**ERASynBio 2nd Call for Proposals 2014 "- Building Synthetic**

**Biology capacity through innovative transnational projects”**

**Project description**

**MiniCell**

**A model-driven approach to minimal cell engineering for medical therapy**

Synthetic biology promises to enable researchers to design therapeutically and industrially valuable organisms. Achieving this promise requires new techniques for designing, synthesizing, and transplanting entire genomes. Here we propose to develop the first model- driven approach to synthetic biology, and use this approach to construct a bacterial chassis capable of synthesizing and delivering human lung therapies *in situ*. Specifically, we propose to develop a whole-cell model of the human lung pathogen *M. pneumoniae*, and use this model to design and construct a reduced, non-pathogenic chassis capable of delivering human lung therapies and/or vaccinations. This project will involve intimate integration of predictive modeling, genomic engineering, and systems and synthetic biology. Model predictions will provide direct input into genomic engineering, and the newly created strains will be characterized to refine the computational model. The project will produce the most accurate computational model of any organism to date, as well as produce the most reduced cell to date. In the future we anticipate this reduced chassis could be extended to synthesize and deliver small molecule and/or protein therapies to diseased lungs *in situ*.

Thematic area: (*Minimal genomes)*

**1. Project concept**

Synthetic biology promises to enable researchers to engineer biological systems for a myriad of applications including producing biofuels and biosensors, delivering commodity chemical and drugs, and decontaminating soil and water. One particularly exciting application is the possibility of creating bacteria factories that can target diseased tissues and produce and secrete therapeutics *in situ*. Currently, the main bottleneck toward realizing this potential is our inability to predict the behavior of engineered bacteria and their human hosts. Computational models that can accurately predict the behavior of engineered bacteria inside their cellular hosts coupled with reduced bacteria with minimal uncharacterized components promise to enable efficient strain design.

*Mycoplasma pneumoniae* (*M. pneumoniae*) is an ideal starting point for designing a minimal cell for use as bacterial therapy chassis. It has a small (860 kb) and comprehensively annotated genome (Wei et. al, in submitted), and a rich collection of functional genomic data describing its transcriptome (*1, 2*), methylome (*3*), proteome (*4-6*), and metabolome (*7*). A flux balance model describing its metabolism (*8*) is also available. In addition, it is closely related to *M. genitalium* and *M. mycoides* whose genomes have been chemical synthetized, transplanted (*9*), and comprehensively modeled (*10*). Furthermore, *M. pneumoniae* is well- suited for bacterial therapy because it is a weak human pathogen and is easily treated with commercially available antibiotics (*11*).

Recently, we and others developed a whole-cell computational model of the most closely related species, *M. genitalium* (11). The model was developed using the *M. genitalium* genome sequence and functional genomic data from other bacteria species including *M. pneumoniae, B. subtilis,* and *E. coli*. The model accounts for the specific function of every annotated gene product, and describes the dynamics of every molecular species over the entire cell cycle. Karr *et al.* demonstrated that the model can accurately predict some experimental measurements, including the phenotypes of single-gene deletion strains. However, the model is unable to accurately predict the effect of multiple deletions in part because the model was not trained using data observed in *M .genitalium*. Nevertheless, whole-cell modeling coupled with species-specific `omics` data promises to provide more accurate predictions which could guide rational biological design of a minimal cell.

The main goal of this proposal is to design and construct a ‘minimal-cell’ using a large compendium of experimental data and a comprehensive whole-cell model.

1) We will construct the first whole-cell model of *M. pneumoniae* for synthetic biology applications. This will be the first whole-cell model developed using comprehensive species-specific data, and therefore will be significantly more accurate than the previous model.

2) We will use the model to define a list of candidate virulence genes to be deleted in order to obtain a reduced non-pathogenic strain that could be used as chassis for delivering lung therapy.

3) We will corroborate the model by: i) using an unbiased mini-transposon mutant library of *M. pneumoniae* (*12*) and High-throughput Insertion Tracking by deep Sequencing (HITS); ii) clone isolation from the library to confirm predicted essentiality for single genes; iii)

deletion of large non-essential genomic regions comprising several genes and virulence factors by genome transplantation technique.

4) We will transplant the *M. pneumoniae* genome to yeast, remove non-essential pathogenicity factors, as well as other non-essential genes, and then transplant it back to *M. pneumoniae.*

5) We will characterize the infectivity and virulence of the reduced strain (chassis) using an

*in vitro* cell culture infection assay.

6) Ultimately, by the end of the project, we will have constructed and characterized a reduced, non-pathogenic chassis capable of attaching to lung cells that could be further extended to synthetize and deliver therapies to diseased human lungs.

We propose to combine our individual expertise in systems modeling (Mt Sinai and CRG), *M. pneumoniae* molecular biology (CRG), genome editing (INRA), and *M. pneumoniae* physiology and *in vitro* infection assays (UGOE) to develop a computational tool for rationally engineering *M. pneumoniae*. Furthermore, we propose to use this tool to design and construct a bacterial chassis capable of delivering drugs to diseased human lung tissue *in situ*.

**2. Background and present state of the art**

Microorganisms have great potential to treat human disease (*13*). To date, most researchers have focused on the therapeutic potential of viruses which are the easiest microorganisms to engineer (*14*). However, viruses can only contain a limited amount of DNA, and present a significant risk for uncontrolled insertion. We believe that bacteria provide several advantages over viruses as a medical therapy delivery vehicle:

• Bacteria are self-sufficient. They can contain all biological machinery needed to synthesize complex therapeutics.

• Complex detector modules can be integrated into bacteria to program them to respond only to diseased tissue.

• The risk of bacterial DNA integration into the host genome is low.

• In most cases bacteria can be effectively controlled using antibiotics, minimizing the risk of uncontrolled proliferation.

• It is possible to engineer killing circuits or auxotrophic dependence to control cell growth. The perfect chassis for bacterial therapy is a bacterium that can be cultured inexpensively,

divides slowly, is commensal or weakly pathogenic and does not cause a strong immune response, and that can be quickly eradicated using readily available antibiotics. *M. pneumoniae* matches most of these design criteria.

**Bacterial therapy**

Engineering bacteria to deliver therapeutic agents is an exciting area of research with great clinical potential. Bacteria have been proposed to deliver cancer therapy (*15*), as well as to deliver RNA therapeutics (*16*). Engineered bacteria are already being designed to prevent HIV infection in women (*17*), and to prevent dental caries (*18*). ActoBiotics

(<http://www.actogenix.com/)>is the only bacterial therapy under active commercial development. ActoBiotics is based on the gut bacterium *Lactococcus lactis*. Three engineered strains are currently in clinical trials: a strain designed to treat ulcerative oral mucositis, and two strains designed to treat inflammatory bowel disease (*19, 20*). In this project, we aim to engineer a bacterial chassis that could be used to deliver lung therapy. Although the lung is a good vaccination route and there are numerous lung diseases, no bacterial lung therapy is currently under development, in part because the lung was considered a sterile organ until recently. In the past few years, studies have shown that healthy lungs harbor a complex microbiome (*21*). Taken together, there is both great need and promising potential to develop a bacterial chassis for treating human lung diseases.

**Minimum genome**

Several theoretical studies have suggested that the minimum number of genes to sustain life is approximately 240 (*22-24*). The most common approach for studying gene essentiality in bacteria is saturation transposon analysis (*25-35*). Most previous studies have used this approach to assess the essentiality of CDSs and known functional RNAs, neglecting regulatory regions, non-coding RNAs and structural elements although these genomic elements can be essential for bacterial growth and can be involved in pathogenicity (*36, 37*).

**Whole-cell modeling**

Recently we and others developed the first ‘whole-cell’ model (*10*), providing a new tool to study cell physiology and to predict the effect of gene depletions and over-expressions on cell physiology. This model provided an invaluable framework for integrative analysis of individual biological processes. However, in the absence of ‘omics’ information for *M. genitalium*, the model was constructed by estimating the values of most of the model’s parameters. Consequently, many of the model’s predictions are incorrect. Going forward, whole-cell models which combine comprehensive, species-specific ‘omics’ information with refined sub-models have the potential to provide substantially more accurate predictions and enable efficient *in silico* genetic circuit design.

**Genome engineering tools**

Recently, several new approaches have emerged for constructing new biological systems including (1) the cloning of both whole natural and synthetic bacterial genome in yeast, (2) the engineering of bacterial genomes in yeast using well-established methods and genetic tools, and (3) the transplantation of isolated bacterial genomes from yeast to a suitable recipient cell to deliver bacterial mutants whose behavior is dictated by the incoming genome (*38, 39*). Preliminary work from the INRA in collaboration with the CRG has resulted in cloning the whole *M. pneumoniae* genome in yeast.

**Novelty, originality and feasibility and important gaps of the current knowledge**

Our proposal touches three important aspects of synthetic biology: whole-cell modeling, minimal cell engineering, and the development of a novel mode of medical therapy delivery. First, we will develop a whole-cell model of *M. pneumoniae* by expanding a previous model of *M. genitalium*, and by training the model using comprehensive species-specific data. Second, we will identify the genes that are important for lung epithelium attachment infection and virulence. Third, we will use genome modification and chromosome transplantation

techniques to remove the virulence and other non-essential genes to develop a reduced *M. pneumoniae* chassis. If successful, this project will be the first example of whole-cell model- driven biological design, and will pave the way for bacterial lung therapy and vaccination.

