# Investigating the symbiotic role of plasmid pMESCI01 in *Mesorhizobium ciceri* bv. biserrulae WSM1271 with *Biserrula pelecinus*

Rachel Jane Maymond Brewer

2017

This thesis is submitted in partial requirement for the degree of

Bachelor of Science with Honours in Molecular Biology

School of Veterinary and Life Sciences

Murdoch University

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

Rachel Brewer

## Acknowledgements

I would like to show my appreciation to those who have supported and guided me throughout the journey of my Honours project.

Firstly, I would like to thank my supervisor Dr. Jason Terpolilli for the opportunity to undergo this project. His guidance, support and encouragement have provided me with confidence in my own abilities allowing me to develop as a research student.

Secondly, I wish to thank ALOSCA for awarding me the ALOSCA Honours Scholarship 2016, as well as the benefactor for awarding me the Murdoch University Soil Microbiology and Soil Carbon Scholarship 2016.

I would like to thank Tim for his advice throughout my project. Specifically, teaching me the ways of the lab and your knowledge of WSM1271 has been invaluable. As well as to Talitha for always taking time to answer my many questions and being there for a good chat. Our group morning coffee runs have been a great motivation to get an early start each day!

I would like to take this opportunity to thank all the members of CRS for your support and assistance whenever needed. With thanks to Sofie for your advice on Eckhardt Gels, as well as to Regina and Yvette for your help with my glasshouse trial. Special mentions to George and Amanuel for your great

company in our office, as well as to Emma for your enthusiasm and study sessions in these last couple of months.

To all my friends and boyfriend Calvin who have always believed in my ability and kept me motivated with ice cream and coffee breaks, and a lot of laughter throughout the years. I promise we can celebrate the completion of this project with a good-old karaoke session!

To my sister, Lucy, for always being so supportive of my goals and listening to me talk about my project constantly for the past few weeks. Thank you for always joining in on my random singing and dancing parties and keeping me company on our long train journeys to university.

To my parents, Martin and Janet, who have always put my sister and I first to ensure we can take on every opportunity presented to us. This has allowed me to get to where I am today and I am truly grateful for everything you have done.

Lastly, I would like to mention my beautiful dog whose love and happiness I will always remember.

Astley

(July 2000 – October 2016)

# Abstract

The establishment of a nitrogen-fixing symbiotic relationship between soil-dwelling bacteria (rhizobia) and legumes is a valuable source of bioavailable nitrogen in the biosphere. The genes required for the establishment and maintenance of this symbiosis may be encoded on plasmids or within the chromosome of the rhizobial microsymbiont as a Symbiosis Island (SI). Previous studies have found that *Mesorhizobium ciceri* bv. biserrulae WSM1271, which forms a nitrogen-fixing symbiosis with the pasture legume *Biserrula pelecinus*, transferred its tripartite SI to at least two presumably non-symbiotic *Mesorhizobium* spp., converting them into *B. pelecinus*-nodulating rhizobia. These newly evolved symbionts were not as effective in nitrogen fixation with *B. pelecinus* as WSM1271, even though they had acquired the entire SI from WSM1271. Importantly, WSM1271 harbours a plasmid (pMESCI01) that is absent from WSM2073 and WSM2075. This plasmid may carry additional determinants required for highly effective symbiotic interactions with *B. pelecinus*. Therefore, analysis of the potential symbiotic role and transmissibility of this plasmid may help explain the differences in nitrogen fixation of these three strains with *B. pelecinus.*

Bioinformatic analysis of pMESCI01 revealed >50% of genes to have hypothetical functions only, with a diverse range of function predictions for the remainder of genes. Crucially, 12 putative nitrogen fixation genes were identified on pMESCI01 that could have a role in the symbiosis between WSM1271 and *B. pelecinus.* Homologs of known plasmid conjugal transfer type IV secretion

systems (T4SS) or relaxases were not detected, but a potential origin of transfer (*oriT*) site was located, leaving open the question of whether pMESCI01 is self-transmissible.

To determine whether pMESCI01 is self-transmissible, the plasmid was marked with an Ω-Sp/Sm cassette, encoding spectinomycin and streptomycin resistance. Possible transconjugants from a subsequent conjugation experiment between WSM1271-pMESCI01::Ω-Sp/Sm and WSM2073Nm could not be screened, owing to unexpected spectinomycin resistance of the WSM2073Nm recipient. However, WSM2073Nm was confirmed to be highly sensitive to streptomycin, so the conjugation experiment can be repeated in future to assess self-transmissibility of pMESCI01 with pMESCI01::Ω-Sp/Sm.

To assess whether pMESCI01 was essential to nitrogen fixation in WSM1271 with *B. pelecinus*, the curing vector pSRK*repABC* was constructed and plasmid-cured derivatives of WSM1271 were produced via a plasmid incompatibility approach. Phenotypic assessment of wild-type WSM1271 compared to the plasmid-cured derivatives inoculated onto *B. pelecinus* in a glasshouse trial revealed no difference in mean shoot dry weights or nodulation of plants across the treatments. This indicated that the plasmid-located genes are likely to not be essential to nitrogen fixation in WSM1271 with this host.

Alignment of the sequence of pMESCI01 with accessory plasmids from other fully-sequenced *M. ciceri* strains revealed a high percentage of homology across the plasmids from these strains, which originate from geographically diverse

locations. This suggests that these *M. ciceri* plasmids are widely dispersed and may be a common feature of this species. While the work in this thesis indicates that pMESCI01 is not essential in the symbiosis between WSM1271 and *B. pelecinus*, genes on the plasmid may still impart important regulatory or metabolic benefits to WSM1271. Future investigations into the possible function of the large number of genes (432 in total) on pMESCI01 may help to address the role of this plasmid in WSM1271 and, by extension, the role of accessory plasmids in other *M. ciceri* strains. Therefore, continued studies of *Mesorhizobium spp.* would be important in determining the evolution, transmissibility and role of these plasmids.

# Contents

# List of Abbreviations

| | |
|---|---|
| **ALA** | Aminolevulinic acid |
| **ANOVA** | One-way analysis of variance |
| **BLAST+** | Basic Local Alignment Search Tool Plus (+) database |
| **BLASTN** | Basic Local Alignment Search Tool of nucleotide sequences |
| **BLASTP** | Basic Local Alignment Search Tool of protein sequences |
| **Bp** | Base pairs |
| **BRIG** | BLAST Ring Image Generator |
| **COGs** | Clusters of Orthologous Groups of Proteins |
| **DMSO** | Dimethyl sulfoxide |
| **Dtr** | DNA transfer and replication |
| **DXO** | Double crossover |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **ExPASY** | Expert Protein Analysis System |
| **FeMo-cofactor** | Iron-molybdenum cofactor |
| **Fix$^-$** | Unable to fix nitrogen |
| ***Gm*** | *Gentamycin* |
| **Kb** | Kilobase |
| ***Km*** | *Kanamycin* |
| **LB** | Lysogeny broth |
| **Mpf** | Mating pair formation |
| **N+** | Nitrogen fed un-inoculated control group |
| **N-** | Nitrogen starved un-inoculated control group |
| **NADH** | Nicotinamide adenine dinucleotide hydride |
| ***Nm*** | *Neomycin* |
| **PCR** | Polymerase chain reaction |
| **RDM** | Rhizobium defined media |
| **SDS** | Sodium dodecyl sulfate |
| ***Sm*** | *Streptomycin* |
| ***Sp*** | *Spectinomycin* |

| | |
|---|---|
| **Suc** | Sucrose |
| **SXO** | Single crossover |
| **T4SS** | Type IV secretion system |
| **T6SS** | Type VI secretion system |
| **TAE** | Tris-acetate-EDTA buffer |
| **TA systems** | Type II toxin-antitoxin system |
| **TBE** | Tris-borate-EDTA buffer |
| **TY** | Tryptone yeast |

# 1. Introduction

## 1.1 Symbiotic nitrogen fixation

Nitrogen is an essential element for the growth and development of all organisms and is often the most limiting nutrient affecting plant growth. Although approximately 78% of the atmosphere consists of dinitrogen ($N_2$), this form of nitrogen is metabolically inaccessible to most organisms owing to its chemical stability (Gruber *et al.*, 2008). Dinitrogen can, however, be reduced or "fixed" into ammonia ($NH_3$) by some bacteria in a process called biological nitrogen fixation, which is a crucial link in the biosphere's nitrogen cycle providing many organisms with a source of bioavailable nitrogen (Gruber and Galloway, 2008; Fowler *et al.*, 2013). Global estimates suggest that biological nitrogen fixation (including free-living, associative and symbiotic) contributes approximately 90 Tg $Nyr^{-1}$ of fixed nitrogen to the biosphere, with just over 50% of fixed nitrogen derived from the symbioses between soil bacteria (rhizobia) and plants, most of which are legumes (Vance, 2001).

This symbiosis can be harnessed in agriculture to provide bioavailable nitrogen to crops without the need for application of industrially synthesised fertilisers. As well as the relatively high carbon foot-print associated with fertiliser production, synthetic fertiliser application is also detrimental to ecosystem health, particularly through eutrophication of ground and surface water (Galloway *et al.*, 1995; Vance, 2001). Utilisation of the rhizobia-legume symbiosis in agriculture can therefore reduce both the economic and environmental cost of farming (Vance, 2001; Graham *et al.*, 2003; Herridge *et al.*, 2008).

## 1.2 Establishment of nitrogen-fixing symbioses

Rhizobia are Gram-negative, rod-shaped soil dwelling bacteria capable of forming an endosymbiotic relationship with species of legumes. Establishment of nitrogen-fixing symbioses enables rhizobia to provide fixed nitrogen to the host in return for a carbon supply from the plant. Rhizobia are phylogenetically diverse bacteria that are taxonomically divided into *Alphaproteobacteria* and *Betaproteobacteria* classes. Currently, 14 genera of rhizobia are grouped within the *Alphaproteobacteria* (known as α-rhizobia) while 2 genera of rhizobia are within *Betaproteobacteria* (or β-rhizobia) (Sprent *et al.*, 2017). Importantly, our understanding of symbiotic nitrogen fixation is informed almost exclusively by studies of the α-rhizobia and their associated legume hosts.

The establishment of the symbiotic relationship between most rhizobia and their legume host is dependent on the synthesis of lipochito-oligosacchride polymers (Nod Factors) by rhizobia. However, it is important to note that Nod Factor-independent symbioses also exists, such as has been described for the photosynthetic *Bradyrhizobium spp.* (Nouwen *et al.*, 2016; Okazaki *et al.*, 2016). Nevertheless, this review will focus on the more common and well-studied Nod Factor-dependent process.

Nod Factor-dependent nodulation involves three stages (Figure 1.1). First, recognition between legume and rhizobia occurs via a molecular signal exchange, with chemicals including flavonoids or isoflavonoids being released by legume roots, which in turn activate the expression of the nodulation genes within the rhizobia. These nodulation genes are required for the synthesis of Nod

Factor by the rhizobia. Second, Nod Factor-induced nodule organogenesis begins when a nascent nodule forms on the roots (or stems as in the case of *Sesbania rostrata*) of the host legume and rhizobia infect and occupy this developing nodule. Third, rhizobia differentiate into bacteroids, which is their nitrogen-fixing physiological state. The nitrogen fixation genes (*nif* and *fix*) within bacteroids are expressed, beginning the process of symbiotic nitrogen fixation within the nodule (Laranjo *et al.*, 2014); (Figure 1.1).

### 1.2.1 Initiation of nodulation

The nodulation process begins with the establishment of a molecular dialogue between rhizobia and their legume host. In this dialogue, the host releases signal molecules, including flavonoids or isoflavonoids, which activate the expression of the transcriptional regulator NodD in rhizobia (Horvath *et al.*, 1987; Laranjo *et al.*, 2014); (Figure 1.1). NodD is a member of the LysR-type family of transcriptional regulators and directs the transcription of nodulation genes (commonly referred to as *nod* genes, although they include genes designated as *nod*, *noe*, and *nol*). There are approximately 14 core *nod* genes common to all Nod Factor-dependent symbioses, including *nodD,* that are responsible for directing the synthesis and secretion of Nod Factor, as well as encoding for decorations of the Nod Factor-backbone; specific roles of nodulation genes are reviewed in detail by Spaink (2000), Vanrhijn *et al.* (1995) and Denarie *et al.* (1996).

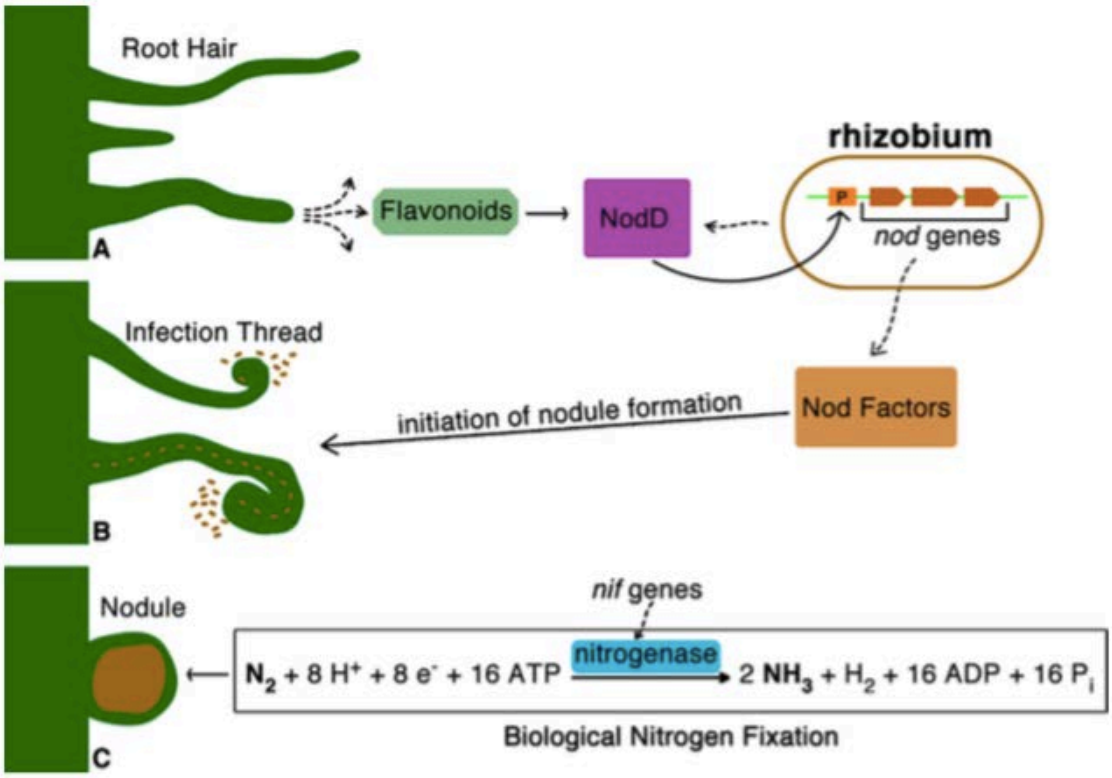


**Figure 1.1:** Outline of the three-stage process for Nod Factor-dependent establishment of a symbiosis between rhizobia and legume requiring root hair curling infection to initiate biological nitrogen fixation. (A) Legume roots exude chemical signal molecules, such as flavonoids, which are taken up by the rhizobia and activate rhizobial NodD proteins. Activated NodD directs the transcription of nodulation (*nod* but also *nol and noe*) genes. (B) Expression of *nod* genes leads to the synthesis of Nod Factor. Nod Factor initiates nodule formation, which in turn simulates infection of the host root hair cells by rhizobia via an infection thread. (C) Once the nodule is formed the rhizobia differentiate within the nodule into bacteroids, *nif* and *fix* (not shown) genes are expressed and the process of biological nitrogen fixation can begin. Figure adapted from Laranjo *et al.* (2014).

Once synthesised, Nod Factor is secreted by rhizobia into the rhizosphere (Figure 1.1). Nod Factor can then bind to host Nod Factor receptors on the surface of plant cells, initiating infection of the roots by rhizobia. In the most well-studied symbioses, this typically occurs through a controlled infection of host root hair cells by rhizobia (reviewed in detail by Oldroyd *et al.* (2008); Murray (2011)). There are also two mechanisms of non-root hair infection;

namely crack-entry infection (Oldroyd and Downie, 2008) and epidermal infection (Sprent, 2009), which are beyond the scope of this review.

Binding of the Nod Factor to Nod Factor receptors triggers root hair curling (Figure 1.1). During this event, rhizobia attached to the root hair tip are trapped within the hair as it curls (Oldroyd and Downie, 2008). Invagination of the root hair cell forms an infection thread that rhizobia occupy through cell division, with the infection thread ramifying through the root cortical cell layer towards the developing nodule primordium (Oldroyd and Downie, 2008). Rhizobia then proceed down the infection thread by cell division, ultimately released as infection droplets into the newly formed nodule (Gage, 2004; Oldroyd and Downie, 2008).

Based on their morphology, nodules can be separated into two broad groups, determinate or indeterminate, with the type of nodule formed dependent on the legume host (Sprent, 2007). Determinate nodules (e.g. *Glycine max* and *Phaseolus vulgaris*) lack a persistent meristem, with growth of the nodule dependent on the enlargement of infected cells. Moreover, determinate nodules contain a single, relatively homogenous nitrogen-fixing zone (Sprent, 2007). In contrast, indeterminate nodules (*e.g. Medicago*, *Pisum* and *Biserrula*) are characterised by persistent apical meristems, with nodules divided into a number of zones, delineating different physiological states of both host and rhizobial cells (Sprent, 2007). Indeterminate nodules are the nodule type discussed within this review.

Indeterminate nodules can be divided into four zones: Zone I) the tip of the nodule, where plant cells are actively dividing, allowing for continual growth of the nodule, and where infecting bacteria exit the infection threads; Zone II) where plant nodule cells endoreduplicate (i.e. increase in chromosomal number without concomitant cell division) resulting in an increase in cell volume, allowing for increased infection by rhizobia released from the infection thread into the nodule, and where rhizobia begin to differentiate into bacteroids; Zone III) where bacteroids reside in their mature nitrogen-fixing state within the microaerobic environment of the nodule; and Zone IV) where bacteroids begin to disintegrate and undergo senescence (Melino *et al.*, 2012); (Figure 1.2).

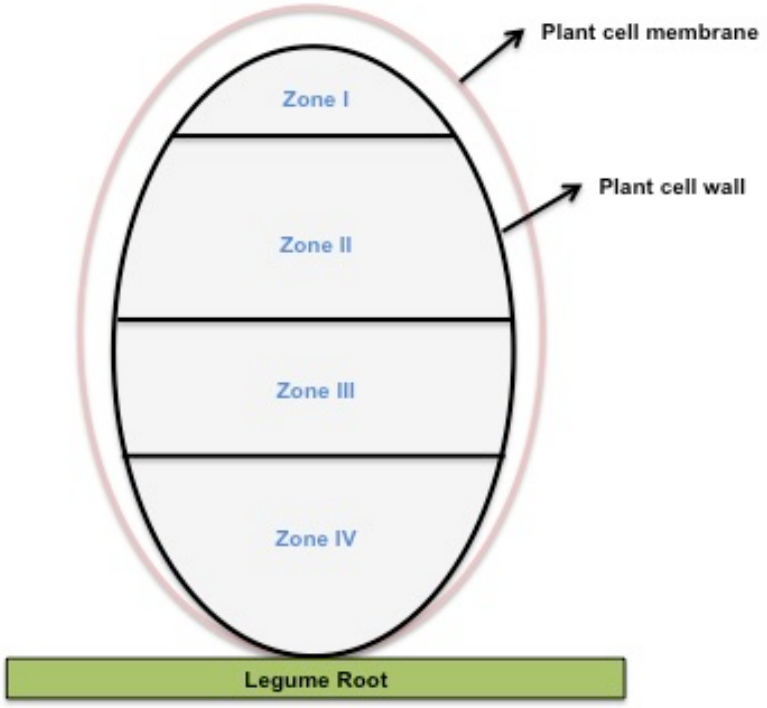


**Figure 1.2:** Representation of the structure and location of the zones of an indeterminate nodule contained within the legume plant cells. Zone I is the tip of the nodule where bacteria exit their infection thread to infect the continually dividing plant cells; Zone II, where the nodule plants cells endoreduplicate allowing increased infection by rhizobia, which then begin to differentiate into bacteroids; Zone III, the location of nitrogen fixation where mature bacteroids reside in the microaerobic environment of the nodule; and Zone IV, where bacteroids undergo senescence. The location of the legume root in relation to the nodule is shown. Figure adapted from Melino *et al.* (2012).

## 1.3 Bacteroid nitrogen fixation

The expressions of a range of genes are essential for the process of bacteroid nitrogen fixation. Common bacteroid nitrogen fixation genes (*nif* and *fix*), along with their known or putative function, are summarised in Table 1.1. To initiate bacteroid nitrogen fixation the transcriptional regulator, NifA, drives the expression of a range of gene including *nifHDK, fixABCX, nifB, nifEN,* and *fdxN* (Buikema *et al.*, 1985; Fischer, 1994). Activation of *nifA* requires a complex cascade of transcriptional regulators; the exact nature of which varies across species of rhizobia (Terpolilli *et al.*, 2012). In *Sinorhizobium meliloti* for example, low $O_2$ (<50nM) activates the sensor protein FixL that in turn activates the transcriptional regulator FixJ, subsequently increasing expression of FixK and NifA (Dixon *et al.*, 2004). This process is regulated by FixT which represses expression of FixK by inhibiting the activity of FixL (Dixon and Kahn, 2004). In contrast in *Bradyrhizobium japonicum*, a secondary redox-responsive RegS-RegR system is required to activate expression of *nifA* (Dixon and Kahn, 2004).

Mature nitrogen-fixing bacteroids express nitrogenase, a very $O_2$-labile enzyme complex, that catalyses the reduction of nitrogen ($N_2$) into ammonia ($NH_3$) (Fischer, 1994; Rubio *et al.*, 2008). Under ideal conditions the stoichiometry of this reaction is:

**Table 1.1:** Common nitrogen fixation genes required for the synthesis and assembly of nitrogenase and iron-molybdenum cofactor (FeMo-cofactor), and the regulation of bacteroid electron transport chain have been included, with their functions and primary reference strain provided where possible. Other genes involved in bacteroid nitrogen fixation that have not been included in this review are reviewed by Buikema *et al.* (1985); Fischer (1994); Rubio and Ludden (2008) and Terpolilli *et al.* (2012). Please note that the most current species names of primary reference species have been shown and may differ to the name referred to in the reference article.

| | **Gene/s** | **Product** | **Function** | **Primary reference species** | **Primary reference** |
|---|---|---|---|---|---|
| **Biological nitrogen fixation genes** | *nifHDK* | NifHDK | Structural genes for synthesis of nitrogenase enzyme; NifH contains Fe-S cluster and hydrolysis site and MgATP site, NifDK complex contains Fe-Mo cofactor +P | *Bradyrhizobium japonicum* | *nifH; (Fuhrmann et al., 1984)* and *nifDK;* (Thony *et al.*, 1985; Brigle *et al.*, 1987) |
| | *nifA* | NifA | Oxygen-sensitive transcriptional regulator of *nif* genes | *Sinorhizobium meliloti* | (Buikema *et al.*, 1985) |
| | *nifB* | NifB | Biosynthesis of FeMo-cofactor | *Sinorhizobium meliloti* | (Buikema *et al.*, 1987) |
| | *nifEN* | NifEN | Biosynthesis of FeMo-cofactor | *Sinorhizobium meliloti* | (Aguilar *et al.*, 1990) |
| | *nifUSV* | NifUSV | NifSV provide substrates required for FeMo-cofactor synthesis, NifU acts as a molecular scaffold in FeMo-cofactor synthesis | *Azobacter vinelandii*<br>*Klebsiella pneumoniae* | *nifUS;* (Jacobson *et al.*, 1989)<br>*nifV; (Hoover et al., 1988)* |
| | *nifF* | NifF | Flavodoxin | *Klebsiella pneumoniae* | (Shah *et al.*, 1983) |
| | *nifJ* | NifJ | Pyruvate flavodoxin oxidoreductase | *Klebsiella pneumoniae* | (Shah *et al.*, 1983) |
| | *nifAL* | NifAL | Regulatory system that controls the expression of *nif* genes to ensure they function only in appropriate conditions | *Klebsiella pneumoniae* | (MacNeil *et al.*, 1978) |
| | *nifQ* | NifQ | Provides molybdenum in FeMo-cofactor synthesis | *Klebsiella pneumoniae* | (Imperial *et al.*, 1984) |
| | *fixABCX* | FixABCX | Electron transport chain to nitrogenase | *Sinorhizobium meliloti* | (Earl *et al.*, 1987) |
| | *fixNOPQ* | FixNOPQ | Cytochrome $cbb_3$-type oxidase that is microaerobically induced | *Bradyrhizobium japonicum* | (Preisig *et al.*, 1993) |
| | *fixLJ* | FixLJ | Transcriptional regulator of *nifA* gene | *Sinorhizobium meliloti* | (David *et al.*, 1988) |
| | *fixK* | FixK | Transcriptional regulator | *Sinorhizobium meliloti* | (Foussard *et al.*, 1997) |
| | *fixGHIS* | FixGHIS | Assembly of cytochrome $cbb_3$-type oxidase | *Bradyrhizobium japonicum* | (Zufferey *et al.*, 1996) |
| | *fdxN* | Ferredoxin | Ferredoxin | *Sinorhizobium medicae* | (Klipp *et al.*, 1989) |

Nitrogenase is encoded by the *nifHDK* operon, driven by the low oxygen tension within the nodules. NifH is a homodimer that contains an iron-sulfur cluster as well as sites for hydrolysis and MgATP. Whereas NifDK is a heterotetrameric complex containing an iron-molybdenum cofactor (FeMo-cofactor) and phosphate group (Rubio and Ludden, 2008). Numerous genes are required for the synthesis of the FeMo-cofactor including *nifS, nifV, nifQ, nifB, nifEN* and *nifU*. The enzymes encoded by *nifS, nifV* and *nifQ* provide the required sulfur group, homocitrate, and molybdenum to the FeMo-cofactor, respectively (Hoover *et al.*, 1988; Imperial *et al.*, 1989; Jacobson *et al.*, 1989). Synthesis of active nitrogenase also requires NifB, which functions to catalyse the first step in the biosynthesis of the FeMo-cofactor, in which 4Fe4S clusters are converted to NifB-cofactor, an Fe-S cluster that serves as a precursor of the FeMo-cofactor (Fischer, 1994; Guo *et al.*, 2016). NifEN and NifU provide a molecular scaffold during the biosynthesis of the FeMo-cofactor (Jacobson *et al.*, 1989; Aguilar *et al.*, 1990). The function of these and other genes required for the synthesis of the FeMo-cofactor are extensively reviewed by Rubio and Ludden (2008).

The *fixABCX* operon encodes a putative electron-transferring flavoprotein ($ETF_N$) complex, which is an alternative electron transport chain pathway found in rhizobia that is the likely route by which electrons are channelled to nitrogenase (Terpolilli *et al.*, 2016). Within the *fixABCX* operon, *fixAB* encodes an electron-transfer flavoprotein ($ETF_N$), *fixC* encodes for a presumed $ETF_N$ oxidoreductase, and *fixX* encodes for a ferredoxin, all of which are consistent with an electron transport chain (Scott *et al.*, 2004). The specific mechanism of this $ETF_N$ complex

has yet to be formally shown in bacteroids, however, a possible mechanism has been proposed by Terpolilli *et al.* (2016).

The gene complexes *fixNOQP* and *fixGHIS* are required for the synthesis and operation of the high-affinity cytochrome *cbb*$_3$-type oxidase that is essential for bacteroid respiration under the low $O_2$ conditions inside the legume root nodules (Preisig *et al.*, 1993; Thony-Meyer, 1997; Pitcher *et al.*, 2004). The *fixNOQP* operon was first demonstrated to be essential to nitrogen fixation in *Bradyrhizobium japonicum* USDA110 (Preisig *et al.*, 1993) and is composed of three major subunits: subunit I is a membrane-bound integral *b*-type cytochrome, and subunits II and III are membrane-anchored *c*-type cytochromes; with *fixN* encoding for typical heme-copper oxidase of subunit I; *fixO* and *fixP* encoding for the membrane-bound *c*-type cytochromes; and *fixQ* encoding for a small membrane protein, the function of which is currently unknown (Zufferey *et al.*, 1996). The *fixGHIS* operon appears to encode for a copper uptake system, which is likely involved in the provision and assembly of the copper cofactor in the cytochrome *cbb*$_3$-type terminal oxidase (Preisig *et al.*, 1996; Zufferey *et al.*, 1996). The complex function of this cytochrome *cbb*$_3$-type oxidase is beyond the scope of this review, but has been extensively reviewed by Thony-Meyer (1997); Pitcher and Watmough (2004).

