WP4 - Tools for Genome Engineering

MycoSynVac

WP4 report of Deliverable

Deliverable D4.1

Tools to Genome Engineering

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634942.

Project Acronym:	MycoSynVac
Project Title:	Engineering of *Mycoplasma pneumoniae* as a broad-spectrum animal vaccine
Grant Agreement no.:	634942
Project Duration:	1 April 2015 - 31 March 2020
Work Package no.:	WP4
Deliverable Number:	D4.1
Deliverable Title:	Tools to Genome Engineering
Lead Beneficiary:	Partner 3 - INRA; Contributors: Partner 1 - CRG
Nature:	R (Report)
Dissemination Level:	PU (Public)
Deliv. Date (Annex I):	1 October 2015

Version	Contributors	Comments
1.0	INRA, CRG	Submission to Project Officer for informal review
1.1	CRG	Submission to Project Officer for informal review of second round
1.2	CRG	Online submission

Table of Contents

1. Introduction

There is a general lack of genetic tools to modify the genome of *Mycoplasma* species. This is true for the specific case of *M. pneumoniae* and several factors that hamper the development of these tools have been identified (Halbedel and Stülke, 2007).

Within this objective of developing tools to genome engineering, there are two approaches.

The first one is to find innovative ways to modify the *M. pneumoniae* genome using different genetic tools. These tools will be used to produce large scale modifications (deletions, mutations and insertions) as efficiently as possible. This genetic tools could either be applied directly on the mycoplasma cell (“mycoplasma option” mostly developed by CRG), or to the mycoplasma genome maintained in yeast (“yeast option”, mostly developed by INRA). The need to work on both approaches in parallel is justified because the yeast option requires that we succeed in transplanting back the modified genome into a mycoplasma recipient cell, something that we have not yet achieved (see below).

The second objective is to develop methods for transplanting *M. pneumoniae* modified genome into a recipient mycoplasma cell. Although such an approach has been described for mycoplasmas belonging to the Spiroplasma group, it has never been performed in the Pneumoniae group that includes *M. pneumoniae.* In fact, JCVI has reported that attempts to transplant the *M. genitalium* genome (this mycoplasma is the species

that is the closest to *M. pneumoniae* among described species) into a recipient cell failed. To circumvent these initial failures, the reasons for this blockage need to be identified and new transplantation methods that take into account these parameters need to be developed.

2. *M. pneumoniae* Genome Engineering

The CRISPR (clustered regularly interspaced short palindromic repeats) and Cas9 protein are the two components of the "prokaryote immune system" that allows bacteria to eliminate exogenous DNA introduced by phages (Figure 1).

In these last years, the CRISPR-Cas system has been extensively used for genome engineering of several eukaryotic organisms including human, mouse, yeast and plants (Doubna and Charpentier, 2014). The type of modification ranges from small mutations up to gene deletion (Torres-Ruiz and Rodriguez-Perales, 2015). The prototype of the CRISPR-Cas system is that of the Gram-positive bacterium *Streptococcus pyogenes*. Cas9 nuclease from *S. pyogenes*, is a RNA-guided endonuclease that catalyzes site-specific cleavage of double stranded DNA using a template RNA (gRNA) that is in part complementary to the DNA to be cleaved. The location of the break is within the target sequence 3 bases from the NGG PAM (Protospacer Adjacent Motif). The PAM sequence, NGG, must follow the targeted region on the opposite strand of the DNA with respect to the region complementary crRNA sequence.

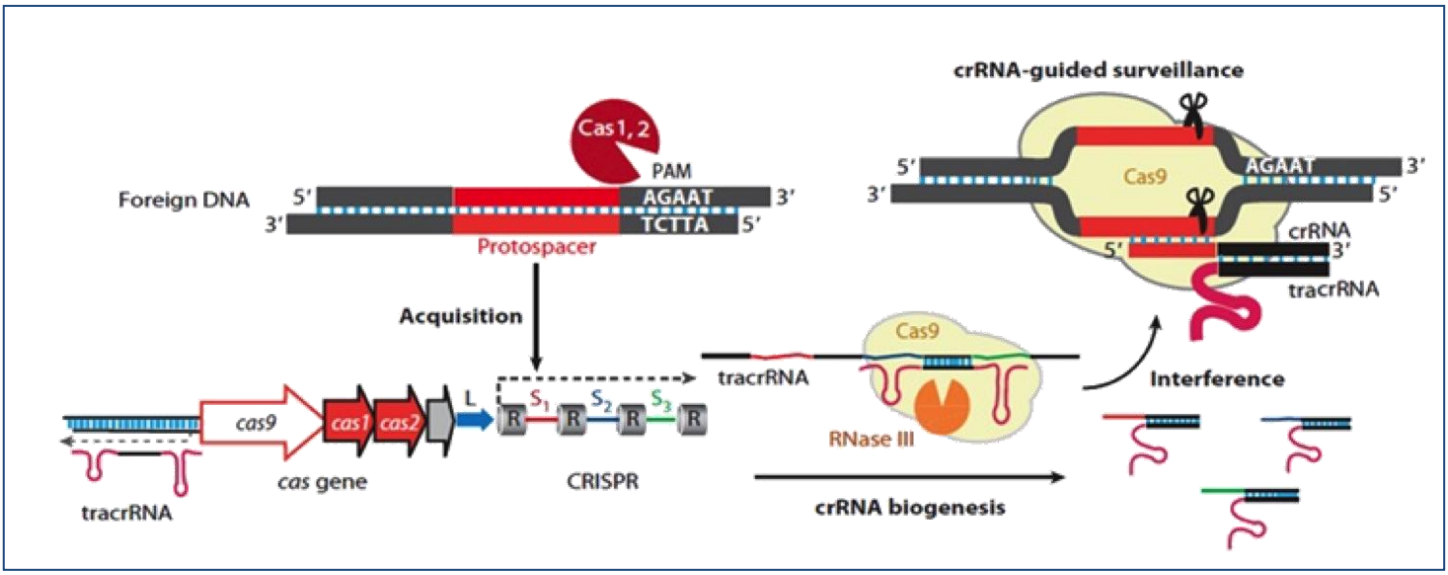

Figure 1. Schematic representation of the CRISPR-CAS9 system. Adapted from Sorek et al. 2013.

In MYCOSYNVAC, our idea is to implement this CRISPR-Cas system in *M. pneumoniae* (to directly delete the genes in *M. pneumoniae* or “mycoplasma option”; to be done by CRG) or in yeast (to engineer the genome of *M. pneumoniae* introduced in a yeast strain or “yeast option”; to be done by INRA). In the “mycoplasma option”, other genetic tools including the CRE-loxP and the Lambda-RED have been also evaluated.

