# Cosmid based mutagenesis causes genetic instability in *Streptomyces coelicolor*, as shown by targeting of the lipoprotein signal peptidase gene

John T Munnoch[1], David A. Widdick[1,2], Govind Chandra[2], Iain C. Sutcliffe[3], Tracy Palmer[4] and Matthew I Hutchings[1]*

[1]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich

[2]Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich

[3] School of Applied Sciences, Northumbria University, Newcastle

[4] Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee.

*** Correspondence: M.hutchings@uea.ac.uk**

# Abstract.

Bacterial lipoproteins are a class of extracellular proteins tethered to cell membranes by covalently attached lipids. Deleting the lipoprotein signal peptidase (*lsp*) gene in *Streptomyces coelicolor* results in growth and developmental defects that cannot be restored by reintroducing the *lsp*. We report resequencing of the genomes of the wild-type M145 and the *cis*-complemented Δ*lsp* mutant (BJT1004), mapping and identifying secondary mutations, including an insertion into a novel putative small RNA, *scr6809*. Disruption of *scr6809* led to a range of developmental phenotypes. However, these secondary mutations do not increase the efficiency of disrupting *lsp* suggesting they are not *lsp* specific suppressors. Instead we suggest that these were induced by introducing the cosmid St4A10Δ*lsp* as part of the Redirect mutagenesis protocol, which transiently duplicates a number of important cell division genes. Disruption of *lsp* using no gene duplication resulted in the previously observed phenotype. We conclude that *lsp* is not essential in *S. coelicolor* but loss of *lsp* does lead to developmental defects due to the loss of lipoproteins from the cell. Significantly, our results indicate the use of cosmid libraries for the genetic manipulation of bacteria can lead to unexpected phenotypes not necessarily linked to the gene or pathway of interest.

# Introduction

Bacterial lipoproteins are essential for building and maintaining the cell envelope and also provide a key interface with the external environment [1–3] Most lipoprotein precursors are exported as unfolded polypeptides via the Sec (general secretory) pathway but others can be exported via the twin arginine transport (Tat) pathway, which is typically utilised for the transport of fully folded proteins [4–6] The signal peptides of lipoproteins closely resemble other types of bacterial Sec and Tat signal peptide but they contain a characteristic lipobox motif, typically $L_{-3}$-A/$S_{-2}$-G/$A_{-1}$-$C_{+1}$, relative to the signal cleavage site, in which the cysteine residue is essential and invariant. The lipobox motif allows putative lipoproteins to be easily identified in bacterial genome sequences [3,7].

Following translocation, lipoprotein precursors are firstly modified by covalent attachment of a diacylglycerol molecule, derived from a membrane phospholipid, to the thiol of the conserved lipobox cysteine residue via a thioether linkage. This reaction is catalysed by an enzyme named Lgt (Lipoprotein diacylglycerol transferase) and results in a diacylated lipoprotein. Lsp (Lipoprotein signal peptidase) then cleaves the signal sequence immediately upstream of the lipidated cysteine to leave it at the +1 position. These early steps in lipoprotein biogenesis are highly conserved and unique to bacteria making them potential targets for antibacterial drug development [2,8]. In Gram-negative bacteria and Gram-positive Actinobacteria, lipoproteins can be further modified by addition of an amide-linked fatty acid to the amino group of the diacylated cysteine residue at the mature N-terminus. This final step is catalysed by the enzyme Lnt (Lipoprotein n-acyltransferase) and results in triacylated lipoproteins. In Gram-negative proteobacteria, Lnt modification is a pre-requisite for the recognition of lipoproteins by the Lol machinery, which transports lipoproteins to the outer membrane [2,9] but its function in monoderm Gram-positive bacteria is not known. Members of the Gram-positive phyla Firmicutes and Mollicutes also N-acylate lipoproteins despite lacking Lnt homologues and *S. aureus* can diacylate or triacylate individual lipoproteins in an environmentally dependent manner [10–14]. These studies suggest that triacylation of lipoproteins in Gram-positive bacteria has an important role in their natural environment but is dispensable *in vitro*. Loss of Lnt activity in *Streptomyces* bacteria has no obvious effect on fitness or lipoprotein localisation *in vitro* but it does have a moderate effect on virulence in the plant pathogen *Streptomyces scabies*, supporting the idea that it has environmental importance [15].

We previously characterised all four steps of the lipoprotein biogenesis pathway in *Streptomyces* spp. (Figure 1) [5,15], which is one of the best studied genera in the Gram-positive phylum Actinobacteria. Our key

findings are (i) that Tat exports ~20% of lipoprotein precursors in streptomycetes; (ii) they N-acylate lipoproteins using two non-essential Lnt enzymes; (iii) *Streptomyces coelicolor* encodes two functional copies of Lgt which cannot be removed in the same strain; (iv) *lsp* mutants can be isolated at low frequencies but they acquire spontaneous secondary mutations which might be *lsp* suppressors. It was recently reported that Lgt is essential in *Mycobacterium tuberculosis*, which is also a member of the phylum Actinobacteria, and that *lgt* deletion in the fast-growing species *Mycobacterium smegmatis* is accompanied by spontaneous secondary mutations [16]. Natural product antibiotics that target the lipoprotein biogenesis pathway include globomycin, made by *Streptomyces globisporus* [2] and antibiotic TA made by *Myxococcus xanthus* [1,16]. Both inhibit Lsp activity and are lethal to *Escherichia coli* but TA resistance arises through spontaneous *IS3* insertion into the *lpp* gene, which encodes an abundant lipoprotein that attaches the *E. coli* outer membrane to the peptidoglycan cell wall [16,17]. Over-expressing *lsp* also confers TA resistance in both *E. coli* and *M. xanthus*, and the latter encodes additional Lsp homologues within the TA biosynthetic gene cluster [17].

Deletion of *S. coelicolor lsp* results in very small and flat colonies that are delayed in sporulation and these *lsp* mutants could not be fully complemented even by reintroducing the *lsp* gene to its native locus. Although both *cis* and *in trans* complementation restored lipoprotein biogenesis and sporulation it did not restore the wild-type growth rate [5]. There are two likely reasons for this: either *lsp* is essential and the mutant strains acquire secondary suppressor mutations, or the Redirect PCR targeting method that we used to delete the *lsp* gene resulted in chromosomal rearrangements and mutations independent of *lsp*. Here we provide evidence to support the second hypothesis and we demonstrate that introduction of the cosmid carrying an ~40 kb region of the *S. coelicolor* chromosome, including *lsp*, from *E. coli* to *S. coelicolor* transiently duplicates cell division and cell wall biosynthesis genes which leads to secondary mutations including disruption of a putative small RNA. We further confirm that *lsp* is non-essential but deletion of the *lsp* gene does lead to growth and developmental delays and the over-production of the antibiotic actinorhodin in *S. coelicolor*, as observed previously. These phenotypes must therefore be due to the loss of lipoproteins from its cytoplasmic membrane.