Establishment of nitrogen fixing symbiotic association therefore requires interaction between host plants and rhizobia. Rhizobia themselves must possess and express a suite of appropriate *nod*, *nif* and *fix* genes to be transformed from soil-dwelling microbes into nodule-inhabiting nitrogen-fixing bacteroids.

## 1.4 Location of symbiosis genes in rhizobial genomes

The location of symbiosis genes (nodulation and nitrogen fixation genes) varies across different genera of rhizobia. Symbiosis genes are known to be located on large accessory plasmids, known as symbiosis plasmids or pSym, in some genera of rhizobia, including *Rhizobium* and *Sinorhizobium* (Johnston *et al.*, 1982; Young *et al.*, 2006; Blanca-Ordonez *et al.*, 2010). While in other rhizobia, such as *Mesorhizobium* or *Bradyrhizobium,* the symbiosis genes are encoded on the bacterial chromosome or Symbiosis Islands (SI) (Sullivan *et al.*, 1998; Kaneko *et al.*, 2000; Kaneko *et al.*, 2002; Sullivan *et al.*, 2002; Nandasena *et al.*, 2006; Barcellos *et al.*, 2007; Giraud *et al.*, 2007; Nandasena *et al.*, 2007; Cytryn *et al.*, 2008; Kasai-Maita *et al.*, 2013). The differences in location of symbiosis genes may be due to differences in the evolution of rhizobia dependent on the transmissibility and stability of plasmids or transfer of Symbiosis Islands.

### 1.4.1 Symbiosis plasmids

Rhizobia may contain accessory plasmids, some of which harbour the nodulation and nitrogen fixation genes required for functional symbioses that are referred to as symbiosis plasmids (pSym). These pSym have predominately been found in *Rhizobium spp.* and *Sinorhizobium spp.* with the number and size of symbiotic plasmids varying between strains (Johnston *et al.*, 1982; Young *et al.*, 2006; Blanca-Ordonez *et al.*, 2010). Interestingly, a single *Mesorhizobium spp., M. huakii* 7653R contains a plasmid, pMHb, that has been identified as a pSym (Peng *et al.*, 2014; Wang *et al.*, 2014). The ability to maintain plasmids throughout rhizobial populations and their transfer ability is important in understanding the role and evolution of plasmids, specifically those that harbour symbiosis genes.

In general, rhizobial plasmids tend to be low-copy number plasmids and as such require efficient mechanisms to ensure stable propagation of plasmids to daughter cells (Mazur *et al.*, 2012; Pinto *et al.*, 2012). This is in contrast to high-copy numbered plasmids, which are distributed stochastically to progeny cells and do not require such complex mechanisms to ensure their stability (Thomas, 2000; Pinto *et al.*, 2012). The known mechanisms that low-copy number plasmids employ to ensure their survival in progeny has been summarised by Pinto *et al.* (2012), which include: 1) replication must occur at a high enough frequency to ensure that a sufficient number of plasmid copies are available to populate daughter cells; 2) an efficient partitioning system that physically distributes newly replicated plasmids into the daughter cells; and 3) toxin-antitoxin systems that ensures the loss of plasmid-less daughter cells due to stable toxin molecules, and survival of plasmid-containing cells with protection by antitoxins.

***Plasmid replication and stability***

Rhizobial plasmids often contain a *repABC*-type replication and partitioning system, which controls the replication, segregation and copy number of the plasmid (Ramirez-Romero *et al.*, 2000; MacLellan *et al.*, 2005; Pinto *et al.*, 2012). The general structure of the *repABC* operon has been determined for *Rhizobium etli* plasmid p42d by Ramirez-Romero *et al.* (2000) and contains genes for replication and partitioning systems, as well as an incompatibility region (Figure 1.3). In the *repABC* operon RepC functions as the replicase to initiate plasmid replication, while RepA and RepB are part of a partitioning system that segregates replicated plasmids into daughter cells; the specific function and

regulation of these genes are substantially reviewed by Ramirez-Romero *et al.* (2000); Bignell *et al.* (2001); MacLellan *et al.* (2006); Cervantes-Rivera *et al.* (2011); Mazur and Koper (2012); Pinto *et al.* (2012).

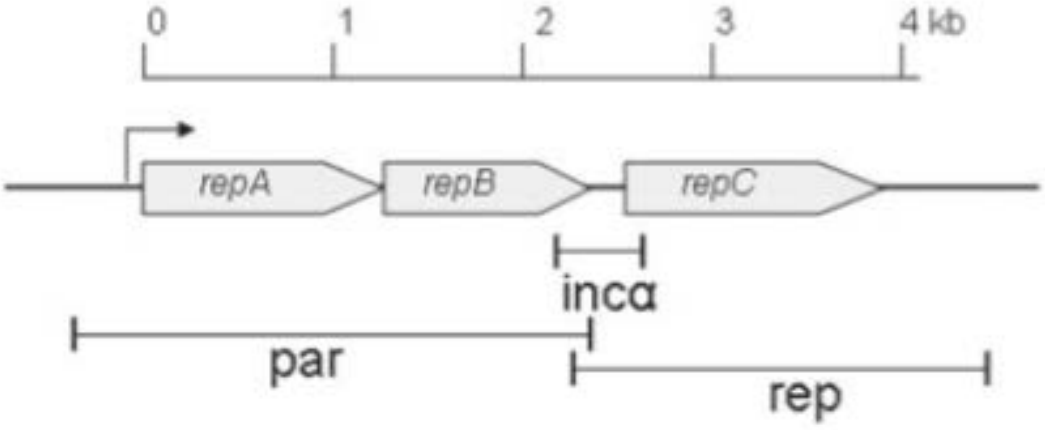


**Figure 1.3**: The *repABC*-type replication and partitioning operon indicating the orientation, structure and order of the genes within this operon. The incompatibility region, incα, is located in between *repB* and *repC* integenic region. Par: partitioning, rep; replication. Figure taken from MacLellan *et al.* (2005).

An important component of *repABC*-type plasmids is inc-α, which mediates plasmid incompatibility. Plasmid incompatibility is the inability of two plasmids that share similar replication and/or partitioning systems to reside in the same cell independently, and that; in the absence of selection pressure one plasmid will be lost from the cell (Ramirez-Romero *et al.*, 2000; MacLellan *et al.*, 2005); (Figure 1.3). The inc-α encodes a small antisense RNA (ctRNA), which modulates *repC* levels and, in turn, replication of the plasmid (Mazur and Koper, 2012).

The phenomenon of plasmid incompatibility is commonly harnessed as a means of removing or "curing" a plasmid from the genome of a strain of interest, often by introducing a vector containing a replication and partitioning site that is

identical to that of the target plasmid. In doing so both the original plasmid and the introduced "curing" vector will contain identical incompatibility determinants causing these plasmids to be incompatible. By including a selectable marker (e.g. antibiotic resistance genes) on the introduced vector this can be selected for in place of the original plasmid, and a strain that is cured of the original plasmid can be obtained (Uraji *et al.*, 2002; Ni *et al.*, 2008; Liu *et al.*, 2012; diCenzo *et al.*, 2014).

Low-copy number plasmids can also be stabilised in a host cell by expression of plasmid-encoded toxin-antitoxin (TA) systems. TA systems are typically composed of two elements, a stable toxin and a labile antitoxin (Goeders *et al.*, 2014). Toxins are usually small proteins that act to inhibit DNA replication by blocking the action of the class II topoisomerase DNA replication enzyme, DNA gyrase, leading to cell death. Whereas, antitoxins function to inhibit, sequester or antagonize the toxin protein, in turn neutralising its effect (Goeders and Van Melderen, 2014). A balanced toxin to antitoxin ratio is essential for cell survival, as perturbation of this ratio can lead to excess of toxin, which may lead to cell death. In the study by diCenzo *et al.* (2014) the plasmid, pSymB, of *Sinorhizobium meliloti* (which contains a *repABC*-type system) was cured through the use of plasmid incompatibility, however, the plasmid TA systems had to be removed prior to successful removal of the plasmid itself.

The addiction phenomenon is a mechanism of ensuring the maintenance of plasmids in bacterial populations by the function of TA systems (Figure 1.4). In practise, the toxin protein is usually more stable than the antitoxin (Goeders and

Van Melderen, 2014). Thus, if a daughter cell does not inherit the plasmid, the antitoxin cannot be expressed, leaving the lingering toxin protein to freely inhibit the action of DNA gyrase, which leads to the death of plasmid-less cells (Goeders and Van Melderen, 2014). This system ensures that bacterial populations maintain their plasmids, which can be especially important for low-copy number plasmids to ensure their stability (Pinto *et al.*, 2012).

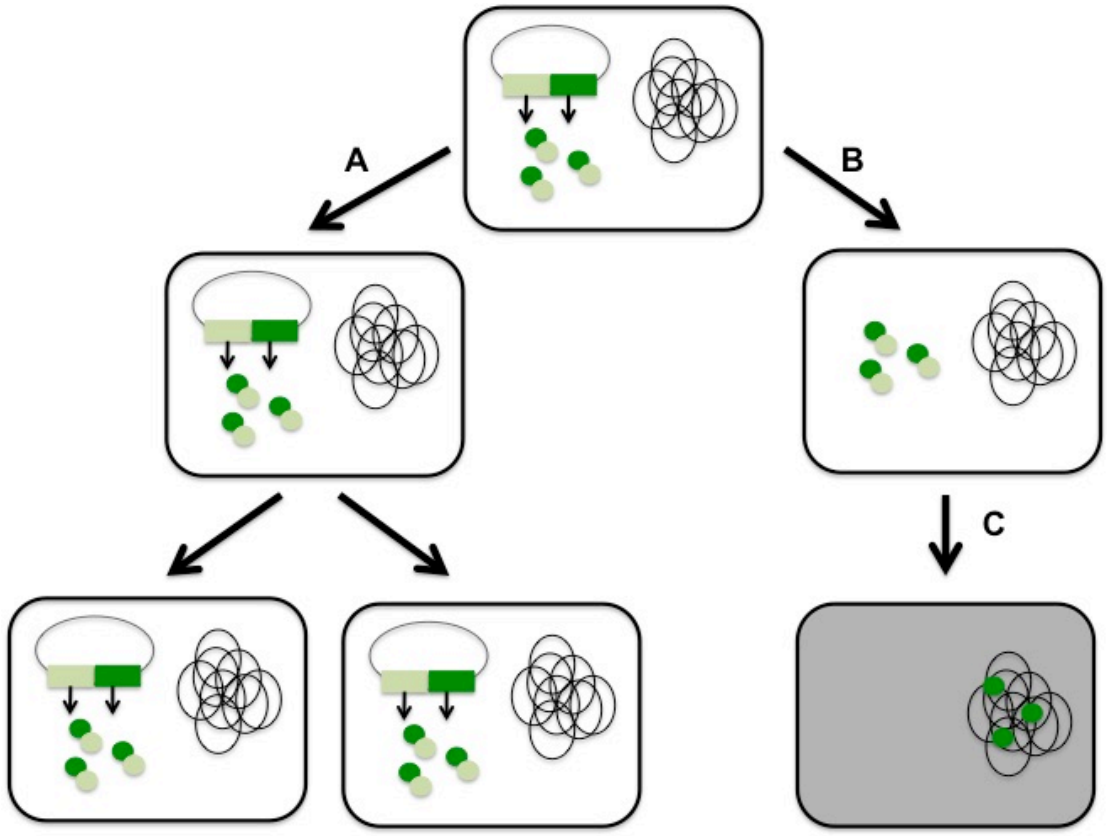


**Figure 1.4:** Example of the action of toxin-antitoxin (TA) systems to ensure stability of plasmids in bacterial populations. A) Daughter cell inheriting plasmid encoding TA system (light and dark green) grows and propagates normally to maintain plasmid. B) Daughter cell does not inherit the plasmid, but lingering TA proteins remain in their cytoplasm. C) Overtime the antitoxin (light green) is degraded while the toxin (dark green) is stable in the cell. The toxin then interacts with the DNA-gyrase, a class II topoisomerase, inhibiting its ability to replicate DNA, eventually leading to cell death. This mechanism leads to the selective killing of plasmid-less cells increasing the prevalence of plasmids throughout bacterial populations. Figure adapted from Goeders and Van Melderen (2014).

***Plasmid transmissibility***

Plasmid transfer is initiated at the *oriT* site where the relaxase creates a single-strand nick of the DNA. The 3' end of the nicked strand acts as a primer for rolling circle replication (RCR), which requires a helicase for the replacement of the transferred strand. Once the plasmid is transferred it is recircularlised within the recipient cell with help of the relaxase, and the complementary strand is then synthesised by the primase (Lawley *et al.*, 2003; Ding *et al.*, 2009). In order for the plasmid to transfer into the recipient cell a type IV secretion system (T4SS) is required, which functions in conjugation, DNA uptake and release, or protein translocation (Lawley *et al.*, 2003; Nelson *et al.*, 2015). To transfer plasmids via a T4SS, the donor cell produces a conjugation pillus, mating pore or channel complex, that brings the recipient cell into direct contact with the donor cell (Ding and Hynes, 2009; Nelson and Sadowsky, 2015).

In characterised rhizobial plasmids, the conjugal transfer system is encoded by cluster of *tra* genes, which are composed of two components: DNA transfer and replication (Dtr) as well as mating pair formation components (Mpf). The Dtr components include genes that encode for relaxases, helicases, primases and other accessory DNA processing proteins, as well as the origin of transfer (*oriT)* sequence (Lawley *et al.*, 2003; Ding and Hynes, 2009). While the Mpf component consists of genes coding for a T4SS (Lawley *et al.*, 2003; Ding and Hynes, 2009; Nelson and Sadowsky, 2015).

Plasmids can be self-transmissible or mobilisable, in which self-transmissible plasmid encode for both Dtr and Mpf components to undergo conjugal transfer.

Mobilisable plasmids, however, are only able to transfer in the presence of a self-transmissible plasmids that provides Mpf components, specifically coupling proteins that recognize the relaxsome *in trans* (Szpirer *et al.*, 2000; Ding and Hynes, 2009; Mazur and Koper, 2012). Theoretically, a mobilisable plasmid requires only an *oriT* site to be mobilised (Ding and Hynes, 2009). The specific mechanisms of self-transmissible and mobilisable plasmids are reviewed thoroughly by Ding and Hynes (2009).

### 1.4.2 Rhizobial Symbiosis Islands

In some genera of rhizobia, such as *Mesorhizobium spp.* or *Bradyrhizobium spp*., nodulation and nitrogen fixation genes are encoded on the bacterial chromosome. In *Mesorhizobium spp*., these chromosomal loci are capable of excision from the chromosome and transfer by conjugation to non-symbiotic *Mesorhizobium spp*. The transfer of these mobile genetic elements, referred to as "Symbiosis Islands" (SI), has been demonstrated in *Mesorhizobium loti* R7A, *M. loti* NZP2037, *M. loti* SU343, *M. ciceri* bv. biserrulae WSM1271, *M. australicum* WSM2073 and *M. opportunistum* WSM2075 (Sullivan and Ronson, 1998; Sullivan *et al.*, 2002; Nandasena *et al.*, 2006; Nandasena *et al.*, 2007; Haskett *et al.*, 2016c). Symbiosis Islands may be monopartite (such as for R7A) or tripartite, such as for WSM1271. The discovery of the WSM1271 Symbiosis Island was made following the introduction of the strain into a field test site in Northam (Australia).

*M.ciceri* bv. biserrulae WSM1271 was isolated from Sardinia (Italy) and was evaluated as an inoculant for the newly introduced pasture legume *Biserrula pelecinus* (Nandasena *et al.*, 2007). Initial tests at the Northam field site revealed

that uninoculated *B. pelecinus* sown into this site were not nodulated by resident rhizobia. This suggested that no compatible symbiotic rhizobia were present at this site, thus plants were subsequently inoculated with WSM1271 (Howieson *et al.*, 1995). Six years after the inoculation with WSM1271, the genetic diversity of 88 strains isolated from the field site were assessed and compared to the wild-type WSM1271 via ERIC and RPO1 fingerprinting (Nandasena *et al.*, 2007). This revealed that while the majority of isolates were genetically indistinguishable from WSM1271 (re-isolates), seven isolates yielded ERIC and RPO1 banding profiles different to that of WSM1271 indicating that they were genetically different to the original inoculant strain WSM1271. These newly evolved isolates were referred to as novel strains or isolates (Nandasena *et al.*, 2007).

The shoot dry weights of *B. pelecinus* inoculated with seven re-isolates and seven novel isolates were compared to plants inoculated with wild-type WSM1271 Nandasena *et al.* (2007); (Figure 1.5). While all re-isolates were as effective as WSM1271 at fixing-nitrogen with *B. pelecinus* as measured by shoot dry weight production, the novel isolates were either less effective than WSM1271 or yielded shoot dry weights that were not different to the nitrogen starved controls ($Fix^-$) (Nandasena *et al.*, 2007). Further comparison of acetylene reduction of WSM1271 and two novel isolates (WSM2073 and WSM2075) with *B. pelecinus* confirmed that while WSM2073 is partially effective (fixing 48% of the WSM1271 rate), WSM2075 is completely $Fix^-$ on this host (Goel, 2009). Furthermore, light and electron microscopy of *B. pelecinus* nodules showed that while WSM1271 and WSM2073 both elicited pink indeterminate nodules with a red colouration, WSM2075 nodules were small and white consistent with their

Fix$^-$ dry weight and acetylene reduction phenotype (Goel, 2009). Thus, while WSM2073 and WSM2075 were both able to nodulate *B. pelecinus*, their symbiotic phenotype differed markedly to that of the original strain, WSM1271. A study of the taxonomy of the novel strains WSM2073 and WSM2075 confirmed these strains represented two different novel *Mesorhizobium* species, which were subsequently named *M. australicum* WSM2073 and *M. opportunistum* WSM2075 (Nandasena *et al.*, 2009).

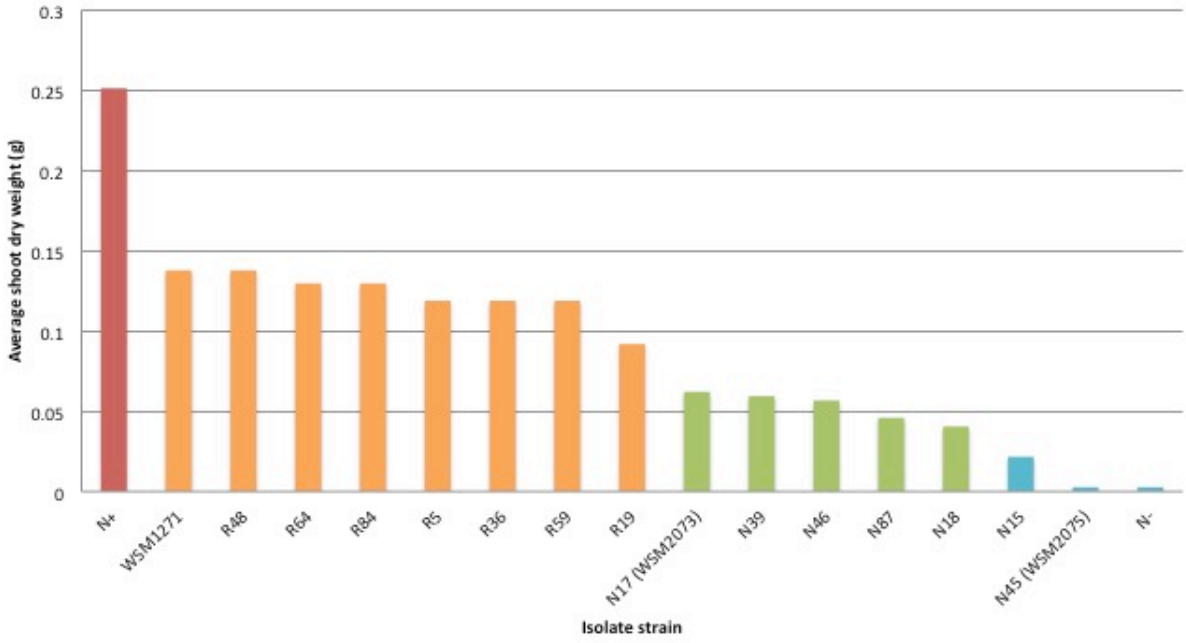


**Figure 1.5:** Mean shoot dry weights of re-isolates (R5, R19, R36, R48, R59, R64, R84) and novel strains (N17 or WSM2073, N39, N46, N87, N18, N15, N45 or WSM2075). The colours of bars indicate significant differences at the level of 0.01 calculated following square root transformations of the average shoot dry weight (g). Shoot dry weights of plants inoculated with re-isolates (R5, R19, R36, R48, R59, R64, R84) were not significantly different to WSM1271, but were less than nitrogen feed control group (N+). Novel strains N17 (WSM2073), N39, N46, N87, and N18 all yielded dry weights less than WSM1271 while N15 and N45 (WSM2075, shown in blue) were equivalent to the N-starved control (N-). Figure adapted from Nandasena *et al.* (2007).

Full sequencing of the genomes of WSM1271, WSM2073 and WSM2075 genomes was subsequently performed (Reeve *et al.*, 2013a; Reeve *et al.*, 2013b;

Nandasena *et al.*, 2014) and comparison of the genome sequences revealed three identical regions were shared between the three strains (Haskett *et al.*, 2016c). The three regions were found to form part of a novel mobile tripartite Symbiosis Island of ~475-kb named ICE*Mc*Sym$^{1271}$ (Haskett *et al.*, 2016c). The perfect sequence identity of the SI across the three strains therefore confirmed that WM2073 and WSM2075 had received their symbiosis genes from the original inoculant strain WSM1271.

Interestingly, this analysis also revealed that WSM1271 harboured an additional plasmid (pMESCI01, 425-kb) representing a total of 6.36% of the WSM1271 genome (Nandasena *et al.*, 2014). Crucially, this plasmid is absent from the genomes of both novel microsymbionts WSM2073 and WSM2075 (Reeve *et al.*, 2013a; Reeve *et al.*, 2013b; Nandasena *et al.*, 2014). The potential role of pMESCI01 in WSM1271 is currently unknown, but an investigation of this plasmid could reveal whether it plays a symbiotic role.

## 1.5 Project aims

*Mesorhizobium spp.* are known to harbour their symbiotic genes on mobile genetic elements referred to as Symbiosis Islands. In the case of *M. ciceri* bv. biserrulae WSM1271, this island (ICE*Mc*Sym$^{1271}$) has been transferred to different species of apparently non-symbiotic soil *Mesorhizobium* (*M. australicum* WSM2073 and *M. opportunistum* WSM2075). Previous studies of the *Mesorhizobium-Biserrula* symbiosis have revealed a difference in effectiveness between *M. ciceri* bv. biserrulae WSM1271, *M. australicum* WSM2073 and *M. opportunistum* WSM2075 on *B. pelecinus.* Specifically, while WSM2073 is

partially effective in nitrogen fixation with *B. pelecinus*, WSM2075 nodulates but does not fix nitrogen on this host (Nandasena *et al.*, 2007; Goel, 2009). Full genome sequencing has revealed that all three strains share identical Symbiosis Islands, however, only WSM1271 carries an additional plasmid (pMESCI01, 425-kb). The genetic complement of pMESCI01 and its potential transmissibility to other strains remains undescribed. However, at least one strain of *Mesorhizobium* (*M. huakii* 7653R) is known to harbour essential plasmid-encoded symbiosis genes (Peng *et al.*, 2014; Wang *et al.*, 2014). It is therefore possible that pMESC01 may carry additional determinants required for highly effective symbiotic interactions with *B. pelecinus*. Understanding the role and transmissibility of pMESCI01 would provide important information regarding *Mesorhizobium spp.* plasmids.

The aims of this thesis are therefore:

1. Conduct a bioinformatic analysis of the genetic cargo harboured by pMESCI01, with a focus on genes with possible symbiotic roles.
2. Construct a marked version of pMESCI01 to assess whether the plasmid is self-transmissible.
3. Develop a pMESCI01-cured version of *M. ciceri* bv. biserrulae WSM1271.
4. Assess the symbiotic phenotype of pMESCI01-cured WSM1271 on *B. pelecinus*.

# 2. Materials and Methods

## 2.1 Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 2.1. *Mesorhizobium spp.* strains were routinely grown in TY (Beringer, 1974), or Rhizobium Defined Medium (RDM) (Ronson *et al.*, 1979) supplemented with vitamins (1 μg/ml each of thiamine and nicotinate, and 20 ng/ml biotin with either 2% (w/v) glucose or 5% (w/v) sucrose and incubated at 28°C. *Escherichia coli* strains were routinely grown in LB (Bertani, 1951) and incubated at 37°C. All media were adjusted to a pH of 7.0 and when solid media was required, agar (Grade A) was added to a final concentration of 1.5% (w/v). Broth cultures were grown on gyratory shakers set at 250 rpm. Antibiotics were used at the following concentrations (μg/ml) when necessary: Neomycin (250) or Kanamycin (50), Gentamycin (20), Streptomycin (100 or 200), Spectinomycin (200), Nitrofurantoin (10), and Ampicillin (100). Selection of transformants within *E. coli* ST18 required the addition of ALA at 60 μg/ml. All antibiotics were purchased from Sigma-Aldrich.

**Table 2.1:** Bacterial strains and plasmids used in this study

| Strain/Plasmid | Features | Reference |
|---|---|---|
| **Strains** | | |
| *Mesorhizobium spp.* | | |
| *M. ciceri* bv. biserrulae WSM1271 | Wild type strain harbours pMESCI01; thiamine auxotroph; $Sm^R$, $Nf^R$ | (Howieson *et al.*, 1995) |
| WSM2073Nm | *M. australicum* WSM2073 derivative harbouring pPROBE; $Nm^R$ | T. Haskett, Murdoch University |
| RB6.5b | WSM1271 derivative harbouring pMESCI01 with SXO insertion of pJQ(Ω-Sp/Sm); thiamine auxotroph; $Gm^R$, $Sp^R$, $Sm^R$, $Nf^R$, $Suc^S$ | This study |
| RB6.6b | WSM1271 derivative harbouring pMESCI01 with SXO insertion of pJQ(Ω-Sp/Sm); thiamine auxotroph; $Gm^R$, $Sp^R$, $Sm^R$, $Nf^R$, $Suc^S$ | This study |
| RB7.2 | WSM1271 derivative harbouring pMESCI01::Ω-Sp/Sm with cassette insertion at coordinates 30,104-30,895; thiamine auxotroph; $Sp^R$, $Sm^R$, $Nf^R$ | This study |

| Strain/Plasmid | Features | Reference |
|---|---|---|
| RB8.3 | WSM1271 derivative harbouring pMESCI01::Ω-Sp/Sm with cassette insertion at coordinates 30,104-30,895;; thiamine auxotroph; $Sp^R$, $Sm^R$, $Nf^R$ | This study |
| RB3.1 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.2 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.3 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.4 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.5 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.6 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.7 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.8 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.9 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.10 | WSM1271 derivative cured of pMESCI01; harbouring pSRK*repABC*; thiamine auxotroph; $Sm^R$, $Nf^R$, $Km^R$, $Suc^S$ | This study |
| RB3.1a | WSM1271 derivative cured of pMESCI01; thiamine auxotroph; $Sm^R$, $Nf^R$ | This study |
| RB3.2a | WSM1271 derivative cured of pMESCI01; thiamine auxotroph; $Sm^R$, $Nf^R$ | This study |
| *Escherchia coli* | | |
| ST18 | S17 *λpir ΔhemA*; ALA auxotroph | (Thoma *et al.*, 2009) |
| DH5α | F- *endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1 Δ(argF-lacZYA)* | Bioline |
| RB6.5 | ST18 derivative harbouring pJQ(Ω-Sp/Sm); $Sp^R$, $Sm^R$, $Gm^R$, $Suc^S$ | This study |
| RB6.6 | ST18 derivative harbouring pJQ(Ω-Sp/Sm); $Sp^R$, $Sm^R$, $Gm^R$, $Suc^S$ | This study |
| T18 | ST18 derivative harbouring pSRK*repABC* | This study |
| **Plasmids** | | |
| pJQ200SK-*SacB* | pACYC containing P15A origin of replication; $Gm^R$, $Suc^S$ | (Quandt *et al.*, 1993) |
| pJET(Ω-Sp/Sm) | pJET containing Ω-Sp/Sm; $Sp^R$, $Sm^R$ | T. Haskett, Murdoch Univrsity |
| pSRKKm-*SacB* | pSRK containing pBBR origin of replication; $Km^R$, $Suc^S$ | (Haskett *et al.*, 2016c) |
| pSRK*repABC* | Plasmid-curing vector; pSRKKm-*SacB* containing *repABC* region of pMESCI01; $Km^R$, $Suc^S$ | This study |
| pJQ(Ω-Sp/Sm) | Plasmid marking vector; pJQ200SK-*SacB* containing Ω-Sp/Sm; $Sp^R$, $Sm^R$, $Gm^R$, $Suc^S$ | This study |
| pMESCI01::Ω-Sp/Sm | Marked pMESCI01 with Ω-Sp/Sm cassette insertion at coordinates 30,104-30,895; $Sp^R$, $Sm^R$, $Nf^R$ | This study |

Abbreviations: $Sp^R$, $Sm^R$, $Gm^R$, $Km^R$, $Nm^R$, $Nf^R$ denote spectinomycin, streptomycin, gentamycin, kanamycin, neomycin, and nitrofurantoin resistance, respectively. $Suc^S$ depicts sucrose sensitivity.