2.1. *M. pneumoniae* genome engineering using the “mycoplasma option”

2.1.1. Using the CRISPR-Cas9 system

Different versions of the Cas9 gene of *S. pyogenes* were cloned in the transposon vector miniTn4001Tet. Different vectors were obtained in *E. coli* to ensure the efficient expression of Cas9 protein in *M. pneumoniae*: i) TnCas9: Cas9 of *S. pyogenes* ii) TnCas9strep: Cas9 of *S. pyogenes* with a Strep tag iii) TnMP200Cas9: Cas9 with a MP200 sequence at the Nter (MP200Cas9). This sequence was shown at **CRG** to enhance expression of the fused protein in *M. pneumoniae* and iv) TnMP200Cas9Strep: MP200Cas9 with a Strep-tag. The four resulting plasmids were transformed in *M. pneumoniae* and Cas9-protein expression was tested by Western Blot (Figure 2).

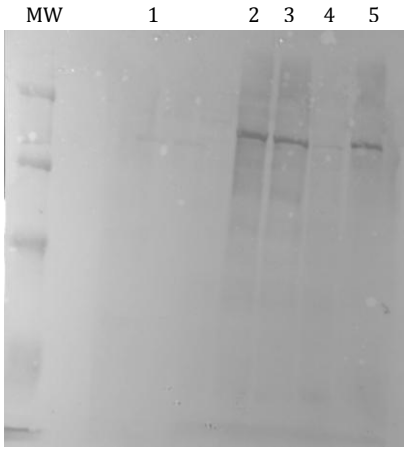

Figure 2. ***Western Blot to detect Cas9 protein***. The protein was detected by using a specific antibody against Cas9 protein. Lane 1: Protein extract of WT *M. pneumoniae* strain; Lane 2:TnCas9 transformed strain; Lane 3: TnCas9Strep transformed strain; Lane 4: TnMP200Cas9 transformed strain. Lane 5: TnMP200Cas9Strep transformed strain.

As shown in lanes 2,3 and 5 the protein is detected in the transformed strains but not in the WT strain (lane 1) confirming that the Cas9 protein is properly expressed in the mycoplasma.

Based on these positive results we built another mini-transposon vector containing the gRNA designed to target the P1 protein of *M. pneumoniae* (MPN142; one of the main virulence factors). To determine if the transcription terminator associated with the gRNA is functional in *M. pneumoniae*, we designed a cassette. This cassette is introduced after the transcription terminator containing a ribosome-binding-site (RBS) and the opening reading frame (ORF) coding for mCherry red fluorescent protein (Figures 3 and 4; Table 1).

All gRNA tested have the following general structure, which is encoded in a fragment cloned with *EcoR*I-*Not*I restriction sites inside a pMTnPar plasmid and introduced by transposition into *M. pneumoniae* M129 cells.

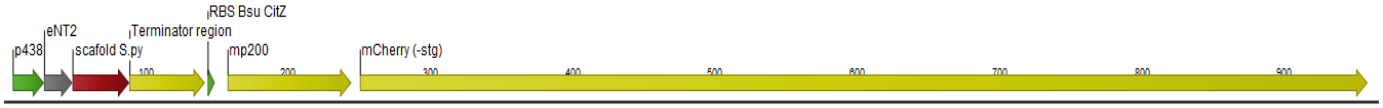

Figure 3. General structure of the RNA cassette tested: promoter p438 + sgRNA + RBS + mp200::mCherry.

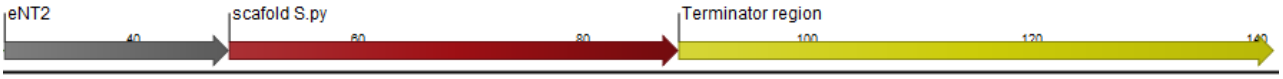

Figure 4. General structure of gRNA tested: protospacer (20nt)+scaffold+terminator region

Table 1. Terminator sequences used for the expression of gRNAs

Vector with gRNA	Terminator seq	Clc file
1 (*S. pyogenes*)	5'-CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT	pMTnGm_P438-sgRNA_eNT2_a+Reporter
2 (*E. coli*)	5'-AGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCG	pMTnGm_P438-sgRNA_eNT2_b+Reporter
3 (*M. pneumoniae*)	5'-GTCAAGAAGATAAAAGAGACTTAGGCCTATGCCTAGGTCTTTTTTATTGTTTT	pMTnGm_P438-sgRNA_eNT2_c+Reporter
4 (No terminator)	5'-CAAGAACGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT	pMTnGm_P438-sgRNA_eNT2_d+Reporter

5 (no sgRNA)	-	pMTnGm_P438+Reporter_ CONTROL

Five different vectors were obtained containing the gRNA under the P438 promoter with each vector having a different terminator sequence followed by the venus fluorescent protein gene (Table 1). After transforming *M. pneumoniae,* the resulting strains were observed by fluorescence microscopy to evaluate the efficiency of different terminator sequences (Figure 5).

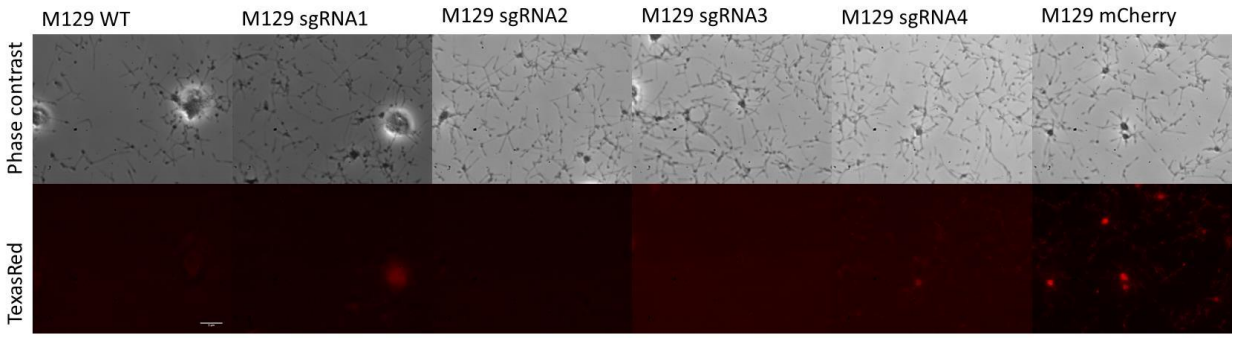

Figure 5. Fluorescence microscopy of different *M. pneumoniae* strains transformed with different constructs.

If the terminator works properly, the cells should not be fluorescent, otherwise transcription will not stop at the terminator sequence and the mCherry will be expressed and cells will be fluorescent. The gRNA followed by a sequence that is not a terminator was used as a control (M129 mCherry sample in Figure 5). In this control all the transformed cells were fluorescent, confirming that the reporter system was working. No fluorescent signal was obtained when different terminator sequences were tested (M129sgRNA1 to M129gRNA4 in Figure 5). This result confirmed that the gRNAs can be expressed in *M. pneumoniae* with the expected length.

The double strand cut produced by the designed CRISP9 system in the genome of *M. pneumoniae* should result in a frameshift in the MPN142 protein upon DNA *in vivo* ligation. Since MPN142 protein is an adhesin, we would expect cells to detach as a result of the cleavage and DNA repair. Also, the gene mutation can be detected by using the Surveyor Mutation Detection kit (IDT Technologies). Unfortunately, the DNA frameshift was not detected which could suggest that the CRISPR-Cas9 system was not working.