# Results

## Mapping secondary mutations in the *cis* complemented *Δlsp* strain BJT1004

We previously reported that the *S. coelicolor* Δ*lsp* mutant BJT1001 cannot be complemented even by restoring *lsp* to its native locus [5]. Since *cis* complementation should effectively restore the genome to wild-type this suggests that other spontaneous mutations have occurred during the genetic manipulations. To test this we Illumina sequenced the genomes of the parent strain *S. coelicolor* M145 and the *cis*-complemented Δ*lsp* strain BJT1004 using two independent companies (GATC Biotech and The Genome Analysis Centre). Across the four sequence samples a total of 51 unique single nucleotide polymorphisms (SNPs) were detected (Supplementary Table S1 online) as well as a chromosomal rearrangement in BJT1004 that is not present in the parent strain M145 (Figure 2(A-B) and Supplementary Text S1 and S2 online). Of the 51 SNPS, 13 are unique to one of the BJT1004 sequences, with 4 residing inside coding regions. However of all of these only one SNP occurs in both BJT1004 sequences and this is in the intergenic region between *sco5331* and *sco5332*. In the single chromosomal rearrangement, the *IS21* insertion element (genes *sco6393* and *sco6394*) has inserted into the intergenic region between the *sco6808* and *sco6809* genes and this was confirmed by PCR (Figure 2(A-C)). Although this might affect the downstream promoter of *sco6808*, which encodes a putative regulator, deletion of *sco6808* (using vector pJM010) had no effect on growth or development under standard laboratory conditions (Figure 3(A)). The intergenic position of *IS21* in BJT1004 suggested it might disrupt a non-coding RNA and examination of RNA sequence data for *S. coelicolor* M145 confirmed the presence of a 189 nt transcript initiating 107 bp upstream of *sco6808* and reading into the last 82 nucleotides of the *sco6809* gene (data from GSM1121652 and GSM1121655 RNA sequencing; Supplementary Text S1 and 2 online; Figure 2(A-B)). Following convention we named this putative small RNA *scr6809* for *S. coelicolor* RNA 6809. Deletion of the *scr6809* (pJM012) sequence (without disrupting either the *sco6808* or *sco6809* coding sequences) resulted in a range of phenotypes from colonies that look like wild-type to non-sporulating bald and white mutants defective in aerial hyphae formation and sporulation, respectively, antibiotic overproducers and small slow growing colonies. Restreaking the Δ*scr6809* colonies (double crossovers) with wild-type appearance again gave rise within the next generation to a range of colonies with different morphologies, including growth and developmental defects (Figure 3(B)). Colonies with mutated morphologies would retain that morphology in

subsequent generations indicating another situation where spontaneous secondary mutations are arising. A previous report showed that a *sco6808* deletion mutant had accelerated production of actinorhodin and undecylprodigiosin as well as precocious spore formation on R5 medium [18]. There was no observable difference between the wild-type and Δ*sco6808* strains under the growth conditions used here but disruption of *sco6808* in strain BJT1004 resulted in an improvement in sporulation (Figure 3(A)). We suspect this difference is based on the recovery of the *scr6809* loci to wildtype as result of the double crossovers between the chromosome of BJT1004 and the Δ*sco6808* deletion cosmid St1A2Δ*sco6808*. This was also seen for St1A2Δ*sco6811* disruptions in each background (not shown).

To determine whether *IS21* insertion into *scr6809* is induced by deletion of *lsp*, we isolated ten more non-clonal *lsp* mutants by introducing cosmid St4A10Δ*lsp* (pJM014) into wild-type strain M145 and then PCR amplified the intergenic region between *sco6808* and *sco6809*. The size of the PCR products matched the predicted wild-type size and indicated that none of these *lsp* mutants contain an *IS21* insertion suggesting that the original observation is not specific to *lsp* mutants (Figure 4). Consistent with this conclusion, the frequency with which *lsp* mutants could be isolated was not increased in BJT1004 relative to M145 suggesting that none of the mapped mutations in BJT1004 suppress fitness defects that arise from deleting Δ*lsp*. Attempts at over-expressing *scr6809* using pJM017 in *S. coelicolor* M145, *S. scabies* 87-22 and *S. venezualae* ATCC 10712 resulted in no observable phenotype but as the same vectors failed to prevent accumulation of developmental phenotypes in the Δ*scr6809* strain, this suggests that functional *scr6809* is not expressed from these vectors.

Cumulatively these results suggested that deletion of *lsp* does not result in secondary mutations and prompted us to hypothesise that these accumulate as a result of duplicating cell division genes on cosmid St4A10 which was used to delete *lsp*. These results further suggest a role for *scr6809* in *S. coelicolor* differentiation, although there is no obvious link to *lsp* and so this was not pursued further here.

# Introduction of wild-type St4A10 results in a pleiotropic phenotype

The Redirect PCR-targeting method uses *E. coli* as a host strain for an *S. coelicolor* cosmid library which can be used to make targeted deletions [19,20]. The Redirect method was used to PCR-target the *lsp* gene *sco2074* on cosmid St4A10, which contains a ~40 kb region of the *S. coelicolor* genome spanning genes *sco2069-2104* (Supplementary Table S2 online). Conjugation of St4A10Δ*lsp* into *S. coelicolor* transiently duplicates all the

genes on that cosmid (except *lsp*) and because this region includes many important cell division genes (*ftsZ*, *ftsQ*, *ftsW*, *ftsI* and *ftsL*) and essential cell wall synthesis genes (*murG*, *murD*, *murX*, *murF* and *murE*) we reasoned that over-expression of these genes, rather than deletion of *lsp*, is responsible for the spontaneous secondary mutations and the resulting pleiotropic phenotype. To test this idea we introduced an origin of transfer into the wild-type St4A10 cosmid backbone and then conjugated this cosmid into strain M145 and selected for single cross-over events where the whole cosmid is integrated into the chromosome, thus duplicating the *S. coelicolor* genes on St4A10. Analysis of these single crossover strains, maintained on kanamycin to select for the cosmid, revealed them to be genetically unstable, with many initially appearing similar to the observed Δ*lsp* phenotype, i.e. small and delayed in sporulation (Figure 5). However, they lack the characteristic Δ*lsp* over-production of the blue antibiotic actinorhodin and colonies from this M145::St4A10 strain also acquired more significant developmental issues upon prolonged maintenance and restreaking on MS agar containing kanamycin (not shown). This suggests they accumulate spontaneous secondary mutations as a direct result of carrying St4A10 and that the observed Δ*lsp* phenotype is at least in part due to duplication of the genes on cosmid St4A10. This is consistent with the fact that complementation of Δ*lsp* restored lipoprotein biogenesis but did not restore wild-type colony morphology [5].

## Targeted deletion of *lsp* results in a small colony phenotype

To test how much deletion of *lsp* contributed to the phenotype of BJT1001 (the Δ*lsp* strain generated using Redirect) we undertook a targeted disruption of *lsp* in wild-type strain M145 using a suicide vector, which does not duplicate or affect any other coding sequences. The *lsp* suicide vector, pJM016 (Table 1), was introduced into wild-type *S. coelicolor* by conjugation and ex-conjugants were selected by growing on MS agar plates containing apramycin. Following introduction of the pJM016, two colony types were observed (Figure 6), one with wild-type appearance while the others were small colonies that over-produce actinorhodin, reminiscent of the *lsp* mutant BJT1001. PCR testing of the genomic DNA of both morphotypes revealed that those with the wild-type colony morphology have a wild-type fully functioning *lsp* gene whereas those with a small colony phenotype have disruptions in *lsp* caused by pJM016. PCR amplification followed by sequencing of the loci in the small colony variant revealed an interesting and unexpected recombination event had occurred: the vector and almost all the *lsp* gene have been removed such that all that remains is the apramycin resistance cassette (Supplementary Text S4 and S5 online). These data confirm that *lsp* is not essential in *S. coelicolor* but loss of

168 Lsp does result in a growth and developmental defect and overproduction of the blue antibiotic actinorhodin as
169 observed previously [5].

# 170 Discussion

171 The pleiotropic nature of the original *S. coelicolor* Δ*lsp* strain BJT1001 resulted primarily from the introduction
172 of cosmid St4A10, most likely caused by the over-expression of cell division and cell wall biosynthesis genes
173 carried on that cosmid (Supplementary Table S2 online). It seems likely, but is not proven, that this led to the
174 secondary mutations we observed in this strain. These secondary mutations do not make it easier to delete *lsp*
175 suggesting they are not *lsp*-specific suppressors. Genetic manipulation has always been challenging in
176 *Streptomyces* bacteria and the Redirect PCR targeting method has been a significant development but this work
177 should be a cautionary tale to others to consider the effects of using large insert cosmid libraries in the genetic
178 manipulation of bacteria. Recent advances in CRISPR/Cas9 editing of *Streptomyces* genomes [21] negate the need
179 for a cosmid library and these techniques will accelerate research into the basic biology of *Streptomyces* and
180 other filamentous actinomycetes. This is vital because the secondary metabolites derived from these bacteria still
181 represent a major underutilised reservoir from which new antibiotics and other bioactive natural products can be
182 discovered. Moreover, the identification here of the novel small RNA *scr6809* and demonstration that its
183 deletion results in a range of growth and developmental defects add to the growing appreciation [22–24] of the
184 significance of small RNAs in streptomycetes.

