Impact of donor–recipient phylogenetic distance on bacterial genome transplantation

Fabien Labroussaa¹,²,†, Anne Lebaudy¹,²,†, Vincent Baby³, Géraldine Gourgues¹,², Dominick Matteau³, Sanjay Vashee⁴, Pascal Sirand-Pugnet¹,², Sébastien Rodrigue³ and Carole Lartigue¹,²,*

¹INRA, UMR 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d’Ornon, France, ²University of Bordeaux, UMR 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d’Ornon, France, ³Université de Sherbrooke, Département de biologie, 2500 boulevard Université Sherbrooke (Québec), J1K 2R1, Canada and ⁴J. Craig Venter Institute, Rockville, MD 20850, USA

ABSTRACT

Genome transplantation (GT) allows the installation of purified chromosomes into recipient cells, causing the resulting organisms to adopt the genotype and the phenotype conferred by the donor cells. This key process remains a bottleneck in synthetic biology, especially for genome engineering strategies of intractable and economically important microbial species. So far, this process has only been reported using two closely related bacteria, Mycoplasma mycoides subsp. capri (Mmc) and Mycoplasma capricolum subsp. capricolum (Mcap), and the main factors driving the compatibility between a donor genome and a recipient cell are poorly understood. Here, we investigated the impact of the evolutionary distance between donor and recipient species on the efficiency of GT. Using Mcap as the recipient cell, we successfully transplanted the genome of six bacteria belonging to the Spiroplasma phylogenetic group but including species of two distinct genera. Our results demonstrate that GT efficiency is inversely correlated with the phylogenetic distance between donor and recipient bacteria but also suggest that other species-specific barriers to GT exist. This work constitutes an important step toward understanding the cellular factors governing the GT process in order to better define and eventually extend the existing genome compatibility limit.

INTRODUCTION

Synthetic biology encompasses a variety of disciplines that aim to design and create new systems either derived from natural organisms (top-down approaches) or built from scratch (bottom-up approaches). From a fundamental perspective, being able to design and build a biological system stands as a very practical measure of understanding of how this system operates and can help to identify molecular components essential for life. The engineering of new and well understood cellular chassis would provide platforms for innovative applications in biotechnology, agriculture or medicine. Different and complementary approaches are currently developed using bacteria that have lost much of their genomes through regressive evolution (1–3). Among them, Mollicutes are considered as the best representatives of natural minimal cells with some of the smallest genomes among free-living organisms (0.58 to ~2.2 Mbp) that can grow in axenic medium (4). Innovative strategies toward the production of minimalist chassis rapidly emerged from global gene inactivation experiments (3, 5) and from comparative genome analyses for a better comprehension of minimal cellular machineries such as the translation apparatus (6). A recent breakthrough was achieved with the design and synthesis of Syn3.0 (0.531 Mbp; 473 genes) (7), a reduced version of the 1.08 Mbp Mycoplasma mycoides subsp. capri (Mmc) natural genome. This work represents a technological milestone with the creation of the smallest genome ever carried by a microorganism capable of autonomous replication. It is also a step forward in defining the minimal set of genes essential for life and the possibility to customize bacterial genomes.

The drastic minimization of Mmc genome succeeded because of the emergence of cutting-edge technologies allowing the synthesis and manipulation of its genetic information on a genome-wide scale. Historically, mycoplasmas have long been considered intractable for genome engineering (4), just like a large majority of non-model microorganisms. Newly developed approaches now rely on the (i) cloning and (ii) engineering of the whole Mmc genome into...
an intermediate host possessing well-developed tools followed by (iii) the back-transplantation of the newly engineered genome into suitable bacterial recipient cells.

Because of its cloning capacities and large repertoire of available genetic tools, the yeast *Saccharomyces cerevisiae* was chosen as the host for whole bacterial genome manipulations and first studies have consisted in the cloning of mycoplasma genomes into *S. cerevisiae* as yeast centromeric plasmids (8,9). So far, seven complete bacterial genomes have been cloned in yeast, including micro-organisms that, unlike most *Mollicutes* (10), use the universal genetic code (11,12). Known key factors limiting in-yeast cloning are the genome size, high G + C %, the absence of ARS-like sequences (autonomous replication sequences) (13) and the presence of any toxic genes (11).

Following in-yeast cloning, new yeast molecular and editing tools have been developed to manipulate cloned bacterial genomes. For instance, the TREC system (Tandem Repeat Endonuclease Cleavage) (14) and TREC-IN derivatives (15) allow the deletion, insertion, or replacement of genes without any scars on bacterial chromosomes cloned in yeast. Such systems have notably been used in mycoplasma to assess gene function (8,16,17) and for genome minimization purposes (15). More recently, the CRISPR/Cas9 system, first developed for yeast mutagenesis purposes (18), has been also adapted for the engineering of bacterial genomes cloned in yeast (19).

Altogether, in-yeast cloning strategies associated with genome engineering tools now pave the way to high-throughput manipulation of natural or synthetic genomes in yeast, allowing the study of virtually any gene of interest on a bacterial chromosome. However, for this extremely powerful set of methods to be widely used, engineered bacterial chromosomes need to be isolated from yeast and transplanted back into a recipient bacterium that provides compatible cellular machineries able to replicate and express its genetic information. This critical whole genome transplantation (GT) step was first established with a whole genome isolated from bacteria (20) and then with natural and synthetic bacterial genomes cloned in yeast (7,8,21). These results demonstrated that a bacterial genome can be designed, synthetized, genetically manipulated on a large scale and transplanted into a recipient cell resulting in a living cell that is genetically programmed by the engineered genome. In these cases, two closely related species, sharing more than 99% identity on 16S rDNA as well as on their core proteome (6; also see Supplementary Table S2), *Mnc* and *Mycoplasma capricolum* subsp. *capricolum* (*Mcap*) were respectively used as donor genome and recipient cell. However, GT technology is currently limited to these two bacterial species and this limitation prevents the demonstration of the strategy for other non-model microorganisms with agricultural, medical or biotechnological interest. In particular, mechanisms underlying the GT process are still unclear and major technological difficulties might arise every time GT technology needs to be applied to other bacterial species.