## 2.2 Bioinformatic analysis

All genome and gene sequences were obtained from Integrated Microbial Genomes online database (IMG) (Markowitz *et al.*, 2012) or via Geneious version 9.0.5 (http://www.geneious.com/) (Kearse *et al.*, 2012), using National Centre for Biotechnology (NCBI) database (Sayers *et al.*, 2011). Nucleotide sequence analysis and amino acid sequence analysis were performed with BLASTN or BLASTP, requiring an 80% or higher sequence alignment to be considered orthologous. Nucleotide sequences were translated for BLASTP analysis using ExPASY tool (http://www.expasy.org) (Artimo *et al.*, 2012). Gene and operon alignments were performed using Geneious alignment tool as well as BRIG alignment program version 0.95 (http://sourceforge.net/projects/brig/) (Alikhan *et al.*, 2011) via BLASTX database (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) (Camacho *et al.*, 2009) as stated. Clusters of orthologous groups of proteins (COG) analysis was conducted on pMESCI01 genome to obtain protein function information provided on the IMG database. The TAfinder online program was used to locate putative Type II toxin-antitoxin (TA) systems on pMESCI01 (Version 2.0; Microbial Bioinformatics Group, Shanghai Jiaotong University [http://202.120.12.133/TAfinder/report.php?job_id=xINgjLLD0d]). Origin or transfer (*oriT*) sequence search on pMESCI01 was conducted using TUBIC BLAST (http://tubic.tju.edu.cn) (Gao *et al.*, 2013).

## 2.3 General Molecular Methods

Agarose gel electrophoresis was performed using TAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA) as described by (Sambrook *et al.*, 1989) unless

otherwise stated. These gels were post-stained with Ethidium Bromide (10 mg/ml) (Sigma-Aldrich) in TAE buffer and de-stained in TAE buffer. All gels were visualised using Molecular Imager® Gel Doc™ XR+ System (BioRad). WSM1271 was cultured in 5 ml TY broth cultures supplemented with appropriate antibiotics and harvested via centrifugation prior to extraction of genomic DNA. Genomic DNA extractions were performed using PrepMan® Ultra Sample Preparation Reagent (Applied Biosystems) as per supplier recommendations. DNA quantitation and purity was determined using NanoDrop™ 2000c (Thermo-Fisher Scientific) set at $A_{280nm}$. DNA purity assessment was conducted and considered pure when absorbance (A) ratios of $A_{260nm/280nm}$ were greater than 1.8 and $A_{230nm/260nm}$ was between 2.0-2.2 (Sambrook *et al.*, 1989). Restriction enzymes *Sac*I-HF (cat #R3156S), *Not*I-HF (cat #R3189S), *Xba*I (cat #R0145S), *Sal*I-HF (cat #R3138S), *Xho*I (cat #R0146S), *Bam*HI-HF (cat #R3136S), and *Eco*RI-HF (cat #R3101S) were sourced from New England Biolabs (USA), and were used as per supplier recommendations. The custom oligonucleotide primer sequences were designed using primer design function of Geneious 9.0.5, and all were synthesized through Integrated DNA Technologies (IDT Singapore) (Table 2.2). PCR amplifications were conducted using the MyCycler™ Thermal Cycler System (BioRad).

The following PCR cycling conditions were used for this study:

**PCR 1:** 98°C for 30 s (x 1); 98°C for 10 s, 70°C for 30 s, 72°C for 90 s (x 5); 98°C for 30 s, 72°C for 30 s (x 30); and 72°C for 5 min and hold at 14°C (x 1).

**PCR 2:** 94°C for 5 min (x 1); 94°C for 30 s, 59°C for 20 s, 70°C for 5 min (x 25); and 70°C for 5 min and hold at 14°C (x 1).

**PCR 3:** 94°C for 5 min (x 1); 94°C for 30 s, 57°C for 20 s, 70°C for 5 min (x 25); and 70°C for 5 min and hold at 14°C (x 1).

**PCR 4:** 98°C for 30 s (x 1); 98°C for 10 s, 72°C for 2 min 30 s (x 35); and 72°C for 5 min and hold at 14°C (x 1).

**PCR 5:** 94°C for 5 min (x 1); 94°C for 30 s, 59°C for 20 s, 70°C for 50 s (x 30); and 70°C for 3 min and hold at 14°C (x 1).

**PCR 6:** 94°C for 5 min (x 1); 94°C for 30 s, 53°C for 30 s, 72°C for 1 min (x 30); and 72°C for 5 min and hold at 14°C (x 1).

**Table 2.2:** Oligonucleotides used in this study

| | Name | Oligonucleotide (5' - 3' orientation) | Reference |
|---|---|---|---|
| **1** | pMESCI01 F P1 | CGCGTAATCAAAGGCTCACG | This study |
| **2** | pMESCI01 R P1 | TGTGCTCTGACCAATCCGAC | This study |
| **3** | pMESCI01 F P2 | CGATGGAAGCAGCATGAACG | This study |
| **4** | pMESCI01 R P2 | ACTTTGATCTTCACCGCCGT | This study |
| **5** | RepABC F BamHI | ATCAGGGATCCGTTGACCTCCGCATGCAAAC | This study |
| **6** | RepABC R XbaI | ATCAG TCTAGAGTCAATCTCACCAGGGCCAG | This study |
| **7** | pMESCI01_UpS_F_SacI | ATCAGGAGCTCGCGAACGGCACCTTCGAG | This study |
| **8** | pMESCI01_UpS_R_NotI | ATCAGGCGGCCGCCGGCATGAAACTCATCCGAATG | This study |
| **9** | pMESCI01_DnS_F_XbaI | ATCAGTCTAGACCCGTTCGCCAGTAGCTTCAAT | This study |
| **10** | pMESCI01_DnS_R_SalI | ATCAGGTCGACTGCGAGATGACAGTTTCGTCCTCAAA | This study |
| **11** | pMESCI01_DXO_StSp_F1 | GGCTGATCCGACTTTCACCA | This study |
| **12** | pMESCI01_DXO_StSp_R1 | CGCGCAGATCAGTTGGAAGA | This study |
| **13** | pMESCI01_DXO_StSp_F2 | GTTTGTTCGCCCAGCTTCTG | This study |
| **14** | pMESCI01_DXO_StSp_R2 | CCGTTTCAGATCCGGACCAA | This study |
| **15** | pJQ200SKR | ACTCACTATAGGGCGAATTGGG | T. Haskett, Murdoch University |
| **16** | pJQ200SKF | AGCGGATAACAATTTCACACAGG | T. Haskett, Murdoch University |
| **17** | MesGMCOF | GCCAAATGGTCGACGCTCTA | T. Haskett, Murdoch University |
| **18** | MesGMCOR | GTCCGACACGAACAGGTTCT | T. Haskett, Murdoch University |
| **19** | PA | AGAGTTTGATCCTGGCTCAG | S. de Meyer, Murdoch University |
| **20** | PH | AAGGAGGTGATCCAGCCGCA | S. de Meyer, Murdoch University |
| **21** | pSRK primer | TGGCCGATTCATTAATGCAGC | T. Haskett, Murdoch University |
| **22** | pSRKKM-SacB(repABC) F1 | GTCACGACGTTGTAAAAC | This study |
| **23** | 16S638R | AATTTCACCTCTACACTCGG | S. de Meyer, Murdoch University |

## 2.4 Construction of pJQ(Ω-Sp/Sm) plasmid-marking vector

### 2.4.1 Preparation of pMESCI01 upstream and downstream regions

WSM1271 was cultured in TY broth supplemented with appropriate antibiotics at 28°C for six days. Cells were harvested through centrifugation and plasmid DNA was extracted by FavorPrep™ Plasmid Extraction kit (Favorgen Biotech Corp, cat #FAPDE300) as per supplier recommendations. PCR amplification of selected adjacent 1-kb upstream and downstream intergenic regions of pMESCI01, located at 28,622-29,621 (upstream) and 29,266-30,655 (downstream) between convergent genes with hypothetical protein function (Mesci_6067-6068) (Figure 2.1). Amplification was performed with primers 7 and 8 with *Sac*I and *Not*I tails at the 5' ends, respectively, and downstream primers 9 and 10 with *Xba*I and *Sai*I tails at 5' ends, respectively. PCR reactions were prepared in 50 µl volumes, containing 25 µl 2X Phusion® High Fidelity DNA polymerase (New England Biolabs, cat #M0531S), 1.0 µl of template plasmid DNA, 0.5 µM of each primer, and 1.5 µl of DMSO (New England Biolabs, cat #B0515). PCR cycling conditions were as described for PCR 1 (Section 2.3). The PCR products were electrophoresed and the molecular weight of products were determined by comparison to a 1-kb NEB ladder (New England Biolabs, cat #N3232L), with gels run at 80 V with 1% (w/v) Agarose. A 40 µl aliquot of amplified upstream PCR products was digested with 20 units each of *Sac*I and *Not*I restriction enzymes, with 10X Cutsmart™ Buffer (New England Biolabs, cat #B7204S), in a 50 µl total reaction. A 40 µl aliquot of amplified downstream PCR products was digested using 20 units each of *Xba*I and *Sal*I restriction enzymes, with 10X Cutsmart™ Buffer (New England Biolabs, cat #B7204S), in a 50 µl total

reaction. Both reactions were incubated in a water bath at 37°C for 4.5 h prior to heat-inactivation at 65°C for 20 min.

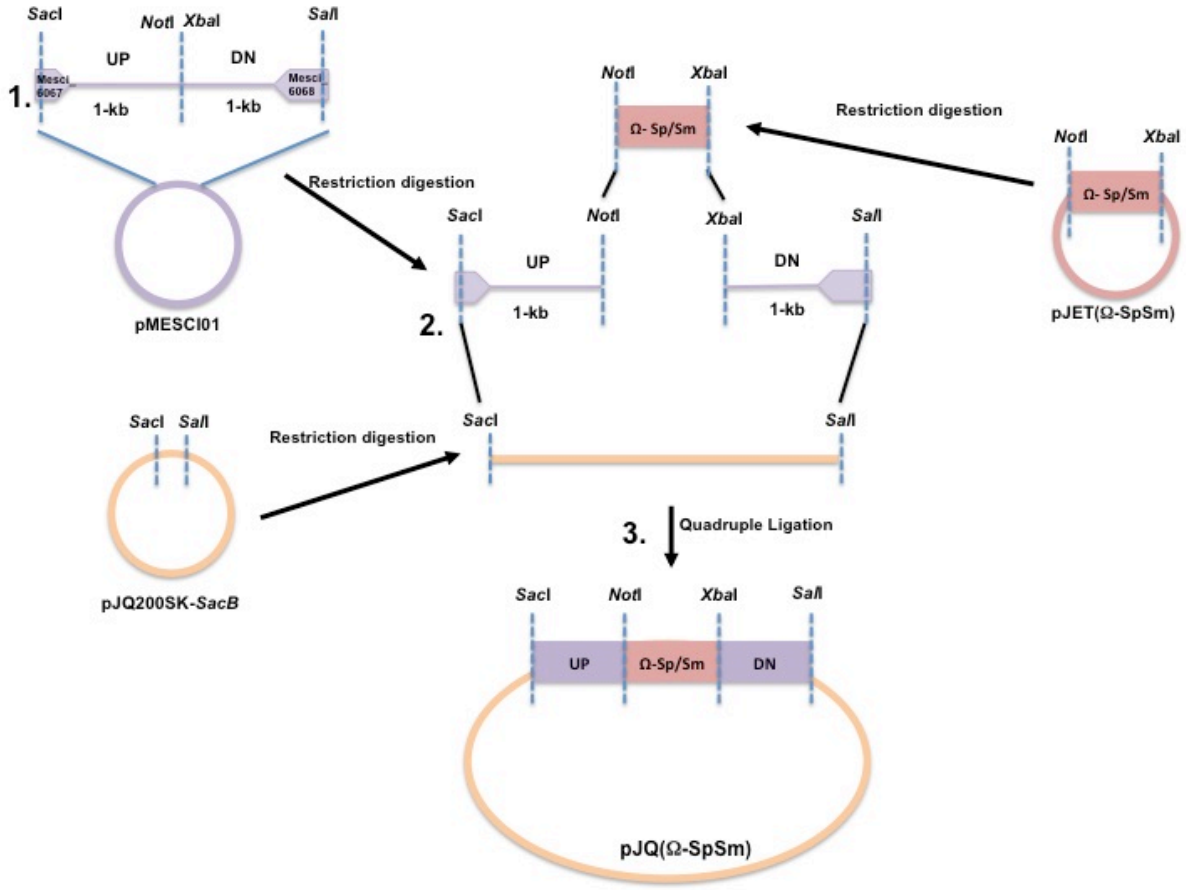


**Figure 2.1:** Approach to construct plasmid-marking vector, pJQ(Ω-Sp/Sm). 1) PCR amplification of the adjacent 1-kb upstream (UP) and 1-kb downstream (DN) regions of pMESCI01 indicating the location of the restriction tails. 2) Restriction digestion of the upstream and downstream fragments, pJET(Ω-Sp/Sm, and pJQ200SK-SacB with *Sac*I, *Not*I, *Xba*I, and *Sal*I as indicated. 3) Quadruple ligation of the four digested fragments to construct the plasmid-marking vector, the pJQ(Ω-Sp/Sm).

### 2.4.2 Preparation of pJET(Ω-Sp/Sm)

*E. coli* ST18 carrying pJET(Ω-Sp/Sm) was cultured in LB broth supplemented with $ALA_{60}$ and appropriate antibiotics at 37°C overnight. Cells were then harvested through centrifugation, and plasmid DNA was extracted using the

FavorPrep™ Plasmid Extraction kit (Favorgen Biotech Corp, cat #FAPDE300) following supplier recommendations. Digestion of 1.5 µg of pJET(Ω-Sp/Sm) was performed with 20 units each of *Xba*I and *Not*I with 10X Cutsmart™ Buffer (New England Biolabs, cat #B7204S), in a total reaction volume of 50 µl. Digestion and heat inactivation were performed as per Section 2.4.1.

### 2.4.3 Preparation of pJQ200SK-*SacB*

*E.coli* ST18 carrying pJQ200SK-*SacB* plasmid was cultured in LB broth supplemented with appropriate antibiotics at 37°C overnight. Cells were then harvested through centrifugation and plasmid DNA was extracted as described previously in Section 2.4.2. Digestion of 3 µg of pJQ200SK-*SacB* was performed with 20units of *Sac*I and 20 units of *Sal*I with 10X Cutsmart™ Buffer (New England Biolabs, cat #B7204S), in a total reaction volume of 50 µl. Digestion and heat inactivation were performed as per Section 2.4.1.

### 2.4.4 Quadruple ligation with pMESCI01 upstream and downstream fragments, and fragments of pJET(Ω-Sp/Sm) and pJQ200SK-*SacB*

Multiple ligation reactions were prepared with ratios of upstream region to downstream region to pJET(Ω-Sp/Sm) to pJQ200SK-*SacB* of 1:1:5:5, respectively. Preparation of multiple 10 µl quadruple ligation reactions with 1 µl NEB T4 Ligase (New England Biolabs, cat #M0202S) and 1 µl 10X Ligase Buffer (New England Biolabs, cat #B0202S), with 1 µl digested upstream product, 1 µl digested downstream product, 1-2 µl pJQ200SK-*SacB*, and 1-2 µl of pJET(Ω-Sp/Sm). These reactions were incubated at 16°C overnight and heat inactivated for 10 min at 65°C prior to transformation into DH5α competent cells.

### 2.4.5 Transformation into *E.coli* and confirmation

A total of 5 µl of the quadruple ligation reaction mixture was added to 50 µl of DH5α competent cells and was then chilled on ice for 30 mins, incubated at 42°C for 30 s, prior to a further 2 min on ice. A total of 1 ml of SOC broth was added and the suspension was then incubated on a gyratory shaker at 250 rpm at 37°C for 2 h. Successful transformants were selected after growth overnight onto LB agar supplemented with $Gm_{20}$ and $Sp_{100}$. Colonies were screened for the presence of the pJQ(Ω-Sp/Sm) vector by colony PCR, using primers 15 and 16 (Table 2.2), which bind to regions flanking the cloning site of pJQ(Ω-Sp/Sm). PCR reactions were prepared in 10 µl volumes, containing 2X GoTaq® Green Mastermix (Promega cat #M712B), and 0.5 µM of each primer. PCR cycling conditions as are described for PCR 2 (Section 2.3). The PCR products were electrophoresed and the molecular weight of products were determined by comparison to a 1-kb NEB ladder (New England Biolabs, cat #N3232L), with gels run at 90V with 1% Agarose.

PCR confirmed transformant colonies containing the pJQ(Ω-Sp/Sm) vector were subjected to restriction analysis. Colonies were cultured in LB broth and plasmid DNA extracted as previously described (Section 2.4.2) prior to digestion with 20 units of *Xho*I restriction enzyme and 10 X Buffer E (Promega, cat #2005A) in a total reaction volume of 25 µl at 37°C for 2 h. The digested products were electrophoresed and the molecular weight of products were determined by comparison to a 1-kb NEB ladder (New England Biolabs, cat #N3232L), with gels run at 90V with 1% Agarose. The plasmid DNA of confirmed DH5α transformants was extracted as previously described in Section 2.4.3, and

transformed into *E.coli* ST18 competent cells prior to subsequent conjugation expreiments. A 1 µl aliquot of plasmid DNA was added to 100 µl of ST18 competent cells with 1 µl of DMSO, and was then chilled on ice for 30 min, incubated at 42°C for 90 s, prior to a further 2 min on ice. A total of 400 µl of SOC broth supplemented with $ALA_{60}$ was added to the suspension and was then incubated on a gyratory shaker at 250 rpm at 37°C for 1 h. The transformation reaction was spread in dilutions onto LB supplemented with $ALA_{60}$ and appropriate antibiotics, and then incubated overnight at 37°C.

## 2.5 Conjugative transfer of pJQ(Ω-Sp/Sm) plasmid-marking vector into WSM1271

The plasmid-marking vector, pJQ(Ω-Sp/Sm), was mobilised from *E. coli* ST18 into WSM1271 via a biparental mating. WSM1271 was grown in 15 ml TY both supplemented with appropriate antibiotics and grown at 28°C for 3 days and harvested at stationary phase ($OD_{600nm}$ ~2). *E.coli* ST18 transformants were grown in 5 ml LB supplemented with $ALA_{60}$ and appropriate antibiotics at 37°C for 4 h, and harvested at log phase ($OD_{600nm}$ ~0.3). Cultures were harvested by centrifugation (5,000 xG, 5min), washed with 1 ml of sterile 0.89% (w/v) saline before being resuspended in 600 µl (WSM1271) or 200 µl (ST18) of sterile 0.89% (w/v) saline. A 50 µl aliquot each of WSM1271 and ST18 suspensions were mixed and spotted onto TY agar supplemented with $ALA_{60}$ and incubated overnight at 28°C. A 50 µl aliquot of WSM1271 and ST18 transformants were separately spotted onto TY agar supplemented with $ALA_{60}$ and incubated over at 28°C overnight as controls. Mating spots were then resuspended in 200 µl of

sterile 0.89% (w/v) saline, and cells were spread in dilutions onto G/RDM supplemented with $Gm_{20}$ and $Sp_{200}$ and incubated at 28°C.

### 2.5.1 Confirmation of single and double crossover transconjugants

Colonies were confirmed for the presence of single crossover (SXO) transconjugants (pMESCI01:Ω-Sp/Sm) following biparental matings by colony PCR, using primers 11 and 12 (Table 2.2), designed to flank the cloning site of the "upstream region" by capturing a region of pMESCI01 upstream to the cloning site and a region of the Ω-Sp/Sm cassette that is downstream to the cloning site. PCR reactions were prepared as described for Section 2.4.5, using PCR cycling conditions as described for PCR 3 (Section 2.3), and products were electrophoresed as per Section 2.4.5.

Confirmed SXO colonies were then grown in TY broth supplemented with $Sp_{100}$ to force for double crossover (DXO), at 28°C at 250 rpm for 4 days. These were harvested by centrifugation resuspended in sterile 0.89% (w/v) saline, and subsequently spread onto sucrose containing RDM (S/RDM) supplemented with $Sp_{200}$ in dilutions to select for DXO colonies. Colonies were then confirmed for the presence of pMESCI01::Ω-Sp/Sm by colony PCR, using primers 11 and 12 (Table 2.2) designed as previously discussed (Section 2.5.1), and primers 13 and 14 designed to flank the cloning site of the "downstream region" by capturing a region of the Ω-Sp/Sm cassette upstream to the cloning site and a region of pMESCI01 that is downstream to the cloning site. PCR reactions were performed and the products electrophoresed as described earlier (Section 2.4.5), using PCR cycling conditions as described for PCR 3 (Section 2.3). Confirmed DXO colonies

were then streaked onto G/RDM supplemented with $Sp_{200}$, as well as G/RDM supplemented with $Sp_{200}$ and $Gm_{20}$ to ensure for loss of $Gm^R$, and incubated at 28°C for 4 days.

## 2.6 Conjugative transfer of marked plasmid pMESCI01::Ω-Sp/Sm into WSM2073

Mobilisation of pMESCI01::Ω-Sp/Sm into WSM2073Nm was conducted via biparental mating with growth of donor WSM1271 to both log ($OD_{600nm}$ ~0.3) or stationary phase ($OD_{600nm}$ ~2), and growth of recipient WSM2073Nm to stationary phase ($OD_{600nm}$ ~2) in 15ml (WSM1271) or 5ml (WSM2073Nm) TY broth supplemented with appropriate antibiotics at 28°C. Cultures were harvested by centrifugation (5,000 x G, 5 min), washed with 1 ml of sterile 0.89% (w/v) saline before being resuspended in 600 µl (WSM2073Nm) or 200 µl (WSM1271) of sterile 0.89% (w/v) saline. A 50 µl aliquot each of WSM1271 and WSM2073Nm suspensions were mixed and spotted onto TY agar and incubated at 28°C overnight. A 50 µl aliquot of WSM1271 transformants and WSM2073Nm were separately spotted onto TY agar and incubated over at 28°C overnight as controls. Mating spots were then resuspended in 1 ml of sterile 0.89% (w/v) saline with 15% (w/v) glycerol solution, and cells were spread in dilutions onto G/RDM supplemented with $Sp_{200}$ and $Nm_{250}$ and incubated at 28°C.

### 2.6.1 Antibiotic screening of WSM2073Nm

Screening of WSM2073Nm was subsequently conducted due to its growth with $Sp_{200}$. WSM2073Nm was grown in 15 ml TY broth supplemented with appropriate antibiotics and harvested by centrifugation (5,000 x G, 5-15 min) when in stationary phase ($OD_{600nm}$ ~2). Following harvesting cells were

resuspended in 1 ml of sterile 0.89% (w/v) saline and 10-fold dilutions of the resuspended cells were spread onto G/RDM supplemented with $Sp_{200}$ and $Nm_{250}$; $Sp_{200}$, $Nm_{250}$, and $Sm_{200}$; or $Nm_{250}$. Subsequent cell numbers were counted and the number of cells/ml was determined.

## 2.7 Construction of pSRK*repABC* plasmid-curing vector

### 2.7.1 Preparation of pMESCI01 *repABC* regions

WSM1271 was cultured in TY broth and plasmid DNA extracted as previously described (Section 2.4.1) prior to PCR amplification of a 5.274-kb region of pMESCI01 containing the *repABC* region of pMESCI01 (Mesci_6410-6412), designed to capture any upstream promoter of this region, was conducted (Figure 2.2). Amplification was performed using primers designed with selected restriction site tails at the 5' end, with primers 5 and 6 (Table 2.2) containing *Bam*HI and *Xba*I on the forward and reverse primers, respectively. PCR reactions were prepared in 25 µl volumes, containing 12.5 µl 2X Phusion® HF DNA polyermase (New England Biolabs, cat #M0531S), 0.5 µl of template plasmid DNA, and 0.5 µM of each primer. PCR cycling conditions are as described for PCR 4 (Section 2.3). The PCR products were electrophoresed and the molecular weight of products were determined by comparison to a 1-kb Axygen® Scientific ladder (Fisher Scientific, cat #14-223-051), with gels run at 70V with 1.2% Agarose; and stained in SYBR® Safe DNA Gel Stain (Invitrogen). Correct PCR product was then purified using FavorPrep™ Gel Purification kit (Favorgen Biotech Corp, cat #FAGCK001-1) as per supplier recommendations. A 40 µl aliquot of amplified PCR products was digested using 20 units of *Bam*HI and 20 units of *Xba*I restriction enzymes, with 10X Cutsmart™ Buffer (New England

Biolabs, cat #B7204S), in a 50 µl total reaction. Digestion and heat inactivation were performed as per Section 2.4.1.

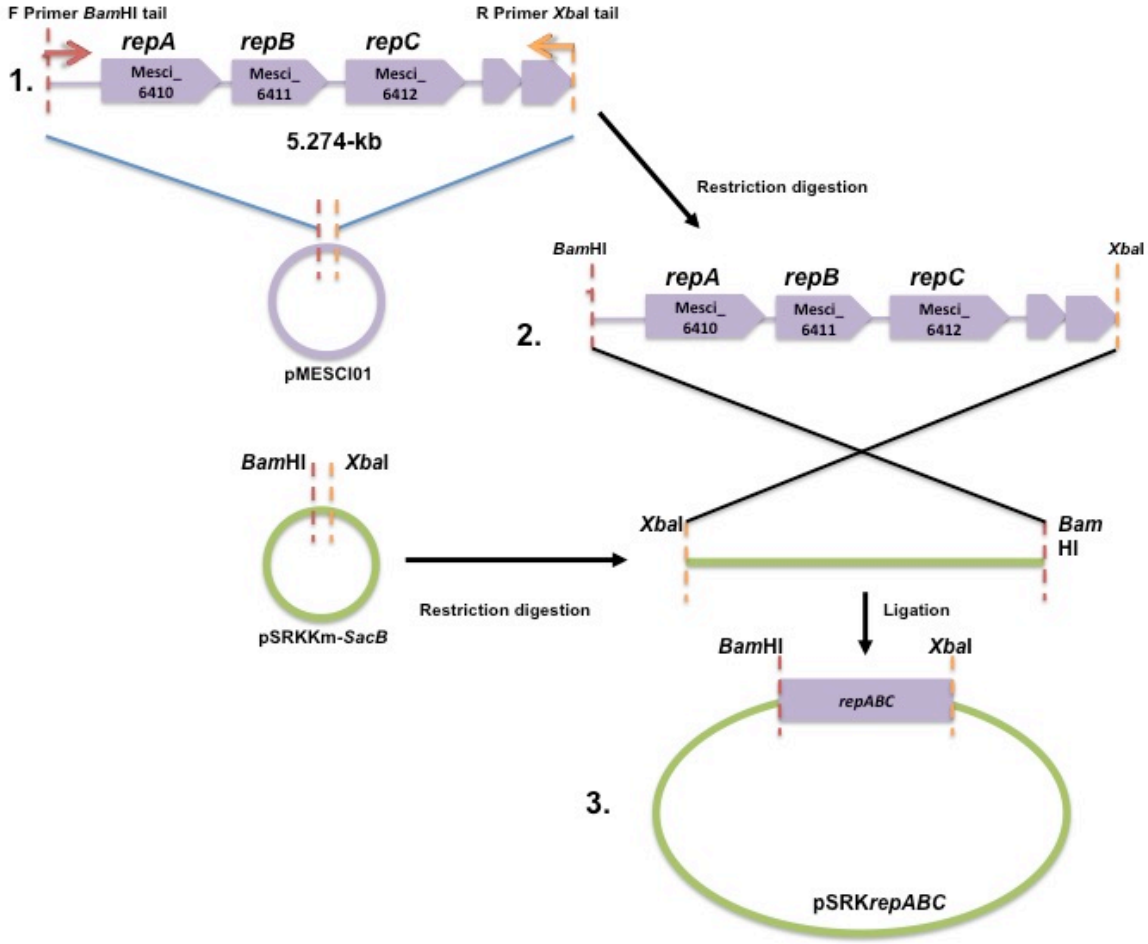


**Figure 2.2:** Construction of the plasmid-curing vector, pSRK*repABC*. 1) PCR amplification of the 5.274-kb *repABC* region of pMESCI01 indicating the addition of *Bam*HI and *Xba*I tails. 2) Restiction digestion of the *repABC* amplified region and the pSRKKm-*SacB* vector with *Bam*HI and *Xba*I. 3) Ligation of the *repABC* region insert with digested pSRKKm-*SacB* vector (note the orientation of the *repABC* insert region).