# 185 Materials and Methods

## 186 Bacterial strains and culture conditions.

187 All primers, plasmids and strains used are listed in Table 1. Strains were routinely grown as previously
188 described [5] following the recipes of Kieser et al., (2000). *E. coli* was grown in LB or LB –NaCl for Hygromycin
189 selection and *S. coelicolor* M145 and its derivatives were grown on Soya Flour Mannitol (SFM) medium to
190 study growth and development or LB culture for genomic isolations.

191 **Table 1. Strains, plasmids and primers.**

Strain

| E. coli | Genotype/description | Plasmid (held) | Resistance | Source |
|---|---|---|---|---|
| TOP10 | F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG | - | - | Invitrogen |
| BW25113 | F-, DE(araD-araB)567, lacZ4787(del)::rrnB-3, LAM-, rph-1, DE(rhaD-rhaB)568, hsdR514 | pIJ790 | CmR | Datsenko & Wanner, 2000 |
| ET12567 | dam- dcm- hsdM- | pUZ8002 | CmR/TetR | MacNeil et al., 1992 |

| Streptomyces | Genotype/description | Plasmid (used) | Resistance | Source |
|---|---|---|---|---|
| M145 | S. coelicolor wild type strain, SCP1-, SCP2- | - | - | Hopwood et al, 1985 |
| BJT1000 | M145 *lsp::apr* | - | AprR | Thompson et al., 2010 |
| BJT1001 | M145 lspFLP | - | - | Thompson et al., 2010 |
| BJT1004 | BJT1000 + Sco *lsp* cis | - | - | Thompson et al., 2010 |
| JTM005 | M145 *sco6808::apr* | pJM010 | AprR | This work |
| JTM007 | M145 *scr6809::apr* | pJM012 | AprR | This work |
| JTM008 | M145 *lsp::apr* | pJM013 | AprR | This work |
| JTM009 | M145 St4A10 *bla::hyg* | pJM014 | KanR/HygR | This work |
| JTM012 | BJT1004 *sco6808::apr* | pJM010 | AprR | This work |

| | | | | |
|---|---|---|---|---|
| JTM015 | BJT1004 *lsp::apr* | pJM013 | AprR | This work |
| JTM018 | M145 *lsp* suicide vector | pJM016 | AmpR/AprR | This work |

192

| Plasmids | Genotype/description | Resistance | Source |
|---|---|---|---|
| pIJ773 | *aac(3)IV oriT* (contains apramycin (apr) resistance cassette) | AprR | Gust et al., 2002 |
| pIJ10700 | contains hygromycin resistance cassette, FRT *oriT-hyg* FRT MkII | HygR | Gust et al., 2002 |
| pIJ790 | *araC-Parab, Y, β, exo, cat, repA1001ts, oriR101* | CmR | Gust et al., 2002 |
| pUZ8002 | RK2 derivative with a mutation in *oriT* | | Kieser et al., 2000 |
| pMS82 | *ori*, pUC18, *hyg, oriT, RK2, int ΦBT1* | HygR | Gregory et al., 2003 |
| pIJ10257 | *oriT, ΦBT1 attB-int, Hygr, ermEp**, pMS81 backbone | HygR | Hong et al., 2005 |
| pGEM-T-Eazy | *bla, lacZα* | AmpR | Promega |
| St1A2 | Supercos-1-cosmid with (39829 bp) fragment containing (*sco6808* and *scr6809*) | KanR/AmpR | Redenbach et al., 1996 |
| St4A10 | Supercos-1-cosmid with (43147 bp) fragment containing (*sco2074 - lsp*) | KanR/AmpR | Redenbach et al., 1996 |
| pJM010 | St1A2 containing *sco6808::apr oriT* (St1A2Δ*sco6808*) | KanR/AmpR/AprR | This work |
| pJM012 | St1A2 containing *scr6809::apr oriT* (St1A2Δ*scr6809*) | KanR/AmpR/AprR | This work |
| pJM013 | St4A10 containing sco2074::apr oriT (St4A10Δ*lsp*) | KanR/AmpR/AprR | This work |
| pJM014 | St4A10 containing *bla::hyg oriT* (St1A2*bla::hyg*) | KanR/HygR | This work |
| pJM015 | pMS82 containing full length *sco2074* and promoter (300 bp of upstream DNA) | HygR | This work |
| pJM016 | *lsp* suicide vector, pGEM-T-Eazy, 411 bp fragment of the *lsp* gene with a BamHI site. The *aac(3)IV* | AmpR/AprR | This work |

| | | | |
|---|---|---|---|
| | containing BamHI fragment from a pIJ773 was sub cloned in. | | |
| pJM017 | pMS82, KpnI/HindIII insert containing, pMC500 MCS and terminators with *scr6809 (see Text S3)* | HygR | This work |

193

| Primer | Sequence | Description | Source |
|---|---|---|---|
| JM0083 | GTCTATGGTTGACGGGTGACTGTCATAGAT CTGCAGATGATTCCGGGGATCCGTCGACC | *sco6808* forward disruption primer (Redirect) | This work |
| JM0084 | GTCATCTTCCGAACGGAGATGGAGGGAGAT CCGGAATCATGTAGGCTGGAGCTGCTTC | *sco6808* reverse disruption primer (Redirect) | This work |
| JM0085 | CGGAGGCCGCCTGTCCTAGC | *sco6808* forward test primer | This work |
| JM0086 | AACGCGCACTCGCTGCGGTC | *sco6808* reverse test primer | This work |
| JM0091 | TCCGACATCTGCAGATCTATGACAGTCACC CGTCAACCAATTCCGGGGATCCGTCGACC | *scr6809* forward disruption primer (Redirect) | This work |
| JM0092 | TGGTACACGGCACCGACTCCGGCTGCCAGA AAGCCATAGTGTAGGCTGGAGCTGCTTC | *scr6809* reverse disruption primer (Redirect) | This work |
| JM0093 | CAGACGCAGGCCTCGCCATC | *scr6809* forward test primer | This work |
| JM0094 | CCCATCGCTACGGCCGCCT | *scr6809* reverse test primer | This work |
| JM0093 | AATCAATCTAAAGTATATATGAGTAAACTT GGTCTGACAGTCAGGCGCCGGGGGCGGTG | *bla (bla::hyg)* gene forward disruption primer (Redirect) for supercos-1 | This work |
| JM0096 | CCCTGATAAATGCTTCAATAATATTGAAAA | *bla (bla::hyg)* gene forward disruption primer (Redirect) for | This work |

| | | | |
|---|---|---|---|
| | AGGAAGAGTAAGTTCCCGCCAGCCTCGCA | supercos-1 | |
| JM0097 | AAGCAGCAGATTACGCGCAG | *bla (bla::hyg)* gene forward test primer (Redirect) for supercos-1 | This work |
| JM0098 | GTGCGCGGAACCCCTATTTG | *bla (bla::hyg)* gene forward test primer (Redirect) for supercos-1 | This work |
| JM0099 | TCGTGCTCAGTCAAGGACCTAGGCTGAGGG | *lsp* (*sco2074*) forward disruption primer (Redirect) | This work |
| | ACTCACGTGATTCCGGGGATCCGTCGACC | | |
| JM0100 | GACAACCAGTCCCTGTGGACAGCCGGACCG | *lsp* (*sco2074*) reverse disruption primer (Redirect) | This work |
| | GAGGGGTCATGTAGGCTGGAGCTGCTTC | | |
| JM0113 | GCAACAGTGCCGTTGATCGTGCTATG | pMS82 cloning forward test primer | This work |
| JM0114 | GCCAGTGGTATTTATGTCAACACCGCC | pMS82 cloning reverse test primer | This work |
| JM0115 | GGATCCCTGTTCGCGGTCGCCCTGTTCGCGT | Forward primer amplifies a 411 bp fragment of the *lsp* gene, adding a | This work |
| | ACCT | bamHI site upstream. | |
| JM0116 | GATGCCGCCGCACACGATCGCCGAGTCGG | Reverse primer amplifies a 411 bp fragment of the *lsp* gene | This work |
| JM0150 | TCGTGCTCAGTCAAGGACCT | Sco Lsp Test For | Thompson et al., 2010 |
| JM0151 | GACAACCAGTCCCTGTGGAC | Sco Lsp Test Rev | Thompson et al., 2010 |
| JM0154 | AAGCTTCGACGAGGCGGACACAGGCAG | *lsp* (*sco2074*) comp (pMS82) for HindII | This work |
| JM0155 | GGTACCTCAGTCCTTGTGGACGGTCCCGTC | *lsp* (*sco2074*) comp (pMS82) Rev KpnI | This work |