In order to circumvent these difficulties, we explored the effect of the phylogenetic distance between donor genomes and the recipient cell on GT efficiency using seven species from the Spiroplasma phylogenetic group with increasing phylogenetic distance from the *Mcap* recipient cell. Our results demonstrate for the first time that GT can be achieved with several *Mcap* related genomes, including a *Mollicutes* belonging to a different genus. Overall, GT efficiency decreases as phylogenetic distance increases, suggesting that a compatibility limit is eventually reached between incoming genome and recipient cell enzymatic machineries. With the cloning of several *Mollicutes* genomes in yeast, this work offers efficient genetic engineering tools in these otherwise difficult-to-manipulate organisms.

**MATERIALS AND METHODS**

**Bacterial and yeast strains, culture conditions**

Competent *Escherichia coli* cells (Electromax DH10B from Invitrogen) [*F−*–*mcrAΔ(mrr-hsdRMS-mcrBC) φ80dlacZ ΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK*–*rpsL*–*mpG] served as the host strain for cloning experiments and plasmid propagation. *E. coli* cells transformed with plasmids were grown at 37°C in Luria-Bertani (LB) broth medium or in LB agar supplemented with 100 μg/ml of ampicillin.

*Saccharomyces cerevisiae* strain W303a (*MATa his3-11, 15 trp1-1 leu2-3,112 ura3-1 ade2-1 can1-100*) *VL6-48N* (*MATa trp1-Δ1 ura3-Δ1 ade2-101 his3-Δ200 lys2 met14 cir*) (22) was cultured at 30°C in YPD medium (Clontech) according to a standard protocol (23). Yeast transformed with mycoplasma genomes was grown in minimal SD Base medium (Clontech), complemented with —HIS DO supplement (Clontech) (SD-HIS medium).

Wild-type *Mcap* (*wtMcap*) strain California KidT (ATCC 27343) was used in this study as well as a restriction-free *Mcap* mutant (*McapΔRE*) obtained by inactivation of the CCATC-restriction enzyme in the wild-type strain (8). *McapΔRE* harbors (i) a puromycin resistance marker for selection in mycoplasma background and (ii) yeast elements (a centromere CEN6, an autonomously replicating sequence ARSH4 and the auxotrophic marker HIS) for selection and propagation of the *Mcap* genome as a centromeric plasmid in yeast. Wild-type *Mycoplasma leachii* strain PG50 (*wtMlea*) (24,25), wild-type *Mnc* strain GM12 (*wtMnc*) (24), wild-type *Mycoplasma mycoides* subsp. *mycoides* strain PG1 (wt*Mmm*) (26), wild-type *Mycoplasma putrefaciens* strain 156 (*wtMputr*), wild-type *Mycoplasma flororum* strain L1 (*wtMflorum*; ATCC 33433) (27) and wild-type *Spiroplasma citri* strain GI13 (*wtScitri*; ATCC 27556) (28) were transformed with the newly constructed vector pMT85tetM-PSlacZ-pRS313 (Supplementary Figure S1) as described by others (9,29), then selected as donor species for GT experiments.

All species were cultured in SP5 medium, deriving from the original SP4 medium (30). SP5 medium is composed of 3.5 g/l of Mycoplasma broth base (Fisher Scientific), 10 g/l of Bacto Tryptone (Fisher Scientific) and 5.3 g/l of Bacto Peptone (Fisher Scientific). The solution was adjusted to pH 7.5, autoclaved for 20 min at 120°C, then supplemented with 0.125% (w/v) glucose, 5% (w/v) CMRL 1066 10× (Invitrogen), 0.11% (w/v) sodium bicarbonate, 1 mM L-glutamine, 3.5% (v/v) yeast extract (Fisher Scientific), 0.2% (w/v) TC yeastolate, 17% (v/v) fetal bovine serum, 0.1 mg/ml ampicillin and 0.002% (w/v) phenol red.
Mycoplasma strains were all cultured at 37°C, whereas *Mflorum* and *Scitri* were cultured at 30 and 32°C, respectively. Tetracycline was added to the medium when needed at concentrations ranging from 2 to 15 µg/ml, depending on the species.

For transplantation experiments, *McapΔRE* recipient cells were grown at 30°C in super optimal broth (SOB) (31) supplemented with 17%(v/v) fetal bovine serum, glucose at 10 g/l, 0.002%(w/v) phenol red, and penicillin at 0.5 µg/ml (SOB (+) medium).

**Phylogenetic reconstruction and similarity matrix**

The phylogenetic tree was generated using concatenated multiple alignments of selected 79 orthologous proteins involved in translation as previously described (6). Multiple alignments were generated using MUSCLE (32), concatenated using Seaview (33) and curated from unreliable sites involving the Maximum Likelihood method using PhyML (35) available on the web server Phylogeny.fr (36). The list of *Mollicutes* analyzed with some of their genomic characteristics is given in Supplementary Table S1. A similarity matrix was produced from the final concatenated alignment using BLASTP 2.4.0 (37).

**Construction of oriC-based replicative plasmids**

The *oriC*-based plasmids pMCO3, pMYCO1, pMYSO1 and pSD4, containing the origin of replication of *Mcap*, *Mmc*, *Mmm* and *Scitri*, respectively, have been previously described (38). *OriC* plasmids pMleaOriC, pMputrOriC and pMF1OriC containing the origin of replication of *Mlea*, *Mputr* and *Mflorum* were constructed for this study as described in Lartigue et al. (38). Primers used for their construction are summarized in Supplementary Table S4. *OriC* segments of these plasmids have a global organization identical to those previously constructed. Each region includes the dnaA gene of each species studied surrounded by natural intergenic regions on each side of the gene. Prior to being used for transformation into *McapΔRE*, the purified plasmids were verified by restriction analyses and the integrity of the dnaA gene sequences was checked by DNA sequencing. Transformations of *McapΔRE* were conducted using 1 or 10 µg of plasmids and transformants were grown on selective medium for 5–8 days depending on the species under study.