**3. Work plan**

The project comprises two main sub-projects: i) developing a comprehensive *M. pneumoniae* model and ii) constructing a minimal non-pathogenic chassis capable of controlled infection (Figure 1). Sub-project 1 includes WP1 and WP2 and sub-project 2 incudes WP1, 3 and 4. Our work plan also includes three work packages dedicated to training the next generation of synthetic biologists (WP5), data and technology dissemination (WP6), and project management including fostering any commercially valuable discoveries (WP7, described in section 6.). The seven individual, but interrelated work packages are:

**WP1:** Develop a whole-cell model of *M. pneumoniae* to guide genome reduction toward a bacterial lung therapy chassis

**WP2:** Experimentally define the *M. pneumoniae* minimal genome

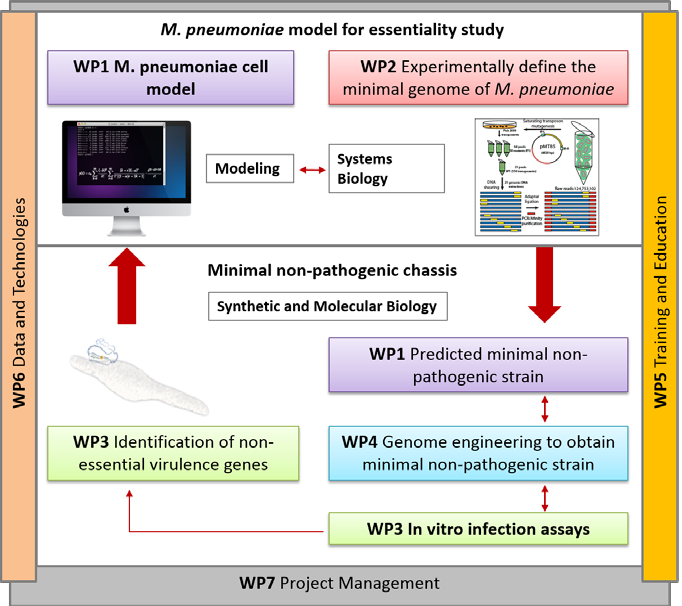
**WP3:** Identify all genes involved in infectivity and pathogenicity

**WP4:** Construct and test reduced non-pathogenic strains predicted by the whole-cell model

**WP5:** Training and education (strategic element)

**WP6:** Data and technology dissemination (strategic element)

**WP7:** Project management and reporting



**Figure 1**. PERT chart illustrating the interactions among the work packages, the main goals of the project, and the interdisciplinary research fields and expertise.

**Work package 1: Develop a whole-cell model of *M. pneumoniae* to guide genome reduction toward a bacterial lung therapy chassis**

**Leader**: Mt Sinai; Partners involved: CRG.

**Objective**: The overall objectives of WP1 are to (1) construct a comprehensive, dynamical, whole-cell model of *M. pneumoniae*, and (2) to use this model to design a chassis optimized for future bacteria lung therapy. The specific objectives of this work package are to:

• Construct a database of *M. pneumoniae* molecular biology which will be used to train a whole-cell model.

• Construct a whole-cell model of *M. pneumoniae* which describes the function of every annotated gene product (ORFs and ncRNAs), accounts for the metabolic cost of every gene including all non-essential genes, and predicts the dynamics of every molecular species over the entire cell cycle.

• Validate the model using publicly available data and data collected by the other work packages.

• Use the model to predict the minimal gene set required for sustained cellular replication in a rich and defined medium.

• Develop reusable software tools for whole-cell modeling and analyzing whole-cell model predictions.

**Task 1.1** Construct a knowledge base of *M. pneumoniae* (months 1-2)

**Task leader:** Mt Sinai; Partners involved: CRG

We will assemble all of the molecular data required to train the whole-cell model into a single database. The database will include the reported genome sequence and functional annotation as well as unpublished data from CRG including the copy number and half-life of each RNA and protein gene product, chemical composition, posttranslational modifications and regulatory networks structures. The database will also contain all of the experimental data collected by the other work packages.

**Task 1.2**. Develop an improved whole-cell modeling platform

**Task leader**: Mt Sinai; Partners involved: CRG

**Task 1.2 a** We will improve our existing whole-cell modeling platform (month 3)

We will simplify our existing platform and make it easier for other researchers to reuse. The platform will be implemented in MATLAB.

**Task 1.2b** Develop a model simulation database (months 3-11)

We will develop a novel hybrid SQL/HDF database which stores model predictions for subsequent analysis. The database will make it easy for other researchers to explore and analyze our simulations via a simple web-based interface without any specific knowledge of computer programming or the particular data structures used to implement the model. The simulation database will be implemented in Python using Django, MySQL, and H5py.

**Task 1.2c** Develop a model user guide (month 11)

We will produce a short, practical guide which describes how other researchers can use the platform to develop their own whole-cell models.

**Task 1.3** Develop a whole cell model of *M. pneumoniae*

**Task leader:** Mt Sinai; Partners involved: CRG.

**Task 1.3a.** Develop sub-models of each cellular process (months 3-6)

We will develop independent sub-models of each of 28 major cellular processes active in *M. pneumoniae*. In particular, we will use new transcriptomic and proteomics data collected by CRG to develop what we expect to be the most accurate transcriptional model of any cell. The sub-models will be adapted from our previous model of *M. genitalium*. The sub-models will be implemented in MATLAB.

**Task 1.3b.** Integrate the sub-models into a single model (months 5-12).

We will integrate the 28 sub-models together using the same 16 state variables we used to integrate our previous *M. genitalium* model*.* The state variables will represent quantities such as the experimentally determined RNA and protein copy numbers, degradation rates, metabolite concentrations, reaction fluxes, and growth rate in defined and rich medium. We will use the reduced surrogate model approach we developed for our previous *M. genitalium* model to identify the model parameters.

**Task 1.3c.** Validate and refine the model (months 8-24).

We will compare the model’s predictions to publicly available experimental data, as well as to unpublished data from the Serrano group. We will iteratively refine the whole-cell model until the model accurately reproduces the experimental data.

**Task 1.3d**. Predict conditionally essential genes and refine model (month 11-36).

We will predict the essentiality of single and multiple knock outs under multiple conditions (antibiotics, temperature, carbon sources, etc.). We will analyze these predictions to identify conditionally essential genes to isolate in WP3 and experimentally characterize in WP2. We will continue to iteratively refine the model using additional data from WP2 and WP3.

**Task 1.4** Predict a minimal, non-pathogenic chassis (months 12-36).

**Task leader:** Mt Sinai; Partners involved: all

We will develop a novel optimization algorithm to efficiently identify the minimal *in silico* gene set required for cellular life. Second, we will develop an approximate algorithm which efficiently identifies the optimal transcriptional and translational regulatory networks which maximize the growth rate of the minimal *in silico* organism. Third, we will simulate the reduced and optimized *in silico* organisms, and compare their behavior to that of the wild- type *in silico* species. Next, we will compare the predicted minimal and optimized genomes to those experimentally defined in WP4. We will continue to refine the model to resolve any discrepancies between the model and experiments.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per participant |
| Mt Sinai | 45 |
| CRG | 5 |

**Deliverables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery  date  (month) | Comments |
|  |  |
| D1.1 | Draft of *M. pneumoniae*  molecular biology database | Mt Sinai | 2 | Database website accessible to consortium members. |
| D1.2 | Model outline | Mt Sinai | 4 | List of cellular processes to model  and major physiological differences compared to *M. genitalium.* |
| D1.3 | Complete process sub-  models | Mt Sinai | 6 | Write and test MATLAB code for  process sub-models. Share code with consortium members via GitHub. |
| D1.4 | Accurate transcriptional  model of *M. pneumoniae* | CRG | 12 | Use Gillespie algorithm and  experimental data to build the model and validate with deep sequencing data. |
| D1.5 | Integrate process sub-  models | Mt Sinai | 12 | Finish model framework. Share  code with consortium members. |
| D1.6 | Preliminary list of predicted  non-essential and conditional essentiality genes | Mt Sinai | 12 | Model will be continuously  optimized and list updated and shared with consortium partner. |
| D1.7 | Tuned model parameters | Mt Sinai | 15 | Share analysis of parameter  values with consortium members. |
| D1.8 | Simulated and validated  model | Mt Sinai | 18 | Share raw data via simulation  database as well as analysis with consortium members. |
| D1.9 | Predicted non-essential  individual genes | Mt Sinai | 18 | List of predicted non-essential  genes to delete. |
| D1.10 | List of conditionally  essentiality genes | Mt Sinai | 24 | List of conditional essentiality  genes. |
| D1.11 | Preliminary *in silico*  designed, optimal strain | Mt Sinai | 24 | Computer simulation |
| D1.12 | Final *in silico* designed, optimal strain | Mt Sinai | 36 | Model will be continuously optimized and list updated and  shared with consortium partner. |

**Significant risks and contingency plans**

Unable to uniquely identify the model parameters. Probability: 5/5; Severity: 1/5 *–* We will identify several model solutions which match the experimental data, simulate each of them, identify differences in their predicted behavior, and design new experiments to discriminate among them.

Model simulations are prohibitively computationally expensive. Probability: 4/5; Severity: 2/5

*–* We will profile the model software to identify places for possible improvement. We will also

use a large cluster to simulate multiple simulations in parallel. If necessary, we will also parallelize the execution of the individual sub-models.

Model does not accurately predict experimental data. Probability: 3/5; Severity: 1/5 *–* We will use the discrepancies between the model predictions and experimental observations to refine the model. This will likely result in discovering previously unknown interactions and quantitative parameters and suggesting new experiments.

Unable to significantly experimentally reduce the *M. pneumoniae* genome. Probability: 3/5; Severity: 1/5 *–* We will begin growth rate optimization from the partially reduced genome, and compare this optimization to directed evolution of the experimentally reduced genome.

**Relationship to other work packages**

The gene essentially experiments in WP2 will be used to define the *in silico* essentiality of the functionally uncharacterized genes included in the whole-cell model. These gene essentiality experiments will also be used to validate the model.

The predicted minimal gene set will be used to guide the genome reduction described in WP4. The minimal genome obtained and optimized in WP4 will also be used to validate the reduced and optimized gene sets in Task 1.4. The information about virulence genes from WP3 will be integrated in the model to define the minimal non-pathogenic strain.