**Table 1. Strains, plasmids and primers**

# Gene deletions and complementation

Gene deletions were carried out following the Redirect method of PCR-targeting [26] as previously described Hutchings *et al.* (2006). Disruption of *lsp (sco2074::apr)* on cosmid St4A10 (pJM013, St4A10Δ*lsp*) using the pIJ773 apramycin disruption-cassette and *sco6808 (sco6808::apr)* and *scr6809 (scr6809::apr)* on cosmid St1A2 (pJM010 - St1A2Δ*sco6808* and pJM012 - St1A2Δ*sco6808* respectively*)* using primers JM0101-2, JM0083-84 and JM0091-2 respectively were confirmed by PCR using primers JM0150-1, JM0085-6 and JM0093-4 respectively. Introduction of the wild-type cosmid St4A10 was facilitated by introducing an *oriT* by disruption and replacement of the Supercos-1 backbone *bla* resistance gene (pJM014 – St4A10*bla::hyg)* using primers JM0095-6 and the hygromycin disruption cassette from pIJ10701, confirmed using primers JM0099-100. The *lsp* suicide vector pJM016 was produced by introducing a 411 bp fragment of the *lsp* gene with an N-terminal *Bam*HI site, amplified with primers JM0117-8 and cloned into pGEM T-Eazy. The *Bam*HI site was then used to subclone the *Bam*HI fragment from a pIJ773 digest, containing an *apr* disruption cassette. An overexpression construct, pJM017 was synthesised by Genscript to include the pMC500 MCS and terminators [28] with *scr6809* (sequence included in Supplementary Text S5 online). All constructs were subsequently conjugated into *S. coelicolor* following the method described by Gust et al. (2002).

# Genomic DNA isolation

Genomic DNA was isolated from M145 and BJT1004 following the Pospiech and Neumann (1995) salting out method as described by Keiser et al. (2000). Mycelium from a 30 ml culture was resuspended in 5 ml SET buffer containing 1 mg/ml lysozyme and incubated at 37°C 30-60 min. To this lysate, 140 $\mu$l of proteinase K solution (20 mg $ml^{-1}$) was added, mixed, then 600 $\mu$l of 10% SDS added, mixed and incubated at 55°C for 2 h, with occasional mixing throughout. After this incubation 2 ml of 5 M NaCl was added, mixed and left to cool to 37°C before adding 5 ml chloroform, mixed at 20°C for 30 min. Samples were centrifuged at 4500 x g for 15 min at 20°C. The supernatant was removed to a fresh tube and DNA precipitated by adding 0.6 volumes of 100% isopropanol. Tubes were mixed by inversion and after at least 3 min DNA spooled out using a sterile Pasteur pipette. The DNA was rinced in 70% ethanol, air dried and dissolved in 1-2 ml TE buffer (10 mM Tris-HCl pH 7.8, 1 mM EDTA) at 55°C.

## Genome resequencing and secondary mutation identification

The isolated DNA from our wild-type *S. coelicolor* M145 parent strain and BJT1004 were sent to both GATC Biotech and The Genome Analysis Centre (TGAC) for 35bp paired end HiSeq Illumina sequencing. Assembly mapping and SNP identification was carried out with MIRA (Chevreux et al., 2004) using the reference genome NC_003888 (Bentley et al., 2002) as a scaffold for mapping each of the resequenced genomes. Putative SNPs were detected in each sample independently reporting the SNP position, the nucleotide change, the number of reads that sequence the region, those containing wild-type or mutated nucleotides and a percentage change. Each set of results was then compared by eye to determine the likely hood that a SNP was real by number of reads containing the mutation and its presence in each sample. Larger mutations (rearrangements) were identified in the same fashion.

## Microscopy

Brightfield images were acquired using a Zeiss M2 Bio Quad SV11 stereomicroscope. Samples were illuminated from above using a halogen lamp images captured with an AxioCam HRc CCD camera. The AxioVision software (Carl Zeiss, Welwyn Garden City, UK) was used for image capture and processing.

# References

1. Hutchings, M. I., Palmer, T., Harrington, D. J. & Sutcliffe, I. C. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold ’em, knowing when to fold 'em. *Trends Microbiol.* **17,** 13–21 (2009).
2. Buddelmeijer, N. The molecular mechanism of bacterial lipoprotein modification--How, when and why? *FEMS Microbiol. Rev.* **39,** 246–261 (2015).
3. Sutcliffe, I. C., Harrington, D. J. & Hutchings, M. I. A phylum level analysis reveals lipoprotein biosynthesis to be a fundamental property of bacteria. *Protein Cell* **3,** 163–170 (2012).
4. Gralnick, J. a, Vali, H., Lies, D. P. & Newman, D. K. Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4669–74 (2006).
5. Thompson, B. J. *et al.* Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium Streptomyces coelicolor. *Mol. Microbiol.* **77,** 943–957 (2010).
6. Shruthi, H., Babu, M. M. & Sankaran, K. TAT-pathway-dependent lipoproteins as a niche-based

adaptation in prokaryotes. *J. Mol. Evol.* **70,** 359–70 (2010).

7. Rahman, O., Cummings, S. P., Harrington, D. J. & Sutcliffe, I. C. Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World J. Microbiol. Biotechnol.* **24,** 2377–2382 (2008).

8. Okuda, S. & Tokuda, H. Lipoprotein Sorting in Bacteria. *Annu. Rev. Microbiol.* **65,** 239–259 (2011).

9. Narita, S. I., Matsuyama, S. I. & Tokuda, H. Lipoprotein trafficking in Escherichia coli. *Arch. Microbiol.* **182,** 1–6 (2004).

10. Jan, G., Fontenelle, C., Le Hénaff, M. & Wróblewski, H. Acylation and immunological properties of Mycoplasma gallisepticum membrane proteins. *Res. Microbiol.* **146,** 739–750 (1995).

11. Asanuma, M. *et al.* Structural evidence of α-aminoacylated lipoproteins of Staphylococcus aureus. *FEBS J.* **278,** 716–28 (2011).

12. Nakayama, H., Kurokawa, K. & Lee, B. L. Lipoproteins in bacteria: Structures and biosynthetic pathways. *FEBS J.* **279,** 4247–4268 (2012).

13. Kurokawa, K. *et al.* The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of Staphylococcus aureus Functions as a Native Ligand for Toll-like Receptor 2. *J. Biol. Chem.* **284,** 8406–11 (2009).

14. Serebryakova, M. V *et al.* The acylation state of surface lipoproteins of mollicute Acholeplasma laidlawii. *J. Biol. Chem.* **286,** 22769–76 (2011).

15. Widdick, D. a *et al.* Dissecting the complete lipoprotein biogenesis pathway in Streptomyces scabies. *Mol. Microbiol.* **80,** 1395–412 (2011).

