

A novel sigma factor reveals a unique regulon controlling cell-specific recombination in *Mycoplasma genitalium*

Sergi Torres-Puig, Alicia Broto, Enrique Querol, Jaume Piñol and Oscar Q. Pich*

Institut de Biotecnologia i Biomedicina and Departament de Bioquímica i Biologia Molecular. Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Received July 10, 2014; Revised March 26, 2015; Accepted April 19, 2015

ABSTRACT

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

INTRODUCTION

Mycoplasma genitalium is a sexually transmitted pathogen implicated in urogenital diseases such as urethritis, cervicitis, pelvic inflammatory disease and infertility (1). In addition to the significance of this microorganism as an emerging human pathogen, research on *M. genitalium* is inti-

mately associated with important scientific milestones such as the construction of the first synthetic chromosome (2,3) or the development of a whole-cell computational model (4). Certainly, the massive genome reduction undergone by *M. genitalium* favors its suitability as a systems biology model. Compared with more complex bacteria, *M. genitalium* lacks the majority of known transcription factors and regulatory pathways. However, despite the apparently reduced gene regulatory toolbox, both environmental stresses and metabolic insults induce complex, specific transcriptional responses in this bacterium (5,6). Therefore, the elements and mechanisms regulating gene expression in *M. genitalium* are essentially unexplored.

Sigma factors are fundamental components of the prokaryotic transcriptional machinery that direct the core RNA polymerase to specific promoter elements. Members of the sigma-70 family recognize promoters with two conserved motifs centered roughly 10 and 35 base pairs upstream of the transcription initiation site (7). A representative member of this family is the primary sigma factor, RpoD, which recognizes the consensus sequence TTGACA-N_{15/19}-TATAAT and directs transcription of most genes in growing cells. However, most bacteria have alternative sigma factors, which control the expression of different subsets of genes usually in response to specific environmental conditions. *Mycoplasma genitalium* and its close relative *Mycoplasma pneumoniae* possess a primary sigma-70 factor (MG249 and MPN352, respectively) but no alternative sigma factors have been characterized. Transcriptional analyses have shown that the archetypal sigma-70 promoter of *M. pneumoniae* features a conserved –10 element but it is devoid of a well-conserved –35 region (8–11). Of note, transcription initiation has also been described in locations lacking an apparent Pribnow box, suggesting the existence of both non-canonical promoter elements and alternative sigma factors controlling transcription in mycoplasmas (11,12).

Sequence similarity and secondary structure prediction analyses suggest that the *M. pneumoniae* MPN626 protein is a putative sigma factor (13). Indeed, MPN626 shows dis-

*To whom correspondence should be addressed. Tel: +34 935 868 953; Fax: +34 935 812 011; Email: oscar.quijada@uab.cat

tant similarity to sigma-70 factors of group IV, which accommodates the members of the extracytoplasmic function (ECF) subfamily of sigma-70 factors (14). Typically, ECF sigma factors are members of cell-surface signaling systems that respond to signals arising from the extracytoplasmic environment. These signals are largely unknown, which limits the understanding of the physiological role of these transcriptional regulators. Of the four conserved regions characteristic of sigma-70 factors (15), only the region implicated in the recognition of the -35 promoter element (region 4) seems to be conserved in the MPN626 protein. Unfortunately, both proteomic and interactomic analyses of this putative sigma factor have been unsuccessful, possibly as a result of its implicit low cellular abundance (16). Therefore, the nature of the MPN626 regulator and its associated regulatory DNA elements remain unknown.

Herein, we investigated the regulatory function of the MG428 protein, which is the *M. genitalium* orthologue of MPN626. Our results demonstrate that MG428 is an alternative sigma factor that activates recombination in this human pathogen. Since recombination plays a pivotal role in the generation of antigenic variation (17–19), MG428 emerges as a novel factor contributing to *M. genitalium* virulence. Orthologues of the MG428 protein can be identified in other species such as *Mycoplasma galisepticum* str. R(low) (MGA_0765), *Mycoplasma capricolum* subs. *capricolum* (MCAP_0855), *Mycoplasma suis* str. Illinois (MSU_0577) or *Acholeplasma* sp. CAG:878 (WP_021921817), suggesting that this protein and its associated regulatory functions might be widespread in the bacterial class Mollicutes.

MATERIALS AND METHODS

Culture conditions and strains

Mycoplasma genitalium was grown in SP-4 medium at 37°C and a 5% CO₂ atmosphere in tissue culture flasks. For the preparation of SP-4 plates, the medium was supplemented with 0.8% agar (Difco BRL). Where indicated, tetracycline (3 µg ml⁻¹), chloramphenicol (17 µg ml⁻¹) or puromycin (3 µg ml⁻¹) was added. Tetracycline and puromycin were protected from light. For the detection of β-galactosidase activity, SP-4 plates were supplemented with X-Gal (150 µg ml⁻¹). All *M. genitalium* strains used in this work are listed in Supplementary Table S1. *Escherichia coli* strain XL-1 Blue was used for cloning and plasmid amplification. This strain was grown in Luria Bertani (LB) or LB agar plates containing 100 µg ml⁻¹ ampicillin, 40 µg ml⁻¹ X-Gal and 24 µg ml⁻¹ isopropyl β-D-1-thiogalactopyranoside (IPTG) when needed. *E. coli* strain BL21 (DE3) was grown in LB and used to express the MG428 recombinant protein.

Plasmids, primers and genetic nomenclature

All plasmids constructed in this study are described in the Supplementary Material. All primers used in this study are listed in Supplementary Table S2. In this work, genes are designated according to their locus tags in the reference strains *M. genitalium* G37 or *M. pneumoniae* M129. However, during the submission of this manuscript, a re-annotation of the *M. genitalium* G37 genome was released

(NC.000908_2). For this reason, a table correlating the old and new locus tags is provided as Supplementary Material (Supplementary Table S3).

DNA manipulation

Plasmid DNA for general DNA manipulations was obtained using Fast Plasmid Mini Kit (5Prime). PCR products were purified from agarose gels using the E.Z.N.A. Gel Extraction Kit (Omega Biotek) and digested with the corresponding restriction enzymes (Fermentas) when necessary. Plasmids for electroporation experiments were obtained using the GenElute HP Midiprep Kit (Sigma) or GeneJet Plasmid Maxiprep Kit (Fermentas). Genomic DNA was isolated from mid-log phase cultures of *M. genitalium* grown in 75 cm² flasks as previously described (20).

Extraction of total RNA

Total RNA was extracted from mid-log phase cultures of *M. genitalium* using the RNAqueous Kit (Life Technologies). Total RNA was treated with Turbo DNase (Life Technologies) following the manufacturer's instructions.

Primer extension

Primer extension analyses were performed with 20 µg of total RNA as previously described (21). Fragments were analyzed using PeakScanner v1.0 software (Applied Biosystems). At least two independent primer extension experiments were performed with each primer.

Sequencing reactions

Sequencing reactions were performed with the BigDye® v3.1 Cycle Sequencing kit using 2.5 µl of genomic DNA, following the manufacturer's instructions. All sequencing reactions were analyzed using an ABI PRISM 3130xl Genetic Analyser at the Servei de Genòmica i Bioinformàtica (UAB).

Full-genome sequencing and data analysis

Genomic DNA was prepared with the Nextera XT DNA Sample Preparation Kit (Illumina) and analyzed using a MiSeq Desktop Sequencer (Illumina) at Servei de Genòmica i Bioinformàtica (UAB). Illumina reads were 251 nucleotides long. Bowtie 2 (22) and BWA-MEM (23) were used to align sequences, which were then piled up with SAMtools (24). Finally, variants were called using VarScan (25). Variants were only selected when the corresponding nucleotide had a phred quality score >20 (26,27).

Transformation of *M. genitalium* and screening for mutants

Transformation of *M. genitalium* was achieved by electroporation as previously described (20). For the generation of double-recombinant and transposon mutants, 30 and 10 µg of DNA were used, respectively. Transposon mutants were screened for by sequencing with the CmUp and CmDown primers. Mutants obtained by allelic exchange were screened for by PCR and sequencing of the resulting amplicons.

Production of MG428 antiserum

The *E. coli* strain carrying the pET21aMG428 plasmid was grown in 2 L of LB medium at 37°C to an OD₆₀₀ of ~0.6 and MG428 expression was induced with 1 mM IPTG. Then, cells were harvested by centrifugation and sonicated. Purification was achieved under denaturing conditions using 6M urea on a HisTrap Fast Flow affinity purification column (GE Healthcare Life Sciences). Eluate was dialyzed three times in 1XPBS. Polyclonal antibodies to the MG428 protein were obtained by repeated immunization of BALB/c mice with 0.5 mg ml⁻¹ purified recombinant MG428.

SDS-PAGE and western blotting

Separation of total mycoplasma proteins by SDS-PAGE followed standard procedures. Total protein concentration was determined with the PierceTM BCA Protein Assay Kit (Thermo Scientific). Proteins were electrophoretically transferred to Polyvinylidene fluoride (PVDF) membranes and probed with polyclonal antiserum against MG428 (1:100) or mCherry from BioVision (1:1000). Monoclonal anti-HsdS (Host specificity determinant S subunit, MG438) antibody (1:2000) was used as a protein loading control (28).

Analysis of MG428–RNA polymerase interaction

Two hundred nanograms of soluble rMG428 protein were mixed with 1, 2, 4 and 8 µg of *E. coli* RNA polymerase (RNAP) core- or holoenzymes (Epicentre) in a buffer containing 25 mM HEPES pH 8, 10 mM MgCl₂, 100 mM potassium glutamate and 2 mM Dithiothreitol (DTT). After incubation at 37°C for 30 min, 5 µl of 87% glycerol was added to the samples prior to loading on a Mini-PROTEAN TGX precast gel (Biorad). Gels were run at 80 V for 90 min in 25 mM Tris, 192 mM glycine pH 8.3 and stained with colloidal Coomassie (29). For the production of soluble rMG428 protein, the *E. coli* strain carrying the pET21aMG428 plasmid was grown in 2 L of LB medium at 37°C to an OD₆₀₀ of ~0.6 and MG428 expression induced with 0.1 mM IPTG. After 2 h of incubation, cells were harvested by centrifugation and sonicated in the presence of lysozyme (Sigma) and DNase I (Sigma). Purification was achieved on a HisTrap Fast Flow affinity purification column (GE Healthcare Life Sciences). Eluate was dialyzed three times in 25 mM HEPES pH 8.

