The Bacillus subtilis Conjugative Plasmid pLS20 Encodes Two Ribbon-Helix-Helix Type Auxiliary Relaxosome Proteins That Are Essential for Conjugation

Andrés Miguel-Arribas, Jian-An Hao, Juan R. Luque-Ortega, Gayetri Ramachandran, Jorge Val-Calvo, César Gago-Córdoba, Daniel González-Álvarez, David Abia, Carlos Alfonso, Ling J. Wu and Wilfried J. J. Meijer

Bacterial conjugation is the process by which a conjugative element (CE) is transferred horizontally from a donor to a recipient cell via a connecting pore. One of the first steps in the conjugation process is the formation of a nucleoprotein complex at the origin of transfer (oriT), where one of the components of the nucleoprotein complex, the relaxase, introduces a site- and strand specific nick to initiate the transfer of a single DNA strand into the recipient cell. In most cases, the nucleoprotein complex involves, besides the relaxase, one or more additional proteins, named auxiliary proteins, which are encoded by the CE and/or the host. The conjugative plasmid pLS20 replicates in the Gram-positive Firmicute bacterium Bacillus subtilis. We have recently identified the relaxase gene and the oriT of pLS20, which are separated by a region of almost 1 kb. Here we show that this region contains two auxiliary genes that we name aux1LS20 and aux2LS20, and which we show are essential for conjugation. Both Aux1LS20 and Aux2LS20 are predicted to contain a Ribbon-Helix-Helix DNA binding motif near their N-terminus. Analyses of the purified proteins show that Aux1LS20 and Aux2LS20 form tetramers and hexamers in solution, respectively, and that they both bind preferentially to oriT of pLS20, although with different characteristics and specificities.

In silico analyses revealed that genes encoding homologs of Aux1LS20 and Aux2LS20 are located upstream of almost 400 relaxase genes of the RelLS20 family (MOB-L) of relaxases. Thus, Aux1LS20 and Aux2LS20 constitute the founding member of the first two families of auxiliary proteins described for CEs of Gram-positive origin.

Keywords: conjugation, relaxosome, auxiliary protein, DNA binding protein, Ribbon-Helix-Helix, antibiotic resistance, Firmicutes, horizontal gene transfer
INTRODUCTION

Bacteria exchange genetic material at gross scale, even between distantly related species, via different routes collectively called horizontal gene transfer (HGT) (for review see, Ochman et al., 2000; Frost et al., 2005; Thomas and Nielsen, 2005; Boto, 2010). Horizontal exchange of DNA provides bacteria instantly with a new set of gene(s) and hence is an important driver for the rapid adaptation and evolution of bacteria. Among the genes that are spread by HGT are those responsible for antibiotic resistance (AR), which poses a serious and increasingly worrisome economic and health problem at a global scale. Three main mechanisms are responsible for HGT: transformation through natural competence, transduction via bacterial phage, and conjugation (Ochman et al., 2000; Frost et al., 2005; Thomas and Nielsen, 2005). Of these, conjugation appears to be the route that is predominantly responsible for spreading AR genes (Mazel and Davies, 1999; Waters, 1999; Norman et al., 2009; Davies and Davies, 2010). Conjugation is the process by which a conjugative element (CE) is transferred from a donor cell to a recipient cell through a dedicated transportation pore connecting both cells. CEs contain all the genes required for processing the DNA, establishing contact with the recipient cell, those encoding the structural proteins of the connecting pore as well as those for transporting the DNA. CEs can be integrated in a bacterial chromosome or be present on plasmids, which are named integrative and conjugative elements (ICEs) and conjugative plasmids, respectively. Due to the enormous numbers and density of microbes and the constant replenishment of bacteria upon the intake of food and liquids, the intestinal gut of humans and animals is a niche that is particularly apt for emerging, pooling, and spreading AR (Sommer et al., 2009, 2010; Forsberg et al., 2012; Penders et al., 2013).

Conjugative elements are commonly present in Gram-positive (G+) and Gram-negative (G−) bacteria and the basic concepts of the transfer process are conserved (Alvarez-Martinez and Christie, 2009; De la Cruz et al., 2010; Smillie et al., 2010; Goessweiner-Mohr et al., 2013). However, whereas in most systems conjugation involves the transfer of a single DNA strand (see below), DNA is transferred in its double-stranded form during conjugation in G+ mycelial Streptomyces bacteria (Goessweiner-Mohr et al., 2013; Thoma and Muth, 2016), which is not further considered here. Conjugation starts with a process named mating pair formation (Mpf) in which a donor cell recognizes and interacts with a suitable recipient cell. Probably, this triggers the signal for processing the DNA of the CE and subsequent transfer of one of its strands, named T-strand, into the recipient cell. The sophisticated, multi-component pore connecting the donor and the recipient cell is named transerosome, which is a type IV secretion system (T4SS). The enzyme responsible for initiating the generation of the T-strand is a relaxase, a phosphodiesterase, that cleaves the DNA in a strand- and site-specific manner at a specific position called the nic site, which is located within the origin of transfer region (oriT). Relaxase-mediated cleavage generates a hydroxyl group at 3′ end of the nic site which functions as a primer for DNA elongation; i.e., the relaxase initiates a rolling-circle type of DNA replication (also named DNA transfer replication [Dtr]). Upon nicking, the relaxase remains covalently attached to the 5′-end of the nicked T-strand which is then transferred, together with the attached T-strand, into the recipient cell. In most cases the active site residue that becomes covalently attached to the T-strand concerns a tyrosine. However, very recently it has been shown that relaxases of the MOB family employ a histidine instead of a tyrosine residue to nick the DNA (Pluta et al., 2017). Due to its crucial role in conjugation, relaxases have attained considerable attention and several of them have been characterized in detail at the biochemical, functional and structural levels. In some cases, for instance ICEBs1 of Bacillus subtilis and the broad host range conjugative plasmid pIP501, the relaxase is the only protein that is required for processing the DNA (Kopec et al., 2005; Lee and Grossman, 2007; Grohmann et al., 2016). However, in the majority of cases additional protein(s), encoded either by the CE or the host, bind to the oriT and are involved in processing of the DNA. The nucleoprotein complex at oriT formed by the relaxase and additional proteins is called the relaxosome, and the additional proteins are named auxiliary or accessory proteins. Although their name may suggest that they play secondary role(s) in the processing reaction, most if not all of the auxiliary proteins studied so far have been shown to be essential for conjugation.

