834ff6ad.

Manhart, L. E. (2013). *Mycoplasma genitalium*: An emergent sexually transmitted disease? *Infect. Dis. Clin. North Am.*, 27(4), pp. 779–792. doi: 10.1016/j.idc.2013.08.003.

Manhart, L. E., Holmes, K. K., Hughes, J. P., Houston, L. S. and Totten, P. A. (2007). *Mycoplasma genitalium* among young adults in the United States: an emerging sexually transmitted infection. *Am. J. Public Health*, 97(6), pp. 1118–25. doi: 10.2105/AJPH.2005.074062.

Maniloff, J. and Dybvig, K. (1988). Mycoplasma viruses. *CRC Crit. Rev. Microbiol.*, 15(4), pp. 339–389. doi: 10.3109/10408418809104462.

Mårdh, P. A. (1983). *Mycoplasma hominis* - a neglected human pathogen. *Eur. J. Clin. Microbiol.*, 2(4), pp. 303–8.

Mariscal, A. M., González-González, L., Querol, E. and Piñol, J. (2016). All-in-one construct for genome engineering using Cre-lox technology. *DNA Res.*, 23(3), pp. 263–270. doi: 10.1093/dnares/dsw015.

Martinelli, L., García-Morales, L., Querol, E., Piñol, J., Fita, I. and Calisto, B. M. (2016). Structure-guided mutations in the terminal organelle protein MG491 cause major motility and morphologic alterations on *Mycoplasma genitalium*. *PLOS Pathog.* Edited by M. F. Balish, 12(4), p. e1005533. doi: 10.1371/journal.ppat.1005533.

Martinelli, L., Lalli, D., García-Morales, L., Ratera, M., Querol, E., Piñol, J., Fita, I. and Calisto, B. M. (2015). A major determinant for gliding motility in Mycoplasma genitalium. *J. Biol. Chem.*, 290(3), pp. 1699–1711. doi: 10.1074/jbc.M114.594762.

Mauri, M. and Klumpp, S. (2014). A model for sigma factor competition in bacterial cells. *PLoS Comput. Biol.*, 10(10), p. e1003845. doi: 10.1371/journal.pcbi.1003845.

Mazin, P. V., Fisunov, G. Y., Gorbachev, A. Y., Kapitskaya, K. Y., Altukhov, I. A., Semashko, T. A., Alexeev, D. G. and Govorun, V. M. (2014). Transcriptome analysis reveals novel regulatory mechanisms in a genome-reduced bacterium. *Nucleic Acids Res.*, 42(21), pp. 13254–13268. doi: 10.1093/nar/gku976.

McGowin, C. L., Anderson-Smits, C., Martin, D., Pyles, R. and Cagle, M. (2011). *Mycoplasma genitalium*: An Emerging Cause of Sexually Transmitted Disease in Women. *PLoS Pathog.* Edited by M. Manchester, 7(5), p. e1001324. doi: 10.1371/journal.ppat.1001324.

McGowin, C. L., Ma, L., Jensen, J. S., Mancuso, M. M., Hamasuna, R., Adegboye, D. and Martin, D. H. (2012). Draft genome sequences of four axenic *Mycoplasma genitalium* strains isolated from Denmark, Japan, and Australia. *J. Bacteriol.*, 194(21), pp. 6010–6011. doi: 10.1128/JB.01478-12.

McGowin, C. L., Ma, L., Martin, D. H. and Pyles, R. B. (2009a). *Mycoplasma genitalium*-encoded MG309 activates NF-B via Toll-like receptors 2 and 6 To elicit proinflammatory cytokine secretion from human genital epithelial cells. *Infect. Immun.*, 77(3), pp. 1175–1181. doi: 10.1128/IAI.00845-08.

McGowin, C. L., Popov, V. L. and Pyles, R. B. (2009b). Intracellular Mycoplasma genitalium infection of human vaginal and cervical epithelial cells elicits distinct patterns of inflammatory cytokine secretion and provides a possible survival niche against macrophage-mediated killing. *BMC Microbiol.*, 9(1), p. 139. doi: 10.1186/1471-2180-9-139.

Mernaugh, G. R., Dallo, S. F., Holt, S. C. and Baseman, J. B. (1993). Properties of adhering and nonadhering populations of *Mycoplasma genitalium. Clin Infect Dis.*, S1, pp. S69-78.

Messick, J. B. (2004). Hemotrophic Mycoplasmas (Hemoplasmas): a review and new insights into pathogenic potential. *Vet. Clin. Pathol.*, 33(1), pp. 2–13.

Miles, R. J. (1992). Catabolism in Mollicutes. *Microbiology*, 138(9), pp. 1773–1783. doi: 10.1099/00221287-138-

9-1773.

Miller, C. J., Kassem, H. S., Pepper, S. D., Hey, Y., Ward, T. H. and Margison, G. P. (2003). Mycoplasma infection significantly alters microarray gene expression profiles. *Biotechniques*, 35, pp. 812–814.

Minion, C. F. and Kapke, P. A. (1998). Transformation of Mycoplasmas. in *Mycoplasma Protoc.*, pp. 227–234. doi: 10.1385/0-89603-525-5:227.

Minion, F. C., Jarvill-Taylor, K. J., Billings, D. E. and Tigges, E. (1993). Membrane-associated nuclease activities in Mycoplasmas. *J. Bacteriol.*, 175(24), pp. 7842–7.

Mitchell, J. E., Oshima, T., Piper, S. E., Webster, C. L., Westblade, L. F., Karimova, G., Ladant, D., Kolb, A., Hobman, J. L., Busby, S. J. W. and Lee, D. J. (2007). The *Escherichia coli* regulator of sigma 70 protein, Rsd, can up-regulate some stress-dependent promoters by sequestering sigma 70. *J. Bacteriol.*, 189(9), pp. 3489–3495. doi: 10.1128/JB.00019-07.

