Characterisation of the symbiosis ICE and accessory plasmid of *Mesorhizobium ciceri* CC1192

Emma Rosa Bonello

2018

This thesis is presented for the degree of Bachelor of Science 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.

Emma Rosa Bonello
Abstract

The symbiosis between soil bacteria (rhizobia) and legumes is an integral component of sustainable agricultural practice, providing a source of bioavailable (or fixed) nitrogen to farming systems without supplementation with industrially synthesised fertiliser. Critical to the success of this association are rhizobial nodulation (nod) and nitrogen fixation (nif and fix) genes, that collectively allow for establishment and maintenance of nitrogen-fixing rhizobia-legume symbioses. In the *Mesorhizobium* genus symbiosis genes are chromosomally-encoded on Integrative and Conjugative Elements (ICEs). ICEs are a form of mobile genetic elements, capable of excising from the host chromosome through the action of the ICE-encoded integrase (IntS) and recombination and directionality factor (RdfS). The excised ICE can then transfer via conjugation into the recipient cell, where the ICE integrates into the host chromosome at specific attachment sites, usually adjacent to amino-acyl tRNA genes. Although recipients receive the complete symbiosis ICE from the donor strain and are subsequently able to nodulate a target legume, these exconjugants frequently fix nitrogen sub-optimally, indicating that a strong interaction between chromosomal and ICE-encoded genes exists. Understanding how different *Mesorhizobium* chromosomes and ICEs interact requires an array of ICE-devoid *Mesorhizobium* recipient strains. However currently, there is only such strain available, *M. loti* R7ANS, produced from the ICE-containing parent strain *M. loti* R7A. Furthermore, *Mesorhizobium* spp. may also harbour an accessory plasmid, which may have a role in nitrogen fixation.

Recently, the genome of *Mesorhizobium ciceri* CC1192, the commercial inoculant for *Cicer arietinum* (chickpea) was reported. The genome of CC1192 consists of 6.29-Mb chromosome with a symbiosis ICE (ICE*McSym*1192) of 419-kb, as well as a 648-kb
accessory plasmid, pMc1192. Although the strain has been used as an inoculant for *C. arietinum* for more than 40 years, very little is known about the symbiosis ICE of this strain or the plasmid that it harbours. Therefore, the aims of this thesis were to determine the symbiotic role of the accessory plasmid, pMc1192 in *M. ciceri* CC1192 and to investigate means of producing an ICE-devoid *M. ciceri* strain by curing ICEMcSym$_{1192}$.

Bioinformatic analysis of pMc1192 revealed it harboured several unique symbiosis genes (*fixLJ, fixK* and *fixS*) and several symbiosis genes (*fixNOQP* and *fixGHI*) found on the chromosome. All of these genes, apart from *fixS*, has been shown to be essential for symbiosis. Plasmid pMc1192 was removed through a plasmid incompatibility approach, and loss was confirmed through PCR screening, Eckhardt gel analysis and sequencing. The symbiotic phenotypes of two plasmid-cured strains, MCC69 and MCC70, were assessed alongside the wild type CC1192 with *C. arietinum*. Mean shoot dry weight, nodule number and nodule weight were not significantly different between the three strains (p > 0.05), indicating that the plasmid-encoded genes were dispensable for symbiosis with *C. arietinum*.

In an attempt to remove ICEMcSym$_{1192}$ from CC1192, RdfS which is the protein that stimulates ICE excision in R7A, and for which a homolog exists in CC1192, were individually overexpressed in CC1192. PCR screening of attachment sites formed upon ICEMcSym$_{1192}$ integration and excision in CC1192 overexpressing strains, as well as replica patching of 1,044 colonies demonstrated that ICEMcSym$_{1192}$ was not lost upon overexpression of either *rdfS*, despite the increased proportion of cells excising the ICE when the CC1192 *rdfS* was overexpressed. An alternate approach to remove ICEMcSym$_{1192}$ is through the replacement of the entire 419-kb region with an antibiotic marker through homologous recombination. To achieve this, the integrase
(intS), responsible for catalysing the excision and integration of the ICE, was first inactivated by a replacement with $\Omega$-aadA to prevent re-integration of ICE$_{McSym^{1192}}$ from the chromosome during later deletion of the ICE. PCR screening of the attachment sites in the IntS-deleted strain, MCC86, indicated that the ICE had become immobile, and therefore stabilised in the CC1192 chromosome. The ICE inactivation vector (pMCC8) was also successfully synthesised, providing a tool for future deletion of ICE$_{McSym^{1192}}$ from CC1192.

The removal of pMc1192 as well as the attempted deletion of ICE$_{McSym^{1192}}$ are steps towards the creation of a minimal M. ciceri CC1192 genome devoid of both ICE and accessory plasmids. This strain will prove a valuable research tool to better understand how the interaction between chromosomal and ICE-encoded genes in strains that have recently acquired symbiosis ICEs lead to sub-optimal nitrogen fixation.
Contents

Abstract ................................................................................................................. iii
Contents ................................................................................................................. vi
Acknowledgements ............................................................................................... viii
Abbreviations .......................................................................................................... x
Chapter 1: Introduction ......................................................................................... 1
  1.1 Symbiosis and nitrogen fixation ................................................................. 1
  1.2 Establishment of nitrogen fixing symbioses ................................................ 3
  1.3 Rhizobial genome architecture ................................................................... 6
  1.4 Symbiosis ICEs ............................................................................................ 8
  1.5 Impact of symbiosis ICE transfer on nitrogen fixation ............................... 12
  1.6 Cicer arietinum and Mesorhizobium ciceri CC1192 ................................... 14
  1.7 Aims ............................................................................................................. 15
Chapter 2: Materials and Methods ...................................................................... 17
  2.1 Bacterial strains and plasmids ..................................................................... 17
  2.2 Bioinformatic analysis of CC1192 ............................................................... 19
  2.3 General molecular methods ......................................................................... 19
  2.4 Transformation of plasmids into E. coli ....................................................... 23
  2.5 Biparental conjugation between vectors and CC1192 ............................... 23
  2.6 Curing pMc1192 from the CC1192 genome ................................................ 24
  2.7 Eckhardt Gel Electrophoresis ...................................................................... 25
    2.7.1 Sequencing of MCC69 and MCC70 ....................................................... 26
  2.8 Construction of rdfS overexpression vector, pMCC7 .................................. 26
    2.8.1 Preparation of rdfS from CC1192 ......................................................... 26
    2.8.2 Preparation of pPR3 and ligation with the rdfS fragment ....................... 27
    2.8.3 Screening of pMCC7 ............................................................................. 27
  2.9 Construction of integrase (intS) deletion vector using pEX18Tc, pMCC8 ........ 28
    2.9.1 Preparation of upstream and downstream of intS ................................. 28
    2.9.2 Preparation of pJET-Ω-aadA and pEX18Tc for cloning ......................... 29
    2.9.3 Ligation reactions producing pMCC8 ..................................................... 29
    2.9.4 Screening of pMCC8 ............................................................................ 30
  2.10 Construction of integrase deletion vector using pJQ200SK, pMCC9 ............ 30
  2.11 Site-directed mutagenesis of intS ............................................................... 31
  2.12 Construction of ICE deletion construct, pMCC10 ..................................... 32
  2.13 Symbiotic performance of plasmid-cured derivatives with Cicer arietinum ... 33
  2.14 Plant harvest and nodule sampling ............................................................ 34
2.15 Nodule sections of plasmid-cured derivatives with *Cicer arietinum* ...............35

Chapter 3: Results ...........................................................................................................36

3.1 Determining the symbiotic role of pMc1192 .................................................................36

3.1.1 Bioinformatic analysis of pMc1192 ........................................................................36

3.1.2 Curing of pMc1192 from CC1192 ...........................................................................39

3.1.3 Assessment of the symbiotic phenotype of plasmid-cured derivatives .................43

3.1.4 Nodulation ...............................................................................................................45

3.1.5 Nodule sections .......................................................................................................47

3.2 Investigating curing ICEMcSym\(^{1192}\) from CC1192 ..................................................49

3.2.1 Overexpression of recombination directionality factor, *rdfS* ..................................49

3.2.2 Curing ICEMcSym\(^{1192}\) by site-directed deletion ..................................................55

3.2.3 Construction of ICE deletion vector ........................................................................59

Chapter 4: Discussion .....................................................................................................60

4.1 The symbiotic role of pMc1192 .....................................................................................60

4.2 Curing of ICEMcSym\(^{1192}\) ..........................................................................................63

4.2.1 Overexpression of *rdfS* does not lead to ICEMcSym\(^{1192}\) loss ............................63

4.2.2 Stabilisation of ICEMcSym\(^{1192}\) through the deletion of *intS* ..............................65

4.2.3 Subsequent deletion of ICEMcSym\(^{1192}\) ................................................................66

Conclusions and Future Direction .....................................................................................68

References ........................................................................................................................70

Appendix Figures ...............................................................................................................85
Acknowledgements

There have been a few people and organisations that have been pivotal in my research project that I feel deserve appropriate recognition.

Firstly, to my primary supervisor, Dr Jason Terpolilli, for putting in an enormous amount of time and energy into my project. You were never too busy to answer any of my questions or to explain a concept until it clicked. I cannot thank you enough for your patience, good-humour and mentorship. I’ve definitely learnt a substantial amount from you, in part due to the “inspirational” quotes and character building.

To my secondary supervisors, Dr Graham O’Hara and Dr Joshua Ramsay, I also wish to extend my gratitude. Graham, your sharp eye for editing has helped me realise the subtleties involved in writing. Plus, your continuous concern for my well-being and how the Rabbitohs’ played really meant a lot. And to Josh, although our interaction was minimal, your expertise and feedback helped me in the times where it mattered the most.

I would like to also thank ALOSCA for awarding me with an Honours Scholarship, which has enabled me to focus more on my project.

A big part of my enjoyment this past year has been due to the Centre of Rhizobium Studies (CRS). The group is made up of incredible, welcoming people; all genuinely friendly and each an expert in their field. From the weekly meetings I have developed a breadth of knowledge that I never would have obtained otherwise. It has been an honour and a privilege to be part of the group, and I look forward to working more with you all in the future.
From the CRS I would especially like to thank a few individuals. Firstly, to both Reg and Yvette for their enormous assistance in the glasshouse – I would not have been able to do it without you, both times. Yvette, thank you also for the extra help in the glasshouse, answering a tonne of questions, the jam, tv show recommendations and for helping with some editing. Talitha, you’ve become a great buddy this past year. I have loved our rides to uni, coffee runs, dog stories and all the conversations in between. I definitely think I would have been way more of a mess had it not been for the wisdom and experiences you’ve shared with me.

I am also incredibly grateful to those who I have been able to laugh with – Jack, Georgina, Lalo, Cady, the friends I made in the honours unit and those outside of uni. Those good times were extremely valuable as things got more intense and helped me keep my cool.

A special shout out to my partner, AJ. You’ve been extremely supportive and helpful for the entire year, despite taking the brunt of my stress. But no matter what emotions I threw at you, you were always there to make me laugh…and get food.

Finally, to my mum, Sandra. Your hard work and persistence to give Ryan and I every opportunity you could is what has allowed me to get to where I am today, which is exactly where I want to be.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGRF</td>
<td>Australian Genome Research Facility</td>
</tr>
<tr>
<td>ALA</td>
<td>5-aminolaevulinic acid</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BLASTN</td>
<td>Basic Local Alignment Search Tool – Nucleotide</td>
</tr>
<tr>
<td>BLASTP</td>
<td>Basic Local Alignment Search Tool – Protein</td>
</tr>
<tr>
<td>COG</td>
<td>Cluster of Orthologous Groups</td>
</tr>
<tr>
<td>G/RDM</td>
<td>Glucose defined media</td>
</tr>
<tr>
<td>Gm</td>
<td>Gentamycin</td>
</tr>
<tr>
<td>ICE</td>
<td>Integrative and conjugative element</td>
</tr>
<tr>
<td>ICEM/ISym$^\text{R7A}$</td>
<td>ICE of <em>Mesorhizobium loti</em> R7A</td>
</tr>
<tr>
<td>ICEMcSym$^{1192}$</td>
<td>ICE of <em>Mesorhizobium ciceri</em> CC1192</td>
</tr>
<tr>
<td>ICEMcSym$^{1271}$</td>
<td>ICE of <em>Mesorhizobium ciceri</em> bv. biserrulae WSM1271</td>
</tr>
<tr>
<td>IMG</td>
<td>Integrated microbial genomes</td>
</tr>
<tr>
<td>Km</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny Broth</td>
</tr>
<tr>
<td>NCBI</td>
<td>National centre for biotechnology information</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolab</td>
</tr>
<tr>
<td>pSym</td>
<td>Symbiotic plasmid</td>
</tr>
<tr>
<td>RDFs/rdf$^S$</td>
<td>Recombination directionality factor</td>
</tr>
<tr>
<td>RDM</td>
<td>Defined media</td>
</tr>
<tr>
<td>S/RDM</td>
<td>Sucrose defined media</td>
</tr>
<tr>
<td>Sm</td>
<td>Streptomycin</td>
</tr>
<tr>
<td>Sp</td>
<td>Spectinomycin</td>
</tr>
<tr>
<td>Tc</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>TY</td>
<td>Tryptone yeast</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Symbiosis and nitrogen fixation

A source of nitrogen (N) is required for the production of many of the cellular components required for life (Kaur et al., 2017). Although nitrogen is abundant in the atmosphere as N$_2$ gas, it is chemically stable and largely unavailable to most organisms in this form (Hoffman et al., 2014). In agriculture, the lack of available nitrogen in soils often limits crop productivity, which may be addressed through the application of industrially synthesised nitrogen-containing fertilisers. However, only 30 – 50% of nitrogen applied in this way is typically assimilated by crops, with a significant proportion of the remainder leached into surface and ground waters causing eutrophication, acidification and apoxia (Cameron et al., 2013; Guignard et al., 2017; Kaur et al., 2017). The remaining fertiliser may be readily volatilised into nitrous oxides, which are well-known potent greenhouse gases (Gregorich et al., 2015). Finally, the production of these synthetic nitrogen-containing fertilizers also involves the burning of fossil fuels, contributing further to greenhouse gas emissions (Burris & Roberts, 1993; Kaur et al., 2017).

An alternative to synthetic nitrogen fertiliser application is to include legumes in agricultural rotations either as a crop, such as *Cicer arietinum* (chickpea) and *Pisum sativum* (peas), or as a pasture, such as *Biserrula pelecinus* (biserrula) and *Trifolium* spp. (clovers). Legumes are able to provide nitrogen to farming systems through their ability to form a symbiosis with rhizobia, soil bacteria capable of converting the inaccessible atmospheric N$_2$ into a biologically available form through the process of symbiotic nitrogen fixation. Nitrogen-fixing symbioses are established when soil rhizobia infect legume roots and differentiate into their nitrogen-fixing form within root nodules. Inside these nodules, the rhizobia are encased within a plant-derived
membrane, termed a symbiosome, and function as an ammonia (NH₃)-secreting organelle (Udvardi & Poole, 2013). NH₃ produced from the enzymatic reduction of N₂ in the symbiosome is delivered to the plant cell and in return the rhizobia receive a supply of carbon, essential cations and some amino acids (Udvardi & Poole, 2013). It is estimated that this relationship provides approximately 21 million tonnes (Tg) per annum of bioavailable nitrogen to agricultural systems (Herridge et al., 2008), representing a major input of low cost nitrogen.

Rhizobia are a phylogenetically diverse group of bacteria, consisting of 15 recognised genera (Andrews & Andrews, 2017), classified within two subclasses; α- and β-proteobacteria (Sprent et al., 2017). The α-proteobacteria is the largest and most thoroughly studied subclass, with rhizobia found in six families, while β-proteobacteria contains only one family with two genera of symbiotic nitrogen-fixing bacteria; *Burkholderia* and *Cupriavidus* (Shamseldin et al., 2017; Sprent et al., 2017). A third subclass, γ-proteobacteria, is occasionally included, however the evidence for any of its members being true symbiotic nitrogen-fixing organisms is not clear (Benhizia et al., 2004; Ibáñez et al., 2009; Shiraishi et al., 2010; De Meyer et al., 2015; Sprent et al., 2017). This review will therefore focus on the rhizobia within α-proteobacteria, which form symbioses with legumes of agriculturally important crops and pastures such as *Glycine max* (soybean), *Medicago sativa* (alfalfa), *Pisum sativum* (field pea), *Vicia faba* (broad beans), *Cicer arietinum* (chickpea) and *Biserrula pelecinus* (biserrula).
1.2 Establishment of nitrogen fixing symbioses

The establishment of a nitrogen-fixing symbiosis requires infection of legume roots by the rhizobial microsymbiont. There are three recognised mechanisms of rhizobial infection; root hair curling, epidermal invasion and crack entry (Sprent, 2009). Legumes of agricultural importance are infected by root hair curling, which will therefore be the focus in this review (Downie, 2010; Ibáñez et al., 2017).