Most conjugation studies are based on CE present in G− bacteria, with knowledge on conjugation-related aspects in G+ bacteria lagging far behind. This is especially the case for auxiliary proteins (see Discussion). In our laboratory we study the conjugative plasmid pLS20 which was originally isolated from the Gram− Firmicute bacterium B. subtilis natto IFO3335 (Tanaka et al., 1977). This strain is used for the fermentation of soybeans to produce “natto,” a popular dish in South Asia, and hence it is conceivable that pLS20 or relatives play a role in the conjugation-mediated HGT in the gut of humans and animals. A derivative of pLS20 containing a chloramphenicol-resistance gene, pLS20cat, has been constructed (Itaya et al., 2006) and its sequence has been determined in our lab and in the lab of M. Itaya (Mitsuhiro Itaya, Keio University, Japan). All conjugation genes are located in one large operon spanning genes 28 till 74 according to our nomenclature (Singh et al., 2013). pLS20cat genes 25-27 are involved in regulating the expression of the conjugation genes (Singh et al., 2013; Ramachandran et al., 2014). Recently, we have identified and characterized the relaxase (gene 58) and the oriT of pLS20cat, which we named RelLS20 and oriTLS20, respectively (Ramachandran et al., 2017). Contrary to many other plasmids, the relaxase gene and oriT are located within its large conjugation operon, and RelLS20 turned out to be the founding member of a novel relaxase family containing > 800 members.

Here, we addressed the question whether pLS20cat contains auxiliary relaxosome genes. We demonstrate that genes 56 and 57, located in between the relaxase gene relLS20 and oriTLS20 are two auxiliary genes that are essential for conjugation and denominated them as auxLS20 and auxLS20, respectively. Both gene products were purified and biochemical analyses showed that one of them formed tetramers and the other hexamers in solution. We also show that the proteins bind to distinct DNA motifs present in oriTLS20. In silico analyses revealed that a large fraction of the relaxase genes coding for the MOB family.

Frontiers in Microbiology | www.frontiersin.org 2 November 2017 | Volume 8 | Article 2138
family of relaxases are preceded by genes encoding homologs of Aux1LS20 and/or Aux2LS20. The findings obtained for Aux1LS20 and Aux2LS20 are placed in perspective with other auxiliary proteins of CE present in G+ and G− organisms.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, Media and Oligonucleotides

*Escherichia coli* and *B. subtilis* strains were grown in Luria-Bertani (LB) liquid medium or on 1.5% LB agar plates. When appropriate, media were supplemented with the following antibiotics: ampicillin (100 µg/ml), erythromycin (1 and 150 µg/ml in *B. subtilis* and *E. coli*, respectively), chloramphenicol (5 µg/ml), spectinomycin (100 µg/ml), and kanamycin (10 and 30 µg/ml in *B. subtilis* and *E. coli*, respectively). *B. subtilis* strains used were isogenic with *B. subtilis* strain 168 and are listed in Supplementary Table S1. Plasmids and oligonucleotides used are listed in Supplementary Tables S2, S3, respectively. All oligonucleotides were purchased from Isogen Life Science, Netherlands.

Transformation

*Escherichia coli* cells were transformed using standard methods (Sambrook et al., 1989). Preparation of competent *B. subtilis* cells and transformation were carried as described before (Bron et al., 1989). Transformants were selected on LB agar plates with appropriate antibiotics. pLS20cat encodes a protein, Rok, isogenic with *B. subtilis* (Sambrook et al., 1989). Preparation of competent *E. coli* was isolated from *Plasmid DNA of the constructed pAXO1 and pDR110 derivatives genes comK inhibits the development of competence by repressing the key transcriptional activator of competence genes (Singh et al., 2012). Therefore, to manipulate genes on pLS20cat we prepared competent cells of a pLS20cat-harboring strain that contains the chromosomal P<spank>spank</span>K-<spank>comK</span> fusion (PK556) using a standard protocol (Singh et al., 2012).

Construction of Plasmids and Strains

The correctness of sequences of all cloned PCR fragments was confirmed by sequence analysis. Amplification by PCR of pLS20cat regions was performed using as template total DNA isolated from pLS20cat harboring strain PKS11. Details regarding the construction of integration vectors based on plasmids pDR110 (*amyE* integration vector with IPTG-inducible P<sub>spank</sub>Promoter) or pAX01 (*lacA* integration vector with xylose-inducible P<sub>xyr</sub>Promoter) are given in Supplementary Table S2. In summary, gene 56 was cloned under the control of the P<sub>xyr</sub>Promoter or the P<sub>spank</sub>Promoter. In addition, genes 56-57-58, genes 57-58, or gene 58 were cloned behind the P<sub>spank</sub>Promoter. Plasmid DNA of the constructed pAX01 and pDR110 derivatives was isolated from *E. coli* cells and then used to transform competent *B. subtilis* cells. Double-crossover integration into the chromosome was checked by PCR in the case of the pAX01-derivatives. When pDR110 derivatives were used to transform competent *B. subtilis* cells, double cross over integration was tested by the loss of amylase activity. The pLS20cat genes 58 (rel<sub>LS20</sub>), 57 (aux1<sub>LS20</sub>) and 56 (aux2<sub>LS20</sub>) were cloned in the *E. coli* expression vector pET28b+ to generate fusion genes containing a C-terminal his<sub>6</sub> extension. Details regarding these cloning strategies are given in Supplementary Table S2. The resulting derivatives of pET28b+ were constructed using *E. coli* strain XL1-Blue. Once verified its correctness, the plasmids were transformed into *E. coli* strain BL21(DE3).

Conjugation Assays

Conjugation was carried out in liquid medium as described previously (Singh et al., 2013). The effect of ectopic expression of a given gene placed under the control of the inducible P<sub>spank</sub> and/or P<sub>xyr</sub> promoter on conjugation was studied by adding the inducer (1 mM IPTG, 1% xylose) to prewarmed LB medium used to dilute overnight cultures of the donor cells.

Analytical Ultracentrifugation Experiments

Sedimentation velocity (SV), sedimentation equilibrium (SE), and dynamic light scattering (DLS) assays and processing of the data, including estimations of molar masses of the relaxosome proteins from the hydrodynamic measurements, were carried out using the same conditions to those used before in the analysis of Rel<sub>LS20</sub> (Ramachandran et al., 2017).

Over Expression and Purification of Recombinant Rel<sub>LS20</sub>, Aux1<sub>LS20</sub>, and Aux2<sub>LS20</sub> Containing a C-Terminal His<sub>6</sub> Tag

Recombinant versions of Rel<sub>LS20</sub>, Aux1<sub>LS20</sub>, and Aux2<sub>LS20</sub> were expressed and purified using similar protocols. In brief, *E. coli* BL21(DE3) cells containing plasmid pAND83 (rel<sub>LS20</sub>His<sub>6</sub>), or pHJA56 (aux1<sub>LS20</sub>His<sub>6</sub>), or pHJA57 (aux2<sub>LS20</sub>His<sub>6</sub>) were inoculated in fresh LB media complemented with 30 µg/ml kanamicin and grown at 37°C with shaking (200 rpm). At an OD<sub>600</sub> of about 0.6, IPTG was added to a final concentration of 1 mM to induce the recombinant protein and growth was continued for 2 h. Next, cells were collected by centrifugation and processed as described before (Singh et al., 2012). The nickel-column purified proteins (>95% pure) were finally dialysed against buffer B (20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM NaCl, 10 mM MgCl<sub>2</sub>, 7 mM β-mercaptoethanol, 50% v/v glycerol) and stored in aliquots at −80°C. Bradford assay and OD<sub>280</sub> determination were used to determine the protein concentrations.