**Work package 2: Experimentally define the *M. pneumoniae* minimal genome**

**Leader:** CRG. Partners involved: all

**Objective:** The overall objective of WP2 is to determine the essential genes and regions of the *M. pneumoniae* genome. The specific objectives of this work package are to:

• Determine the essential regions using an unbiased mini-transposon mutant library of *M. pneumoniae* (*12*) and High-throughput Insertion Tracking by Deep Sequencing (HITS) (*40*),

• Develop a metric for predicting the essentiality of a genomic region based on transcriptomic and proteomic data

• Experimentally validate the essentiality metric by testing the viability of individual transposon library clones.

• Validate the whole cell model by characterizing conditionally essentiality genes

**Task 2.1** Isolate genomic DNA from pools of the transposon mutants library (months 1)

**Task leader:** CRG; Partners involved: UGOE, Mt Sinai

The mini-transposon mutants library of *M. pneumoniae* was previously obtained by “haystack mutagenesis” (*41*). The library is based on an ordered collection of 64 pools that have 2976 pMT85 transposon mutants capable of colony formation. With this number of mutants, the probability of finding a desired mutant in the library is 99.999%. The 64 pools will be ordered into 11 groups and genomic DNA from these 11 samples will be extracted with the Illustrabacteria genomic Kit (GE). The genomic DNA from these 11 samples will then be sequenced by the HITS approach.

**Task 2.2** High-throughput Insertion Tracking by Deep Sequencing (HITS) (months 2-5)

**Task leader:** CRG

Genomic DNAs will be sheared to 100 bp DNA using a Covaris S2 device fragmenter. Paired-end Illumina libraries will be created as described by Bentley et al. (*42*) and enrichment of transposon/chromosomal junction regions will be performed by PCR amplification with a 5´-biotinylated transposon enrichment primer and adapter specific PCR PE 2.0 enrichment primer. The resulting transposon libraries will be PCR amplified and quantified on an Agilent Bioanalyzer chip (Agilent Technologies). Double-stranded templates will be cluster amplified and sequenced on an Illumina GAII.

**Task 2.3** Analyze and sequence data (months 6-7)

**Task leader:** CRG

Standard Illumina paired-end sequencing will be used and clusters will be generated using the Illumina cluster generation protocol. Raw reads will be filtered according to IR sequences. Filtered reads will be mapped to the *M. pneumoniae* reference genome (NC\_000912, NCBI) using the Maq mapping software.

**Task 2.4** Develop and validate an essentiality metric using gold standards for essential and non-essential genes (months 8-17)

**Task leader:** CRG; Partners involved: UGOE

Our essential gold standard set will be composed of the RNA polymerase subunits, the single sigma 70 factor, the tRNA synthases, the DNA polymerase complex, the ribosomal RNA and the central glycolytic enzymes needed for ATP production (*7*). The non-essential gold standard set will be composed of a group of *M. pneumoniae* genes not found in the closely related species *M. genitalium*. For each of the two gold standard sets, we will calculate the average density of insertions and the calculated values will be considered as previously described by Christen et al. (*43*). Then for each of the 693 validated and re- annotated ORFs, 355 ncMPNs, non-transcribed regions, as well as 5´-UTRs longer than 70 bp, we will determine the probability values to be essential, PE and non-essential, PNE. This analysis will allow us to classify different genomic regions as essential or non-essential.

The scoring function will be validated using two approaches: i) confirming insertions by PCR amplification from genomic DNA of individual pools and posterior sequencing. This approach will not only confirm insertions, but also discard putative non-specific amplifications by PCRs in library preparation; and ii) isolating transposon mutants from the pools to assess the viability of the different clones and to determine if the criteria are robust.

**Task 2.5** Validate model essentiality predictions under multiple conditions (months 12-26)

**Task leader:** CRG; Partners involved: UGOE

The growth rate of selected mutants will be analyzed under different conditions to validate model predictions in WP1. The different clones will be isolated by UGOE.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per participant |
| CRG | 20 |
| UGOE | 8 |

**Deliverables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery  date  (month) | Comments |
|  |  |
| D2.1 | Mini-transposon mutants  library sequenced by HITS | CRG | 5 | Sequencing will be done at the CRG  Ultrasequencing facility. Raw data will be saved on a server. |
| D2.2 | Pipeline for sequencing  data analysis | CRG | 7 | This pipeline will be published and  publicly accessible. |
| D2.3 | Identification of  transposon insertion sites | CRG | 9 | Data will be summarized in a table  and reported |
| D2.4 | Confirmation of  transposon insertion sites | UGOE | 10 | Data will be summarized in a table  and reported. PCRs gels confirming the gene disruptions will be stored. |
| D2.5 | Table of predicted  essential genomic elements | CRG | 17 | The table of identified genomic  elements will be shared with the consortium and reported |
| D2.6 | List of selected clones,  after testing mutants growth | UGOE | 26 | The list of confirmed clones will be  shared with the consortium and reported |

**Significant risks and contingency plans**

The specific primer used for HIST hybridizes to off-target sites. Probability: 2/5; Severity: 1/5

*– We will perform controls with the wild-type strain to characterize any non-specific hybridization. We will also design primers additional primers to discern among specific and non-specific hybridization.*

Inserted DNA pool will be contaminated with insertions into DNA from dead cells. Probability:

5/5; Severity: 1/5 *– We will perform statistical analyses to discriminate true insertions from background insertions in DNA from dead cells. We will also define a metric to highlight the mostly likely essentiality regions. Additionally, we will independently validate the non- essential genes by clone isolation.*

**Relationship to other work packages**

As described in WP1, the gene essentially experiments will be used in WP1 to define the *in silico* essentiality of the functionally uncharacterized genes included in the whole-cell model. The gene essentiality experiments will also be used to validate the model developed in WP1.

**Work package 3: Identify all genes involved in infectivity and pathogenicity**

**Leader:** UGOE; Partners involved: all

**Background:**

**Main virulence and pathogenicity factors identified in *M. pneumoniae*:**

**Hydrogen peroxide production:** Similar to many *Mycoplasma* species, *M. pneumoniae* has been shown to produce hydrogen peroxide during glycerol metabolism. Furthermore, hydrogen peroxide has been shown to be toxic to cultured eukaryotic cells (*44, 45*).

**CARDS toxin:** *M. pneumoniae* produces the community-acquired respiratory distress syndrome (CARDS) toxin. This toxin is unique to *M. pneumonia*, but belongs to a superfamily of toxins (ADP-ribosylating enzymes) that includes pertussis, diphtheria and cholera toxins. Mice infected only with the CARDS toxin develop severe lesions. Furthermore, during infection the amount of CARDS toxin has been shown to correlate with the degree of lung inflammation (*46*).

**Adherence protein P1:** The *M. pneumoniae* adhesin P1 is associated with the terminal organelle (*47*). The P1 adhesin as well as the cytadherence accessory proteins frequently mutate through recombination with repeated DNA segments (*48*).

**Ig Binding protein:** An Ig-binding protein M has been recently identified in *M. genitalium* (MG281) (*49*). This transmembrane protein has been shown to bind with high affinity human immunoglobulin G, predominantly through attachment to the conserved portions of the variable region of the kappa and lambda light chains. This binding strongly inhibits antibody binding to its cognate antigen. A homolog has been found in *M. pneumoniae* (MPN400) that has likely the same properties. Therefore, MPN400 is a putative virulence factor, potentially playing a role in host immune system evasion.

**Immune evasion.** *M. pneumoniae* has several duplicated regions that in principle could recombine to generate antigenic variation.

**Objective:** The overall objective of WP3 is to identify the *M. pneumoniae* genes involved in pathogenicity and lung adhesion. The specific objectives of this work package are:

• Enumerate all previously identified *M. pneumoniae* pathogenicity factors by mining the literature (see examples above).

• Identify additional candidate pathogenicity factors by comparison to other *Mycoplasmas*, including *M. genitalium,* and other pathogenic bacteria.

• Identify all non-essential pathogenicity factors using the essentiality data from WP2, the observed operon organization, and the identified pathogenicity factors. Confirm non- essential pathogenicity factors by genome engineering.

• Experimentally confirm predicted non-essential pathogenicity factors by transposon insertion and/or genome engineering using cell-culture assays and/or haemadsorption assays.

**Task 3.1** Enumerate the *M. pneumoniae* genes that encode putative virulence factors (month

1)

**Task leader:** UGOE; Partners involved: INRA

Mine the literature to identify *M. pneumoniae* genes that have been experimentally implicated in virulence (see examples above). Where possible, collect and annotate the primary data.

**Task 3.2** Perform comparative genomic analysis of multiple *Mycoplasma* species to identify conserved essential and non-essential genes

**Task leader**: UGOE; Partners involved: INRA and CRG

**Task 3.2a.** Identify the set of genes shared between *M. pneumoniae* and *M. genitalium* as well as the core genome of the genus *Mycoplasma* (months 1-3)

Because the *M. genitalium* genome is almost entirely present in *M. pneumoniae,* the additional genes present in *M. pneumoniae* are likely non-essential. We will identify these likely non-essential genes comparing the genomes of both species. In addition, we will identify the conserved core genome of the *Mycoplasma* genus. Genes that are part of the core genome are likely important for the biology of *Mycoplasmas*, whereas those genes that are conserved in *M. genitalium* and *M. pneumoniae* but not part of the core genome may reflect the specific evolution of this particular lineage and the adaptation to a common host.

**Task 3.2b.** Identify genes that are not essential *in vitro* conditions, but are essential for infection (months 1-4)

We will combine the essentiality information collected by WP2 with information from 22 sequenced and characterized clinical isolates.

**Task 3.3.** Isolate selected mutants affected in *in vitro* non-essential genes as well as infection conditionally essential genes (month 12-36)

**Task leader:** UGOE

Mutants defective in *in vitro* non-essential and conditionally essential genes related with pathogenicity will be isolated from an ordered transposon insertion library. For this purpose, PCR screens with primers for the gene of interest and the transposon will first be performed with pools of mutants. In the second step, the individual clones of the pools will be screened and the mutants will finally be isolated and verified by Southern blot analysis.