The CRISPR-Cas9 is dependent of the DNA repair system that is quite reduced in term of components in *M. pneumoniae.* It is possible that DNA end-joining repair system does not work well in this species and the CRISP9 DNA-cleavage kills the cells. In this case, the combination Cas9 protein with the gRNA will be toxic, and only insertions of the second mini-transposon with the gRNA in the Cas9 gene would be tolerated. This

would promote the inactivation of the Cas9 protein and then it would explain why we do not detect the cut. We did a western blot of the strain transformed with both transposons (Cas9 and gRNA) and this supports the idea that the Cas9 protein is not expressed (Figure 6).

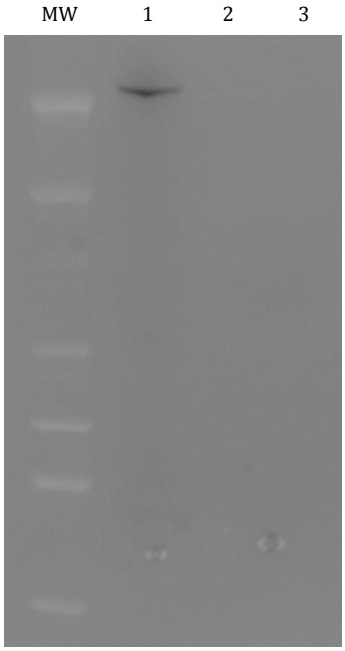

Figure 6. Western Blot to detect the Cas9 protein in *M. pneumoniae* strains. Lane 1: TnCas9 strain, Lane2: M129 strain and Lane 3; gRNA1 strain. It is shown that in the strain containing both mini-transposons (lane 3) the protein is not detected.

Genomic DNAs of the viable cells were extracted and they are currently being sequenced by Illumina MiSeq. This will allow to corroborate that the insertions of the mini-transposon containing the gRNA are in the Cas 9 protein-encoding gene.

These results suggested to the CRG team a new strategy for site directed mutagenesis in *M. pneumoniae* that combines transposon mutagenesis with the Cas9 system. Minitranspon libraries generate randomly insertions in all non-essential regions of the genome, resulting in a mixture of clones. Isolating these clones is a tedious work since it implies filtering and several steps of spreading of the bacteria on agar plates followed by the sequencing of isolated clones. Furthermore *M. pneumoniae* strain aggregates so it is very difficult to isolate and obtain a pure clone. A method that would allow isolating the clone of interest would considerably facilitate obtaining *M. pneumoniae* mutants.

The idea consists in generating a library of transposon mutants and then transform with a vector that contains the Cas9 and the gRNA. The gRNA will target the region where we want to have the transposon insertion. The vector will be a suicide vector (the Cas9 will be lost after several rounds of division). Only the cells having the transposon insertion in the desired site will be viable since the gRNA won't be able to recognize the PAM sequence and the Cas9 will not cut. The cells that have the insertions in other regions of the genome will be eliminated by the system, since their DNA will be cut and not repaired. At the **CRG**, we are currently evaluating the efficiency of this system. Also we are planning to combine this technique with the Cre-lox system (see below) to perform large targeted deletions of the genome.

2.1.2. Using the *CRE-LOX* system

Cre-Lox recombination is known as a site-specific recombinase technology, and is widely used to carry out deletions, insertions, translocations and inversions at specific sites in the DNA of cells. It allows DNA modification to be targeted to a specific cell type or be triggered by a specific external stimulus. It is implemented both in eukaryotic and prokaryotic systems.

The idea of MycoSynVac project is obtaining a strain that does not have any antibiotic resistance markers in the genome. Since insertions, deletions and site directed mutagenesis normally requires introduction of antibiotic resistance and since we need to do several rounds of mutagenesis, we need to develop a system to remove antibiotic markers at each step. It is for this purpose that we develop the Cre-Lox system in *M. pneumoniae*.

To use the Cre-Lox system, we need a transitory expression to allow for multiple rounds of mutagenesis. Thus, it is necessary to implement an induction system for the expression of the Cre, or alternatively, design a suicide vector that is lost after several culture passages.

Since suicide vectors will be required also for the expression of the Cas9 protein (see above) we decided to focus in the design of two alternative vectors. One of the vectors was supplied by a UAB collaborator (Molecular biology group; Figure 7). It is a pBSK vector backbone with the Cre gene and the gentamycin resistance marker gene. At the CRG we have designed the other vector where the gentamycin marker is replaced by puromycin resistance gene. These two vectors would be compared for their stability and maintenance in *M. pneumoniae* transformed cells.

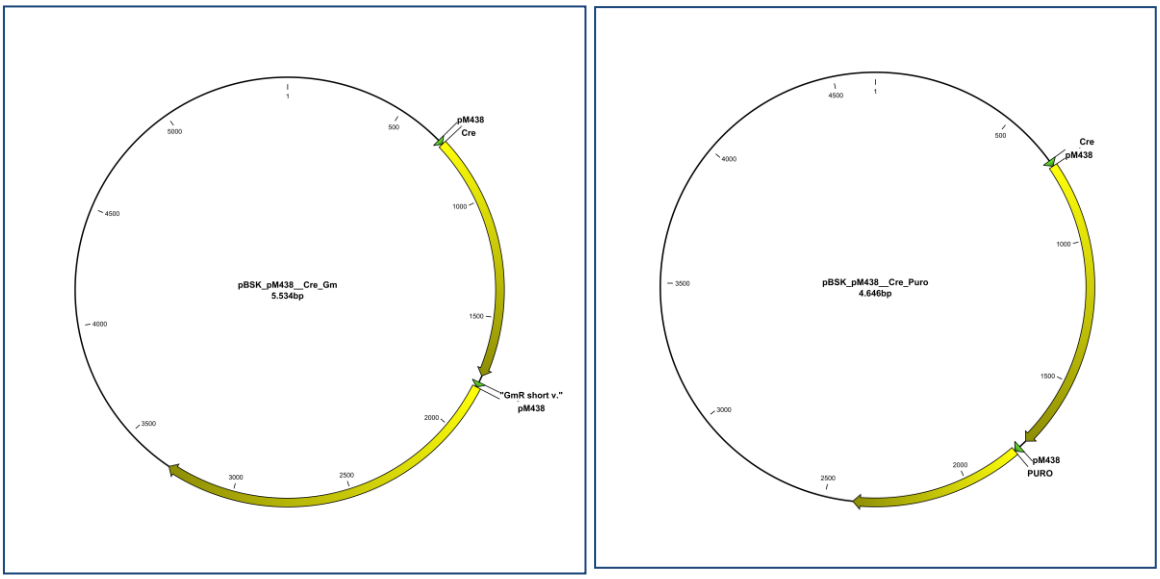

Figure 7. Vectors maps of obtained suicide vectors.