Liquid chromatography–mass spectrometry (LC–MS) analysis

Samples were analyzed on a Maxis Impact Q-TOF spectrometer (Bruker, Bremen), coupled to a nano-HPLC system (Proxeon, Denmark). Further information on the manipulation of the sample for LC analysis can be found in Supplementary Table S6. Proteins were identified using Mascot (Matrix Science, London, UK) by search on a database constructed with *E. coli* K-12 substrain MG1655 and *M. genitalium* G37 sequences in the SwissProt Database (4738 sequences). MS/MS spectra were searched with a precursor mass tolerance of 10 ppm, fragment tolerance of 0.05 Da, trypsin specificity with a maximum of two missed

cleavages, cysteine carbamidomethylation set as fixed modification and methionine oxidation as variable modification. Significance threshold for the identifications was set to give a False Discovery Rate <1% at the peptide level measured by searching a Decoy database.

qRT-PCR analysis

Reverse transcription was performed with iScript reverse transcriptase (Bio-Rad) and random primers. Primers used for qPCR (Supplementary Table S2) were devised using Primer3 software. Calculated amplification efficiencies are shown in Supplementary Table S4. qPCR was performed with iTaq polymerase (Bio-Rad) and SYBR green in a CFX384 PCR instrument (Bio-Rad). Amplification was obtained through a 3-step PCR protocol including denaturation at 95°C for 20 s, annealing at 56°C for 20 s and extension at 72°C for 20 s. Detection was performed at the end of the extension step and melting curve analyses were performed at the end of each run. Expression of MG_177, MG_418 and MG_430, that code for the alpha subunit of a DNA-directed RNA polymerase, the ribosomal protein L13 and the 2,3-bis-phosphoglycerate-independent phosphoglycerate mutase, respectively, were used as housekeeping genes for normalization of qRT-PCR data. Relative gene expression was calculated using the Pfaffl method (30), which considers the amplification efficiencies of the target and the reference genes, and accurate normalization was achieved by geometric averaging of multiple reference genes (31). Differential gene expression was judged based on the common arbitrary 2-fold cutoff. Data presented in the manuscript correspond to the analysis of RNAs isolated from three independent biological replicates for each strain under study.

Quantitative assessment of the recombination capacity

Mycoplasma genitalium cultures were grown and prepared for electroporation as previously described (20). Mycoplasma cell suspensions were adjusted to ~10⁹ cells ml⁻¹, which was monitored by diluting the resuspended cells to an OD₆₀₀ ~2. Next, 100 µl cell aliquots were mixed with 30 µg of the suicide plasmid pΔMG_218-lacZ and electroporated. After the addition of 1 ml of SP4, electroporated cells were incubated at 37°C for 4 h, and 200 µl aliquots seeded onto SP4 plates containing puromycin and X-Gal. After 21 days of incubation at 37°C, plates were screened for the presence of puromycin resistant, blue colonies. A double recombination event between homologous sequences from the pΔMG_218-lacZ plasmid and the chromosome of *M. genitalium* promotes the replacement of the MG_218 gene by the puromycin resistance marker. The MG_218 gene, which codes for the cytoadherence-associated protein HMW2, was shown to be dispensable for growth under laboratory culture conditions in previous studies (32,33). Deletion of the MG_218 gene was chosen because previous work from our laboratory showed that transformants were obtained at a high frequency (32). The lacZ gene present in the pΔMG_218-lacZ plasmid confers β-galactosidase activity and facilitates the detection of the transformant colonies. The transformation efficiency (number of transformants

per viable cell) was calculated as an indicator of the recombination capacity of selected *M. genitalium* strains. Results presented in the manuscript correspond to at least three independent biological repeats.

Phase contrast and fluorescence microscopy

Mycoplasma genitalium was grown in filtered (0.22 μm) SP-4 medium on IBIDI chamber slides for 16 h, washed once with 1x phosphate buffered saline (PBS) and visualized on a Nikon Eclipse TE 2000-E microscope at 37°C and 5% CO₂. All strains were grown and visualized under homogeneous conditions. Phase contrast and TRITC epifluorescence images were captured with a Digital Sight DS-SMC Nikon camera controlled by NIS-Elements BR software. Images were analyzed with Image J software and GDSC plug-in.

RESULTS

Construction of a ΔMG_428 mutant and its complemented strain

To investigate the role of the MG428 protein, we constructed a MG₄₂₈ null mutant by allelic exchange (Supplementary Figure S1). For control purposes, the MG₄₂₈ wild-type (WT) allele was subsequently reintroduced into the ΔMG_428 mutant by transposon delivery to create its respective complemented strain. To this end, we used the Tn*CatMG₄₂₈* minitransposon (MiniTnp), which carries the MG₄₂₈ gene under the control of its own promoter. Primer extension analyses revealed that this promoter is located upstream of the MG₄₂₇ gene (Supplementary Figure S2). Consistently, transformation efficiencies obtained in electroporation experiments of the ΔMG_428 mutant with the Tn*CatMG₄₂₈* MiniTnp were markedly low ($\sim 10^{-7}$ transformants per viable cell). On average, we only obtained seven chloramphenicol-resistant colonies per electroporation experiment. In contrast, electroporation of the ΔMG_428 mutant with a Tn*Cat* MiniTnp without the MG₄₂₈ gene yielded ~ 10 thousand chloramphenicol-resistant colonies ($\sim 10^{-4}$ transformants per viable cell). The dramatic reduction in the number of viable transformants observed after electroporation with the Tn*CatMG₄₂₈* MiniTnp suggested that the expression of the MG₄₂₈ gene supplied *in trans* was deleterious to *M. genitalium*.

Twenty-five chloramphenicol-resistant colonies from four different electroporation experiments were picked, propagated and the MiniTnp insertion point determined by sequencing with primers specific to the chloramphenicol acetyltransferase (*cat*) gene. We found the MiniTnp inserted within specific loci rather than randomly inserted throughout the chromosome (Figure 1A and Supplementary Table S5). Of note, all the MiniTnp insertions were found in the antisense orientation with respect to the disrupted gene, indicating that this orientation was also selected. Furthermore, we determined that the two isolated mutants with a MiniTnp insertion within the MG₃₉₀ gene carried a truncated copy of the reintroduced MG₄₂₈ gene (Figure 1B) reinforcing the hypothesis that expression of the MG₄₂₈ gene supplied *in trans* was toxic. A possible mechanism to generate MG₄₂₈ truncated copies is described in Supplementary Figure S3.

Analysis of MG428 expression by Western blot

Expression of the MG428 protein was analyzed by Western blot using anti-MG428 polyclonal antibodies (Figure 2A). The MG428 protein was not detected in lysates of the WT strain, indicating that it was expressed at very low levels in *M. genitalium*. Likewise, MG428 expression was not detected in lysates of the ΔMG_428 mutant. In contrast, a band of the predicted molecular mass of the MG428 full-length protein (17 kDa), was clearly detected in the complemented mutants. This result indicates that transcriptional fusion of the MG₄₂₈ gene to its own promoter, located upstream of the MG₄₂₇ gene, leads to increased levels of MG428 expression as compared to the WT strain. In light of this finding, we analyzed the intergenic region between the MG₄₂₇ and MG₄₂₈ genes and we identified a possible Rho-independent terminator (Supplementary Figure S4). The presence of this putative terminator, which is supported by several transcription terminator prediction softwares (34–36), could explain the reduced levels of MG428 expression observed in the WT strain. On the other hand, MG428 levels differed considerably among the complemented strains, indicating that the genetic context of the transposon insertion had a significant impact on the expression of the MG₄₂₈ ectopic copy. Finally, in agreement with the presence of a truncated copy of the MG₄₂₈ gene, a 9 kDa band was detected in the Tn::MG₃₉₀₋₁ mutant.

Analysis of gene expression by qRT-PCR

The existence of transcriptional changes between the WT strain and the different MG₄₂₈ mutants described above was analyzed by qRT-PCR (Figure 3). As expected, no MG₄₂₈ transcript could be detected in the ΔMG_428 mutant. However, in agreement with the western blot analysis, we found increased levels of MG₄₂₈ transcript (~ 20 -fold) in the complemented mutants as compared to the WT strain. No transcriptional changes of the MG₄₂₇ gene were observed in any of the strains tested, suggesting that MG428 overexpression has no effect on MG₄₂₇ transcript levels.