### 2.7.2 Preparation of pSRKKm-*SacB*

*E.coli* ST18 carrying pSRKKm-*SacB* cells were cultured in LB and plasmid DNA extracted as described in Section 2.4.3. Digestion of 73.9 ng/ µl (13 µl) of pSRKKm-*SacB* was performed with 20 units of *Bam*HI and 20 units of *Xba*I

restriction enzymes, with 10X Cutsmart™ Buffer (New England Biolabs, cat #B7204S), in a 50 µl total reaction, then incubated at 37°C in a water bath for 1 h, followed by dephosphorylation with rSAP (Recombinant Shrimp Alkaline Phosphatase, New England Biolabs, cat #M0371L) incubated at 37°C for 1 h. These digested products were then purified using FavorPrep™ Gel Purification kit (Favorgen Biotech Corp, cat #FAGCK001-1) as per supplier recommendations.

### 2.7.3 Ligation of digested *repABC* and pSRKKm-*SacB* products

Multiple ligation reactions were prepared with insert to vector ratio at 21:16. Preparation of multiple 10 µl ligation reactions with 1 µl NEB T4 ligase (New England Biolabs, cat #M0202S), and 1 µl 10X Ligase Buffer (New England Biolabs, cat #B0202S), with 0-6.5 µl digested *repABC* insert product and 2 µl (~80ng) digested pSRKKm-*SacB* vector product. These reactions were incubated at 16°C overnight and heat inactivated for 10 min at 65°C prior to transformation into DH5α competent cells.

### 2.7.4 Transformation into *E.coli* and confirmation

A total of 5 µl of the ligation reaction mixture was added to 50 µl competent cells DH5α and was transformed as described in Section 2.4.5. Transformant cells were then spread in dilutions onto LB solid media supplemented with $Km_{50}$ and incubated at 37°C overnight. Colonies were then confirmed for the presence of the plasmid-curing vector (pSRK*repABC*) by colony PCR, using primers number 21 and 22 (Table 2.2) designed to capture the *repABC* region and a region of pSRKKm-*SacB* downstream to the *repABC* insert. PCR cycling conditions are as described for PCR 2 (Section 2.3). PCR reactions were performed and the

products electrophoresed as described earlier (Section 2.4.5), with molecular weight of products instead determined as described in Section 2.7.1. PCR confirmed DH5α transformants containing the pSRK*repABC* vector were then cultured in LB broth and plasmid DNA extracted as previously described (Section 2.4.2) prior to digestion with 20 units of EcoRI restriction enzyme, 10 X Buffer E (Promega, cat #2005A) in a total reaction volume of 25 µl at 37°C overnight. The digested products were electrophoresed and the molecular weights of products were determined as previously in Section 2.7.1. Confirmed transformants were then transformed into *E.coli* ST18 commercially competent cells as described previously in Section 2.4.5, and the transformation reaction was spread in dilutions onto LB supplemented with $ALA_{60}$ and $Km_{50}$, and then incubated overnight at 37°C.

## 2.8 Conjugative transfer of pSRK*repABC* plasmid-curing vector into WSM1271

The plasmid-curing vector (pSRK*repABC*) was mobilised from *E.coli* ST18 into WSM1271 by biparental mating to cure pMESCI01 via plasmid incompatibility. WSM1271 was grown in TY broth and *E.coli* ST18 transformants were grown in LB broth, and biparental mating performed as described previously in Section 2.5. Mating spots were then resuspended in 500 µl of sterile 0.89% (w/v) saline, and cells were spread in dilutions onto G/RDM supplemented with $Nm_{250}$ and incubated at 28°C. Transconjugants were then passaged onto G/RDM supplemented with $Nm_{250}$ and incubated at 28°C.

### 2.8.1 Confirmation of pMESCI01-cured transconjugants and Eckhardt Gel electrophoresis

Colonies were confirmed for the absence of pMESCI01 by colony PCR, using primers 1 and 2, or 3 and 4, or 17 and 18 (Table 2.2) designed to capture either locations of pMESCI01 or WSM1271 chromosome. PCR cycling conditions as were described for PCR 5 (Section 2.3), with PCR reactions were performed and the products electrophoresed as described earlier (Section 2.4.5).

Confirmed colonies then underwent further confirmation via Eckhardt Gel electrophoresis confirmation. Wild-type WSM1271, putative plasmid-cured derivatives of WSM1271, and *Rhizobium leguminosarum* 3841 (selected to provide the plasmid molecular weight markers), were grown in 5 ml TY broth to achieve an $OD_{600nm}$ of ~0.3. Cells were cooled on ice and a 200 µl aliquot of each was added to 1 ml of cold 0.3% sarcosyl in 1 X TBE buffer (89mM Tris, 89mM boric acid, 2mM EDTA) and further incubated on ice for 10 min. Cells were harvested by centrifugation (5,000 x G, 5 min) at 4°C, and resuspended in 25 µl lysis solution (10 mg/ml lysozyme in E1 solution; 10% (w/v) sucrose, 5 µl 20 mg/ml RNAse stock solution (Invitrogen, cat #12091-021), in 1 X TBE buffer). Lysed cells were electrophoresed and the molecular weight of products were determined by comparison to the lysed *R. leguminosarum* 3841. Gels were run at 4°C for 16 h at 70V with 0.75% Agarose with 1% SDS (pH 8) in 1X TBE buffer. Gels were post-stained with Ethidium Bromide (10 mg/ml) (Sigma-Aldrich) in 1X TBE buffer and de-stained in 1X TBE buffer.

### 2.8.2 16S rDNA sequencing of confirmed pMESCI01-cured derivatives of WSM1271

The 16S rDNA region of pMESCI01-cured WSM1271 derivative colonies was amplified, using primers 19 and 20 (Table 2.2), which bind within the 16S region. PCR reactions were prepared in 50 µl volumes containing 2X GoTaq® Green Mastermix (Promega cat #M712B), and 0.5 µM of each primer. PCR conditions as were described for PCR 6 (Section 2.3), with PCR products electrophoresed as described earlier (Section 2.4.5). A 50 µl aliquot of the PCR product was purified prior to sequencing using FavorPrep™ PCR Purification kit (Favorgen Biotech Corp, cat #FAGCK001-1). A 10 µl aliquot of purified PCR product with 5 µM primer 23, and a 10 µl aliquot of purified PCR product with 5 µM primer 19, were sequenced by Sanger sequencing through the Australian Genome Research Facility (AGRF).

### 2.8.3 Removal of pSRK*repABC* vector from confirmed pMESCI01-cured WSM1271 derivatives

Confirmed pMESCI01-cured WSM1271 derivatives were then passaged onto S/RDM to force removal of the pSRK*repABC* vector incubated at 28°C for 4 days. Subsequently these were then replica patched onto G/RDM supplemented with $Nm_{250}$ and S/RDM to ensure for loss of $Km^R$, and incubated at 28°C for 4 days.

## 2.9 Symbiotic assessment of pMESCI01-cured derivatives of WSM1271 with *B. pelecinus*

### 2.9.1 Set-up of glasshouse trial and watering routine

Symbiotic assessment of two pMESC01-cured WSM1271 derivatives inoculated with *B. pelecinus* cv. Casbah, compared to WSM1271 wild-type inoculated *B.*

*pelecinus* cv. Casbah, and uninoculated nitrogen-fed (N+) and nitrogen-starved (N-) controls was conducted following a 61 day glasshouse trial. Seeds were scarified and surface sterilised with 30 s of light scarification followed by sterilisation in 70% (v/v) ethanol for 1 min, 4% (v/v) hydrogen peroxide for 3 min, and six rinses in sterile water. Prior to sowing, all seeds were pre-germinated on 0.9% (w/v) sterile water agar and inocubated at 28°C overnight until clear radicals were observed. Eight sterilised seeds were sown into each free-draining treatment pot, containing a 1:1 ratio of sterilised river-sand to yellow-sand mixture. The inoculant strains were grown on G/RDM agar slopes and incubated at 28°C for 4 days and then washed with 3 ml sterile water, and 1 ml aliquot was diluted into 50 ml of sterile water to make the final inoculant solution. A volume of 1 ml of inoculant solution was provided to each seed, with a total of 8 ml per pot. All plants were initially watered with 35 ml of a nitrogen-containing starter solution (CRS Plant Growth Nutrient Solution with 0.25g/L $KNO_3$) (Howieson *et al.*, 2016). In the first three weeks post-inoculation and sowing, all plants were watered with 40 ml CRS Plant Growth Nutrient Solution (Howieson and Dilworth, 2016) weekly and N+ plants were provided 5 ml of nitrogen source (10g/L $KNO_3$) weekly. Following week three all plants were provided 20 ml of CRS Plant Growth Nutrient Solution per week and N+ plants were provided 10 ml of the nitrogen source until harvest. Weekly observations of plant growth were made.

### 2.9.2 Plant harvesting and nodule preparation

All plants were harvested at 61 days post-inoculation and observations of plants, including colour and relative sizes, were made. The shoots from each plant were

separated from the roots and then dried at 60°C and weighed to determine individual plant shoot dry weights. Counts of nodules, as well as observations of their colour and distribution were conducted. A selection of nodules across the treatments were removed from plant roots for nodule sectioning and fixed with 3% glutaraldehyde and 25 mM phosphate buffer at pH 7 at 4°C overnight. Nodule sectioning and subsequent light microscopy slides were prepared as described by Richardson *et al.* (1960) and Spurr (1969). Nodules were prepared for nodule crushing and surface sterilised in 70% (v/v) ethanol for 30 s, 4% (v/v) hydrogen peroxide for 30 s, followed by six rinses in sterile water. Nodules were crushed and resuspended in sterile water and streaked onto G/RDM supplemented with $Sm_{200}$ and incubated at 28°C.

### 2.9.3 Statistical analysis of shoot dry weights

The one-way analysis of variance (ANOVA) was then determined across all treatments using RStudio (https://www.rstudio.com). A Tukey multiple comparisons of means (TukeyHSD) were then determined following ANOVA using RStudio. The presumed contaminated N- plants were removed from these sample sets.

# 3. Results

## 3.1 Bioinformatic analysis of pMESCI01

Plasmid pMESCI01 is 425,539-bp in size (6.36% of the genome of WSM1271) and harbours 432 protein-coding genes (Table 3.1). To investigate the potential role of these genes, an analysis of the Clusters of Orthologous Groups of proteins (COG) of pMESCI01 was conducted. COG analysis classifies proteins based on sequence similarity into known or putative functional groups, which can assist in determining the potential function of protein-encoding genes (Tatusov *et al.*, 2001). Genes on pMESCI01 were grouped into 18 distinct COG categories, with 51.2% (221 genes) either classified as hypothetical proteins or un-annotated genes with no known putative functions. A further 2.6% (11 genes) were categorised within the non-specific "general function prediction", which includes proteins with predicted biochemical activity where the specific COG function is unclear (Tatusov *et al.*, 2001); (Table 3.1).

The highest percentage of genes (7.2%) in pMESCI01 to be grouped within a functional COG category were within "energy production and conversion", which included a possible NADH dehydrogenase (COG_1252). The plasmid appears to also contain genes for transport and metabolism of amino acids (5.1%), carbohydrates (4.6%), lipids (3.2%), coenzymes (2.8%), and inorganic ions (2.3%), as well as genes functioning in the biosynthesis, transport or catabolism of secondary metabolites (0.7%). Genes putatively involved in transcription (3.0%), translation (0.5%), and post-translational modification, protein turnover

**Table 3.1:** Summary of genes encoded on pMESCI01 and categorised based on Clusters of Orthologous Groups of proteins (COG). Total number of genes and percentage of genes are shown for each COG category. The highlighted groups show those genes of unknown function or not within a determined COG group.

| COG Function | Number of Genes | % Genes |
|---|---|---|
| Amino acid transport and metabolism | 22 | 5.1% |
| Carbohydrate transport and metabolism | 20 | 4.6% |
| Cell cycle control, cell division, chromosome partitioning | 2 | 0.5% |
| Cell motility | 9 | 2.1% |
| Cell wall/membrane/envelope biogenesis | 8 | 1.9% |
| Coenzyme transport and metabolism | 12 | 2.8% |
| Defence mechanisms | 9 | 2.1% |
| Energy production and conversion | 31 | 7.2% |
| General function prediction only | 11 | 2.6% |
| Inorganic ion transport and metabolism | 10 | 2.3% |
| Intracellular trafficking, secretion, and vesicular transport | 16 | 3.7% |
| Lipid transport and metabolism | 14 | 3.2% |
| Posttranslational modification, protein turnover, chaperones | 7 | 1.6% |
| Replication, recombination and repair | 3 | 0.7% |
| Secondary metabolites biosynthesis, transport and catabolism | 3 | 0.7% |
| Signal transduction mechanisms | 19 | 4.4% |
| Transcription | 13 | 3.0% |
| Translation, ribosomal structure and biogenesis | 2 | 0.5% |
| Not in COG | 202 | 46.8% |
| Function unknown | 19 | 4.4% |
| **Grand Total** | **432** | |

or chaperones (1.6%) were also present. In addition, genes potentially involved in cell wall, membrane or envelopes biogenesis (1.9%), and motility (2.1%) were also located on pMESCI01. Putative signal transduction (4.4%), intracellular trafficking, secretion, or vesicular transport (3.7%), and defence mechanisms (2.1%) genes were also discovered. A small percentage of genes coding for DNA replication, recombination and repair (0.7%), as well as cell cycle control, cell division and chromosome partitioning (0.5%) were also found on pMESCI01.

Thus, although more than half of the genes on pMECI01 have no putative role, those grouped into COG categories appear to be involved in a diverse range of cellular functions. Importantly, this bioinformatic analysis revealed genes potentially involved in nitrogen fixation, plasmid replication, transfer and stability as well as bacterial secretion systems are also encoded on pMESCI01. In order to determine the role of pMESCI01 in WSM1271 further bioinformatic analysis of these genes was conducted.

### 3.1.1 Genes putatively involved in nitrogen fixation

Plasmid pMESCI01 encodes a total of 12 potential nitrogen fixation genes, which include *fixGHIS* (Mesci_6042-6045), *fixNOQP* (Mesci_6046-6049), *fixLJ* (Mesci_6462-6463), *fixK* (Mesci_6459), and *nifB* (Mesci_6052) (Figure 3.1).

The *fixNOQP* and *fixGHIS* operons are required for synthesis and operation of the high-affinity $cbb_3$-type cytochrome oxidase, which is expressed in nitrogen fixing bacteroids, (see Section 1.3, and review by Pitcher and Watmough (2004)). To determine whether the WSM1271 genome contained additional copies of these

nitrogen fixation genes, BLASTP searches of the WSM1271 chromosome (including ICE*Mc*Sym[1271]) were performed using pMESCI01 FixNOQP and FixGHIS as the query sequences. These searches revealed that WSM1271 harbours an additional copy of the *fixGHIS* and *fixNOQP* operons located within the α region of ICE*Mc*Sym[1271], with amino acid sequence identity of >81% to that of the pMESCI01 copies (Table 3.2).

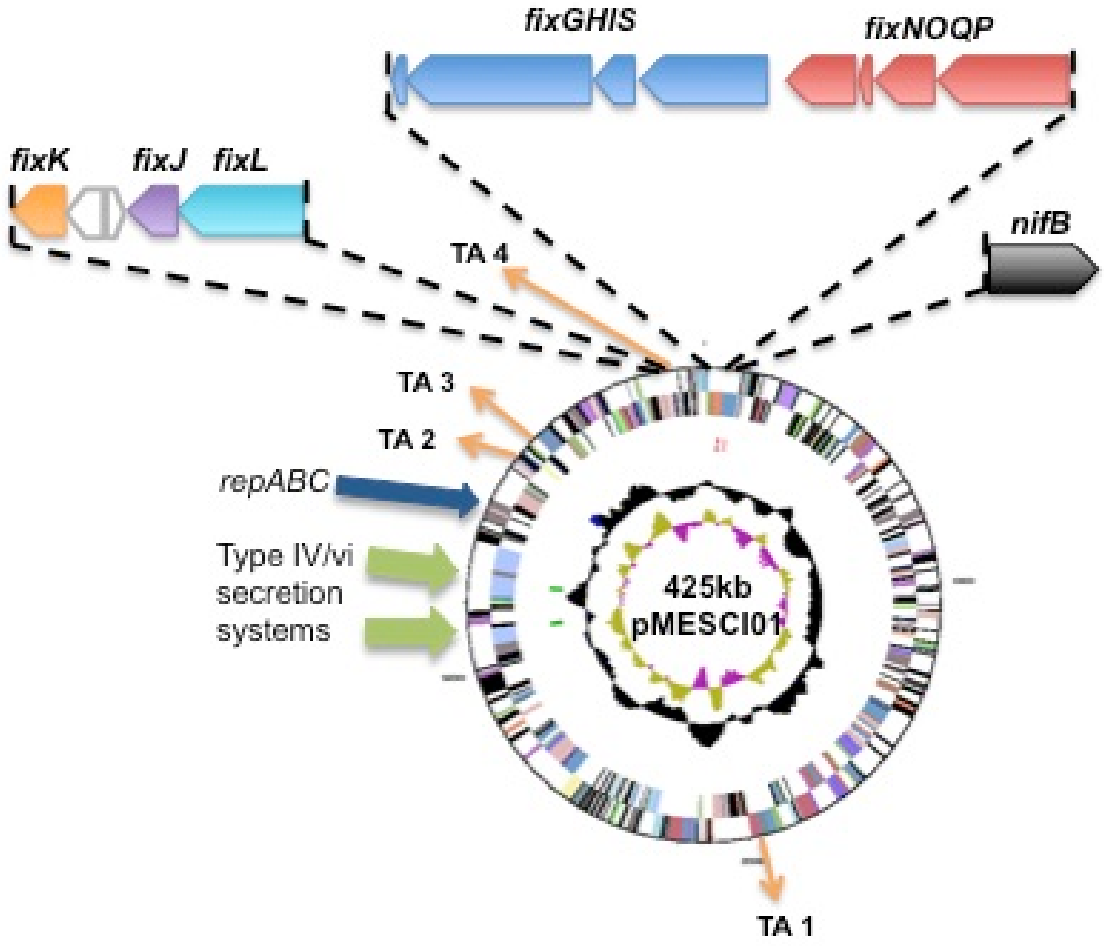


**Figure 3.1:** Plasmid map of pMESCI01 indicating the location, orientation and relative size of putative nitrogen fixation genes*; fixGHIS, fixNOQP, fixK, fixLJ,* and *nifB*; the annotated Type IV/vi secretion systems; replication and partitioning site of pMESCI01 (*repABC*); and the four putative Type II toxin-antitoxin systems (TA1, TA2, TA3, TA4). The IMG/JGI plasmid-mapping tool was used to create the plasmid map of pMESCI01, which has been annotated to highlight genes of interest.

Further pairwise alignments of the entire ICE*Mc*Sym[1271] with pMESCI01-encoded *fixGHIS* and *fixNOQP* operons demonstrated that the orientation and order of both operons in both replicons of WSM1271 were the same. BLASTN comparison of the nucleotide sequence of these operons found pMESCI01-encoded *fixGHIS* and *fixNOQP* to be 84.8% and 86.7% identical to the ICE*Mc*Sym[1271] operons, respectively. Overall this analysis indicates that both copies of *fixGHIS* and *fixNOQP* on pMESCI01 are likely to be present in operons and that, although sharing >81% identity, are reasonably distinct at the sequence level. It is therefore possible that both plasmid and ICE*Mc*Sym[1271]-encoded sets of genes are functional and therefore could be required for nitrogen fixation in WSM1271.

**Table 3.2:** Percentage identity of putative pMESCI01-located nitrogen fixation genes *fixGHIS* and *fixNOQP*, *fixLJ, and fixK* following BLASTP comparison with genomes of WSM1271 ICE*Mc*Sym[1271], and naïve (i.e. consisting of the chromosomal only, and lacking ICE*Mc*Sym[1271]) WSM2073 and WSM2075.

| **Protein name** | **Gene Locus tag** | **BLASTP % identity to ICE*Mc*Sym[1271]** | **BLASTP % identity to naïve WSM2073** | **BLASTP % identity to naïve WSM2075** |
|---|---|---|---|---|
| FixG | Mesci_6045 | 83% | Absent | Absent |
| FixH | Mesci_6044 | 84% | Absent | Absent |
| FixI | Mesci_6043 | 83% | Absent | Absent |
| FixS | Mesci_6042 | 85% | Absent | Absent |
| FixN | Mesci_6049 | 81% | Absent | Absent |
| FixO | Mesci_6048 | 90% | Absent | Absent |
| FixQ | Mesci_6047 | 81% | Absent | Absent |
| FixP | Mesci_6046 | 81% | Absent | Absent |
| FixL | Mesci_6463 | Absent* | 82% | 80% |
| FixJ | Mesci_6462 | Absent* | 77% | 78% |
| FixK | Mesci_6459 | Absent* | 79% | 76% |

*Note also absent from chromosome of WSM1271 (excluding ICE*Mc*Sym[1271])

Intriguingly, pMESCI01 was also found to encode copies of the nitrogen fixation regulators *fixK* (Mesci_6459), *fixL* (Mesci_6463), and *fixJ* (Mesci_6462) (Section 1.3). BLASTN and BLASTP analysis of the genome of WSM1271 revealed that the chromosome (including ICE*Mc*Sym$^{1271}$) did not harbour any additional copies of any of these genes (Table 3.2). In contrast, WSM2073 and WSM2075 both carry copies of these genes on their chromosome, located outside of ICE*Mc*Sym$^{1271}$. Therefore, in WSM1271, pMESCI01 appears to carry a unique copy of these regulatory genes, but the function of this regulatory complex in WSM1271 is unknown.

Two copies of the *nifB* gene, known to be essential for the synthesis of FeMo-cofactor (Section 1.3) were also discovered in the genome of WSM1271, with one copy encoded on pMESCI01 (Figure 3.1) and the other on ICE*Mc*Sym$^{1271}$. While the two proteins share 78.2% identity, wider analysis of the gene neighbourhood of both genes showed that the surrounding genes differed in putative function and orientation. *nifEN* was not found in pMESCI01, but a single copy of this operon is located on ICE*Mc*Sym$^{1271}$, which is known to be required alongside *nifB* for the synthesis of the FeMo-cofactor (Fischer, 1994; Guo *et al.*, 2016); (Section 1.3).

### 3.1.2 Plasmid replication, transfer, and stability

Identifying replication, transmissibility, and stability functions is an important part of characterising the nature of the plasmids. Bioinformatic analysis identified a possible *repABC*-type plasmid replication and partitioning region located on pMESCI01 (Mesci_6410-6412) (Figure 3.1). The *repABC*-type system

is known to allow for stable propagation of usually low-copy number plasmids in rhizobia (Mazur and Koper, 2012); (Section 1.4.1). Between *repB* and *repC* a putative plasmid transfer (*oriT*) site was also identified using TUBIC BLAST (Gao *et al.*, 2013), with 84% identity to the known *oriT* site of the Ti plasmid of *Agrobacterium radiobacter* K84. Further BLASTN analysis of this potential *oriT* sequence against other *Mesorhizobium* spp. harbouring plasmids revealed that this region is located in the same relative position (i.e. between *repB* and *repC*) and shares high sequence similarity with these replicons. Specifically, the pMESCI01 *oriT* shared 100% identity with the *oriT* of pMc1284 (*M. ciceri* bv. biserrulae WSM1284) and pMc1192 (*M. ciceri* CC1192), 97% identity to pML2037 (*M. loti* NZP2037) and pM0123d (*M. loti* MAFF303099), as well as 92% identity to pMHa (*M. huakuii* 7653R).

A vital component of the DNA transfer and replication (Dtr) system in plasmids is the relaxase enzyme, which can initiate plasmid replication and transfer by nicking the plasmid DNA at the *oriT* site. A BLASTP search of pMESCI01 with the amino acid sequence of 29 known relaxase enzymes did not yield any significant hits (Appendix Table A1). Thus, although a putative *oriT* site was located on pMESCI01, a candidate pMESCI01-encoded relaxase could not be located.

To investigate whether pMESCI01 encoded functions for self-transmissibility, an investigation into the presence of Dtr and mating pair formation (Mpf) genes was conducted (Section 1.4.1). Initially, an analysis of two genes annotated as "Type IV/vi secretion systems" (Mesci_6367 and Mesci_6379) (Figure 3.1) found on pMESCI01 was performed. BLASTP comparison of Mesci_6367 revealed this

sequence to have 71% identity to the T6SS protein ImpK of *M. muleience,* whereas Mesci_6379 was revealed to have 90% identity to the *M. gingshengii* T6SS protein ImpK. These similarities indicate both Mesci_6367 and Mesci_6379 are likely to function as part of a T6SS, rather than a T4SS. However additional functional analysis is required to confirm this designation.

Further analysis was conducted to determine whether pMESCI01 contained any genes required for a known T4SS. BLASTP interrogation of the known T4SS genes of the self-transmissible plasmid pRLG203 for *Rhizobium leguminosarum* bv. viciae WSM2304 (Ding *et al.*, 2013) with pMESCI01 was performed. This comparison found no significant identity across any T4SS genes in the pMESCI01 sequence. The inability to detect any homologs for known T4SS genes suggests that pMESCI01 either has a novel conjugation system or that the plasmid is not self-transmissible.

A TAfinder search for plasmid stability Type II toxin-antitoxin (TA) system genes was also conducted using the TAfinder online program (Version 2.0; Microbial Bioinformatics Group, Shanghai Jiaotong University). This search yielded four putative TA systems located on pMESCI01 (Figure 3.1), potentially indicating that these systems could function to allow stable inheritance of the pMESCI01 plasmid (Section 1.4.1).

In summary, the bioinformatic analysis suggests that pMESCI01 may lack the ability to be self-transmissible due to the absence of a relaxase and a T4SS. It is, however possible the plasmid may encode an unknown conjugation system that

allows for self-transmissibility. Furthermore, pMESCI01 harbours several genes that could be involved in nitrogen fixation (*fixGHIS, fixNOQP, fixLJ, fixK,* and *nifB*). Therefore, functional analysis of pMESCIO1 is required to determine whether the plasmid is self-transmissible and whether it has a role in symbiosis.

## 3.2 Marking pMESCI01 to assess plasmid self-transmissibility

To allow for the assessment of the self-transmissibility, pMESCI01 was marked with an antibiotic cassette (Ω-Sp/Sm) encoding resistance to spectinomycin and streptomycin. The cassette was delivered into pMESCI01 by first creating a plasmid marking vector with a pJQ200SK-*SacB* vector backbone utilizing a quadruple ligation approach.