**Plasmid construction used to mark donor *Mollicutes* genomes**

A transposon-based plasmid (pMT85tetM-PSlacZ-pRS313, Supplementary Figure S1, accession number KX011460) designed for marking the genomes of all *Mollicutes* species selected as donor for GT experiments was constructed using the Gibson Assembly® method (39). Details for this construction are described in the Supplementary material. Ten to thirty micrograms of pMT85tetM-PSlacZ-pRS313 were used to transform wt*Mmm*, wt*Mlea*, wt*Mputr* species by PEG-mediated transformation (40) and wt*Scitri* by electroporation (41). A marked *Mflorum* donor genome was obtained by transformation of wt*Mflorum* as described by Baby et al. (V. Baby, F. Labroussaa, J. Brodeur, D. Matteau, G. Gourgues, C. Lartigue and S. Rodrigue, in preparation).

Transformants were grown in SP5 medium containing suitable concentration of tetracycline, filter cloned three times to ascertain the clonality of the population and stored at −80°C. Genomic DNA from subclones was extracted with the Wizard Genomic DNA Purification Kit (Promega) and further analyzed by direct sequencing to localize the transposon insertion site. Primers used for direct sequencing and localization of insertions into *Mollicutes* genomes are summarized in Supplementary Table S5.

**Isolation of intact *Mollicutes* genomes in agarose plugs**

The protocol for whole, intact genomic DNA isolation from bacteria cultures was done as previously described by Lartigue et al. (20) with some modifications. First, the amount of bacterial cells present in each agarose plug was accurately measured using a newly developed quantitative real-time polymerase chain reaction (qPCR) assay, (see Supplementary Figure S2 and Supplementary Data for details). This allowed the preparation of a reproducible series of plugs (2-fold serial dilutions) ranging from 8 to 0.25 µg of genomic DNA. The resuspension buffer used for *Mflorum* cells preparation was also modified. Briefly, a 50-ml *Mflorum* culture was centrifuged at 5000 g for 15 min at 4°C and resuspended into 25 ml of a cold resuspension buffer (8 mM Hepes, 272 mM Sucrose; pH 7.4). After repeating the centrifugation step, cells were resuspended into 500 µl of the same resuspension buffer and then mixed with a 2% low-melting-point agarose solution. This modification prevented cell aggregation when *Mflorum* cells were resuspended before their inclusion into agarose plugs. This protocol was applied to all the other species used in this study.

Preparation of yeast agarose plugs was performed according to Lartigue et al. (8).

**Genome transplantation into *McapΔRE* recipient cells**

Prior to transplantation experiments, entrapped *Mollicutes* genomes (isolated from bacterial cultures or yeast cells) were released from the agarose plug with β-agarase I (3 units/plug) per the manufacturer’s recommendations (New England Biolabs). All transplantation experiments in this study were performed with as previously developed using a 5% polyethylene glycol (PEG) mediated protocol (8,20), except for a few modifications found to improve transplantation/ transformation efficiencies. *McapΔRE* recipient cells were cultivated at 30°C instead of 37°C in SOB(+) and this temperature was also used during the 90 min recovery period after incubation of cells with PEG. Also, SP5 medium was used instead of SP4. This protocol was also used for the transformation of plasmids carrying each *oriC* of the different bacterial species transplanted. Depending of the transformation efficiencies observed, 1 or 10 µg of plasmid was used instead of 20 or 100 µl of melted agarose plug solutions, respectively used for from-bacteria or from-yeast transplantations.
Whole genome sequencing

For Mcap, Mmc, Mmm, Mlea, Mputr and Mflorum, gDNA from the donor strains and two from-bacteria transplants was extracted using the Wizard genomic DNA purification kit (Promega) for mycoplasmas and the Puregene yeast/bact. kit B (QIAGEN) for Mflorum using the manufacturer’s recommendations. For Mmc, Mlea and Mputr, four transplants obtained from-yeast GT assays were also processed. Illumina libraries were prepared as in Baby et al. (V. Baby, F. Labroussaa, J. Brodeur, D. Matteau, G. Gourgues, C. Lartigue and S. Rodrigue, in preparation) and sequenced on a HiSeq2000 instrument at the McGill University and Genome Québec Innovation Centre (Montreal, Canada). The resulting 50-bp paired-end reads were aligned to the available RefSeq genomes; Mcap NC_007633.1, Mmc NZ_CP001621.1, Mmm NC_005364.2, Mlea NC_014751.1, Mputr NC_015946.1. For Mflorum, the expected sequence containing the yeast elements and tetM gene was used and variant detection was performed as in Baby et al. (V. Baby, F. Labroussaa, J. Brodeur, D. Matteau, G. Gourgues, C. Lartigue and S. Rodrigue, in preparation). Briefly, the reads were trimmed using Trimmomatic (v. 0.32) (42) and aligned using BWA-mem (v. 0.7.10) (43). GATK (v. 3.3.0) (44) was used to detect variants, and VCFtools (v. 0.1.12b) (45) was used to compare the donor and transplant genomes.

Statistics

Transformation rates are the mean of three independent transformation experiments performed in triplicates. Transformation rates correspond to the average of four independent GT events obtained during the same GT experiment. Three independent GT experiments were conducted for each species. We used the Tukey’s HSD test for all pairwise comparisons among oriC-based plasmid transformation rates, from-bacteria and from-yeast GT experiments. From-bacteria and from-yeast GT data were treated independently for statistical calculations. We used a parametric paired t-test (46) to assess statistical difference during chloramphenicol treatment for each dilution point. qPCR data were log-transformed to meet model assumptions. The Benjamini–Hochberg (47) method was used to adjust the false discovery rate associated with the test.