Miyata, M. (2008). Centipede and inchworm models to explain Mycoplasma gliding. *Trends Microbiol.*, 16(1), pp. 6–12. doi: 10.1016/j.tim.2007.11.002.

Miyata, M. (2010). Unique centipede mechanism of Mycoplasma gliding. *Annu. Rev. Microbiol.*, 64(1), pp. 519–537. doi: 10.1146/annurev.micro.112408.134116.

Miyata, M. and Hamaguchi, T. (2016). Integrated information and prospects for gliding mechanism of the pathogenic bacterium *Mycoplasma pneumoniae*. *Front. Microbiol.*, 7, p. 960. doi: 10.3389/fmicb.2016.00960.

Miyata, M., Yamamoto, H., Shimizu, T., Uenoyama, A., Citti, C. and Rosengarten, R. (2000). Gliding mutants of *Mycoplasma mobile*: relationships between motility and cell morphology, cell adhesion and microcolony formation. *Microbiology*, (6), pp. 1311–20. doi: 10.1099/00221287-146-6-1311.

Miyazaki, R., Minoia, M., Pradervand, N., Sulser, S., Reinhard, F. and van der Meer, J. R. (2012). Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element. *PLoS Genet.*, 8(7), p. e1002818. doi: 10.1371/journal.pgen.1002818.

Montagnier, L., Berneman, D., Guétard, D., Blanchard, A., Chamaret, S., Rame, V., Van Rietschoten, J., Mabrouk, K. and Bahraoui, E. (1990). Infectivity inhibition of HIV prototype strains by antibodies directed against a peptide sequence of Mycoplasma. *C. R. Acad. Sci. III.*, 311(12), pp. 425–30.

Mortier-Barrière, I., Velten, M., Dupaigne, P., Mirouze, N., Piétrement, O., McGovern, S., Fichant, G., Martin, B., Noirot, P., Le Cam, E., Polard, P. and Claverys, J.-P. (2007). A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell*, 130(5), pp. 824–836. doi: 10.1016/j.cell.2007.07.038.

Muffler, A., Fischer, D., Altuvia, S., Storz, G. and Hengge-Aronis, R. (1996). The response regulator RssB controls stability of the as subunit of RNA polymerase in *Escherichia coli*. *EMBO J.*, 15(6), pp. 1333–1339.

Musatovova, O., Dhandayuthapani, S. and Baseman, J. B. (2006). Transcriptional heat shock response in the smallest known self-replicating cell, *Mycoplasma genitalium*. *J. Bacteriol.*, 188(8), pp. 2845–2855. doi: 10.1128/JB.188.8.2845-2855.2006.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A. and Hattori, M. (2006). The 160-Kilobase genome of the bacterial endosymbiont *Carsonella*. *Science*, 314(5797), pp. 267–267. doi: 10.1126/science.1134196.

Nakano, S., Kuster-Schock, E., Grossman, A. D. and Zuber, P. (2003). Spx-dependent global transcriptional control is induced by thiol-specific oxidative stress in *Bacillus subtilis*. *Proc. Natl. Acad. Sci.*, 100(23), pp. 13603–13608. doi: 10.1073/pnas.2235180100.

Namiki, K., Goodison, S., Porvasnik, S., Allan, R. W., Iczkowski, K. A., Urbanek, C., Reyes, L., Sakamoto, N. and Rosser, C. J. (2009). Persistent exposure to Mycoplasma induces malignant transformation of human prostate cells. *PLoS One*, 4(9), p. e6872. doi: 10.1371/journal.pone.0006872.

Napierala, M. S. and Weiss, H. A. (2009). Association of *Mycoplasma genitalium* and HIV infection: a systematic review and meta-analysis. *AIDS*, 23(5), pp. 611–620. doi: 10.1097/QAD.0b013e328323da3e.

Neimark, H. and Kocan, K. M. (1997). The cell wall-less rickettsia *Eperythrozoon wenyonii* is a Mycoplasma. *FEMS Microbiol. Lett.*, 156(2), pp. 287–91.

Newberry, K. J., Nakano, S., Zuber, P. and Brennan, R. G. (2005). Crystal structure of the *Bacillus subtilis* anti-alpha, global transcriptional regulator, Spx, in complex with the C-terminal domain of RNA polymerase. *Proc. Natl. Acad. Sci.*, 102(44), pp. 15839–15844. doi: 10.1073/pnas.0506592102.

Nicholas, R. A. J., Ayling, R. D. and McAuliffe, L. (2009). Vaccines for Mycoplasma diseases in animals and man. *J. Comp. Pathol.*, 140(2–3), pp. 85–96. doi: 10.1016/j.jcpa.2008.08.004.

Nicholas, R. and Baker, S. (1998). Recovery of Mycoplasmas from animals. in *Mycoplasma Protoc.*, pp. 37–43. doi: 10.1385/0-89603-525-5:37.

Noormohammadi, A. H., Markham, P. F., Kanci, A., Whithear, K. G. and Browning, G. F. (2000). A novel mechanism for control of antigenic variation in the haemagglutinin gene family of *Mycoplasma synoviae*. *Mol. Microbiol.*, 35(4), pp. 911–23.

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., Bork, P. and Gavin, A.-C. (2012). Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Mol. Syst. Biol.*, 8(571), p. 571. doi: 10.1038/msb.2012.4.