Legume root hair infection is initiated by a chemical dialogue between host roots and rhizobia in the rhizosphere of the plant. Legume roots exude a range of phenolics including flavonoids and isoflavonoids that act as chemoattractants to rhizobia (Janczarek et al., 2015). These phenolics diffuse through the bacterial cell membrane before binding and activating the transcriptional regulator NodD, which in turn activates transcription of nodulation genes (or nod genes, which encompass nod, noe and nol) that direct the synthesis and secretion of lipochito-oligosaccharides termed Nod factors (Mulligan & Long, 1985; Feng et al., 2003; Li et al., 2008). Synthesised Nod factors released by rhizobia bind specific receptor-like kinases expressed on the cell surface of legume root hairs (Oldroyd et al., 2011; 2013; Udvardi & Poole, 2013), activating downstream signalling in the plant including calcium influx/efflux, induction of plant genes and the release of hormones that trigger nodule organogenesis (Oldroyd, 2013; Janczarek et al., 2015). Root hair curling is also initiated by Nod factors binding to the plant’s root hairs, causing them to form a ‘shepherd’s crook’ structure and entrapping the rhizobia attached to the tip of the root hair (Sahlman & Fåhraeus, 1963; Ridge & Rolfe, 1985; Esseling et al., 2003) (Figure 1.1).
Figure 1.1 Summary of infection of legume root hairs by rhizobia. (a) Flavonoids released by the plant root signal to rhizobia in the rhizosphere, which in turn produce nodulation factors (Nod factors) recognized by the plant. Nod factor perception activates the symbiosis signalling pathway, leading to calcium spiking, initially in epidermal cells but later also in cortical cells preceding their colonization. (b) Root hair cells curl and engulf rhizobia attached to the root hair tip. An infection thread forms by invagination of the plant cell membrane and rhizobia entrapped inside the curled root hair then proceed down the growing infection thread by cell division. (c) The infection thread grows toward the developing root nodule. (d) The infection thread containing rhizobia ramifies within the nodule tissue, before the bacteria are released into membrane-bound compartments inside the cells of the nodule, where they differentiate into their nitrogen-fixing bacteroid state. Image from Oldroyd (2013).

Sustained production of Nod factor by the entrapped rhizobia further stimulates nodule development. The root hair invaginates and creates an infection thread, a channel delimited by plant-derived cell membranes extending from the cortical cell layer down to the nodule primordium (Oldroyd et al., 2011). Entrapped rhizobia grow and divide down this infection thread, being released as infection droplets encased in plant-derived membrane once they reach the developing nodule (Bassett et al., 1977). The plant membrane that surrounds the released bacteria persists and at this point the entire structure is referred to as the symbiosome (Udvardi & Poole, 2013).

Within the symbiosome, rhizobia differentiate into their nitrogen-fixing or bacteroid form, leading to expression of the nif clusters encoding proteins directing the synthesis and assembly of the nitrogenase enzyme complex (Fischer, 1994). This complex contains two subunits: a small protein containing iron (Fe) encoded by nifH and a large molybdenum-Iron (MoFe) protein consisting of two α and β subunits encoded by nifD and nifK, respectively (Fischer, 1994; Dixon & Kahn, 2004). In addition, at least 12
other Nif products are required for the generation of the active nitrogenase (Rubio & Ludden, 2008). Under ideal conditions, the nitrogenase reduces atmospheric $N_2$ into $NH_3$ according to the following reaction (Dixon & Kahn, 2004):

$$N_2 + 8e^- + 8H^+ + 16 \text{ ATP} \rightarrow 2 \text{ NH}_3 + \text{ H}_2 + 16 \text{ ADP} + 16\text{ Pi}$$

Both the Fe and MoFe components of nitrogenase are $O_2$ labile (Dixon & Kahn, 2004; Udvardi & Poole, 2013), thus $O_2$ tension within root nodules must be kept low for continual nitrogenase activity (Preisig et al., 1993; Mus et al., 2016). This is achieved and maintained through the high rate of $O_2$ consumption by nodule mitochondria, the presence of mucus and a protective outer cell layer to decrease the potential for $O_2$ diffusion and synthesis of leghemoglobin by the plant that binds $O_2$ with high affinity, thus keeping $O_2$ tension low (Udvardi & Poole, 2013; Mus et al., 2016).

Expression of nif genes is under the control of the master transcriptional regulator, NifA (Tsoy et al., 2016). In Sinorhizobium meliloti, which forms a nitrogen-fixing association with Medicago spp., expression of nifA is itself regulated by a two component $O_2$-responsive regulatory system comprising of a sensor (FixL) and activator (FixJ) protein. Under low $O_2$ tension, FixL triggers a phosphorylation cascade, leading to activation of FixJ which in-turn upregulates the transcriptional activator, FixK (Batut et al., 1989; Dixon & Kahn, 2004; Tsoy et al., 2016). FixK then directs transcription of fixNOQP and fixGHIS operons (Dixon & Kahn, 2004). FixNOQP encodes a $cbb_3$-type cytochrome oxidase, which functions as a terminal electron acceptor in the low $O_2$ environment of the bacteroid (Preisig et al., 1996b), while fixGHIS encodes a copper-uptake system important for maturation of FixNOQP (Preisig et al., 1996a; Koch et al., 2000). The combined action of FixNOQP and FixGHIS provide bacteroids with ATP necessary to fuel the nitrogenase, while a
source of electrons is likely to be facilitated by the electron transferring flavoprotein complex FixABCX (Sperotto et al., 2004; Terpolilli et al., 2016). These fix genes have all been shown to be essential for the continual function of the nitrogenase and therefore symbiotic nitrogen fixation (Earl et al., 1987; Preisig et al., 1993; 1996a).

1.3 Rhizobial genome architecture

Bacterial genomes can be defined as consisting of both core and accessory genes. Core genes are those fundamental for cellular function and are shared across all members of a species (Harrison et al., 2010; Segerman, 2012; Young, 2016). In contrast, accessory genes are those not essential for cellular function but which may enhance survival through the provision of advantageous phenotypes for various niches (Young et al., 2006; 2016). Therefore, rhizobial nod, nif and fix genes, collectively referred to as symbiosis genes, are generally considered part of the accessory genome (Young, 2016; Poole et al., 2018).

Rhizobial genomes are relatively large compared to most bacteria, ranging in size from ~5 – 10 Mb, and tend to consist of multiple replicons in addition to the single chromosome, such as accessory plasmids (Koonin & Wolf, 2008; Masson-Boivin et al., 2009; Poole et al., 2018). Rhizobial accessory plasmids vary greatly in size (~50-kb to 1.6-Mb) and number (Zahran, 2017). Analysis of a wide range of rhizobial genera has shown that the symbiosis genes are often clustered tightly together within the genome (Provorov et al., 2017). In Rhizobium spp. and Sinorhizobium spp. (also known as Ensifer), the symbiosis genes tend to be encoded on accessory plasmids, referred to as symbiotic plasmids (pSyms) (Zahran, 2017) (see Figure 1.2). These genera may also contain a unique replicon known as a chromid; a hybrid between a chromosome and a plasmid (Harrison et al., 2010). Chromids are characterised by the
replication and maintenance system of a plasmid, a nucleotide composition similar to the chromosome as well as the presence of core genes similar to that of closely related species (Harrison et al., 2010). An example of this architecture can be found in *Rhizobium leguminosarum* bv. *viciae* 3841, which contains six non-chromosomal replicons in addition to the chromosome. One of these replicons (pRL10) harbours the symbiosis genes (Young et al., 2006) while the other plasmids are either other accessory elements (pRL8 and pRL7) or chromids (pRL12, pRL11 and pRL9) (Harrison et al., 2010).

**Figure 1.2** Overview of the genomic structure of Rhizobia based on the location of the symbiosis genes. (1) Those usually with the genes on a symbiosis plasmid (pSym), such as *Rhizobium* and *Sinorhizobium*, may also harbour a chromid along with several accessory plasmids. Chromids are a hybrid between a chromosome and a plasmid, containing the replication and maintenance system of a plasmid, the nucleotide composition similar to the chromosome and core genes from closely related species (Harrison et al., 2010). (2) Genera with the symbiosis genes encoded within the chromosome, including *Mesorhizobium*, *Bradyrhizobium* and *Azorhizobium*. In *Mesorhizobium* and *Azorhizobium* these genes are harboured on an Integrative and Conjugative Element (ICE), which are capable of excising from the chromosome and forming a plasmid-like replicon that can transfer via conjugation to a recipient cell. *Mesorhizobium* symbiosis ICES may be monopartite or tripartite in structure. Figure from Poole et al. (2018).
In contrast to *Rhizobium* spp. and *Sinorhizobium* spp., symbiotic genes in *Mesorhizobium* spp., *Bradyrhizobium* spp. and *Azorhizobium* spp. are most often clustered together on the chromosome. In *Mesorhizobium* spp. and *Azorhizobium caulinodans*, the chromosomally-encoded symbiosis genes are mobile and can transfer through specific mechanisms involving conjugation (Ramsay *et al.*, 2006; Haskett *et al.*, 2016a; Ling *et al.*, 2016). These mobile elements are therefore referred to as symbiosis integrative and conjugative elements (ICEs) and allow for the transfer and integration of large (~600-kb) regions of DNA horizontally into new genomes (Sullivan *et al.*, 2002; Laranjo *et al.*, 2014; Burrus, 2017; Delavat *et al.*, 2017).

### 1.4 Symbiosis ICEs

The most well characterised symbiosis ICE is that of *Mesorhizobium loti* R7A, which forms a nitrogen-fixing symbiosis with the pasture legume *Lotus corniculatus* (Sullivan *et al.*, 2002; Delavat *et al.*, 2017). The symbiosis ICE of *Mesorhizobium loti* R7A (denoted as ICEMISymR7A) is a 502-kb element that integrates into the sole phe-tRNA gene of the R7A chromosome in a non-disruptive manner (Sullivan & Ronson, 1998). The first ORF of this ICE encodes an integrase, IntS, that catalyses the site-specific recombination reaction between the homologous sequence on the chromosome (referred to as the *attB* site) and on the excised circular ICE (the *attP* site) (Ramsay *et al.*, 2006) (Figure 1.3). IntS is a tyrosine recombinase that mediates the breakage and re-joining of the DNA in an energy- and nucleotide-conserving manner (Grindley *et al.*, 2006). Integration of the element into the chromosome then forms the *attL* and *attR* sites bordering the ICE (Ramsay *et al.*, 2006), allowing ICEMISymR7A to be stably maintained in the genome of the cell during cell division.
Figure 1.3 The excision and integration of ICE\textit{MI}Sym$^{R7A}$ as described by Ramsay et al. (2006). The 502-kb ICE\textit{MI}Sym$^{R7A}$ integrates adjacent to a \textit{phe}-tRNA by the action of the integrase (Int\textit{S}), forming att\textit{L} and att\textit{R} sites on either side of the integrated ICE. The excision of the element requires the presence of a recombination directionality factor (Rdf\textit{S}), which stimulates the excisionase function of the integrase. Excision of the ICE causes the att\textit{L} and att\textit{R} sites to recombine, resulting in att\textit{B} on the chromosome where the ICE\textit{MI}Sym$^{R7A}$ once was, and att\textit{P} on the excised element. Figure adapted from Ramsay et al. (2006).

Before horizontally transferring from a donor cell to a recipient, a symbiosis ICE must first excise from the chromosome. Excision of ICE\textit{MI}Sym$^{R7A}$ requires Int\textit{S} and the presence of a recombination directionality factor (Rdf\textit{S}) that stimulates the excisionase function of the integrase (Ramsay et al., 2006) (see Figure 1.3). An ICE-encoded relaxase (Rlx\textit{S}) then binds to the excised double stranded circular ICE, creating a single-stranded break at the origin of transfer (ori\textit{T}), before guiding the single strand through a mating pore, formed by an ICE-encoded type-IV secretion
system, and into the recipient cell (Burrus et al., 2002; Wozniak & Waldor, 2010). In the recipient cell, the single stranded ICE acts as the template for replication by DNA polymerase, before integration into the recipient host genome specifically at the phe-tRNA gene (Burrus et al., 2002; Delavat et al., 2017). The circularised single strand that remains in the host cell is replicated via rolling-circle amplification and then re-integrates into the host’s chromosome (Delavat et al., 2017).

Recently, a second type of symbiosis ICE was identified in M. ciceri bv. biserrulae WSM1271 consisting of three separate chromosomal regions (α, β and γ) that recombine into a single circular element prior to conjugal transfer (Haskett et al., 2016a). Like ICEMISymR7A, regions β and γ reside adjacent to highly conserved genes (guaA for β, while γ is flanked by met-tRNA and phe-tRNA) (Haskett et al., 2016a).

Unusually, the α region is integrated into an intergenic region of the chromosome (Haskett et al., 2016a). Excision of this tripartite ICE (ICEMcSym1271) is similar to that of the monopartite ICEMISymR7A, but requires the action of three separate integrases (IntS, IntG and IntM) and three recombination directionality factors (RdfS, RdfG and RdfM) acting on three sets of attachment sites (Haskett et al., 2016a), with RdfS directing excision of the element (Haskett et al., 2016a; 2018). In log-phase cultures, both ICEMISymR7A and ICEMcSym1271 exist predominantly in their chromosomally-integrated form, with excised ICEMISymR7A detected in less than 0.04% of cells in a population (Ramsay et al., 2006), while the proportion of WSM1271 with an excised ICEMcSym1271 is much lower at 0.001% of cells (Haskett et al., 2016a). Both of these proportions increase 10- to 100-fold in stationary phase cells, due to additional activation of ICE-excision by quorum sensing (Ramsay et al., 2009; 2013; 2015; Haskett et al., 2018).
Although ICEMISym\textsuperscript{R7A} is preferentially integrated into the chromosome, it is possible to remove or cure the element from the R7A genome (Ramsay et al., 2006). Mobilisation of a vector overexpressing the recombination directionality factor, rdfS, into R7A resulted in only a small number of slow-growing transconjugants observed on selective media. PCR screening detected intact \textit{attB} sites (indicative of a reformed chromosomal attachment site) in all colonies, but no \textit{attP} (circularised form of ICE), \textit{attL} or \textit{attR} sites (demarcating the ICE within the chromosome) (see Figure 1.3), suggesting that the element had been lost from the genome of R7A (Ramsay et al., 2006). All colonies were devoid of screened ICE-encoded genes (\textit{nodB}, rdfS and genes encoding biotin, thiamine and nicotinate synthesis), auxotrophic for the aforementioned vitamins and were unable to nodulate \textit{L. corniculatus}, confirming that ICEMISym\textsuperscript{R7A} had been cured from R7A and giving rise to the naïve symbiosis ICE-devoid strain, R7ANS (Ramsay et al., 2006). R7ANS has subsequently proved a very valuable research tool, acting as a recipient strain for \textit{in vitro} symbiosis ICE transfer experiments and allowing for the demonstration of ICE mobility and the measurement of ICE transfer rates (Ramsay et al., 2006; Haskett et al., 2016a). Although the overexpression of rdfS causes the loss of the symbiosis ICE from the R7A host genome, an analogous experiment conducted with WSM1271 failed to cure ICEMcSym\textsuperscript{1271} from this strain when either rdfS or all three rdf genes ( rdfS, rdfG and rdfM) were overexpressed (Haskett et al., 2018). However, the inability to cure WSM1271 of ICEMcSym\textsuperscript{1271} in this way may be due to the complex nature of the excision mechanism of this tripartite element.
1.5 Impact of symbiosis ICE transfer on nitrogen fixation

The ability of symbiosis ICEs to conjugally transfer between rhizobia has the potential to have a significant impact in the field where *Mesorhizobium* strains are used as commercial inoculants for legumes of agricultural importance, such as the pasture legume *Biserrula pelecinus* and the grain legume *Cicer arietinum*. The *B. pelecinus* inoculant strain, *M. ciceri* bv. *biserrulae* WSM1271, was introduced to a site originally devoid of compatible rhizobia. Six years after introduction, 8% (7/88) of nodules on *B. pelecinus* contained strains of *Mesorhizobium* identified as genetically distinct to WSM1271 (Nandasena et al., 2007a). Crucially, when the efficacy of these novel strains was tested in a glasshouse trial, all were shown to be less effective than WSM1271 at fixing nitrogen on *B. pelecinus* (Nandasena et al., 2007a). A similar phenomenon was observed when *M. ciceri* bv. *biserrulae* strain WSM1497, the current Australian commercial inoculant for *B. pelecinus* (Bullard et al., 2005), was introduced at other field sites in the Western Australian Wheatbelt, where almost 50% (193/387) of nodule occupants isolated several years after introduction were genetically distinct to WSM1497 (Nandasena et al., 2007b). As with WSM1271, all of the novel isolates tested in the WSM1497 study were also shown to be less effective than WSM1497 at fixing nitrogen on *B. pelecinus* (Nandasena et al., 2007a). These data were consistent with the inoculant strains, WSM1271 and WSM1497, transferring their symbiosis genes to other *Mesorhizobium* strains present in the soil, resulting in the evolution of novel *B. pelecinus*-nodulating organisms.

Subsequent phylogenetic analysis of two novel isolates from the WSM1271 study revealed these strains to be two new species, *Mesorhizobium australicum* WSM2073 (which fixes nitrogen at approximately 50% of the WSM1271 rate on *B. pelecinus*) and *Mesorhizobium opportunistum* WSM2075 (which nodulates *B. pelecinus* but does
not fix nitrogen) (Nandasena et al., 2009). Genome sequencing confirmed both strains harboured the complete tripartite ICEMcSym\textsuperscript{1271} (Reeve et al., 2013a; 2013b), consistent with WSM2073 and WSM2075 acquiring the ICE from the inoculant strain. Therefore, introduction of Mesorhizobium inoculant strains for B. pelecinus resulted in the transfer of their symbiosis ICEs to novel recipient strains that were less efficient at fixing nitrogen with B. pelecinus. It is possible that these novel B. pelecinus-nodulating organisms were non-symbiotic soil Mesorhizobium spp. capable of acquiring the symbiosis ICEs from the inoculant strains.