Gel Retardation Assays

Gel retardation assays were essentially carried out as described before (Singh et al., 2012). Thus, different DNA fragments were amplified by PCR using pLS20cat as template. The resulting PCR fragments were purified and 170 ng of DNA [200 or 362 bp] (with or without 220 ng of control DNA [176 bp]) were incubated on ice in binding buffer [20 mM Tris HCl pH 8, 1 mM EDTA, 5 mM MgCl<sub>2</sub>, 0.5 mM DTT, 100 mM KCl, 10% (v/v) glycerol, 0.05 mg ml<sup>−1</sup> BSA] without or with purified Aux1<sub>LS20</sub> or Aux2<sub>LS20</sub> to a fixed final concentration of 90 nM (Supplementary Figure S3) or using twofold increasing concentrations ranging from 0.09 to 5.76 µM (Figure 3) in a total volume of 16 µl. The
negative control, corresponding to bp numbers 63,774–63,950 of accession number NC_015148.1, has an AT-content that is very similar to the AT content of the oriT fragment (61.4 vs. 61.1%). This DNA corresponds to sequences located inside a gene (gene 24), lowering the possibility that it harbors particular features for recruiting a transcriptional regulator or other DNA binding protein. In addition, it is predicted to lack a static bend.

After careful mixing, samples were incubated for 20 min at 30°C, placed back on ice for 10 min, then loaded onto 2% agarose gel in 0.5XTBE. Electrophoresis was carried out in 0.5XTBE at 50 V at 4°C. Finally, the gel was stained with ethidium bromide, destained in 0.5XTBE and photographed with UV illumination.

**In Silico Analyses**

**Identification of MobL Members**

RelL520 was used as a query sequence to execute a psi-blast (version 2.6.1+) search against the NCBI nr protein database (July, 2017), allowing up to 10 rounds of reiteration with an e-value threshold of 1e-15 (Altschul et al., 1997, 2005; Schaffer et al., 2001) producing 1445 hits. The program “USEARCH” (version v10.0.240_i86linux32) was then used to identify and remove redundant sequences showing 100% identity (Edgar, 2010), resulting in 1249 unique hits showing high similarity to RelL520.

**Identification of Putative Auxiliary Proteins**

Protein sequences of Aux1L520 and Aux2L520 were used as query against the NCBI nr protein database (July 2017) using psi-blast (version 2.6.1+), with an e-value threshold of 1e-6 and 1e-7, respectively, until no new hits were retrieved. The sequence identifiers obtained from psi-blast, were cross-referenced with the sequence identifiers preceding the relaxase gene, members, obtained from the nucleotide entries from they were translated.

**Prediction of Secondary Structure for Aux1L520 and Aux2L520 Homologs**

Corresponding sequences were submitted to the RaptorX property web server (Wang et al., 2016) and predictions for β-strands and α-helices along the sequences were plotted with “R” (R Core Team, 2017).

**RESULTS**

**Identification of Putative Relaxosome Genes of pLS20cat by in Silico Analysis**

Recently, we have shown that pLS20cat gene 58 is essential for conjugation and that it encodes the relaxase, which we named RelL520 (Ramachandran et al., 2017). In these studies we also identified the nic site of RelL520 and delineated the functional oriT, named oriTL520, to a region of 362 bp. Remarkably, oriTL520 and relL520 are separated by a region of 865 bp, which has been annotated to contain two relatively small putative genes, designated genes 56 and 57 (Singh et al., 2013, see Figure 1 for a schematic view of this region). Often, but not always, conjugative plasmid-located relaxase genes are accompanied by small auxiliary relaxosome genes that generally are located upstream of the relaxase gene. This prompted us to investigate whether genes 56 and 57 might encode auxiliary relaxosome genes of pLS20cat.

In silico analyses of pLS20cat genes 56 and 57 showed that, firstly, relL520 is translationally coupled to the preceding gene 57 [i.e., the stop (TAA) and start codon (ATG) of genes 57 and relL520, respectively, overlap; see Figure 1], and only a small intergenic region of 183 bp separates gene 57 from its preceding gene 56. Second, gene 56 and 57 are both small genes (79 and 147 codons, respectively). And third, the proteins encoded by these genes are both putative DNA binding proteins predicted to contain a Ribbon-Helix-Helix (RHH) motif in their N-terminal regions. An overview of the secondary structure prediction of both proteins and their homology with CopG, a paradigm of RHH DNA binding protein (Gomis-Ruth et al., 1998; Del Solar et al., 2002), is shown in Supplementary Figure S1. This figure shows that both Aux1L520 and Aux2L520 contain several lysine and arginine residues near the end of their predicted helix 1 and beginning of helix 2. The corresponding region in known RHH structures has been shown to be close to the phosphate backbone of the DNA (for example see, Schildbach et al., 1999). In summary, in silico analyses suggested that the two small genes 56 and 57 preceding the relaxase gene relL520 may encode auxiliary relaxosome proteins.

**pLS20cat Genes 56 and 57 Are Essential for Conjugation**

Previously, we engineered a derivative of pLS20cat, pLS20Δ56-58, in which the putative genes 56-57 together with the relaxase gene relL520 (gene 58) have been deleted, and demonstrated that this plasmid was deficient in conjugation. Conjugation of pLS20Δ56-58 was restored when all three genes (56-58), were ectopically expressed from the IPTG-inducible Pspank promoter at the chromosomal amyE locus, but not in the absence of gene 58, showing that RelL520 was essential for conjugation (Ramachandran et al., 2017). We used a similar approach to study whether genes 56 and/or 57 were essential for conjugation. Thus, we constructed strain GR153, which harbors pLS20Δ56-58 and also contains relL520 (gene 58), but not 56 and 57, under the control of the Pspank promoter at the amyE locus. We then employed this strain as donor to determine the conjugation efficiencies using a standard protocol (see Materials and Methods). Strains PKS11, GR149 and GR150 were included as controls. As shown in Table 1, the efficiency of conjugation observed for the wild type plasmid pLS20cat was in the range of 10−3, which is similar to values reported previously under similar conditions (Singh et al., 2013; Ramachandran et al., 2014, 2017). As reported before (Ramachandran et al., 2017), conjugation was observed for pLS20Δ56-58 only when genes 56-58 were expressed from the chromosome (Table 1, strain GR149 and GR150). Importantly, no transconjugants were obtained when strain GR153 (amyE::Pspank−relL520, pLS20Δ56-58) was used as donor in conjugation experiments, regardless of whether they were grown in the presence or absence of IPTG. These results