16. Tschumi, A. *et al.* Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial lgt mutant. *J. Bacteriol.* **194,** 3938–49 (2012).

17. Xiao, Y. & Wall, D. Genetic redundancy, proximity, and functionality of lspA, the target of antibiotic TA, in the Myxococcus xanthus producer strain. *J. Bacteriol.* **196,** 1174–1183 (2014).

18. Yang, Y. H. *et al.* Finding new pathway-specific regulators by clustering method using threshold standard deviation based on DNA chip data of Streptomyces coelicolor. *Appl. Microbiol. Biotechnol.* **80,** 709–717 (2008).

19. Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1541–6 (2003).

20. Redenbach, M. *et al.* A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome. *Mol. Microbiol.* **21,** 77–96 (1996).

21. Cobb, R. E., Wang, Y. & Zhao, H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. *ACS Synth. Biol.* **4,** 723–8 (2014).

22. Vockenhuber, M.-P. *et al.* Deep sequencing-based identification of small non-coding RNAs in *Streptomyces coelicolor*. *RNA Biol.* **8,** 468–477 (2011).

23. Moody, M. J., Young, R. A., Jones, S. E. & Elliot, M. A. Comparative analysis of non-coding RNAs in the antibiotic-producing *Streptomyces* bacteria. *BMC Genomics* **14,** 558 (2013).

24. Romero, D. A. *et al.* A comparison of key aspects of gene regulation in Streptomyces coelicolor and Escherichia coli using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing. *Mol. Microbiol.* **94,** 963–987 (2014).

25. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. Practical Streptomyces Genetics. *John Innes Cent. Ltd.* 529 (2000). doi:10.4016/28481.01

26. Gust, B., Kieser, T. & Chater, K. PCR targeting system in Streptomyces coelicolor A3 (2). *John Innes Cent.* **3,** 1–39 (2002).

27. Hutchings, M. I., Hong, H.-J., Leibovitz, E., Sutcliffe, I. C. & Buttner, M. J. The sigma(E) cell envelope stress response of Streptomyces coelicolor is influenced by a novel lipoprotein, CseA. *J. Bacteriol.* **188,** 7222–9 (2006).

28. Gao, C., Mulder, D., Yin, C. & Elliot, M. a. Crp Is a Global Regulator of Antibiotic Production in Streptomyces. *MBio* **3,** 1–12 (2012).

29. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **417,** 141–147 (2002).

# Acknowledgments.

We thank Marie Elliot for sharing unpublished RNA sequence data for *Streptomyces coelicolor* M145. This work was supported by a NERC PhD studentship to JTM and BBSRC grants BB/F009429/1 and BB/F009224/1 to MIH and TP, respectively.

# Author contribution

JTM designed and carried out the experiments, JTM ICS TP MIH designed experiments, JTM DAW ICS TP MIH analysed data and wrote the manuscript, DAW prepared DNA for sequencing, GC analysed sequencing data.

# Additional information

**Supplementary information** accompanies this paper at:

**Competing financial interests:** The authors declare no competing financial interests.

# Figure legends

**Figure 1. Lipoprotein biogenesis in *Streptomyces coelicolor.*** Approximately 80% of precursor lipoproteins in *S. coelicolor* are translocated via the general secretory (Sec) pathway with around 20% being translocated by the twin arginine transport (Tat) pathway (A). Following translocation across the cytoplasmic membrane they are diacylated on the thiol of the lipobox (+1) cysteine residue by Lgt1 or Lgt (B) and then the signal sequence is cleaved by Lsp immediately upstream of that modified cysteine (C). Lnt1 then adds a third acyl chain to the amino group on the +1 cysteine to produce a triacylated lipoprotein (D). Lnt2 is not essential for triacylation *in vitro* but appears to increase its efficiency. The function of the N-acyl modification is not yet known.

**Figure 2. *IS21* insertion into *scr6809.*** The *sco6811-08* region of the *S. coelicolor* M145 genome contains 4 genes (*sco6811* (Purple), *sco6810* (green) , *sco6809* (yellow) and *sco6809* (blue) and a putative sRNA *scr6809* (large red) along with 3 putative promoters (broken arrows). Representations of the WT loci (A) and that sequenced from BJT1004 (B) indicate where an IS21 element (*sco6393* and sco6394) has inserted within *scr6809*. PCR verification of the IS21 insertion with primers JM0093 and JM0094 (small red arrows) was carried out (C) using M145 (lane 1), BJ1001 (lane 2) and BJT1004 (lane 3) genomic DNA. Lanes marked L contain the size ladders (Invitrogen 1kb plus DNA ladder), lane 1 contains the PCR product using wild-type M145 DNA (514 bp), lane 2 contains the PCR product using Δ*lsp* strain BJT1001 DNA and lane 3 contains the PCR product using genomic DNA from the cis complemented Δ*lsp* strain BJT1004 (both 2884 bp.

**Figure 3. Analysis of the *IS21* disrupted genomic region in *S. coelicolor*.** Colony morphology (A) shows that deletion of *sco6808* has no obvious effect on growth or development in wild-type M145 but does partially

restore sporulation in BJT1004 (recovery of *scr6809*). Disruption of *scr6809* in M145 results in a range of pleiotropic morphological and developmental phenotypes (B).

**Figure 4. New *lsp* mutants generated using Redirect do not contain the *IS21* mutation.** Colony morphology of mutants *lsp::apr* 1-10 (corresponding to strains JTM008.01 – JTM008.10), both single (n=7, colonies 1-2, 4-7 and 9-10) and double crossovers (n=3, colonies: 3 and 7-8) show a range of phenotypes (A). PCR of the *lsp* loci indicates colonies are either a single (WT and/or mutant band) or double (mutant band only) crossovers (WT = 687 bp, mutant = 1447 bp).

**Figure 5. Introduction of wild-type St4A10 causes a pleiotropic phenotype.** Conjugation of M145 with St4A10 *bla::hyg* results in non-wildtype phenotypes similar to those observed in the St4A10 *lsp::apr* single crossovers.

**Figure 6. Targeted disruption of *lsp* using a suicide vector results in a small colony phenotype that over-produces actinorhodin.** To test how much of the BJT1001 phenotype is due to loss of *lsp* we disrupted the *lsp* gene using a suicide vector which does not affect or duplicate any other target genes. Plate images (A) show two distinct phenotypes following insertion of the suicide vector (pJM016) into M145, either a wildtype appearance (M145 pJM016 (1), n=3 corresponding to JTM018.03-04 and 08) or a small colony phenotype over producing actinorhodin (M145 pJM016 (2), n=7, corresponding to strains JTM018.01-2, 05-07 and 09-10) similar to our original observation of the *lsp* phenotype [5].