Complete genome sequencing also revealed that WSM1271 harboured a small 426-kb plasmid, pMESC\textsuperscript{101}, which was absent from the genomes of WSM2073 and WSM2075 (Nandasena et al., 2009; 2014). It is possible this plasmid could play a role in symbiosis with B. pelecinus and its absence from the genomes of the WSM2073 and WSM2075 recipients might explain in part their reduced symbiotic performance on B. pelecinus. However, In vitro mating experiments showed that ICEMcSym\textsuperscript{1271} could be transferred from WSM1271, WSM2073 and WSM2075 separately into the ICE-devoid R7ANS recipient strain (Haskett et al., 2016a). Importantly, transfer of each tripartite ICE to R7ANS resulted in exconjugants with a partially effective symbiotic phenotype, irrespective of whether the ICE donor strain was effective (WSM1271), partially effective (WSM2073) or ineffective (WSM2075) at fixing nitrogen on B. pelecinus (Haskett et al., 2016a). This highlights that the Mesorhizobium chromosomal background of the symbiosis ICE recipient is a critical factor in determining the nitrogen-fixation effectiveness of the novel microsymbionts.
1.6 *Cicer arietinum* and *Mesorhizobium ciceri* CC1192

Work with *Cicer arietinum* (chickpea) has indicated that the same phenomenon of symbiosis ICE transfer may have occurred following inoculation with the commercial inoculant strain of this crop legume, *M. ciceri* CC1192 (Elias & Herridge, 2015). In their survey of *C. arietinum* growing on farms in Moree (New South Wales), Elias and Herridge (2015) reported that 47% of the 628 rhizobial isolates obtained from *C. arietinum* were genetically distinct to CC1192 (Elias & Herridge, 2015). However, unlike the situation for *B. pelecinus* where all novel strains were less effective at fixing nitrogen than the inoculant strains, ten of the 29 novel isolates assessed in a glasshouse experiment performed equivalently to CC1192 on *C. arietinum* (Elias & Herridge, 2015). Although it is currently unclear whether these novel isolates had acquired symbiosis genes from CC1192, it is possible that in some instances, effective novel isolates may be capable of evolving in the field.

Recently, the complete genome sequence of *M. ciceri* CC1192 was reported, showing this strain harbours a monopartite 419-kb symbiosis ICE (ICEMcSym\textsuperscript{1192}) integrated in the 6.29-Mb chromosome (Haskett \textit{et al.}, 2016c). Like ICE\textit{M/I}Sym\textsuperscript{R7A}, ICEMcSym\textsuperscript{1192} encodes genes known to be important for ICE integration (\textit{intS}), excision (\textit{rdfS}), replication and transfer (\textit{rlxS}, type IV secretion system). However, in contrast to ICEMcSym\textsuperscript{R7A} which is integrated at the sole \textit{phe}-tRNA in the R7A chromosome, ICEMcSym\textsuperscript{1192} is integrated at one of five \textit{ser}-tRNA genes in CC1192 (Haskett \textit{et al.}, 2016c). ICEMcSym\textsuperscript{1192} is mobile, with \textit{in vitro} mating experiments showing ICEMcSym\textsuperscript{1192} can transfer to R7ANS (T. Haskett, unpublished results). At present, the symbiotic effectiveness of these exconjugants is not known, however, the demonstration of ICE transfer in the laboratory indicates that the novel isolates
reported by Elias and Herridge (2015) could have arisen from resident soil *Mesorhizobium* acquiring ICEMcSym\(^{1192}\) from the CC1192 inoculant.

In addition to ICEMcSym\(^{1192}\), the genome of CC1192 also contains a single 648-kb plasmid, called pMc1192 (Haskett *et al.*, 2016c). This plasmid is similar in size to pMESCI01 (426 kb) in the genome of *B. pelecinus*-nodulating WSM1271. In fact, recent genome sequencing of two other *M. ciceri* strains, WSM1497 and WSM1284, has revealed these also each carry a single plasmid of comparable size to CC1192 (532-kb and 554-kb, respectively) (Haskett *et al.*, 2016b; Brewer *et al.*, 2017). The exact role/s of these *M. ciceri* plasmids is currently not known, although some preliminary evidence for plasmid pMESCI01 suggests it may not be essential for symbiotic nitrogen fixation of WSM1271 with *B. pelecinus* (R. Brewer, unpublished results).

1.7 Aims

Symbiosis genes in *Mesorhizobium* spp. are chromosomally-encoded on symbiosis ICEs capable of conjugative transfer to recipient cells. When *Mesorhizobium* spp. are supplied as inoculants with legumes in the field, this transfer results in the evolution of novel microsymbionts, many of which are sub-optimally effective at fixing nitrogen on their respective host. The development of populations of sub-optimally effective microsymbionts may create a barrier to future inoculation success, competing with the inoculant for nodulation of the legume (Thies *et al.*, 1991; Rao *et al.*, 1994). This could therefore pose a threat to the long-term efficacy of *Mesorhizobium* inoculants. Managing the impact of symbiosis ICE transfer therefore requires a detailed understanding of the molecular determinants of highly effective nitrogen-fixing
symbioses. The accessory plasmids that appear to be present in *M. ciceri* genomes may play an important, yet hitherto unrecognised role in symbiotic function.

*M. loti* R7A was the first strain reported to harbour a symbiosis ICE, ICEM/Sym\textsuperscript{R7A}, which was subsequently cured from this strain by over-expressing the recombination directionality factor *rdfS*, producing the symbiosis ICE-devoid *M. loti* strain, R7ANS (Ramsay *et al.*, 2006). This naïve strain has proved an invaluable tool for *in vitro* symbiosis ICE mobility experiments (Haskett *et al.*, 2016a). The monopartite symbiosis ICE of the commercial *C. arietinum* inoculant strain *M. ciceri* CC1192, ICE\textsuperscript{McSym}\textsuperscript{1192}, harbours ICE-integration and excision genes homologous to those in R7A, including *intS* and *rdfS*. Thus, it is possible that CC1192 could similarly be cured of its symbiosis ICE, producing an *M. ciceri* ICE-devoid strain. The availability of two different species of ICE-devoid strains (*M. loti* and *M. ciceri*) would prove a powerful tool to help unravel how symbiotic performance of exconjugants is affected by the interaction between chromosomal and ICE-encoded genes.

Therefore, the three aims of this study are;

1. To remove the CC1192 plasmid, pMc1192, from the *M. ciceri* CC1192 genome.
2. To assess the symbiotic phenotype of the plasmid-cured *M. ciceri* CC1192 strain.
3. To investigate approaches to cure ICE\textsuperscript{McSym}\textsuperscript{1192} from the *M. ciceri* CC1192 genome.
Chapter 2: Materials and Methods

2.1 Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 2.1. *Escherichia coli* cultures were grown at 37°C in lysogeny broth (LB) (Bertani, 1951), supplemented with 1.5% (w/v) agar as the solidifying agent when required. *E. coli* ST18 cultures were supplemented with 60 µg/ml 5-aminolevulinic acid (ALA). *Mesorhizobium ciceri* CC1192 was grown at 28°C in Tryptone yeast extract (TY) broth (Beringer, 1974), on 1.5% (w/v) TY agar occasionally supplemented with 146 mM sucrose, or on defined media (RDM) (Ronson & Primrose, 1979) with 1.75% (w/v) agar containing either 15 mM glucose (G/RDM) or 146 mM sucrose (S/RDM) (Hynes *et al.*, 1989). *Rhizobium leguminosarum* bv. *viciae* 3841 was either grown in TY broth or on TY agar. Broth cultures were incubated on gyratory shakers at 250 rpm. RDM was supplemented with 1 µg/ml thiamine, 1 µg/ml nicotinate and 20 ng/ml biotin unless screening for vitamin auxotrophy (Sullivan *et al.*, 2002). Where appropriate for selection, media was supplemented with the following antibiotics for *M. ciceri* (µg/ml); neomycin (250), spectinomycin (100), streptomycin (50), gentamycin (40) and tetracycline (1). Media for *E. coli* was supplemented with the following antibiotics (µg/ml) where appropriate; kanamycin (50), tetracycline (10), gentamycin (10) and spectinomycin (50). Media for *R. leguminosarum* was always supplemented with 500 µg/ml streptomycin. All antibiotics were purchased from Sigma-Aldrich (St. Louis, USA).
**Table 2.1** Bacterial strains and plasmids used in the removal of the pMc1192 and the attempted removal of ICEMcSym\(^{192}\).

<table>
<thead>
<tr>
<th>Strain/Plasmid</th>
<th>Description</th>
<th>Reference/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST18</td>
<td>S17-1ΔhemA thi pro hsdRM with chromosomal integrated [RP4-2 Tc::Mu:Km'::Tn7, Tra' Tri' Str']</td>
<td>Thoma &amp; Schobert, 2009</td>
</tr>
<tr>
<td>DH5α</td>
<td>F - deoR endA1 recA1 relA1 gyrA96 hsdR17 (r(^{k-}), m(^{c+})) supE44 thi-1 phoA Δ(lacZYA argF)U169 Φ80lacZΔM15λ –</td>
<td>Bioline</td>
</tr>
<tr>
<td><strong>Mesorhizobium ciceri</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC1192</td>
<td>Wild type strain</td>
<td>Corbin <em>et al.</em>, 1977</td>
</tr>
<tr>
<td>MCC69</td>
<td>CC1192 derivative; ΔpMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>MCC70</td>
<td>CC1192 derivative; ΔpMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>MCC86</td>
<td>CC1192 derivative; intS-Ω-aadA; Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Rhizobium leguminosarum</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3841</td>
<td>Wild type strain; contains 6 plasmids: pRL12, pRL11, pRL10, pRL9, pRL8, pRL7; Sm(^R)</td>
<td>Young <em>et al.</em>, 2006</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pEX18Tc</td>
<td>ColE1 oriV; sacB; Tc(^R)</td>
<td>Hoang <em>et al.</em>, 1998</td>
</tr>
<tr>
<td>pJQ200SK</td>
<td>P15a oriV; sacB; Gm(^R)</td>
<td>Quandt &amp; Hynes, 1993</td>
</tr>
<tr>
<td>pJET-nptII</td>
<td>pJET derivative; contains nptII cassette; Amp(^R), Km(^R)/Nm(^R)</td>
<td>Tim Haskett</td>
</tr>
<tr>
<td>pJET-Ω-aadA</td>
<td>pJET derivative; contains Ω-aadA; Amp(^R), Sm(^R), Sp(^R)</td>
<td>Tim Haskett</td>
</tr>
<tr>
<td>pPR3</td>
<td>pSV1/p15a oriV; nptII promoter upstream of gfp; Km(^R)</td>
<td>Rodpoothong <em>et al.</em>, 2009</td>
</tr>
<tr>
<td>pJR204</td>
<td>pFAJ1708 derivative; RK2 oriV; contains R7A rdfs under nptII promoter; Tc(^R)</td>
<td>Ramsay <em>et al.</em>, 2006</td>
</tr>
<tr>
<td>pMCC6</td>
<td>pSacB derivative; pBBR1 oriV; plasmid-curing vector containing repABC region of pMESC101; sacB; Km(^R)/Nm(^R)</td>
<td>R. Brewer, unpublished</td>
</tr>
<tr>
<td>pMCC7</td>
<td>pPR3 derivative; rdfs of CC1192 cloned into EcoRI fragment downstream of nptII promoter; Km(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>pMCC8</td>
<td>pEX18Tc derivative; intS deletion construct harbouring upstream and downstream region of intS and Ω-aadA cassette; sacB; Sm(^R), Sp(^R), Tc(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>pMCC9</td>
<td>pJQ200SK derivative; intS deletion construct harbouring upstream and downstream region of intS and Ω-aadA cassette; sacB; Gm(^R), Sm(^R), Sp(^R)</td>
<td>This study</td>
</tr>
<tr>
<td>pMCC10</td>
<td>pJQ200SK derivative; ICE deletion construct harbouring upstream and downstream region of ICEMcSym(^{192}) and nptII cassette; sacB; Km(^R)/Nm(^R), Gm(^R)</td>
<td>This study</td>
</tr>
</tbody>
</table>

Abbreviations: Amp\(^R\), Km\(^R\), Gm\(^R\), Sm\(^R\), Sp\(^R\), Nm\(^R\), Tc\(^R\) represent ampicillin, kanamycin, gentamycin, streptomycin, spectinomycin, neomycin or tetracycline resistance, respectively. aadA: aminoglycoside resistance. gfp: green fluorescent protein. nptII: neomycin phosphotransferase II enzyme. Ω: omega transcriptional terminator. sacB: levanosucrase.
2.2 Bioinformatic analysis of CC1192

Genome sequences were retrieved from The National Centre for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/) via Geneious version 11.0.5 (Kearse et al., 2012). Genes were analysed based on predicted protein sequences through the Integrated Microbial Genomes (IMG) database (Markowitz et al., 2012). Comparative sequence analyses were performed using the Basic Local Alignment Search Tool (BLAST), to compare either nucleotide sequences (BLASTN) for sequenced constructs, or protein sequences (BLASTP). Proteins were classified into Clusters of Orthologous groups (COGs) through IMG.

2.3 General molecular methods

Plasmid DNA was extracted using the Favorgen Plasmid Extraction Kit (Favorgen; cat No. FAPDE300) and genomic DNA was extracted using the Prepman Ultra Kit (Thermo-Fisher; 4318930). Digested PCR fragments or DNA treated with BamHI was purified using the FavorPrep Gel/PCR purification kit (Favorgen; FAGCK001-1). All purified DNA was eluted in PCR grade water. Yield and purity of DNA was measured using a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo-Fisher Scientific).

Agarose gel electrophoresis was routinely performed as per Sambrook et al. (1989), using either 1.2% or 2% (w/v) agarose in 1 x Tris/Acetic Acid/EDTA (TAE) buffer (Biorad; 1610773) with samples run at 80 V. Gels were stained with Ethidium Bromide (0.5 µg/mL) (Sigma-Aldrich; E8751) for 20-30 min. Agarose gel images were captured on a Gel Doc XR+ Imager (Bio-Rad).
Restriction enzymes EcoRI-HF (cat No. RS101S), KpnI-HF (R3142S), PstI-HF (R3140S), NcoI-HF (R3193S), XhoI-HF (R0146S), SpeI-HF(R3133S), XbaI (R0145S), BamHI-HF (R3136S) and SalI-HF (R3138S) as well as T4 DNA ligase (M0202S), Phusion High-Fidelity PCR Master Mix with HF buffer (M05315), Shrimp Alkaline Phosphatase (M0371L), 1-kb DNA ladder (N3232L) and 100-bp DNA ladder (N3231S) were all purchased through New England Biolabs (USA). All restriction enzymes were provided with a 10 x Cutsmart buffer (B7204S). GoTaq Green Master Mix (M7122) and a 100-bp ladder (G2101) were sourced from Promega (USA).

Oligonucleotide primers were designed using Geneious version 11.0.5 based on sequences from the NCBI database and produced by Integrated DNA technologies (Singapore) (Table 2.2). PCR cycling conditions for this study are as follows:

**PCR 1:** 1 cycle of 94°C for 5 min; 30 cycles of 94°C for 30 s, 59°C for 30 s, 70°C for 40 s with a final extension at 70°C for 3 min.

**PCR 2:** 1 cycle of 98°C for 30 s; 35 cycles of 98°C for 10 s, 72°C for 20 s with a final extension at 72°C for 2 min.

**PCR 3:** 1 cycle of 94°C for 5 min; 30 cycles of 94°C for 30 s, 59°C for 20 s, 70°C for 30 s with a final extension at 70°C for 3 min.

**PCR 4:** 1 cycle of 98°C for 30 s; 35 cycles of 98°C for 10 s, 72°C for 30 s with a final extension at 72°C for 5 min.

**PCR 5:** 1 cycle of 94°C for 5 min; 30 cycles of 94°C for 30 s, 59°C for 20 s, 70°C for 4 min 45 s with a final extension at 70°C for 5 min.

**PCR 6:** 1 cycle of 98°C for 30 s; 30 cycles of 98°C for 10 s, 69°C for 20 s, 72°C for 45 s with a final extension at 72°C for 5 min.
**PCR 7:** 1 cycle of 98°C for 30 s; 25 cycles of 98°C for 10 s, 72°C for 35 s with a final extension at 72°C for 5 min.