**Task 3.4** *In vitro* testing of mutation impact (Tn insertion and/or genome engineering) on *M. pneumoniae* cell pathogenicity and cell adhesion using haemadsorption and cell-culture assays (month 13-36)

**Task leader:** UGOE; Partners involved: CRG, INRA

The selected strains obtained from the transposon library and by genome engineering (mutants that are affected in known or predicted virulence factors) will be used to assay adhesive growth and pathogenicity. To test adhesion, we will test whether the mutants adhere to the plastic surface in cultivation vials (which is typically observed with *M. pneumoniae*). Moreover, haemadsorption will be tested. Pathogenicity will be tested by infecting cultures of HeLa cells with wild type and mutant cells. Growth of the cell culture in the presence of bacteria is the established indicator for the loss of pathogenicity (*44, 45, 50,*

*51*).

**Task 3.5** Identify chromosome regions for deletion by genome engineering taking into account both essentiality data and virulence predictions (months 21-36)

**Task leader**: INRA; Partners involved: CRG

The chromosome regions that carry the non-essential virulence genes to be targeted will be analyzed carefully to find the optimal strategy for gene inactivation requiring the lowest number of steps. A map carrying the targeted virulence genes, the operon organization inferred from transcriptomics and the essentiality data will be produced. From this map a list of virulence genes and other non-essential genes that will be removed from the initial genome will be produced.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per participant |
| UGOE | 46 |
| INRA | 6 |
| CRG | 24 |

**Deliverables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery  date  (month) | Comment |
| D3.1 | List of candidate genes to  be deleted | UGOE | 1 | List of genes not shared between  *M. pneumoniae* and *M. genitalium* as well as of a list of *M. pneumoniae* genes not present in the *Mycoplasma* genus genome |
| D3.2 | List of putative *M.*  *pneumoniae* virulence factors | UGOE | 2 | Based on literature research |
| D3.3 | List of *M. pneumoniae*  mutants to be isolated for virulence tests | UGOE | 25 | List to order specified transposon  insertion library |
| D3.4 | List of experimentally  confirmed virulence factors | UGOE | 36 | In vitro assay data on the impact of  mutated genes on adhesion and pathogenicity |
| D3.5 | Concept and design for  genome deletions | INRA | 36 | A map carrying the targeted  virulence genes, the operon organization inferred from transcriptomics and the essentiality data |

**Significant risks and contingency plans**

Inability to isolate all of the anticipated mutants. Probability: 4/5; Severity: 1/5 – If we are unable to isolate all of the planned mutants, this may indicate that the corresponding genes are essential. This information will be spread in the consortium and implemented in the model.

Lack of feasibility of the genome engineering plan. Probability: 3/5; severity: 1/5 – If the genes to be deleted are too much scattered over the genome, this may results in a genome engineering strategy with too many steps. To overcome this difficulty, it is possible (i) after deleting large regions of the chromosome to add essential gene that have been removed or (ii) to develop (WP4) new methods for the simultaneous deletion of several genes.

**Relationship to other work packages**

This WP involves close collaboration with all other WPs. The whole-cell model (WP1) and the experimental essentiality data will provide an important foundation to decide which mutants to isolate. In this WP, we will also test the genome-engineered strains produced in

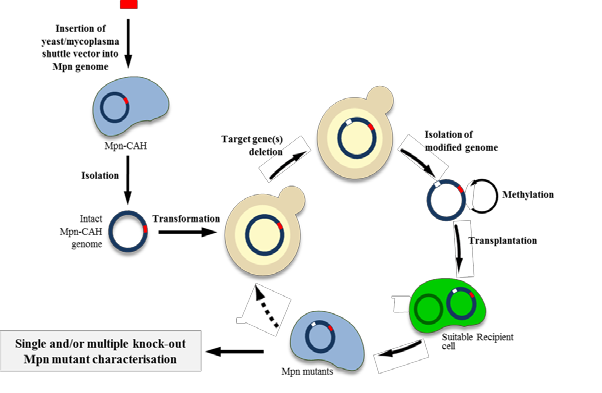
WP4 for adhesion and virulence. The genome engineering strategy used in this WP will be determined by the methods developed in WP4.

**Work package 4: Construct and test reduced non-pathogenic strains predicted by the whole-cell model**

**Leader:** INRA

**Background**

Efforts to identify *M. pneumoniae* virulence factors and to produce minimal cells that can be used as chassis are currently hindered by the lack of efficient molecular genetics methods. Most of the genetics tools for model organisms such as *E. coli* or *B. subtilis* cannot be used with *Mycoplasmas.* Only one example of gene disruption via homologous recombination has been reported in *M. pneumoniae* (*52*), and furthermore the method is not efficient enough to achieve multiple/wide gene knock-outs and thus is unable to support genome-scale functional characterization. It has been suggested that the weak frequency of homologous recombination in *M. pneumoniae* is the result of insufficient expression of genes involved in initial recombination and resolution of Holliday junctions. Indeed, proteins such as the RuvAB Holliday junction DNA helicase have not detected by proteomics (*53*); Lluch et al in preparation). Thus, to date, all *Mycoplasma* genetic modification has been performed using labor-intensive random transposon mutagenesis (*54-56*).



**Figure 2: “Genomic platform” for producing *M. pneumoniae* mutants using synthetic biology tools.** First, wt *M. pneumoniae* (Mpn) cells are transformed with a yeast/*Mycoplasma* integrative vector (red rectangle). This vector generally contains (i) a **C**entromere, an **A**utonomously replicating sequence and an auxotrophic marker for **H**istidine for selection and propagation of centromeric plasmids in yeast and (ii) a selectable-resistant marker for screening in *Mycoplasma* cells. Second, newly marked genomes are gently isolated from Mpn-CAH cells and transferred into yeast spheroplasts using conventional yeast transformation procedure. After cloning, the large repertoire of yeast genetic tools is used to modify the incoming genome. Third, the engineered genome is then isolated and transplanted back into suitable recipient cell to generate Mpn mutant strains that will be further genotypically and phenotypically characterized. In some cases, methylation of the donor DNA before the transplantation step might be necessary to protect it from the recipient cell’s restriction system(s). This cycle can be repeated starting from the newly engineered genome (dashed arrow) to create a multiple knock-out strains. Figure adapted from Lartigue *et al*, 2009.

**Objective:** Recent achievements in *Mycoplasma* genome transplantations, whole genome synthesis and in-yeast genome engineering have opened new possibilities for modifying bacterial genetic material at genomic scale. In this work package, we propose to extend synthetic biology new technologies to *M. pneumoniae* to establish a powerful platform for functional genomics research and biotechnological applications. The new genomic engineering platform will then be used to reduce the *M. pneumoniae* genome. Guided by the whole cell model (WP1), virulence (WP3) and non-essential (WP2) will be to construct a viable minimal non-pathogenic *M. pneumoniae* chassis (Figure 2).

**Task 4.1** Deletion of gene(s) in *M. pneumoniae* genome cloned in yeast (months 2-26)

**Task leader:** INRA; Partners involved: CRG

**Task 4.1a** Targeted deletion of genes selected in WP3.

**Task leader:** INRA; Partner: CRG

We will take advantage of yeast’s capacity to take-up and recombine specific DNA fragments to a target site to either replace, exchange or insert genes into the *M. pneumoniae* genome. First, we will produce a linear DNA fragment which contains the inter alia yeast marker and has identical ends (~60 bp) to its target site. Depending on its size, this fragment will be produced either by polymerase chain reaction, chew-back- anneal or gene synthesis. We will then introduce the fragment into yeast spheroplasts by standard methods. Yeast transformants will then be assessed by PCR, PFGE and DNA sequencing (NGS). To accomplish this we will optimize the methods to directly obtain bacterial DNA without interference from yeast DNA.

This genetic manipulation method frequently introduces unwanted sequences/markers into the genome. These unwanted sequences can be removed using the *cre-loxP* system (*57*) or Tandem Repeat Endonuclease Cleavage (TREC). TREC produces seamless modifications in episomal DNA molecules maintained in yeast (*58*). The performance of TREC is comparable with *cre/lox* system, but offers the significant advantage of not leaving scars.

**Task 4.2** Identification of a suitable recipient cell for *M. pneumoniae* genome (months 1-12)

**Task leader**: INRA; Partner: CRG

To date, genome transplantation has been achieved for a growing number of *Mycoplasma* genomes, but not for *M. pneumoniae*. One reason could be the specific restriction- modification systems present in this bacterium (*3*). In fact sequence comparison reveals that *M genitalium* has only one DNA methylase orthologous to that of *M. pneumoniae*. An obvious choice is another *M. pneumoniae* strain (i.e. FH that differs through missense mutations throughout the genome) or *M. genitalium.* As a backup we will also try *M. gallisepticum,* a species that combines several advantages including efficient methods of transformation with PEG, availability of *oriC* plasmid for control experiments (*59*) and ease of growth.

Building upon our existing experience in *Mycoplasma* genome transplantation (*38, 39*) we will develop methods of genome transplantation for *M. pneumoniae* genomes. To establish the protocol, we will use genomes (carrying an antibiotic resistance marker) isolated from *M. pneumoniae* cells rather than genomes isolated from yeast. This will ensure that we obtain highly concentrated gDNA preparation and avoids potential restriction barrier issues.