---

1https://www.R-project.org/
FIGURE 1 | The pLS20cat relaxosome region. (A) Genetic organization of the pLS20cat genes 55–59. The genes, their sizes and orientation of transcription are indicated with arrows. Genes 55 and 59 are colored gray. Genes 56, 57, and 58 (rel$_{LS20}$) are colored green, orange, and yellow, respectively. The same color code is used in “B,” as well as in Figure 4 (see below). The 362 bp oriT$_{LS20}$ region is indicated with a blue box labeled oriT. Base pair numbering is given on the top. (B) DNA sequence of genes 56 and 57 and their deduced protein sequences. Stop codons are indicated with an asterisk and likely Ribosomal Binding sites (RBS) are highlighted with a red box. Note that genes 57 and rel$_{LS20}$ are translationally coupled. Only the first 11 codons of the rel$_{LS20}$ gene are given.

TABLE 1 | pLS20cat genes 56 and 57 are required for conjugation.

| Strain     | Genotype | Plasmid         | Inductor | Conjugation efficiency$^*$ |
|------------|----------|-----------------|----------|---------------------------|
| PKS11      | 168      | pLS20cat        | –        | $7.9 \times 10^{-3}$      |
| GR149      | 168      | pLS20Δ56-58     | –        | $<10^{-7}$                |
| GR150      | 168, amyE::P$^{spank}$ 56-58 | pLS20Δ56-58 | –        | $<10^{-7}$                |
|            |          |                 | +        | $3.2 \times 10^{-3}$      |
| GR153      | 168, amyE::P$^{spank}$ 58 | pLS20Δ56-58 | –        | $<10^{-7}$                |
|            |          |                 | +        | $<10^{-7}$                |
| GR197      | 168, amyE::P$^{spank}$ 58 | pLS20Δ56-58 | –        | $<10^{-7}$                |
|            |          |                 | +        | $<10^{-7}$                |
| GR200      | 168, amyE::P$^{spank}$ 58, lacA::P$_{xyl}$-56 | pLS20Δ56-58 | –        | $<10^{-7}$                |
|            |          |                 | +        | $5.8 \times 10^{-4}$      |
| GR225      | 168, amyE::P$^{spank}$ 57-58, lacA::P$_{xyl}$-56 | pLS20Δ56-58 | –        | $<10^{-7}$                |

$^*$Conjugation efficiencies were calculated as transconjugants/donor, and correspond to the mean value of at least three independent experiments. When indicated, the inducer IPTG was added at a final concentration of 1 mM in the case of strains GR150 and GR197. In the case of GR200 and GR225, the final concentrations of the inducers was 1 mM (IPTG) and 1% (xylene).
showed that pLS20cat gene 56 and/or 57 are necessary for conjugation.

We next tested whether only one or both genes were required for conjugation. For this, we constructed the pLS20Δ56-58-harboring strains GR197 and GR200 in which relSLS20 together with either gene 57 (strain GR197) or gene 56 (strain GR200) could be induced from the bacterial genome. When used as donor, no transconjugants were obtained for each strain regardless whether they were grown in the absence or presence of the inductor(s) (see Table 1), demonstrating that both genes are essential for conjugation.

In the above conjugation experiments, one or a combination of genes 56, 57, relSLS20 was complemented by expressing them from the IPTG-inducible Pspank promoter for all the strains except for strain GR200. In this strain relSLS20 is controlled by Pspank at the amyE locus and gene 56 by the xylose-inducible PxyC promoter at the lacA locus. To rule out the possibility that transconjugants were not obtained for donor strain GR200 because the genes were expressed from different promoters at a different locus, we constructed strain GR225 in which gene 56 was placed under the control of the PxyC promoter at lacA, and genes 57 and 58 under the control of the Pspank promoter at amyE. Transconjugants were obtained for this strain when cells were grown in the presence of both inducers (Table 1), demonstrating that the gene products expressed from the two different promoters and chromosomal loci were all functional. These results demonstrate therefore that besides relSLS20 genes 56 and 57 are also required for conjugation. Taking into account these results, together with the structural organization of these genes with respect to relSLS20 and oriTLS20, the in silico analyses presented above, and additional evidence presented below, we conclude that pLS20cat gene 56 and 57 encode auxiliary relaxosome proteins which we name Aux1LS20 and Aux2LS20, respectively.

In Vitro Analysis of the Relaxosome Proteins Aux1LS20 and Aux2LS20, and RelSLS20

Oligomerization State Determined by Analytical Ultracentrifugation and DLS Techniques

To characterize the auxiliary relaxosome proteins in vitro, we purified Aux1LS20 (Mw 10,601 Da) and Aux2LS20 (Mw 18,605 Da) from E. coli, each fused to a His6 tag at its C-terminus. We first determined the oligomerization state of the proteins, and also investigated putative interactions among them and with RelSLS20, using two complementary analytical ultracentrifugation approaches, i.e., SV and SE (Figures 2A–D), together with DLS experiments using the same experimental conditions.

Sedimentation profiles obtained by SV assays showed Aux1LS20 as a single species with an experimental sedimentation coefficient of 2.5 S (s20, w = 2.9 S) compatible with a moderately elongated tetrameric form of the protein (f/fθ = 1.5) (Figure 2A). Subsequent analysis of Aux1LS20 gave a D-value of 52.5 ± 0.3 μm2/s. The obtained S- and D-values, once introduced in the Svedberg equation, yielded an apparent molar mass of 46,290 Da. SE data, best-fit analysis to single species model gave an average molecular mass of 42,200 Da ± 300 Da, confirming that Aux1LS20 is a tetramer in solution (Figure 2B).

In the case of Aux2LS20, analysis of the sedimenting boundaries showed a sedimentation profile with a main peak corresponding to 90.0% of the total proteins at 4.4 S (s20, w = 5.1 S), together with a second peak at 3.3 S (s20, w = 3.8 S) encompassing 7% of the sample (Figure 2C). The S-value of the main peak is compatible with the theoretical behavior of a spherical Aux2LS20 tetramer (f/fθ = 1.2), as well as with a moderately elongated hexamer (f/fθ = 1.6). DLS analysis of Aux2LS20 yielded a D of 38.2 ± 1.0 μm2/s, which combined with the obtained S-value of 4.4 in the Svedberg formula resulted in an apparent molar mass of 113,400 Da that is very close to the molecular mass of Aux2LS20 hexamers (111,630 Da). SE experiments were decisive for establishing the oligomerization state of Aux2LS20, as the best fit of the SE data gave an average molecular mass of 111,300 ± 1,200 Da, unequivocally demonstrating that Aux2LS20 forms hexamers in solution (Figure 2D). In summary, the outcome of three complementary experimental approaches showed that Aux1LS20 and Aux2LS20 form tetramer and hexamers in solution, respectively.