Next, we analyzed transcription of all genes specifically targeted by the Tn*CatMG₄₂₈* MiniTnp. *recA* messenger levels in the ΔMG_428 and the Tn::MG₃₉₀₋₁ mutants were similar to those of the WT strain. Thus, no changes in *recA* transcription were detected in the absence of a full copy of the MG₄₂₈ gene. In contrast, *recA* was found to be dramatically up-regulated (17.7- and 11.4-fold, respectively) in the two complemented mutants carrying a full copy of the MG428 gene that were selected for further analysis. Similarly, we assessed transcription of *ruvA*, *ruvB* and an ORF of unknown function (MG_RS02200) that has been recently annotated immediately upstream of the *ruvAB* genes (see Supplementary Table S3). As previously observed for *recA*, messenger levels of the MG_RS02200 ORF and the *ruvAB* genes and in the ΔMG_428 and the Tn::MG₃₉₀₋₁ mutants were comparable to those of the WT strain. However, mRNA levels of these three genes were found to be markedly higher in the complemented mutants as compared to the WT strain. A similar pattern was observed when MG₂₂₀ transcription was examined. Of note, MG₂₂₀ up-regulation in the Tn::*recA-2* mutant was found to be

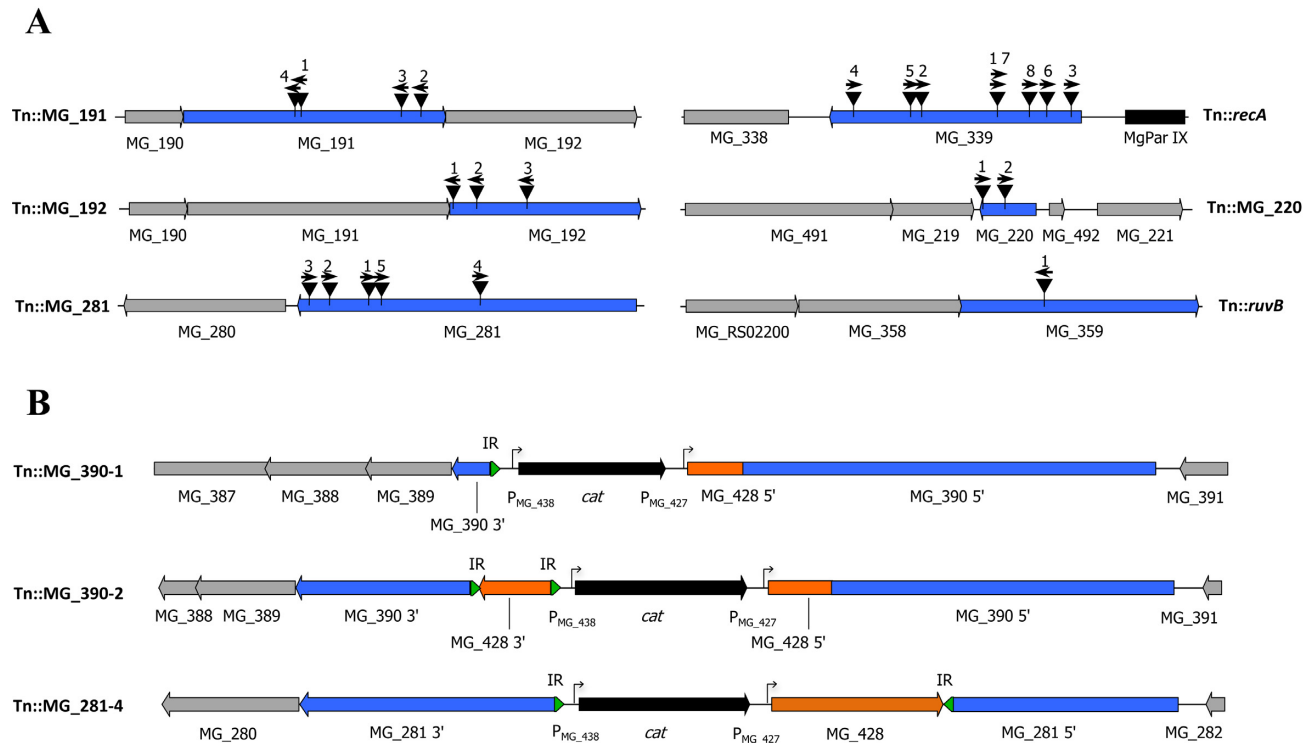


Figure 1. Insertion points of the Tn*CatMG_428* MiniTnp in the *M. genitalium* genome. (A) Schematic representation depicting the genes preferentially targeted by Tn*CatMG_428* in the genome of a Δ MG_428 mutant. Genes disrupted by Tn*CatMG_428* are shown in blue. Black filled triangles represent the transposon insertion points. Arrows above the triangles indicate the orientation of the transposon insertion. (B) Schematic representation showing the presence of a truncated MG_428 ectopic copy in the genome of the Tn::MG390-1 and -2 mutants. The MG_428 coding region is highlighted in orange and the inverted repeats (IR) of the Tn*CatMG_428* MiniTnp are shown in green. The Tn::MG_281-4 mutant, carrying a full copy of the MG_428 gene, is also shown for comparison.

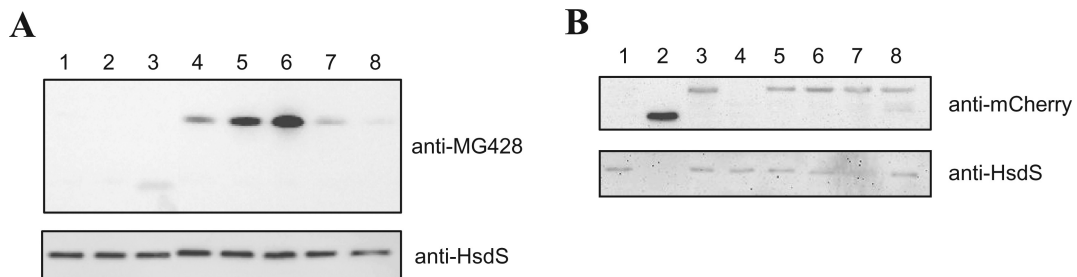


Figure 2. Analysis of protein expression by western blotting. (A) Immunoblot analysis of MG428 expression in the WT strain and several representative mutants. Lane 1, WT; lane 2, Δ MG_428; lane 3, Tn::MG_390-1; lane 4, Tn::MG_281-1; lane 5, Tn::*recA*-2; lane 6, Tn::MG_220-1; lane 7, Tn::MG_191-2 and lane 8, Tn::MG_192-1. (B) Immunoblot analysis of mCherry expression. Lane 1, WT; lane 2, Cat:Ch; lane 3, RecA:Ch; lane 4, MG_428:Ch; lane 5, Δ MG_428-RecA:Ch; lane 6, RecA:Ch-10; lane 7, RecA:Ch-22 and lane 8, RecA:Ch-35. Because the Cat-mCherry fusion is expressed at very high levels compared to the RecA-mCherry fusion, the amount of total protein loaded for the Cat:Ch mutant was reduced 20 times. HsdS protein was detected with a monoclonal antibody and used as a loading control.

extraordinarily pronounced (165-fold). MG_390 transcription was also activated to some extent (more than 3-fold) upon MG428 overexpression. Likewise, the MG_389 gene was also found to be up-regulated in the complemented mutants, suggesting that MG_390 and MG_389 are co-transcribed. We observed that MG_389 transcription was also elevated in the Tn::MG_390-1 mutant, which carried the Tn*CatMG_428* inserted within the MG_390 locus. The increased MG_389 transcript levels in this mutant are likely the result of the activity of cryptic promoters located within the MiniTnp. In contrast, transcription of the MG_281,

MG_191 and MG_192 genes was not found to be altered in the Δ MG_428 or the complemented mutants (data not shown), indicating that expression of these genes is not under the control of the MG428 protein.

MG428 interacts with the RNAP core enzyme

Interaction of the MG428 protein with the *E. coli* RNAP was analyzed by native gel electrophoresis. A fixed amount of soluble rMG428 protein was mixed with increasing concentrations of RNAP core enzyme, which is free of de-

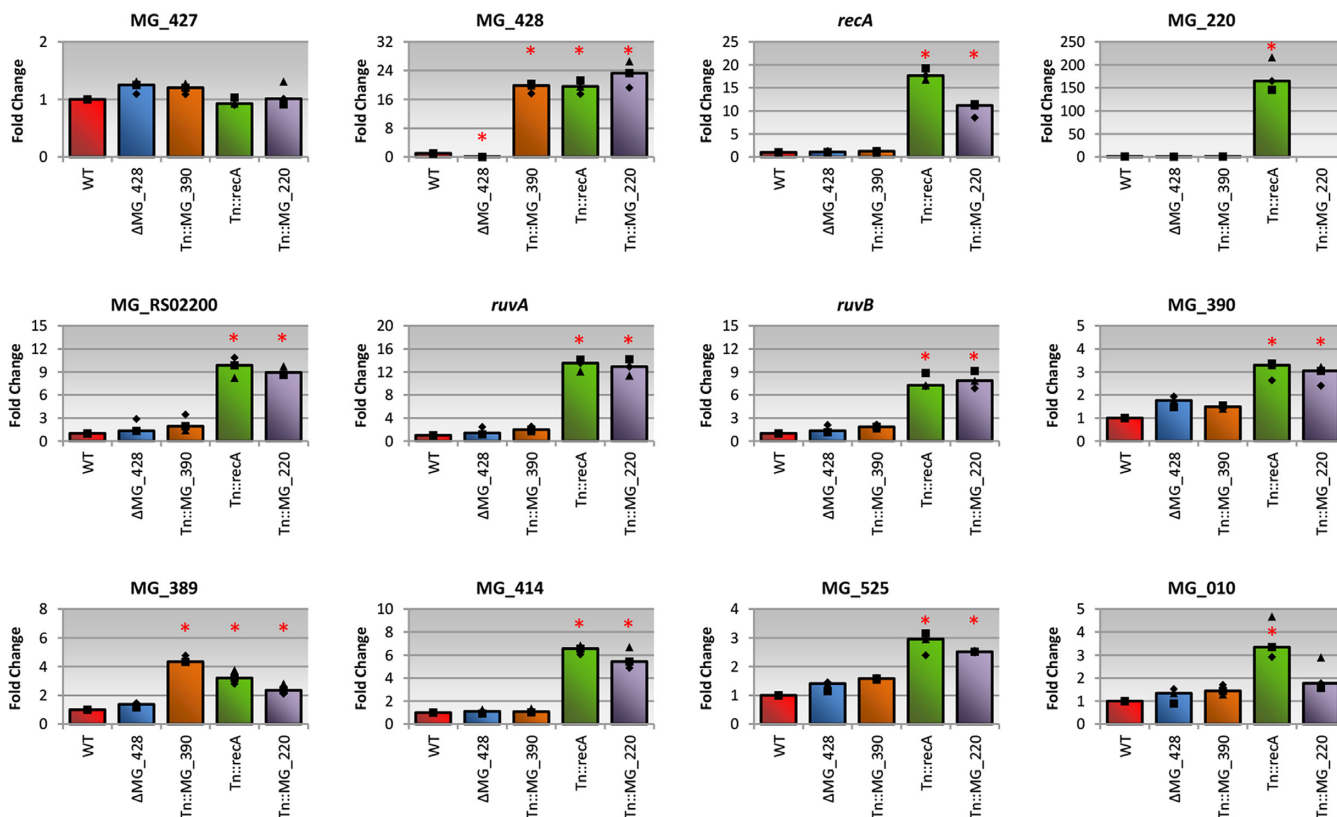


Figure 3. Analysis of gene expression by qRT-PCR. Transcriptional analysis of selected *M. genitalium* genes in the WT and several mutant strains. Three independent biological repeats were performed and the respective fold-changes in gene expression are indicated with diamonds, squares and triangles. Mean fold-changes for each target gene are represented by color bars. Statistical significance of mean fold-changes above the cutoff (>2) was assessed with Student's *t* test. Statistically significant values ($P < 0.05$) are indicated with a red asterisk. Transcription of MG_220 (309 bp) in the Tn::MG_220 mutant could not be assessed by qRT-PCR due to the presence of a Tn*Cat*MG_428 MiniTnp insertion in this gene.