In addition, a mini-transposon vector containing the cloramphenicol (Cm) resistance marker flanked by two lox sites was cloned in *E. coli* (Figure 8) and then transformed in *M. pneumoniae* (LoxCmMP strain). After selection with chloramphenicol, the mycoplasma cells were transformed with a suicide vector containing the Cre protein.

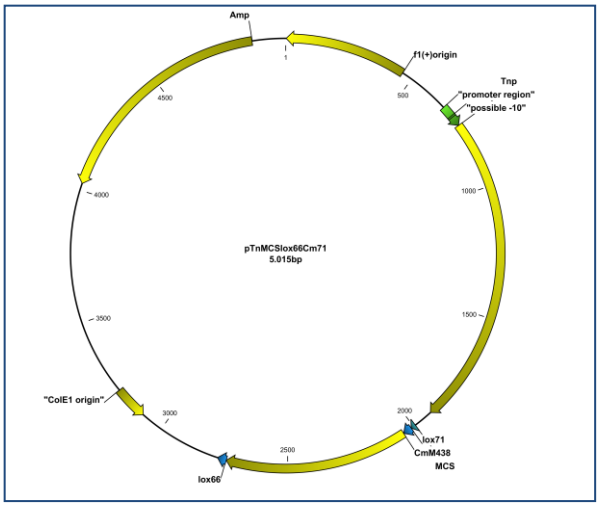

Figure 8. Vector to include the Cm flanked by Lox sites in *M. pneumoniae* genome.

M. pneumoniae M129 and LoxCmMP strains were transformed with the two vectors resulting in four strains. Plasmid maintenance and Cre activity were tested by following growth curves in the presence of different antibiotics (Figure 9).

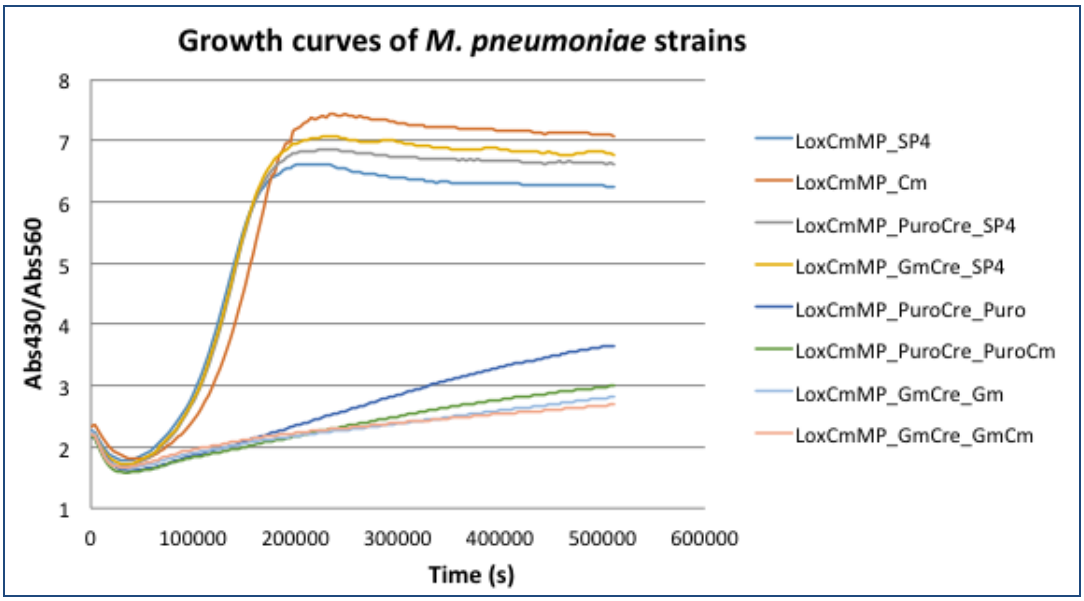

Figure 9. Growth curves of *M. pneumoniae* strains in presence of different antibiotics

The growth curves of the LoxCmMP, LoxCmMP-GmCre (LoxCmMP transformed with the suicide CreGm vector) and LoxCmMP-PuroCre (LoxCmMP transformed with the CrePuro vector) showed that the suicide vector, despite being lost in the cell division process, allows the expression of Cre protein. Expression of the Cre protein results in the deletion of the Cm resistance marker included in the genome. We also find that the vector with the puromycin resistance is more stable than the gentamycin vector. This increased stability results in a better expression of the Cre protein and a more efficient deletion of the resistance marker.

At the moment, isolated clones are being sequenced to confirm the excision of the resistance marker. Also, the suicide vector with the puromycin resistance marker will be used as template to obtain a suicide vector for the expression of the Cas9 and guided RNA as discussed above.

2.1.3. Using the LAMBDA-RED system

Lambda-Red is a system composed of three different proteins that the bacteriophage Lambda employs to integrate its genome inside the genome of the bacterial host. This system has been widely employed with

biotechnological purposes given its ability to generate precise insertions, deletions and point mutations in the genome of different bacterial strains. It is in principle independent of the host recombination machinery and can precisely recombine single stranded oligonucleotides, or double-stranded DNA cassettes at specific points of the chromosome. The specificity in these genome modifications is provided by the homology regions that should be included in the design of exogenous recombination substrates (dsDNA or ssDNA) that could be as little as 35 pb, although the required length of the homology region is variable among different bacterial strains.

As mentioned above, the system is composed of three different proteins, Gam prevents the degradation of linear DNA by endogenous nucleases such as RecBCD and SbcCD, whereas Exo degrades the recombinogenic dsDNA substrate in a 5' to 3' manner leaving single stranded DNA (ssDNA) as substrate for recombination. Lastly, Beta binds to the recombinogenic ssDNA (either produced by Exo or provided as oligonucleotide) and mediates recombination with the chromosomal homology target on the lagging strand of the replication fork. Thus, lambda-red system promotes recombination by including the recombination substrate as an Okazaki fragment during the replication, which explains its independency from host recombination machinery, and its ability to mediate recombination in different bacterial strains.

In order to perform specific deletions into the chromosome of *M. pneumoniae* we constructed different transposons in which the three proteins of the lambda-red system were included with different regulatory regions (Figure 10).