Oakeshott, P., Hay, P., Taylor-Robinson, D., Hay, S., Dohn, B., Kerry, S. and Jensen, J. S. (2004). Prevalence of *Mycoplasma genitalium* in early pregnancy and relationship between its presence and pregnancy outcome. *BJOG An Int. J. Obstet. Gynaecol.*, 111(12), pp. 1464–1467. doi: 10.1111/j.1471-0528.2004.00276.x.

Ochman, H., Lawrence, J. G. and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405(6784), pp. 299–304. doi: 10.1038/35012500.

Österberg, S., Peso-Santos, T. Del and Shingler, V. (2011). Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.*, 65(1), pp. 37–55. doi: 10.1146/annurev.micro.112408.134219.

Paget, M. S. (2015). Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules*, 5(3), pp. 1245–65. doi: 10.3390/biom5031245.

Paget, M. S. B. and Helmann, J. D. (2003). The sigma70 family of sigma factors. *Genome Biol.*, 4(1), p. 203. doi: 10.1186/gb-2003-4-1-203.

Paik, S. H., Chakicherla, A. and Hansen, J. N. (1998). Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *J. Biol. Chem.*, 273(36), pp. 23134–42. doi: 10.1074/JBC.273.36.23134.

Palmer, H. M., Gilroy, C. B., Claydon, E. J. and Taylor-Robinson, D. (1991). Detection of Mycoplasma genitalium in the genitourinary tract of women by the polymerase chain reaction. *Int. J. STD AIDS*, 2(4), pp. 261–263. doi: 10.1177/095646249100200407.

Párraga-Niño, N., Colomé-Calls, N., Canals, F., Querol, E. and Ferrer-Navarro, M. (2012). A comprehensive proteome of *Mycoplasma genitalium*. *J. Proteome Res.*, pp. 3305–3316. doi: 10.1021/pr300084c.

Pennisi E. (2010). Synthetic genome brings new life to bacterium. *Science*, 328(5981), pp. 958–959.

Perez-Brocal, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J. M., Silva, F. J., Moya, A. and Latorre, A. (2006). A small microbial genome: the end of a long symbiotic relationship? *Science*, 314(5797), pp. 312–313. doi: 10.1126/science.1130441.

Pérez-González, J. A., Vara, J. and Jiménez, A. (1983). Acetylation of puromycin by *Streptomyces alboniger* the producing organism. *Biochem. Biophys. Res. Commun.*, 113(3), pp. 772–7.

Peterson, S. N., Bailey, C. C., Jensen, J. S., Borre, M. B., King, E. S., Bott, K. F. and Hutchison, C. A. (1995). Characterization of repetitive DNA in the *Mycoplasma genitalium* genome: possible role in the generation of antigenic variation. *Proc. Natl. Acad. Sci. U. S. A.*, 92(25), pp. 11829–33.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids*

Res., 29(9), p. e45. doi: 10.1093/nar/29.9.e45.

Pflaum, K., Tulman, E. R., Beaudet, J., Liao, X. and Geary, S. J. (2016). Global changes in *Mycoplasma gallisepticum* phase-variable lipoprotein gene *vlhA* expression during in vivo infection of the natural chicken host. *Infect. Immun.* Edited by S. M. Payne, 84(1), pp. 351–355. doi: 10.1128/IAI.01092-15.

Philibert, P., Khiri, H., Pénaranda, G., Camus, C., Drogoul, M.-P. and Halfon, P. (2014). High prevalence of asymptomatic sexually transmitted infections among men who have sex with men. *J. Clin. Med.*, 3(4), pp. 1386–1391. doi: 10.3390/jcm3041386.

Phillips, D. M., Pearce-Pratt, R., Tan, X. and Zacharopoulos, V. R. (1992). Association of Mycoplasma with HIV-1 and HTLV-I in Human T Lymphocytes. *AIDS Res. Hum. Retroviruses*, 8(11), pp. 1863–1868. doi: 10.1089/aid.1992.8.1863.

Piccinelli, G., Gargiulo, F., Biscaro, V., Caccuri, F., Caruso, A. and De Francesco, M. A. (2017). Analysis of mutations in DNA gyrase and topoisomerase IV of *Ureaplasma urealyticum* and *Ureaplasma parvum* serovars resistant to fluoroquinolones. *Infect. Genet. Evol.*, 47, pp. 64–67. doi: 10.1016/j.meegid.2016.11.019.

Pich, O. Q., Burgos, R., Ferrer-Navarro, M., Querol, E. and Piñol, J. (2006a). *Mycoplasma genitalium* mg200 and mg386 genes are involved in gliding motility but not in cytadherence. *Mol. Microbiol.*, 60(6), pp. 1509–1519. doi: 10.1111/j.1365-2958.2006.05187.x.

Pich, O. Q., Burgos, R., Ferrer-Navarro, M., Querol, E. and Piñol, J. (2008). Role of *Mycoplasma genitalium* MG218 and MG317 cytoskeletal proteins in terminal organelle organization, gliding motility and cytadherence. *Microbiology*, 154(10), pp. 3188–3198. doi: 10.1099/mic.0.2008/020636-0.

Pich, O. Q., Burgos, R., Planell, R., Querol, E. and Piñol, J. (2006b). Comparative analysis of antibiotic resistance gene markers in *Mycoplasma genitalium*: application to studies of the minimal gene complement. *Microbiology*, 152(2), pp. 519–527. doi: 10.1099/mic.0.28287-0.

Pich, O. Q., Burgos, R., Querol, E. and Piñol, J. (2009). P110 and P140 cytadherence-related proteins are negative effectors of terminal organelle duplication in *Mycoplasma genitalium*. *PLoS One*. Edited by C. Herman, 4(10), p. e7452. doi: 10.1371/journal.pone.0007452.