tectable sigma factors (Figure 4A). At the lowest concentrations of RNAP used, the rMG428 protein and the core enzyme migrated as two discrete bands on the native gel. However, at the highest concentration of core enzyme used, the band corresponding to the unbound rMG428 protein was not detected, suggesting the formation of a rMG428-RNAP complex with an electrophoretic mobility similar to that of the core enzyme alone. Indeed, LC-MS/MS analysis of this band recovered from the native gel confirmed the presence of all the expected RNAP subunits plus the rMG428 protein (Supplementary Table S6). In contrast, this interaction was greatly reduced when the rMG428 protein was mixed with increasing concentrations of the RNAP holoenzyme, which consists of the RNAP core enzyme saturated with sigma-70 factor (Figure 4B). Supporting this statement, the band corresponding to the unbound rMG428 protein was clearly present in the native gel at all concentrations of holoenzyme used.

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

A putative promoter sequence with sigma-70 architecture was identified within the upstream region (UR) of the genes or operons up-regulated by the MG428 protein. This putative promoter was composed of two conserved elements of six residues separated by 18 or 19 nucleotides

and showed the consensus 5'-TTGTCA-N_{18/19}-ATTWAT-3' (Figure 5A). Of note, a conserved sequence with sigma-70 promoter architecture was also recognized immediately upstream of all of the *M. pneumoniae* genes homologous to the members of the *M. genitalium* MG428-regulon. The conserved sequence identified in *M. pneumoniae* showed the consensus 5'-TTGGCR-N_{18/19}-ATTYAT-3' (Figure 5B). We conducted primer extension analyses of the *recA*, *ruvA* and MG_220 genes in the WT strain and the Tn::*recA*-2 mutant (Figure 6). In repeated experiments, we did not detect transcription initiation within the URs of these genes using RNA from the WT strain. In contrast, single, unequivocal transcriptional start sites (TSS) were identified using RNA from the Tn::*recA*-2 mutant. As expected, all the identified TSSs were located immediately downstream of the anticipated promoter sequences (Figure 5A). For the *ruvA* gene, we found that transcription initiated in the promoter region of the MG_RS02200 ORF, which precedes the *ruvAB* genes. This specific TSS was confirmed with a second primer complementary of the MG_RS02200 ORF region (Supplementary Figure S5). Therefore, MG428-dependent activation of the *ruvAB* genes is driven from the MG_RS02200 promoter.

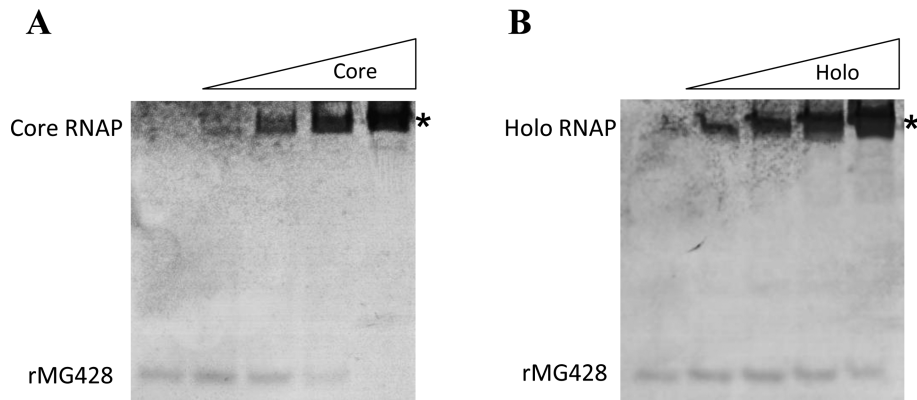


Figure 4. Analysis of MG428-RNA polymerase interaction. A fixed amount of soluble rMG428 protein (200 ng) was incubated with increased concentrations (0, 2.5, 5, 10 and 20 mM) of RNAP core enzyme (A) or RNAP holoenzyme (B). Mixtures were separated on native discontinuous 4–15% polyacrylamide gels and stained with colloidal Coomassie. Bands indicated with an asterisk (*) were cut off the gel and analyzed by LC–MS (see also Supplementary Table S6).

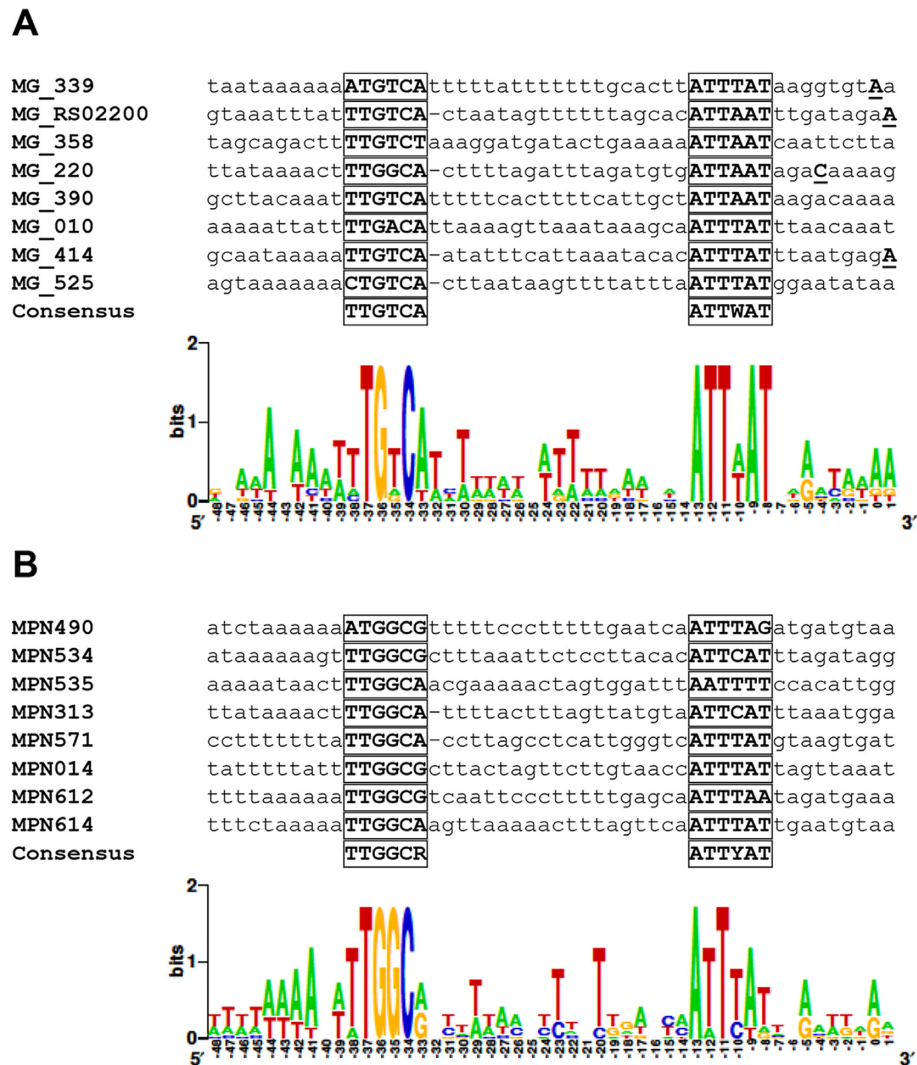


Figure 5. Identification of a conserved region with sigma-70 promoter architecture within the UR of the MG428-regulated genes. (A) Sequence logo generated with the UR of all the MG428-regulated genes or operons identified in this study. (B) Sequence logo generated with the UR of the respective *M. pneumoniae* homologs. The logo consists of stacks of letters, one stack for each position in the sequence analyzed. The overall height of each stack indicates the sequence conservation at that position (measured in bits). Nucleotides corresponding to the putative –35 and –10 promoter elements are boxed. Underlined and bolded nucleotides indicate experimentally determined TSSs (see also Figure 6). Due to the different length of the spacer region between the hexanucleotide promoter elements, one nucleotide gap was arbitrarily located immediately after the –35 box when necessary.