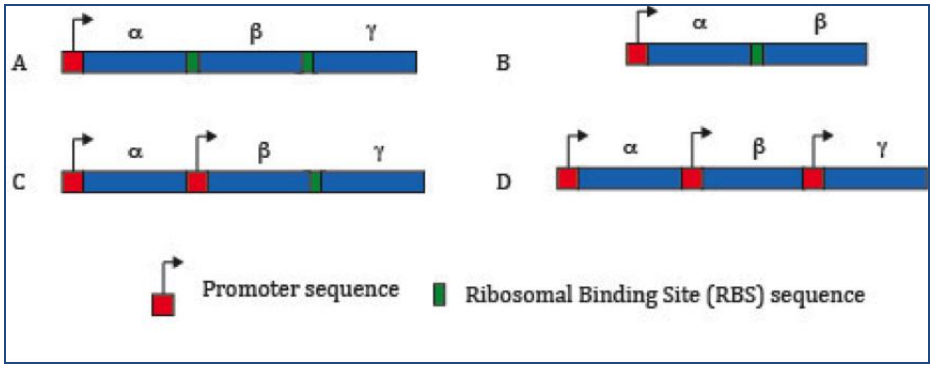

Figure 10. Different lambda-red constructs tested for *M. pneumoniae* genome edition

In addition, we constructed a suicide (i.e. non-replicative) vector in which a tetracycline resistance cassette was flanked by DNA regions homologous to the regions that are flanking the gene coding for P1 adhesin in *M. pneumoniae* genome (Figure 11)

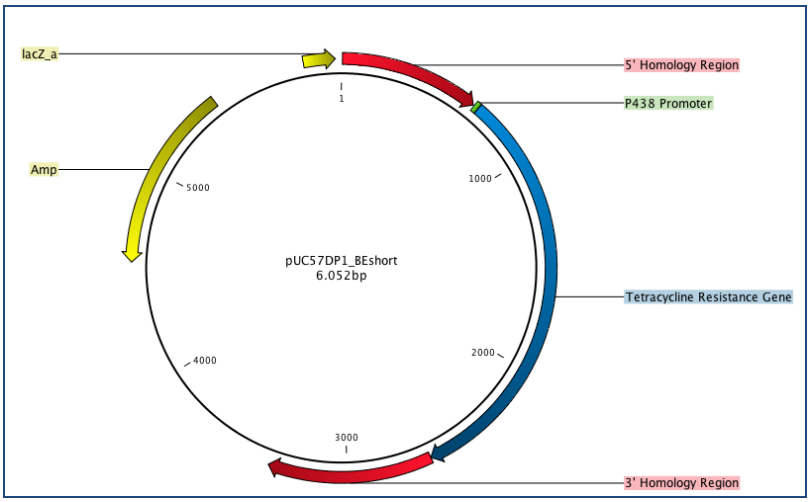

Figure 11. Vector map of the suicide plasmid to delete the P1-adhesin gene.

After electroporation of this suicide vector in *M. pneumoniae* strains expressing the different lambda red system constructs, we were able to find seven tetracycline resistant colonies in the strain expressing lambda red construct C.

Preliminary PCR results showed that three of these colonies were double recombinants (i.e. lacking P1 adhesin) although it seems that the colonies should be truly isolated with a syringe since a PCR product corresponding to the expected size of non-recombinants was also found. Anyway, this experiment is the proof of concept that lambda red system could be useful for *M. pneumoniae* genome edition, though additional improvements such as flanking the antibiotic resistance gene with lox sites should be done.

2.2. *M. pneumoniae* genome engineering using the "yeast option"

During the past months, we have demonstrated that CRISPR/Cas9 can be used for one-step and seamless deletion of a complete gene within the genome of *M. mycoides* subsp. *capri* (Mmc) cloned in yeast. This work started in October 2014 with the recruitement of a PhD student, Iason Tsarmpopoulos, and was pursued in collaboration with Sanjay Vashee (JCVI, US) and Joerg Jores (ILRI, Kenya) in the frame of a project granted by the US National Science Foundation [grant number IOS-1110151]. This work is in fact a proof of concept using the genome of *M. mycoides* subsp. *capri* (Mmc) cloned in yeast as a model; it now can be applied right away to *M. pneumoniae* (see below). Only a brief description of the work is provided here, a paper describing this study will soon be submitted for publication. The MYCOSYNVAC grant as partial financial support is mentioned in the acknowledgement section (see Appendix).

Recently, CRISPR/Cas9 tools have been adapted to yeast. Using a two-plasmid system, DiCarlo *et al.* (2013) reported targeted gene mutagenesis in *Saccharomyces cerevisiae* with efficiency rates close to 100%. We have adapted this CRISPR/Cas9 tool for an efficient one-step seamless deletion of genes in a mycoplasma genome cloned in yeast. The first step was to modify the plasmid used by DiCarlo et al to introduce AarI restriction sites on each side of the sequence that is going to provide the guide RNA for targeting the action of the Cas9 nuclease (Figure 12). This modification of the original construct allows changing quickly the target gene. As a proof of concept, we targeted the *glpO* gene of MmC, which encodes a major virulence factor of this mycoplasma. However, the strategy not only involved the inactivation of the gene, which would only require a short mutation but the complete, seamless deletion of the entire *glpO* gene (about 1,100 bp). In order to achieve this result, in the phase II (Figure 12 C) of the DNA repair a sequence (recombination template) overlapping the extremities of the target gene was also provided during transformation of yeast.

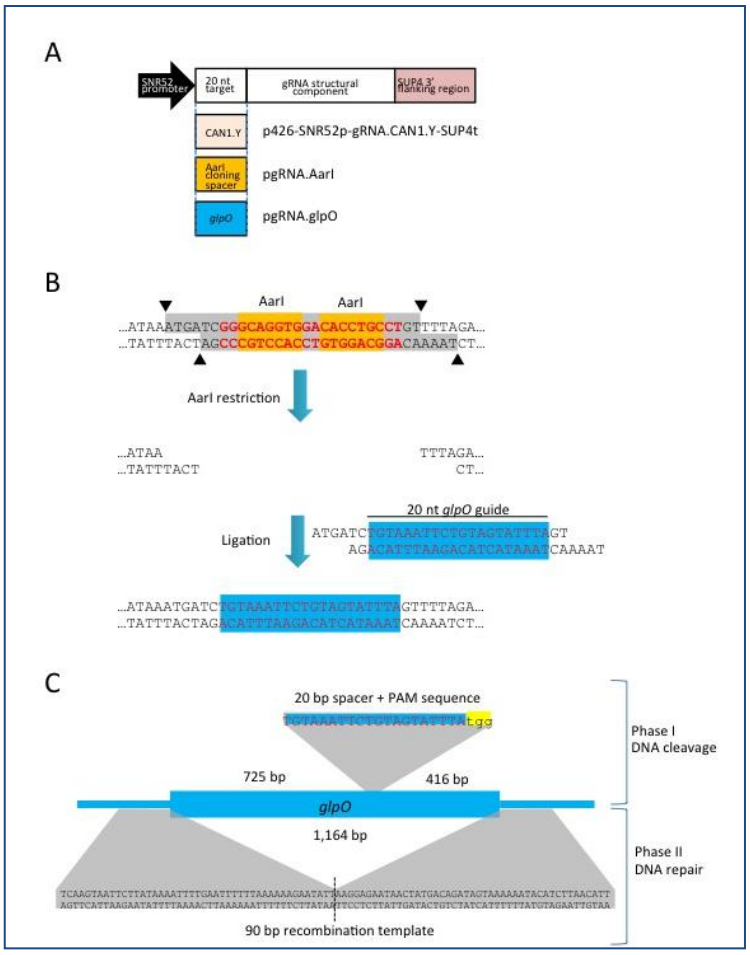

Figure 12. Construction of gRNA expression vector and glpO deletion design. A. Design of gRNA expression constructs. Expression of chimeric gRNA is controled by snoRNA SNR52 promoter and terminator from the 3' region of the yeast SUP4 gene. The CAN1.Y target 20 nt sequence from the original plasmid from DiCarlo et al. was replaced to generate the other plasmids B. Schematic for seamless cloning of the *glpO* guide sequence oligonucleotides into the customized p426-SNR52p-AarI-SUP4t plasmid containing the expression cassette for the gRNA. The type IIS AarI restriction enzyme recognition and cleavage sites are indicated in orange and by arrowheads, respectively. The *glpO* guide oligonucleotides are annealed and contain overhangs for ligation into the pair of AarI sites in pgRNA.AarI. C. Localization of the 20 nt-guide sequence within *glpO* gene where the DNA is first cleaved. Adjacent PAM sequence tgg is highlighted in yellow. Sequence of the 90 bp-recombination template, required for subsequent DNA repair, prepared from two complementary oligonucleotides with 45 bp-arms designed against *glpO* adjacent genomic sequences.