### 3.2.1 Construction and confirmation of plasmid-marking vector

The selected insertion site for the Ω-Sp/Sm cassette was the ~1.4-kb intergenic region between convergently transcribed genes Mesci_6067 and Mesci_6068, both of which encode hypothetical proteins. To construct the plasmid-marking vector two regions, a 1-kb upstream and 1-kb downstream adjacent regions of coordinate 29621 of pMESCI01, were separately amplified with selected restriction sites attached (Section 2.4.1, Appendix Figure A1). The amplified upstream and downstream region, pJQ200SK-*SacB*, and pJET(Ω-Sp/Sm) were digested with *Sac*I, *Not*I, *Xba*I or *Sal*I (Section 2.4.2-3; Appendix Figure A2-3).

The four digested fragments were ligated in a single step, yielding the plasmid-marking vector, pJQ(Ω-Sp/Sm), which was subsequently transformed into *E. coli* DH5α. PCR screening of 13 transformant colonies revealed two colonies (T5 and T6) contained the expected ~5.5-kb product (Figure 3.2), with some non-specific

product also evident particularly in the lane containing T6. To confirm the PCR results, these two colonies were further analysed by restriction digestion of plasmid DNA with *Xho*I. Both colonies yielded the expected banding patterns of ~4.4-kb (Appendix Figure A4), confirming the construction of a plasmid-marking vector, pJQ(Ω-Sp/Sm). The two successful transformants confirmed to contain the plasmid-marking vectors were named RB6.5 and RB6.6. Once confirmed, both RB6.5 and RB6.6 were separately transformed into ST18 competent cells, prior to conjugation with WSM1271.

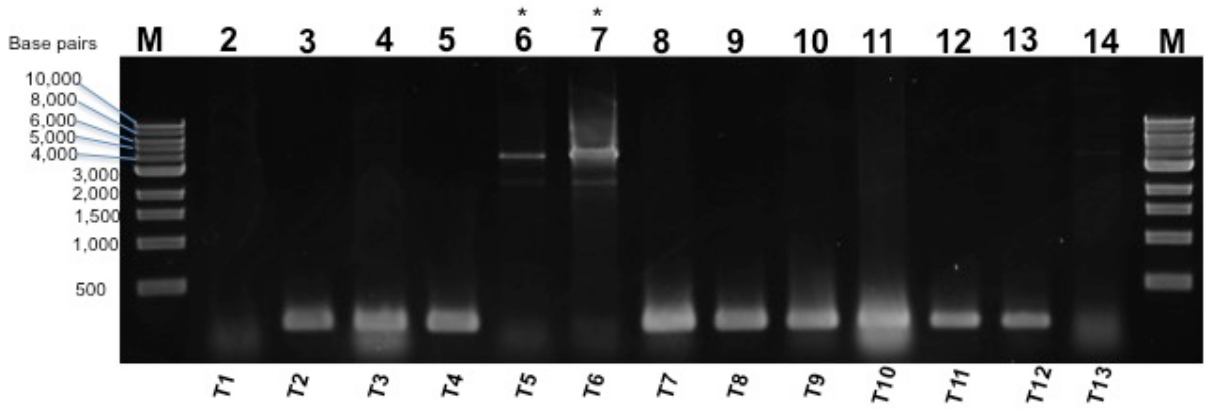


**Figure 3.2:** Image of electrophoresed PCR products of 13 transformant colonies of *E. coli* DH5 (T1-13) putatively containing the plasmid-marking vector, pJQ(Ω-Sp/Sm). Lanes 2-14 show PCR products of transformant colonies (T1-13), with T5 and T6 (lanes 6-7) showing successful banding pattern of ~5.5-kb (indicated with *). The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

### 3.2.2 Conjugation and selection for marked pMESCI01 in WSM1271

Confirmed transformants containing the plasmid-marking vector, pJQ(Ω-Sp/Sm), RB6.5 and RB6.6, were separately conjugated into WSM1271 and $Gm^R$ colonies of putative single crossover (SXO) transconjugants were selected. Colony PCR screening of both putative RB6.5 and RB6.6 SXO transconjugants revealed the expected banding pattern of ~2-kb for seven of the ten RB6.5 SXO

transconjugants and ~2-kb for eight of the ten RB6.6 SXO transconjugants (Appendix Figure A5). One successful colony from each mating was chosen to select for double crossover (DXO) transconjugants, and named RB6.5b and RB6.6.b, respectively, to the plasmids RB6.5 and RB6.6.

Double crossover (DXO) transconjugants were then selected for in order to remove the pJQ200SK-*SacB* backbone from within the marked pMESCI01, leaving only the Ω-Sp/Sm cassette in pMESCI01. A total of 20 (ten each from RB6.5b and RB6.6b) putative DXO transconjugant colonies growing on S/RDM were screened by colony PCR with primers specific to a region upstream or downstream of the Ω-Sp/Sm cassette insertion site in pMESCI01. The screening revealed a total of 14 successful DXO transconjugants were present, with the expected upstream region product of ~4-kb and downstream region product of ~4.4-kb (with a total of six and eight confirmed DXO transconjugants coming from RB6.5b and RB6.6b, respectively) (Figure 3.3). These results indicated that these 14 colonies were contained the desired marked pMESCI01 referred to as pMESCI01::Ω-Sp/Sm.

Five DXO confirmed colonies from both matings, RB6.5b and RB6.6b, were then replica patched to confirm $Gm^S$, as would be expected with DXO transconjugants that have lost the pJQ200SK-*SacB* backbone. One confirmed DXO colony from each mating was selected following replica patching and named RB7.2 and RB8.3, representing WSM1271 derivatives containing marked plasmid, pMESCI01::Ω-Sp/Sm.

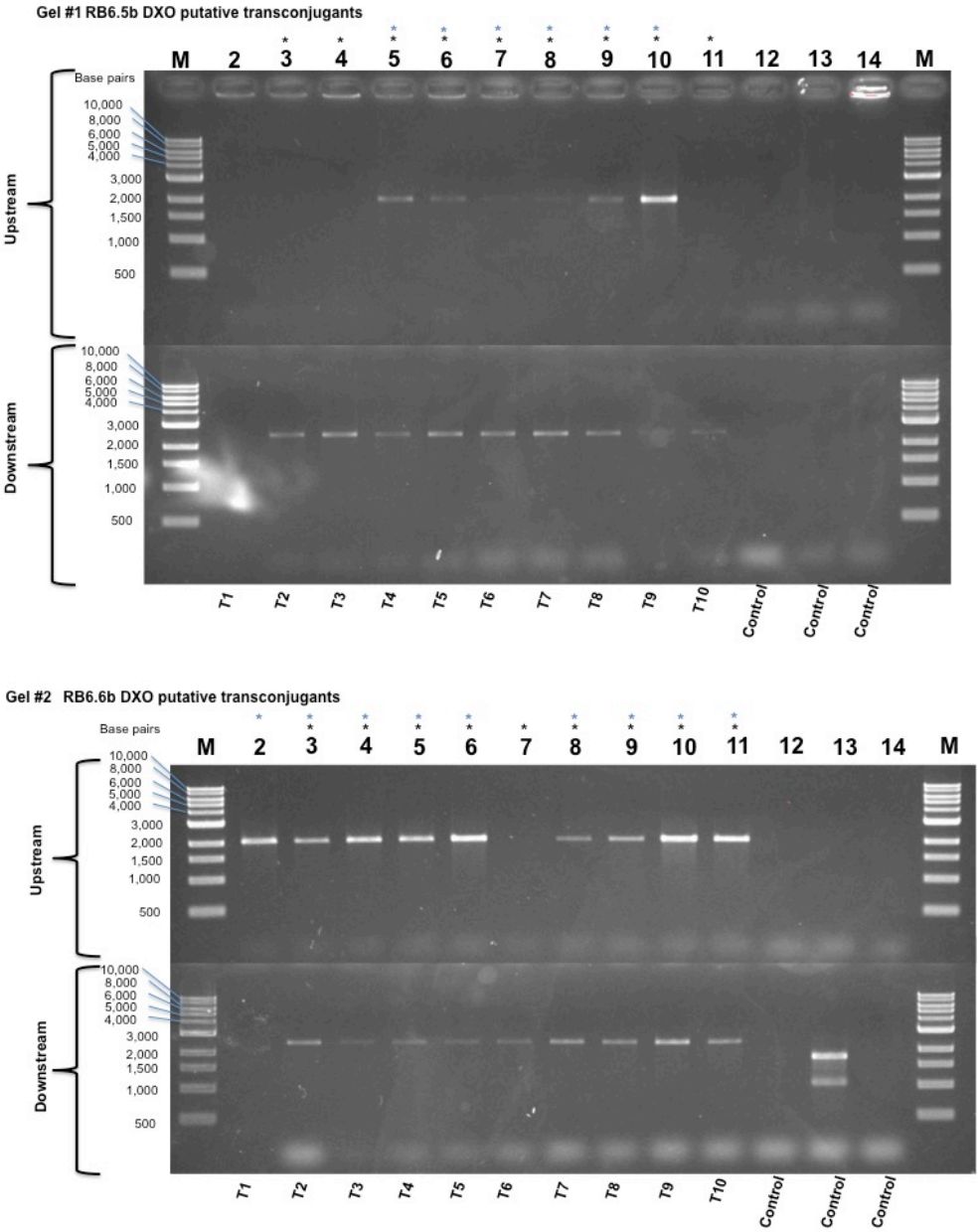


**Figure 3.3:** Image of electrophoresed PCR products of 20 putative DXO transconjugants of WSM1271 with pMESCI01::Ω-Sp/Sm. Gel 1 lanes 2-11 of both upstream and downstream show PCR products of transconjugants of RB6.5b DXO selection. Gel 2 lanes 2-11 of both upstream and downstream show PCR products of transconjugants of RB6.6b DXO selection. The "upstream" image of both gels shows PCR of an upstream site within pMESCI01 to insertion location of Ω-Sp/Sm. Those revealing the expected banding pattern of ~4-kb are shown with blue *. The "downstream" image of both gels shows PCR of a downstream site within pMESCI01 the to insertion location of Ω-Sp/Sm. Those revealing the expected banding pattern of ~4.4-kb are shown with black *. Those indicated with both the blue and black * are confirmed as being successful DXO transconjugants. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

### 3.2.3 Conjugal transfer of pMESCI01::Ω-Sp/Sm to WSM2073

The confirmed WSM1271 derivatives containing pMESCI01::Ω-Sp/Sm, RB7.2 and RB8.3, were separately conjugated into WSM2073Nm using donor cells harvested from (1) log phase and (2) stationary phase growth, to determine whether pMESCI01 is self-transmissible.

Putative transconjugants of WSM2073Nm containing pMESCI01::Ω-Sp/Sm were plated to select for transfer of the marked plasmid with $Sp_{200}$ and $Nm_{250}$. After incubation, copious growth was observed on all selection plates, including plates containing only the parental control (WSM2073Nm). This suggested that WSM2073Nm was in fact $Sp^R$ at all dilutions plated.

Subsequent antibiotic screening of WSM2073Nm was conducted to determine suitable selection for repetition of this experiment (Section 2.6.1). Screening revealed that WSM2073Nm grown on G/RDM supplemented with $Nm_{250}$ and $Sp_{200}$ contained 3 x $10^9$ cells/ml, equal to cell numbers obtained when WSM2073Nm was grown on G/RDM supplemented only with $Nm_{250}$. This confirmed that WSM2073Nm was $Sp^R$. However, when WSM2073Nm was plated onto G/RDM supplemented with $Nm_{250}$, $Sp_{200}$ and $Sm_{200}$, viable counts were reduced by nine orders of magnitude, with counts of 80 cells/ml.

Therefore, screening for self-transmissibility of pMESCI01 in WSM2073Nm could be assessed, using Sm and Nm to select for transconjugants. However, time constraints precluded this experiment from being performed in this thesis.

## 3.3 Curing of pMESCI01 from WSM1271

In order to determine whether genes on pMESCI01 are essential to nitrogen fixation of WSM1271 with *B. pelecinus*, a plasmid-cured version of WSM1271 was created using a plasmid incompatibility approach.

### 3.3.1 Construction and confirmation of a plasmid-curing vector

To cure pMESCI01 from WSM1271 a plasmid-curing vector was constructed. A 5.274-kb region containing the putative *repABC* region of pMESCI01 (Mesci_6410-6412) was successfully amplified by PCR with a forward primer containing a *Bam*HI tail and reverse primer with an *Xba*I tail. The amplified *repABC* region and a pSRKKm-*SacB* vector were then digested with *Bam*HI and *Xba*I and ligated together to construct the plasmid-curing vector, pSRK*repABC* (Section 2.7). This vector was then transformed into *E. coli* DH5α.

PCR screening of 20 transformant colonies revealed five colonies successfully contained the plasmid-curing vector, yielding the expected ~5.5-kb PCR product (Figure 3.4). These five colonies (T3, T5, T6, T15 and T18) were further screened by restriction digestion analysis of purified plasmid DNA with *Eco*RI. Transformant "T18" (Figure 3.5) yielded the expected banding patterns of ~5.1-kb and ~8.0-kb, confirming the presence of the plasmid-curing vector, pSRK*repABC,* in this colony, which were subsequently transformed into *E.coli* ST18.

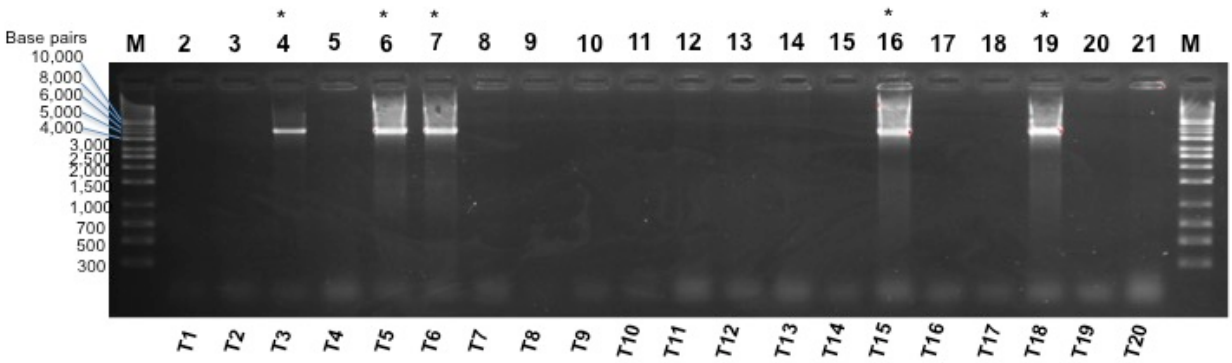


**Figure 3.4:** Image of electrophoresed PCR products of 20 transformant colonies of *E. coli* DH5 (T1-20) putatively containing the plasmid-curing vector, pSRK*repABC.* Lanes 2-21 show PCR products of transformant colonies (T1-20), with transformants T3, T5, T6, T15 and T18 showing successful banding pattern of ~5.5-kb (indicated with *). The Axygen 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

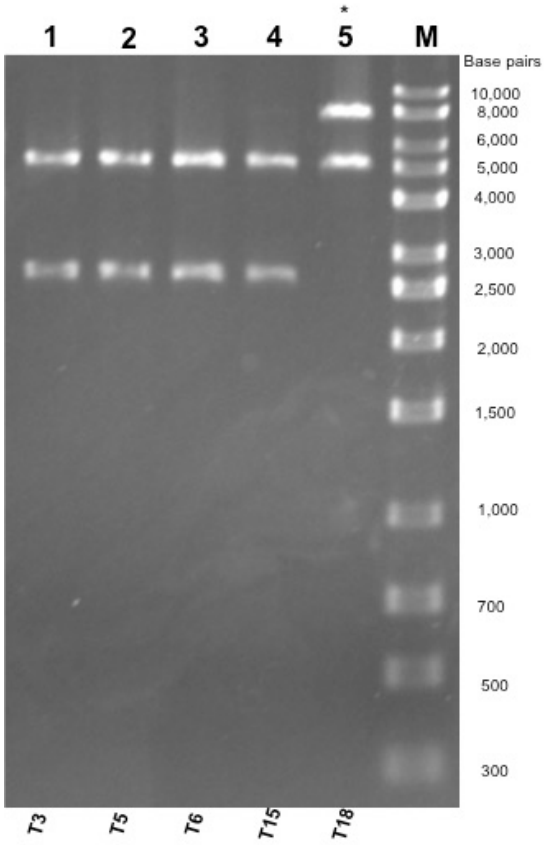


**Figure 3.5:** Image of electrophoresed DNA products digested with *Eco*RI of 5 PCR confirmed transformants (Figure 3.4) containing the plasmid-curing vector, pSRK*repABC.* Lanes 1-5 show digested T3, T5, T6, T15 and T18 consecutively, with transformant T18 showing successful banding patterns of ~5.1-kb and ~8.0-kb (indicated with *). The Axygen 1-kb marker sizes are shown to the right of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

### 3.3.2 Transfer of the plasmid-curing vector into WSM1271 and selection for plasmid-cured derivatives of WSM1271

The plasmid-curing vector, pSRK*repABC,* was introduced into WSM1271 via conjugal transfer with *E.coli* ST18 confirmed vector-containing transformant "T18" and $Nm^R$ transformants representing putative pMESCI01-cured WSM1271 colonies were selected. PCR screening of ten putative plasmid-cured WSM1271 transconjugants with pMESCI01 and chromosomal primers revealed that pMESCI01 was absent in all WSM1271 transconjugants tested (Figure 3.6, Appendix Figure A6-8). The putative plasmid-cured WSM1271 isolates were renamed; RB3.1, RB3.2, RB3.3, RB3.4, RB3.5, RB3.6, RB3.7, RB3.8, RB3.9, RB3.10.

Subsequently, five of the putatively plasmid-cured WSM1271 isolates (RB3.1, RB3.2, RB3.8, RB3.9 and RB3.10) were screened by Eckhardt gel electrophoresis. These Eckhardt gels revealed all transconjugants were successfully cured of pMESCI01 (Figure 3.7).

Prior to further analysis with these confirmed plasmid-cured WSM1271 colonies, 16S rDNA sequencing was performed to confirm the chromosomal background was WSM1271. A PCR and Eckhardt gel-confirmed plasmid-cured colony, RB3.1, was selected for this sequencing, which showed that RB3.1 has a 549-bp 16S region identical to that of WSM1271 (Appendix Figure A9).

The plasmid-curing vector (pSRK*repABC*) was subsequently removed from PCR and Eckhardt gel-confirmed plasmid-cured colonies, including RB3.1 and RB3.2, by plating of cultures containing each of the two strains separately onto 5%

sucrose-containing RDM plates. Confirmation for the loss of the plasmid-curing vector was subsequently conducted to counter-select for the absence of pSRK*repABC* harbouring *sacB*, by replica patching onto S/RDM and G/RDM, in which colonies that grew on both sucrose and glucose indicated loss of the plasmid-curing vector (Section 2.8.3). The confirmed pMESCI01-less and curing vector-less WSM1271 strains RB3.1 and RB3.2 were renamed RB3.1a and RB3.2a.

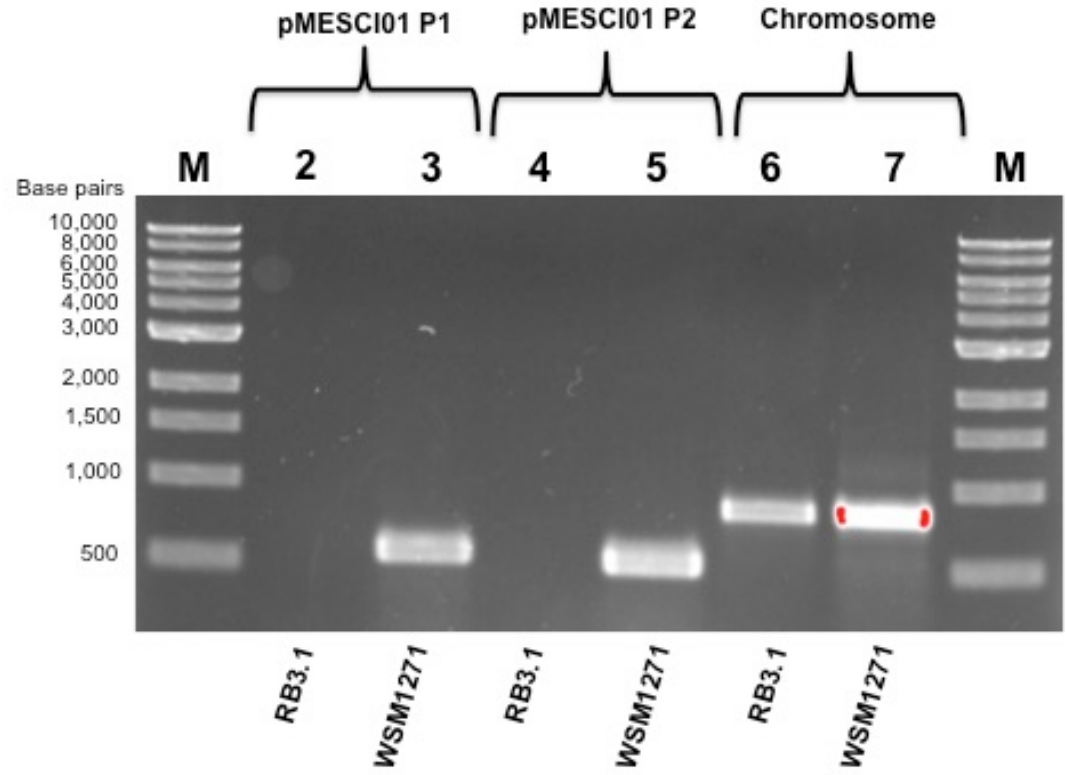


**Figure 3.6:** Image of electrophoresed PCR products of putative plasmid-cured WSM1271 transconjugant RB3.1. Lane 2, PCR product of RB3.1 with pMESCI01 located primer pair 1 (P1); lane 3, PCR product of WSM1271 wild type control with pMESCI01 (P1); lane 4, PCR product of RB3.1 with pMESCI01 located primer set 2 (P2); lane 5, PCR product of WSM1271 with pMESCI01 (P2); lane 6, PCR product of RB3.1 with WSM1271 chromosome located primer set; lane 7, PCR product of WSM1271 with WSM1271 chromosome located primer set. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

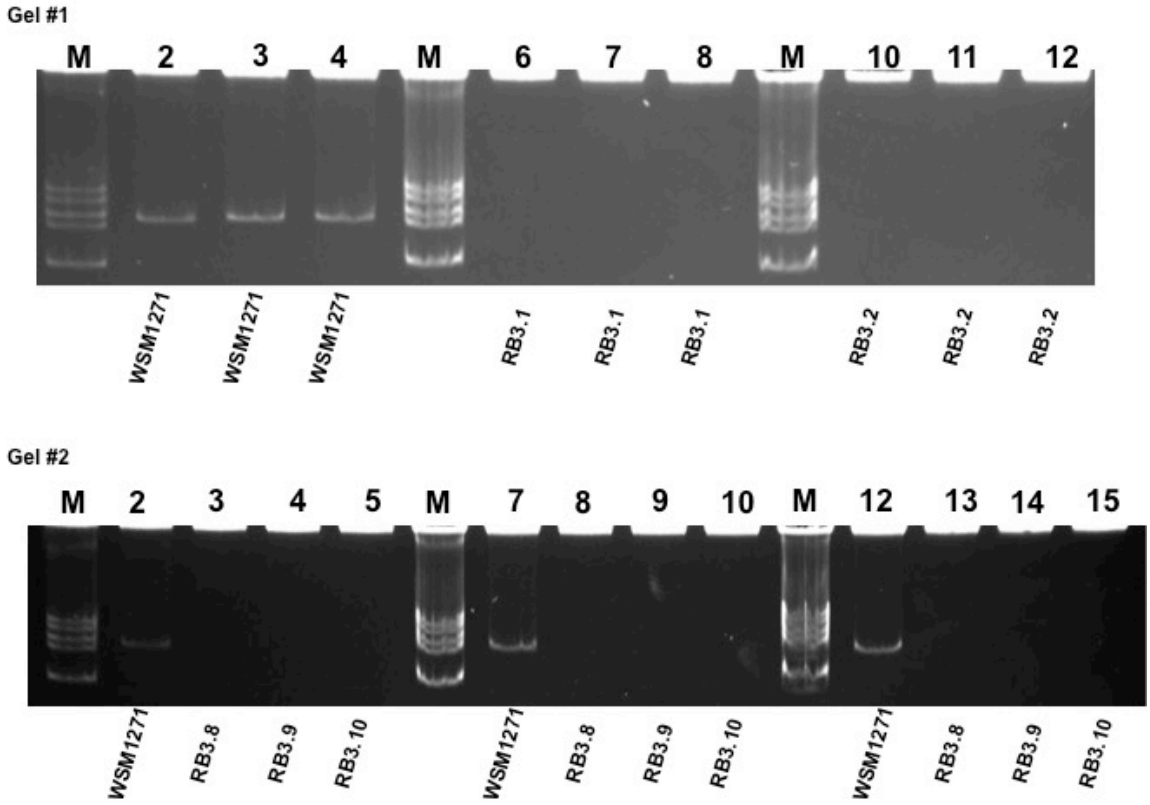


**Figure 3.7:** Image of Eckhardt gel electrophoresed products of putative plasmid-cured WSM1271 transconjugants (RB3.1, RB3.2, RB3.8, RB3.9, and RB3.10). Gel 1; lanes 2-4, WSM1271 wild type at different optical densities (O.D); lanes 6-8, RB3.1 at different O.D; lanes 10-12, RB3.2 at different O.D. Gel 2; lanes 2, 7, 12, WSM1271 wild-type at different O.D; lanes 3, 7, 13, RB3.8 at different O.D; lanes 4, 9, 14, RB3.9 at different O.D; lanes 5, 10, 15, RB3.10 at different O.D. A *Rhizobium leguminosarum* 3841 at corresponding O.D was used as the marker (M). 0.75% (w/v) Agarose with 1% (v/v) SDS in 1 X TBE gel was used, with gels run for 16 hours at 4°C.

## 3.4 Symbiotic analysis of plasmid-cured WSM1271

In order to determine whether pMESCI01 has a role in symbiosis, wild-type WSM1271 and confirmed plasmid-cured WSM1271 strains, RB3.1a and RB3.2a, were inoculated separately onto *B. pelecinus* in a glasshouse experiment. Plants were grown in free-draining pots and harvested 61 days post-inoculation at which time plant shoot dry weights were analysed and nodule counts were performed.

### 3.4.1 Nodule counts and observations

The average nodule number for plants inoculated with wild type WSM1271 and both plasmid-cured WSM1271 treatments was within the range of 100-150 nodules per plant. Additionally, no differences in nodule location were observed with nodules on all three treatments located on the upper-to-middle zone of the primary root. Further analysis of nodule morphology by light microscopy revealed no difference between nodules from all three treatments (Figure 3.8). Nodules across all treatments exhibited a branched, multi-lobed appearance (Figure 3.8). Overall, all nodules were of comparable size with nodule sections confirming the plant cells appeared to be well occupied by bacteroids (Figure 3.8).