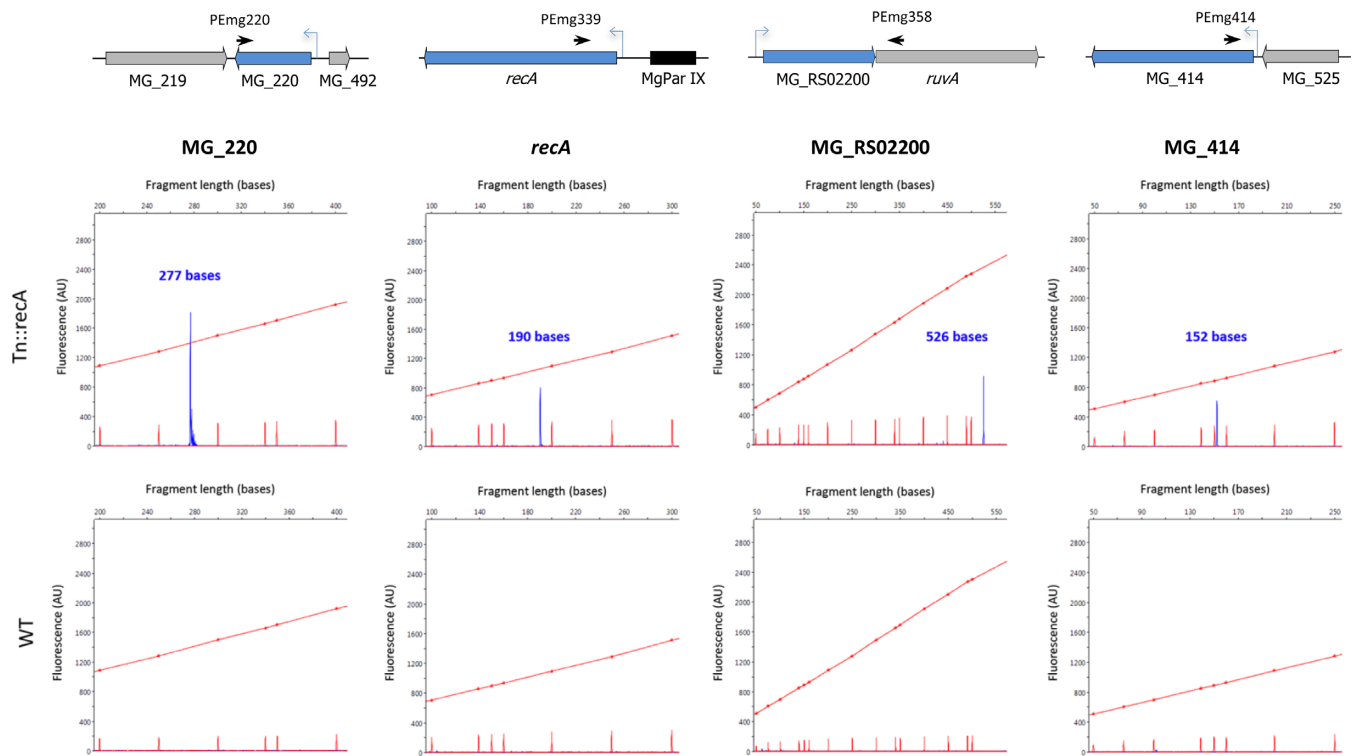


Figure 6. Identification of the TSSs of several MG428-regulated genes. Primer extension analysis of the MG_220, *recA*, MG_RS02200 and MG_414 genes in the WT strain and the Tn::*recA*-2 mutant. All electropherograms were generated with Peak Scanner v1.0 (Applied Biosystems) analysis software. Red peaks represent ROX size standards while blue peaks correspond to the primer extension products. Schematic representations of the genome regions analyzed are shown and the presence of the identified promoters indicated with blue arrows. The approximate location of the primers used in these analyses is also indicated by arrows.

Identification of additional MG428-regulated genes using the novel promoter sequence

A number of sequences similar to the divergent promoter sequence described above were present all over the genome of the *M. genitalium* G37 strain. In particular, sixty-nine instances of the conserved sequence were identified when no more than one mismatch was allowed (Supplementary Table S7). Since the relative importance of each residue in the conserved sequence is still unknown, we allowed the mismatch to occur at any position of the sequence. To enrich for *bona fide* promoter sequences, the previous list was filtered omitting those sequences that were located >300 nucleotides upstream from the translational start point of a known gene. As a result, we obtained a catalog of 19 selected candidate promoters (Table 1).

In addition to *recA*, *ruvA*, MG_220, MG_390 and the MG_RS02200 ORF, we anticipated that other genes located immediately downstream of the candidate promoters would also be under the control of the MG428 protein. To test this hypothesis, transcription of the MG_010, MG_414 and MG_525 genes was analyzed by qRT-PCR (Figure 3). Our results confirmed that these three genes were significantly up-regulated when the MG428 protein was overexpressed. In agreement with our previous data, transcript levels of these three genes in the Δ MG_428 and the Tn::MG_390-1 mutants were comparable to those of the WT strain. Furthermore, we conducted primer extension analysis of the MG_414 gene in the WT strain and the Tn::*recA*-2 mu-

tant (Figure 6). As expected, a TSS was detected nine nucleotides downstream of the predicted promoter sequence in the Tn::*recA*-2 mutant but not in the WT strain. Altogether, these results provide compelling evidence that the identified promoter sequence is involved in MG428-regulation.

Single cell analysis of MG428 and RecA expression by fluorescence microscopy

To monitor the expression of MG428 and RecA in the population of *M. genitalium*, mutants carrying MG428-, RecA- or a control Cat-mCherry fusion (Supplementary Figure S6) were obtained by allelic exchange. Several clones of each mutant strain were analyzed by fluorescence microscopy (Figure 7 and Table 2). As expected, expression of the Cat-mCherry fusion under the control of an endogenous *M. genitalium* promoter was detected in virtually the entire population. In contrast, only a 0.46% of the cells carrying the MG428-mCherry fusion exhibited detectable fluorescence. Likewise, the RecA-mCherry fluorescence was detected in a 0.66% of the population. Interestingly, mCherry fluorescence was usually observed in cell pairs (Figure 7), suggesting a spatial association between mycoplasmas with increased levels of MG428 or RecA. In addition, we explored RecA-mCherry expression in a Δ MG_428 mutant background. This strain was obtained by deleting the MG_428 gene from the mutant carrying the RecA-mCherry fusion (Supplementary Figure S6). We could not detect any fluorescent cell in any of the clones analyzed, demonstrating

Table 1. Occurrence of sequences with no more than one mismatch with respect to the consensus sequence TGTGCA-N_{18/19}-ATTWAT in the chromosome of *M. genitalium*^a

Distance from 5' end ^b	Strand	Sequence	Locus at 3' side ^c	Distance to ATG ^d	Protein name	Sequence in <i>Mpn</i> homolog ^e	Locus in <i>Mpn</i>
12029	+	TTGACA-N ₁₉ -ATTTAT	MG_010	7	DNA primase-related protein	TTGGCG-N ₁₉ -ATTTAT	MPN014
67783	-	TTGTCA-N ₁₈ -ATTAAA	MG_059	110	SsrA-binding protein	-	MPN074
210096	+	TTGTCA-N ₁₉ -ATTTTT	MG_184	263	DNA modification MTase	TTGGCT-N ₁₉ -ATTTTT	MPN198
237286	+	TTGTAA-N ₁₉ -ATTAAT	MG_200	29	DNaJ domain-containing protein	-	MPN119
266402	-	TTGGCA-N ₁₈ -ATTAAT	MG_220	16	Hypothetical protein	TTGGCA-N ₁₈ -ATTCAT	MPN313
325946	-	TTGTCA-N ₁₉ -ATTAAA	MG_268	36	Putative deoxyguanosine kinase	-	MPN386
393865	-	TTGTAA-N ₁₉ -ATTAAT	MG_314	200	Hypothetical protein	-	MPN449
395623	-	TTGTAA-N ₁₉ -ATTAAT	MG_316	8	ComEC/Rec2-related protein	TTGGCA-N ₁₈ -AATTAA	MPN451
420161	+	TTGTCA-N ₁₈ -ATTTTT	MG_517	141	Glycosyl transferase	TTGGTG-N ₁₈ -TTTTAA	MPN483
420535	-	TTGTCA-N ₁₉ -AGTTAT	MG_516	230	Hypothetical protein	TTGGTA-N ₁₉ -ATTAGC	MPN482
428158	-	ATGTCA-N ₁₉ -ATTTAT	MG_339	15	Recombinase A	ATGGCG-N ₁₉ -ATTTAG	MPN490
456270	+	TTGTCA-N ₁₈ -ATTAAT	MG_RS02200	26	Hypothetical protein	TTGGCG-N ₁₉ -ATTCAT	MPN534
456490	+	TTGTCT-N ₁₉ -ATTAAT	MG_358	233	Holliday junction DNA helicase	TTGGCA-N ₁₉ -AATTTT	MPN535
489580	-	TAGTCA-N ₁₈ -ATTAAT	MG_386	2	Gliding motility protein	-	MPN567
493564	-	TTGTCA-N ₁₉ -ATTAAT	MG_390	17	ABC transporter	TTGGCA-N ₁₈ -ATTTAT	MPN571
519400	-	TTGTCA-N ₁₈ -ATTTAT	MG_414	15	Hypothetical protein	TTGGCG-N ₁₉ -ATTTAA	MPN612
521461	-	CTGTCA-N ₁₈ -ATTTAT	MG_525	17	Hypothetical protein	TTGGCA-N ₁₉ -ATTTAT	MPN614
531509	-	TTGTAA-N ₁₉ -ATTAAT	MG_423	261	Putative ribonuclease_J	-	MPN621
550394	-	TTGTAA-N ₁₉ -ATTTAT	MG_446	171	30S ribosomal protein S16	-	MPN660

^aIn this Table, only sequences located less than three-hundred nucleotides upstream from the translational start site of a known gene coded in the same strand are shown. For a full list of sequences, see Supplementary Table S7.

^bCoordinates of the G37 genome where the identified sequences are found.

^cFirst known locus found on the same strand at the 3' side of conserved sequence.

^dDistance in base pairs between the 3' end of the sequence and the 5' end of the candidate regulated locus.