*Mycoplasma* cells will be embedded in agarose plugs, treated with proteinase K and detergent, and the gDNA will be released from agarose by melting and agarase treatment. To allow for the entry of the *M. pneumoniae* donor genome into the suitable recipient cells, we will develop a cell fusion protocol. To date, only PEG-mediated protocols has proven to be effective for whole bacterial genome transplantation. A recent study showed that a treatment of *M. pneumoniae* cells with 10% PEG200 increases cells attachment and probably cell fusion (60). After transplantation we will use tetracycline to select putative transplants carrying the tetracycline resistance marker. Finally, we will perform a series of genotypic and phenotypic analyses to confirm that no recombination between the donor and endogenous chromosomes occurred, and that the transplantation reaction was successful.

**Task 4.3** Transplantation of modified *M. pneumoniae* genome from yeast into suitable recipient cells (months 12-26)

**Task leader:** INRA; Partners involved: CRG

We will use the protocol developed in task 4.2 to transplant modified *M. pneumoniae* genomes, except that gDNA will be released from yeast agarose plugs digested with zymolyase and treated with detergent and proteinase K, prior transplantation into suitable recipient cells.

The major difficulty in the yeast to bacteria transplantation procedure is to avoid the destruction of the incoming genome by the specific endonucleases expressed by the recipient cell. The identification of all methyltransferases that modify the *M. pneumoniae* genome (*3*) is a major step for circumventing this nuclease problem. Strategies that have been proposed to overcome the restriction barrier and that should applicable are: (1) *in vitro* methylation of donor genomic DNA with cellular extracts prepared from the recipient cells, (2) *in vitro* methylation of donor genomic DNA with recipient cell recombinant methyltransferases produced in *E. coli* and (3) *in vivo* methylation of donor genomic DNA in engineered yeast that express *M. pneumoniae* identified methyltransferase gene(s) (*3*). *In vitro* methylation has been tested, and appears to effectively protect the donor genome from recipient cell restriction enzymes (*38*). We will follow the second and third strategies.

**Task 4.4** Construction of a minimal non-pathogenic *M. pneumoniae* strain (months 24-36)

**Task leader:** INRA; Partners involved: CRG

Data obtained in WP 1, 2 and 3 and methods developed in WP4 will be combined to construct a minimal non-pathogenic *M. pneumoniae* strain

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per  participant |
| CRG | 18 |
| INRA | 18 |

**Deliverables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery  Date (month) | Comments |
| D4.1 | Bacteria to bacteria transplantation | INRA | 6 | Protocol will be set- |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | protocol |  |  | up and reported |
| D4.2 | Identification of best protocol to  delete selected genes in yeast | INRA | 7 | Protocol will be set-  up and reported |
| D4.3 | Efficient method for sequencing  whole bacterial genome from yeast | CRG | 12 | Protocol will be set-  up and reported |
| D4.4 | *M. pneumoniae* methylation protocol  assessed | CRG | 13 | Protocol will be set-  up and reported |
| D4.5 | Yeast to bacteria transplantation  protocol set-up | INRA | 18 | Protocol will be set-  up and reported |
| D4.6 | First engineered yeast strains  containing the *M. pneumoniae* genome with deleted virulence genes | INRA | 24 | Panel of yeast strains |
| D4.7 | *M. pneumoniae* strains containing in  yeast engineered genomes | INRA | 30 | Panel of *M.*  *pneumoniae* strains |
| D4.8 | First engineered *M. pneumoniae*  chassis | INRA | 30 | Minimal non-  pathogenic strain |

**Significant risks and contingency plans**

Large genomic regions cannot be deleted in yeast by TREC. Probability: 3/5; Severity: 1/5 – New genome editing methods (CRISPR/TALEN) will be assayed in yeast cells carrying *M. pneumoniae* genomes as well as in *M. pneumoniae* cells. Older methods including CRE/LOX, or lambda red will also be tested in *M. pneumoniae* cells. If *M. pneumoniae* genome cannot be transplanted from yeast to a suitable recipient cell, we will modify the *M. pneumoniae* genome directly.

Clustered, regularly interspaced, short palindromic repeats (CRISPR)-associated Cas9 endonuclease complexed with dual-RNAs has been recently used to introduce precise mutations into Streptococcus pneumoniae and *E. coli* (*61*). We will test this promising method in *M. pneumoniae* cells after verifying that Cas9 nuclease can be expressed in *M. pneumoniae* cells. In addition to the CRISPR/CAS system, we will evaluate the Cre-Lox and lambda phage Red recombination systems in *M. pneumoniae* (*62*). In the last few years, the CRG has accumulated immense knowledge on *M. pneumoniae* promoters and transcriptional regulation, and has already successfully expressed heterologous proteins in *M. pneumoniae*.

Expression of the methyltransferase of *M. pneumonie* in yeast is toxic or it shows low activity. Probability: 3/5; Severity: 1/5 – We will recombinantly express MPN198 (the main *M. pneumoniae* methyltransferase) in *E. coli*, and then use this protein to methylate *M. pneumonie* DNA *in vitro*.

Recipient strain rejects transplanted genome. Probability: 3/5; Severity: 1/5 *–* The *M. pneumoniae* FH strain is an ideal recipient because it has a very similar proteome. However, *M. pneumoniae* contains an uncharacterized type II restriction-modification system that could interfere with transplantation. Alternatively, we will try to transplant the *M. pneumoniae* genome into *M. genitalium* which only has the same type I restriction-modification system present in *M. pneumoniae*.

**Relation to other work packages**

Genes predicted to be involved in virulence (WP3) and some genes identified as non- essential (WP2) will be deleted from *M. pneumoniae* genome guided by the whole-cell model (WP1) in order to produce viable reduced non-pathogenic *M. pneumoniae* strain(s). Molecular and biochemical analyses will be accomplished in WP3 to check that the strain(s)

behavior matches the model’s predictions (WP1).

**Work package 5: Training and education (strategic element) Leader:** INRA; Partners involved: all

**Objective:** MiniCell consortium is dedicated to fulfilling ERASynBio’s mission to train the next generation of synthetic biologists. Combining the project’ expertise in molecular biology and computational approaches, all partners will contribute to training on two levels. First, the partners will organize a workshop on synthetic biology open to the scientific community, and second, they will create opportunities for trainees to visit and study with other consortium members.

**Task 5.1** Organize a scientific workshop titled "From modelling to building a minimal cell using synthetic biology tools"

**Task leader:** INRA; Partners involved: all

During the second year of the project the INRA group, supported by the CRG, will organize a two day workshop in Bordeaux. The workshop will focus on comparing and evaluating the two main approaches that scientists have pursued to build minimal cellular chasses: (1) approaches based on common model organisms with large genomes (e.g. *E. coli, B. subtilis) and* (2) approaches based on organisms with naturally small genomes (e.g. approach followed in this proposal). We will gather specialists from both approaches. The program will cover lectures and discussions with renowned scientists as well as methodological training, organized as short practical projects. There will also be demonstrations of key technologies. For each of these mini-projects, a team of 2 participants will work under the direction of the local faculty. The course will cover the following aspects: metabolic modeling, comparative genomics of minimal cells, identification of essential genes, bacterial genome engineering, bacterial genome cloning in yeast, genome transplantation, and the ethics and regulatory issues concerning the building of minimal cells.

Furthermore, attendees will discuss the ethical implications of synthetic biology on society and research. Alain Blanchard and co-workers are participating in public debates with authorities and will chair discussions among workshop participants.

The call for workshop applications will be advertised on the MiniCell website, as well as on other public workshop/seminar portals (e.g. iAnn, <http://iann.pro/iannviewer>). PhD students and postdoctoral fellows in bioinformatics, biology, biochemistry, and other related fields will be asked to submit brief applications. The consortium will invite 10 trainees to attend based on the anticipated impact of the workshop on their careers.

**Task 5.2** Trainee exchanges among partner laboratories

**Task leader:** CRG; Partner: all

Trainees involved in the MiniCell project will be encouraged to do at least one two week secondment exchange with one of the partner laboratories to enable trainees to learn new techniques and/or knowledge from their host lab, and vice-versa. The CRG will coordinate the secondment planning during the first six months of the project, and will oversee the visits including collecting reports of each visit. The secondment exchanges will also promote communication among the partner laboratories.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per  participant |
| INRA | 1 |
| CRG | 0.5 |
| UGOE | 0.5 |
| Mt Sinai | 0.5 |

**Deliverables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery date  (month) | Comments |
| D5.1 | Secondment plan | CRG | 6 | Coordinate exchanges among all partners. Report in midterm review. |
| D5.2 | Scientific  Workshop | INRA | 24 | "From modelling to building a minimal  cell using synthetic biology tools" |
| D5.3 | Short report about  each secondment | CRG | 36 | Secondment reports will be included in  the final report. |

**Work package 6: Data and technology dissemination (strategic element) Leader:** Mt Sinai; Partners involved: all

**Objective:** By choosing the strategic element Data and Technologies MiniCells follows the vision of ERASynBIO to make existing system biology data usable. MiniCells adds value on the data by collecting them and creating a knowledge base; integrating them and developing a model for predicting phenotypes; and creating a simulation database saving the output of our model. The modeling platform as well as the databases will be open-source and shared with the academic community by the end of the project. In the case of industrial partners a commercial version (license based) is considered to be developed. Having such a ready to use and modifiable software tool available allows other researchers to build and optimize their models for their own research and to advance systems biology research.

**Task 6.1** Providing an open software tool for whole-cell modelling to the scientific community

(task 1.2)

**Task leader:** Mt Sinai

A reusable whole-cell modeling platform: The platform will provide a common set of tools and standards for researchers developing whole-cell models, including sub-model templates, data structures for integrating the sub-models, an algorithm for simulating whole-cell models, and a system for logging model predictions. The main objective of the platform will be to make it easier for researchers to build and execute whole-cell models. The platform will distributed open-source via our project website, GitHub, and SimTK. The modeling platform

will be developed at the beginning of the project, and continually refined throughout the project. A whole-cell model user guide will be developed and will help other scientists to execute, modify, and extend the whole-cell model for their own research. The user guide will include detailed examples and sample source code. The user guide will be publicly accessible via our project website.