**PCR 8:** 1 cycle of 94°C for 5 min; 30 cycles of 94°C for 30 s, 59°C for 20 s, 70°C for 4 min 30 s with a final extension at 70°C for 5 min.
Table 2.2 Primers used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligonucleotide sequence (5’ – 3’)*</th>
<th>Replicon specificity</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 pMc1192-Fa</td>
<td>CGTTCCGAGACTTTGAACCCAGGA</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>2 pMc1192-Ra</td>
<td>CCTCAAAAGCTGGCAGATCGAAC</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>3 pMc1192-Fb</td>
<td>GATCAATGGTGGCCGAGAAGAC</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>4 pMc1192-Rb</td>
<td>CGCTGTTTCGACCGTTTGTTC</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>5 pMc1192-Fc</td>
<td>TTTCCCGAAGGAGATTGCAC</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>6 pMc1192-Rc</td>
<td>AAGCGCGATCGAGATATGAT</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>7 CC1192_attB_F</td>
<td>GTTGCCGAGACTGGTTTGTTC</td>
<td>CC1192</td>
<td>This study</td>
</tr>
<tr>
<td>8 CC1192_attB_R</td>
<td>TTGGTTTCTCCCTCGAAGCAGG</td>
<td>pMc1192</td>
<td>This study</td>
</tr>
<tr>
<td>9 CC1192_attP_F</td>
<td>GCCGATTGTCACAGGCTACT</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>10 CC1192_attP_R</td>
<td>CGGAGCAGATACAGATGATCC</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>11 rdfS&lt;sub&gt;5’&lt;/sub&gt;_EcoRI</td>
<td>ATGAAGAATTCGGAGGCAGAATGACGAGCA</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>12 rdfS&lt;sub&gt;3’&lt;/sub&gt;_EcoRI</td>
<td>TACTGAATTTCTCATGAGCCCTCCCTCGTG</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>13 nptII</td>
<td>ATGGCGAGGGGAGGACATCCAGATCTAG</td>
<td>pPR3</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>14 nptIII</td>
<td>ACTCCAGTGAAAGTCTTCCTCTCCT</td>
<td>pPR3</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>15 IntS_Up_F</td>
<td>ATCAAGGATTCCTCACTGAGCCTCCCTCCGT</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>16 IntS_Up_R</td>
<td>ATCAAGGATTCCTCACTGAGCCTCCCTCCGT</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>17 IntS_Dn_F</td>
<td>ATCAAGGATTCCTCACTGAGCCTCCCTCCGT</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>18 IntS_Dn_R</td>
<td>ATCAAGGATTCCTCACTGAGCCTCCCTCCGT</td>
<td>ICEMcSym&lt;sup&gt;192&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>19 pEX18TcF</td>
<td>GCCCTCTGAGCTATTACGCAC</td>
<td>pEX18Tc</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>20 pEX18TcR</td>
<td>TGGTTGGAATTGTGACAGCCGA</td>
<td>pEX18Tc</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>21 pJQ200SKF</td>
<td>AGCCGATAACAAATTCACACAGG</td>
<td>pJQ200SK</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>22 pJQ200SKR</td>
<td>ACTCACTATAGGGCAGATTGGG</td>
<td>pJQ200SK</td>
<td>T. Haskett</td>
</tr>
<tr>
<td>23 Sp7</td>
<td>GAAGTAATCGCAACATCCGCA</td>
<td>Ω-aadA cassette</td>
<td>T.H. Rogers</td>
</tr>
<tr>
<td>24 IntS_DXO_R</td>
<td>CTCAACTGTTACCTCCACTCC</td>
<td>CC1192</td>
<td>This study</td>
</tr>
<tr>
<td>25 ICE_Up_F</td>
<td>ATCAAGGACGAGGGCAGAGAAGGAC</td>
<td>ICE192</td>
<td>This study</td>
</tr>
<tr>
<td>26 ICE_Up_R</td>
<td>ATCAAGTCTAGATGGCAGAGGGAGGAGG</td>
<td>ICE192</td>
<td>This study</td>
</tr>
<tr>
<td>27 ICE_Dn_F</td>
<td>ATCAAGTCTAGATGGCAGAGGGAGGAGG</td>
<td>ICE192</td>
<td>This study</td>
</tr>
<tr>
<td>28 ICE_Dn_R</td>
<td>ATCAAGTCTAGATGGCAGAGGGAGGAGG</td>
<td>ICE192</td>
<td>This study</td>
</tr>
</tbody>
</table>

*Bold sequences indicate restriction sites as follows: GAATTCC (EcoRI), GGATCC (BamHI), CCAATTG (NcoI), GGCGGCGCG (NotI), GTTCGAG (SalI), TCTAGA (XbaI), CTCGAG (XhoI), ACTAGT (SpeI).
2.4 Transformation of plasmids into *E. coli*

*E. coli* DH5α cells (Bioline; cat No. Bio-85027) were thawed on ice for 30 min before a 5 µl aliquot of the ligation reaction products was added to 50 µl of competent cells for incubation on ice for 30 min, followed by heat shock treatment at 42°C for 50 s. The suspension was replaced on ice for a further 2 min and then 450 µl of SOC medium, prepared as per Sambrook *et al.* (1989), was added. The culture was incubated for 2 h at 37°C on a gyratory shaker at 250 rpm and aliquots of the culture were then spread onto LB agar containing appropriate antibiotic selection (see Table 2.1).

For transformation into *E. coli* ST18, a 200 µl aliquot of ST18 cells was thawed on ice before 2 µl of purified plasmid DNA (Section 2.3) and 2 µl of DMSO were added. The mixture was incubated on ice for 30 min, before heat shock at 42°C for 90 s and a further 2-min incubation on ice. An 800 µl aliquot of SOC with ALA was added to the cells and the cell suspension incubated at 37°C on a gyratory shaker for 90 min. Aliquots of the culture were then spread onto LB agar supplemented with ALA and appropriate antibiotic selection.

2.5 Biparental conjugation between vectors and CC1192

All constructs were transformed into *E. coli* ST18 cells prior to conjugation with CC1192 (Section 2.4) to allow *E. coli* growth suppression on selection plates devoid of ALA. CC1192 was grown to stationary phase in TY broth while two *E. coli* ST18 strains, either harbouring the relevant construct or the empty vector (as a control), were grown to log phase (OD₆₀₀nm ~ 0.180 – 0.820) in 5 ml LB broth with ALA and the appropriate antibiotics for construct maintenance (see Table 2.1).
Cultures were washed by harvesting cells via centrifugation at 5,500 x g for 5 min, discarding the supernatant and resuspending cell pellets in 1 ml sterile 0.89% (w/v) saline. The resuspension was centrifuged again and the cell pellets finally resuspended in 250 µl sterile 0.89% (w/v) saline. A 50 µl aliquot each of CC1192 and ST18 harbouring the construct were combined and spotted onto a TY plate containing ALA. As controls, 50 µl of each parent strain (CC1192 or ST18) was also spotted separately onto the TY plate. This procedure was repeated for conjugation between CC1192 and ST18 harbouring the empty vector control. The spots were allowed to dry before inversion and incubation at 28°C for one day if introducing a broad host range plasmid (e.g. pPR3 constructs), or two days if introducing a narrow host range plasmid (e.g. pEX18Tc or pJQ200SK constructs). Spots were scraped off the plate after the incubation period, resuspended in 1 ml sterile 0.89% (w/v) saline and dilutions spread onto TY agar (no added ALA) with relevant antibiotics for selection of the desired transconjugants.

2.6 Curing pMc1192 from the CC1192 genome

The plasmid-curing vector, pMCC6, was created previously by cloning the repABC region from pMESCI01 into the pSacB vector (R. Brewer, unpublished results). The construct was transformed into E. coli ST18 (Section 2.4) before conjugation into CC1192 (Section 2.5). The presence of the CC1192 plasmid, pMc1192, was screened via colony PCR using primer pairs 1 & 2, 3 & 4 and 5 & 6, binding to three separate regions dispersed around the plasmid as well as primers 7 & 10 to screen for the attL site for confirmation of strain (Table 2.2). The end of a pipette tip was used to transfer cells from individual colonies to a 10 µl reaction containing 1 x GoTaq Green Master
Mix and 0.5 µl of each primer using PCR 1 conditions, before being visualised on a gel (Section 2.3).

Three colonies were inoculated in non-selective TY broth to facilitate the loss of the curing vector. On S/RDM, 100 µl of stationary phase broth was plated and dried as a primary streak before dilution streaking. Single colonies arising were replica patched firstly onto G/RDM with neomycin (Nm) and then on G/RDM only, to ensure complete loss of the curing vector. Two of the colonies unable to grow on GRDM supplemented with Nm were stocked and designated as MCC69 and MCC70.

2.7 Eckhardt Gel Electrophoresis

Eckhardt gel electrophoresis was adapted from that of Eckhardt (1978). Briefly, the two plasmid cured derivatives, MCC69 and MCC70, wild type CC1192 and the reference strain Rlv3841 (see Table 2.1) were grown in triplicate to optical densities (OD_{600nm}) of approximately 0.3 as measured by light spectrophotometry (Hitachi U-1900). A 200 µl aliquot of cells from each culture was chilled separately on ice for 10 min before 1 ml of cold 0.3% (w/v) sarcosyl in 1 x Tris/Borate/EDTA (TBE) buffer (89 mM tris, 82 mM boric acid and 2 mM EDTA) was added, mixed by inversion and left on ice for a further 10 min. The mixture was centrifuged (20,817 x g for 5 min at 4°C) and aspirated, before gentle resuspension in 25 µl lysis solution (0.1 mg/mL lysozyme in E1 solution; 10% (w/v) sucrose, 10 µg/mL RNase A solution (Invitrogen; 12091-012) in 1 x TBE). Immediately, 20 µl of the sample was loaded onto 0.75% (w/v) agarose with 1% (w/v) sodium dodecyl sulphate and left for 30 min to settle. The gel was then run at 70 V for 16 h at 4°C before staining in 1 µg/mL Ethidium
Bromide in 1 x TBE for 60 min followed by de-staining in 1 x TBE for 15 min and visualisation as per section 2.3.

2.7.1 Sequencing of MCC69 and MCC70

Genomic DNA from MCC70 was extracted using a DNeasy Blood and Tissue Kit (Qiagen; 69504) according to the manufacturer’s instructions and sent for Illumina MiSeq Next-Generation Sequencing. Library preparations were prepared by the J.P.R. laboratory (Curtin University), as were the de novo assembly using SPAdes (Bankevich et al., 2012) and alignment of the raw MiSeq reads to the CC1192 chromosome (accession no. NZ_CP015062) and the plasmid, pMc1192 (accession no. CP015063) using BowTie 2 (Langmead & Salzberg, 2012).

2.8 Construction of rdfS overexpression vector, pMCC7

2.8.1 Preparation of rdfS from CC1192

Genomic DNA was extracted from stationary phase CC1192 cultures as per section 2.3. PCR amplification of the 270-bp rdfS (locus tag: A4R28_RS20865) was performed with primers 11 and 12 (Table 2.2), imparting EcoRI restriction sites to either end of the fragment, in a 50 µl volume containing 25 µl of 1 x Phusion High Fidelity DNA polymerase Master Mix, 2.0 µl of CC1192 gDNA, 0.5 µM of each primer and 1.5 µl of DMSO (NEB). PCR cycling conditions were as stated for PCR 2 (Section 2.3).

The PCR product was electrophoresed alongside a 100-bp DNA ladder (Promega) to confirm amplicon size. The amplified DNA was purified (Section 2.3) and the
products digested by EcoRI in a 50 µl reaction containing 948 ng of the purified PCR product and 20 units EcoRI-HF in 1 x Cutsmart buffer, incubated overnight at 37°C. The reaction was heat inactivated at 65°C for 20 min before purification.

2.8.2 Preparation of pPR3 and ligation with the *rdfS* fragment

Plasmid pPR3 was extracted from overnight *E. coli* cultures (Section 2.3) and 3 µg was digested with 40 units of EcoRI-HF in a 1 x Cutsmart Buffer, in a 50 µl reaction for 2 h at 37°C. Three units of shrimp alkaline phosphatase were then added and incubated for another hour. The reaction was then heat-inactivated at 65°C for 20 min. Three 10 µl ligation reactions were performed, each containing 30 ng of EcoRI-HF digested pPR3 with 20 units of T4 DNA ligase in a 1 x ligase buffer supplied by the manufacturer. Either 1 µl, 2 µl or 4 µl of the EcoRI-HF-digested *rdfS* insert was added separately to each reaction. All reactions were incubated at room temperature for 30 min after which ligation products from all three reactions were combined into a single tube for transformation into *E. coli* DH5α cells (Section 2.4).

2.8.3 Screening of pMCC7

The presence of *rdfS* in pPR3 was confirmed by colony PCR. In a 10 µl reaction, a pipette tip was used to transfer cells from an individual colony on a LB agar plate to a PCR reaction consisting of 1 x GoTaq Green Master Mix and 0.5 µM of primers 13 and 14 (Table 2.2) to amplify the pPR3 multiple cloning site for a 472-bp product. Cycling conditions were as per PCR 3 (Section 2.3). A second colony PCR using primer 12 (binding within the cloned *rdfS*) and 13 (binding in the pPR3 multiple
cloning site) was performed to determine the orientation of the insert, with the same cycling conditions as aforementioned to amplify a 395-bp fragment. An amplicon from the initial PCR of the correct size and was Sanger sequenced by the Australian Genome Research Facility (AGRF), and the resultant sequence verified by alignment to the available genome sequence using Geneious software.

2.9 Construction of integrase (intS) deletion vector using pEX18Tc, pMCC8

2.9.1 Preparation of upstream and downstream of intS

Genomic DNA was extracted from stationary phase CC1192 cultures as per section 2.3. Amplification of the 1,028-bp upstream region of intS (A4R28_RS20660), (including 57-bp of the 5’ region of the gene) was performed with primers 15 and 16 (Table 2.2), designed to add a BamHI site to the 5’ end of the fragment and NcoI restriction site to the 3’ end. The 1,076-bp downstream region was amplified with primers 17 and 18 (including 69-bp of the 3’ region of the gene) with a NotI and SalI restriction sites added to the 5’ and 3’ end of the fragment, respectively. Reactions were performed with Phusion High Fidelity DNA polymerase, as per section 2.8.1, using PCR 4 conditions (Section 2.3). Resultant PCR products were visualised by agarose gel electrophoresis and then purified (Section 2.3).

A 40 µl aliquot of both upstream and downstream regions were each digested in 1 x Cutsmart buffer with either 20 units each of BamHI-HF and NcoI-HF (upstream region) or 20 units each of NotI-HF and SalI-HF (downstream region) in a 50 µl reaction volume for 2 h at 37°C. Both reactions tubes were subsequently purified (Section 2.3).
2.9.2 Preparation of pJET-Ω-aadA and pEX18Tc for cloning

The Ω-aadA fragment, encoding streptomycin (Sm) and spectinomycin (Sp) resistance and containing the Ω-transcriptional terminator, was prepared from pJET-Ω-aadA by digestion of 5 µg of the vector with 50 units each of NcoI-HF and NotI-HF in 1 x Cutsmart buffer in a 100 µl reaction. Plasmid pEX18Tc was similarly prepared for cloning by digestion of 5 µg of purified DNA with 50 units each of BamHI-HF and SalI-HF in 1 x Cutsmart Buffer in a 100 µl total volume. Both reactions were incubated at 37ºC for 3 h, followed by heat-inactivation at 80ºC for 20 min and purification (Section 2.3). The treated DNA was visualised by agarose gel electrophoresis (Section 2.3) alongside the untreated DNA and a 1-kb ladder to ensure adequate cutting and the production of the expected fragment sizes.

2.9.3 Ligation reactions producing pMCC8

To produce the intS inactivation vector pMCC8, three separate ligation reactions were performed with different ratios of the various digested constituents. Reaction 1 contained 2 µl of cut pJET-Ω-aadA and 1 µl each of the upstream region, downstream region and cut pEX18Tc. Reaction 2 contained 4 µl of the cut pJET-Ω-aadA, 2 µl each of the upstream and downstream regions and 1 µl of cut pEX18Tc. Reaction 3 contained 2 µl of cut pJET-Ω-aadA vector, 1 µl each of both the upstream and downstream regions, and 2 µl of cut pEX18Tc. To each reaction 20 units of T4 DNA ligase and 1 x ligase buffer was added before all reactions were incubated overnight at 16ºC. The reaction products were transformed into E. coli DH5α as described in section 2.5, with resultant E. coli cultures spread onto LB media containing Tc and Sp to select for transformants with the correct construct.
2.9.4 Screening of pMCC8

The presence of the intS inactivation vector pMCC8 in transformant E. coli was confirmed firstly by colony PCR (as per section 2.8.3) using primers 19 and 20. Four colonies yielding the expected PCR product of 4,618-bp were then selected and cultured in LB broth (with Tc and Sp). Plasmid DNA was extracted from these cultures (Section 2.3) and screened by restriction digestion whereby approximately 350 ng of DNA was digested by 10 units each of KpnI-HF and PstI-HF in 1 x Cutsmart buffer in 20 µl reactions for 3 h at 37°C. Digest products were visualised by agarose gel electrophoresis (Section 2.3) alongside a 1-kb ladder to ensure fragments were of the expected size. One of the correct constructs was then selected and transformed into ST18 (Section 2.4) and spread onto LB containing ALA, Sp and Tc.

2.10 Construction of integrase deletion vector using pJQ200SK, pMCC9

The upstream and downstream regions (Section 2.9.1) and Ω-aadA (Section 2.9.2) prepared earlier for ligation into pEX18Tc were used here for creation of pMCC9. Purified pJQ200SK was digested with 50 units each of BamHI-HF and SalI-HF in a 1 x Cutsmart buffer in a 50 µl reaction for 3 h at 37°C and then purified (Section 2.3). A ligation reaction was carried out as per section 2.9.3 but containing the pJQ200SK backbone rather than pEX18Tc. The ligation mixture was transformed into E. coli DH5α (Section 2.4) before being plated on LB containing Sp and gentamycin (Gm).