Fig. 1

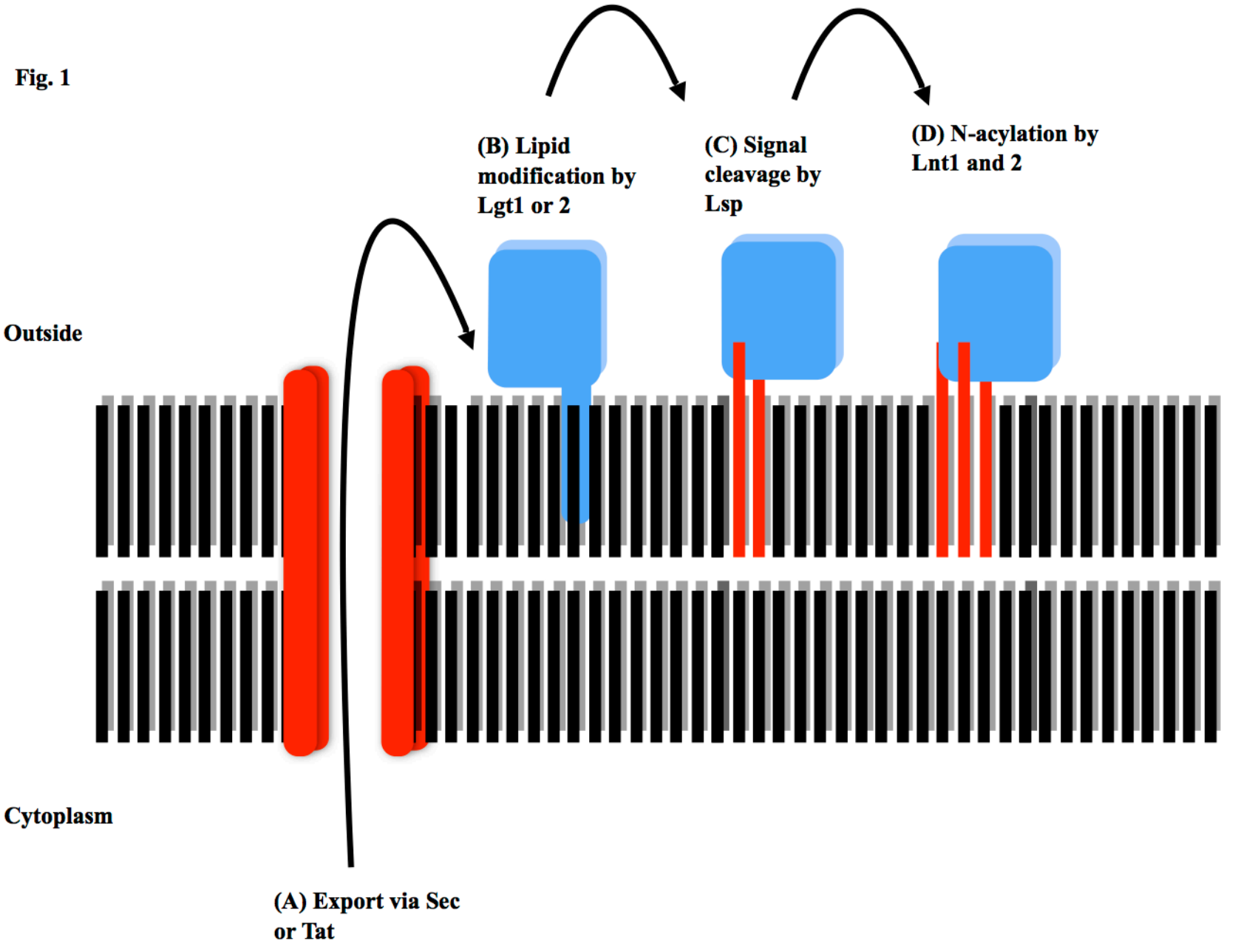

(B) Lipid modification by Lgt1 or 2
(C) Signal cleavage by Lsp
(D) N-acylation by Lnt1 and 2
Outside
Cytoplasm
(A) Export via Sec or Tat

Fig. 2

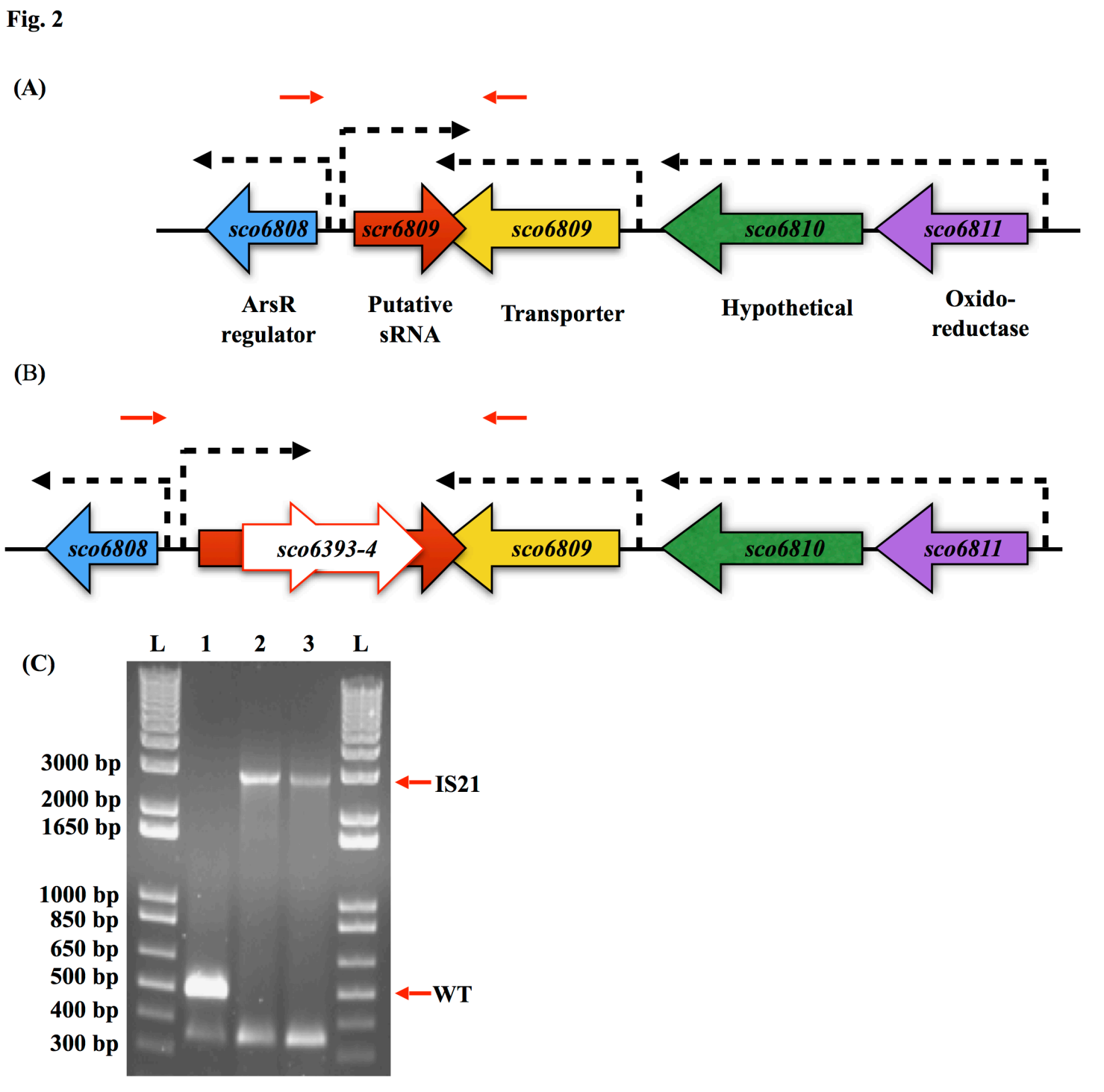

(A)
sco6808
scr6809
sco6809
sco6810
sco6811
ArsR regulator
Putative sRNA
Transporter
Hypothetical
Oxido-reductase
(B)
sco6808
sco6393-4
sco6809
sco6810
sco6811
(C)
L
1
2
3
L
3000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
500 bp
400 bp
300 bp
IS21
WT