^eSequences from *M. pneumoniae* (*Mpn*) are shown if they did not contain more than three mismatches with respect to the consensus sequence TTGGCR-N_{18/19}-ATTYAT. As for *M. genitalium*, the distance in base pairs between the 3' end of the sequence and the 5' end of the candidate regulated locus was less than three-hundred nucleotides.

W: A or T; Y: C or T; R: G or A.

Table 2. Frequency of cells with detectable levels of MG428- and RecA-mCherry fusions

Strain	Analyzed cells	Fluorescent cells	Percentage of fluorescent cells
WT	1172	0	0
Cat:Ch	2249	2009	89.34
RecA:Ch	4529	30	0.66
MG428:Ch	2836	13	0.46
ΔMG_428-RecA:Ch	4551	0	0
RecA:Ch-10	3645	0	0
RecA:Ch-22	3296	19	0.58
RecA:Ch-35	2720	0	0

that the presence of the MG_428 gene was necessary for RecA expression at high levels.

On the other hand, we also created several RecA:Ch mutants with scrambled sequences centered at the -10, -22 and -35 regions of the *recA* promoter (Supplementary Figure S7). No fluorescent cells were observed in the RecA:Ch-10 or the RecA:Ch-35 mutants (Table 2). This result demonstrates that the identified -10 and -35 promoter elements are essential for MG428-dependent activation of *recA*. In contrast, mutation of a non-conserved region located between the -10 and -35 *recA* promoter elements (mutant RecA:Ch-22) had very little or no impact on the percentage of fluorescent cells (0.58%). Of note, the RecA-mCherry fusion was detected by Western blot in all the tested mutants, including those mutants for which no fluorescent cells were observed (Figure 2B). Therefore, basal levels of RecA expression can be detected by western blot but not by fluorescence microscopy. In agreement with previous data, expression of the MG_428-mCherry fusion was not detected by Western blot with anti-mCherry antibodies.

Quantitative assessment of the recombination capacity

As we described earlier, several genes up-regulated by the MG428 protein coded for proteins with a known role in homologous recombination (HR). This fact prompted us to investigate the capacity of the ΔMG_428 mutant and some of the complemented strains to be genetically modified by HR (Figure 8 and Supplementary Figure S8). The transformation efficiency by homologous recombination (TE-HR) of the WT strain with the suicide plasmid pΔMG_218-lacZ was $\sim 1 \times 10^{-6}$ transformants per viable cell. In contrast, repeated attempts to transform the ΔMG_428 strain by HR with the same plasmid failed. Similarly, no transformant colonies could be observed after electroporation of the mutants carrying the Tn*CatMG_428* MiniTnp inserted within the *recA* or the *ruvB* genes (Figure 8). Instead, the presence of a WT copy of the MG_428 allele in the Tn::MG_281-1 mutant fully restored the capacity of this strain to be modified by HR (Figure 8). Therefore, our data indicate that deletion of the MG_428 gene leads to severe recombination defects that can be complemented *in trans* by the reintroduction of an ectopic copy of this gene.

On the other hand, we constructed null mutants of the MG_010, MG_220, MG_390, MG_414, MG_525 and MG_RS02200 genes by allelic exchange and assessed their capacity to be modified by HR (Figure 8). TE-HR of the ΔMG_010 and ΔMG_414 mutants was similar to that of the WT strain. However, mutants lacking the MG_390 or MG_525 genes showed a ten-fold TE-HR reduction as compared to the WT. Furthermore, repeated attempts to transform the ΔMG_220 or ΔMG_RS02200 mutants by HR were unsuccessful. Of note, all these mutants were transformed with a Tn*Cat* MiniTnp with efficiencies similar to that of the WT strain, indicating that DNA entry was not

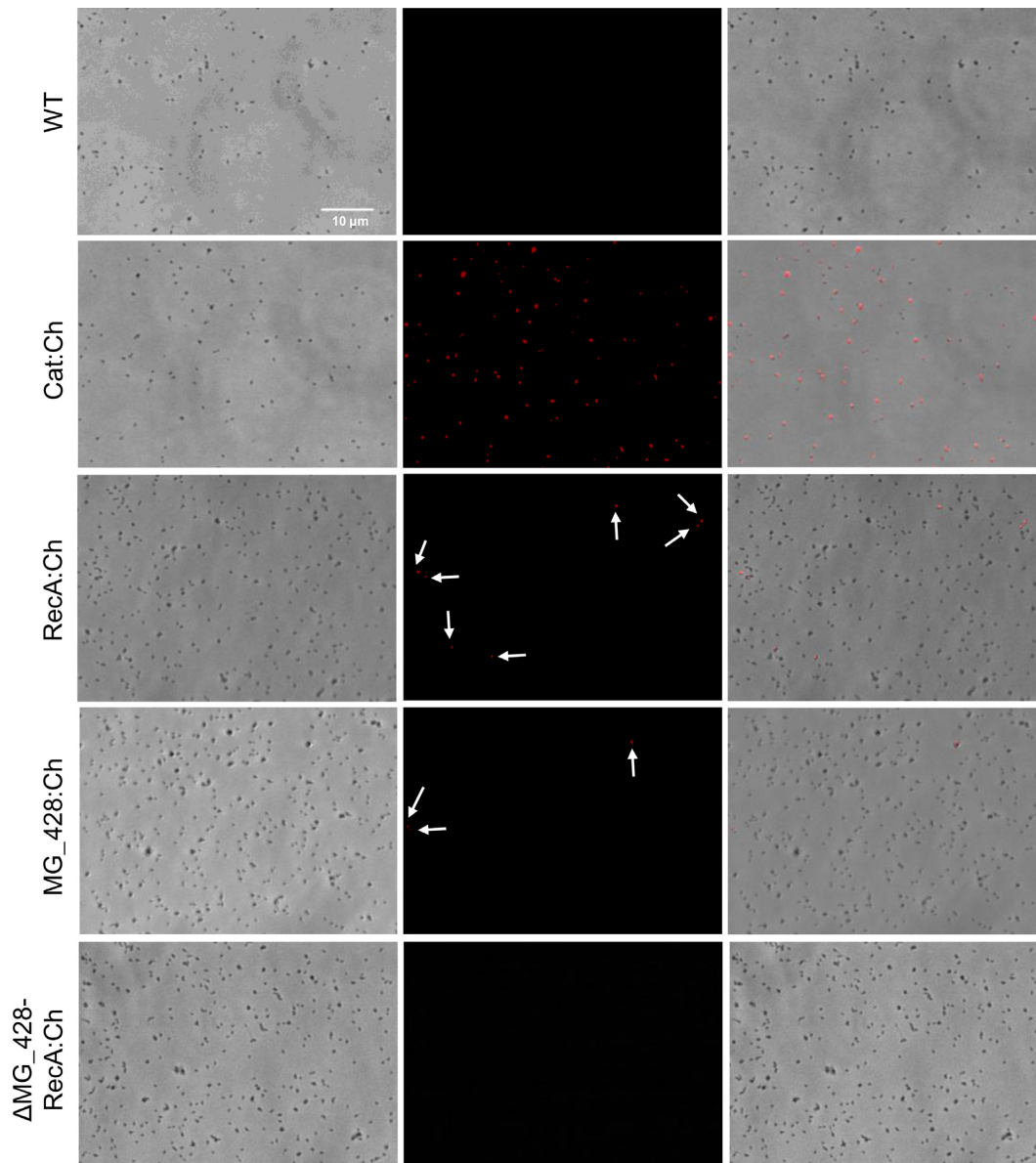


Figure 7. Single cell analysis of Cat-, RecA- and MG428-mCherry expression by fluorescence microscopy. Each row contains a series of three images corresponding to the phase contrast, the Texas Red channel and the resulting overlay, respectively. Arrows indicate the presence of fluorescent cells expressing either RecA- or MG428-mCherry fusions. All pictures are shown at the same magnification.

compromised (data not shown). The recombination capacity of the mutants expressing a RecA-mCherry fusion was also explored (Figure 8). TE-HR of the RecA:Ch mutant was comparable to that of the WT strain. Moreover, transformation of the RecA:Ch-22 mutant with the pΔMG_218-lacZ plasmid was similar to that of the RecA:Ch mutant. In contrast, the RecA:Ch-10 and RecA:Ch-35 mutants could not be transformed by HR.

Analysis of genome variation by next-generation sequencing

HR plays an important role in the generation of genetic variants in the population (37). Therefore, since MG428 functions as a novel regulator of HR in *M. genitalium*, we investigated whether high levels of MG428 expression

correlated with increased genetic diversity. We found that DNA repeats of several MgPar regions exhibited high nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the WT strain or the Tn::MG_390-1 mutant (Figure 9 and Supplementary Figure S9). Essentially, MgPar regions consist of repeated sequences that are homologous, but not identical, to sequences within the MG_191 and MG_192 genes. Similarly, DNA repeats within the MG_191 gene also exhibited high nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the two control strains (Figure 9 and Supplementary Figure S9). Therefore, the relatively high levels of MG428 expression in the Tn::MG_192-1 mutant (Figure 2A), correlate with increased nucleotide sequence variation within DNA repeats. This result is in agreement with a recent report describing

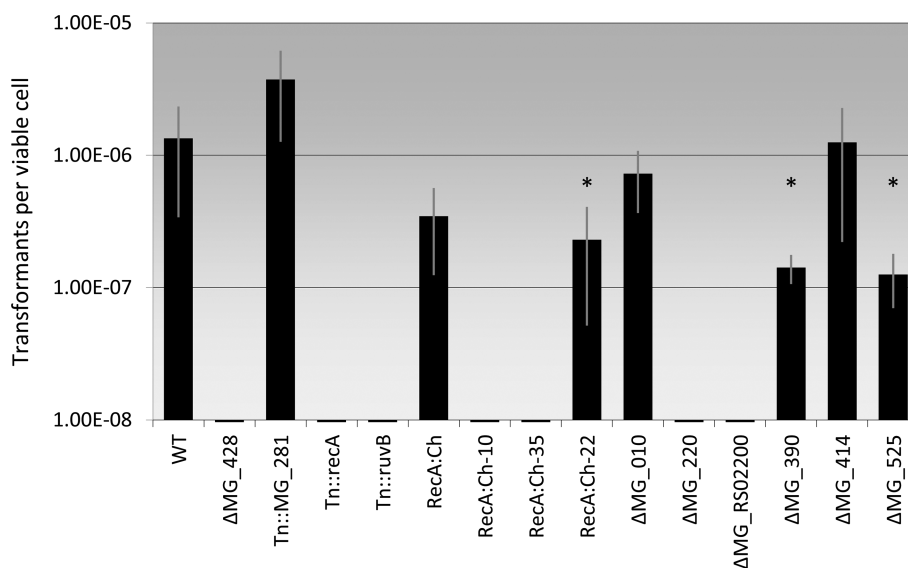


Figure 8. Transformation efficiency by homologous recombination (TE-HR) of different *M. genitalium* strains. Bars show the averages and the standard deviations of at least three independent biological replicates for each strain. TE-HR of strains Δ MG_428, Tn::*recA*-2, Tn::*ruvB*, Δ MG_220 and Δ MG_RS02200 was below the limit of detection ($<10^{-8}$). TE-HR that were found to be statistically different than that of the WT strain are indicated with an asterisk. Statistical significance was assessed with Student's *t* test ($P < 0.05$).