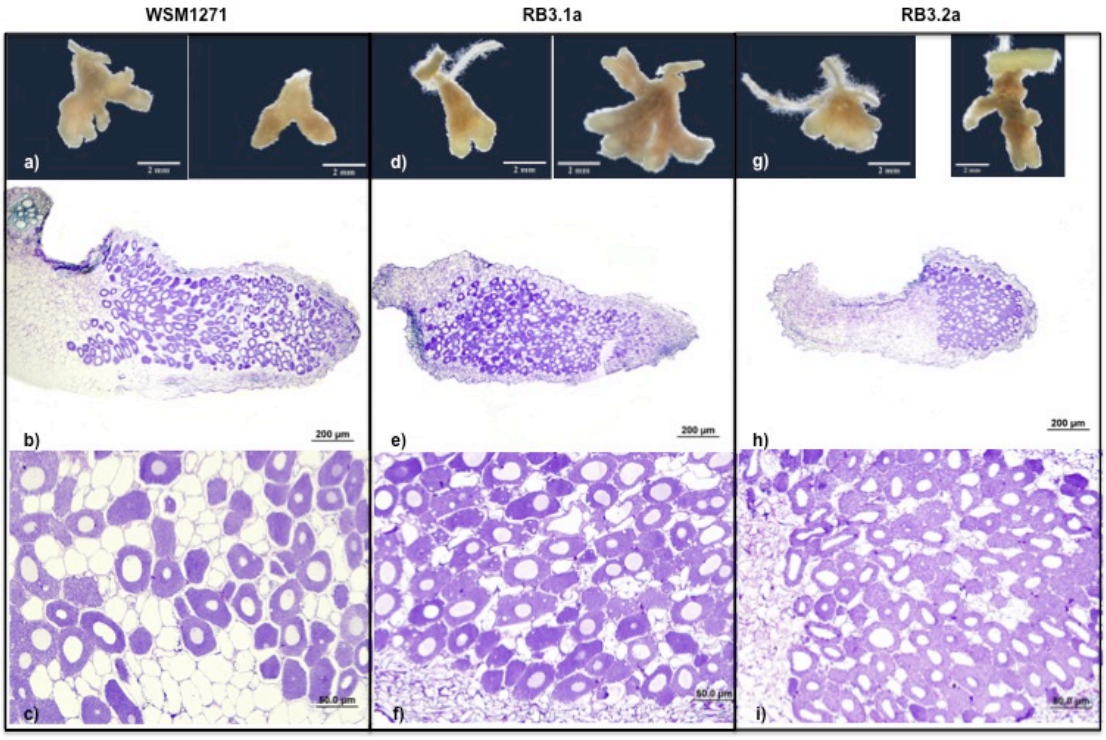


**Figure 3.8:** Images of nodules and nodule sections from *B. pelecinus* inoculated with 1) wild-type WSM1271 (left column a-c), 2) plasmid-cured derivatives of WSM1271 RB3.1a (centre column d-f) or RB3.2a (right column g-i); harvested 61 days post inoculation.

While no nodules were evident on nitrogen fed (N+) controls, some nodules were observed on ten plants across four pots of the uninoculated, nitrogen starved (N-) treatments. These nodules were low in number (<10 per plant) consisting of small, single white protrusions located on the lower region of the root system. Subsequent excision, surface-sterilisation, crushing and streaking of these nodules onto non-selective TY agar plates was conducted. However following incubation for 6-8 days at 28°C, no growth was observed on these plates. The lack of pink colouration of these nodules and their location on the lower parts of the root system indicate that the nodules may have arisen from a post-inoculation contamination event in some, but not all, pots. Nevertheless, all contaminated N- control plants were removed from subsequent shoot dry weight analysis.

### 3.4.2 Plant shoot dry weights and observations

All inoculated *B. pelecinus* were large and green in colour with no differences between plasmid-cured WSM1271 and wild-type WSM1271 inoculated treatment plants. Statistical analysis of shoot dry weights of all plants revealed no significant difference between plants inoculated with plasmid-cured strains, RB3.1a and RB3.2a, versus the wild-type WSM1271 ($P > 0.5$, Tukey multiple comparison of means following one-way analysis of variance, Figure 3.9). In contrast, shoot dry weights of inoculated plants were significantly larger than un-inoculated N- as well as N+ controls plants.

The mean shoot dry weights of N+ control plants were significantly lower than the inoculated treatments ($P < 0.05$; Figure 3.9), possibly indicating nitrogen

limitation of the N+ control plants. At approximately three weeks, the amount of nitrogen provided (as $KNO_3$) was doubled to the $N^+$ plants, which may have lead to increased growth rate of these plants. However, although smaller than the inoculated treatments, none of the N+ plants showed visible signs of chlorosis, as was visible in the N- controls.

Overall these results reveal that pMESCI01 is not essential to effective nitrogen fixation in the symbiosis between WSM1271 and *B. pelecinus*.

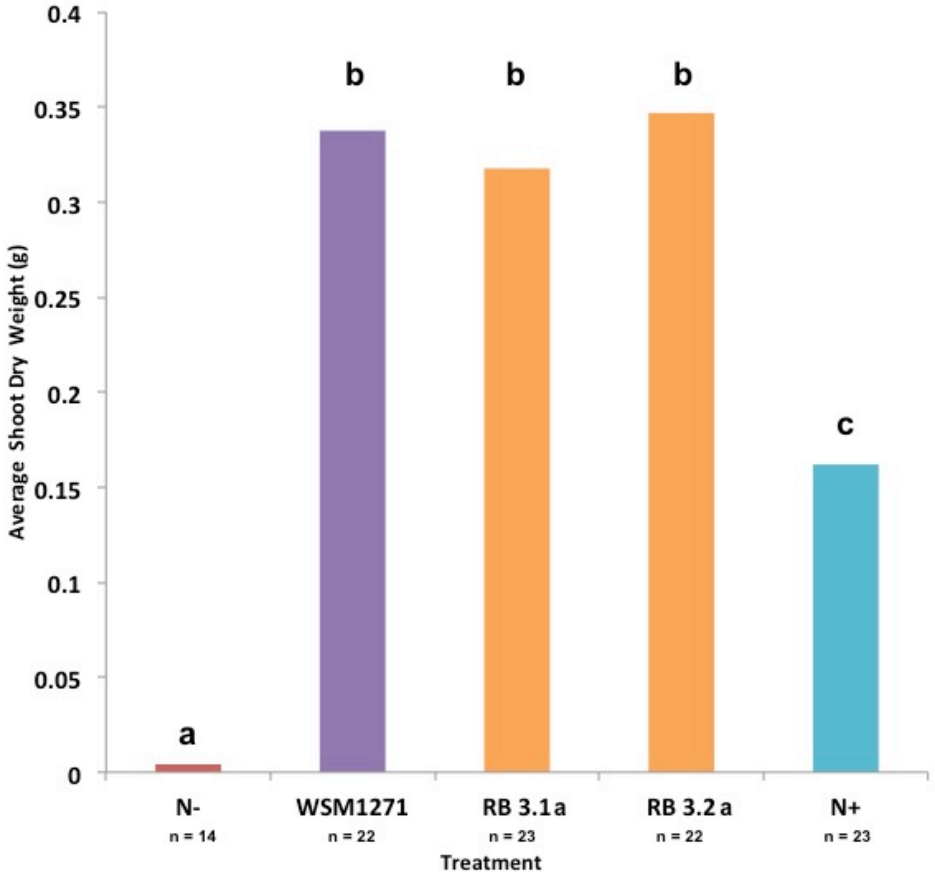


**Figure 3.9:** Mean shoot dry weights of *Biserrula pelecinus L*. grown in free-draining pots for 61 days. Plants were inoculated with either WSM1271 wild-type (purple) or plasmid-cured versions of WSM1271 (RB3.1a or RB3.2a) (orange), or un-inoculated nitrogen starved (N-) (red) or un-inoculated nitrogen fed (N+) (orange) controls. Treatments that share a letter are not significantly different following Tukey multiple comparisons of means following one-way analysis of variance ($P<0.5$). The number of plants per treatment (n) is shown.

## 3.5 Comparison of pMESCI01 sequence to other *Mesorhizobium spp.* plasmids

In order to investigate sequence similarity between pMESCI01 and plasmids from other completely sequenced *Mesorhizobium* spp., a comparative bioinformatic analysis was conducted. This analysis was performed by comparing pMESCI01 with plasmids from *M. ciceri* CC1192 (pMc1192), *M. ciceri* bv. biserrulae WSM1284 (pMc1284) and *M. ciceri* bv. biserrulae WSM1497 (pMc1497) by BRIG alignment (Section 2.2). This analysis revealed a surprisingly high similarity between all four plasmids (Figure 3.10). Plasmid pMc1192 was the most similar with approximately 77% of the plasmid sharing perfect identity with pMESCI01. Analysis also found that 70% of pMc1284 and 65% of pMc1497 sequences also identical to pMESCI01.

In particular, a region spanning approximately 35-kb containing response regulators and transferases as well as the putative nitrogen fixation genes of pMESCI01 (*fixGHIS, fixNOQP, fixLJ, fixK,* and *nifB*) (coordinates ~415,000-14,000 in pMESCI01) was 100% identical between pMESCI01 and pMc1192. However, this same region appeared to be absent and did not align to the pMc1284 and pMc1497 plasmids (Figure 3.10). The *repABC* region of pMESCI01 was 100% identical across all plasmids (coordinates ~5,200-36,000 in pMESCI01) (Figure 3.10). Other regions of high similarity across all plasmids included genes encoding a putative T6SS, including the predicted ImpK encoding Mesci_6367 and Mesci_6379 (Section 3.1.2) (coordinates ~300,000-340,000 in pMESCI01); a region with genes encoding for mostly metabolism or energy function, as well as

transporter, regulators and some hypothetical proteins (coordinates ~220,000-270,000 in pMESCI01); a region containing genes mostly encoding hypothetical proteins as well as cytochrome c oxidase $aa_3$-type subunits (coordinates ~75,000-150,000 in pMESCI01); and another region encoding mostly for energy and transporter functions that appears to be well-conserved across all plasmids analysed (coordinates ~165,000-190,000 in pMESCI01) (Figure 3.10).

Thus, a relatively large proportion of regions of pMESCI01 appear to be shared by plasmids of other *Mesorhizobium spp.*.

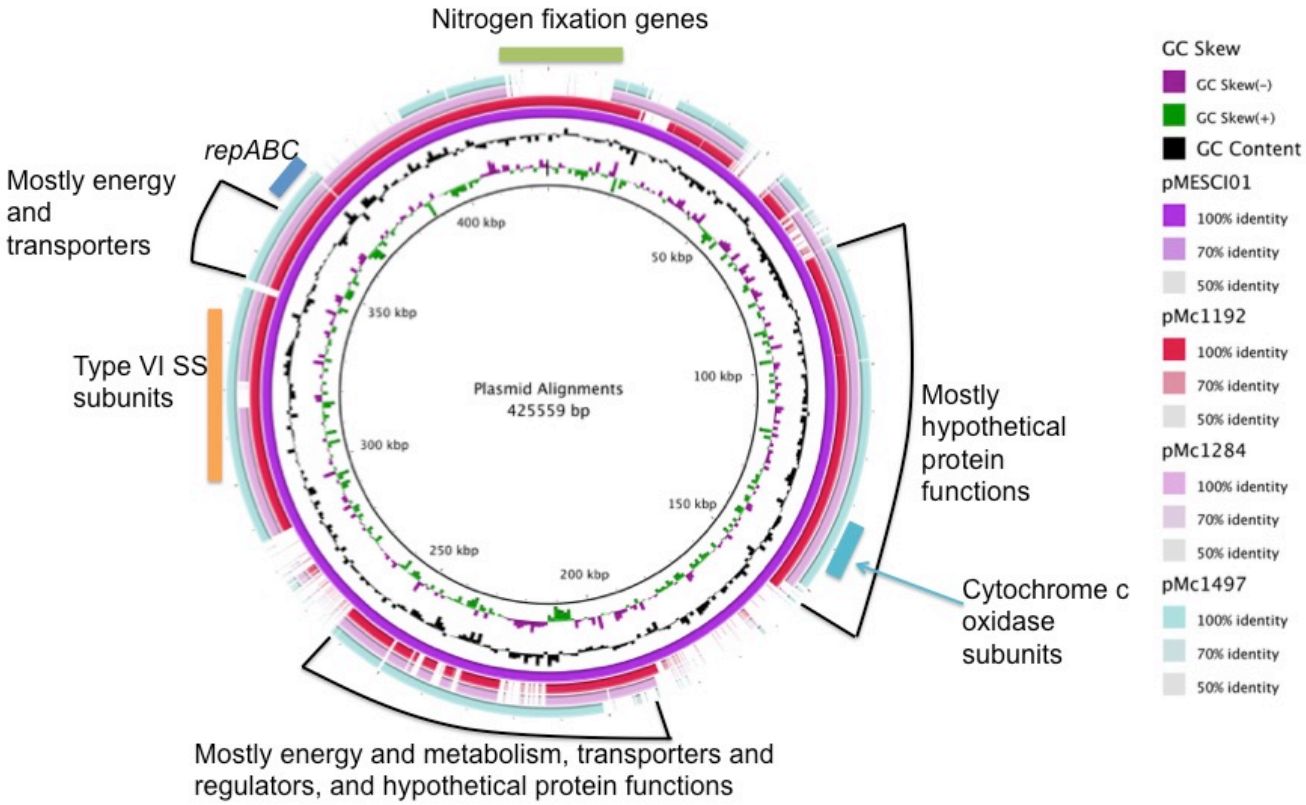


**Figure 3.10:** Plasmid comparison of *M. ciceri* bv. biserrulae, pMESCI01, to other *M. ciceri spp.* plasmids; *M. ciceri* CC1192 (pMc1192), *M. ciceri* bv. biserrulae WSM1284 (pMc1284), and *M. ciceri* bv. biserrulae WSM1497 (pMc1497). BRIG alignment was generated using BLAST+ database.

# 4. Discussion

The objective of this study was to investigate the role of the *M. ciceri* bv. biserrulae WSM1271 plasmid pMESCI01 in symbiosis with *Biserrula pelecinus.* Four main approaches were undertaken to address this objective: a bioinformatic analysis of pMESCI01 encoded genes; construction of a marked version of pMESCI01 to assess plasmid self-transmissibility; curing of pMESCI01 from the WSM1271 genome; and assessment of the phenotype of plasmid-cured WSM1271 derivatives in symbiosis with *B. pelecinus.*

## 4.1 Marking of pMESCI01 with Ω-Sp/Sm cassette to assess self-transmissibility

Bioinformatic analysis of pMESCI01 revealed that although the plasmid encodes a replication and partitioning system (*repABC*) and an origin of transfer (*oriT*), it appears to lack genes known to be necessary for self-transmissibility, including a relaxase or T4SS (Section 3.1.2). This suggests that pMESCI01 is not self-transmissible. To provide a functional means to test the self-transmissibility of pMESCI01 the plasmid was marked with the Ω-Sp/Sm cassette (pMESCI01::Ω-Sp/Sm), yielding plasmid-marked derivatives of WSM1271 named RB7.2 and RB8.3 (Section 3.2.2). The development of a marked version of pMESCI01 is a useful tool to assess plasmid self-transmissibility. Transfer of pMESCI01::Ω-Sp/Sm to the plasmid-less WSM2073Nm was attempted by selecting with $Sp_{200}$ and $Nm_{250}$, however, the $Sp^R$ of WSM2073Nm meant that the rate of plasmid transfer could not be assessed. WSM2073 has been previously described as sensitive to $Sp_{50}$ (Nandasena *et al.*, 2007) so it is possible that WSM2073Nm, which carries $Nm^R$ encoded on its broad host range plasmid (pPROBE), could

have subsequently developed spontaneous resistance to Sp. However, as Nandasena *et al.* (2007) used an antibiotic impregnated filter-disc to assess the sensitivity of WSM2073 to Sp, as opposed to being added directly to the medium as was done in this work, it is possible that the resistance of WSM2073Nm is simply an artefact of the different approaches. However, WSM2073Nm is highly sensitive to Sm, with a reduction in viable counts for WSM2073Nm of nine orders of magnitude on G/RDM supplemented with $Sp_{200}$, $Nm_{250}$ and $Sm_{200}$, compared to G/RDM supplemented with $Nm_{250}$ alone. Therefore, a subsequent experiment selecting WSM2073Nm-harbouring pMESCI01 transconjugants with the addition of $Sm_{200}$ will likely allow for a more accurate assessment of the self-transmissibility of this plasmid.

## 4.2 pMESCI01 is not essential for nitrogen fixation in WSM1271 with *B. pelecinus*

Plasmid-cured derivatives of WSM1271 were successfully developed via a plasmid incompatibility approach, in which pMESCI01 was removed from the WSM1271 genome, facilitated by the plasmid curing vector pSRK*repABC* constructed in this study (Section 3.3.2). As plasmids pMC1192 (*M. ciceri* CC1192), pMc1497 (*M. ciceri* WSM1497) and pMc1284 (*M. ciceri* bv. biserrulae WSM1284) were shown to harbour an identical *repABC* region to that of WSM1271 it is possible that pSRK*repABC* could also be used to cure the plasmids in these strains.

Subsequent symbiotic analysis of the plasmid-cured derivatives of WSM1271 (RB3.1a and RB3.2a) compared to wild-type WSM1271 inoculated onto *B.*

*pelecinus* yielded no differences in nodule number, morphology or size (Figure 3.8) or shoot dry weights of plants (Figure 3.9). The mean shoot dry weights of the N+ controls plants were significantly less than WSM1271 (Section 3.4.2), which was unexpected, as previous reports by Nandasena *et al.* (2007) found that the mean shoot dry weight of the N+ controls were significantly higher than that of WSM1271 inoculated *B. pelecinus*. Following a doubling in nitrogen provided as $KNO_3$ to the N+ controls at three weeks from 5ml to 10ml (Section 2.9.1), the rate of plant growth appeared to increase. Thus, it is highly likely that the N+ controls were nitrogen-limited. The nitrogen requirements or preferences of *B. pelecinus* remain largely unquantified. A study to determine the optimum nitrogen source and feed rate would be invaluable as it is possible that *B. pelecinus* may exhibit better growth with an alternative source of nitrogen, such as $NH_4NO_3$, or with a different rate or frequency of nitrogen application. Nevertheless, symbiotic testing of plasmid-cured pMESCI01 derivatives demonstrated that the plasmid is not essential for effective nitrogen fixation of WSM1271 with *B. pelecinus*.

Bioinformatic analysis did reveal the presence of 12 putative nitrogen fixation genes on pMESCI01 (*fixGHIS*, *fixNOQP*, *nifB*, *fixLJ*, and *fixK*). Both the *fixGHIS* and *fixNOQP* operons have a second copy located on ICE*Mc*Sym$^{1271}$. These ICE*Mc*Sym$^{1271}$ located operons are likely functional in WSM1271 as there was no partial reduction in mean shoot dry weights of plasmid-cured strains inoculated with *B. pelecinus* in comparison to WSM1271 wild-type inoculated plants (Section 3.4.2). A study by Schluter *et al.* (1997) inactivated duplicate copies of *fixN* from separate plasmid copies of the *fixNOQP* operons in *Rhizobium*

*leguminosaruem* bv. viciae VF39. Single plants inoculated with either *fixN* mutant exhibited decreased shoot dry weights in comparison to the parent strain, while plants inoculated with double *fixN* mutants (i.e. inactivated plasmid and chromosomal copies) were completely Fix$^-$, indicating that both copies of *fixN* were functional and that at least one copy of *fixN* was essential for nitrogen fixation. It is therefore likely that the pMESCI01-located putative *fixGHIS* and *fixNOQP* operons are non-functional in WSM1271 or redundant for nitrogen fixation with *B. pelecinus* due to there being no reduction in plant mean shoot dry weights.

The *nifB* gene was also located on both pMESCI01 and ICE*Mc*Sym$^{1271}$, with these two copies sharing ~78% identity at the protein level (Section 3.1.1). NifB functions in the synthesis of FeMo-cofactor alongside NifEN (Section 1.3). The operon that encodes NifEN was absent from pMESCI01, but located on ICE*Mc*Sym$^{1271}$ (Section 3.1.1). Thus, although the associated *nifEN* operon appears to be present on ICE*Mc*Sym$^{1271}$ the divergence between each NifB copy may suggest that the pMESCI01-encoded copy may not associate with ICE*Mc*Sym$^{1271}$-encoded NifEN. It is also likely that the plasmid-located *nifB* is non-functional in the WSM1271 background due to there being no partial reduction in shoot dry weights observed between plasmid-cured and wild-type WSM1271 inoculated *B. pelecinus* (Section 3.4.2).

Single copies of the *fixLJ* and *fixK* regulators were located on pMESCI01, as well as within the chromosome (outside of the ICE*Mc*Sym$^{1271}$) of WSM2073 and WSM2075, but were absent entirely from the WSM1271 chromosome (including

ICE*Mc*Sym$^{1271}$) (Section 3.1.1). Due to no reduction in shoot dry weights observed between plasmid-cured and wild-type WSM1271 inoculated *B. pelecinus* (Section 3.4.2) these plasmid-located regulators are likely non-functional or redundant in the WSM1271 background. While the mechanism of action of this regulatory system, which ultimately activates the transcriptional regulator NifA, differs across rhizobia (Terpolilli *et al.*, 2012), there are no reports on its function in *Mesorhizobium spp.* (Buikema *et al.*, 1985; Fischer, 1994; Dixon and Kahn, 2004); (Section 1.3). It is therefore possible that an as yet undescribed regulatory system present within ICE*Mc*Sym$^{1271}$ is required for the regulation of NifA in the low oxygen environment of the nodule (Section 1.3).

Although symbiotic testing with plasmid-cured WSM1271 derivatives showed that the putative symbiosis genes encoded on pMESCI01 are not essential to nitrogen fixation with *B. pelecinus*, it does not mean that the genes are necessarily non-functional or not expressed. In order to determine whether any of these putative nitrogen fixation genes are expressed in the WSM1271 background, transcriptomics of wild-type WSM1271 and plasmid-cured WSM1271 bacteroids could be performed. Transcriptomics of these bacteroids in low $O_2$ conditions may also reveal the presence of novel NifA regulators in ICE*Mc*Sym$^{1271}$.

A search conducted to locate Type II TA systems on pMESCI01 revealed four putative TA systems to be present on this plasmid (Section 3.1.2). TA systems are known to prevent the loss of plasmids from cell progeny through the addiction phenomenon (Section 1.4.1), and in some rhizobia functional TA systems make

plasmid curing a laborious process, such as curing of pSymB in *S. meliloti* (diCenzo *et al.*, 2014). The relative ease by which pMESCI01 was cured from the WSM1271 genome suggests that the multiple TA systems identified may be non-functional in WSM1271. Alternatively, the systems could be functional in wild-type WSM1271, but not when an additional vector, such as pSRKKm-*SacB*, is introduced carrying an identical origin of replication.

The plasmid-curing vector, pSRK*repABC*, constructed in this study could be used for curing plasmids with similar *repABC* region. Specifically, the approach established in this study can be used to remove the plasmid from other *Mesorhizobium* spp., such as commercial inoculant CC1192 plasmid pMc1192 or WSM1497 plasmid pMc1497, which were shown in this study to contain 100% similar *repABC* regions (Section 3.5). Inactivation of the plasmids of these agriculturally important strains following the approach established in this study would provide vital information to the function of these plasmids.

## 4.3 Other possible roles for pMESCI01

In addition to the putative symbiosis, plasmid replication and TA system genes, pMESCI01 also encodes 409 other genes, many of whose possible functions are unknown (Section 3.1). Moreover, there is a large diversity in the COG categories across these genes with 51.6% having no predicted function. Some of these genes may be important in WSM1271. For example, 7.2% of genes were predicted to be involved in energy conversion and production or metabolism (Section 3.1). These genes could encode an important metabolic function. No differences in growth of WSM1271 wild-type and plasmid-cured strains was

observed in TY or in minimal media with glucose or sucrose as the sole carbon source and $NH_4Cl$ as the sole nitrogen source. However, it is possible that the plasmid could encode anabolic or catabolic enzymes of some other nutrient or metabolite. Comparing the growth of wild-type WSM1271 with plasmid-cured derivatives of WSM1271 would help to uncover possible roles for pMESCI01. One approach would be by measuring growth rates on different nutrients or metabolites, such as can be achieved with the BioLog assay system (Biolog Inc., Hayward, CA, USA; http://www.biolog.com/microID.html). This technique can be used in the identification of bacterial species, specifically their metabolic functions, and could reveal whether any putative metabolic genes located on pMESCI01 are beneficial for growth of WSM1271.

A putative Type VI secretion system (T6SS) was located on pMESCI01 following bioinformatic analysis (Section 3.1.2) which, if functional, could play a role in defence of WSM1271, as T6SS typically secrete effector molecules that exhibit antimicrobial activity (Ryu, 2015). To analyse the function of these putative T6SS genes, as well as genes with hypothetical functions, construction of a library of gene-inactivation mutants of WSM1271 could be developed. Phenotypic assessment of these gene-inactivated WSM1271 mutants could determine the function of the inactivated genes, potentially revealing a potential role for pMESCI01 in WSM1271.

Interestingly, RNA sequencing of WSM1271 has revealed a low-level of expression of pMESCI01-encoded genes compared to chromosomally-encoded genes for TY-grown cells in stationary phase (T. Haskett, unpublished data). This

suggests that the genes encoded on pMESCI01 have little impact on WSM1271 in these conditions. However, it is possible that genes located on the plasmid are expressed in WSM1271 bacteroids, even though they are expressed at low levels in these free-living conditions. Trancriptomics of nitrogen-fixing plasmid-cured WSM1271 bacteroids compared to wild-type WSM1271 would therefore be informative. Another possibility is that the genes of pMESCI01 may be redundant in that their function may not be required for survival of WSM1271. Construction of a library of pMESCI01 gene-inactivated mutants of WSM1271 would also be beneficial in identifying any essential genes encoded by pMESCI01.

## 4.4 Conclusions and Future Directions

While pMESCI01 is not essential for nitrogen fixation of WSM1271 in symbiosis with *Biserrula pelecinus*, the plasmid does encode a further 420 genes, which could have functions important to WSM1271 beyond symbiotic nitrogen fixation. Therefore, further work characterising the potential genetic and metabolic roles of these genes in WSM1271 is required.

Comparative bioinformatic analysis through BRIG alignment revealed pMESCI01 to be highly similar to plasmids harboured by other *Mesorhizobium spp.* isolated from geographically diverse origins. BRIG alignment revealed >65% identity across pMESCI01 compared to plasmids of *M. ciceri* CC1192 (pMc1192) from Israel (Haskett *et al.*, 2016b), *M. ciceri* bv. biserrulae WSM1284 (pMc1284) from Sardinia (Haskett *et al.*, 2016a) and *M. ciceri* bv. biserrulae WSM1497 (pMc1497) from Mykonos, (Howieson *et al.*, 1995; Nandasena *et al.*, 2001); (Figure 3.10). It is interesting to speculate on how strains isolated from geographically diverse

locations happen to share highly similar plasmids. One possibility is that the plasmid has been vertically transmitted (i.e. from mother to daughter cells) from an ancestral cell to progeny cells, which subsequently spread over time throughout the Mediterranean region. The second possibility is that the plasmid is highly transmissible and capable of easily spreading via horizontal transfer to plasmid-less recipients.

This study revealed that a relaxase and T4SS were absent from pMESCI01, which are components thought to be essential for plasmid self-transmissibility. Although pMESCI01 appears to lack the components necessary to be self-transmissible, it is possible that it can be mobilised in the presence of a self-transmissible plasmid, due to the presence of a putative *oriT* site on pMESCI01 (Section 3.1.2). To be mobilisable, a plasmid theoretically only requires an *oriT* site that is then nicked by a relaxase encoded by a second self-transmissible plasmid (Ding and Hynes, 2009); (Section 1.4.1). A second plasmid (carrying self-transmissibility components) would therefore be required within the genome of WSM1271, as well as the genome of the *M. ciceri* strains, for their plasmids to be mobilised. The absence of a secondary plasmid across all these sequenced strains is consistent with these plasmids having undergone vertical transmission. It is possible that there may be as yet unidentified *Mesorhizobium* strains that do harbour additional plasmids (containing self-transmissibility components) which could be responsible for mobilising the replicons in *Mesorhizobium* spp.. However, it is equally possible that pMESC01, along with the plasmids from other strains of *M. ciceri* analysed in this thesis, are capable of self-transmission. Therefore, there is currently insufficient information to

determine whether vertical or horizontal transmission of these *M. ciceri* plasmids is the most likely explanation for their apparent wide distribution. Performing the additional mating experiments indicated previously for marked pMESCI01::Ω-Sp/Sm (Section 4.1), as well as isolating and sequencing more *Mesorhizobium* strains are therefore critical to being able to answer this question and unravel the nature and role of these plasmids.

# 5. Bibliography


Aguilar, O.M., J. Taormino, B. Thony, T. Ramseier, H. Hennecke and A.A. Szalay, 1990. The *nifEN* genes participating in FeMo cofactor biosynthesis and genes encoding dinitrogenase are part of the same operon in *Bradyrhizobium* species. Molecular and General Genetics, 224(3): 413-420.

Alikhan, N.-F., N.K. Petty, N.L.B. Zakour and S.A. Beatson, 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics, 12(1): 402.

Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, E. De Castro, S. Duvaud, V. Flegel, A. Fortier and E. Gasteiger, 2012. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Research, 40(W1): W597-W603.