Previously, we determined that purified RelSLS20 forms monomers in solution (Ramachandran et al., 2017). To study possible interactions between the relaxosome proteins in solution we used combinations of Aux1LS20, Aux2LS20 and RelSLS20 and subjected these to SE experiments (Supplementary Figure S2). No additional peaks with increased S-values reflecting new protein hetero-complexes were obtained in any of the combinations tested implying that the relaxosome proteins of pLS20cat do not interact in solution, at least not under the conditions tested.

Aux1LS20 and Aux2LS20 Bind Specifically to oriTLS20

Electrophoretic Mobility Shift Assays (EMSA) were performed to study the DNA binding properties of Aux1LS20 and Aux2LS20. The results presented in Figure 3 show that both auxiliary proteins bound DNA, and that both bound preferentially to oriTLS20. Nevertheless, there were distinct differences in binding characteristics between the two proteins. The addition of Aux1LS20 resulted in the appearance of only one retarded species of oriTLS20, and even at the highest concentration tested Aux1LS20 did not bind to the negative control DNA (Figure 3, left panel). One retarded oriTLS20 species was also observed for Aux2LS20 at low concentrations. However, higher Aux2LS20 concentrations resulted in the appearance of additional shifted species of oriTLS20. In addition, at higher concentrations Aux2LS20 bound also to the negative control DNA, and at the highest concentration tested a smear of retarded species was observed (Figure 3, right panel). These results show that both proteins bind preferentially to oriTLS20, but Aux1LS20 appears to bind oriTLS20 with a higher specificity than Aux2LS20.

To delineate further the binding sites of Aux1LS20 and Aux2LS20 we generated thirteen overlapping 200 bp DNA fragments (F21–F33) covering the oriTLS20 region with a sliding window of 25 bp, and used them in EMSA. The results presented in Supplementary Figure S3 show that Aux1LS20 bound to fragments F22–F29, which share the 25 bp sequence 5’-CAAATAATCTGTGCACGAAAA-3’ located in the 5’...
half of oriT_{LS20}. This sequence contains the inverted repeat 5'-TGGTACCA-3', which could be the binding site of Aux1_{LS20}. In the case of Aux2_{LS20} retarded species of oriT_{LS20} with strong and weak intensity were observed for fragments F21–F25 and F26–F28, respectively. No shifts were observed for fragments F29–F33 at the protein concentration used. This shows that Aux2_{LS20} binds the 5' half region of oriT_{LS20} upstream of Aux1_{LS20}. The sequence motif 5'-TGTGCAT-3' is present three times in a directed repeated orientation in the 5' half of oriT_{LS20}. While fragments F21–F25 each contain the three 5'-TGTGCAT-3' motifs, fragment F26 contains only two, and the motif is present only once on fragments F27 and F28. This suggests that the motif 5'-TGTGCAT-3' may be the preferred binding site for Aux2_{LS20}. It is worth mentioning that two of the 5'-TGTGCAT-3' motifs are embedded within a larger motif (5'-TTTATGTGCATT-3').

Over 400 Members of the MOB_L Family of Relaxase Genes Contain Upstream Genes Encoding Homologs of Aux1_{LS20} and/or Aux2_{LS20}

Previously, we reported that the pLS20cat-encoded Rel_{LS20} constitutes the founding member of a novel, large family of relaxases that we named MOB_L, which contained 817 members that were almost exclusively encoded in bacteria belonging to the phylum Firmicutes (Ramachandran et al., 2017). We wanted to know whether other MOB_L relaxase genes were also preceded by genes encoding putative homologs of Aux1_{LS20} and/or Aux2_{LS20}. To study this we first determined the current number of MOB_L relaxase genes, applying the same method as that used in our previous study; i.e., we performed a psi-blastp search of the NCBI nr database using Rel_{LS20} as a query. After removing redundant sequences this search now resulted in 1,453 hits that showed high similarity with Rel_{LS20} (threshold value $P = 1e-15$). Next, the corresponding DNA accession number of each identified MOB_L relaxase was retrieved, which was subsequently used to generate a database that contains the accession number of each MOB_L member together with that of the protein encoded by the gene located upstream and downstream of the relaxase gene. We then performed the same procedures for Aux1_{LS20} and Aux2_{LS20}; i.e., we identified proteins sharing a high level of similarity with Aux1_{LS20} and Aux2_{LS20} and generated databases that contained these accession numbers together with those of the proteins encoded by the flanking genes. Finally, the three databases were crossed to identify those MOB_L members that are preceded by a gene encoding a putative homolog of Aux1_{LS20} and/or Aux2_{LS20}. This approach revealed 387 MOB_L.
relaxase genes that were preceded by a gene encoding a putative Aux2_L20 homolog; and of these 87 contained an additional Aux1_L20 homolog encoding gene upstream. Without exception, the identified MOB1 relaxase genes having upstream gene(s) encoding putative homologs of Aux1_L20 and/or Aux2_L20 are all present in bacteria belonging to the phylum Firmicutes. Although stringent settings were used to identify proteins sharing high similarity with Aux1_L20 or Aux2_L20, this does not automatically imply that the identified proteins will contain a Ribbon-Helix-Helix motif in their N-terminal region, which is a characteristic feature of both Aux1_L20 and Aux2_L20 (see above). We therefore carried out secondary structure prediction for all the putative Aux1_L20 and Aux2_L20 homologs identified (see Materials and Methods). The results of these analyses, which are presented in Supplementary Table S4, show that 86 of the 87 (98.9%), and 384 of the 387 (99.2%) putative relaxase genes having upstream gene(s) encoding putative homologs of Aux1_L20 and Aux2_L20, respectively, contain a typical Ribbon-Helix-Helix signature in their N-terminal region, and thereby support the view that they are auxiliary proteins of the corresponding relaxase. In summary, these analyses provide compelling evidence that almost 400 MOB1 relaxase genes are preceded by a gene encoding an Aux2_L20 homolog, and that in 87 of these cases this putative auxiliary gene is preceded by another auxiliary gene encoding an Aux1_L20 homolog. Consequently, pLS20 encoded Aux1_L20 and Aux2_L20 are the founding members of two families of Ribbon-Helix-Helix type auxiliary proteins that are encoded by Firmicutes bacteria.