**Task 6.2** Sharing a knowledge base of *M. pneumoniae* with the academic community (task

1.1)

**Task leader:** Mt Sinai; Partners involved: CRG

A knowledge base of *M. pneumoniae* molecular biology: The knowledge base will organize all of the information used to train the *M. pneumoniae* whole-cell model including its genomic sequence, the functional annotation of each gene product, the copy number and half-lives of each RNA and protein species, the chemical composition of *M. pneumoniae*. The knowledge base will contain all experimental data collected as part of the projects as well as other publicly available *M. pneumoniae* molecular data. In order to make the knowledge base accessible to the consortium (during the project) and the academic community (at the end of the project) it will be available on the MiniCell website and accessible through a web interface.

**Task 6.3** Sharing a simulation database of *M. pneumoniae* with the academic community

(task 1.2b)

**Task leader:** Mt Sinai; Partners involved: CRG

We will develop a novel hybrid SQL/HDF database which stores the output of our model predictions (e.g. predicted growth rate, reaction fluxes, etc.). The database will make it easy for other researchers to explore and analyze our simulations via a simple web-based interface without any specific knowledge of computer programming or the particular data structures used to implement the model. The simulation database will be implemented in Python using Django, MySQL, and H5py.

In order to make the knowledge base accessible to the consortium (during the project) and the academic community (at the end of the project) it will be available on the MiniCell website and accessible through a web interface.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per  participant |
| Mt Sinai | 0.5 |

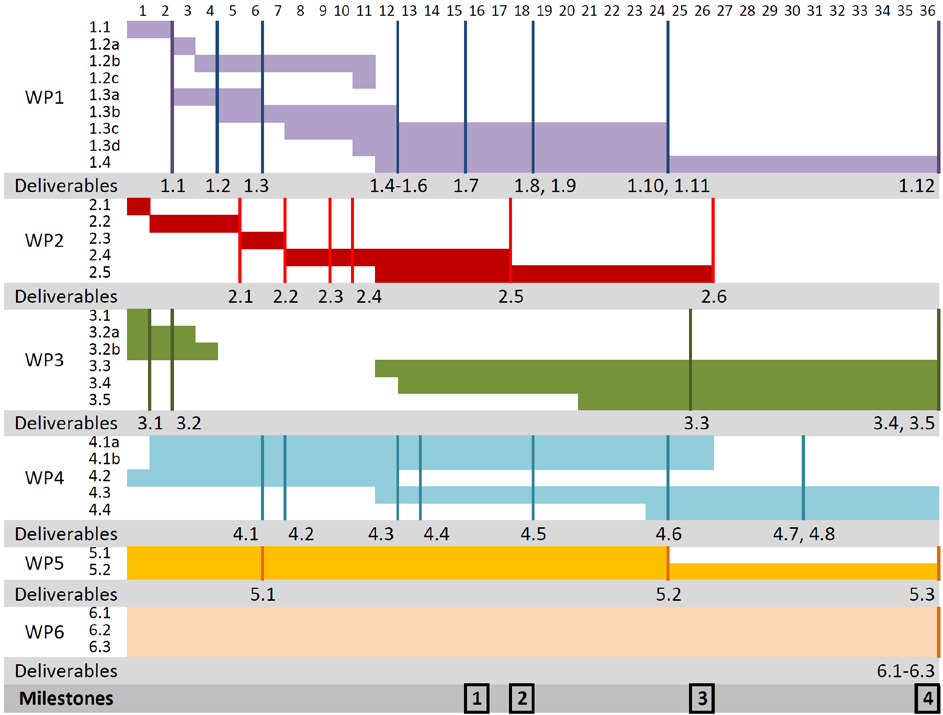
**Deliverables**

|  |  |  |  |
| --- | --- | --- | --- |
| Number | Deliverable Title | Lead  Partner | Delivery date  (month) |
|  |
| D6.1 | Whole-cell modelling platform plus user guide available for  the scientific community at MiniCell website, GitHub, and  SimTK | Mt  Sinai | 36 |
| D6.2 | Final version of knowledge base publically available on  MiniCell website | Mt  Sinai | 36 |
| D6.3 | Final version of simulation database publically available on | Mt | 36 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | MiniCell website | Sinai |  |

**Overall work plan**

The Gantt chart below illustrates the project timeline, including all deliverables and milestones.



|  |  |  |  |
| --- | --- | --- | --- |
| Number | Milestone Title | Lead  Partner | Delivery  date (month) |
|  |
| MS1 | New synthetic biology tools to engineer the *M.*  *pneumoniae* genome at large scale using yeast | INRA | 16 |
| MS2 | Predicted non-essential genomic regions | Mt Sinai | 18 |
| MS3 | Experimental confirmation of predicted non-essential and  conditionally essentiality genes | CRG | 26 |
| MS4 | Refined model based on experimental data | Mt Sinai | 36 |

**4. Impact**

**Dissemination, publication, and intellectual property rights management**

The CRG communication office will supervise the dissemination of all results to the public, including coordinating with the partners’ institutions’ communication offices. The CRG communication office has extensive experience disseminating scientific results to the public. In 2013 the communication office published 56 press releases and coordinated more than

560 appearances in print, radio and TV media. The CRG communication office also runs a primary education program and organizes public scientific lectures.

The consortium will publish project results in open-access journals wherever possible; in cases where this is not possible, the partners will pay article-processing charges (APC) to make publications open-access. The whole-cell model training data, model software, and model predictions will be available open-source at our project website as well as at SimTK and GitHub. All experimental data generated by the project will be publicly available and downloadable at our project website.

Prior to publication we will take adequate measures to identify and protect any potentially commercially valuable intellectual property. Before starting the project all partners will sign a consortium agreement which will define the intellectual property rights protection strategy, access rights, ownership, disclosure policies, and material exchanges. In addition, through existing collaborations with MSD, we will explore the possibility of commercializing the resulting bacterial therapy chassis.

**Impact of results, including potential economic applications and societal benefits**

The long-term goal of the MiniCell project is to provide a new avenue to treat human disease and increase health through bacterial therapy. Within the proposed three year timeframe of the MiniCell project we aim to develop a reduced, predictable bacterial chassis, which in the future, could be extended to deliver drugs or vaccines to patients. Here we describe the technology readiness of the MiniCell project in terms of four interrelated research sections:

1. Developing the first whole-cell computational model based on organism-specific experimental data, and using the model to guide genome engineering. Most immediately, the model will help researchers examine connections among individual cellular pathways and processes. The model will also accelerate the functional characterization of important, but poorly understood genomic elements including structural regions, non-coding RNAs, and small proteins. Second, together with data from the transposon essentiality and infectivity analyses, we will use the model to identify essential and non-essential and virulence factors. These virulence factors are potential vaccine targets which could have multiple applications in human health and agriculture. The non-essential virulence factors are genes which could be removed to create a more predictable reduced strain. In addition to having high scientific interest, these findings have high patentable value. **TRL 1 Basic principles observed and reported.**

2. Experimentally defined the minimal genome. Until recently the concept of the minimal genome has been focused on determining the essentiality of conventional ORFs (proteins larger than 100 amino acids (aa)). Here we will determine the essentiality of the entire genome, including all of the “dark matter” comprised of structural regions, non-coding RNAs,

regulatory elements and small ORFs (<100 aa), producing the first exhaustive analysis of genome essentiality. This work will have a significant impact on synthetic biology as it will enable us to comprehensively predict the minimal genome required for cellular life. This work is expected to be publishable in a high-impact journal because it would be the first comprehensive genome essentiality study to focus on small proteins and non-coding regulatory elements. **TRL 3 Analytical and experimental critical function and/or characteristic proof-of-concept**

3. Identification of critical virulence and pathogenic factors. Non-essential antigens and virulence factors would be valuable targets for novel human and animal vaccines. In addition, accurate characterization of the virulence factors is needed to develop an effective and non- pathogenic bacterial chassis capable of supplying human patients small molecules and/or protein therapies. **TRL 4 Component/subsystem validation in laboratory environment.**

4. Minimal chassis for bacterial delivery. We will develop the first model-driven bacterial chassis designed to deliver therapy human therapy. In the future this chassis could be used to safely synthesize and deliver biological and/or chemical compounds to patients. **TRL 2**

**Technology concept and/or application formulated.**

The expected results will reach technology readiness levels 1 to 4, indicating that the proposed research is innovative and in the early stages of valuable commercial application. As mentioned previously, through their ongoing collaborations with MSD, the partners will explore potential commercial applications of the proposed bacterial chassis. The consortium will invite MSD researchers to the final consortium meeting.

**Data management**

The Karr group (Mt Sinai) will manage all project data. All partners will transfer any existing data, including all relevant metadata, which can be used to train the proposed whole-cell model as well as all data collected by the project to Mt Sinai. Mt Sinai will assemble this data into a single, hyperlinked knowledge base. Mt Sinai will also develop and manage a database of all whole-cell model simulations. Both databases will be publicly accessible via simple web-based interfaces available on the MiniCell website. Mt Sinai will maintain both databases responsibly and sustainably, including regularly backing up their software and content. The databases will not store any confidential data.

In addition, the Karr group will manage the development and open-source dissemination of the proposed whole-cell modeling platform. The platform will be implemented in MATLAB and available open-source through the project website and GitHub, ensuring that the software is maximally useful and reusable by other researchers.

**5. Consortium**

This interdisciplinary project requires the collaboration of multiple research groups with expertise in computational systems biology, *Mycoplasma* molecular biology, genome engineering, and host infectivity.

**Whole-cell modeling**

Recently the Karr group (Mt Sinai) led the development of the first whole-cell model of the closely related species *M. genitalium*. They have the expertise and knowledge to develop a whole-cell model of *M. pneumoniae*, and to use the model to guide *M. pneumoniae* genome

reduction and optimization. The Karr group will work in close partnership with the Serrano group (CRG) to develop the *M. pneumoniae* whole-cell model. This will combine their individual expertise in computational systems biology and molecular biology to produce the most accurate model of an individual organism to date.