RESULTS

Mcap is currently the only recipient bacterium that has been used successfully in GT experiments. The McapΔRE strain, a derivative of Mcap that was obtained by inactivating the gene encoding the CCATC-restriction endonuclease, was chosen as the recipient cell since this deletion eliminates the need for a DNA methylation step before GT (8). Here, seven bacterial species of the Spiroplasma phylogenetic group were selected as donor-genome strains to investigate the importance of phylogenetic distance on GT (Figure 1). A first group comprising Mmc, Mmm as well as Mlea, all belonging to the Mycoides cluster that also includes Mcap, are more closely related to the recipient cell (>99% identity on the 79 core proteins; Supplementary Table S2). A second group of species, more phylogenetically distant from Mcap (between 85 and 96% identity on the 79 core proteins; Supplementary Table S2), included Mputr, Mflorum and Scitri. All species can be cultured under the same growth conditions and use the alternative mycoplasma genetic code for gene expression, thus constituting plausible candidates for GT given our experimental framework.

Heterologous transformation of Mcap with oriC plasmids

Considering that few critical cellular processes are likely to ensure the ‘boot up’ of the new chromosome by the recipient cell, we surmised that the cellular machinery of Mcap should allow the replication of plasmids carrying the origin of replication region (oriC region) of donor genomes. We thus sought to evaluate whether McapΔRE can be transformed with plasmids containing the oriC region of the seven individual potential donor genomes. Genomic oriC regions are highly syntenic among the considered species and all oriC plasmids were designed similarly to include the dnaA gene and the intergenic regions located on each side. Four oriC plasmids, pMCO3 (Mcap), pMYCO1 (Mmc), pMYSO1 (Mmm) and pSD4 (Scitri) were previously shown to contain all functional elements required for efficient replication (38). Three additional plasmids were built using the same backbone to contain the complete ∼1.9–2.0 kbp oriC regions of Mlea, Mputr or Mflorum. All oriC plasmids were found to replicate in McapΔRE suggesting that the enzymatic machinery of the bacteria should be able to initiate the replication of genomes with a diversity of oriC. However, significant differences in transformation rates were observed (Figure 2). Strikingly, no direct relationship was observed between the transformation efficiencies of oriC plasmids and their phylogenetic distances to the recipient cell. Plasmids pMYSO1 and pMYCO1 containing oriC regions of Mmm and Mmc were transformed with higher efficiencies (>10^6 transformants/μg of plasmid) than the pMCO3 plasmid (∼10^2 transformants/μg of plasmid) even if this last one was derived from the genome of the recipient cell. Overall, these data support the choice of McapΔRE as a suitable recipient cell for GT experiments within the Spiroplasma phylogenetic group.

Insertion of genetic markers in donor Mollicutes genomes

GT can be accomplished using chromosomes directly isolated from bacteria (from-bacteria GT) or from yeast containing the cloned bacterial genome (from-yeast GT) (Figure 3). However, isolated donor genomes must harbor a marker (tetM, conferring resistance to tetracycline) for selection of bacterial transplants. For genomes cloned in yeast, additional elements are important to allow maintenance and selection of the chromosome once inserted in its intermediate host (i.e. ARSH4, CEN6 and HIS3). Mcap, Mmc and Mflorum genomes containing these yeast and bacterial elements were obtained from other studies ((8); V. Baby, F. Labroussaa, J. Brodeur, D. Matteau, G. Gourgues, C. Lartigue and S. Rodrigue, in preparation). We then constructed a novel transposon-based plasmid, named pMT85tetM-PSlacZ-pRS313 (Supplementary Figure S1) to allow a single-step integration of all elements. This plasmid was used to similarly mark the genomes of

Downloaded from https://academic.oup.com/nar/article-abstract/44/17/8501/2468051 by guest on 30 July 2018
Figure 1. Phylogenetic tree of Mollicutes. The phylogenetic tree was generated using concatenated multiple sequence alignments of selected 79 orthologous proteins involved in translation as described in Grosjean et al. (6). The genetic code used is indicated on the right side. Phylogenetic group: Spiroplasma; Pneumoniae; Hominis; Acholeplasma-Phytoplasma. Mcap used as the recipient cell is indicated with the (R) symbol. Species used as donor genomes, all belonging to the Spiroplasma phylogenetic group, are marked with an asterisk.

Figure 2. Transformation of McapΔRE with different oriC-based replicative plasmids. Plasmids containing complete oriC regions of all Mollicutes donor genomes used in GT experiments were tested. Letters indicate statistically different transformation rates (pair-wise comparisons).

Mmm, Mlea, Mputr and Scitri. For each of these species, a clone in which the transposon inserted into intergenic sequences or genes not likely to interfere with GT was used for the following experiments (Supplementary Table S5).

Impact of gDNA concentration on genome transplantation

Bacterial cultures from the seven Mollicutes species, followed by a qPCR assay developed to accurately estimate cell concentration, were used to prepare 2-fold dilution series of DNA plugs with similar donor genome gDNA concentrations (~400, 200, 100, 50, 25 and 12.5 ng/µl) (Supplementary Figure S2). Each plug of every series was then used in an independent GT experiment using McapΔRE as the recipient cell. Transplants were obtained for every species tested (Figure 4), except for Scitri, a species that remained refractory to transplantation in McapΔRE under all tested conditions (Figure 4B). A similar pattern was even observed between all species tested when transplantation rates obtained with all the gDNA concentrations used for
Figure 3. Global scheme representing the main steps conducted in this study. Donor-genome cells were first transformed with an integrative transposon-based plasmid. Newly marked genomes were then gently isolated and either directly back-transplanted into restriction free *Mcap* recipient cells for from-bacteria transplantation experiments or transformed into yeast spheroplasts for further from-yeast transplantation experiments. All putative transplants originating from from-bacteria and from-yeast transplantation experiments were genotypically and phenotypically characterized systematically.

Figure 4. Effect of gDNA quantity in from-bacteria GT assays. Different quantities of donor gDNA were tested in GT experiments using *Mcap*ΔRE as the recipient cell. Numbers of transplants obtained in relation to the gDNA quantity used are reported for (A) highly transplantable species (*Mcap, Mlea, Mmc*) or (B) poorly transplantable species (*Mmm, Mputr, Mflorum, Scitri*).