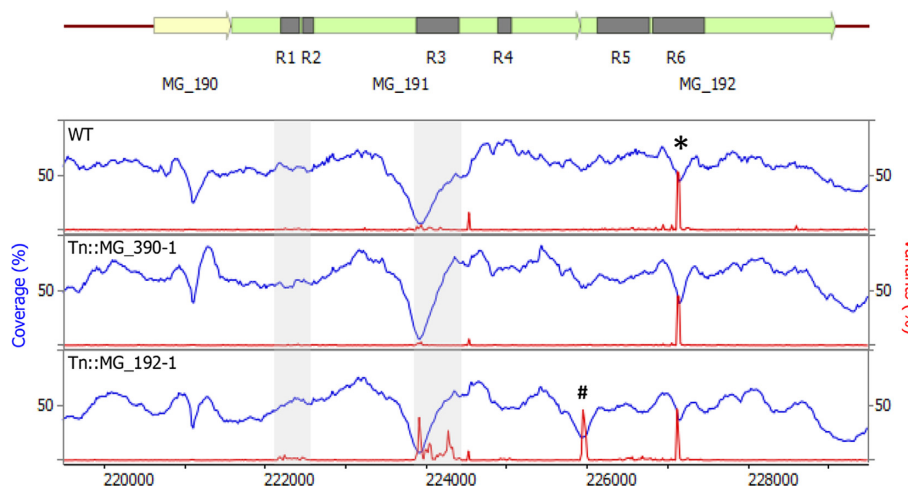


Figure 9. Whole-genome analysis of selected MG_428 complemented mutants by next generation sequencing. A scheme of the *M. genitalium* MgPa operon is shown along with its respective coverage data (in blue) and variant frequencies (in red). Discrete DNA repeats are boxed and numbered (R1 to R6). Genomic regions exhibiting a high rate of nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the WT strain are shadowed. Variations labeled with an asterisk (*) correspond to length polymorphisms of a trinucleotide repeat. A variation labeled with a pound sign (#) corresponds to the insertion point of the Tn*CatMG_428* MiniTnp in the Tn::MG_192-1 mutant (see Supplementary Table S5). Frequency of variants in the transposon insertion point is ~50%. This result can be explained because reads of the positive strand show only sequence variants downstream of the transposon insertion point. Conversely, reads of the negative strand show only sequence variants upstream of the transposon insertion point. Since the graphic coverage shows combined reads from both the positive and the negative strand, this leads to a frequency of variation of about 50% in the regions flanking the transposon insertion site. Genomic regions exhibiting a high rate of nucleotide sequence variation in the Tn::MG_192-1 mutant as compared to the WT strain are shadowed.

the possible function of the MG428 protein in the generation of antigenic diversity in *M. genitalium* (38).

DISCUSSION

In this report, we show that MG428 is a master regulator of recombination in *M. genitalium*. Transcription of the genes coding for RecA and the Holliday junction resolvases RuvA and RuvB was found to be extensively up-regulated by the MG428 protein. These three enzymes are

nearly ubiquitous in bacteria, evidencing their fundamental role in HR (39). However, recombination requires the participation of other enzymes that are less conserved across bacterial species. For example, the RecBCD and RecFOR protein complexes provide two complementary pathways to load RecA onto ssDNA (40). Although several *Mycoplasma* species possess homologs of the recombination genes *recD*, *recO* and *recR*, these genes appear to be absent in the chromosomes of *M. genitalium* and *M. pneumoniae* (41,42). In-

Table 3. Summary of experimentally validated MG428-regulated genes and function

Gene	Function	Mean transcript fold-increase ^a
MG_220	Unknown	164.87
MG_339 (<i>recA</i>)	Recombination and repair	17.65
MG_358 (<i>ruvA</i>)	Recombination and repair	13.54
MG_RS02200	Unknown	9.87
MG_359 (<i>ruvB</i>)	Recombination and repair	7.27
MG_414	Unknown	6.56
MG_010	Replication (putative)	3.34
MG_390	Protein/peptide secretion (putative)	3.29
MG_389	Unknown	3.20
MG_525	Unknown	2.95

^aData shown correspond to the Tn::*recA*-2 mutant compared to the WT strain.

Interestingly, we found that up to seven genes with unknown function were under the control of the MG428 protein (Table 3). Deletion of two of these unknown genes (MG_220 and MG_RS02200) abrogated the capacity of *M. genitalium* to be genetically modified by HR. Moreover, our results indicate that the MG_390 and MG_525 genes may also participate in the recombination pathway of this bacterium. This opens a new avenue in the understanding of recombination in this model organism, as some of these unknown genes may represent a novel set of recombination enzymes. Of note, inactivation of two MG428-regulated genes (MG_414 and MG_010) had very little or no impact in the recombination capacity of *M. genitalium*, indicating that MG428 likely regulates other biological functions beyond recombination.

Unexpectedly, we found that insertion of a MiniTnp that leads to MG428 overexpression was highly selective for genes under the control of the MG428 protein. To our knowledge, this straightforward strategy to identify regulated genes is unprecedented and relies on the potential toxicity observed upon MG428 overexpression. In a previous report, Glass *et al.* described that 31% of the mutants isolated in a global transposon mutagenesis study carried transposons inserted within the MG_428, MG_339, MG_414 and MG_525 genes (43). That is, three genes under the control of the MG428 protein and the MG_428 gene itself. In addition, the authors stated that mutants with a disrupted MG_414 or MG_525 genes grew significantly faster than the WT strain. These results were later confirmed in a comprehensive study comparing growth rates of different *M. genitalium* single-gene mutants (44). Altogether, these findings suggest that inactivation of MG_428 or specific MG428-regulated genes generate mutant cells with better fitness *in vitro*. Conservation of genes that are apparently detrimental for *in vitro* growth of this minimal bacterium, advocates for an important role of the MG_428 regulatory pathway in survival within the host. Work is in progress to identify other genes under the control of the MG428 protein using high throughput technologies.

In agreement with the low cellular abundance characteristic of transcriptional regulators, MG428 expression could not be detected by western blot in the WT strain. However, we found that MG428 was expressed at relatively high levels in a small subpopulation of cells. Similarly, only a few mycoplasmas exhibited detectable fluorescence upon the expression of a RecA-mCherry fusion protein from its native chromosomal locus. This is in contrast to observations in *E. coli*, where the majority of cells expressing a RecA-GFP

fusion exhibited detectable fluorescence under standard laboratory conditions (45). Moreover, RecA-mCherry fluorescent cells were only observed in the presence of the MG_428 gene, indicating that *recA* activation is primarily under the control of the MG428 protein. Altogether, these data indicate that recombination is cell-specific in *M. genitalium* and strictly dependent on the MG428 protein. At the present time, the factors governing the activation of the MG428 regulatory pathway are unknown. Interestingly, it has been described that transcript levels of the *Saccharomyces cerevisiae Rad51* gene, which codes for the functional homolog of RecA in eukaryotes, fluctuate throughout the cell cycle (46). Activation of the MG428 regulatory pathway could also be coordinated with the cell division cycle in *M. genitalium* and accordingly, we observed a spatial association between mycoplasmas with high levels of MG428 or RecA expression. Alternatively, MG428-activation could be linked to a yet unidentified mechanism of cell-to-cell DNA transfer in *M. genitalium*. Of note, conjugal transfer with the incorporation of the incoming DNA into the recipient chromosome by homologous recombination has been recently described in *Mycoplasma agalactiae* (47). The existence of a mechanism promoting horizontal gene transfer and the participation of the MG428-regulated proteins in this process would reinforce a role of MG428 in genome evolution and adaptation in *M. genitalium*.

Aside from the insight into the biological function of the MG428 protein, our data provide compelling evidence that this regulator is an alternative sigma factor of *M. genitalium*. Accordingly, we found that MG428 interacts with RNAP and that this binding is markedly reduced in the presence of the primary sigma-70 factor RpoD. Furthermore, we found that MG428 directs RNAP to specific promoter regions of *M. genitalium*. Of note, MG428-dependent promoters display a novel Pribnow box and a conserved -35 element. The identified -35 element resembles the canonical TTGACA box, which is rarely found in mycoplasma promoters (8). Unlike previous reports demonstrating that the -35 region is of minor importance for promoter function in *M. genitalium* and *M. pneumoniae* (5,9), herein we show that this element is essential for the activation of MG428-dependent promoters. Similarly, the divergent Pribnow box (ATTWAT) identified in this study is also necessary for MG428-regulation. Altogether, our findings demonstrate that transcription in *M. genitalium* is under the control of at least two sigma factors that direct RNAP to different promoter regions.