Fig. 3

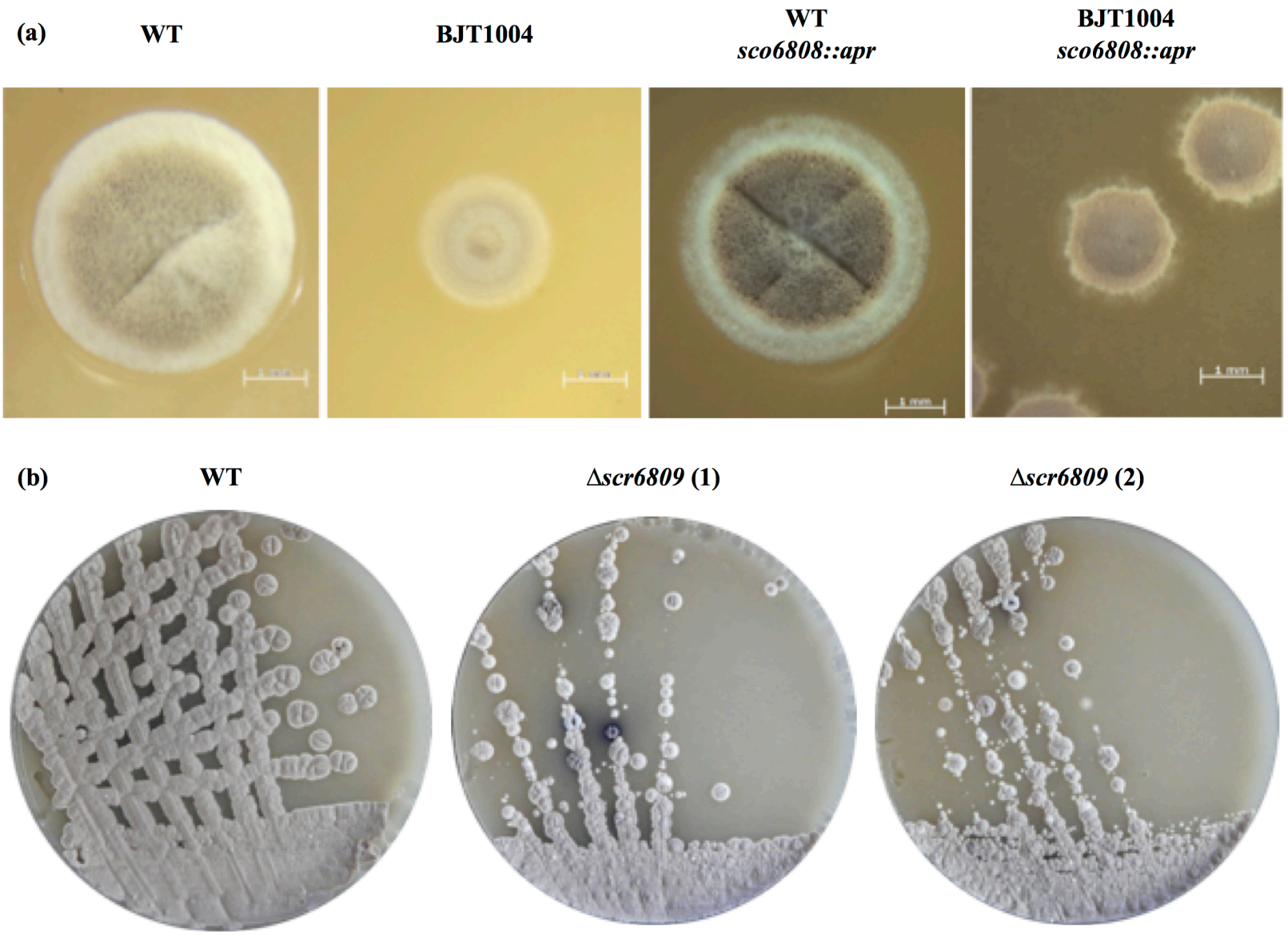

(a)
WT
BJT1004
WT
sco6808::apr
BJT1004
sco6808::apr
(b)
WT
Δscr6809 (1)
Δscr6809 (2)

**Fig. 4**

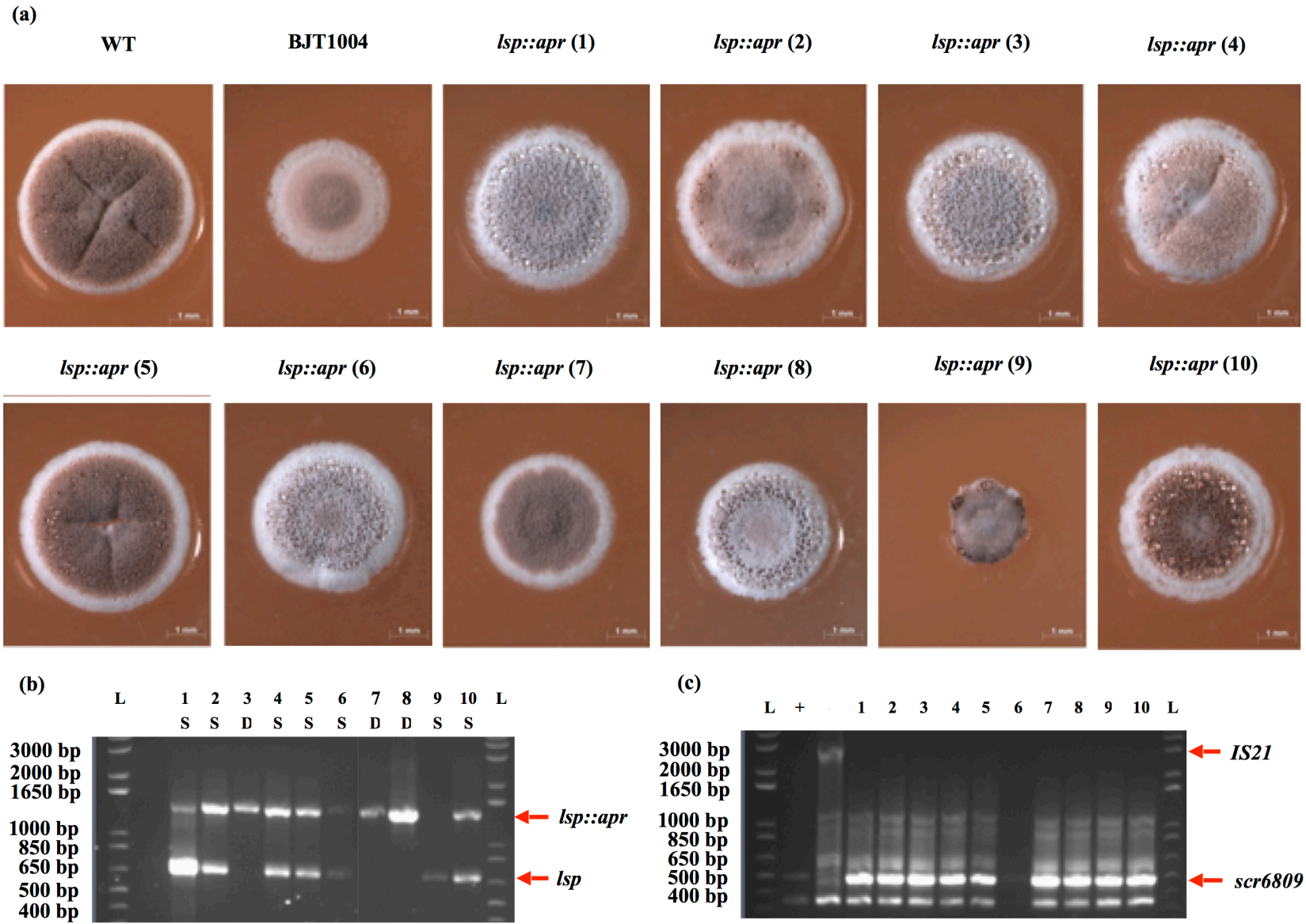

(a)
WT
BJT1004
lsp::apr (1)
lsp::apr (2)
lsp::apr (3)
lsp::apr (4)
lsp::apr (5)
lsp::apr (6)
lsp::apr (7)
lsp::apr (8)
lsp::apr (9)
lsp::apr (10)
(b)
L 1 2 3 4 5 6 7 8 9 10 L
S S D S S S D D S S
3000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
500 bp
400 bp
lsp::apr
lsp
(c)
L + 1 2 3 4 5 6 7 8 9 10 L
3000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
500 bp
400 bp
IS21
scr6809