The presence of the intS inactivation vector pMCC9 in transformant E. coli was confirmed by colony PCR as described in section 2.8.3 using primers 19 and 20 (Table 2.2), amplifying the region spanning the multiple cloning site of pJQ200SK. Colonies containing the desired 4,573-bp product were further confirmed by restriction
digestion performed as per section 2.9.4, with 110 ng plasmid DNA and 10 units each of KpnI-HF and NcoI-HF. One of the constructs containing the correct fragments was then transformed into ST18 as outlined in section 2.4.

2.11 Site-directed mutagenesis of intS

The intS deletion vector, pMCC9, was introduced into CC1192 via a biparental conjugation (Section 2.4) and a single crossover plasmid integrant was selected by plating on TY with Sp and Gm. Before selection for a double crossover intS deletion mutant, whereby most intS would be replaced with the Ω-aadA cassette, four single crossover transconjugants were grown to stationary phase in TY with Sp. Aliquots of 1 µl, 10 µl and 100 µl of each broth was spread onto TY supplemented with sucrose, Sp and Sm. Sucrose resistant transconjugants were replica-patched firstly onto TY with Sp and Gm, and then onto TY and Sp. Transconjugants unable to grow on Gm were suspected to contain the intS double crossover replacement.

DNA was extracted from three of these transconjugants as well as CC1192 and screened for the insertion of the Ω-aadA cassette in the integrase similar to section 2.8.1 using primers 23 and 24 in PCR 6 conditions (Section 2.3). The amplified product of transconjugant one was Sanger Sequenced (ARGF) to confirm the presence and location of the insertion. Further screening of the attB, attP and attL sites was performed similarly to section 2.6 using primers 7 and 8, 9 and 10 or 7 and 10, respectively, in PCR 1 conditions (Section 2.3) to confirm that the ICE had been stabilised.
2.12 Construction of ICE deletion construct, pMCC10

The regions immediately adjacent to ICEMcSym\textsuperscript{1192} were amplified from CC1192 genomic DNA. Primers 25 and 26 were used to amplify the 1,566-bp upstream region encompassing the entire \textit{ser}-tRNA into which ICEMcSym\textsuperscript{1192} integrates as well as incorporating a NotI and XbaI restriction site at the 5’ and 3’ end of the fragment, respectively. Primers 27 and 28 were used to amplify the 1,477-bp downstream region omitting the \textit{attR} site and adding XhoI and SpeI restriction sites to the 5’ and 3’ region of the PCR product, respectively. The PCR reaction was carried out as per section 2.8.1 using PCR 7 conditions (Section 2.3). The products were visualised on a gel and then purified (Section 2.3). Restriction digestion was performed as per section 2.9.1 using 20 units each of NotI-HF and XbaI (upstream region) or 20 units each of XhoI-HF and SpeI-HF (downstream region).

\textit{pJET-nptII} was extracted from DH5\textit{a} (Section 2.3) before digestion of 4.7 µg of plasmid DNA with 50 units of XbaI and XhoI-HF (Section 2.9.2) to obtain the \textit{nptII} fragment. Plasmid \textit{pJQ200SK} was similarly prepared, but with 3.7 µg digested with 50 units of NotI-HF and SpeI-HF (Section 2.9.2). Both vectors were separated by gel electrophoresis and run alongside uncut vector and a 1-kb ladder before visualisation (Section 2.3) to ensure the presence of the appropriate fragments.

The ligation was set up as per section 2.9.3, but with \textit{pJET-nptII} replacing \textit{pJET-\Omega-aadA} and \textit{pJQ200SK} rather than \textit{pEX18Tc}. The three reactions were combined and transformed into \textit{E. coli} DH5\textit{a} (Section 2.4) before plating onto LB supplemented with Nm and Gm. The presence of the ICE deletion vector in DH5\textit{a} was confirmed by colony PCR as per section 2.8.3, using primers 21 and 22, both binding to the \textit{pJQ200SK} vector and flanking the cloned region, in PCR 8 conditions (Section 2.3). The presence of a 4,433-bp product in colonies indicated proper ligation. These
colonies were then used in a restriction digestion of 90 ng of DNA as per section 2.9.4, but using 10 units of XhoI-HF.

2.13 Symbiotic performance of plasmid-cured derivatives with *Cicer arietinum*

A 44-day glasshouse house trial was conducted to compare the symbiotic phenotype of plasmid-cured strains, MCC69 and MCC70, with the wild type CC1192 on *C. arietinum* cv. Neelam. These treatments were also compared to uninoculated control plants either fed nitrogen (N+) or nitrogen starved (N-). A 50:50 soil mixture of yellow sand and Gingin quartz was made and adjusted to pH 6.5 – 7 by addition of 11.25 mM Iron III Sulphate per ~ 28 kg of sand mix. Small olive pots (0.8 L) containing the sand mix were then steam sterilized and flushed twice with boiling water before surface sterilized seeds were sown. Seeds were sterilized in 70 % (v/v) ethanol for 1 min and bleach for 3 min, followed by six rinses in sterile water. The sterilized seeds were spread onto 0.9% (w/v) water agar with 1 ml of sterile water before incubation at 22°C for two days.

The inoculant strains were streaked in a cross-hatch pattern onto TY and incubated at 28°C for 6 days. The bacteria were washed-off the plates with sterile 0.9% (w/v) sucrose solution and resuspended to an OD$_{600\text{nm}}$ of 1.0. Ten pots were allocated to each treatment with two seeds being initially sown per pot and being reduced to one plant per pot two weeks post-sowing. For inoculated treatments, 1 ml of inoculum was added to each seed at sowing. Miles and Misra counts of each inoculum treatment (Howieson & Dilworth, 2016) established the number of live cells added to each seed for the different treatments. Bacterial counts were as follows: CC1192, 2.8 x $10^8$ cells; MCC69, 2.6 x $10^8$; MCC70, 1.2 x $10^8$ cells. Following inoculation the surface of each
pot was covered with sterile beads and provided with a sterile watering tube as per Howieson and Dilworth (2016).

Plants were grown at a mean temperature of 22°C and watered three times a week according to plant requirements. Plants were also fed 20 ml of nitrogen-free nutrient solution according to Howieson and Dilworth (2016) once a week for the first three weeks, and then twice a week until harvest. The N-fed plants received 2 ml of 0.01 M KNO$_3$ at the same frequency as stated for the nutrient solution.

2.14 Plant harvest and nodule sampling

Nodules from the aforementioned glasshouse trial were assessed for size, colour and location, and representatives from each treatment selected for verification of nodule occupants. These nodules were excised and surface-sterilized in the same manner as the seeds (section 2.13), but with 30 s and 1 min exposure to the ethanol and bleach treatments, respectively. Sterilised nodules were crushed and streaked onto TY agar. Colony PCR of isolated organisms was performed according to section 2.6 using PCR 1 conditions (Section 2.3) and primer pairs 1 & 2, 3 & 4 and 5 & 6 (see Table 2.2) to screen for plasmid presence or absence. In addition, primers 7 & 10 and 8 & 9, amplifying the left (att$L$) and right (att$R$) junctions of ICEMcSym$^{1192}$ in CC1192, respectively, were used for confirmation of strain. PCR products were run on a gel to visualise the amplified fragments (Section 2.3). Plant shoots were collected by cutting at the hypotcotyl, while nodules on all remaining treatments were removed from plant root systems. Shoots and root nodules were dried separately in 60°C for 1-2 days before weighing.
2.15 Nodule sections of plasmid-cured derivatives with *Cicer arietinum*

A glasshouse experiment was set up as described in section 2.13 to compare the morphology of nodules on *C. arietinum* cv. Ambar inoculated separately with MCC60, MCC70 and CC1192. Nodules were collected from plants of each treatment after 56 days of growth and fixed with 3% (v/v) glutaraldehyde and 25 mM phosphate buffer pH 7 at 4°C overnight. Nodule sections and subsequent light micrographs were prepared as previously described (Richardson *et al.*, 1960; Spurr, 1969) by the Murdoch University Veterinary and Life Sciences pathology laboratory.
Chapter 3: Results

3.1 Determining the symbiotic role of pMc1192

3.1.1 Bioinformatic analysis of pMc1192

The genome of *M. ciceri* CC1192 possesses a 648-kb plasmid named pMc1192, which harbours a *repABC*-type (gene locus tags: A4R28_31005 - A4R28_30995) origin of plasmid replication. The plasmid carries a total of 613 putative protein encoding genes, with Clusters of Orthologous groups (COG) analysis categorising 338 of these proteins (55% of total) into 22 out of 25 possible functional groups (Table 3.1). Of these genes, more than 40% were grouped into four categories: Carbohydrate transport and metabolism, Transcription, Amino acid transport and metabolism, and Energy production and conversion. The remaining 275 (45%) putative protein encoding genes could not be assigned a COG, indicating that these proteins have unknown functions or represent hypothetical genes.
Table 3.1 Summary of the Cluster of Orthologous Groups of proteins (COGs) for the 613 putative protein encoding genes harboured on pMc1192.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Number of genes†</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid transport and metabolism</td>
<td>33</td>
<td>5.38</td>
</tr>
<tr>
<td>Carbohydrate transport and metabolism</td>
<td>37</td>
<td>6.04</td>
</tr>
<tr>
<td>Cell cycle control, cell division, chromosome partitioning</td>
<td>3</td>
<td>0.49</td>
</tr>
<tr>
<td>Cell motility</td>
<td>9</td>
<td>1.47</td>
</tr>
<tr>
<td>Cell wall/membrane/envelope biosynthesis</td>
<td>16</td>
<td>2.61</td>
</tr>
<tr>
<td>Chromatin structure and dynamics</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Coenzyme transport and metabolism</td>
<td>15</td>
<td>2.45</td>
</tr>
<tr>
<td>Cytoskeleton</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Defence mechanisms</td>
<td>12</td>
<td>1.96</td>
</tr>
<tr>
<td>Energy production and conversion</td>
<td>33</td>
<td>5.38</td>
</tr>
<tr>
<td>Extracellular structures</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Function unknown</td>
<td>21</td>
<td>3.43</td>
</tr>
<tr>
<td>General function prediction only</td>
<td>22</td>
<td>3.59</td>
</tr>
<tr>
<td>Inorganic ion transport and metabolism</td>
<td>12</td>
<td>1.96</td>
</tr>
<tr>
<td>Intracellular trafficking, secretion, and vesicular transport</td>
<td>26</td>
<td>4.24</td>
</tr>
<tr>
<td>Lipid transport and metabolism</td>
<td>16</td>
<td>2.61</td>
</tr>
<tr>
<td>Mobilome: prophages, transposons</td>
<td>3</td>
<td>0.49</td>
</tr>
<tr>
<td>Nucleotide transport and metabolism</td>
<td>1</td>
<td>0.16</td>
</tr>
<tr>
<td>Post-translational modification, protein turnover, chaperones</td>
<td>16</td>
<td>2.61</td>
</tr>
<tr>
<td>Replication, recombination and repair</td>
<td>6</td>
<td>0.98</td>
</tr>
<tr>
<td>RNA processing and modification</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Secondary metabolites biosynthesis, transport and catabolism</td>
<td>4</td>
<td>0.65</td>
</tr>
<tr>
<td>Signal transduction mechanisms</td>
<td>18</td>
<td>2.94</td>
</tr>
<tr>
<td>Transcription</td>
<td>34</td>
<td>5.55</td>
</tr>
<tr>
<td>Translation, ribosomal structure and biogenesis</td>
<td>1</td>
<td>0.16</td>
</tr>
<tr>
<td>Not in COG</td>
<td>275</td>
<td>44.86</td>
</tr>
<tr>
<td><strong>Total Genes</strong></td>
<td><strong>613</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

† Determined by Integrated Microbial Genomes annotation and genome browsing platform. Each category shows the number and percentage of genes in the replicon in each category.

A search of the pMc1192 sequence for potential symbiosis-related genes revealed the presence of *fixNOQP* (A4R28_30260 - A4R28_30275), encoding a putative *cbb*-type microaerobic terminal oxidase (Preisig *et al.*, 1993; 1996b); *fixGHIS* (A4R28_30280 - A4R28_30295), encoding a possible microaerobic copper-uptake system (Kahn *et al.*, 1989; Preisig *et al.*, 1996a; Koch *et al.*, 2000); and *fixLJ* (A4R28_30350 - A4R28_30355) and *fixK* (A4R28_30370) encoding a putative low O2-responsive, two-
component regulatory system and transcriptional activator (Batut et al., 1989; Cosseau & Batut, 2004; Tsoy et al., 2016). All three systems have been shown to be essential for symbiotic nitrogen fixation in other legume-rhizobia interactions (Batut et al., 1989; Preisig et al., 1993; 1996a). BLASTP analysis of the entire CC1192 genome showed that a second copy of genes encoding FixNOQP and FixGHI were present on the chromosome within ICEMcSym\textsuperscript{1192}. Percentage identities > 89% for these chromosomal copies indicate that the proteins are highly similar (Table 3.2). In contrast, fixLJ and fixK were absent from ICEMcSym\textsuperscript{1192} with only low percentage identity matches (29-34%) identified outside this region on the chromosomal backbone. pMc1192 appears to also harbour the sole copy of fixS in CC1192, with BLASTP and subsequent BLASTN searches returning no significant matches for chromosomal or ICE-encoded genes. Therefore, pMc1192 harbours some genes that may be essential to symbiotic nitrogen fixation of CC1192 with C. arietinum.
Table 3.2 Symbiosis-related genes encoded on pMc1192 and the location and similarity of the chromosomal copies within the CC1192 genome.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Locus Tag</th>
<th>Length (aa)</th>
<th>Locus Tag</th>
<th>Length (aa)</th>
<th>Location</th>
<th>Identity/Coverage*</th>
</tr>
</thead>
<tbody>
<tr>
<td>FixN</td>
<td>A4R28_30260</td>
<td>536</td>
<td>A4R28_RS20715</td>
<td>539</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>491/536 (92%)</td>
</tr>
<tr>
<td>FixO</td>
<td>A4R28_30265</td>
<td>243</td>
<td>A4R28_RS20710</td>
<td>243</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>229/243 (94%)</td>
</tr>
<tr>
<td>FixQ</td>
<td>A4R28_30270</td>
<td>49</td>
<td>A4R28_RS20705</td>
<td>49</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>44/49 (90%)</td>
</tr>
<tr>
<td>FixP</td>
<td>A4R28_30275</td>
<td>287</td>
<td>A4R28_RS20700</td>
<td>287</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>269/287 (94%)</td>
</tr>
<tr>
<td>FixG</td>
<td>A4R28_30280</td>
<td>519</td>
<td>A4R28_RS20695</td>
<td>519</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>469/519 (90%)</td>
</tr>
<tr>
<td>FixH</td>
<td>A4R28_30285</td>
<td>165</td>
<td>A4R28_RS20690</td>
<td>165</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>147/165 (89%)</td>
</tr>
<tr>
<td>FixI</td>
<td>A4R28_30290</td>
<td>762</td>
<td>A4R28_RS20685</td>
<td>759</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>680/756 (90%)</td>
</tr>
<tr>
<td>FixS</td>
<td>A4R28_30295</td>
<td>55</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>FixL</td>
<td>A4R28_30350</td>
<td>518</td>
<td>A4R28_RS19445</td>
<td>677</td>
<td>CC1192 backbone</td>
<td>107/368 (29%)</td>
</tr>
<tr>
<td>FixJ</td>
<td>A4R28_30355</td>
<td>204</td>
<td>A4R28_RS28800</td>
<td>219</td>
<td>CC1192 backbone</td>
<td>57/167 (34%)</td>
</tr>
<tr>
<td>FixK</td>
<td>A4R28_30370</td>
<td>215</td>
<td>A4R28_RS20720</td>
<td>197</td>
<td>ICEMcSym\textsuperscript{1192}</td>
<td>60/189 (32%)</td>
</tr>
</tbody>
</table>

*Percentage identity of plasmid-encoded genes compared to those found on ICEMcSym\textsuperscript{1192} or on the chromosomal backbone (i.e. located on the chromosome but outside of ICEMcSym\textsuperscript{1192}).

