Curtin Medical School

Molecular and Genetic Characterisation of Bacterial Cell-Cell Signalling

Tahlia Bastholm

0000-0002-3729-2356

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University

December 2023

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Tahlia Bastholm
15/12/2023

Acknowledgement of Country

We acknowledge that Curtin University works across hundreds of traditional lands and custodial groups in Australia, and with First Nations people around the globe. We wish to pay our deepest respects to their ancestors and members of their communities, past, present, and to their emerging leaders. Our passion and commitment to work with all Australians and peoples from across the world, including our First Nations peoples are at the core of the work we do, reflective of our institutions' values and commitment to our role as leaders in the Reconciliation space in Australia.

Abstract

Many bacteria utilise bacterial cell-cell signalling called quorum sensing as a means of regulating various phenotypes. Small lipid-diffusible signalling molecules produced by the bacteria are called autoinducers and increase in concentration as population density increases. At a certain autoinducer concentration threshold, the signalling molecules are recognised by transcriptional activators that go on to activate, repress or derepress target genes regulating population density-dependent phenotypes. The LuxRI QS systems of Gram-negative bacteria comprise LuxI-family proteins that act as the autoinducer synthase and LuxR-family proteins that act as the autoinducer receptor and transcriptional activator of target genes. The autoinducer produced in these systems are called *N*-acyl homoserine lactones (AHLs) and include a homoserine lactone ring and a fatty acid chain with varying length, degrees of saturation, and group substitution at the third carbon position.

Quorum sensing regulates many phenotypes in a broad range of bacteria, including symbiosis and nitrogen-fixation phenotypes in plant-associated bacteria of the rhizosphere. *Mesorhizobium* spp. generally form symbiotic relationships with leguminous plants and different species use different quorum sensing systems to regulate phenotypes such as growth rate, plant nodulation and horizontal gene transfer.

In this thesis, we characterise an extremely conserved quorum sensing system in *Mesorhizobium* spp., called the *Mesorhizobium* quorum sensing (MQS) loci, that uniquely incorporates two AHL-synthase genes, both of which are essential to produce the AHL molecule to activate the cognate LuxR-family receptor. Various genetic and molecular techniques alongside LC-MS/MS and RNA sequencing were used to determine that the AHL produced by MQS is a unique di-saturated C12 molecule 2,4-trans-C12-HSL, and that the system regulates a suite of at least seven non-coding RNAs. The genomic location of these RNAs hint at possible roles in phosphate and polysaccharide regulation.

A distinct quorum sensing system of *Mesorhizobium* spp. (TraRI) regulates the horizontal gene transfer of mobile genetic elements harbouring genes for plant symbiosis. Genetic and molecular techniques were used to further characterise the epigenetic protein concentration-dependent regulation of this system regulating the transfer of an integrative and conjugative element ICE*Ml*Sym^{R7A} in *M. japonicum* R7A.

The remainder of this thesis describes cross-talk of quorum sensing systems and an additional layer of quorum-sensing regulation called quorum quenching, whereby quorum sensing is inhibited by mechanisms such as the degradation of AHLs, hindrance of AHL synthesis and LuxR-family receptor antagonism. Genetic and molecular techniques along with recombinant protein purification were used to determine the specificity and cross-talk capability of *M. japonicum* quorum sensing systems, and characterisation of a *Mesorhizobium* spp. encoded AHL degradation enzyme which degrades the TraI AHLs and represses horizontal gene transfer.

Overall, this thesis includes evidence that *Mesorhizobium* spp. encode a novel, unique quorum sensing system where two AHL-synthase genes are essential to produce a di-unsaturated AHL molecule – 2,4-trans-C12-HSL – and the system regulates a network of non-coding RNAs. A separate quorum sensing system of *Mesorhizobium* spp. called TraRI is known to regulate horizontal gene transfer, is itself epigenetically regulated by DNA-binding protein concentrations. The two aforementioned quorum sensing systems do not interact with each other, although some species encode an AHL degradation enzyme able to degrade TraI AHLs. Understanding quorum sensing mechanisms such as those described here not only increases our collective knowledge of biological nitrogen fixation and symbiosis, but of all quorum sensing systems regulating growth, pathogenicity, motility, virulence, metabolism and more in many bacteria.