GT were compared. Transplantation rates increased when gDNA concentrations were comprised between 0 and 2 μg of gDNA and then progressively decreased when higher gDNA concentrations were used, giving bell-shape curves for all species (Figure 4). Surprisingly, the same optimal gDNA amount (~2 μg), corresponding to plugs containing ~100 ng/μl of bacterial DNA, was clearly needed for successful GT for all transplanted species. We categorized these species in two sub-groups based on their respective transplantation rates. *Mcap, Mlea* and *Mmc* were considered as highly transplantable with around or over 100 transplants obtained when gDNA quantities were in the range of 1–4 μg (Figure 4A). On the other hand, *Mputr, Mmm* and *Mflorum* were considered as poorly transplantable species, especially for the two latter with <10 transplants obtained regardless of the gDNA quantity used (Figure 4B).

**Importance of the phylogenetic distance in the GT process**

We next used the previously determined optimal gDNA quantity to accurately measure the transplantation rate for each of the seven targeted species (Figure 5A). All species considered, the transplantation of *Mcap* into its own genetic background resulted in the highest transplantation rate (Table 1). In contrast to results obtained for oriC plasmids transformation, the other genomes were successfully transplanted into *Mcap*ΔRE and a negative correlation between transplantation efficiencies and phylogenetic distance was observed (Figure 5A). The only exception to this general trend was *Mmm*, for which transplants were only obtained at a rate of ~3.43 transplants/μg of gDNA. This rate is significantly lower than its closest phylogenetic neighbors *Mlea* (~1.96 × 10^2 transplants/μg of gDNA) and *Mmc* (~7.44 × 10^1 transplants/μg of gDNA). Altogether, these results showed that we have extended GT technology to five new *Mollicutes* species, with *Mflorum* being the *Mcap* most-distant compatible genome in our experimental system (92.47% identity on the 79 core proteins).

**Transplantation assays using bacterial genomes cloned in yeast**

We next tested the ability of these seven donor *Mollicutes* genomes to be transplanted when they are cloned and isolated from yeast. *Mcap, Mmc* and *Mflorum* genomes were already cloned in yeast at the beginning of this work (9); V. Baby, F. Labroussaa, J. Brodeur, D. Matteau, G. Gourgues, C. Lartigue and S. Rodrigue, in preparation). For
Table 1. Optimal genomic DNA and transplantation rates obtained for each donor species tested, in either from-bacteria or from-yeast GT experiments

| Species transplanted                  | From-bacteria transplantation | From-yeast transplantation |
|---------------------------------------|-------------------------------|-----------------------------|
|                                       | Optimal gDNA concentration (µg) | Transplantation rate (transplants/µg DNA ± SEM) | Optimal gDNA concentration (µg) | Transplantation rate (transplants/µg DNA ± SEM) |
| M. capricolum subsp. capricolum CK    | 1.92                          | 2.38 × 10³ ± 3.93 × 10¹     | nd                           | nd                                        |
| M. leachii PG50                       | 2.29                          | 1.96 × 10³ ± 8.94 × 10¹     | 0.66                         | 2.11 × 10² ± 1.61 × 10²                     |
| M. mycoides subsp. capri GM12         | 2.02                          | 7.44 × 10¹ ± 9.66           | 0.71                         | 2.41 × 10¹ ± 1.69 × 10¹                     |
| M. mycoides subsp. mycoides PG1       | 2.30                          | 3.43 ± 2.86                 | nd                           | 0                                         |
| M. putrefaciens strain 156            | 1.99                          | 1.88 ± 1.27                 | 0.57                         | 2.67 × 10¹ ± 5.66                         |
| M. florum L1                          | 2.09                          | 3.02 ± 5.23                 | 0.52                         | 2.14 × 10¹ ± 6.84                         |
| S. citri GII3                         | nd                            | 0                           | nd                           | 0                                         |

SEM (standard error of the mean).

Figure 5. Transplantation rates obtained for (A) from-bacteria and (B) from-yeast GT assays. Optimal gDNA quantities were used to determine an optimal transplantation rate for all species tested. Nd, not determined. Letters indicate statistically different transformation rates (pair-wise comparisons).

other species (M. mycoides, M. leachii, M. putrefaciens, S. citri), intact donor genomes were released from agarose plugs and transformed into yeast spheroplasts as previously described (8,9) (Figure 3). After transformation, the integrity of all four individual genomes was assessed by multiplex PCR and pulsed-field gel electrophoresis (PFGE) analysis (data not shown).

All seven Mollicutes genomes cloned in yeast were isolated in agarose plugs then transplanted into M. capricolum recipient cells (Figure 3). The M. capricolum genome cloned in yeast was previously transplanted at a rate equivalent to ~10³ transplants/µg of gDNA, and was not re-attempted in this study (17). From-yeast transplantation of M. leachii, M. mycoides, M. putrefaciens and M. florum whole genomes were also successful (Figure 5B). The general trend is in good agreement with the negative correlation previously observed for GT using chromosomes directly isolated from bacteria, with overall rates decreasing as the phylogenetic distance from M. capricolum increased (Figure 5A and B). No transplants were obtained for S. citri, but more surprisingly we did not observe any GT events for M. mycoides.

Confirmation of transplant phenotypes and genotypes

Putative transplants were first observed as colonies on plates to provide a first assessment of the nature of the species obtained (Figure 6). The colony morphology of each transplant was in all aspects similar to the donor species. The colonies of all species tested displayed the typical Mollicutes ‘fried-egg’ phenotype, except for M. putrefaciens colonies showing a fuzzier form, brown color and uneven margin. Phenotypic distinctions could also distinguish the different species tested. Colonies from M. capricolum and M. mycoides showed very large colonies in comparison to M. leachii and M. florum that all had even and defined margins. Colony sizes were also different (Figure 6, 100 µm scale) with M. mycoides and M. leachii showing very large colonies in comparison to M. mycoides. All colonies were convex except for M. florum transplants that showed a rounded elevation near the margin associated with a flat center.