***M. pneumoniae* functional genomics and infectivity analysis**

The Serrano group (CRG) has characterized the *M. pneumoniae* methylome, transcriptome, proteome, and metabolome, and will collaborate with the Karr group to integrate these data into a comprehensive model of *M. pneumoniae*. The Stülke group (UGOE) has expertise in *M. pneumoniae* metabolism and its role in virulence. The Stülke group has also developed a genetic tool set for working with *M. pneumoniae* that will be applied in this project. The CRG, in collaboration with the UGOE, will experimentally identify the essential regions of the *M. pneumoniae* genome. The essentiality information will be used to validate the whole-cell model developed by Mt Sinai. The Blanchard group (INRA) will engineer reduced, non- pathogenic *M. pneumoniae* strains by in-yeast genome transplantation based on model predictions by Mt Sinai. The CRG and UGOE will characterize the growth and infectivity pheonotypes of these reduced strains.

**Genome engineering**

The Blanchard group (INRA) has extensive expertise in *Mycoplasma* physiology and genome engineering and transplantation. Recently, they developed new technologies for in-yeast genome engineering, enabling seamless removal or the replacement of entire genomic regions. The Blanchard group will create multiple deletion strains, and send these strains to the CRG and UGOE who will characterize their growth and infectivity phenotypes. The Blanchard group also has strong expertise in comparative genomics, has maintained the MolliGen mollicutes comparative genomics database for over ten years ([http://www.molligen.org](http://www.molligen.org/)), and has significantly contributed to the identification minimal *Mycoplasma* gene sets.

**6. Project management and reporting (WP7)**

The project will be coordinated and managed by the CRG International and Scientific Affairs

(ISA) office, with support from the CRG communication, legal and technology transfer offices.

**Objective**

• Manage the project to achieve all of the objectives with the highest quality within the proposed time frame and budget

• Monitor results and coordinate reporting

• Ensure all ethical issues are properly addressed

• Plan, organize, and support all project’s and training activities and events

• Oversee dissemination of the project’s results to the scientific and non-scientific communities

• Identify and supervise results with commercial and/or medical potential

**Task 7.1** Coordinate partners: supervise all activities performed by the partners, foster collaborations, and suggest corrective and/or alternative strategies in case of any unexpected bottlenecks or ethical issues; prepare a consortium agreement based on the DESCA model and a confidentiality agreement with MSD before scientific exchange.

**Task 7.2** Manage consortium: Organize all consortium meeting and events; use electronic communication tools (intranet, mailing lists, phone conferences, etc.) to facilitate effective communication among the partners and between the coordinator and ERASynBio. The MiniCell partners will meet monthly via video-conferencing to discuss progress. In additional, the project website will include a forum to enable partners to discuss the project with the broader scientific community.

**Task 7.3** Report progress to ERASynBio and national funding agencies: Assemble midterm and final project reports; present project progress at ERASynBio status meetings.

**Task 7.4** Support in organizing a scientific workshops; coordinate trainee exchange visits among the partner laboratories and participation in the workshop.

**Task 7.5** Coordinate research dissemination: Construct a project website to disseminate results to the scientific community and general public, including the proposed open-source software and databases; facilitate publication of research findings in peer-review journals (encouraging joint publications); facilitate presentation of research findings at conferences. The project website will follow the EC communication guidelines.

**Task 7.6** Identify results with potential commercial value: The CRG technology transfer office will review the project’s progress each year in order to identify discoveries with potential commercial value. The CRG technology transfer office will discuss all potentially valuable discoveries with the technology transfer offices of the partners’ institutions.

**Personnel**

|  |  |
| --- | --- |
| Participant short name | Person-months per  participant |
| CRG | 1 |

**Deliverables**

|  |  |  |  |
| --- | --- | --- | --- |
| Number | Deliverable title | Lead  Partner | Delivery  Date (month) |
| D.PM1 | Final consortium agreement | CRG | 3 |
| D.PM2 | MiniCell website (public and intranet) | CRG | 6 |
| D.PM3 | Midterm report (progress, commercial potential of  results, training activities, dissemination) | CRG | 18 |
| D.PM4 | Scientific workshop | CRG | 24 |
| D.PM5 | Final report (progress, commercial potential of results, training activities, dissemination) | CRG | 36 |

**Risk and contingency**

The CRG has the infrastructure and human resources to oversee and carry out all the proposed activities and overcome any obstacles. Furthermore, the CRG has extensive expertise in coordinating and managing scientific projects from FP6 and FP7. To ensure smooth progress, prior to the start of the project all partners will sign a consortium agreement (based on the DESCA model) which outlines each partner’s rights, responsibilities, and benefits.

**7. Ethical and legal issues**

The MiniCell project does not involve the use of human subjects, human materials, non- human primates, or other mammals.

Biosafety is a main concern in all our laboratories. As Synthetic Biology is emerging, the regulatory issues need to adapt to the new possibilities, hence the new risks of these technologies. We participate to the debates with the relevant authorities to help the decision- making process. C. Lartigue (INRA, Blanchard group) has presented her work during a special session dedicated to Synthetic Biology convened by the “Comité scientifique du Haut Conseil des Biotechnologies (HCB)”, in December 2010. In December 2012, A. Blanchard (INRA) contributed to the "SynBio Workshop (Paris 2012) - Risk assessment challenges of synthetic biology" co-organized by the HCB and other equivalent bodies from other European countries; this meeting resulted in a publication (*63*).

Ethical issues and implication of the INRA group in the public debate

The ethical implications of our Synthetic Biology projects have to be considered very carefully. Following publication from the work done at the J. Craig Venter Institute, C. Lartigue participated to the public debate, including in French newspapers (Sud-Ouest, Le Monde), about the new possibilities offered by the chemical synthesis of whole genomes, the cloning of bacterial genomes in yeast and the transplantation in recipient cells. We consider that it is our responsibility to continue to contribute to this debate and to provide all the information to help decision makers understand the issues correctly. In September 2011, the INRA group met members of the INRA-CIRAD ethical comity to discuss the ethical impact of our projects; recently, this committee has produced a report in 2014 for these French research organizations; this note is available on the web:<http://institut.inra.fr/Missions/Promouvoir-ethique-et-deontologie/Avis-du-comite-d-ethique/5e-> avis-Biologie-de-synthese. In December 2013, A. Blanchard co-organized with the Regional Council of Aquitaine a symposium dedicated to “Synthetic biology and Green Chemistry” which gathered more than 200 participants from both academics and industry. Among the speakers, we invited Thierry Magnin (Université catholique de Lyon, Member of the UNESCO bioethics group) to give his own perspective on the ethical issues facing the development of synthetic biology. Finally, in 2014 a student in sociology, Benjamin Rimbaud, during the course of his PhD visited our laboratory during two weeks to study the organization of a laboratory working in Synthetic Biology. Therefore, by integrating ethical and social questions in the research process, we are following the recommendations listed by ERASynBio in its strategic vision document (*64*).

**France:** The Blanchard group has legal authorization to genetically modify mollicutes including using synthetic biology tools (Agreement No. 5524; July 25, 2011). All *M. pneumoniae* genomic engineering will be conducted in a biosafety level 2 laboratory.

**Spain:** The CRG has approval from the National Biosafety Commission to perform genetically modification in low risk installations (Notification No: A/ES/12/I-33) including *M. pneumoniae* in type 2 facilities (Notification No. A/ES/12/37). These activities comply with all applicable European (European Directive 2000/54/EC) and national legislation (RD Biosafety Law 951/97, Biological Agents RD 664/97) and CRG procedures.

**Germany:** The Stülke lab at UGOE has legal approval by the Gewerbeaufsichtsamt Göttingen (June 18, 2014) to conduct research using *M. pneumoniae* in an S2 lab. An S2 lab is available in the department (AZ 501.40611/0102/117).

**USA:** The Karr group will develop the *M. pneumoniae* whole-cell computational model, and therefore has no biosafety risks. This model is expected to significantly decrease the number of laboratory experiments needed to engineer *M. pneumoniae* for bacterial lung therapy, thereby reducing the use of harmful chemicals, *in vitro* experiments, animal subjects, and patient samples, as well as minimizing the laboratory handling of pathogenic bacteria.

**Bibliography**

1. M. Guell *et al.*, Transcriptome complexity in a genome-reduced bacterium. *Science* **326**, 1268 (Nov 27, 2009).

2. E. Yus *et al.*, Transcription start site associated RNAs in bacteria. *Molecular systems biology* **8**,

585 (2012).

3. M. Lluch-Senar *et al.*, Comprehensive methylome characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. *PLoS genetics* **9**, e1003191 (2013).

4. S. Kuhner *et al.*, Proteome organization in a genome-reduced bacterium. *Science* **326**, 1235 (Nov 27, 2009).

5. T. Maier *et al.*, Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Molecular systems biology* **7**, 511 (2011).

6. V. van Noort *et al.*, Cross-talk between phosphorylation and lysine acetylation in a genome- reduced bacterium. *Molecular systems biology* **8**, 571 (2012).

7. E. Yus *et al.*, Impact of genome reduction on bacterial metabolism and its regulation. *Science*

**326**, 1263 (Nov 27, 2009).

8. J. A. Wodke *et al.*, Dissecting the energy metabolism in Mycoplasma pneumoniae through genome-scale metabolic modeling. *Molecular systems biology* **9**, 653 (2013).

9. G. A. Benders, Cloning whole bacterial genomes in yeast. *Methods Mol Biol* **852**, 165 (2012).

10. J. R. Karr *et al.*, A whole-cell computational model predicts phenotype from genotype. *Cell*

**150**, 389 (Jul 20, 2012).

11. E. Chiner, J. Signes-Costa, A. L. Andreu, L. Andreu, [Mycoplasma pneumoniae pneumonia:

and uncommon cause of adult respiratory distress syndrome]. *An Med Interna* **20**, 597 (Nov,

2003).