After selection of yeast transformants, DNA extraction was performed on 12 pools of 20 colonies each followed by a PCR screening with primers located on both sides of the *glpO* gene (Figure 13). Eight pools tested showed a 483 bp amplification product corresponding to Δ*glpO* mutants, in addition to the 1,640 bp amplification product corresponding to the wild-type genomic structure. Two positive pools were selected and individual clones were screened to isolate mutants using the same process. A single 483 bp-amplified product indicating that the *glpO* gene had been removed was observed for 4 and 5 individual clones within pools P7 and P8, respectively. Mixed profiles were observed for some other clones suggesting individual clones were not pure. Similar results were obtained in two replicates of the experiment. The seamless deletion of *glpO* was confirmed in three clones by sequencing of the PCR products, indicating that homologous recombination had occurred as expected.

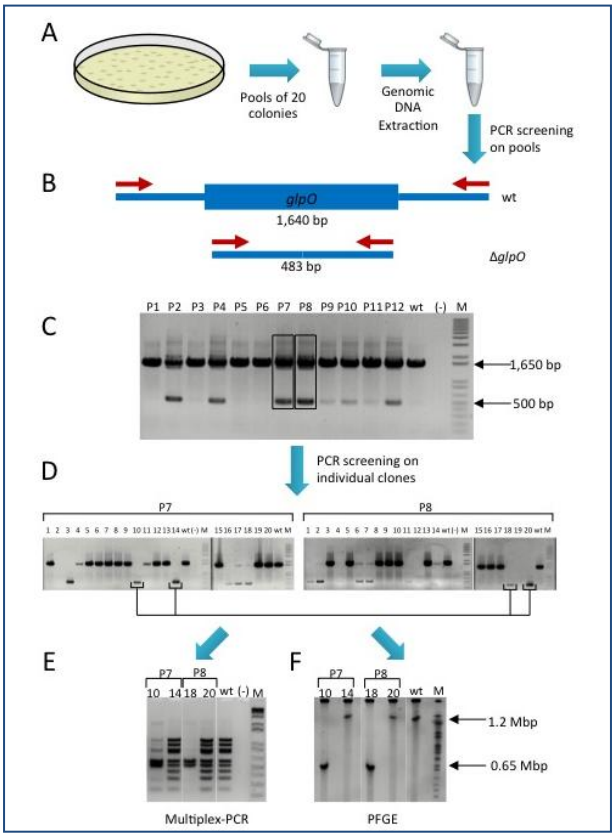

Figure 13. Screening yeast with glpO-deleted mycoplasma genomes. A. Genomic DNA from pools of 20 yeast co-transformed with the p426-SNR52p-gRNA.*glpO*-SUP4t plasmid and recombination template was extracted for PCR screening of *glpO* deletion. B. Diagram of the expected *glpO* region in Mmc (wt) and *glpO*-deleted mutants (Δ*glpO*). Lengths of PCR products are indicated. C. Example of gel electrophoresis of PCR products. Pools with bands of about 500 bp indicated the presence of Δ*glpO* mutants. D. Gel electrophoresis of PCR

products obtained from individual clones present in positive pools 7 and 8. E. Gel electrophoresis of multiplex PCR to check mycoplasma genome integrity of mutants P7.10, P7.14, P8.18 and P8.20. M, 100 bp-ladder (Promega); wt, positive control DNA from Mmc; (-), H2O negative control. F. PFGE analysis of mutants P7.10, P7.14, P8.18 and P8.20 after PspXI restriction; M, CHEF *S. cerevisiae* chromosomal DNA (Biorad); wt, positive control from wt Mmc.

The integrity of the genome of selected yeast clones was verified by multiplex PCR (Figure 13E) and by PFGE (Figure 13F).

In addition, the mycoplasma genomes from selected yeast clones were back-transplanted into mycoplasma recipient cells. Transplants were obtained after 5 days for both yeast clones, demonstrating the viability of the resulting recombinants. A final PCR test was conducted on three clones to confirm that the *glpO* gene was actually deleted in the transplanted mycoplasmas. All three showed the expected PCR product, demonstrating the deletion of the *glpO* gene. Since the GlpO enzyme is responsible for H_2O_2 production as a by-product of the glycerol metabolism, the impact of *glpO* deletion on the production of H_2O_2 in presence of glycerol was investigated as described previously. A concentration of 5-10 $mg.L^{-1}$ of H_2O_2 was measured after a 100 min of incubation of wt Mmc with 100 µM glycerol, no H_2O_2 production could be detected for the three clones were *glpO* had been deleted. This result was in complete accordance with previous studies showing that *glpO* is directly responsible for the release of highly cytotoxic H_2O_2 by mycoplasmas and that our strategy has resulted in the full inactivation of this enzymatic activity, as expected.

The main advantages of the method compared to previously methods used for gene deletion in yeast are the following:

- This is a one step method that does not require selection marker because of its high efficiency ;
- It is much faster than previously-described method such as (2 weeks versus about 4 weeks for TREC method) ;
- This method offers the possibility to inactivate in one experiment multiple targets.

Our goal is now to apply the system developed in our lab for the engineering of *M. pneumoniae* genome already cloned in yeast. We do not expect any specific difficulty because it is highly similar that what we just performed using the Mmc genome. We have chosen to delete the gene MPN142 that encodes a member of the RepMP5 family proteins and that is known as a virulence factor involved in the adhesion of *M. pneumoniae* to the lung cells. The genetic constructions have been obtained and the experiment is under way. Once again,

the interest of this approach is highly dependent on our ability to back-transplant the *M. pneumoniae* into recipient mycoplasma cells. This is why we work in parallel on the transplantation of *M. pneumoniae* genomes.

3. Transplantation of the *M. pneumoniae* genome into recipient mycoplasma cells

Building upon existing experience on mycoplasma genome transplantation (Lartigue et al 2007 and 2009) we started to develop protocols for genome transplantation (GT) in *M. pneumoniae*.