3.1.2 Curing of pMc1192 from CC1192

To determine whether genes on pMc1192 are essential for symbiotic nitrogen fixation of CC1192 with *C. arietinum*, a plasmid-cured version of CC1192 was created using plasmid incompatibility. Plasmid incompatibility is where plasmids with the same replication and partitioning system cannot be maintained in the same cell (Ramírez-Romero *et al.*, 2000; Uraji *et al.*, 2002), therefore, should two plasmids with the same replication and maintenance system be in a cell, selection for one plasmid can lead to loss of the other. The curing vector pMCC6, produced as part of a previous study, carries the repABC genes cloned from pMESCI01, a repABC-type accessory plasmid from the *Biserrula pelecinus*-nodulating *M. ciceri* bv. *biserrulae* WSM1271 (R. Brewer, unpublished results). As the amino acid sequence of RepABC from pMc1192 is 100% identical to that of pMESCI01, pMCC6 was mobilised into *M. ciceri* CC1192.
A total of ten Nm resistant transconjugants were PCR screened for the loss of the plasmid using three primer pairs, designed to bind three equidistant regions on the plasmid (approximately 210-kb apart). A CC1192 chromosomal-specific primer \((attL)\) was also added as a control. Isolates two-through-nine yielded the expected 501-bp product for the chromosomal primer pair, with no products observed for any of the plasmid-specific primers (Figure 3.1). This suggested that pMc1192 had been cured from eight of the screened CC1192 transconjugants. The loss of the plasmid curing vector pMCC6 was induced in transconjugants 2 and 4 though sucrose counter-selection. This was made possible by the \(sacB\) gene on pMCC6, which encodes levansucrase, an enzyme that confers sucrose sensitivity to cells. Sucrose-tolerant colonies were subsequently replica-patched on media with and without Nm to ensure pMCC6 was lost. No difference in growth or phenotype between the wild type or either plasmid-cured derivatives were observed. Two sucrose resistant and Nm sensitive transconjugants were stored as strains MCC69 and MCC70.
Figure 3.1 Image of electrophoresed PCR products amplified from 10 transconjugant colonies screening for presence or absence of PMc1192 after introduction of the plasmid curing vector, pMCC6, into *M. ciceri* CC1192. **Ladder:** 100-bp NEB ladder with band sizes (bp, highest to lowest) of 1517, 1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100; **Lanes 1-10:** *M. ciceri* CC1192 transconjugants colonies screened as follows; **attL:** attL site amplified with primers 7 and 10 for a 501-bp product; **a, b & c:** amplification of three regions around PMc1192 using primers 1 and 2 (a), 3 and 4 (b) or 5 and 6 (c) for a 500-bp, 525-bp and 507-bp product, respectively; **NTC:** No template control; **CC1192:** *Mesorhizobium ciceri* CC1192.

Strains MCC69 and MCC70 were next screened by Eckhardt gel electrophoresis, a well-established method for separating and visualising bacterial plasmids (Eckhardt, 1978; Hirsch et al., 1980). While bands of approximately 0.65-Mb were visible in all lanes carrying whole-cell extracts of wild-type CC1192 harvested at three separate optical densities, these bands were absent in MCC69 and MCC70 extracts (Figure
3.2). The absence of a band in MCC69 and MCC70 samples was consistent with the loss of the plasmid from the genomes of these two isolates. Finally, MCC70 was selected for whole genome sequencing to confirm the loss of pMc1192. From 933, 302 reads, the genome of MCC70 was assembled into 29 contigs totalling 6,168,350 bp, or 89% of the complete genome of CC1192 (entire genome is 6,943,628 bp). There was 98.10% nucleotide similarity between the assembled MCC70 genome and the CC1192 chromosome using Bowtie2. In contrast, a similar alignment of MCC70 sequence reads to pMc1192 returned an alignment of only 0.12%. As the 648,231 bp plasmid, pMc1192, constitutes 9.34% of the entire CC1192 genome, this lack of sequence alignment confirms that pMc1192 was successfully removed from the MCC70 genome. The 0.12% alignment is likely to be the fix gene duplicates identified in section 3.1.1. Therefore, MCC69 and MCC70 represent plasmid-cured derivatives of \textit{M. ciceri} CC1192.

![Image of an Eckhardt gel electrophoresis of wild-type CC1192 and two putative plasmid-cured isolates, MCC69 and MCC70, at three different cell optical densities (OD\textsubscript{600nm}); Set 1: 0.157, 0.106, 0.080; Set 2: 0.289, 0.241, 0.200; Set 3: 0.497, 0.376, 0.359. The OD\textsubscript{600nm} of the ladder was 0.194. A band corresponding to 0.65-Mb is present at all three optical densities for CC1192, while this band is absent in the lanes of plasmid-cured derivatives. \textbf{Ladder:} Plasmids of \textit{Rhizobium leguminosarum} \textit{bv. viciae} 3841; pRL12 (0.87-Mb), pRL11 (0.68-Mb); pRL10 (0.48-Mb), pRL9 (0.35-Mb) and pRL8/pRL7 (0.14- and 0.15-Mb); \textbf{Asterisks:} Indicate the presence of pMe1192.]

Figure 3.2
3.1.3. Assessment of the symbiotic phenotype of plasmid-cured derivatives

To ascertain whether the symbiotic genes found on pMc1192 were essential for symbiotic performance, the nodulation and nitrogen fixation phenotype of the two plasmid-cured strains, MCC69 and MCC70, were compared to that of the parent strain CC1192 on *Cicer arietinum*. Plants were harvested 44 days after inoculation with pure cultures of each strain, and plant shoot dry weights analysed.

All inoculated pots contained large plants with dark green shoots, consistent with all inoculated treatments fixing nitrogen with *C. arietinum* (Figure 3.3). The N-fed (N+) treatment produced plants with dark green shoots and were visually larger than the inoculated treatments. Although most of the N+ control plants had no nodules, 8 and 14 large red nodules were located on two out of ten individual plants at the very base of the root system. The location of these nodules very low down the root system is indicative of late nodulation and therefore is unlikely to have impacted plant growth. One N+ plant had fungal contamination on the soil and root system and was therefore excluded in dry weight analysis. In contrast, the N-starved (N-) plants had light green leaves and were significantly smaller than the N+ or inoculated plants.
Figure 3.3 Symbiotic phenotype of *Mesorhizobium ciceri* CC1192 and plasmid-cured strains, MCC69 and MCC70, inoculated with *Cicer arietinum*. Uninoculated plants were either nitrogen starved (N-) or nitrogen fed (N+). Plants were grown for 44 days in free-draining pots.

Statistical analysis by ANOVA of the mean shoot dry weight of each treatment determined that there was no significant difference between MCC69, MCC70 or CC1192 treatments (Figure 3.4). A Tukey’s post-hoc test determined that the N+ and N- controls were significantly different from each other and to the inoculated treatments (p < 0.05), but there was no significant difference between the inoculant strains.
Figure 3.4 Mean shoot dry weight (g/plant) of *Cicer arietinum* grown in free-draining pots inoculated with the commercial inoculant, *M. ciceri* CC1192, or either plasmid-cured derivative, MCC69 or MCC70. Uninoculated controls were either fed nitrogen (N+) or left without (N-). Error bars represent the standard error of each group. Different letters represent statistically significant differences according to Tukey’s post-hoc test (p > 0.05).

### 3.1.4 Nodulation

Nodule appearance, location, weight and number were also compared to determine whether loss of pMc1192 altered nodulation characteristics. Visual examination revealed that all three inoculated treatments had induced large, pink, indeterminant nodules on *C. arietinum* that tended to cluster in groups near the hypocotyl and be individually dispersed through the lateral roots. The mean nodule number and nodule weight (per plant) was also measured to establish if there was any difference in the extent or size of nodules. ANOVA found no significant difference (p > 0.05) in the nodule number or weight between inoculated treatments (Figure 3.5 – 3.6).
Figure 3.5 Mean nodule number on *Cicer arietinum* inoculated with either *M. ciceri* CC1192 or either of two plasmid-cured derivatives, MCC69 or MCC70. Error bars display the standard error of the groups, while different letters represent statistically significant (p < 0.05) differences according to ANOVA.

Figure 3.6 Mean nodule weights of *C. arietinum* plants inoculated either with the commercial inoculant, CC1192, or one of the two plasmid cured derivatives, MCC69 and MCC70. Error bars display the standard errors of groups, while different letters indicate statistically significant differences (p < 0.05) using an ANOVA.
At the time of harvesting, nodules were collected from a representative plant to confirm the identity of nodule occupants in each treatment. Colonies arising from nodule crushes were screened with two chromosomal markers, $\text{attL}$ and $\text{attR}$, as well as the three plasmid-specific primer pairs that bind to three regions around pMc1192. The presence of the $\text{attL}$ and $\text{attR}$ sites in all isolates confirm the presence of CC1192, while the absence of the approximately 500-bp amplified products when screening for the plasmid in plasmid-cured derivatives, and their presence in the wild type CC1192, indicate that there was no cross-contamination in the glasshouse experiment (Appendix figure 1).

### 3.1.5 Nodule sections

Assessment of *Cicer arietinum* nodule sections from inoculation with CC1192 and the two plasmid-cured strains, MCC69 and MCC70, showed that all three treatments resulted in root nodules containing a large number of infected cells with high densities of bacteroids (stained dark purple) (Figure 3.7 A, D and H, respectively). These infected cells were interspersed amongst non-infected cells, as is commonly seen within indeterminate-type nodules. Thus, no difference in the extent of bacteroid invasion and development was evident between nodules produced by wild-type, MCC69 or MCC70 treatments.
Figure 3.7 Nodule sections of *Cicer arietinum* inoculated with CC1192 (A – C), MCC69 (D – F) or MCC70 (G – I) under 4X magnification (A, D, G) or 40X magnification (B, C, E, F, H, I). Examples of an infected cell is demonstrated by the red solid arrows (B, C), with the bacteroids within each cell shown by the purple cellular contents. The opaque border of the arrow in B displays the presence of starch granule accumulation within the infected cells. The asterisks (*) identifies the pronounced zone of senescence in MCC69 and MCC70, and the pink striped arrow (F, I) indicates the region in higher magnification, resulting either from plasmid loss or sectioning.
Examination of the nodule tissue closest to the infection zone under higher magnification showed similar profiles for nodules occupied by CC1192, MCC69 and MCC70, respectively (Figure 3.7 B, E and H). However, the morphology of infected cells proximal to the plant vascular bundles in treatments MCC69 and MCC70 were paler than CC1192, with bacteroid-filled nodule tissue appearing to have separated away from the plant cell wall in the plasmid cured treatments, but not CC1192 (Figure 3.7 C, F and I). Furthermore, observation of these regions of the nodules at low power (Figure 3.7 A, D and G, respectively) showed that this lightly-stained area was much larger in the nodules occupied by either plasmid-cured strain compared with nodules occupied by CC1192. This may potentially indicate a difference in nodule senescence between wild type and plasmid-cured strains.

3.2 Investigating curing ICEMcSym\textsuperscript{1192} from CC1192

3.2.1 Overexpression of recombination directionality factor, rdf\textit{S}

The first approach to cure ICEMcSym\textsuperscript{1192} from the CC1192 genome was to mobilise the pJR204 (Table 2.1) construct constitutively expressing rdf\textit{S} (MESLO_RS0130415) from \textit{M. loti} R7A into CC1192, as this plasmid had successfully cured ICEM\textit{S}\textsuperscript{R7A} from the \textit{M. loti} genome (Ramsay \textit{et al.}, 2006). When pJR204 was introduced into CC1192, a large number of Tc resistant transconjugants grew on selective media. Screening of 20 of these transconjugants by PCR confirmed that they were all positive for the symbiosis ICE-specific \textit{attB} site (excised form of ICE) and \textit{attL} site (integrated form of ICE), indicating that ICEMcSym\textsuperscript{1192} was still present in the genome of all CC1192 transconjugants. Therefore, pJR204 was not capable of inducing ICEMcSym\textsuperscript{1192} loss from CC1192.
Subsequent protein sequence analysis showed that the R7A RdfS in pJR204 was only 79% identical to the CC1192 RdfS (A4R28_RS20865). Therefore, it is possible that the two proteins differ enough such that the R7A RdfS cannot catalyse the excision and subsequent loss of ICEMcSym1192 from the CC1192 genome. Therefore, rdfS from CC1192 was cloned into the EcoRI site of the broad-host range expression vector pPR3 under the control of the constitutive nptII promoter. PCR screening was performed to identify the presence and orientation of rdfS in the construct (Appendix figures 2 and 3). Plasmid DNA was then extracted and the cloned region was Sanger sequenced. A BLASTN alignment of the returned sequence against the CC1192 genome confirmed that the insert was 100% identical to rdfS from CC1192, indicating the complete gene had been cloned (Appendix figure 4). This vector was designated pMCC7.

As overexpression of R7A rdfS led to the loss of ICEMcSymR7A as well as a significantly reduced number of transconjugant colonies, a similar outcome was expected in CC1192, and thus screened for. The vector constitutively expressing CC1192 rdfS, pMCC7, was conjugated into CC1192 alongside the empty expression vector, pPR3, in triplicate and the number of transconjugants quantified. Both vectors had similar conjugation frequencies, with pPR3 producing $2.48 \times 10^{-4}$ transconjugants/donor (SE = $4.6 \times 10^{-5}$) and pMCC7 producing $2.71 \times 10^{-4}$ transconjugants/donor (SE = $1.21 \times 10^{-4}$). A two-sample t-test confirmed that the aforementioned frequencies were not statistically different ($p > 0.05$), demonstrating that overexpression of rdfS in CC1192 is not lethal to cells carrying their symbiosis ICE.

Next, attB and attP sites were screened via PCR to ascertain whether excision was stimulated by introduction of pMCC7. The intensities of the two att PCR products of
the eight colonies screened appeared greater than those of the wild-type CC1192 (Figure 3.8), consistent with RdfS promoting excision of ICE_{McSym}^{1192}. Interestingly, the maintenance of pPR3 appeared to lead to a lower level of excision than seen in CC1192, indicated by fainter bands (Figure 3.8).

![Figure 3.8](image)

**Figure 3.8** Image of an agarose gel electrophoresis of PCR amplification of attB (chromosomal) and attP (circularised excised ICE) attachment sites of CC1192 harbouring pMCC7 and grown in broths to stationary phase. **Ladder:** 100-bp NEB ladder; **1-8:** Colonies screened; attB: Screening for attB site (formed in the chromosome upon ICE excision) with primers 7 and 8 for a 520-bp product; attP: Screening of the attP site (formed on ICE when excised and circularised) with primers 9 and 10 for an expected 493-bp fragment; **CC1192(pPR3):** CC1192 harbouring the empty pPR3 vector; **CC1192:** *Mesorhizobium ciceri* CC1192; **NTC:** No template control.
The \textit{attL} and \textit{attR} sites of the same eight transconjugants were also screened via PCR to confirm the excision of the ICE. The intensity of PCR products in all transconjugants screened was comparable to the CC1192 control, indicating that all colonies harboured an integrated ICE\textit{McSym}^{1192} (Figure 3.9). While the increased intensity of \textit{attB} and \textit{attP} products in CC1192 transconjugants relative to CC1192 indicates that RdS stimulates excision of ICE\textit{McSym}^{1192}, the positive PCR products for \textit{attL} and \textit{attR} show that the ICE has not been cured by this method.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3_9.png}
\caption*{Figure 3.9 Agarose gel electrophoresis of PCR of \textit{attL} and \textit{attR} sites demarcating the ICE\textit{McSym}^{1192} harbouring pMCC7, from stationary phase broth. \textbf{Ladder:} 100-bp NEB ladder; 1-8: Colonies screened; \textbf{attL}: Screening for the \textit{attL} site (formed on the left border of ICE\textit{McSym}^{1192} and CC1192 chromosome) with primers 7 and 10 to amplify a 501-bp product; \textbf{attR}: \textit{attR} site (formed on the right border of ICE\textit{McSym}^{1192} and CC1192 chromosome) amplified with primers 8 and 9 for an expected 510-bp fragment; \textbf{CC1192(pPR3)}: CC1192 harbouring empty pPR3 vector; \textbf{CC1192: Mesorhizobium ciceri} CC1192; \textbf{NTC:} No template control.}
\end{figure}
Previous work with R7A showed that the loss of ICEMI/SymR7A transformed the strain into a vitamin auxotroph for biotin, nicotinate and thiamine, as the genes required for their synthesis lie on ICEMI/SymR7A (Sullivan et al., 2002; Ramsay et al., 2006). Bioinformatic analysis of CC1192 was undertaken to determine whether the genes required for the synthesis of the aforementioned vitamins are similarly located on ICEMcSym1192. The full complement of biotin (bioBFDAZ), nicotinate (nadABC) and thiamine (thiCOSGED) operons were located on ICEMcSym1192, as identified by BLASTP comparison to the operons in R7A (Table 3.3). As no other copies of these biosynthetic genes were detected in the CC1192 genome, the loss of ICEMcSym1192 was expected to lead to biotin, nicotinate and thiamine auxotrophy. This characteristic could therefore be used as a screening tool to search for ICEMcSym1192 cured CC1192 transconjugants.
Table 3.3 Comparison of nicotinate, biotin and thiamine biosynthetic operons encoded in \textit{M. loti} R7A and \textit{M. ciceri} CC1192 genomes.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>R7A</th>
<th>CC1192</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quinolinate synthetase</td>
<td>nadA</td>
<td>A4R28_RS20955</td>
</tr>
<tr>
<td>L-aspartate Oxidase</td>
<td>nadB</td>
<td>A4R28_RS20960</td>
</tr>
<tr>
<td>Nicotinate-nucleotide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrophosphorylase</td>
<td>nadC</td>
<td>A4R28_RS20965</td>
</tr>
<tr>
<td>Biotin synthase</td>
<td>bioB</td>
<td>A4R28_RS20970</td>
</tr>
<tr>
<td>8-amino-7-oxononanooate</td>
<td>bioF</td>
<td>A4R28_RS20975</td>
</tr>
<tr>
<td>synthase</td>
<td>bioD</td>
<td>A4R28_RS20980</td>
</tr>
<tr>
<td>Dethiobiotin synthase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenosylmethionine-8-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>amino-7-oxononanooate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aminotransferase apoenzyme</td>
<td>bioA</td>
<td>A4R28_RS20985</td>
</tr>
<tr>
<td>Biotin biosynthesis protein</td>
<td>bioZ</td>
<td>A4R28_RS20990</td>
</tr>
<tr>
<td>hydroxymethylpyrimidine synthase</td>
<td>thiC</td>
<td>A4R28_RS22280</td>
</tr>
<tr>
<td>Glycine Oxidase</td>
<td>thiO</td>
<td>A4R28_RS22285</td>
</tr>
<tr>
<td>Sulfur carrier protein</td>
<td>thiS</td>
<td>A4R28_RS22290</td>
</tr>
<tr>
<td>Thiazole-phosphate synthase</td>
<td>thiG</td>
<td>A4R28_RS22295</td>
</tr>
<tr>
<td>Thiamine-phosphate</td>
<td>thiE</td>
<td>A4R28_RS22300</td>
</tr>
<tr>
<td>pyrophosphorylase</td>
<td>thiD</td>
<td>A4R28_RS22305</td>
</tr>
<tr>
<td>hydroxymethylpyrimidine kinase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† All orthologs identified in CC1192 were encoded within ICE\textit{McSym}^{1192}. 