12. S. Halbedel, J. Busse, S. R. Schmidl, J. Stulke, Regulatory protein phosphorylation in

Mycoplasma pneumoniae. A PP2C-type phosphatase serves to dephosphorylate HPr(Ser-P). *J Biol Chem* **281**, 26253 (Sep 8, 2006).

13. E. P. Caragata, T. Walker, Using bacteria to treat diseases. *Expert opinion on biological therapy* **12**, 701 (Jun, 2012).

14. R. Waehler, S. J. Russell, D. T. Curiel, Engineering targeted viral vectors for gene therapy.

*Nature reviews. Genetics* **8**, 573 (Aug, 2007).

15. C. K. Baban, M. Cronin, D. O'Hanlon, G. C. O'Sullivan, M. Tangney, Bacteria as vectors for gene therapy of cancer. *Bioengineered bugs* **1**, 385 (Nov-Dec, 2010).

16. H. Lage, J. H. Fruehauf, Delivery of therapeutic RNA molecules to cancer cells by bacteria.

*Therapeutic delivery* **2**, 441 (Apr, 2011).

17. T. L. Chang *et al.*, Inhibition of HIV infectivity by a natural human isolate of Lactobacillus jensenii engineered to express functional two-domain CD4. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11672 (Sep 30, 2003).

18. J. D. Hillman, Genetically modified Streptococcus mutans for the prevention of dental caries.

*Antonie van Leeuwenhoek* **82**, 361 (Aug, 2002).

19. K. Van Huynegem, M. Loos, L. Steidler, Immunomodulation by genetically engineered lactic acid bacteria. *Front Biosci (Landmark Ed)* **14**, 4825 (2009).

20. L. Steidler, P. Rottiers, B. Coulie, Actobiotics as a novel method for cytokine delivery. *Annals of the New York Academy of Sciences* **1182**, 135 (Dec, 2009).

21. M. A. Sze, J. C. Hogg, D. D. Sin, Bacterial microbiome of lungs in COPD. *International journal of chronic obstructive pulmonary disease* **9**, 229 (2014).

22. M. L. Shuler, P. Foley, J. Atlas, Modeling a minimal cell. *Methods Mol Biol* **881**, 573 (2012).

23. R. Gil, F. J. Silva, J. Pereto, A. Moya, Determination of the core of a minimal bacterial gene set. *Microbiol Mol Biol Rev* **68**, 518 (Sep, 2004).

24. A. R. Mushegian, E. V. Koonin, A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci U S A* **93**, 10268 (Sep 17, 1996).

25. J. I. Glass *et al.*, Essential genes of a minimal bacterium. *Proc Natl Acad Sci U S A* **103**, 425 (Jan 10, 2006).

26. C. A. Hutchison *et al.*, Global transposon mutagenesis and a minimal Mycoplasma genome.

*Science* **286**, 2165 (Dec 10, 1999).

27. G. L. Murray *et al.*, Genome-wide transposon mutagenesis in pathogenic Leptospira species.

*Infect Immun* **77**, 810 (Feb, 2009).

28. N. R. Salama, B. Shepherd, S. Falkow, Global transposon mutagenesis and essential gene analysis of Helicobacter pylori. *J Bacteriol* **186**, 7926 (Dec, 2004).

29. M. A. Jacobs *et al.*, Comprehensive transposon mutant library of Pseudomonas aeruginosa.

*Proc Natl Acad Sci U S A* **100**, 14339 (Nov 25, 2003).

30. B. Christen *et al.*, The essential genome of a bacterium. *Mol Syst Biol* **7**, 528 (2011).

31. C. T. French *et al.*, Large-scale transposon mutagenesis of Mycoplasma pulmonis. *Mol*

*Microbiol* **69**, 67 (Jul, 2008).

32. B. J. Akerley *et al.*, A genome-scale analysis for identification of genes required for growth or survival of Haemophilus influenzae. *Proc Natl Acad Sci U S A* **99**, 966 (Jan 22, 2002).

33. J. E. Griffin *et al.*, High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog* **7**, e1002251 (Sep, 2011).

34. L. A. Gallagher *et al.*, A comprehensive transposon mutant library of Francisella novicida, a bioweapon surrogate. *Proc Natl Acad Sci U S A* **104**, 1009 (Jan 16, 2007).

35. G. C. Langridge *et al.*, Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. *Genome Res* **19**, 2308 (Dec, 2009).

36. M. Rajewska, K. Wegrzyn, I. Konieczny, AT-rich region and repeated sequences - the essential

elements of replication origins of bacterial replicons. *FEMS microbiology reviews* **36**, 408 (Mar, 2012).

37. J. P. Bardill, B. K. Hammer, Non-coding sRNAs regulate virulence in the bacterial pathogen

Vibrio cholerae. *RNA biology* **9**, 392 (Apr, 2012).

38. C. Lartigue *et al.*, Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325**, 1693 (Sep 25, 2009).

39. C. Lartigue *et al.*, Genome transplantation in bacteria: changing one species to another.

*Science* **317**, 632 (Aug 3, 2007).

40. S. M. Wong, J. D. Gawronski, D. Lapointe, B. J. Akerley, High-throughput insertion tracking by deep sequencing for the analysis of bacterial pathogens. *Methods Mol Biol* **733**, 209 (2011).

41. S. Halbedel, J. Stulke, Tools for the genetic analysis of Mycoplasma. *Int J Med Microbiol* **297**,

37 (Feb, 2007).

42. D. R. Bentley *et al.*, Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53 (Nov 6, 2008).

43. A. L. Osterman, S. Y. Gerdes, *Microbial Gene Essentiality*. Methods in Molecular Biology

(Humana Press Inc, Totowa, NJ, 2008), vol. 416.

44. C. Hames, S. Halbedel, M. Hoppert, J. Frey, J. Stulke, Glycerol metabolism is important for cytotoxicity of Mycoplasma pneumoniae. *Journal of bacteriology* **191**, 747 (Feb, 2009).

45. S. R. Schmidl *et al.*, A trigger enzyme in Mycoplasma pneumoniae: impact of the glycerophosphodiesterase GlpQ on virulence and gene expression. *PLoS pathogens* **7**, e1002263 (Sep, 2011).

46. T. R. Kannan *et al.*, Synthesis and distribution of CARDS toxin during Mycoplasma pneumoniae infection in a murine model. *The Journal of infectious diseases* **204**, 1596 (Nov

15, 2011).

47. J. D. Kornspan, M. Tarshis, S. Rottem, Adhesion and biofilm formation of Mycoplasma pneumoniae on an abiotic surface. *Archives of microbiology* **193**, 833 (Nov, 2011).

48. E. B. Spuesens *et al.*, Sequence variations in RepMP2/3 and RepMP4 elements reveal intragenomic homologous DNA recombination events in Mycoplasma pneumoniae. *Microbiology* **155**, 2182 (Jul, 2009).

49. R. K. Grover *et al.*, A structurally distinct human mycoplasma protein that generically blocks

antigen-antibody union. *Science* **343**, 656 (Feb 7, 2014).

50. S. R. Schmidl *et al.*, The stability of cytadherence proteins in Mycoplasma pneumoniae requires activity of the protein kinase PrkC. *Infection and immunity* **78**, 184 (Jan, 2010).

51. S. Grosshennig, S. R. Schmidl, G. Schmeisky, J. Busse, J. Stulke, Implication of glycerol and phospholipid transporters in Mycoplasma pneumoniae growth and virulence. *Infection and immunity* **81**, 896 (Mar, 2013).

52. R. Krishnakumar *et al.*, Targeted chromosomal knockouts in Mycoplasma pneumoniae.

*Applied and environmental microbiology* **76**, 5297 (Aug, 2010).

53. J. D. Jaffe, H. C. Berg, G. M. Church, Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* **4**, 59 (Jan, 2004).

54. D. C. Krause *et al.*, Transposon mutagenesis reinforces the correlation between Mycoplasma pneumoniae cytoskeletal protein HMW2 and cytadherence. *Journal of bacteriology* **179**,

2668 (Apr, 1997).

55. M. J. Willby, D. C. Krause, Characterization of a Mycoplasma pneumoniae hmw3 mutant:

implications for attachment organelle assembly. *Journal of bacteriology* **184**, 3061 (Jun,

2002).

56. M. F. Balish, S. M. Ross, M. Fisseha, D. C. Krause, Deletion analysis identifies key functional domains of the cytadherence-associated protein HMW2 of Mycoplasma pneumoniae. *Molecular microbiology* **50**, 1507 (Dec, 2003).

57. U. Gueldener, J. Heinisch, G. J. Koehler, D. Voss, J. H. Hegemann, A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* **30**, e23 (Mar 15, 2002).

58. V. N. Noskov, T. H. Segall-Shapiro, R. Y. Chuang, Tandem repeat coupled with endonuclease cleavage (TREC): a seamless modification tool for genome engineering in yeast. *Nucleic Acids Res* **38**, 2570 (May, 2010).

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

60. G. Prakash, M. G. Gabridge, Influence of the fusogenic agent polyethylene glycol on

attachment of Mycoplasma pneumoniae to other cells. *Infection and immunity* **32**, 969 (May,

1981).

61. W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31**, 233 (Mar, 2013).

62. S. K. Sharan, L. C. Thomason, S. G. Kuznetsov, D. L. Court, Recombineering: a homologous recombination-based method of genetic engineering. *Nature protocols* **4**, 206 (2009).

63. R. M. K. Pauwels, C. Golstein, D. Breyer ; P. Herman, M Kaspari, JC Pages, H. Pfister, F. van der Wilk. B. Schönig, Risk assessment challenges of Synthetic Biology. *Event report: SynBio Workshop (Paris 2012) J. Verbr. Lebensm.* **8**, 215 (2013).

64. . (2014).