3.1. Choice of the recipient species for genome transplantation assays

A species phylogenetically close to *M. pneumoniae* is probably the best candidate to be used as recipient species during GT experiments because there is a need for compatibility between the incoming genome and the proteins/enzymes already present in the recipient cell. Several strains/species could be selected (Table 2), and according to this choice the transplantation would be either intra- or inter-species.

Table 2. Potential recipient mycoplasma species

➢ ***Mycoplasma pneumoniae M129 (ATCC 29342)***	***Intra-species genome transplantation assays***
➢ ***Mycoplasma pneumoniae M129-B170 (ATCC29343),*** an avirulent/nonadherent mutant, derived from M129-B7 strain	
➢ ***Mycoplasma pneumoniae FH (ATCC 1553)***	
➢ ***Mycoplasma genitalium G37 (NCTC 10195) or other strains***	***Inter-species genome transplantation assays***
➢ ***Mycoplasma gallisepticum***	

Recent data from INRA indicate that “intra-species GT” would be the best option to succeed a genome transplantation of *M. pneumoniae* genome. Indeed, genome transplantation has been only achieved so far within the Spiroplasma phylogenetic group using *M. capricolum* as recipient cell and INRA team has demonstrated that the donor and the recipient cell should share above 90% similarity of the *Mollicutes* core

proteome (unpublished data). All our attempts to choose another recipient cell with a lower similarity of core proteome have failed. Therefore, we calculated the core proteome similarity for the pairs of *Mycoplasma* species that are close to *M. pneumonaie* (Table 3). The inter-species transplantation using *M. genitalium* as a recipient cell would be close to this limit (Table 3) and transplantation using *M. gallisepticum* would be below this lower limit of "compatibility". If *M. genitalium* could be evaluated, our experience has shown that with this low level of similarity, the transplantation efficiency is very low, which would be extremely risky to follow.

Table 3. Average similarity calculated for 94 core proteins that are shared by most mollicutes

	Mean Core proteome similarity*
M. pneumoniae M129/ M. pneumoniae FH	Close to 100 %
M. pneumoniae M129 /M. genitalium	91,36 %
M. pneumoniae M129/M. gallisepticum	80,32 %

However, the intra-species genome transplantation offers the difficulty to differentiate the genomes from the donor and from the recipient cells (and also possible recombination products between the incoming genome and the genome from the recipient cell). We also know from previous CRG data that at least the two M129 and M129-B170 strains express a restriction nuclease that needs to be taken into account; such an enzymatic activity will block each attempts of genome transplantation if the incoming genome is not modified with the corresponding methylase. This nuclease activity is not known for the FH strain. In addition for all the *M. pneumoniae* strains, there is the possibility that membrane nucleases could also represent a difficulty in these experiments.

3.2. Genome transplantation assays

Genome transplantation experiments using tetM-marked *M.pneumoniae* genomes as donor genomes and *Mycoplasma pneumoniae M129* as recipient cells have started. In a first set of experiments performed at CRG in a collaboration with INRA, protocols already established for other *Mycoplasma* species were used, including a PEG-70% based protocol formerly developed for *M. leachii* (unpublished) and a PEG-5% based protocol

formerly developed for *M. capricolum* (Lartigue et al., 2009). All the preliminary attempts failed. While those results are not totally surprising because similar results have been reported at JCVI, they indicate that a protocol entirely adapted to *M. pneumoniae* species should be developed. Several parameters should be tested including the PEG percentage and molecular weight, wash buffers, recipient cells growth temperature and/or growth medium.

To be more efficient in the development of these protocols, efficient molecular tools such as *oriC* plasmids are required. Indeed, such tools will allow not only to evaluate the efficacy of the transplantation process but also to considerably increase the number of experiments performed by week. Plasmids are easier to produce in high quantity than intact genomes isolated from *M. pneumoniae* cells. A similar strategy was also followed with mycoplasmas for which a transplantation scheme has been successfully developed.

3.2.1 OriC-based protocols

Efficient replicable oriC plasmids has been reported for *M. gallisepticum*, using the *oriC* region upstream from the *soj* gene (Lee et al., 2008) (Figure 14). Based on this study, **INRA** team constructed two *oriC* plasmids for *M. pneumoniae.* The first plasmid, similarly to *M. gallisepticum*, contains only the intergenic region between *soj* and *dnaN* (~740pb). The second one contains a larger region going from *dnaA* to *dnaN* genes *(~4037bp)* (Figure 14 and 15).

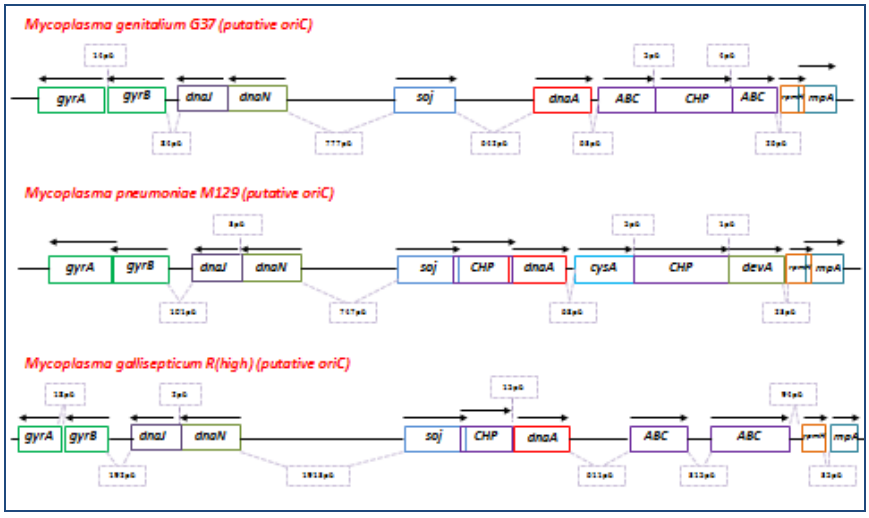

Figure 14. Comparison of chromosomal origin of replication region from mycoplasmas closely related to *M. pneumoniae*

A notable difference between *M. gallisepticum* and *M. pneumoniae oriC* is the presence of numerous CTAT methylation sites in the soj-dnaN intergenic region of *M. pneumoniae* genome (Lluch-Senar et al; 2013) (Figure

15). The methyltransferase responsible for this methylation pattern has been identified (Mpn198) during the characterization of *M. pneumoniae* methylome. Mpn 198 is encoded by an orphan gene that is essential for *M. pneumoniae* growth. No restriction enzyme has been associated to this type II methyltransferase, reinforcing a potential key role in *M. pneumoniae* cell cycle regulation. We are currently producing Mpn198 as a recombinant protein in *E. coli*.

Once produced, oriC plasmids will be in vitro methylated with the pure methyltransferase and/or with crude *M. pneumoniae* cellular extracts, prior to be transformed into *M. pneumoniae* using electroporation and PEG-based protocols.