Locus ID | Length (aa) | Identity/coverage |
R7A       |             |                  |
MesloDRAFT_5957 | 324 | 259/322 (80%)   |
MesloDRAFT_5956 | 513 | 353/512 (69%)   |
MesloDRAFT_5955 | 293 | 223/293 (76%)   |
MesloDRAFT_5953 | 331 | 264/330 (80%)   |
MesloDRAFT_5952 | 379 | 336/378 (89%)   |
MesloDRAFT_5951 | 211 | 175/211 (83%)   |
MesloDRAFT_5950 | 421 | 383/421 (91%)   |
MesloDRAFT_5949 | 327 | 289/327 (88%)   |
MesloDRAFT_5928 | 611 | 576/611 (94%)   |
MesloDRAFT_5927 | 333 | 285/333 (86%)   |
MesloDRAFT_5926 | 85  | 58/65 (89%)     |
MesloDRAFT_5925 | 257 | 236/257 (92%)   |
MesloDRAFT_5924 | 201 | 174/201 (87%)   |
MesloDRAFT_5923 | 256 | 146/245 (60%)   |
A4R28_RS20955  | 324 |                  |
A4R28_RS20960  | 513 |                  |
A4R28_RS20965  | 293 |                  |
A4R28_RS20970  | 336 |                  |
A4R28_RS20975  | 379 |                  |
A4R28_RS20980  | 211 |                  |
A4R28_RS20985  | 421 |                  |
A4R28_RS20990  | 327 |                  |
A4R28_RS22280  | 611 |                  |
A4R28_RS22285  | 333 |                  |
A4R28_RS22290  | 65  |                  |
A4R28_RS22295  | 257 |                  |
A4R28_RS22300  | 201 |                  |
A4R28_RS22305  | 256 |                  |

54
Replica patching of 1,044 CC1192 colonies harbouring the rdfS expression vector (pMCC7), on media supplemented with Nm and either with or without the aforementioned vitamins, was performed to determine whether loss of ICEMcSym\textsuperscript{1192} from rdfS overexpression was substantially less-frequent than for ICEMISym\textsuperscript{R7A}. All colonies patched were able to grow on vitamin deficient media, consistent with ICEMcSym\textsuperscript{1192} not being lost in CC1192 strains over-expressing rdfS. Therefore, although overexpression of rdfS appears to increase the excision rate of ICEMcSym\textsuperscript{1192} from the CC1192 chromosome, this does not readily lead to loss of the element from the cell.

### 3.2.2 Curing ICEMcSym\textsuperscript{1192} by site-directed deletion

Another approach to curing ICEMcSym\textsuperscript{1192} from CC1192 is through a site-directed deletion using homologous recombination to replace the entire element with an antibiotic marker. This approach requires regions flanking the ICE to be cloned into an appropriate suicide vector, a vector that is not maintained within the cell and is lost unless integration into the host genome occurs by homologous recombination. The construct is then mobilised into the target strain. The desired mutation is selected for in a two-step process involving a (1) single crossover, whereby one homologous region recombines and integrates the entire vector into the chromosome, and then a (2) double-crossover, where the second homologous region recombines, excising the vector backbone and the ICE, leaving only the antibiotic cassette. However, the action of the integrase (IntS) and RdfS could catalyse excision of the complete ICE in between steps 1 and 2, or the re-integration of the ICE before the element was lost after the double-cross over reaction. This increases the screening complexity and could
result in a final strain that still maintains a completely intact ICE (See Figure 1.3). Therefore, to cure CC1192 of its ICE in this fashion, the element must first be locked into the chromosome or stabilised so that excision and integration no longer occur. One approach to achieve this is to delete the gene encoding the ICE integrase (intS).

To stabilise ICE\textsubscript{McSym}\textsuperscript{1192}, an antibiotic replacement deletion construct for the integrase (\textit{intS}, A4R28\_RS20660) was created containing the upstream and downstream regions of the gene and the \textit{Ω-aadA} cassette within pEX18\textsubscript{Tc}. Colony PCR and restriction digestion confirmed the fragments had ligated correctly (Appendix figure 5 and 6). This construct was named pMCC8.

This \textit{intS} deletion construct was introduced into CC1192 and a very large number of colonies arose on selective media, indicating that the vector was not suiciding in CC1192, but rather was maintained as a replicative plasmid. To determine whether this was the case, a biparental conjugation was performed between CC1192 and broad-host range vector pPR3 as a positive control, as well as potential suicide plasmids pEX18\textsubscript{Tc} and pJQ200SK. On selective media, CC1192(pPR3) and CC1192(pEX18\textsubscript{Tc}) matings both yielded large numbers of transconjugants, while no CC1192(pJQ200SK) transconjugants were observed. Therefore, in CC1192, pEX18\textsubscript{Tc} is able to replicate autonomously and is not a suitable vector for site-directed mutagenesis. However, pJQ200SK did appear to be lost in CC1192 and therefore was a viable alternative to pEX18\textsubscript{Tc}. The \textit{intS} deletion construct to stabilise the ICE\textsubscript{McSym}\textsuperscript{1192} was then remade with the pJQ200SK backbone. Colony PCR and restriction digestion showed that the regions had been cloned into the vector in the correct orientation (Appendix figures 7 and 8). This construct was subsequently named pMCC9.
The integrase deletion construct, pMCC9, was introduced into CC1192 and a single-crossover was first selected for, before selection for a double crossover, whereby the integrase would be replaced with the Ω-aadA cassette. Screening for loss of the plasmid backbone was achieved through the utilization of sacB (encoding sucrose sensitivity) and aacCI (encoding gentamycin resistance) on pJQ200SK. Replica plating of 75 colonies confirmed their Gm sensitivity, and their resistance to Sp, Sm and sucrose, consistent with the replacement of intS by Ω-aadA and the loss of the pJQ200SK backbone in a double cross-over.

Three colonies were screened via PCR to confirm the integration and location of the cassette, with all lanes containing the expected fragment of 1,470 bp (Figure 3.10).

Figure 3.10 Image of electrophoresed PCR products using primers 23 and 24 to screen for the replacement of the integrase (intS) with the Ω-aadA cassette. Replacement of the integrase would generate a 1,470-bp product using primers 23 and 24, whilst the intact integrase would produce no product. 1 – 3: Three independent potential CC1192 intS:Ω-aadA. CC1192: M. ciceri CC1192. NTC: No template control. Ladder: 1-kb NEB ladder.
Colony 1 was chosen for Sanger sequencing and pairwise alignment of the sequenced products and the predicted product confirmed that the Ω-aadA cassette had inserted and replaced intS (Appendix figure 9). This strain became referred to as MCC86.

To determine whether the integrase deletion stabilised ICEMcSym\(^{1192}\), PCR screening for the presence of attB (chromosomal attachment site) and attP (attachment site on circularised excised ICE) sites that form upon ICE excision, as well as the attR site demarcating the right boarder of the ICE and chromosome, was performed. The presence of attB and attP products amplified from CC1192 and the absence of products from the three CC1192 intS::Ω-aadA samples (Figure 3.11) is consistent with the deletion of intS supressing ICE excision, and stabilisation of this element in the genome of CC1192.

**Figure 3.11** Image of agarose gel electrophoresis of PCR products from three colonies screened for intS::Ω-aadA. Colonies were screened for the presence of attB and attP, indicative of ICE excision, as well as the attR site demarcating the border of ICEMcSym\(^{1192}\) and the CC1192 chromosome. **attB**: 520 bp product formed excised ICE; **attP**: 493 bp product formed on chromosome after ICE excision; **attR**: 510 bp product over integrated ICE junction; **1 – 3**: Colonies shown to have intS::Ω-aadA; **CC1192**: *Mesorhizobium ciceri* CC1192; **NTC**: No template control; **Ladder**: 100 bp NEB ladder consisting of the following fragment sizes (bp): 1517, 1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100.
3.2.3 Construction of ICE deletion vector

To next delete the ICEMcSym$^{1192}$, a 1.5-kb upstream and downstream region bordering ICEMcSym$^{1192}$ was cloned adjacent to an antibiotic resistance gene, \textit{nptII}, and ligated into the pJQ200SK backbone. PCR screening produced the expected 4,433-bp product in over half of the colonies screened (Appendix figure 10), while restriction digestion produced the expected fragment sizes in three out of four of these colonies (Appendix figure 11). This ICE deletion construct was named pMCC10, for future introduction into CC1192 \textit{intS:: Ω-aadA} (MCC86).
Chapter 4: Discussion

This aims of this thesis was to characterise the symbiotic role of the plasmid, pMc1192, in *M. ciceri* symbiosis with *C. arietinum*, by firstly creating a plasmid-cured derivative, and then secondly to compare the symbiotic phenotype of plasmid-cured CC1192 to that of wild type CC1192. And finally, to investigate curing CC1192 of its symbiosis integrative and conjugative element (ICE), referred to as ICEMcSym1192.

4.1 The symbiotic role of pMc1192

Bioinformatic analysis of pMc1192 revealed it harboured fixNOQP and fixGHIS, two operons shown to be essential for nitrogen fixation in *Bradyrhizobium janponicum*, which forms a symbiotic association with *Glycine max* (soybean) (Preisig *et al.*, 1996b; Koch *et al.*, 2000). FixNOQP is a cbb3-type cytochrome oxidase that functions in the low O$_2$-tension environment in legume root nodules, while FixGHIS is a Cu$^+$-uptake system that supports the activity of FixNOQP (Preisig *et al.*, 1996b; Koch *et al.*, 2000). Both fixNOQP and fixGHI are also encoded on ICEMcSym1192, but interestingly pMc1192 appears to harbour the sole copy of fixS in the CC1192 genome. Symbiotic phenotyping of two independent plasmid-cured CC1192 derivatives, MCC69 and MCC70, compared with CC1192 showed that there was no significant difference in mean plant shoot dry weights, nodule number or nodule mass when the three strains were inoculated separately on *C. arietinum*. This indicates that the fix genes found on the plasmid are not essential for symbiotic nitrogen fixation on this legume host. However, nodule sections of plasmid-cured strains did appear to possess a zone of senescence proximal to the vascular bundle much larger than that of wild-type CC1192. This could indicate that plasmid encoded genes may play some role in
maintaining infected root nodule cells, although this observation could equally be explained as an artefact of the nodule fixing and sectioning process. As only very few nodules were assessed, more sections of nodules from plasmid-cured and wild-type strains need to be examined before conclusions on the possible role of plasmid-encoded genes in senescence may be drawn.

Previous studies have shown that deletion of fixNOQP or fixGHI in Bradyrhizobium japonicum 110spc4 led to a Fix− phenotype with G. max (i.e. strains were incapable of fixing nitrogen) (Preisig et al., 1996a; 1996b), indicating their critical role in nitrogen fixation. In the Preisig et al. (1996a) mutagenesis study, fixS was excluded, so its significance to symbiotic nitrogen fixation was not determined in B. japonicum. However, deletion of ccoS (an ortholog of fixS) in the photosynthetic bacterium Rhodobacter capsulatus blocked electron transfer in photosynthetic conditions, owing to cofactors heme b, heme b3 and Cu⁺ not being incorporated into CcoN (ortholog of FixN) (Koch et al., 2000). The lack of a phenotype in CC1192 with the loss of the plasmid-encoded fixS suggests that either ICEMcSym1192-encoded FixNOQP does not require FixS or that another, as yet identified, gene in CC1192 may encode a protein that provides the required activity.

Also present on pMc1192 is fixLJ, encoding a low O₂-responsive, two-component regulatory system, and fixK, encoding a transcriptional activator in S. meliloti 1021 (Batut et al., 1989; Tsoy et al., 2016). This regulatory system is known to activate genes required for microaerobic respiration (including fixNOQP) in response to low O₂ (Batut et al., 1989; Tsoy et al., 2016) and mutations in any of these genes has previously been shown to lead to a Fix− phenotype in S. meliloti inoculated onto Medicago sativa (David et al., 1988; Batut et al., 1989). Although copies of fixLJ and fixK are present on the CC1192 chromosome, they were dissimilar at the protein level.
to those on the plasmid (≤ 34% identity). Despite the low level of homology, it is possible that the chromosomal copies of *fixLJ* and *fixK* may function in place of the plasmid versions. Alternatively, the FNR system, a second mechanism of O₂ sensing and transcriptional regulation present in α-proteobacteria, consisting of a single protein (Fnr) that acts both as an O₂-sensor and transcriptional activator, may be present in CC1192 (Cosseau & Batut, 2004). Further work investigating the role of the chromosomally encoded putative O₂-sensor and transcriptional regulator system is needed to determine whether CC1192 encodes a Fnr system, which could function in place of, or in conjunction with, the chromosomal *fixLJ* and *fixK* system. Nevertheless, the glasshouse results in this study show that the loss of *fixLJ* and *fixK* on pMc1192 has no effect on the symbiotic ability of the strain.

The experiments with plasmid cured-CC1192 also indicate that none of the 613 putative protein encoding genes on this replicon are essential to symbiotic nitrogen fixation with *C. arietinum*, nor are they essential for free-living growth in standard laboratory culturing conditions. It is therefore interesting to speculate what role, if any, pMc1192-encoded genes could have in CC1192. In *Rhizobium leguminosarum* bv. *phaseoli* CFN41, which forms a symbiosis with *Phaseolus vulgaris* (common bean) and harbours six plasmids (pA to pF, with pD known to be the pSym), removal of accessory plasmid pC, pF, pB or part of pE resulted in a reduced ability to nodulate *P. vulgaris* when co-inoculated with the wild-type strain, while removal of pA actually increased the strain’s competitiveness (Brom *et al.*, 1992). Plasmid-cured strains have also been noted to be able to utilise less carbon sources than their wild-type parent strain (Baldani *et al.*, 1992; Brom *et al.*, 2000). Therefore, although pMc1192 had no observable impact on the effectiveness of symbiotic nitrogen fixation on *C. arietinum* or in free-living growth conditions, the plasmid may provide some competitive
advantage in nodulation or endow the strain with the ability to metabolise a range of different metabolites.

4.2 Curing of ICEMcSym\textsuperscript{1192}

4.2.1 Overexpression of rdfS does not lead to ICEMcSym\textsuperscript{1192} loss

Overexpression of rdfS did not result in the loss of ICEMcSym\textsuperscript{1192} from the genome of CC1192, with either the native (i.e. CC1192-derived) or heterologous (i.e. R7A-derived) version of the gene, as patching of transconjugants onto vitamin-deficient media did not yield vitamin auxotrophs. However, rdfS over-expression did lead to increased product intensity in PCR screening, consistent with an apparent increase in the proportion of cells in the sample population harbouring excised versions of the ICE. Therefore, while rdfS over-expression does not result in ICE-cured derivatives of CC1192, RdfS does appear to be important in catalysing excision of ICEMcSym\textsuperscript{1192}, which agrees with its putative role as a recombination directionality factor (Ramsay \textit{et al.}, 2006; Wozniak & Waldor, 2010; Haskett \textit{et al.}, 2016c).