Barcellos, F.G., P. Menna, J.S.D. Batista and M. Hungria, 2007. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium (Ensifer) fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. Applied and Environmental Microbiology, 73(8): 2635-2643.

Beringer, J., 1974. R Factor Transfer in *Rhizobiurn zegurninosarum*. Journal of General Microbiology 84: 188-198.

Bertani, G., 1951. Studies On Lysogenesis I.: The Mode of Phage Liberation by Lysogenic *Escherichia coli*1. Journal of Bacteriology, 62(3): 293.

Bignell, C. and C.M. Thomas, 2001. The bacterial ParA-ParB partitioning proteins. Journal of Biotechnology, 91(1): 1-34.

Blanca-Ordonez, H., J.J. Oliva-Garcia, D. Perez-Mendoza, M.J. Soto, J. Olivares, J. Sanjuan and J. Nogales, 2010. pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. Journal of Bacteriology, 192(23): 6309-6312.

Brigle, K.E., M.C. Weiss, W.E. Newton and D.R. Dean, 1987. Products of the iron-molybdenum cofactor-specific biosynthetic genes, *nifE* and *nifN*, are structurally homologous to the products of the nitrogenase molybdenum-iron protein genes, *nifD* and *nifK*. Journal of Bacteriology, 169(4): 1547-1553.

Buikema, W.J., J.A. Klingensmith, S.L. Gibbons and F.M. Ausubel, 1987. Conservation of structure and location of *Rhizobium meliloti* and *Klebsiella pneumoniae nifB* genes. Journal of Bacteriology, 169(3): 1120-1126.

Buikema, W.J., W.W. Szeto, P.V. Lemley, W.H. Orme-Johnson and F.M. Ausubel, 1985. Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. Nucleic Acids Research, 13(12): 4539-4555.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T.L. Madden, 2009. BLAST+: architecture and applications. BMC Bioinformatics, 10(1): 421.

Cervantes-Rivera, R., F. Pedraza-López, G. Pérez-Segura and M.A. Cevallos, 2011. The replication origin of a *repABC* plasmid. BMC Microbiology, 11(1): 158.

Cytryn, E.J., S. Jitacksorn, E. Giraud and M.J. Sadowsky, 2008. Insights learned from pBTAi1, a 229-kb accessory plasmid from *Bradyrhizobium sp.* strain BTAi1 and prevalence of accessory plasmids in other *Bradyrhizobium sp.* strains. International Society for Microbial Ecology 2(2): 158-170.

David, M., M.L. Daveran, J. Batut, A. Dedieu, O. Domergue, J. Ghai, C. Hertig, P. Boistard and D. Kahn, 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell, 54(5): 671-683.

Denarie, J., F. Debelle and J.C. Prome, 1996. *Rhizobium* lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annual Review of Biochemistry, 65: 503-535.

diCenzo, G.C., A.M. MacLean, B. Milunovic, G.B. Golding and T.M. Finan, 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. Public Library of Science Genetics, 10(10): e1004742.

Ding, H. and M.F. Hynes, 2009. Plasmid transfer systems in the rhizobia. Canadian Journal of Microbiology, 55(8): 917-927.

Ding, H., C.B. Yip and M.F. Hynes, 2013. Genetic characterization of a novel rhizobial plasmid conjugation system in *Rhizobium leguminosarum* bv. viciae strain VF39SM. Journal of Bacteriology, 195(2): 328-339.

Dixon, R. and D. Kahn, 2004. Genetic regulation of biological nitrogen fixation. Nature Reviews Microbiology, 2(8): 621-631.

Earl, C.D., C.W. Ronson and F.M. Ausubel, 1987. Genetic and structural analysis of the *Rhizobium meliloti fixA, fixB, fixC,* and *fixX* genes. Journal of Bacteriology, 169(3): 1127-1136.

Fischer, H.M., 1994. Genetic regulation of nitrogen fixation in rhizobia. Microbiological Reviews, 58(3): 352-386.

Foussard, M., A.M. Garnerone, F. Ni, E. Soupene, P. Boistard and J. Batut, 1997. Negative autoregulation of the *Rhizobium meliloti fixK* gene is indirect and requires a newly identified regulator, FixT. Molecular Microbiology, 25(1): 27-37.

Fowler, D., M. Coyle, U. Skiba, M.A. Sutton, J.N. Cape, S. Reis, L.J. Sheppard, A. Jenkins, B. Grizzetti, J.N. Galloway, P. Vitousek, A. Leach, A.F. Bouwman, K. Butterbach-Bahl, F. Dentener, D. Stevenson, M. Amann and M. Voss, 2013. The global nitrogen cycle in the twenty-first century. Philosophical Transactions of the Royal Society of London, 368(1621): 20130164.

Fuhrmann, M. and H. Hennecke, 1984. *Rhizobium japonicum* nitrogenase Fe protein gene (*nifH*). Journal of Bacteriology, 158(3): 1005-1011.

Gage, D.J., 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiology and Molecular Biology Reviews, 68(2): 280-300.

Galloway, J.N., W.H. Schlesinger, H. Levy, A. Michaels and J.L. Schnoor, 1995. Nitrogen-Fixation - Anthropogenic Enhancement-Environmental Response. Global Biogeochemical Cycles, 9(2): 235-252.

Gao, F., H. Luo and C.-T. Zhang, 2013. DoriC 5.0: an updated database of *oriC* regions in both bacterial and archaeal genomes. Nucleic Acids Research, 41(D1): D90-D93.

Giraud, E., L. Moulin, D. Vallenet, V. Barbe, E. Cytryn, J.C. Avarre, M. Jaubert, D. Simon, F. Cartieaux, Y. Prin, G. Bena, L. Hannibal, J. Fardoux, M. Kojadinovic, L. Vuillet, A. Lajus, S. Cruveiller, Z. Rouy, S. Mangenot, B. Segurens, C. Dossat, W.L. Franck, W.S. Chang, E. Saunders, D. Bruce, P. Richardson, P. Normand, B. Dreyfus, D. Pignol, G. Stacey, D. Emerich, A. Vermeglio, C. Medigue and M. Sadowsky, 2007. Legumes symbioses: absence of *nod* genes in photosynthetic Bradyrhizobia. Science, 316(5829): 1307-1312.

Goeders, N. and L. Van Melderen, 2014. Toxin-antitoxin systems as multilevel interaction systems. Toxins (Basel), 6(1): 304-324.

Goel, S., 2009. Factors Restricting Nitrogen Fixation In Root Nodule Bacteria Of The Pasture Legume *Biserrula pelecinus* cv. Casbah. In: School of Biological Sciences and Biotechnology. Murdoch Univeristy, Perth, Australia.

Graham, P.H. and C.P. Vance, 2003. Legumes: Importance and constraints to greater use. Plant Physiology, 131(3): 872-877.

Gruber, N. and J.N. Galloway, 2008. An Earth-system perspective of the global nitrogen cycle. Nature, 451(7176): 293-296.

Guo, Y., C. Echavarri-Erasun, M. Demuez, E. Jimenez-Vicente, E.L. Bominaar and L.M. Rubio, 2016. The nitrogenase FeMo-Cofactor precursor formed by NifB protein: a diamagnetic cluster containing eight iron atoms. Angewandte Chemie International Edition in English, 55(41): 12764-12767.

Haskett, T., P. Wang, J. Ramsay, G. O'Hara, W. Reeve, J. Howieson and J. Terpolilli, 2016a. Complete Genome Sequence of *Mesorhizobium ciceri* bv. biserrulae Strain WSM1284, an Efficient Nitrogen-Fixing Microsymbiont of the Pasture Legume Biserrula pelecinus. Genome Announc, 4(3).

Haskett, T., P. Wang, J. Ramsay, G. O'Hara, W. Reeve, J. Howieson and J. Terpolilli, 2016b. Complete Genome Sequence of *Mesorhizobium ciceri* Strain CC1192, an Efficient Nitrogen-Fixing Microsymbiont of *Cicer arietinum*. Genome Announcements, 4(3).

Haskett, T.L., J.J. Terpolilli, A. Bekuma, G.W. O'Hara, J.T. Sullivan, P. Wang, C.W. Ronson and J.P. Ramsay, 2016c. Assembly and transfer of tripartite integrative and conjugative genetic elements. Proceedings of the National Academy of Sciences of the United States of America, 113(43): 12268-12273.

Herridge, D.F., M.B. Peoples and R.M. Boddey, 2008. Global inputs of biological nitrogen fixation in agricultural systems. Plant and Soil, 311(1-2): 1-18.

Hoover, T.R., J. Imperial, P.W. Ludden and V.K. Shah, 1988. Homocitrate cures the NifV- phenotype in *Klebsiella pneumoniae*. Journal of Bacteriology, 170(4): 1978-1979.

Horvath, B., C.W. Bachem, J. Schell and A. Kondorosi, 1987. Host-specific regulation of nodulation genes in *Rhizobium* is mediated by a plant-signal, interacting with the *nodD* gene product. European Molecular Biology Organisation, 6(4): 841-848.

Howieson, J. and M.J. Dilworth (Eds.), 2016. Working with rhizobia. Canberra: Australian Centre for International Agriculture Research.

Howieson, J.G., A. Loi and S.J. Carr, 1995. *Biserrula pelecinus* L. - a legume pasture species with potential for acid, duplex soils which is nodulated by unique root-nodule bacteria. Australian Journal of Agricultural Research, 46(5): 997-1009.

Imperial, J., T.R. Hoover, M.S. Madden, P.W. Ludden and V.K. Shah, 1989. Substrate reduction properties of dinitrogenase activated in vitro are dependent upon the presence of homocitrate or its analogues during iron-molybdenum cofactor synthesis. Biochemistry, 28(19): 7796-7799.

Imperial, J., R.A. Ugalde, V.K. Shah and W.J. Brill, 1984. Role of the *nifQ* gene product in the incorporation of molybdenum into nitrogenase in *Klebsiella pneumoniae*. Journal of Bacteriology, 158(1): 187-194.

Jacobson, M.R., V.L. Cash, M.C. Weiss, N.F. Laird, W.E. Newton and D.R. Dean, 1989. Biochemical and genetic analysis of the *nifUSVWZM* cluster from *Azotobacter vinelandii*. Molecular and General Genetics, 219(1-2): 49-57.

Johnston, A.W.B., G. Hombrecher, N.J. Brewin and M.C. Cooper, 1982. Two transmissible plasmids in *Rhizobium leguminosarum* strain 300. Journal of General Microbiology, 128(Jan): 85-93.

Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada and S. Tabata, 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti* (supplement). DNA Research, 7(6): 381-406.

Kaneko, T., Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada and S. Tabata, 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110 (supplement). DNA Research, 9(6): 225-256.

Kasai-Maita, H., H. Hirakawa, Y. Nakamura, T. Kaneko, K. Miki, J. Maruya, S. Okazaki, S. Tabata, K. Saeki and S. Sato, 2013. Commonalities and differences among Symbiosis Islands of three *Mesorhizobium loti* strains. Microbes and Environments, 28(2): 275-278.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz and C. Duran, 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12): 1647-1649.

Klipp, W., H. Reilander, A. Schluter, R. Krey and A. Puhler, 1989. The *Rhizobium meliloti fdxN* gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with *nifA* and *nifB*. Molecular General Genetics, 216(2-3): 293-302.

Laranjo, M., A. Alexandre and S. Oliveira, 2014. Legume growth-promoting rhizobia: an overview on the *Mesorhizobium* genus. Microbiology Research, 169(1): 2-17.

Lawley, T.D., W.A. Klimke, M.J. Gubbins and L.S. Frost, 2003. F factor conjugation is a true type IV secretion system. FEMS Microbiology Letters, 224(1): 1-15.

Liu, X.K., D.S. Wang, H.G. Wang, E.L. Feng, L. Zhu and H.L. Wang, 2012. Curing of plasmid pXO1 from *Bacillus anthracis* using plasmid incompatibility. Public Library of Science One, 7(1).

MacLellan, S.R., L.A. Smallbone, C.D. Sibley and T.M. Finan, 2005. The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. Molecular Microbiology, 55(2): 611-623.

MacLellan, S.R., R. Zaheer, A.L. Sartor, A.M. MacLean and T.M. Finan, 2006. Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. Molecular Microbiology, 59(5): 1559-1575.

MacNeil, T., D. MacNeil, G.P. Roberts, M.A. Supiano and W.J. Brill, 1978. Fine-structure mapping and complementation analysis of *nif* (nitrogen fixation) genes in *Klebsiella pneumoniae*. Journal of Bacteriology, 136(1): 253-266.

Markowitz, V.M., I.-M.A. Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, B. Jacob, J. Huang and P. Williams, 2012. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Research, 40(D1): D115-D122.

Mazur, A. and P. Koper, 2012. Rhizobial plasmids—replication, structure and biological role. Open Life Sciences, 7(4): 571-586.

Melino, V.J., E.A. Drew, R.A. Ballard, W.G. Reeve, G. Thomson, R.G. White and G.W. O'Hara, 2012. Identifying abnormalities in symbiotic development between *Trifolium spp.* and *Rhizobium leguminosarum* bv. trifolii leading to sub-optimal and ineffective nodule phenotypes. Annuals of Botany, 110(8): 1559-1572.

Murray, J.D., 2011. Invasion by invitation: rhizobial infection in legumes. Molecular Plant-Microbe Interactions, 24(6): 631-639.

Nandasena, K., R. Yates, R. Tiwari, G. O'Hara, J. Howieson, M. Ninawi, O. Chertkov, C. Detter, R. Tapia, S. Han, T. Woyke, S. Pitluck, M. Nolan, M. Land, K. Liolios, A. Pati, A. Copeland, N. Kyrpides, N. Ivanova, L. Goodwin, U. Meenakshi and W. Reeve, 2014. Complete genome sequence of *Mesorhizobium ciceri* bv. *biserrulae* type strain (WSM1271(T)). Standards in Genomic Sciences, 9(3): 462-472.

Nandasena, K.G., W. O'Hara G, R.P. Tiwari and J.G. Howieson, 2006. Rapid in situ evolution of nodulating strains for *Biserrula pelecinus L*. through lateral transfer of a symbiosis island from the original mesorhizobial inoculant. Applied and Environmental Microbiology, 72(11): 7365-7367.

Nandasena, K.G., G.W. O'Hara, R.P. Tiwari, E. Sezmis and J.G. Howieson, 2007. In situ lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus L*. Environmental Microbiology, 9(10): 2496-2511.

Nandasena, K.G., G.W. O'Hara, R.P. Tiwari, A. Willems and J.G. Howieson, 2009. *Mesorhizobium australicum sp. nov.* and *Mesorhizobium opportunistum sp. nov*., isolated from *Biserrula pelecinus L*. in Australia. International Journal of Systematic and Evolutionary Microbiology, 59(Pt 9): 2140-2147.

Nandasena, K.G., G.W. O'Hara, R.P. Tiwari, R.J. Yates and J.G. Howieson, 2001. Phylogenetic relationships of three bacterial strains isolated from the pasture legume *Biserrula pelecinus L.* International Journal of Systematic and Evolutionary Microbiology, 51(Pt 6): 1983-1986.

Nelson, M.S. and M.J. Sadowsky, 2015. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. Frontiers in Plant Science, 6: 491.

Ni, B., Z. Du, Z. Guo, Y. Zhang and R. Yang, 2008. Curing of four different plasmids in *Yersinia pestis* using plasmid incompatibility. Letters in Applied Microbiology, 47(4): 235-240.

Nouwen, N., J. Fardoux and E. Giraud, 2016. NodD1 and NodD2 are not required for the symbiotic interaction of *Bradyrhizobium* ORS285 with Nod factor independent *Aeschynomene* legumes. Public Library of Sciences One, 11(6): e0157888.

Okazaki, S., P. Tittabutr, A. Teulet, J. Thouin, J. Fardoux, C. Chaintreuil, D. Gully, J.F. Arrighi, N. Furuta, H. Miwa, M. Yasuda, N. Nouwen, N. Teaumroong and E. Giraud, 2016. *Rhizobium*-legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. International Society for Microbiology Ecology Journal, 10(1): 64-74.

Oldroyd, G.E. and J.A. Downie, 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. Annual Review of Plant Biology, 59: 519-546.

Peng, J., B. Hao, L. Liu, S. Wang, B. Ma, Y. Yang, F. Xie and Y. Li, 2014. RNA-Seq and microarrays analyses reveal global differential transcriptomes of *Mesorhizobium huakuii* 7653R between bacteroids and free-living cells. Public Library of Science One, 9(4): e93626.

Pinto, U.M., K.M. Pappas and S.C. Winans, 2012. The ABCs of plasmid replication and segregation. Nature Reviews Microbiology, 10(11): 755-765.

Pitcher, R.S. and N.J. Watmough, 2004. The bacterial cytochrome *cbb*$_3$ oxidases. Biochimica et Biophysica Acta, 1655(1-3): 388-399.

Preisig, O., D. Anthamatten and H. Hennecke, 1993. Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis. Proceedings of the National Academy of Sciences of the United States of America, 90(8): 3309-3313.

Preisig, O., R. Zufferey and H. Hennecke, 1996. The *Bradyrhizobium japonicum fixGHIS* genes are required for the formation of the high-affinity *cbb*$_3$-type cytochrome oxidase. Archives of Microbiology, 165(5): 297-305.

Quandt, J. and M.F. Hynes, 1993. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene, 127(1): 15-21.

Ramirez-Romero, M.A., N. Soberon, A. Perez-Oseguera, J. Tellez-Sosa and M.A. Cevallos, 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. Journal of Bacteriology, 182(11): 3117-3124.

Reeve, W., K. Nandasena, R. Yates, R. Tiwari, G. O'Hara, M. Ninawi, O. Chertkov, L. Goodwin, D. Bruce, C. Detter, R. Tapia, S.S. Han, T. Woyke, S. Pitluck, M. Nolan, M. Land, A. Copeland, K. Liolios, A. Pati, K. Mavromatis, V. Markowitz, N. Kyrpides, N. Ivanova, L. Goodwin, U. Meenakshi and J. Howieson, 2013a. Complete genome sequence of *Mesorhizobium opportunistum* type strain WSM2075(T). Standards in Genomic Sciences, 9(2): 294-303.

Reeve, W., K. Nandasena, R. Yates, R. Tiwari, G. O'Hara, M. Ninawi, W. Gu, L. Goodwin, C. Detter, R. Tapia, C. Han, A. Copeland, K. Liolios, A. Chen, V. Markowitz, A. Pati, K. Mavromatis, T. Woyke, N. Kyrpides, N. Ivanova and J. Howieson, 2013b. Complete genome sequence of *Mesorhizobium australicum* type strain (WSM2073(T)). Standards in Genomic Sciences, 9(2): 410-419.

Richardson, K., L. Jarett and E. Finke, 1960. Embedding in epoxy resins for ultrathin sectioning in electron microscopy. Stain Technology, 35(6): 313-323.

Ronson, C.W. and S.B. Primrose, 1979. Carbohydrate metabolism in *Rhizobium trifolii*: identification and symbiotic properties of mutants. Microbiology, 112(1): 77-88.

Rubio, L.M. and P.W. Ludden, 2008. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annual Review of Microbiology, 62: 93-111.

Ryu, C.M., 2015. Against friend and foe: type 6 effectors in plant-associated bacteria. Journal of Microbiology, 53(3): 201-208.

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

Sayers, E.W., T. Barrett, D.A. Benson, E. Bolton, S.H. Bryant, K. Canese, V. Chetvernin, D.M. Church, M. DiCuccio and S. Federhen, 2011. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, 39(suppl 1): D38-D51.

Schluter, A., T. Patschkowski, J. Quandt, L.B. Selinger, S. Weidner, M. Kramer, L. Zhou, M.F. Hynes and U.B. Priefer, 1997. Functional and regulatory analysis of the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39. Molecular Plant-Microbe Interactions Journal, 10(5): 605-616.

Scott, J.D. and R.A. Ludwig, 2004. *Azorhizobium caulinodans* electron-transferring flavoprotein N electrochemically couples pyruvate dehydrogenase complex activity to $N_2$ fixation. Microbiology, 150(Pt 1): 117-126.

Shah, V.K., G. Stacey and W.J. Brill, 1983. Electron transport to nitrogenase. Purification and characterization of pyruvate:flavodoxin oxidoreductase. The *nifJ* gene product. The Journal of Biological Chemistry, 258(19): 12064-12068.

Spaink, H.P., 2000. Root nodulation and infection factors produced by rhizobial bacteria. Annual Review of Microbiology, 54: 257-288.

Sprent, J.I., 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. New Phytologist, 174(1): 11-25.

Sprent, J.I., 2009. Evolution of Nodulation. In: Legume Nodulation: A Global Perspective. Wiley-Blackwell, United Kingdom: pp: 54.

Sprent, J.I., J. Ardley and E.K. James, 2017. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. New Phytologist.

Spurr, A.R., 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. Journal of Ultrastructure Research, 26(1-2): 31-43.

Sullivan, J.T. and C.W. Ronson, 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a *phe-tRNA* gene. Proceedings of the National Academy of Sciences of the United States of America, 95(9): 5145-5149.

Sullivan, J.T., J.R. Trzebiatowski, R.W. Cruickshank, J. Gouzy, S.D. Brown, R.M. Elliot, D.J. Fleetwood, N.G. McCallum, U. Rossbach, G.S. Stuart, J.E. Weaver, R.J. Webby, F.J. De Bruijn and C.W. Ronson, 2002. Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. Journal of Bacteriology, 184(11): 3086-3095.

Szpirer, C.Y., M. Faelen and M. Couturier, 2000. Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob. Molecular Microbiology, 37(6): 1283-1292.

Tatusov, R.L., D.A. Natale, I.V. Garkavtsev, T.A. Tatusova, U.T. Shankavaram, B.S. Rao, B. Kiryutin, M.Y. Galperin, N.D. Fedorova and E.V. Koonin, 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Research, 29(1): 22-28.

Terpolilli, J.J., G.A. Hood and P.S. Poole, 2012. What determines the efficiency of $N_2$-fixing *Rhizobium*-legume symbioses? Advances in Microbial Physiology, 60: 325-389.

Terpolilli, J.J., S.K. Masakapalli, R. Karunakaran, I.U. Webb, R. Green, N.J. Watmough, N.J. Kruger, R.G. Ratcliffe and P.S. Poole, 2016. Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. Journal of Bacteriology, 198(20): 2864-2875.

Thoma, S. and M. Schobert, 2009. An improved *Escherichia coli* donor strain for diparental mating. FEMS Microbiology Letters, 294(2): 127-132.

Thomas, C.M., 2000. Paradigms of plasmid organization. Molecular Microbiology, 37(3): 485-491.

Thony, B., K. Kaluza and H. Hennecke, 1985. Structural and functional homology between the alpha-subunits and beta-subunits of the Nitrogenase MoFe protein as revealed by sequencing the *Rhizobium japonicum nifK* gene. Molecular and General Genetics, 198(3): 441-448.

Thony-Meyer, L., 1997. Biogenesis of respiratory cytochromes in bacteria. Microbiology and Molecular Biology Reviews, 61(3): 337-376.

Uraji, M., K. Suzuki and K. Yoshida, 2002. A novel plasmid curing method using incompatibility of plant pathogenic Ti plasmids in *Agrobacterium tumefaciens*. Genes & Genetic Systems, 77(1): 1-9.

Vance, C.P., 2001. Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. Plant Physiology, 127(2): 390-397.

Vanrhijn, P. and J. Vanderleyden, 1995. The *Rhizobium*-Plant Symbiosis. Microbiological Reviews, 59(1): 124-142.

Wang, S., B. Hao, J. Li, H. Gu, J. Peng, F. Xie, X. Zhao, C. Frech, N. Chen, B. Ma and Y. Li, 2014. Whole-genome sequencing of *Mesorhizobium huakuii* 7653R provides molecular insights into host specificity and symbiosis island dynamics. BMC Genomics, 15: 440.

Young, J.P., L.C. Crossman, A.W. Johnston, N.R. Thomson, Z.F. Ghazoui, K.H. Hull, M. Wexler, A.R. Curson, J.D. Todd, P.S. Poole, T.H. Mauchline, A.K. East, M.A. Quail, C. Churcher, C. Arrowsmith, I. Cherevach, T. Chillingworth, K. Clarke, A. Cronin, P. Davis, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, S. Whitehead and J. Parkhill, 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biology, 7(4): R34.

Zufferey, R., O. Preisig, H. Hennecke and L. Thony-Meyer, 1996. Assembly and function of the cytochrome *cbb*$_3$ oxidase subunits in *Bradyrhizobium japonicum*. The Journal of Biological Chemistry, 271(15): 9114-9119.

# 6. Appendix

**Table A1:** BLASTP searches with known relaxase genes of 29 plasmids across multiple genera of rhizobia and model plasmids with pMESCI01. The percentage query cover and identity of each gene is shown for each BLASTP search performed, as well as the strain each plasmid is found. These known relaxase genes used for comparison were selected from representative relaxase groups listed in Ding *et al.* (2013).

| **Strain** | **Plasmid** | **BLASTP Query cover %** | **BLASTP Identity %** |
|---|---|---|---|
| *Agrobacterium vitis* S4 | pTi54 | 32 | 24 |
| *Agrobacterium vitis* S4 | pAtS4c | 32 | 24 |
| *Agrobacterium radiobacter* K84 | pAtK84b | 32 | 24 |
| *Agrobacterium vitis* S4 | pAtS4e | 32 | 24 |
| *Rhizobium leguminosarum* bv. viciae 3841 | pRL8KI | 29 | 25 |
| *Sinorhizobium meliloti* SM11 | pSmeSM11b | 28 | 25 |
| *Sinorhizobium fredii* NGR234 | pNGR234a | 32 | 24 |
| *Sinorhizobium medicae* WSM419 | pSmed01 | 32 | 25 |
| *Sinorhizobium medicae* WSM419 | pSmed02 | 32 | 25 |
| *Rhizobium etli* CFN42 | pRetCFN42a | 32 | 25 |
| *Rhizobium leguminosarum* bv. viciae 3841 | pRL7JI | 31 | 26 |
| *Agrobacterium rhizogenes* | pRi1724 | 31 | 26 |
| *Agrobacterium rhizogenes* | pRi2659 | 31 | 24 |
| *Rhizobium leguminosarum* bv. trifoli WSM1325 | pR132503 | 31 | 25 |
| *Agrobacterium vitis* S4 | pAtS4b | 24 | 24 |
| *Rhizobium etli* CIAT652 | pRetCIAT652b | 24 | 24 |
| *Sinorhizobium meliloti* 1021 | pSymA | 24 | 24 |
| *Sinorhizobium fredii* NGR234 | pNGR234b | 23 | 24 |
| *Agrobacterium tumefacians* | pAgK84 | 34 | 26 |
| *Sinorhizobium meliloti* | pSmeSM11a | 29 | 25 |
| *Rhizobium leguminosarum* bv. viciae 3841 | pRL10JI | 33 | 24 |
| *Rhizobium leguminosarum* bv. viciae 3841 | pRL11JI | 33 | 23 |
| *Rhizobium leguminosarum* bv. viciae 3841 | pRL12JI | 33 | 23 |
| *Rhizobium leguminosarum* bv. trifoli WSM1325 | pR132504 | 33 | 25 |
| *Rhizobium leguminosarum* bv. viciae | pRleVF39b | NA | NA |
| *Rhizobium etli* CFN42 | pRetCFN42d | 21 | 26 |
| *Rhizobium etli* CIAT652 | pRetCIAT625c | 23 | 24 |
| *Rhizobium leguminosarum* bv. trifoli WSM1325 | pR132502 | 33 | 23 |
| *Rhizobium leguminosarum* bv. trifoli WSM1325 | pR132501 | 33 | 23 |

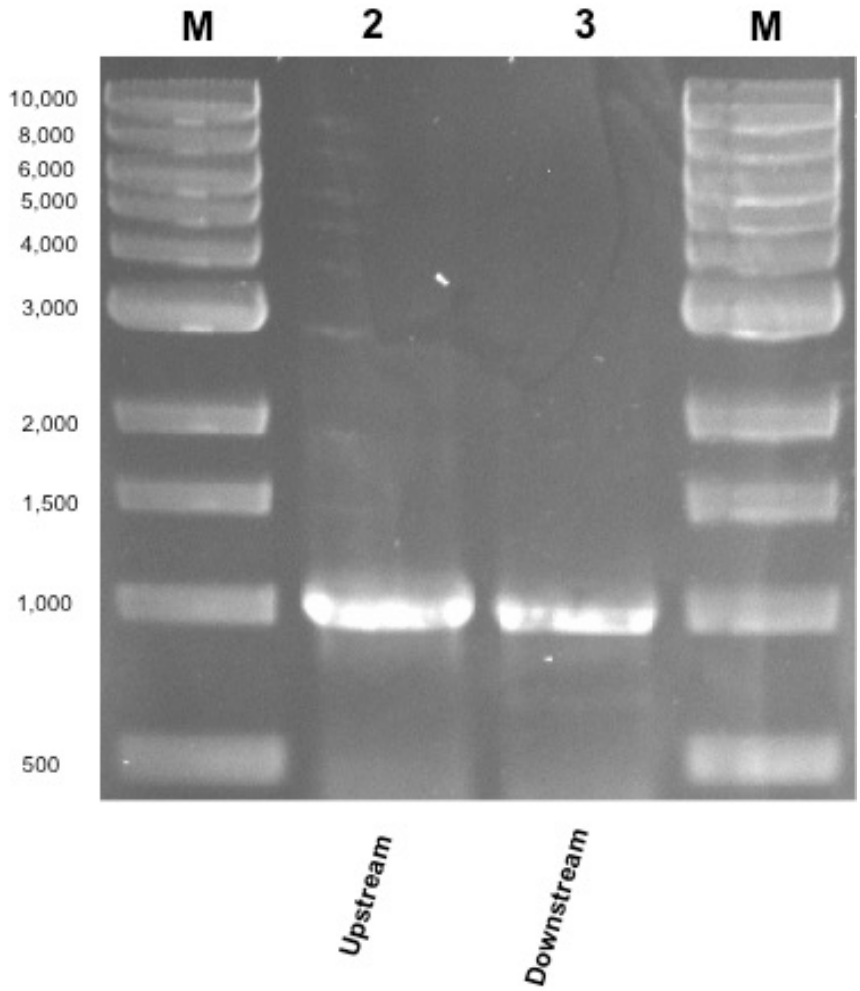


**Figure A1:** Image of electrophoresed PCR products of a 1-kb upstream region and 1-kb downstream region of pMESCI01. Lane 2, upstream region; lane 3, downstream region. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1% (w/v) Agarose in 1 X TAE gel was used.