Acknowledgements and contributions

I would like to express my sincere appreciation for my supervisory team Josh Ramsay, Jason Terpolilli and Clive Ronson. I would especially like to thank Josh, I am so grateful for your unwavering support and guidance over the last few years, and I feel so lucky that I landed in your group as a young naïve researcher. Any success gained by myself and your other students in the future is made possible by your mentorship and commitment to us from the beginning.

I thank all Ramsay lab members, past and present, for everything you've taught me in the lab, in the office and at the tav. Special thanks to Callum Verdonk for sharing the experience with me from honours to now. We made it! Thanks Ratish Permala and Riley Murphy for taking me under your wing when I was a newbie and making me feel so welcome. Thanks Karina Yui Eto and Elena Colombi who became amazing mentors and friends, both in and out of the lab. To everyone in the dungeon; Drew, Amy, Lamps, Lesh, Aleesha, Mark, Chris, Jared, Lee and Kirsten, I truly could not have asked for a better group of people to share my life with for 6+ years.

Thank you also to my thesis chair Carl Mousley, who always checked in with me through the years and made sure I was doing well – PhD or otherwise!

Tim Haskett, who's PhD laid foundation for a lot of this thesis, thanks for putting up with my dumb questions throughout the years and just being a great mate.

I would also like to thank the B311 PC2 crew; Heather, Luma, Angie, Yumi, Lily, Chris and Randy who have been nothing but supportive of my double life as a PhD student and a full-time tech for the last 2.5 years. I would not be where I am now without the kindness and understanding of this team.

Thanks to Ayeshia, Sara and Zal for the ongoing support, especially in the toughest last few months. A special thank you to Melissa Eccles for the years of dinners, laughs, ongoing support, and friendship.

A huge thanks to Calum, Will, Ben, Caitlan, Georgina, Kit, Mac, and Emma from the groups at Otago and Murdoch, with whom I've gained lifelong collaborators and friends.

Thanks to Ash, Kyla, Lani and Erin who have supported me not only during this time but for the last 15+ years.

I would like to extend a huge thank you to my mum and dad, Toni and Mark, who have supported me through anything and everything I wanted to do in my life. You always gave me everything you had and more so I could succeed. Thanks also to Aunty Renae and Uncle Eric, who have supported me more than you know. My best friend and sister Caitlin, I seriously would not be here, finishing this thing without you. And finally, thank you Spencer and Dustin, for everything.

This research was supported by the Curtin Health Innovation Research Institute (CHIRI) and the Curtin Medical School, Curtin University.

This research was funded by the Australian Government, Australian Research Council Future Fellowship Project ID FT170100235, including HDR stipend.

This research was supported by an Australian Government Research Training Program (RTP) Fee Offset Scholarship.

For technical assistance I would like to thank:

Nigel Halliday, Alex Truman and Paul Williams at the University of Nottingham for synthesis of the 5-cis-C12-HSL molecule and carrying out LC-MS experiments. Data is presented in section 3.2.6.

Rob Trengove at Curtin University for assistance with LC-MS/MS experiments, data presented in section 3.2.7

John Sullivan at the University of Otago for carrying out the Lotus *in planta* experiments, data presented in sections 4.2.1.5 and 4.3.1.6., and for constructing various constructs and strains prior to my starting this project.

Yvette Hill, Jason Terpolilli, Graham O'Hara, Kit Burns and Georgina Stagg at Murdoch University for assistance and expertise in chickpea glass house experiments. Data presented in sections 4.2.1.1 – 4.2.1.4.

Tim Haskett, Hayley Knights and Philip Poole at the University of Oxford for carrying out the RNAseq experiments when I was unable to travel due to the COVID-19 pandemic. Data presented in sections 5.2.1 and 5.2.2.