Transplants were further analyzed using (i) species-specific PCR, (ii) PFGEs and (iii) whole genome sequencing to ensure that they had the expected genotype. For each bacterial species, ten transplants were initially investigated by PCR using species-specific primers (Supplementary Table S6). All tested transplants harbored the amplicons corresponding to the expected donor genome and no sign of the M. capricolum recipient genome was detected (Supplementary Figure S3). Two of these ten transplants per species were subjected to PFGE analysis to validate the proper global organization of the transplanted chromosome. Again, profiles, in accordance with each respective donor genotype, were obtained (Supplementary Figure S4 and Table S7). Finally, the genome of the original donor bacteria was sequenced along with the same two transplants subjected to PFGE analysis. With the exception of a few single nucleotide polymorphisms and indels, the transplants’ genomes fully matched the expected sequence and sequences from the recipient chromosome were not detected at significant rates in unmapped reads (Supplementary Table S3). Taken together,
these data confirm that the recipient Mcap strain was effectively changed into five new species, adopting the phenotype and genotype conferred by the donor genome.

**DISCUSSION**

Many recent genome engineering techniques such as TREC, TREC-IN and CRISPR-Cas9 use *S. cerevisiae* as a ‘living workbench’ to allow on-demand editing of almost any genes or group of genes from bacterial genomes. However, the back-transplantation of such engineered genomes into a suitable recipient cell is an essential process to access and study this modified genetic information. Understanding the mechanisms governing GT is thus of major importance for the study of intractable microorganisms presenting interests in biotechnology, but also in biomedical applications with the development of innovative vaccines (48,49,50).

The success of a GT experiment depends on several factors, many remaining poorly understood or completely unknown. In order to learn more about the mechanisms underlying the GT process, we broached the question of the degree of relatedness necessary between donor and recipient cells for successful transplantation and pursued genetic factors that possibly govern this compatibility. We selected Mcap/Δ1RE as the recipient cell for several reasons: (i) this species, Mcap, was used as recipient cell in the initial work showing the transplantation of whole bacterial genomes (8,20), (ii) it supports the replication of plasmids carrying the chromosomal origin of replication of different donor genomes (Figure 2) (38) and (iii) the mutant McapΔRE does not express any restriction enzymes, which removes a known barrier in GT experiments (8).

The genomes that were evaluated for GT all belong to the same phylogenetic group within the *Mollicutes* and represent three different genera (Figure 1). The seven selected genomes were first modified by simultaneous integration of bacterial (*tetM*) and yeast (ARS4, CEN6 and HIS) genetic markers into their genomes. Following this marker integration, the genomes were successfully cloned as independent replicons in yeast. This study added four genomes (Mlea, Mmm, Mputr and Scitri) to the list of cloned genomes in yeast, confirming that this eukaryote host is extremely versatile for cloning bacterial genomes. The Scitri genome, with a size of 1.84 Mbp, is with the *Haemophilus influenzae* genome (51) the largest bacterial genome cloned into yeast. Although difficulties have been reported with genomes expressing toxic products for the yeast cell, such as nucleases (11), the genome cloning reported here was straightforward, which was probably facilitated by the alternative genetic code of this group of organisms (Figure 1). With the availability of these cloned genomes in yeast, we compared the efficiency of GT between donor genomes directly extracted from bacteria (from-bacteria GT) and donor genomes extracted from yeast (from-yeast GT) (Figure 3).

Using from-bacteria GT, we first evaluated the optimal concentration of donor gDNA necessary to obtain the highest transplantation rate. The development of a qPCR assay allowed us to determine precisely the amount of genomic DNA added to the recipient cells during the GT experiment. Using the same number of McapΔRE recipient cells (∼5 × 10⁹), we found that optimal yields of transplantation were obtained using about 2/μg of donor gDNA; the only exception being Scitri, for which no transplant could be obtained whatever the gDNA quantity used (Figure 4 and Table 1). Beyond this optimal 2 μg-amount of donor gDNA, the number of transplants dropped dramatically. This result is reminiscent of standard bacterial transformation as it is known that above a given concentration of DNA resulting in the highest number of transformants, there is a drop in the transformation efficiency. This decrease has been attributed potentially to impurities in the reaction or to the saturation of DNA transporters at cell surface in other cases (52–54). Whole GT is an artificial process, limited to the laboratory, that occur concomitantly to PEG-mediated cell fusion events (8,51,55). Besides the possible requirement of cell syncytia for the reaction, it is yet unknown if specific DNA channels and/or transmembrane proton motive forces must be mobilized for DNA translocation.
Using an optimal quantity of gDNA, transplanta-
tion efficiencies were compared in-between donor species. For all genomes that were successfully transplanted into McapΔRE, a negative correlation between the transplanta-
tion efficiency and the phylogenetic distance was ob-
served. Highest transplantation efficiencies were obtained with genomes closest to McapΔRE and proportionally di-
minished with the increase in phylogenetic distance (Mcap > Mflorum > Mmm > Mputr > Mm). Very similar re-
sults were obtained for from-yeast GT with the Mm, Mflorum, Mm, Mputr and Mflorum donor genomes although a higher number of recipient cells was required for the reaction (1.5 × 10^10 cells/transplantation). However, for Mmm and Sc-
iti, no colonies were obtained in the from-yeast GT exper-
iments despite numerous attempts. Regardless, the success
obtained with the other four genomes implies that it is now possible to perform genome engineering for these species of interest and that GT can be obtained between bacteria from different genera. In this system, we established that the transplantation limit was located between Mflorum and spiroplasmas, since the Sciti genome was the only genome that could not be transplanted into McapΔRE. Although phylogenetic distance is probably the main factor in the es-
stablishment of this barrier, it is also possible that other fac-
tors such as genome size and growth rates of the donor cells might contribute to this phenomenon. Sciti, Mcap-most
distant phylogenetic species, has the largest genome size
(1.84 Mbp) of the tested donor species and has the slowest
growth rate under selective pressure (doubling time > 4 h),
it is currently difficult to draw any clear conclusions on the relative importance of these factors. However, we attempted
to transplant the genome of another Spiroplasma species, Spiroplasma floricola (56), that has a genome size close to
Mm (~1.3 versus ~1.1 Mbp respectively) and is consid-
ered as a fast grower among Mollicutes (doubling time < 1 h).
Despite these properties, no transplants were obtained using
S. floricola as the donor genome (data not shown). This result reinforces the idea that the main constraint in compatibil-
ity between the donor genome and the recipient cell is indeed the phylogenetic distance. The current fron-
tier stands between Mflorum (92.47% identity on the core
proteome) and spiroplasmas species (~86% identity). The identification of Mflorum as the McapΔRE most-distant
compatible GT organism remains extremely interesting and
should greatly facilitate our search for genetic factors re-
sponsible for genome compatibility between the recipient
cell and potential donor genomes. Investigating proteins in-
olved in cellular processes supposed to play a key role dur-
ing genome boot-up such as replication, transcription and
translation but differently conserved between compatible
and non-compatible genomes, should allow us to identify
potential barriers for GT technology. To our knowledge,
this is also the first time that a GT was successfully achieved
between bacteria belonging to different genera.