Piir, K., Paier, A., Liiv, A., Tenson, T. and Maiväli, U. (2011). Ribosome degradation in growing bacteria. *EMBO Rep.*, 12(5), pp. 458–62. doi: 10.1038/embor.2011.47.

Pilo, P., Vilei, E. M., Peterhans, E., Bonvin-Klotz, L., Stoffel, M. H., Dobbelaere, D. and Frey, J. (2005). A metabolic enzyme as a primary virulence factor of Mycoplasma mycoides subsp. mycoides small colony. *J. Bacteriol.*, 187(19), pp. 6824–6831. doi: 10.1128/JB.187.19.6824-6831.2005.

Pour-El, I., Adams, C. and Minion, F. C. (2002). Construction of mini-Tn4001tet and its use in *Mycoplasma gallisepticum. Plasmid*, 47(2), pp. 129–37. doi: 10.1006/plas.2001.1558.

Proctor, B. L. and Gaulden, M. E. (1986). Chemical stability of mitomycin C in culture medium with and without fetal calf serum as determined by high pressure liquid chromatography and mass spectrometry. *Arch. Environ. Contam. Toxicol.*, 15(2), pp. 235–240. doi: 10.1007/BF01059972.

Razin, S. (1994). DNA probes and PCR in diagnosis of Mycoplasma infections. *Mol. Cell. Probes*, 8(6), pp. 497–511. doi: 10.1006/mcpr.1994.1071.

Razin, S. (1996). *Mycoplasmas*. *Med. Microbiol.*

Razin, S., Kahane, I., Banai, M. and Bredt, W. (1981). Adhesion of Mycoplasmas to eukaryotic cells. *Ciba Found. Symp.*, 80, pp. 98–118.

Razin, S., Yogev, D. and Naot, Y. (1998). Molecular biology and pathogenicity of Mycoplasmas. *Microbiol. Mol. Biol. Rev.*, 62(4), pp. 1094–156.

Reddy, S. P., Rasmussen, W. G. and Baseman, J. B. (1996). Isolation and characterization of transposon Tn4001-generated, cytadherence-deficient transformants of Mycoplasma pneumoniae and Mycoplasma genitalium. *FEMS Immunol. Med. Microbiol.*, 15(4), pp. 199–211. doi: 10.1111/j.1574-695X.1996.tb00086.x.

Reeves, P. (1965). The bacteriocins. *Bacteriol. Rev.*, 29, pp. 24–45.

Relich, R. F. and Balish, M. F. (2011). Insights into the function of *Mycoplasma pneumoniae* protein P30 from orthologous gene replacement. *Microbiology*, 157(10), pp. 2862–2870. doi: 10.1099/mic.0.052464-0.

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

Renzette, N., Gumlaw, N., Nordman, J. T., Krieger, M., Yeh, S. P., Long, E., Centore, R., Boonsombat, R. and Sandler, S. J. (2005). Localization of RecA in *Escherichia coli* K-12 using RecA-GFP. *Mol. Microbiol.*, 57(4), pp. 1074–1085. doi: 10.1111/j.1365-2958.2005.04755.x.

Reznikoff, W. S., Siegele, D. A., Cowing, D. W. and Gross, C. A. (1985). The regulation of transcription initiation in bacteria. *Annu. Rev. Genet.*, 19(1), pp. 355–387. doi: 10.1146/annurev.ge.19.120185.002035.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.*, 17(10), pp. 1030–1032. doi: 10.1038/13732.

Rikihisa, Y., Kawahara, M., Wen, B., Kociba, G., Fuerst, P., Kawamori, F., Suto, C., Shibata, S. and Futohashi, M. (1997). Western immunoblot analysis of *Haemobartonella muris* and comparison of 16S rRNA gene sequences of *H. muris*, *H. felis*, and *Eperythrozoon suis*. *J. Clin. Microbiol.*, 35(4), pp. 823–9.

Riley, S. P., Bykowski, T., Cooley, A. E., Burns, L. H., Babb, K., Brissette, C. A., Bowman, A., Rotondi, M., Miller, M. C., DeMoll, E., Lim, K., Fried, M. G. and Stevenson, B. (2009). *Borrelia burgdorferi* EbfC defines a newly-identified, widespread family of bacterial DNA-binding proteins. *Nucleic Acids Res.*, 37(6), pp. 1973–1983. doi: 10.1093/nar/gkp027.

Rocha, E. P. C. and Blanchard, A. (2002). Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution. *Nucleic Acids Res.*, 30(9), pp. 2031–2042.

Rocha, E. P. C., Cornet, E. and Michel, B. (2005). Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.*, pp. 0247–0259. doi: 10.1371/journal.pgen.0010015.

Rosengarten, R. and Wise, K. S. (1991). The Vlp system of *Mycoplasma hyorhinis*: combinatorial expression of distinct size variant lipoproteins generating high-frequency surface antigenic variation. *J. Bacteriol.*, 173(15), pp. 4782–93.

Röske, K., Blanchard, A., Chambaud, I., Citti, C., Helbig, J. H., Prevost, M. C., Rosengarten, R. and Jacobs, E. (2001). Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect. Immun.*, 69(12), pp. 7642–51. doi: 10.1128/IAI.69.12.7642-7651.2001.

Ross, J. D. C. and Jensen, J. S. (2006). *Mycoplasma genitalium* as a sexually transmitted infection: implications for screening, testing, and treatment. *Sex. Transm. Infect.*, 82(4), pp. 269–71. doi: 10.1136/sti.2005.017368.

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, 132, pp. 365–86.