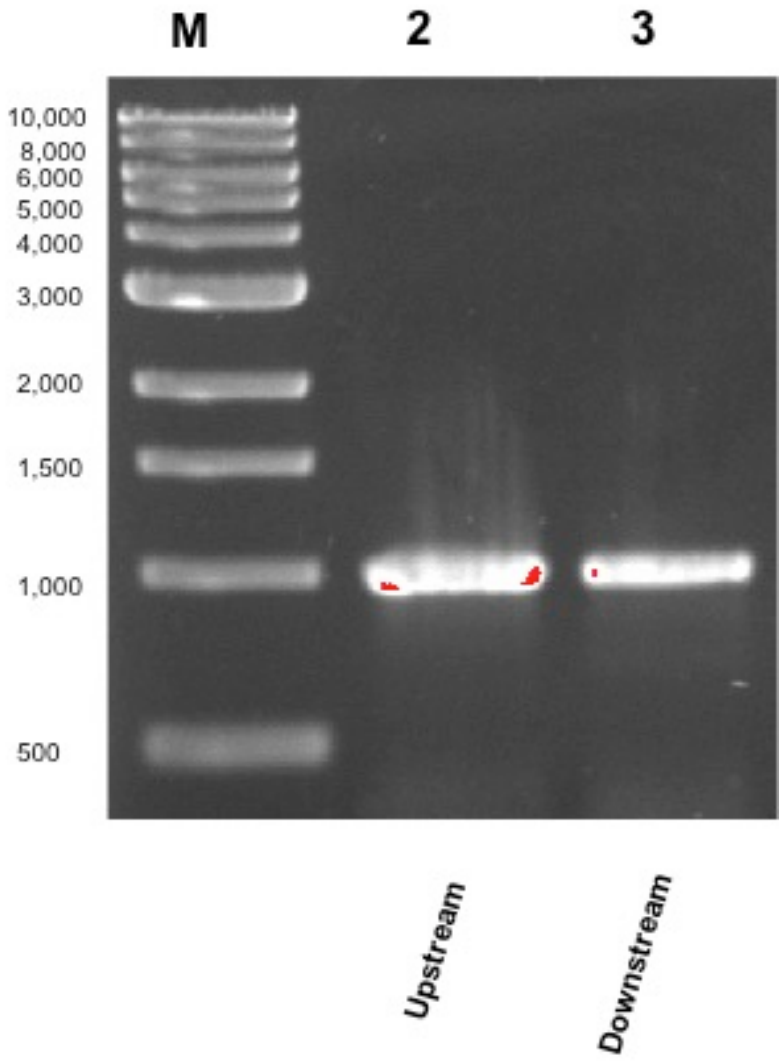


**Figure A2:** Image of electrophoresed digested products of a 1-kb upstream region and 1-kb downstream region of pMESCI01. Lane 2, upstream region digested with *Sac*I and *Not*I; lane 3, downstream region digested with *Xba*I and *Sal*I. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1% (w/v) Agarose in 1 X TAE gel was used.

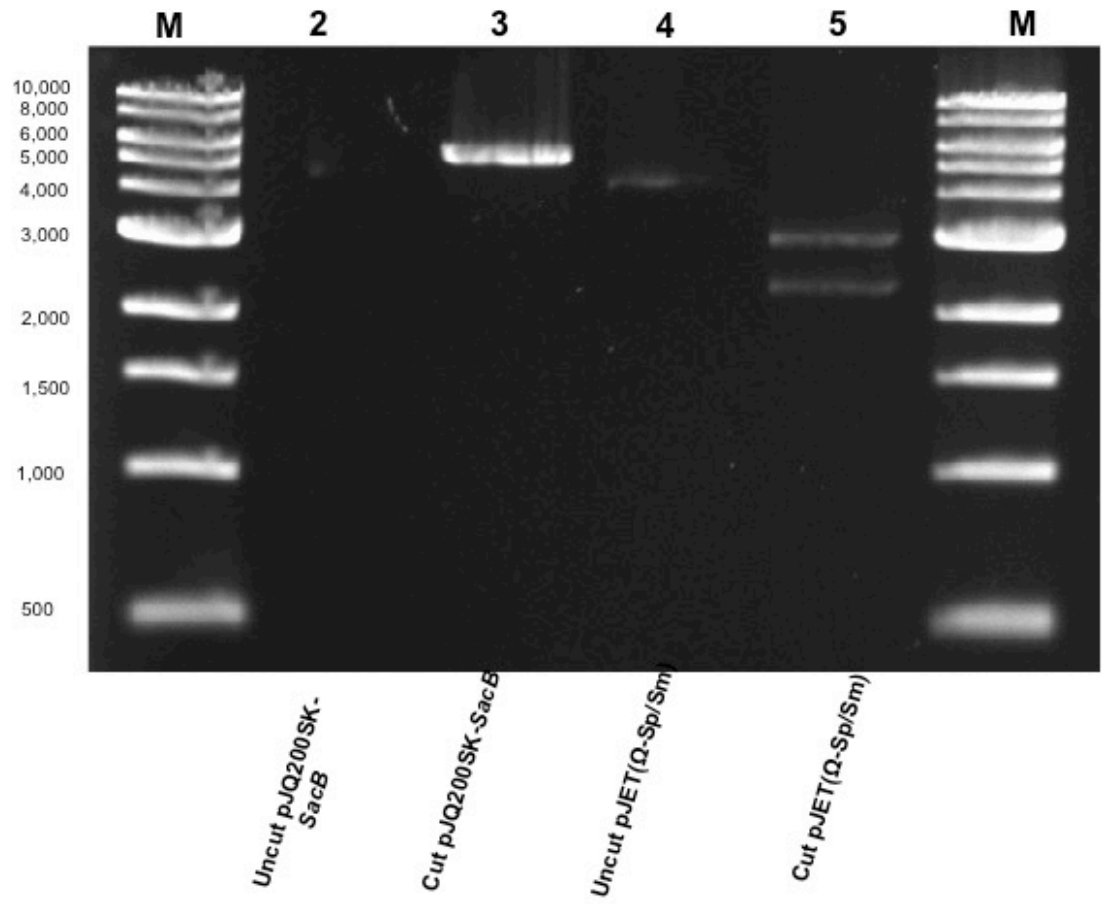


**Figure A3:** Image of electrophoresed digested or undigested products of pJQ200SK-*SacB* and pJET(Ω-Sp/Sm). Lane 2, undigested (uncut) pJQ200SK-*SacB*; lane 3, pJQ200SK-*SacB* digested (cut) with *Sal*I and *Sac*I; lane 3, uncut pJET(Ω-Sp/Sm); lane 4, cut pJET(Ω-Sp/Sm) with *Not*I and *Xba*I. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1% (w/v) Agarose in 1 X TAE gel was used.

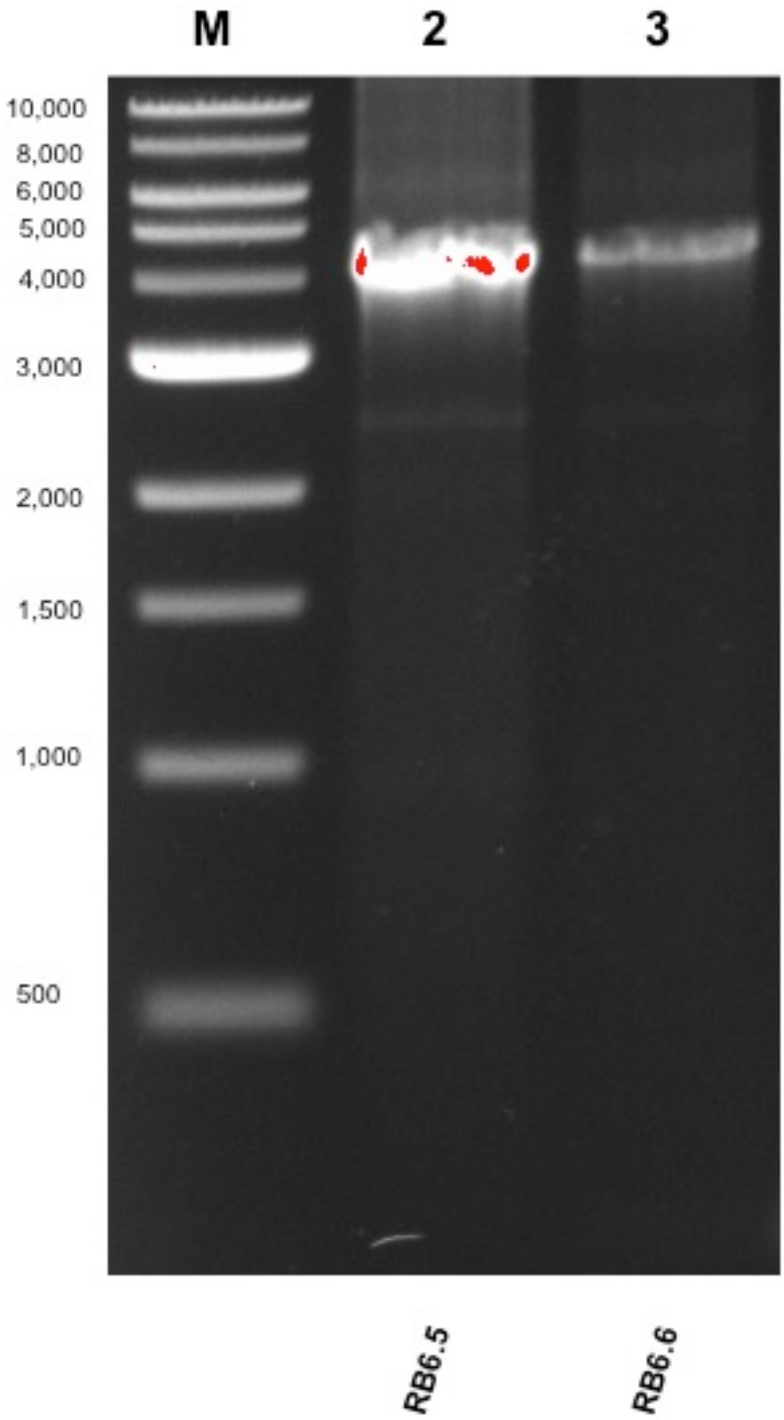


**Figure A4:** Image of electrophoresed digested products of PCR confirmed transformants containing plasmid-marking vector, pJQ(Ω-Sp/Sm), RB6.5 and RB6.6. Lane 2, RB6.5 digested with *Xho*I; lane 3, RB6.6 digested with *Xba*I The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1% (w/v) Agarose in 1 X TAE gel was used.

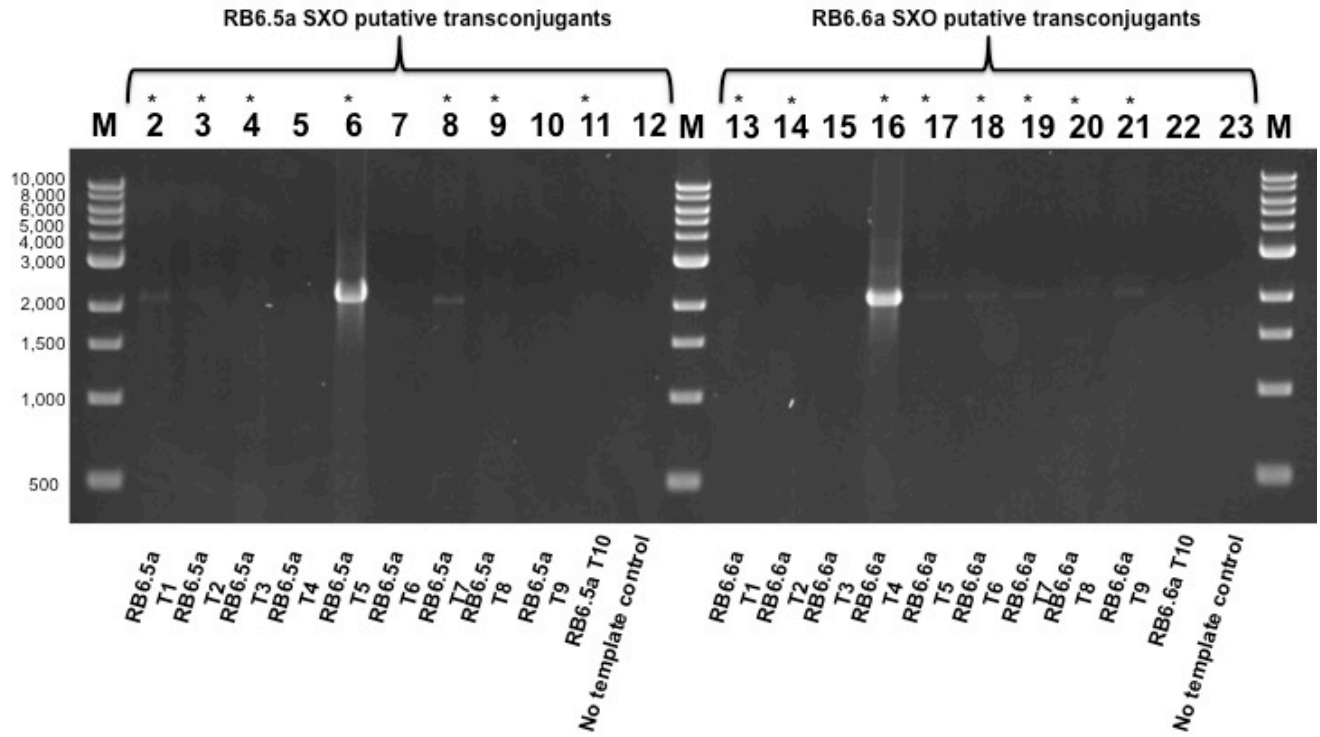


**Figure A5:** Image of electrophoresed PCR products of putative SXO marked pMESCI01 transformants from independent sources (RB6.5a or RB6.6a). Lane 2-11, PCR products of putative SXO marked pMESCI01 transformants from RB6.5a; lanes 13-22 PCR products of putative SXO marked pMESCI01 transformants from RB6.6a; lanes 12 and 23 contain no template controls. Transformants showing successful banding pattern of ~2-kb indicated with *. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

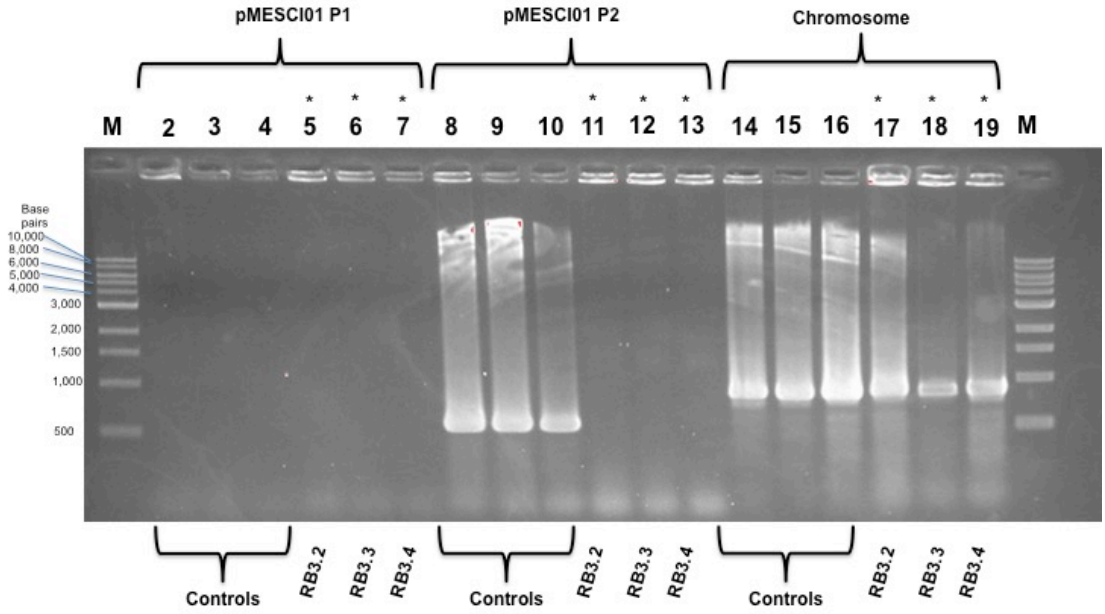


**Figure A6:** Image of electrophoresed PCR products of putative plasmid-cured WSM1271 transconjugants RB3.2, RB3.3, and RB3.4. Lane; 5, PCR product of RB3.2 with pMESCI01 located primer pair 1 (P1); lane; 6, PCR product of RB3.3 with P1; lane; 7, PCR product of RB3.4 with P1; lane 11, PCR product of RB3.2 with pMESCI01 located primer set 2 (P2); lane 12, PCR product of RB3.3 with P2; lane 13, PCR product of RB3.4 with P2; lane 17, PCR product of RB3.2 with WSM1271 chromosome located primer set; lane 18, PCR product of RB3.3 with chromosomal primer set; lane 19, PCR product of RB3.4 with chromosomal primer set; lanes 2-4, PCR product of WSM1271 with P1; lanes 8-10, PCR product of WSM1271 with P2; lanes 14-16 PCR product of WSM1271 with chromosomal primer set. The correct transconjugants are indicated with *. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

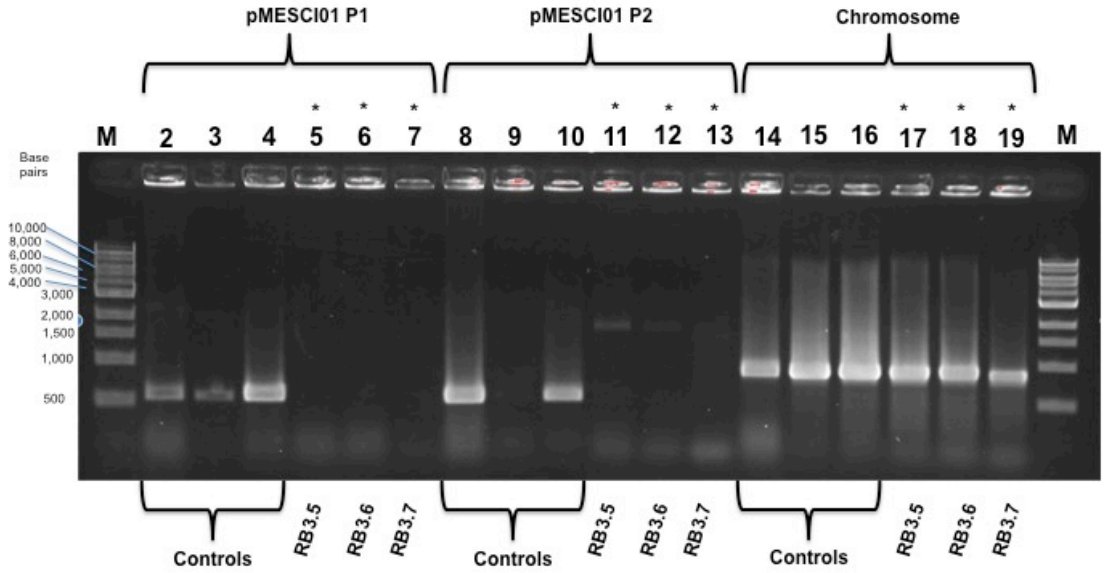


**Figure A7:** Image of electrophoresed PCR products of putative plasmid-cured WSM1271 transconjugants RB3.5, RB3.6, and RB3.7. Lane; 5, PCR product of RB3.5 with pMESCI01 located primer pair 1 (P1); lane; 6, PCR product of RB3.6 with P1; lane; 7, PCR product of RB3.7 with P1; lane 11, PCR product of RB3.5 with pMESCI01 located primer set 2 (P2); lane 12, PCR product of RB3.6 with P2; lane 13, PCR product of RB3.7 with P2; lane 17, PCR product of RB3.5 with WSM1271 chromosome located primer set; lane 18, PCR product of RB3.6 with chromosomal primer set; lane 19, PCR product of RB3.7 with chromosomal primer set; lanes 2-4, PCR product of WSM1271 with P1; lanes 8-10, PCR product of WSM1271 with P2; lanes 14-16 PCR product of WSM1271 with chromosomal primer set. The correct transconjugants are indicated with *. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

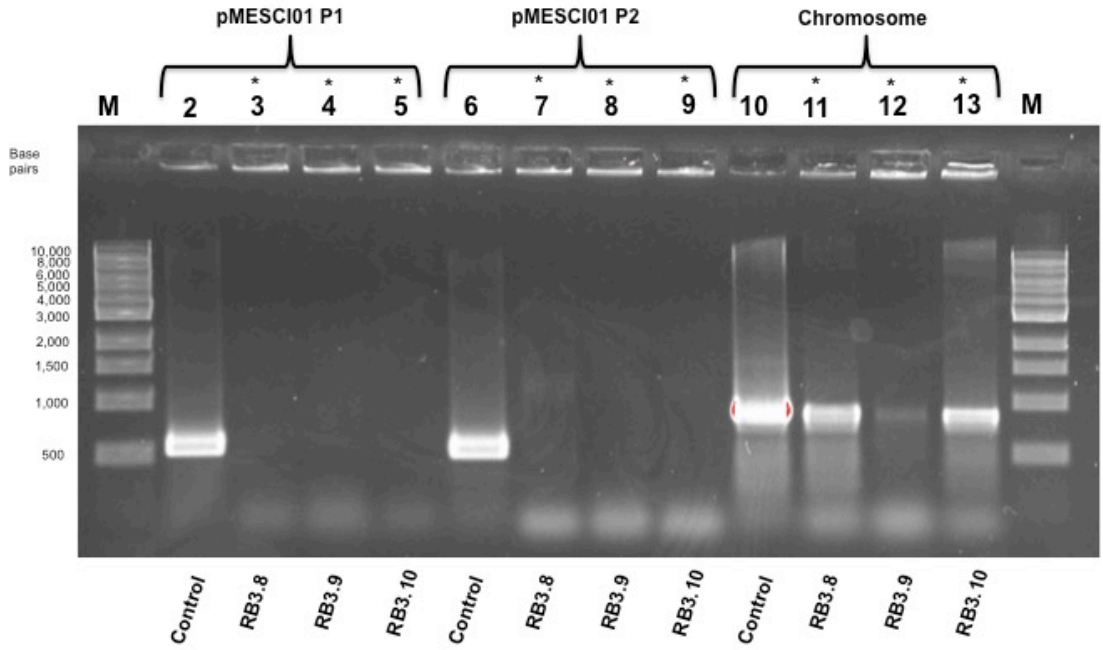


**Figure A8:** Image of electrophoresed PCR products of putative plasmid-cured derivatives of WSM1271 (RB3.8, RB3.9, RB3.10). Lane; 3, PCR product of RB3.8 with pMESCI01 located primer pair 1 (P1); lane; 4, PCR product of RB3.9 with P1; lane; 5, PCR product of RB3.10 with P1; lane 7, PCR product of RB3.8 with pMESCI01 located primer set 2 (P2); lane 8, PCR product of RB3.9 with P2; lane 9, PCR product of RB3.10 with P2; lane 11, PCR product of RB3.8 with WSM1271 chromosome located primer set; lane 12, PCR product of RB3.9 with chromosomal primer set; lane 13, PCR product of RB3.10 with chromosomal primer set; lane 2, PCR product of WSM1271 with P1; lane 6, PCR product of WSM1271 with P2; lane 10 PCR product of WSM1271 with chromosomal primer set. The correct transconjugants are indicated with *. The NEB 1-kb marker sizes are shown to the left of the gel (bp), a 1.2% (w/v) Agarose in 1 X TAE gel was used.

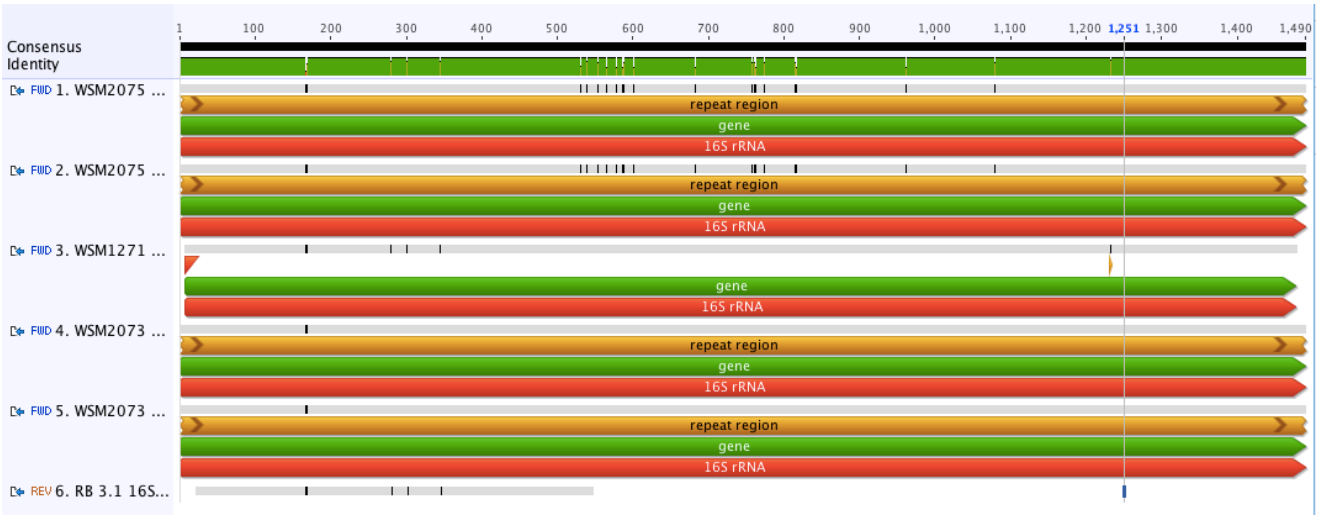


**Figure A9:** Geneious alignment of the 16S rDNA sequences of plasmid-cured derivative of WSM1271, RB3.1 (row 6), compared to the two 16S sequences of WSM2075 (row 1 and 2), WSM1271 (row 3), and the two sequences of WSM2073 (row 4 and 5).