Fig. 5

WT

St4A10 *Bla::hyg*

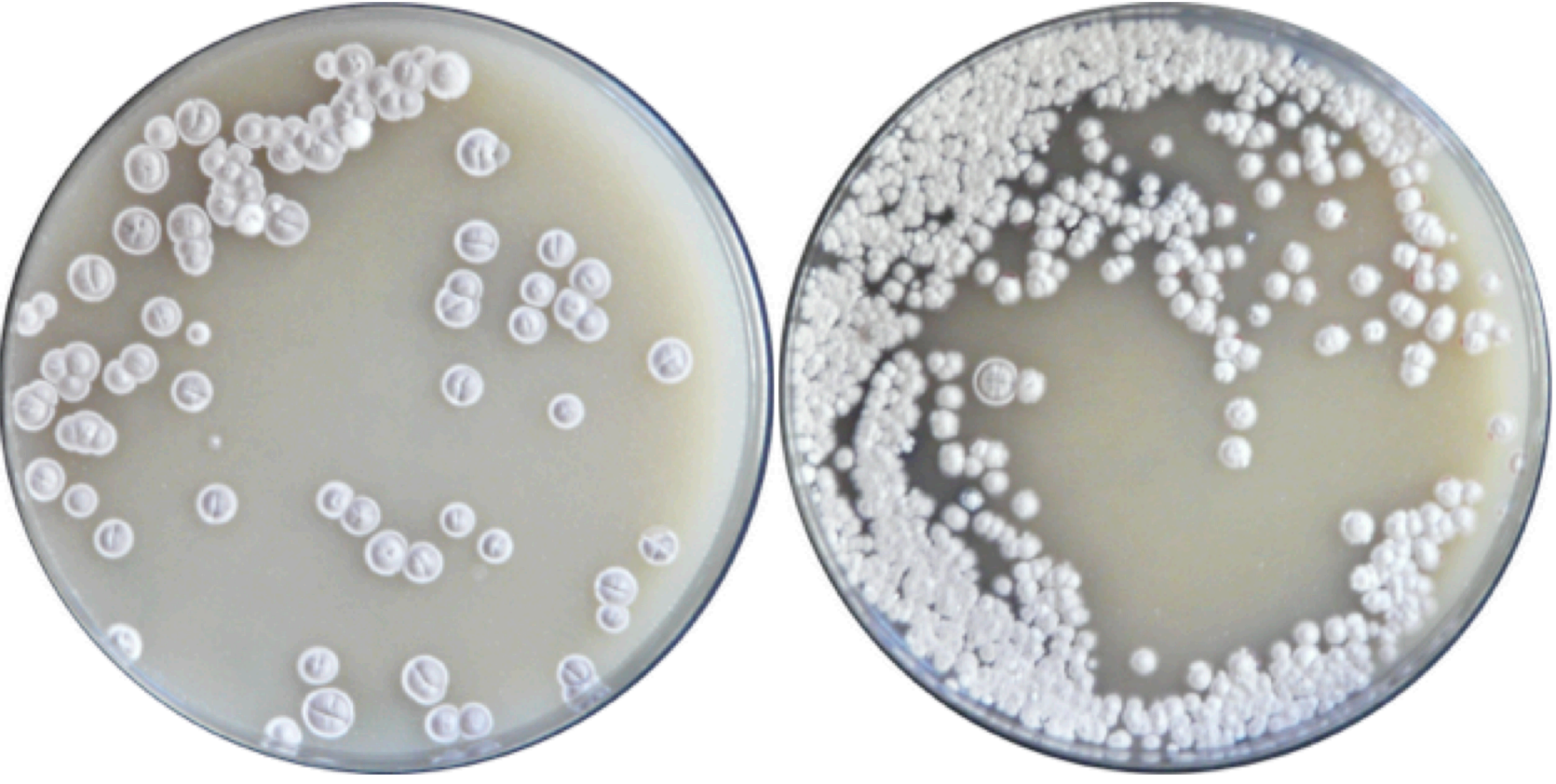
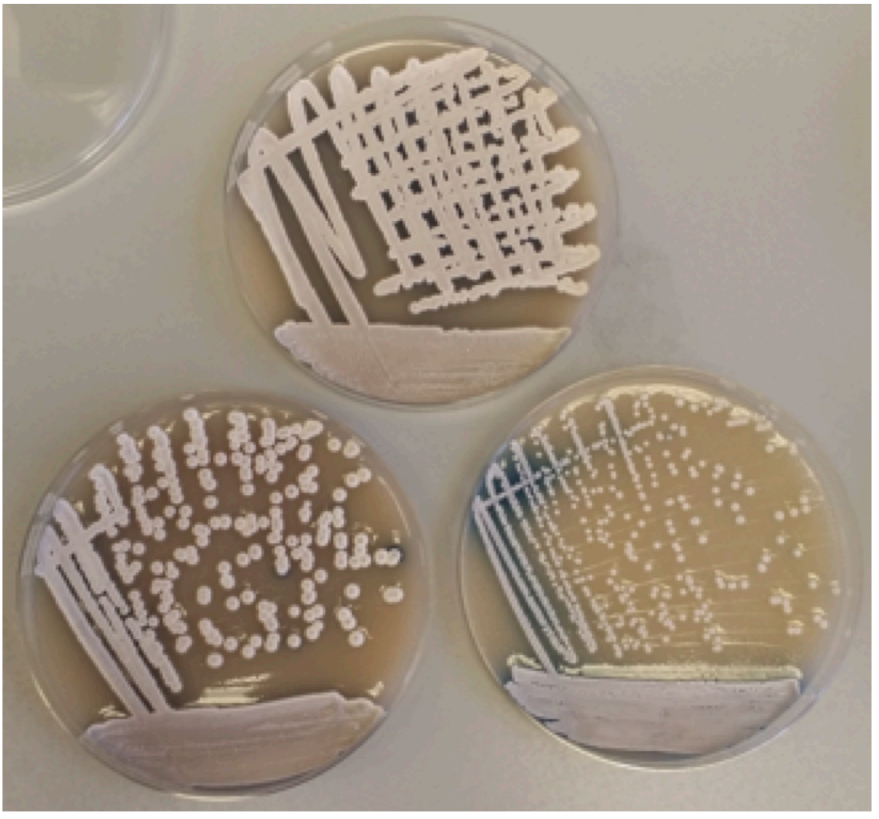
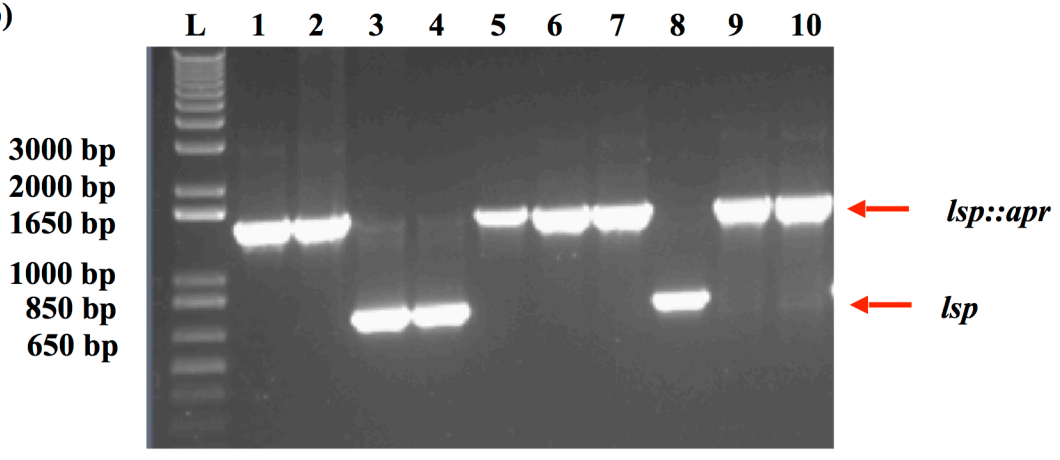


Fig. 6

(a)


M145
M145 pJM016 (1)
M145 pJM016 (2)

(b)


L
1
2
3
4
5
6
7
8
9
10
3000 bp
2000 bp
1650 bp
1000 bp
850 bp
650 bp
lsp::apr
lsp