Saikolappan, S., Sasindran, S. J., Yu, H. D., Baseman, J. B. and Dhandayuthapani, S. (2009). The *Mycoplasma genitalium* MG_454 gene product resists killing by organic hydroperoxides. *J. Bacteriol.*, 191(21), pp. 6675–82. doi: 10.1128/JB.01066-08.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press. New York*. doi: 574.873224 1/1989.

Sanghvi, J. C., Regot, S., Carrasco, S., Karr, J. R., Gutschow, M. V, Bolival, B. and Covert, M. W. (2013). Accelerated discovery via a whole-cell model. *Nat. Methods*, 10(12), pp. 1192–5. doi: 10.1038/nmeth.2724.

Sasaki, Y., Honda, M., Makino, M. and Sasakil, T. (1993). Mycoplasmas stimulate replication of human immunodeficiency virus type 1 through selective activation of CD4 + T lymphocytes. *AIDS Res. Hum. Retroviruses*, 9(8), pp. 775–780. doi: 10.1089/aid.1993.9.775.

Schafer, E. R., Oneal, M. J., Madsen, M. L. and Minion, F. C. (2007). Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to hydrogen peroxide. *Microbiology*, 153(11), pp. 3785–3790. doi: 10.1099/mic.0.2007/011387-0.

Schmidl, S. R., Gronau, K., Pietack, N., Hecker, M., Becher, D. and Stülke, J. (2010). The phosphoproteome of the minimal bacterium *Mycoplasma pneumoniae*: analysis of the complete known Ser/Thr kinome suggests the existence of novel kinases. *Mol. Cell. Proteomics*, 9, pp. 1228–1242. doi: 10.1074/mcp.M900267-MCP200.

Schott, C., Cai, H., Parker, L., Bateman, K. G. and Caswell, J. L. (2014). Hydrogen peroxide production and free radical-mediated cell stress in *Mycoplasma bovis* pneumonia. *J. Comp. Pathol.*, 150(2–3), pp. 127–137. doi: 10.1016/j.jcpa.2013.07.008.

Schulz, A. and Schumann, W. (1996). *hrcA*, the first gene of the *Bacillus subtilis dnaK* operon encodes a negative regulator of class I heat shock genes. *J. Bacteriol.*, 178(4), pp. 1088–93.

Schümperli, D., McKenney, K., Sobieski, D. A. and Rosenberg, M. (1982). Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. *Cell*, 30(3), pp. 865–871. doi: 10.1016/0092-8674(82)90291-4.

Semashko, T. A., Arzamasov, A. A., Fisunov, G. Y. and Govorun, V. M. (2017). Transcription profiling data set of different states of *Mycoplasma gallisepticum*. *Genomics Data*, 11, pp. 49–54. doi: 10.1016/j.gdata.2016.11.021.

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. *J. Bacteriol.*, 183(5), pp. 1621–30. doi: 10.1128/JB.183.5.1621-1630.2001.

Shahid, M. A., Marenda, M. S., Markham, P. F. and Noormohammadi, A. H. (2014). Development of an oriC vector for use in *Mycoplasma synoviae*. *J. Microbiol. Methods*, 103, pp. 70–76. doi: 10.1016/j.mimet.2014.05.014.

Shapira, S., Chou, J., Richaud, F. and MJ., C. (1983). New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzymatically active carboxy-terminal portion of beta-galactosidase. *Gene*, pp. 71–82. doi: 6319233.

Shapiro, A. L., Viñuela, E. and V. Maizel, J. (1967). Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem. Biophys. Res. Commun.*, 28(5), pp. 815–820. doi: 10.1016/0006-291X(67)90391-9.

Sharma, S., Citti, C., Sagné, E., Marenda, M. S., Markham, P. F. and Browning, G. F. (2015). Development and host compatibility of plasmids for two important ruminant pathogens, *Mycoplasma bovis* and *Mycoplasma agalactiae*. *PLoS One*. Edited by J. R. Battista, 10(3), p. e0119000. doi: 10.1371/journal.pone.0119000.

Shimizu, T. (2016). Inflammation-inducing factors of Mycoplasma pneumoniae. *Front. Microbiol.*, 7, p. 414. doi: 10.3389/fmicb.2016.00414.

Shimizu, T., Kida, Y. and Kuwano, K. (2007). Triacylated lipoproteins derived from *Mycoplasma pneumoniae* activate nuclear factor-kappaB through toll-like receptors 1 and 2. *Immunology*, 121(4), pp. 473–83. doi: 10.1111/j.1365-2567.2007.02594.x.

Simmons, W. L. and Dybvig, K. (2015). Catalase enhances growth and biofilm production of Mycoplasma pneumoniae. *Curr. Microbiol.*, 71(2), pp. 190–194. doi: 10.1007/s00284-015-0822-x.

Sirand-Pugnet, P., Lartigue, C., Marenda, M., Jacob, D., Barré, A., Barbe, V., Schenowitz, C., Mangenot, S., Couloux, A., Segurens, B., de Daruvar, A., Blanchard, A. and Citti, C. (2007). Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS Genet.*, 3(5), p. e75. doi: 10.1371/journal.pgen.0030075.

Slavec, B., Berčič, R. L., Cizelj, I., Narat, M., Zorman-Rojs, O., Dovč, P. and Benčina, D. (2011). Variation of *vlhA* gene in *Mycoplasma synoviae* clones isolated from chickens. *Avian Pathol.*, 40(5), pp. 481–489. doi: 10.1080/03079457.2011.604840.

Sluijter, M., Estevão, S., Hoogenboezem, T., Hartwig, N. G., van Rossum, A. M. C. and Vink, C. (2012). The

RuvA homologues from *Mycoplasma genitalium* and *Mycoplasma pneumoniae* exhibit unique functional characteristics. *PLoS One*, 7(5). doi: 10.1371/journal.pone.0038301.