**DISCUSSION**

In this study we have demonstrated that the pLS20cat genes 56 (aux1_L20) and 57 (aux2_L20) encode the auxiliary relaxosome proteins of pLS20cat. Combined with our previously published results (Ramachandran et al., 2017), we have identified the relaxosome module of pLS20cat that includes oriT_L20 and the downstream genes aux1_L20, aux2_L20, and rel1_L20. This module is embedded within the large conjugation operon of pLS20cat (Singh et al., 2013). In addition, we have provided strong evidence that Aux1_L20 and Aux2_L20 constitute the founding member of corresponding families of Ribbon-Helix-Helix type auxiliary proteins whose genes precede a large fraction of the MOB1 type relaxase genes. Thereby, our results provide a better understanding of the relaxosome components present on Gram+ mobile elements in general and particularly those belonging to the phylum Firmicutes.

The results presented here, together with those obtained previously (Ramachandran et al., 2017), show that aux1_L20 and aux2_L20 encode trans-acting proteins that are essential for conjugation. We also showed that Aux1_L20 and Aux2_L20 form tetramers and hexamers in solution, respectively, and we detected no interaction between the three pLS20 relaxosome proteins under the conditions tested. We cannot exclude the possibility that they interact when they form a nucleoprotein complex at oriT_L20. Aux1_L20 bound with high specificity to a region of 25 bp located about 100 bp upstream of the nic site that contains the inverted repeated sequence 5′-TGGTACCA-3′.
The preferred binding site of Aux2LS20 resulted to be a 140 bp fragment located at the 5’ half of oriTLS20 and that contains three times the sequence 5’-TGTGCAT-3’. In our previous study (Ramachandran et al., 2017), we showed that a derivative of oriTLS20 that includes the nic site and the binding site for Aux1LS20, but lacks the 5’-located 100 bp containing two of the three 5’-TGTGCAT-3’ motifs was not functional in vivo. The topology of DNA can have a large effect on the binding characteristics of DNA binding proteins and which in turn may affect their function (Gimenes et al., 2008; Fogg et al., 2012). The oriT regions of several conjugative plasmids contain an intrinsic bend that is thought to be important for optimal binding and functionality of the relaxosome proteins (for review see, De la Cruz et al., 2010). We have demonstrated that the oriTLS20 region is also intrinsically bent, and that the bend is located in the 5’ half of oriTLS20 (Ramachandran et al., 2017), which we show here corresponds to the region where Aux1LS20 and Aux2LS20 (preferentially) bind. When we combine the results obtained here and in our previous study a picture emerges that is schematically presented in Figure 4. Aux1LS20 and Aux2LS20 bind to the left half of oriTLS20 that is intrinsically bent and we envisage that the formation of this nucleoprotein complex contributes to optimal functioning of Rel1LS20. In other systems, auxiliary proteins have been described to stimulate relaxase-mediated nicking at oriT by recruiting the relaxase to oriT, probably by facilitating the relaxase to access the nic site, and/or by acting as molecular wedges to melt double-stranded DNA (reviewed in, Alvarez-Martinez and Christie, 2009). Thus, it is conceivable that the auxiliary proteins of pLS20 fulfill similar function(s).

Most of our knowledge on auxiliary proteins is related to those encoded by conjugative plasmids replicating in G− bacteria; in particular, the auxiliary proteins of F and related plasmids have been studied in detail at the functional, biochemical and structural levels (for review see, Alvarez-Martinez and Christie, 2009; De la Cruz et al., 2010; Wong et al., 2012). Upon binding, TraY and TraM of plasmid F bent the DNA and therewith play important roles in organizing the relaxosome complex at oriT and influencing the nicking reaction of the relaxase. In addition, they both play a role in gene expression by regulating the activity of their own promoters. TraM also has a key role in delivering the relaxosome to the conjugative pore by interacting with its cognate T4CP (Wong et al., 2011; Peng et al., 2014). Future studies are needed to determine whether the auxiliary proteins of pLS20 fulfill similar functions to those of F, although it is doubtful that Aux1LS20 and Aux2LS20 play a role in gene regulation due to the different genetic organization. In the case of F, the monocistronic traM gene is located directly downstream of its oriT. TraM is followed by another monocistronic gene, traI, which in turn is followed by a large multicistronic operon in which traY is the first gene (Zatyka and Thomas, 1998). In the case of pLS20, though, the relaxosome genes are embedded within the large conjugation operon and are under the control of the main conjugation promoter Pc, that is located almost 26 kbp upstream of aux1LS20 (Singh et al., 2013; Ramachandran et al., 2014). At present, we cannot fully exclude the possibility that the relaxosome genes of pLS20cat are controlled by an additional promoter that is regulated by Aux1LS20 or Aux2LS20. RNAseq data showed, however, that repression of the main conjugation promoter results in silencing of the relaxosome genes, as well as other genes in the conjugation operon of pLS20cat (Singh et al., 2013).

Far less is known about auxiliary proteins encoded by conjugative plasmids of Gram+ origin. The monomeric Helix-Turn-Helix protein TraN of the Enterococcus faecalis conjugative plasmid pIP501 binds to its oriT region, which suggested that it might be an auxiliary protein of pIP501. However, recent results revealed that traN is not essential for conjugation, and it is now believed that it may be a repressor of conjugation by regulating either the expression of the conjugation operon or activity of the relaxase TraA (Goessweiner-Mohr et al., 2014; Grohmann et al., 2016).
The auxiliary proteins PcfF encoded by the *Enterococcus faecalis* plasmid pCF10, and LtrF of *Lactococcus lactis* plasmid pRS01 share 47% sequence identity. As far as we know, these are the only auxiliary proteins encoded by conjugative plasmids of Gram+ origin that have been studied in some detail (Chen et al., 2007, 2008). The *pcfF* and *ltrF* genes are essential for conjugation and purified PcfF and LtrF bind their cognate oriTs. Moreover, evidence supports a model in which PcfF recruits the relaxase PcfG to oriT, and that PcfF, probably in conjunction with the relaxase PcfG, interacts with its cognate T4CP and hence plays an important role in delivering the relaxosome to the conjugative pore.