In contrast to CC1192, overexpression of rdfS in \textit{M. loti} R7A causes the loss of ICEMlSym\textsuperscript{R7A}, producing the non-symbiotic derivative, R7ANS (Ramsay \textit{et al.}, 2006). The introduction of a plasmid constitutively expressing rdfS in R7A resulted in a small number of slow growing colonies on selective media that had lost ICEMlSym\textsuperscript{R7A} (Ramsay \textit{et al.}, 2006). In contrast, when a similar construct was introduced into CC1192, the same number of transconjugants/donor grew on selective media as the empty vector control, indicating that rdfS did not adversely impact cell growth. In R7A, RdfS induces ICE excision by binding within the \textit{attL} site, where the ICE is integrated in the chromosome (C. Verdonk, unpublished data). This binding
causes the DNA to bend, thus allowing excision to occur (Ramsay et al., 2006). ICE\textsubscript{Ml}Sym\textsuperscript{R7A} integrates within the sole phe-tRNA gene in the R7A genome, whereas ICE\textsubscript{Mc}Sym\textsuperscript{1192} integrates into one of five ser-tRNAs in the genome of this strain (Sullivan & Ronson, 1998; Haskett et al., 2016c). It is possible that overexpression and binding of RdfS within the \textit{attL} site in R7A blocks subsequent transcription of the single phe-tRNA, disrupting protein synthesis and making over-expression of \textit{rdfS} lethal. ICE-cured transconjugants would therefore presumably be those that had naturally lost ICE\textsubscript{Ml}Sym\textsuperscript{R7A}. As only one of the five ser-tRNA genes contain the integrated ICE\textsubscript{Mc}Sym\textsuperscript{1192}, overexpression of \textit{rdfS} and binding to the \textit{attL} site would not abolish ser-tRNA translation, therefore having no lethal consequences for the cell. In this case, there would be no selection for cells that had lost ICE\textsubscript{Mc}Sym\textsuperscript{1192}, as what may have occurred in R7A. Another possibility is that ICE\textsubscript{Mc}Sym\textsuperscript{1192} may harbour stability factors (such a toxin-antitoxin system), absent from ICE\textsubscript{Ml}Sym\textsuperscript{R7A}, that could prevent its loss from the cell following overexpression of \textit{rdfS}. Further analysis of the genetic complement of ICE\textsubscript{Mc}Sym\textsuperscript{1192} is required to determine if this is the case.

From the end-point PCR performed in this thesis, it is difficult to determine the degree of increased excision in CC1192 transconjugants overexpressing \textit{rdfS}. To quantify the increase in excision rate, it would be valuable to employ a qPCR-based assay, similar to that used to quantify the rates of excision of ICE\textsubscript{Ml}Sym\textsuperscript{R7A} and ICE\textsubscript{Mc}Sym\textsuperscript{1271} (Ramsay et al., 2006; Haskett et al., 2018).

Although the rate of ICE\textsubscript{Mc}Sym\textsuperscript{1192} excision increased upon overexpression of \textit{rdfS}, the loss of the element was not detected. The large number of colonies patched for loss of the ICE indicates that this event is less frequent than in R7A. As ICE\textsubscript{Mc}Sym\textsuperscript{1192} was unable to be cured through the overexpression of \textit{rdfS}, another approach was taken.
4.2.2 Stabilisation of ICEMcSym\textsuperscript{1192} through the deletion of \textit{intS}

An alternative approach to curing ICEMcSym\textsuperscript{1192} is to completely delete the element by homologous recombination. Owing to the ability of the ICE to exist in both integrated and excised forms, ICEMcSym\textsuperscript{1192} must first be stabilised into the host chromosome by deleting the integrase (\textit{intS}). The initial attempt to delete the integrase in CC1192 with a pEX18Tc-based construct was not successful as the vector was stably maintained in CC1192. The ability of a plasmid to replicate within a host genome depends on its origin of replication (\textit{ori}) and whether the host cellular machinery recognises this site (Snyder \textit{et al.}, 2013). To use a plasmid as an inactivation vector, the plasmid must be incapable of replication within the genome of interest, as to allow for selection of plasmid integration into the genome (Quandt & Hynes, 1993; Snyder \textit{et al.}, 2013).

Plasmid pEX18Tc carries a ColE1 origin of replication, which is generally considered to be a narrow host range vector (Hoang \textit{et al.}, 1998; Snyder \textit{et al.}, 2013). ColE1-based vectors have been previously used for site-directed mutagenesis in other \textit{Mesorhizobium} spp.; \textit{M. ciceri} bv. \textit{biserrulae} WSM1271 (Haskett \textit{et al.}, 2018), \textit{M. loti} R7A (Ramsay \textit{et al.}, 2006) and \textit{Mesorhizobium} sp. WSM3873 (A. Asrat Bekuma, unpublished results). However, from the results of this thesis, the ColE1-based pEX18Tc vector replicated in CC1192, indicating the vector is not an appropriate tool for site-directed mutagenesis in this strain. Deletion of \textit{intS} in CC1192 was subsequently achieved using pJQ200SK, with a p15a-based origin of replication (Quandt & Hynes, 1993). This vector was therefore shown to be a viable alternative for future mutagenesis work in CC1192.
Screening of *intS* deletion mutants of CC1192 by end-point PCR showed the presence of the integrated form of ICEM*Mc*Sym\textsuperscript{1192}, but not the excised forms. Based on work in R7A and other bacteria, ICE integrases catalyse both excision and integration of these elements (Ramsay *et al.*, 2006; Murphy & Boyd, 2008; Bañuelos-Vazquez *et al.*, 2017). Therefore, this data is consistent with IntS being the sole integrase in CC1192 capable of catalysing excision of the element. However, further screening of *intS* mutants generated in this study should be performed to confirm the complete stabilisation of the ICE. This could be achieved by measuring ICE transfer rates in the integrase-deleted and wild-type strains donating their ICE to R7ANS. Alternatively, quantification of the proportion of *attB/attP* (reformed chromosomal and excised ICE attachment sites, respectively) by qPCR would provide additional evidence for the stability of ICEM*Mc*Sym\textsuperscript{1192} in the chromosome of CC1192.

### 4.2.3 Subsequent deletion of ICEM*Mc*Sym\textsuperscript{1192}

Should ICEM*Mc*Sym\textsuperscript{1192} prove to be stabilised in CC1192 *intS::Ω-aadA* (MCC86), the ICE deletion vector pMCC10 successfully produced in this thesis can be used to attempt to delete the entire ICE in a single step. The likelihood of success for this approach is dependent upon (1) the presence of core genes on the ICE and (2) the feasibility of deleting a 419-kb region of DNA.

ICEM*Mc*Sym\textsuperscript{1192} harbours 350 putative protein-encoding genes, of which it is unknown whether any are core genes, encoding products essential for cellular function (i.e. for growth in standard laboratory conditions) (Harrison *et al.*, 2010; Haskett *et al.*, 2016c; Young, 2016). The presence of unique core genes on ICEM*Mc*Sym\textsuperscript{1192} could inhibit the recovery of ICE-cured derivatives. Although deletion of a symbiosis ICE has not been
previously reported, deletion of large symbiosis-encoding replicons has been achieved in *Sinorhizobium meliloti* Rm1021 (diCenzo et al., 2013). The genome of Rm1021 consists of a chromosome (3.7-Mb) plus two large megaplasmids (pSymA, 1.4-Mb; pSymB, 1.7-Mb). pSymA was cured relatively easily by sequential transposon mutagenesis (Oresnik et al., 2000), whereas pSymB could not be cured by this means. A series of deletions in pSymB coupled with complementation identified two genes (a unique copy of tRNA\(^{\text{arg}}\) and engA encoding a GTPase likely involved in ribosome biogenesis) essential for growth under standard conditions on pSymB (diCenzo et al., 2013). In addition, an active toxin-antitoxin system was identified in pSymB and deleted, that also presumably hindered earlier attempts at plasmid curing (Milunovic et al., 2014). The presence of either an active toxin-antitoxin system or essential genes on ICE\(\text{McSym}\)\(^{1192}\) can inhibit the recovery of ICE-cured derivatives (Burru, 2017). However, it is not currently known whether these exist on ICE\(\text{McSym}\)\(^{1192}\).

The ability to delete such a large region in a single step is a second factor that will determine the success of the ICE\(\text{McSym}\)\(^{1192}\) deletion. Previously, a 430-kb region in the megaplasmid of *Sinorhizobium meliloti* SU47, pRmeSU47b, was successfully deleted using homologous recombination between 1.5-kb Tn5-derivative insertions (Charles & Finan, 1991). Although deletion of a symbiosis ICE has not been reported yet in rhizobia, deletion of ICEs has been reported in other genera of bacteria. For example, the 40-kb cag-ICE of *Helicobacter pylori*, encoding genes for a type IV protein secretion system for translocating the CagA protein into human epithelial cells, was deleted through homologous recombination (Odenbreit et al., 2001). Similarly, the 89-kb pathogenicity island of *Streptococcus suis*, a major swine pathogen, was deleted from the genome after the inactivation of an active toxin-antitoxin system on the ICE using the cre-loxP system (Yao et al., 2015). The cre-loxP system relies on
the introduction of two 34-bp loxP sites flanking the region to be deleted before introduction of a plasmid-borne cre recombinase that catalyses the site-specific recombination reaction (Suzuki et al., 2008). Although effective at deleting large regions of DNA (diCenzo et al., 2013), this system does leave a small 34-bp scar after the recombination reaction, corresponding to a single lox site (Suzuki et al., 2008; Yao et al., 2015). Therefore, double crossover homologous recombination is more desirable in this instance.

**Conclusions and Future Direction**

The removal of pMc1192 from the CC1192 genome has shown that the plasmid harbours no genes essential to nitrogen fixation or free-living growth in standard laboratory media. A similar result was obtained when pMESCI01 was cured from the *Biserrula pelecinus*-nodulating strain *M. ciceri* bv. *biserrulae* WSM1271 (R. Brewer, unpublished results). Two other *M. ciceri* strains have now been completely sequenced (WSM1497, WSM1284) and these also harbour similar sized plasmids (pWSM1497, 532-kb; pMc1284, 554-kb, respectively) (Haskett et al., 2016b; Brewer et al., 2017). Given that these four strains were isolated from a range of geographical locations (Sardinia, Mykonos and Israel) it appears that these plasmids may be a common feature of *M. ciceri* strains. It would therefore be informative to investigate if these plasmids impart any competitive or metabolic benefit to the host strains that might explain their presence in the genomes of these organisms.

The experiments with CC1192 rdfS (recombination directionality factor) and intS (integrase) suggest that the mechanism of action for ICE excision and integration are similar to that of R7A. The deletion of the integrase appeared to stabilise the element,
while overexpression of rdfS led to an increased number of cells with excised products. These results are both consistent with the R7A model, whereby the integrase catalyses the site-specific recombination reactions, with the presence of rdfS stimulating the excision reaction (Ramsay et al., 2006). However, the overexpression of rdfS did not stimulate the loss of ICE McSym1192, as it did in R7A (Ramsay et al., 2006), suggesting a difference in the mechanism of action. Further work on the mechanism of action of RdS proteins are required to determine the molecular basis of these differences.

Finally, symbiosis ICE transfer in the field often results in the evolution of novel strains that are suboptimal for nitrogen fixation (Nandasena et al., 2007a). However, we currently lack the tools to determine the genetic basis of this phenomenon. The creation of a plasmid-cured CC1192 derivative, as well as the putatively ICE-stabilised strain, both constitute important steps towards a *Mesorhizobium ciceri* CC1192 naïve strain devoid of its mobile genetic elements. This strain will be an invaluable tool that can be paired with *M. loti* R7ANS (Ramsay et al., 2006) to better understand how symbiotic performance of novel strains is affected by the interaction between chromosomal and ICE-encoded genes.
References


Nandasena, K., G. O'hara, R. Tiwari, E. Sezmiş & J. Howieson. (2007a) *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**: 2496-2511.


Preisig, O., R. Zufferey & H. Hennecke. (1996a) The *Bradyrhizobium japonicum* *fixGHIS* genes are required for the formation of the high-affinity *cbb3*-type cytochrome oxidase. *Archives of Microbiology* **165**: 297-305.


Segerman, B. (2012) The genetic integrity of bacterial species: the core genome and
the accessory genome, two different stories. *Frontiers in Cellular and Infection
Microbiology* **2**: 1-8.

Shamseldin, A., A. Abdelkhalek & M. Sadowsky. (2017) Recent changes to the
classification of symbiotic, nitrogen-fixing, legume-associating bacteria: a

Gammaproteobacteria *Pseudomonas* sp. and the Betaproteobacteria
*Burkholderia* sp. *Systematic and Applied Microbiology* **33**: 269-274.

Snyder, L., J. Peters, T. Henkin & W. Champness. (2013) *Molecular genetics of
bacteria*. American Society for Microbiology.

transfer flavoprotein *fixABCX* gene products from *Azospirillum brasilense*
show a NifA-dependent promoter regulation. *Current Microbiology* **49**: 267-
273.


symbiosis island that integrates into a *phe*-tRNA gene. *Proceedings of the
National Academy of Sciences* **95**: 5145-5149.

Sullivan, J., J. Trzebiatowski, R. Cruickshank, J. Gouzy, S. Brown, R. Elliot, D.
sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *Journal of Bacteriology* **184**: 3086-3095.


Young, J. (2016) Bacteria are smartphones and mobile genes are apps. *Trends in Microbiology* **24**: 931-932.

Appendix Figures

Appendix figure 1 Image of electrophoresed DNA from colony PCR of bacteria isolated from nodule crushes of MCC69 and MCC70 for strain confirmation. Products within a, b, and c indicate the presence of a plasmid whilst products in attL and attR confirm the strain as CC1192; **Ladder**: 100 bp NEB ladder; **attL**: attL site of CC1192 amplified with primers 7 and 10 for a fragment size of 501 bp; **attR**: attR site of CC1192 amplified by primers 8 and 9 to produce a 510 bp product; **a, b & c**: three primer pairs that amplify around pMc1192 (primers 1/2, 3/4 or 5/6) with fragment sizes of 500-bp, 525-bp and 507-bp, respectively; **MCC69/MCC70**: plasmid cured derivatives of CC1192; **CC1192**: *Mesorhizobium ciceri* CC1192; **NTC**: No template Control.
Appendix figure 2 Image of electrophoresed PCR product amplified using primers 13 and 14 for the presence of CC1192 *rdfS* cloned into pPR3. A 472-bp product indicates that the insert had integrated into the MCS which was seen in colonies 1, 2-6 and 8-10, whilst a 185-bp product indicated no insert had integrated. **Ladder:** 100-bp Promega ladder; 1-18: Colonies screened; 19: *E. coli* ST18 harbouring pPR3; **NTC:** No template control.
Appendix figure 3 Image of agarose gel containing electrophoresed products from an orientation check of CC1192 *rdfS* cloned into pPR3 using primers 12 and 13. The 395-bp product in lanes 1, 2, 5 and 7 indicates the sought orientation. **Ladder:** 100-bp Promega ladder; **1-8:** Colonies screened; **9:** pPR3 without an insert; **NTC:** No template control.

Appendix figure 4 Sequence alignment between pMCC7 (sequence 1) and sequenced *rdfS* product (sequence 2). Consensus Identity (green) shows 100% alignment between the two sequences. Note, the start of the *rdfS* product was trimmed for quality.
Appendix figure 5 Image of electrophoresed PCR products of the quadruple ligation to create the intS deletion construct containing the upstream and downstream regions immediately adjacent to the integrase gene within CC1192, an \(\Omega\)-aadA interposon cassette and the backbone of pEX18Tc. Primers 19 and 20 were used to amplify a 4,618-bp fragment if all components ligated together. **Ladder:** 1-kb NEB ladder; **1-28:** Colonies that grew on plates selecting for tetracycline and spectinomycin resistance; **NTC:** No template control.
Appendix figure 6 Image of the electrophoresed fragments from a restriction digestion with KpnI and PstI of potential integrase deletion constructs. The expected fragment sizes, seen in all cut colonies, were 1,138-bp, 3,280-bp and 6,322-bp. Adjacent lanes contained the purified plasmid untreated. Ladder: 1-kb NEB ladder.
Appendix figure 7 Image of an electrophoresed gel agarose of PCR products from the region spanning the BamHI and SalI restriction sites where the upstream region, Ω-\textit{aadA} cassette, and downstream region ligated. The expected product was 4,573-bp. The clean NTC indicates that the smaller products obtained were the result of off-target binding. NTC: No template control; Ladder: 1-kb NEB ladder.
Appendix figure 8 Image of an electrophoresed gel containing fragments of a restriction digestion of eight colonies that appeared to have successfully ligated the components for the integrase deletion construct; composed of the pJQ200SK backbone, the upstream region, the Ω-aadA cassette and the downstream region. The expected product sizes after digestion with KpnI and NcoI were 3,379-bp, 2,967-bp, 2,241-bp and 1,157-bp. The untreated construct was run directly before the digested construct. **Ladder:** 1-kb NEB ladder.

Appendix figure 9 Sequence alignment of intS replacement DXO in the CC1192 chromosome (top sequence, highlighted yellow) and the forward and reverse sequences using primers 23 and 24 binding to the integrase deletion construct, pMCC9, and the CC1192 chromosome, respectively (bottom two sequences, highlighted purple). The consensus identity (green, top) shows the 99% similarity between the three sequences. Note, the sequenced regions were trimmed for quality control.
Appendix figure 10 Image of an electrophoresed gel containing PCR products screening for the quadruple ligation to create the ICE deletion construct consisting of a pJQ200SK backbone and a nptII gene flanked with the 1.5-kb upstream and downstream regions of the ICEMcSym1192. Primers 21 and 22 were used to amplify the DNA ligated within the pJQ200SK backbone. A 4,433-bp product would be expected if the regions ligated. Ladder: 1-kb NEB ladder. 1 – 25: Colonies arising on selective media after the transformation into DH5α. NTC: No template control.
Appendix figure 11 Image of an agarose gel electrophoresis of the restriction digestion with XhoI of the ICE deletion construct harbouring the approximate 1.5-kb regions upstream and downstream of ICEMcSym1192 flanking the nptII cassette within the pJQ200SK backbone. The expected fragment sizes were 1,540-bp, 2,636-bp and 5,412-bp, which was seen in constructs 6, 8 and 13, indicating successful ligation of all four fragments. Ladder: 1-kb NEB ladder.