Sluijter, M., Hoogenboezem, T., Hartwig, N. G. and Vink, C. (2008). The *Mycoplasma pneumoniae* MPN229 gene encodes a protein that selectively binds single-stranded DNA and stimulates Recombinase A-mediated DNA strand exchange. *BMC Microbiol.*, 8, p. 167. doi: 10.1186/1471-2180-8-167.

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 non-functional orthologue in *Mycoplasma pneumoniae*. *Mol. Microbiol.*, 77(5), pp. 1261–1277. doi: 10.1111/j.1365-2958.2010.07288.x.

Sluijter, M., Spuesens, E. B. M., Hartwig, N. G., Van Rossum, A. M. C. and Vink, C. (2009). The *Mycoplasma pneumoniae* MPN490 and *Mycoplasma genitalium* MG339 genes encode RecA homologs that promote homologous DNA strand exchange. *Infect. Immun.*, 77(11), pp. 4905–4911. doi: 10.1128/IAI.00747-09.

Sulyok, K. M., Kreizinger, Z., Wehmann, E., Lysnyansky, I., Bányai, K., Marton, S., Jerzsele, Á., Rónai, Z., Turcsányi, I., Makrai, L., Jánosi, S., Nagy, S. Á. and Gyuranecz, M. (2016). Mutations associated with decreased susceptibility to seven antimicrobial families in field and laboratory-derived Mycoplasma bovis strains. *Antimicrob. Agents Chemother.*, 61(2), p. AAC.01983-16. doi: 10.1128/AAC.01983-16.

Suzuki-Minakuchi, C., Hirotani, R., Shintani, M., Takeda, T., Takahashi, Y., Matsui, K., Vasileva, D., Yun, C.-S., Okada, K., Yamane, H. and Nojiri, H. (2015). Effects of tThree different nucleoid-associated proteins encoded on IncP-7 plasmid pCAR1 on host *Pseudomonas putida* KT2440. *Appl. Environ. Microbiol.* Edited by M. A. Elliot, 81(8), pp. 2869–2880. doi: 10.1128/AEM.00023-15.

Svenstrup, H. F., Fedder, J., Abraham-Peskir, J., Birkelund, S. and Christiansen, G. (2003). *Mycoplasma genitalium* attaches to human spermatozoa. *Hum. Reprod.*, 18(10), pp. 2103–2109. doi: 10.1093/humrep/deg392.

Svenstrup, H. F., Nielsen, P. K., Drasbek, M., Birkelund, S. and Christiansen, G. (2002). Adhesion and inhibition assay of *Mycoplasma genitalium* and *M. pneumoniae* by immunofluorescence microscopy. *J. Med. Microbiol.*, 51(5), pp. 361–373. doi: 10.1099/0022-1317-51-5-361.

Sweeney, E. L., Dando, S. J., Kallapur, S. G. and Knox, C. L. (2017). The human *Ureaplasma* Species as causative agents of chorioamnionitis. *Clin. Microbiol. Rev.*, 30(1), pp. 349–379. doi: 10.1128/CMR.00091-16.

Szczepanek, S. M., Boccaccio, M., Pflaum, K., Liao, X. and Geary, S. J. (2014). Hydrogen peroxide production from glycerol metabolism is dispensable for virulence of *Mycoplasma gallisepticum* in the tracheas of chickens. *Infect. Immun.*, 82(12), pp. 4915–4920. doi: 10.1128/IAI.02208-14.

Tardy, F., Mick, V., Dordet-Frisoni, E., Marenda, M. S., Sirand-Pugnet, P., Blanchard, A. and Citti, C. (2015). Integrative conjugative elements are widespread in field isolates of Mycoplasma Species pathogenic for ruminants. *Appl. Environ. Microbiol.* Edited by C. A. Elkins, 81(5), pp. 1634–1643. doi: 10.1128/AEM.03723-14.

Taylor-Robinson, D. (1997). Antibiotic susceptibilities of Mycoplasmas and treatment of mycoplasmal infections. *J. Antimicrob. Chemother.*, 40(5), pp. 622–630. doi: 10.1093/jac/40.5.622.

Taylor-Robinson, D., Davies, H. A., Sarathchandra, P. and Furr, P. M. (1991). Intracellular location of Mycoplasmas in cultured cells demonstrated by immunocytochemistry and electron microscopy. *Int. J. Exp. Pathol.*, 72(6), pp. 705–14.

Taylor-Robinson, D., Gilroy, C. B., Horowitz, S. and Horowitz, J. (1994). *Mycoplasma genitalium* in the joints of two patients with arthritis. *Eur. J. Clin. Microbiol. Infect. Dis.*, 13(12), pp. 1066–1069. doi: 10.1007/BF02111830.

Taylor-Robinson, D., Gilroy, C. B. and Keane, F. E. (2003). Detection of several Mycoplasma species at various anatomical sites of homosexual men. *Eur. J. Clin. Microbiol. Infect. Dis.*, 22(5), pp. 291–3. doi: 10.1007/s10096-003-0910-x.

Taylor, P. (1998). Recovery of human Mycoplasmas. in *Mycoplasma Protoc.*, pp. 25–35. doi: 10.1385/0-89603-525-5:25.

Taylor, R. R., Varsani, H. and Miles, R. J. (1994). Alternatives to arginine as energy sources for the non-

fermentative *Mycoplasma gallinarum. FEMS Microbiol. Lett.*, 115(2–3), pp. 163–7.

Teachman, A. M., French, C. T., Yu, H., Simmons, W. L. and Dybvig, K. (2002). Gene transfer in *Mycoplasma pulmonis. J. Bacteriol.*, 184(4), pp. 947–51.