Elena Colombi at Curtin University for assistance with cloning and MUG assays. Data presented in sections 5.2.3 and 5.2.4.

Former honours student Beatrice Panganiban carried out cloning, mutant construction and CV026 assays in figure 6.10 and section 6.2.10 under the supervision and experimental design of Tahlia Bastholm.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

C4-HSL	*N*-tetranoyl-*L*-homoserine lactone
C5-HSL	*N*-pentanoyl-*L*-homoserine lactone
C6-HSL	*N*-hexanoyl-*L*-homoserine lactone
C7-HSL	*N*-heptanoyl-*L*-homoserine lactone
C8-HSL	*N*-octanoyl-*L*-homoserine lactone
C10-HSL	*N*-decanoyl-*L*-homoserine lactone
C12-HSL	*N*-dodecanoyl-*L*-homoserine lactone
3-OH-C6-HSL	*N*-(3-hydroxyhexanoyl)-*L*-homoserine lactone
3-OH-C8-HSL	*N*-(3-hydroxyoctanoyl)-*L*-homoserine lactone
3-OH-C10-HSL	*N*-(3-hydroxydecanoyl)-*L*-homoserine lactone
3-OH-C12-HSL	*N*-(3-hydroxydodecanoyl)-*L*-homoserine lactone
3-oxo-C4-HSL	*N*-(3-oxotetranoyl)-*L*-homoserine lactone
3-oxo-C6-HSL	*N*-(3-oxohexanoyl)-*L*-homoserine lactone
3-oxo-C8-HSL	*N*-(3-oxooctanoyl)-*L*-homoserine lactone
3-oxo-C12-HSL	*N*-(3-oxododecanoyl)-*L*-homoserine lactone
3-oxo-5-cis-C12-HSL	*N*-((5*Z*)-5-3-oxododecanoyl)-*L*-homoserine lactone
3-oxo-9-cis-C16-HSL	*N*-((9*Z*)-9-3-oxohexadecanoyl)-*L*-homoserine lactone
3-oxo-11-cis-C16-HSL	*N*-((11*Z*)-11-3-oxohexadecanoyl)-*L*-homoserine lactone
3-OH-7-cis-C14-HSL	*N*-((7*Z*)-7-3-hydroxytetradecanoyl)-*L*-homoserine lactone
5-cis-C12-HSL	*N*-((5*Z)*-5-dodecanoyl) homoserine lactone
7-cis-C14-HSL	*N*-((7*Z)*-7-tetradecanoyl) homoserine lactone
9-cis-C16-HSL	*N*-((9*Z)*-9-hexadecanoyl) homoserine lactone
2,4-trans-C12-HSL	*N*-((2*E*, 4*E)*-2,4-dodecanoyl) homoserine lactone
°C	Degrees Celsius
6H	Hexahistidine
Ab	Arabinose
ACP	Acyl carrier protein

AHL	*N*-acyl homoserine lactone
AI-2	Autoinducer-2
AIP	Autoinducing peptide
Ala	5-Aminolevulinic acid
ANOVA	Analysis of variance
Ap	Ampicillin
APS	arabinose-containing polysaccharide
ARA	Acetylene reduction assay
as	Antisense
att	Attachment site
BDSF	Burkholderia diffusible signalling factor
BLAST	Basic local alignment search tool
BLASTn	BLAST (nucleotide)
BLASTp	BLAST (protein)
BLASTx	BLAST (translated nucleotide)
bp	Base pair
C protein	Controller protein
c-di-GMP	Cyclic diguanosine monophosphate
CFU	Colony forming units
CoA	Coenzyme A
Ct	Cycle threshold
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dpi	Days post inoculation
DSF	Diffusible signalling factor
DTT	Dithiothreitol
EIC	Extracted ion chromatogram
EPS	Exopolysaccharide
et al.	Et alia
g	Gravitational force
G/RDM	Glucose/RDM
Gm	Gentamicin