Among the factors that probably play a role in the success
of GT experiments, the compatibility between the origin of
replication (oriC) from the donor genome and the cellular
machinery in the recipient cell was tested using plasmids
carrying the chromosomal oriC of genome donor species. Al-
though McapΔRE was transformed by all oriC plasmids
tested, differences in transformation efficiencies were ob-
served and were not directly related to the phylogenetic dis-
tance between donor species and the recipient cell. In par-
sicular, no direct correlation was observed between trans-
formation rates obtained with pMCO3 (Mcap), pMYSO1
(Mmm) and pSD4 (Sciti) and transplantation rates of their
concerning whole genomes. These results suggest that,
even though the initiation of replication of the transplanted
genome might be triggered by the recipient cell machinery,
others cellular functions are likely to be required for a suc-
cessful GT. We hypothesize that the compatibility between
the transcription and translation apparatus of the recipient
cell and incoming genome is essential. This is based on the
requirements for the host machinery to be able to read the
genetic information carried by the input genome, to direct
protein synthesis and possibly replicate the donor genome,
all likely critical processes in the initial boot-up of the trans-
planted genome. Once key genes from the donor genome are
properly expressed by the machinery from the recipient cell,
the expression of other genes should normally follow using
the newly synthesized transcription-translation system.

All previously described factors, related to the phylo-
genetic distance between the donor cell and the recipient
cell, cannot explain the low yield of transplantation ob-
tained when Mmm was used as the donor genome (Fig-
ure 5 and Table 1). Indeed, Mmm is much more related to
Mcap (~99% core proteome similarity) than Mputr
and Mflorum but GT transplantation rates remained very low
or null in comparison to the other species. These data sug-
 gest that some species-specific genetic factors, present or
absent on Mmm genome could prevent GT. Genome com-
parison between Mmm and all other transplanted genomes
highlighted the presence of few genes of unknown func-
tions (hypothetical membrane proteins) but also some in-
volved in capsular biosynthesis pathway. Indeed, Mmm
has the capacity to produce capsular galactan (57,58) but also
exopolysaccharides as secreted or attached molecules (59)
that may contaminate the purified gDNA and present some
toxicity for the recipient cells. However, a large majority
of sequences were associated with insertion sequences (IS).
Overall, ~13% of the Mmm genome is composed of IS ele-
ments belonging to different families (26). We hypothesize
that the transposition of these IS may be triggered in the
recipient cylinder (Mcap) that lacks these elements and cellu-
lar mechanisms involved in their regulation (60). The stress
caused by the GT process might also be responsible for
IS movement since stress is known to induce transposi-
tion events (61). Such transpositions may result in lethal
events for the recipient cell and therefore explain the drop in
the transplanted efficiency observed for Mmm in our from-
bacteria GT or our incapacity to transplant its genome
in from-yeast GT. Inversely, no sequences common to all
transplanted genomes but absent on Mmm were identified
in silico.

This work provides a useful platform for the study of sev-
eral mycoplasmas, ideal organisms toward the creation
of minimal cells and essential biological units for our compre-
hension of life. As more and more bacterial genomes are
now cloned in yeast and can be engineered using powerful
genetic tools, one of the main bottlenecks of synthetic bi-
ology strategies remains our capacity to extent GT technics
to a growing number of species. Our work shows that the

Downloaded from https://academic.oup.com/nar/article-abstract/44/17/8501/2468051 by guest
on 30 July 2018
transplantation of a bacterial genome into a recipient cell, that was until now limited to two very closely related ruminant mycoplasmas, can be extended to more distant species including species from a different genus. This achievement paves the way for future studies devoted to better understand and control the cellular factors that define the compatibility between a donor genome and a recipient cell. In addition, several genomes for which in-yeast cloning and back transplantation is now possible originate from mycoplasmas species pathogenic for ruminants, and our work could, thus, lead to significant breakthroughs in the biology of these pathogenic micro-organisms as well as in the development of new and innovative treatments.

**ACCESSION NUMBER**

GenBank KX011460.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

We thank Enora Bodin for technical assistance on genotype confirmation of bacterial transplants. We thank Prof Alain Blanchard for helpful discussions and for reviewing the manuscript. We also thank Dr Florence Tardy (Anses, Laboratoire de Lyon, UMR Mycoplasmoses des Ruminants, Lyon, France) for providing the wtMputr strain 156. Part of the experiments was performed at the Genomic and Transcriptomic Facility of Bordeaux. This work was also made possible with the support of M. Arieis and Prof E. Gnant. Authors contributions: F.L., A.L., V.B., G.G. and C.L. performed the experiments. D.M., V.B and S.R. performed Mollicutes genomes sequencing and subsequent analyses. F.L., S.V., P.S.P., S.R. and C.L. wrote the paper.