Torres-Puig, S., Broto, A., Querol, E., Piñol, J. and Pich, O. Q. (2015). A novel sigma factor reveals a unique regulon controlling cell-specific recombination in *Mycoplasma genitalium. Nucleic Acids Res.*, 43(10), pp. 4923–4936. doi: 10.1093/nar/gkv422.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U. S. A.*, 76(9), pp. 4350–4.

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. Nat. Commun.*, 8, p. 14665. doi: 10.1038/ncomms14665.

Tsai, S., Wear, D. J., Shih, J. W. and Lo, S.-C. (1995). Mycoplasmas and oncogenesis: persistent infection and multistage malignant transformation. *Proc. Natl. Acad. Sci.*, 92(22), pp. 10197–10201.

Tu, A.-H. T., Voelker, L. L., Shen, X. and Dybvig, K. (2001). Complete Nucleotide Sequence of the Mycoplasma Virus P1 Genome. *Plasmid*, 45(2), pp. 122–126. doi: 10.1006/plas.2000.1501.

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. *J. Infect. Dis.*, 139(4), pp. 478–82.

Tully, J. G., Taylor-Robinson, D., Cole, R. M. and Rose, D. L. (1981). A newly discovered mycoplasma in the human urogenital tract. *Lancet*, (8233). doi: 10.1016/S0140-6736(81)92461-2.

Tully, J. G., Taylor-robinson, D., Rose, D. L., Cole, R. M. and D Joseph M Bove, A. N. (1983). *Mycoplasma genitalium*, a new species from the human urogenital tract. *Int. J. Syst. Bacterology*, 830010(2), pp. 387–396.

Typas, A., Barembruch, C., Possling, A. and Hengge, R. (2007). Stationary phase reorganisation of the *Escherichia coli* transcription machinery by Crl protein, a fine-tuner of σS activity and levels. *EMBO J.*, 26(6), pp. 1569–1578. doi: 10.1038/sj.emboj.7601629.

Ueno, P. M., Timenetsky, J., Centonze, V. E., Wewer, J. J., Cagle, M., Stein, M. A., Krishnan, M. and Baseman, J. B. (2008). Interaction of *Mycoplasma genitalium* with host cells: evidence for nuclear localization. *Microbiology*, 154(10), pp. 3033–3041. doi: 10.1099/mic.0.2008/020735-0.

Uno, M., Deguchi, T., Saito, A., Yasuda, M., Komeda, H. and Kawada, Y. (1997). Prevalence of *Mycoplasma genitalium* in asymptomatic men in Japan. *Int. J. STD AIDS*, 8(4), pp. 259–260. doi: 10.1258/0956462971919877.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, 3(7), p. RESEARCH0034. doi: 10.1186/gb-2002-3-7-research0034.

Vink, C. (2014). Uncoupling of the apyrimidinic/apurinic endonucleolytic and 3'→5' exonucleolytic activities of the Nfo protein of Mycoplasma pneumoniae through mutation of specific amino acid residues. *Microbiology*, 160, pp. 1087–1100. doi: 10.1099/mic.0.077578-0.

Vink, C., Rudenko, G. and Seifert, H. S. (2012). Microbial antigenic variation mediated by homologous DNA recombination. *FEMS Microbiol. Rev.*, pp. 917–948. doi: 10.1111/j.1574-6976.2011.00321.x.

Voelker, L. L. and Dybvig, K. (1999). Sequence analysis of the *Mycoplasma arthritidis* bacteriophage MAV1 genome identifies the putative virulence factor. *Gene*, 233(1–2), pp. 101–7.

Waites, K. B. and Talkington, D. F. (2004). *Mycoplasma pneumoniae* and its role as a human pathogen. *Clin. Microbiol. Rev.*, 17(4), p. 697–728, table of contents. doi: 10.1128/CMR.17.4.697-728.2004.

Wang, J., Karnati, P. K., Takacs, C. M., Kowalski, J. C. and Derbyshire, K. M. (2005). Chromosomal DNA transfer in *Mycobacterium smegmatis* is mechanistically different from classical Hfr chromosomal DNA transfer.

Mol. Microbiol., 58(1), pp. 280–288. doi: 10.1111/j.1365-2958.2005.04824.x.

Weber, S. de S. D. S., Sant'Anna, F. H. and Schrank, I. S. (2012). Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution. *DNA Res.*, 19(2), pp. 103–115. doi: 10.1093/dnares/dsr045.

Weiner, J., Herrmann, R. and Browning, G. F. (2000). Transcription in *Mycoplasma pneumoniae. Nucleic Acids Res.*, 28(22), pp. 4488–96. doi: 10.1093/nar/28.22.4488.

Weiner, J., Zimmerman, C.-U. U., Göhlmann, H. W. H. and Herrmann, R. (2003). Transcription profiles of the bacterium *Mycoplasma pneumoniae* grown at different temperatures. *Nucleic Acids Res.*, 31(21), pp. 6306–6320. doi: 10.1093/nar/gkg841.

Weiss, A. and Shaw, L. N. (2015). Small things considered: the small accessory subunits of RNA polymerase in Gram-positive bacteria. *FEMS Microbiol. Rev.* Edited by G. Wegrzyn, 39(4), pp. 541–554. doi: 10.1093/femsre/fuv005.

Williams, M. V and Pollack, J. D. (1990). A mollicute (Mycoplasma) DNA repair enzyme: purification and characterization of uracil-DNA glycosylase. *J. Bacteriol.*, 172(6), pp. 2979–85.

Winner, F., Rosengarten, R. and Citti, C. (2000). *In vitro* cell invasion of *Mycoplasma gallisepticum. Infect. Immun.*, 68(7), pp. 4238–44.