Several auxiliary proteins of conjugative plasmids of Gram− origin are described to contain a RHH motif. These include, TraY and TraM of F plasmid, TraW of R388, VirC2 of *Agrobacterium tumefaciens*, NikA of R64, TraJ of RP4, MobC of RSF1010, MbeC of ColE1, MobC of RA3 (Bowie and Sauer, 1990; Zhang and Meyer, 1997; Moncalian and De la Cruz, 2004; Ragonese et al., 2007; Yoshida et al., 2008; Lu et al., 2009; Varsaki et al., 2009; Godziszewska et al., 2016). For some of them structure-based mutational analyses have demonstrated the importance of the RHH motif in oriT binding as well as relaxase recruitment (Yoshida et al., 2008; Varsaki et al., 2009; Godziszewska et al., 2016). Interestingly, Aux1LS20 and Aux2LS20 are also predicted to contain an RHH DNA-binding domain in their N-terminal region (Supplementary Figure S1). In addition, our *in silico* analyses predict that the auxiliary PcfF and LtrF proteins of Gram+ *E. faecalis* pCF10 and *L. lactis* pRS01 plasmids, respectively, also contain an RHH motif in their N-terminal region (our unpublished results). The presence of a likely RHH motif in Aux1LS20 and Aux2LS20 is therefore in line with the conclusion that they are auxiliary proteins. More importantly, the observation that the auxiliary proteins encoded by plasmids pLS20, pRS01, and pCF10, replicating in Gram+ bacteria, all contain a predicted RHH motif indicates that this is a conserved motif in auxiliary proteins encoded by CE+ origin that have been studied in some detail. This strategy also supported the general importance of the RHH motif in the binding of the T4CPs and the recruitment of the relaxase. The auxiliary proteins PcfF and LtrF of *Gram+* origin that are essential for conjugation, and that they form the founding members of families of auxiliary relaxosome proteins that are encoded in Firmicutes bacteria.

### AUTHOR CONTRIBUTIONS

All authors listed have made substantial, direct experimental and/or intellectual contribution to the work. AM-A, J-AH, GR, CG-C, DG-A, and JV-C generated all plasmids and strains, purified proteins and executed all the experiments except the ultracentrifugation studies, which were performed by JL-O and CA. DA performed *in silico* analyses contributed to the general design and analyses of the results. LW and WM designed the experimental plan and were principally responsible for analyzing the results and writing the paper. WM supervised AM-A, J-AH, GR, CG-C, DG-A and JV-C.

### FUNDING

Work in the Meijer lab was funded by grants Bio2013-41489-P and BIO2016-77883-C2-1-P of the Ministry of Economy and Competitiveness of the Spanish Government to WM, which also funded AM-A, CG-C, and JV-C. Part of the economic support of the two aforementioned grants was provided by the “Agencia Estatal de Investigación (AEI)” and “Fondo Europeo de Desarrollo Regional (FEDER).” This research was also supported by institutional grants from the “Fundación Ramón Areces” and “Banco de Santander” to the Centro de Biología Molecular “Severo Ochoa”. LW’s work was supported by Wellcome Trust grant WT098374AIA to Jeff Errington. JL-O and CA were supported by grant BFU2014-52070-C2-2-P of the Ministry of Science and Innovation to CA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. J-AH received a State Scholarship Fund from the China Scholarship Council.

### ACKNOWLEDGMENTS

We thank Jose Belio for help with preparing the Figures, and Margarita Salas and Jeff Errington for their support on our work. We also want to acknowledge helpful discussion with other lab members, and particularly want to mention Praveen K. Singh.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2017.02138/full#supplementary-material
REFERENCES

Altenschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Altenschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schäffer, A. A., et al. (2005). Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* 272, 5101–5109. doi: 10.1111/j.1742-4658.2005.04945.x

Alvarez-Martinez, C. E., and Christie, P. J. (2009). Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. doi: 10.1128/MMBR.00023-09

Boto, L. (2010). Horizontal gene transfer in evolution: facts and challenges. *Proc. Biol. Sci.* 277, 819–827. doi: 10.1098/rspb.2009.1679

Bowie, J. U., and Sauer, R. T. (1990). TraY proteins of F and related episomes are members of the Arc and Mnt represer family. *J. Mol. Biol.* 211, 5–6.

Bron, S., Peijnenburg, A., Peeters, B. P. H., Haima, P., and Venema, G. (1989). “Cloning and plasmid (in)stability in *Bacillus subtilis*,” in *Genetic Transformation and Expression*, eds L. O. Butler, C. R. Harwood, and B. E. B. Moseley (Andover: Intercept Ltd.), 205–219.

Chen, Y., Staddon, J. H., and Dunny, G. M. (2007). Specificity determinants of conjugal DNA processing in the *Enterococcus faecalis* plasmid pCF10 and the Lactococcus lactis plasmid pS01. *Mol. Microbiol.* 63, 1549–1564.

Chen, Y., Zhang, X., Manias, D., Yeo, H. J., Dunny, G. M., and Christie, P. J. (2008). *Enterococcus faecalis* PecC, a spatially localized substrate receptor for type IV secretion of the pCF10 transfer intermediate. *J. Bacteriol.* 190, 3632–3645. doi: 10.1128/JB.00199-07

Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* 74, 417–433. doi: 10.1128/MMBR.00016-10

De la Cruz, F., Frost, L. S., Meyer, R. J., and Zechner, E. L. (2010). Conjugative DNA metabolism in Gram-negative bacteria. *FEBS Rev.* 53, 18–40. doi: 10.1111/j.1742-4658.2009.01915.x

Del Solar, G., Hernandez-Arriaga, A. M., Gomis-Ruth, F. X., Coll, M., and Espinosa, M. (2002). A genetically economical family of plasmid-encoded transcriptional repressors involved in control of plasmid copy number. *J. Bacteriol.* 184, 4943–4951.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Fogg, J. M., Randall, G. L., Pettit, B. M., Summers de, W. L., Harris, S. A., and Moncalian, G. (2004). DNA binding properties of protein TrwA, and its N-terminal relaxase domain of the Gram-positive plasmid pIP501 show specific oriT binding and behave as dimers in solution. *Biochem. J.* 387, 401–409.

Lee, C. A., and Grossman, A. D. (2007). Identification of the origin of transfer (oriT) and DNA relaxase required for conjugation of the integrative and conjugative element ICEBs1 of *Bacillus subtilis*. *J. Bacteriol.* 189, 7254–7261.

Liu, J., den Dulk-Ras, A., Hooykaas, P. J., and Glover, J. N. (2009). *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9643–9648. doi: 10.1073/pnas.0812199106

Mazel, D., and Davies, J. (1999). Antibiotic resistance in microbes. *Cell Mol. Life Sci.* 56, 742–754.

Moncalian, G., and De la Cruz, F. (2004). DNA binding properties of protein TraW, a possible structural variant of the Arc repressor superfamily. *Biochim. Biophys. Acta* 1701, 15–23.

Norman, A., Hansen, L. H., and Sørensen, S. J. (2009). Conjugative plasmids: vessels of the communal gene pool. *Philos. Trans. R. Soc. Lond B Biol. Sci.* 364, 2275–2289. doi: 10.1098/rstb.2009.0037

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304.