**FUNDING**

French National Funding Research Agency [ANR-13-JSV5-0004-01 ‘SYNBIOMOLL’]; National Science Foundation [IOS-1110151]; Conseil Régional d’Aquitaine [20030304002FA, 2004035003FA]; European Union, FEDER [2003227]; Investissements d’Avenir, Convention attributive d’aide [ANR-10-EQPX-16-01]. Funding for open access charge: French National Funding Research Agency [ANR-13-JSV5-0004-01 ‘SYNBIOMOLL’]; National Science Foundation [IOS-1110151].

**REFERENCES**

1. Mushegian, A.R. and Koonin, E. V. (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U.S.A.*, 93, 10268–10273.
2. Gil,R., Sabater-Muñoz,B., Latorre,A., Silva,F.J. and Moya,A. (2002) Extreme genome reduction in *Buchnera* spp.: toward the minimal genome needed for symbiotic life. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 4454–4458.
3. Hutchison, C.A., Peterson, S.N., Gill,S.R., Cline, R.T., White,O., Fraser,C.M., Smith,H.O. and Venter,J.C. (1999) Global transposon mutagenesis and a minimal Mycoplasma genome. *Science*, 286, 2165–2169.
4. May,M., Balish,M.F. and Blanchard,A. (2014) *Mycoplasmas*: In: Rosenberg,E, DeLong,EF., Lory,S., Stackebrandt,E and Thompson,F (eds) *The Prokaryotes*. Springer-Verlag, Berlin Heidelberg, pp. 515–550.
5. Glass,J.J., Assad-Garcia,N., Alperovich,N., Yooseph,S., Lewis,M.R., Maruf,M., Hutchison,C.A., Smith,H.O. and Venter,J.C. (2006) Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. U.S.A.*, 103, 425–430.
6. Grosjean,H., Breton,M., Sirand-Pugnet,P., Tardy,F., Thiaucourt,F., Citti,C., Barré,A., Yoshizawa,S., Fourmy,D., de Crécy-Lagard,V. et al. (2014) Predicting the minimal translation apparatus: lessons from the reductive evolution of mollicutes. *PLoS Genet.*, 10, e1004363.
7. Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deeink, T.J., Ellisman, M.H., Gill,J., Kannan,K., Karas,B.J., Ma,L., et al. (2016) Design and synthesis of a minimal bacterial genome. *Science*, 351, aad6253.
8. Lartigue, C., Vashee, S., Algire,M.A., Chuang, R.-Y., Benders,G.A., Ma,L., Noskov,V.N., Denisova,E.A., Gibson,D.G., Assad-Garcia, N. et al. (2009) Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science*, 325, 1063–1068.
9. Benders, G.A., Noskov,V.N., Denisova,E.A., Lartigue,C., Gibson,D.G., Assad-Garcia,N., Chuang,R.-Y., Carrera,W., Moodie,M., Algire,M.A. et al. (2010) Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.*, 38, 2558–2569.
10. Citti,C., Maréchal-Drouard,L., Saillard,C., Weil,J.H. and Bové,J.M. (1992) Spiroplasma citri UGG and UGA tryptophan codons: sequence of the two tryptophanyl-tRNAs and organization of the corresponding genes. *J. Bacteriol.*, 174, 6471–6478.
11. Karas,B.J., Tagwerker,C., Yonemoto,I.T., Hutchison,C.A. and Smith,H.O. (2012) Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth. Biol.*, 1, 22–28.
12. Tagwerker,C., Dupont,C.L., Karas,B.J., Ma,L., Chuang,R.-Y., Benders,G.A., Ramon,A., Novotny,M., Montague,M.G., Vennepally, P. et al. (2012) Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.*, 40, 10375–10383.
13. Noskov,V.N., Karas,B.J., Young,L., Chuang,R.-Y., Gibson,D.G., Lin,Y.-C., Stam,J., Yonemoto,I.T., Suzuki,Y., Andrews-Pfannkoehc, C. et al. (2012) Assembly of large, high G+C bacterial DNA fragments in yeast. *ACS Synth. Biol.*, 1, 267–273.
14. Noskov,V.N., Segall-Shapiro,T.H. and Chuang,R.-Y. (2010) Tandem repeat coupled with endonuclease cleavage (TREC): a seamless modification tool for genome engineering in yeast. *Nucleic Acids Res.*, 38, 2570–2576.
15. Chandran,S., Noskov,V.N., Segall-Shapiro,T.H., Ma,L., Whitei,C., Laritgure,C., Jores,J., Vashsee,S. and Chuang,R.-Y. (2014) TREC-IN: gene knock-in genetic tool for genomes cloned in yeast. *BMC Genomics*, 15, 1180–1190.
16. Schiek,E., Laritgure,C., Frey,J., Voza,N., Hegermann,J., Miller,R.A., Valguarnera,E., Muriuki,C., Meens,J., Nene,V. et al. (2016) Galactofuranose in *Mycoplasma mycoides* is important for membrane integrity and conceals adhesins but does not contribute to serum resistance. *Mol. Microbiol.*, 99, 55–70.
17. Laritgue,C., Lebaudy,A., Blanchard,A., El Yacoubi,B., Rose,S., Grosjean,H. and Doughwaite,S. (2014) The flavoprotein Mcap0476 (RlmFO)catalyzes m5U1939 modification in *Mycoplasma pneumoniae*. *J. Bacteriol.*, 196, 3515–3525.
18. 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.
19. Tsarpopoulos,L., Gouguers,G., Blanchard,A., Vashsee,S., Jores,J., Laritgue,C. and Sirand-Pugnet,P. (2016) In-yeast engineering of a bacterial genome using CRISPR/Cas9. *ACS Synth. Biol.*, 5, 104–109.
20. Laritgue,C., Glass,J.I., Alperovich,N., Pieper,R., Parmar,P.P., Hutchison,C.A., Smith,H.O. and Venter,J.C. (2007) Genome transplantation in bacteria: changing one species to another. *Science*, 317, 632–638.