Witte, G., Hartung, S., Büttner, K. and Hopfner, K.-P. (2008). Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol. Cell*, 30(2), pp. 167–178. doi: 10.1016/j.molcel.2008.02.020.

Wodke, J. A. H., Alibés, A., Cozzuto, L., Hermoso, A., Yus, E., Lluch-Senar, M., Serrano, L. and Roma, G. (2015). MyMpn: a database for the systems biology model organism *Mycoplasma pneumoniae. Nucleic Acids Res.*, 43(Database issue), pp. D618-23. doi: 10.1093/nar/gku1105.

Wodke, J. a H., Pucha, J., Lluch-Senar, M., Marcos, J., Yus, E., Puchałka, J., 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. *Mol. Syst. Biol.*, 9(653), p. 653. doi: 10.1038/msb.2013.6.

Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.*, 51(2), pp. 221–271.

Wood, G. E., Iverson-Cabral, S. L., Patton, D. L., Cummings, P. K., Cosgrove Sweeney, Y. T. and Tottena, P. A. (2013). Persistence, immune response, and antigenic variation of *Mycoplasma genitalium* in an experimentally infected pig-tailed macaque (*Macaca nemestrina*). *Infect. Immun.*, 81(8), pp. 2938–2951. doi: 10.1128/IAI.01322-12.

Wozniak, R. A. and Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.*, 8. doi: 10.1038/nrmicro2382.

Yadav, T., Carrasco, B., Hejna, J., Suzuki, Y., Takeyasu, K. and Alonso, J. C. (2013). *Bacillus subtilis* DprA recruits RecA onto single-stranded DNA and mediates annealing of complementary strands coated by SsbB and SsbA. *J. Biol. Chem.*, 288(31), pp. 22437–22450. doi: 10.1074/jbc.M113.478347.

Yadav, T., Carrasco, B., Myers, A. R., George, N. P., Keck, J. L. and Alonso, J. C. (2012). Genetic recombination in *Bacillus subtilis*: a division of labor between two single-strand DNA-binding proteins. *Nucleic Acids Res.*, 40(12), pp. 5546–5559. doi: 10.1093/nar/gks173.

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. *Mol. Syst. Biol.*, 8, p. 585. doi: 10.1038/msb.2012.16.

Yus, E., Maier, T., Michalodimitrakis, K., van Noort, V., Yamada, T., Chen, W.-H., Wodke, J. A. H., Guell, M., Martinez, S., Bourgeois, R., Kuhner, S., Raineri, E., Letunic, I., Kalinina, O. V., Rode, M., Herrmann, R., Gutierrez-Gallego, R., Russell, R. B., Gavin, A.-C. *et al.* (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science*, 326(5957), pp. 1263–1268. doi: 10.1126/science.1177263.

Zarei, O., Rezania, S. and Mousavi, A. (2013). *Mycoplasma genitalium* and cancer: a brief review. *Asian Pacific J. Cancer Prev.* doi: 10.7314/APJCP.2013.14.6.3425.

Zhang, B., Tsai, S., Shih, J. W., Wear, D. J. and Lo, S. C. (1998). Absence of mycoplasmal gene in malignant mammalian cells transformed by chronic persistent infection of Mycoplasmas. *Proc. Soc. Exp. Biol. Med.*, 218(1), pp. 83–9.

Zhang, N., Wang, R., Li, X., Liu, X., Tang, Z. and Liu, Y. (2014). Are *Ureaplasma spp.* a cause of nongonococcal urethritis? A systematic review and meta-analysis. *PLoS One*. Edited by R. L. Schmidt, 9(12), p. e113771. doi: 10.1371/journal.pone.0113771.

Zhang, S., Tsai, S., Wu, T. T., Li, B., Shih, J. W.-K. and Lo, S.-C. (2004). *Mycoplasma fermentans* infection promotes immortalization of human peripheral blood mononuclear cells in culture. *Blood*, 104(13), pp. 4252–4259. doi: 10.1182/blood-2004-04-1245.

Zhang, W. and Baseman, J. B. (2011a). Transcriptional regulation of MG_149, an osmoinducible lipoprotein gene from *Mycoplasma genitalium*. *Mol. Microbiol.*, 81(2), pp. 327–39. doi: 10.1111/j.1365-2958.2011.07717.x.

Zhang, W. and Baseman, J. B. (2011b). Transcriptional response of *Mycoplasma genitalium* to osmotic stress. *Microbiology*, 157(Pt 2), pp. 548–56. doi: 10.1099/mic.0.043984-0.

Zhang, W. and Baseman, J. B. (2014). Functional characterization of osmotically inducible protein C (MG_427) from *Mycoplasma genitalium. J. Bacteriol.*, 196(5), pp. 1012–1019. doi: 10.1128/JB.00954-13.

Zhang, Y., Lin, J. and Gao, Y. (2012). *In silico* identification of a multi-functional regulatory protein involved in Holliday junction resolution in bacteria. *BMC Syst. Biol.*, 6(Suppl 1), p. S20. doi: 10.1186/1752-0509-6-S1-S20.

Zhao, G., Zhang, H., Chen, X., Zhu, X., Guo, Y., He, C., Anwar Khan, F., Chen, Y., Hu, C., Chen, H. and Guo, A. (2017). *Mycoplasma bovis* NADH oxidase functions as both a NADH oxidizing and O2 reducing enzyme and an adhesin. *Sci. Rep.*, 7(1), p. 44. doi: 10.1038/s41598-017-00121-y.

Zuber, P. (2013). *Function and control of the Spx-family of proteins within the bacterial stress response*.