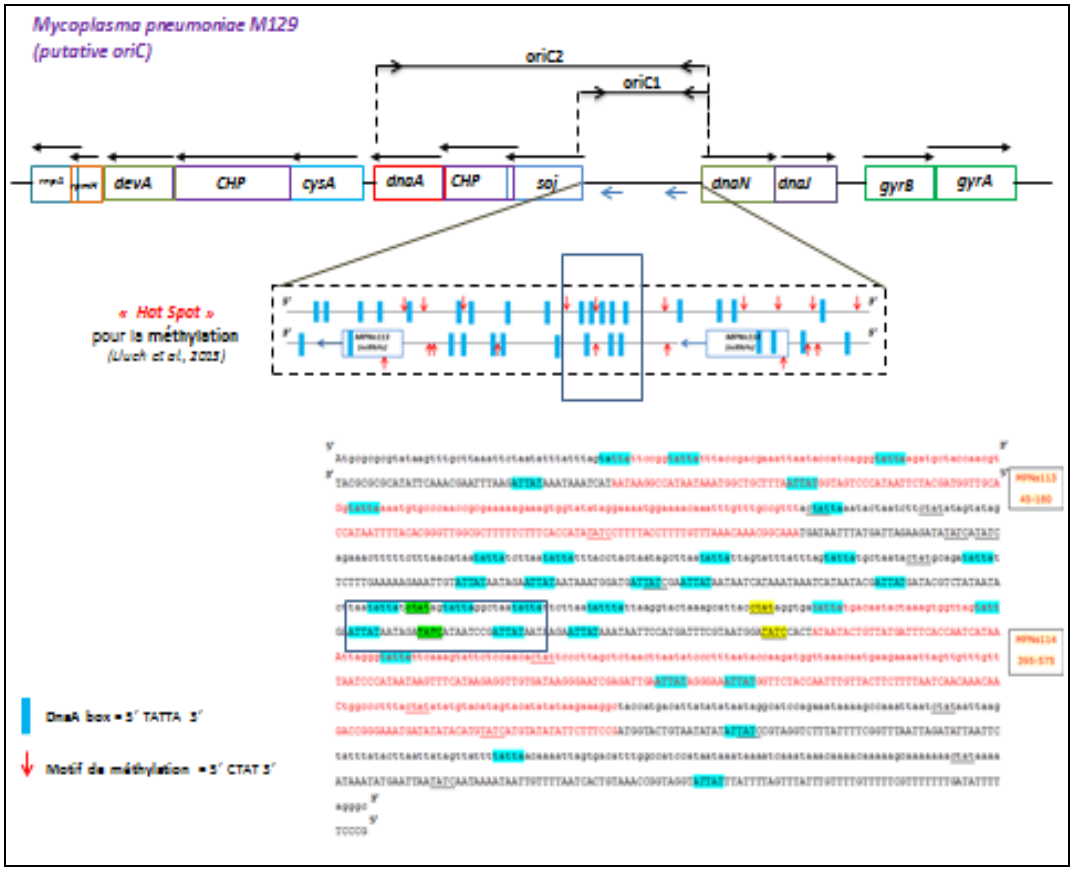

Figure 15. *M. pneumoniae* chromosomal origin of replication region

3.2.2 Genome transplantation potential barriers

While we are constructing and testing molecular tools for *M. pneumoniae*, we also performed genome and literature analyses to identify potential barriers to GT in *M. pneumoniae* and determine means to circumvent those obstacles (Figure 16)

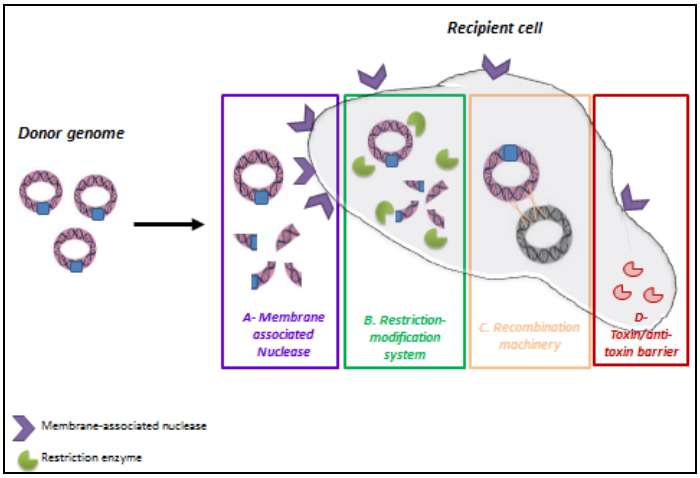

Figure 16. Potential barriers to genome transplantation in *M. pneumoniae* A number of barriers can block the genome transplantation into a recipient cell. Restriction-modification systems are well known and have been demonstrated to play a key role in mycoplasma genome transplantation (Lartigue et al, 2007). In addition, *Mycoplasma* species are known to produce either secreted or membrane-bound unspecific nucleases that are highly active and are thought to provide a source of nucleotides for the mycoplasma cells. Within seconds these enzymes can hydrolyze DNA and need to be neutralized during the genome transplantation. The recombination machinery is a potential difficulty in particular in intra-species transplantation. Indeed, if there is a recombination event between the incoming genome and the genome of the recipient cell it will be difficult to obtain pure “transplants”. Finally, several bacteria produce toxin-antitoxin systems that can be genome-encoded. In that case, the incoming genome needs to encode the antitoxin or it will hydrolyzed.

4. Cited References

1. DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013) Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343.
2. Doudna JA, Charpentier E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. **346**(6213):1258096.
3. Halbedel S, Stülke J. (2007) Tools for the genetic analysis of Mycoplasma. Int J Med Microbiol. **297**(1):37-44.
4. Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA 3rd, Smith HO, Venter JC. (2007). Genome transplantation in bacteria: changing one species to another. Science. **317**(5838):632-8.
5. Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, Noskov VN, Denisova EA, Gibson DG, Assad-Garcia N, Alperovich N, Thomas DW, Merryman C, Hutchison CA 3rd, Smith HO, Venter JC, Glass JI. (2009) Creating bacterial strains from genomes that have been cloned and engineered in yeast. Science. **325**(5948):1693-6.
6. Lee SW, Browning GF, Markham PF. (2008) Development of a replicable oriC plasmid for Mycoplasma gallisepticum and Mycoplasma imitans, and gene disruption through homologous recombination in M. gallisepticum. Microbiology. **154**(Pt 9):2571-80.
7. Lluch-Senar M1, Luong K, Lloréns-Rico V, Delgado J, Fang G, Spittle K, Clark TA, Schadt E, Turner SW, Korlach J, Serrano L (2013) Comprehensive methylome characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. PLoS Genet. **9**(1):e1003191.
8. Sorek R, Lawrence CM, Wiedenheft B. (2013) CRISPR-mediated adaptive immune systems in bacteria and archaea. Annu Rev Biochem. **82**:237-66.
9. Torres-Ruiz, R., Rodriguez-Perales, S. (2015) CRISPR-Cas9: A Revolutionary Tool for Cancer Modelling. *International journal of molecular sciences* ***16***: 22151-22168.