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

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to B.M. Carpenter (USUHS) for her advice on statistical analysis, the staff of the Servei de Genòmica i Bioinformàtica (UAB) for performing sequencing reactions and S. Bartolomé (UAB) for his assistance in performing quantitative RT-PCR analysis. We are also thankful to D. Scott Merrell (USUHS) for carefully reading and commenting on the manuscript. LC-MS analysis was performed at the Proteomics Laboratory at Vall d'Hebron Institute of Oncology (VHIO) that belongs to ProteoRed (PRB2-ISCIII). This work was presented in part at the 114th General Meeting of the American Society for Microbiology held in Boston, Massachusetts, in May 2014 (Poster H-2619).

FUNDING

Ministerio de Economía y Competitividad [BIO2013-48704R, BFU2013-50176-EXP]. Funding for open access charge: Ministerio de Economía y Competitividad [BIO2013-48704R, BFU2013-50176-EXP].

Conflict of interest statement. None declared.

REFERENCES

- McGowin,C.L. and Anderson-Smits,C. (2011) Mycoplasma genitalium: an emerging cause of sexually transmitted disease in women. *PLoS Pathog.*, **7**, e1001324.
- Gibson,D.G., Glass,J.I., Lartigue,C., Noskov,V.N., Chuang,R.Y., Algire,M.A., Benders,G.A., Montague,M.G., Ma,L., Moodie,M.M. et al. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52–56.
- Gibson,D.G., Benders,G.A., Andrews-Pfannkoch,C., Denisova,E.A., Baden-Tillson,H., Zaveri,J., Stockwell,T.B., Brownley,A., Thomas,D.W., Algire,M.A. et al. (2008) Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science*, **319**, 1215–1220.
- Karr,J.R., Sanghvi,J.C., Macklin,D.N., Gutschow,M.V., Jacobs,J.M., Bolival,B. Jr, Assad-Garcia,N., Glass,J.I. and Covert,M.W. (2012) A whole-cell computational model predicts phenotype from genotype. *Cell*, **150**, 389–401.
- Zhang,W. and Baseman,J.B. (2011) Transcriptional regulation of MG_149, an osmoinducible lipoprotein gene from Mycoplasma genitalium. *Mol. Microbiol.*, **81**, 327–339.
- Zhang,W. and Baseman,J.B. (2011) Transcriptional response of Mycoplasma genitalium to osmotic stress. *Microbiology*, **157**, 548–556.
- Paget,M.S. and Helmann,J.D. (2003) The sigma70 family of sigma factors. *Genome Biol.*, **4**, 203.
- Weiner,J. 3rd, Herrmann,R. and Browning,G.F. (2000) Transcription in Mycoplasma pneumoniae. *Nucleic Acids Res.*, **28**, 4488–4496.
- Halbedel,S., Eilers,H., Jonas,B., Busse,J., Hecker,M., Engelmann,S. and Stulke,J. (2007) Transcription in Mycoplasma pneumoniae: analysis of the promoters of the ackA and ldh genes. *J. Mol. Biol.*, **371**, 596–607.
- Guell,M., van Noort,V., Yus,E., Chen,W.H., Leigh-Bell,J., Michalodimitrakis,K., Yamada,T., Arumugam,M., Doerks,T., Kuhner,S. et al. (2009) Transcriptome complexity in a genome-reduced bacterium. *Science*, **326**, 1268–1271.
- Lluch-Senar,M., Delgado,J., Chen,W.H., Llorens-Rico,V., O'Reilly,F.J., Wodke,J.A., Unal,E.B., Yus,E., Martinez,S., Nichols,R.J. et al. (2015) Defining a minimal cell: essentiality of small ORFs and ncRNAs in a genome-reduced bacterium. *Mol. Syst. Biol.*, **11**, 780.
- Weber Sde,S., Sant'Anna,F.H. and Schrank,I.S. (2012) Unveiling Mycoplasma hyopneumoniae promoters: sequence definition and genomic distribution. *DNA Res.*, **19**, 103–115.
- Bornberg-Bauer,E. and Weiner,J. 3rd (2002) A putative transcription factor inducing mobility in Mycoplasma pneumoniae. *Microbiology*, **148**, 3764–3765.
- Helmann,J.D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.*, **46**, 47–110.
- Osterberg,S., del Peso-Santos,T. and Shingler,V. (2011) Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.*, **65**, 37–55.
- Maier,T., Schmidt,A., Guell,M., Kuhner,S., Gavin,A.C., Aebersold,R. and Serrano,L. (2011) Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.*, **7**, 511.
- Ma,L., Jensen,J.S., Myers,L., Burnett,J., Welch,M., Jia,Q. and Martin,D.H. (2007) Mycoplasma genitalium: an efficient strategy to generate genetic variation from a minimal genome. *Mol. Microbiol.*, **66**, 220–236.
- Wood,G.E., Iverson-Cabral,S.L., Patton,D.L., Cummings,P.K., Cosgrove Sweeney,Y.T. and Totten,P.A. (2013) Persistence, immune response, and antigenic variation of Mycoplasma genitalium in an experimentally infected pig-tailed macaque (Macaca nemestrina). *Infect. Immun.*, **81**, 2938–2951.
- Iverson-Cabral,S.L., Astete,S.G., Cohen,C.R. and Totten,P.A. (2007) mgpB and mgpC sequence diversity in Mycoplasma genitalium is generated by segmental reciprocal recombination with repetitive chromosomal sequences. *Mol. Microbiol.*, **66**, 55–73.
- Pich,O.Q., Burgos,R., Planell,R., Querol,E. and Pinol,J. (2006) Comparative analysis of antibiotic resistance gene markers in Mycoplasma genitalium: application to studies of the minimal gene complement. *Microbiology*, **152**, 519–527.
- Lluch-Senar,M., Vallmitjana,M., Querol,E. and Pinol,J. (2007) A new promoterless reporter vector reveals antisense transcription in Mycoplasma genitalium. *Microbiology*, **153**, 2743–2752.
- Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
- Li,H. (2012) Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics*, **28**, 1838–1844.
- Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- Koboldt,D.C., Chen,K., Wylie,T., Larson,D.E., McLellan,M.D., Mardis,E.R., Weinstock,G.M., Wilson,R.K. and Ding,L. (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, **25**, 2283–2285.
- Ewing,B. and Green,P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.*, **8**, 186–194.
- Ewing,B., Hillier,L., Wendl,M.C. and Green,P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.*, **8**, 175–185.
- Calisto,B.M., Pich,O.Q., Pinol,J., Fita,I., Querol,E. and Carpena,X. (2005) Crystal structure of a putative type I restriction-modification S subunit from Mycoplasma genitalium. *J. Mol. Biol.*, **351**, 749–762.
- Dyballa,N. and Metzger,S. (2009) Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *J. Visual. Exp.: JoVE*, doi:10.3791/1431.
- Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, e45.
- Vandesompele,J., De Preter,K., Pattyn,F., Poppe,B., Van Roy,N., De Paepe,A. and Speleman,F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, **3**, RESEARCH0034.
- Pich,O.Q., Burgos,R., Ferrer-Navarro,M., Querol,E. and Pinol,J. (2008) Role of Mycoplasma genitalium MG218 and MG317 cytoskeletal proteins in terminal organelle organization, gliding motility and cytodherence. *Microbiology*, **154**, 3188–3198.
- Dhandayuthapani,S., Rasmussen,W.G. and Baseman,J.B. (1999) Disruption of gene mg218 of Mycoplasma genitalium through homologous recombination leads to an adherence-deficient phenotype. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 5227–5232.
- Kingsford,C.L., Ayanbule,K. and Salzberg,S.L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.*, **8**, R22.
- Gautheret,D. and Lambert,A. (2001) Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J. Mol. Biol.*, **313**, 1003–1011.

36. Lesnik, E.A., Sampath, R., Levene, H.B., Henderson, T.J., McNeil, J.A. and Ecker, D.J. (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res.*, **29**, 3583–3594.
37. Vink, C., Rudenko, G. and Seifert, H.S. (2012) Microbial antigenic variation mediated by homologous DNA recombination. *FEMS Microbiol. Rev.*, **36**, 917–948.
38. Burgos, R. and Totten, P.A. (2014) MG428 is a novel positive regulator of recombination that triggers *mgpB* and *mgpC* gene variation in *Mycoplasma genitalium*. *Mol. Microbiol.*, **94**, 290–306.
39. Rocha, E.P., Cornet, E. and Michel, B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.*, **1**, e15.
40. Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. and Rehrauer, W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, **58**, 401–465.
41. Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M. *et al.* (1995) The minimal gene complement of *Mycoplasma genitalium*. *Science*, **270**, 397–403.
42. Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.C. and Herrmann, R. (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.*, **24**, 4420–4449.
43. Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchison, C.A. 3rd, Smith, H.O. and Venter, J.C. (2006) Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 425–430.
44. Sanghvi, J.C., Regot, S., Carrasco, S., Karr, J.R., Gutschow, M.V., Bolival, B. Jr and Covert, M.W. (2013) Accelerated discovery via a whole-cell model. *Nat. Methods*, **10**, 1192–1195.
45. Renzette, N., Gumlaw, N., Nordman, J.T., Krieger, M., Yeh, S.P., Long, E., Centore, R., Boonsombat, R. and Sandler, S.J. (2005) Localization of RecA in *Escherichia coli* K-12 using RecA-GFP. *Mol. Microbiol.*, **57**, 1074–1085.
46. Basile, G., Aker, M. and Mortimer, R.K. (1992) Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Mol. Cell. Biol.*, **12**, 3235–3246.
47. Dordet-Frisoni, E., Sagne, E., Baranowski, E., Breton, M., Nouvel, L.X., Blanchard, A., Marena, M.S., Tardy, F., Sirand-Pugnet, P. and Citti, C. (2014) Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *mBio*, **5**, e01958.