Penders, J., Stobberingh, E. E., Savelkoul, P. H., and Wolfs, P. F. (2013). The human microbiome as a reservoir of antimicrobial resistance. *Front. Microbiol.* 4:387. doi: 10.3389/fmicb.2013.00387

Peng, Y., Lu, I., Wong, J. J., Edwards, R. A., Frost, L. S., and Mark Glover, J. N. (2014). Mechanistic basis of plasmid-specific DNA binding of the F plasmid regulatory protein, TraM. *J. Mol. Biol.* 426, 3783–3795. doi: 10.1016/j.jmb.2014.09.018

Pluta, R., Boer, D. R., Lorenzo-Díaz, F., Russi, S., Gomez, H., Fernandez-Lopez, C., et al. (2017). Structural basis of a histidine-DNA nicking/joining mechanism for gene transfer and promiscuous spread of antibiotic resistance. *Proc. Natl. Acad. Sci. U.S.A.* 114, E6526–E6535. doi: 10.1073/pnas.1702971114

R Core Team (2017). *R: A Language, and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ragonese, H., Haisch, D., Villareal, E., Choi, J. H., and Matson, S. W. (2007). The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT during an interaction with TraM. *Mol. Microbiol.* 63, 1173–1184.

Ramachandran, G., Miguel-Arribas, A., Abia, D., Singh, P. K., Crespo, L., Gago-Cordoba, C., et al. (2017). Discovery of a new family of relaxases in Fimicibacteria. *PLoS Genet.* 13:e1006586. doi: 10.1371/journal.pgen.1006586

Ramachandran, G., Singh, P. K., Luque-Ortega, J. R., Yuste, L., Alfonso, C., Rojo, F., et al. (2014). A complex genetic switch involving overlapping divergent promoters and DNA looping regulates expression of conjugation genes of a gram-positive plasmid. *PLoS Genet.* 10:e1004733. doi: 10.1371/journal.pgen.1004733

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schäffer, A. A., Aravind, L., Madden, T. L., Shavirin, S., Spouge, J. L., Wolf, Y. I., et al. (2001). Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res.* 29, 2994–3005.

Schilbach, J. F., Karazi, A. W., Baumann, B. E., and Sauer, R. T. (1999). Origins of DNA-binding specificity: role of protein contacts with the DNA backbone. *Proc. Natl. Acad. Sci. U.S.A.* 96, 811–817.

Singh, P. K., Ramachandran, G., Duran-Alcalde, L., Alonso, C., Wu, L. J., and Meijer, W. J. (2012). Inhibition of *Bacillus subtilis* natural competence by a native, conjugative plasmid-encoded comK repressor protein. *Environ. Microbiol.* 14, 2812–2825. doi: 10.1111/j.1462-2920.2012.02819.x

Singh, P. K., Ramachandran, G., Ramos-Ruíz, R., Peiro-Pastor, R., Abia, D., Wu, L. J., et al. (2013). Mobility of the native *Bacillus subtilis* conjugative plasmid pLS20 is regulated by intercellular signaling. *PLoS Genet.* 9:e1003892. doi: 10.1371/journal.pgen.1003892

Smillie, C., Garcia-Barea, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452.
Sommer, M. O., Church, G. M., and Dantas, G. (2010). The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* 1, 299–303. doi: 10.4161/viru.1.4.12010

Sommer, M. O., Dantas, G., and Church, G. M. (2009). Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* 325, 1128–1131. doi: 10.1126/science.1176950

Tanaka, T., Kuroda, M., and Sakaguchi, K. (1977). Isolation and characterization of four plasmids from *Bacillus subtilis*. *J. Bacteriol.* 129, 1487–1494.

Thoma, L., and Muth, G. (2016). Conjugative DNA-transfer in *Streptomyces*, a mycelial organism. *Plasmid* 87–88, 1–9. doi: 10.1016/j.plasmid.2016.09.004

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721.

Varsaki, A., Moncalian, G., Garcillan-Barcia, M. P., Drainas, C., and De la Cruz, F. (2009). Analysis of ColE1 MbeC unveils an extended ribbon-helix-helix family of nicking accessory proteins. *J. Bacteriol.* 191, 1446–1455. doi: 10.1128/JB.01342-08

Wang, S., Li, W., Liu, S., and Xu, J. (2016). RaptorX-Property: a web server for protein structure property prediction. *Nucleic Acids Res.* 44, W430–W435. doi: 10.1093/nar/gkw306

Waters, V. L. (1999). Conjugative transfer in the dissemination of beta-lactam and aminoglycoside resistance. *Front. Biosci.* 4:D433–D456.

Wong, J. J., Lu, J., Edwards, R. A., Frost, L. S., and Glover, J. N. (2011). Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res.* 39, 6773–6788. doi: 10.1093/nar/gkz296

Wong, J. J., Lu, J., and Glover, J. N. (2012). Relaxosome function and conjugation regulation in F-like plasmids - a structural biology perspective. *Mol. Microbiol.* 85, 602–617. doi: 10.1111/j.1365-2958.2012.08131.x

Yoshida, H., Furuya, N., Lin, Y. J., Guntert, P., Komano, T., and Kainosho, M. (2008). Structural basis of the role of the NikA ribbon-helix-helix domain in initiating bacterial conjugation. *J. Mol. Biol.* 384, 690–701. doi: 10.1016/j.jmb.2008.09.067

Zatyka, M., and Thomas, C. M. (1998). Control of genes for conjugative transfer of plasmids and other mobile elements. *FEBS Microbiol. Rev.* 21, 291–319.

Zhang, S., and Meyer, R. (1997). The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. *Mol. Microbiol.* 25, 509–516.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors JL-O and CA.

*Copyright © 2017 Miguel-Arribas, Hao, Luque-Ortega, Ramachandran, Val-Calvo, Gago-Córdoba, González-Álvarez, Abia, Alfonso, Wu and Meijer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

---

**Conflict of Interest Statement:**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors JL-O and CA.

**Copyright © 2017 Miguel-Arribas, Hao, Luque-Ortega, Ramachandran, Val-Calvo, Gago-Córdoba, González-Álvarez, Abia, Alfonso, Wu and Meijer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**

---

**Conflict of Interest Statement:**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors JL-O and CA.

**Copyright © 2017 Miguel-Arribas, Hao, Luque-Ortega, Ramachandran, Val-Calvo, Gago-Córdoba, González-Álvarez, Abia, Alfonso, Wu and Meijer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**

---

**Conflict of Interest Statement:**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors JL-O and CA.

**Copyright © 2017 Miguel-Arribas, Hao, Luque-Ortega, Ramachandran, Val-Calvo, Gago-Córdoba, González-Álvarez, Abia, Alfonso, Wu and Meijer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**

---

**Conflict of Interest Statement:**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors JL-O and CA.

**Copyright © 2017 Miguel-Arribas, Hao, Luque-Ortega, Ramachandran, Val-Calvo, Gago-Córdoba, González-Álvarez, Abia, Alfonso, Wu and Meijer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**