HF	High fidelity
HPLC	High-performance liquid chromatography
ICE	Integrative and conjugative element
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
IR	Inverted repeat
kb	Kilobase
kDa	Kilodaltons
Km	Kanamycin
LB	Lysogeny broth
LC	Liquid chromatography
LC-MS	Liquid chromatography–mass spectrometry
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
m/z	Mass to charge ratio
MAST	Motif Alignment and Search Tool
mbar	Millibar
MCS	Multiple cloning site
MEME	Multiple Em for Motif Elicitation
MES	2-(N-morpholino) ethanesulfonic acid
MGB	Minor-groove binding
MQS	Mesorhizobium Quorum Sensing
MRM	Multiple reaction monitoring
mRNA	Messenger RNA
MS	Mass spectrometry
MS2	Tandem mass spectrometry
MUG	4-methylumbelliferyl-β-D-galactopyranoside
ncRNA	Non-coding RNA
nm	Nanometers
Nm	Neomycin
NMR	Nuclear magnetic resonance
NS	Not significant
OD	Optical density

P	Promoter
PAGE	polyacrylamide gel electrophoresis
PAS	Per-Arnt-Sim
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMA1	Phenotypic microarray 1
PMA2	Phenotypic microarray 2
PolyP	Polyphosphate
ppGpp	Guanosine tetraphosphate
PQS	Pseudomonas quinolone signal
psi	Pounds per square inch
PVC	Polyvinyl chloride
qPCR	Quantitative PCR
QQ	Quorum quenching
Qrr	Quorum regulatory RNA
QS	Quorum sensing
RDM	Rhizobium defined medium
RFU	Relative fluorescence units
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RNAseq	RNA sequencing
rpm	Revolutions per minute
S/RDM	Sucrose/RDM
SAM	S-Adenosyl methionine
SDR	Short-chain dehydrogenase/reductase
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SIM	Single ion monitoring
Sm	Streptomycin
SOB	Super optimal broth
Sp	Spectinomycin
Sp.	Species (singular)
Spp.	Species (plural)

TAE	Tris base, acetic acid, Ethylenediaminetetraacetic acid
Tc	Tetracycline
T-DNA	Transfer DNA
tRNA	Transfer RNA
TSS	Transcription start site
TY	Tryptone yeast
UPD	Uridine diphosphate glucose
UV	Ultraviolet
v/v	Volume/volume ratio
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X^R	X-antibiotic resistant
X^S	X-antibiotic sensitive
Å	Angstrom
α	Alpha
β	Beta
Ω	Omega

Chapter 1

Introduction

1.1 Quorum sensing

Many bacteria use chemical signalling systems called quorum sensing (QS) to communicate intercellularly and regulate various phenotypes. Bacterial cells produce small membrane-diffusible molecules called autoinducers, which proportionally increase in concentration alongside culture density and can therefore act as an indirect measure of population density. At low population density, cells produce a basal level concentration of autoinducers. Once the concentration reaches a minimum threshold, the autoinducers bind cytoplasmic receptor proteins – generally also DNA-binding transcriptional regulators – which subsequently induce, derepress or repress genes controlling certain phenotypes under QS control. Most QS circuits exhibit positive feedback, where activation of the receptor protein also upregulates autoinducer synthesis, synchronising the activation of QS in other adjacent cells and leading to population-level synchronisation of gene expression and phenotype regulation (1-4).

1.2 The LuxR-LuxI paradigm

The first QS system was discovered in 1970 in a bioluminescent bacterium *Vibrio fischeri* (previously *Photobacterium fischeri*), which lives symbiotically within the light organ of various marine animals including the bobtail squid (4, 5). The bioluminescence phenotype of *V. fischeri* is encoded by the *lux* operon, and genes encoding QS responsible for *lux* gene-regulation were named similarly with the autoinducer synthase named LuxI and the receptor named LuxR (6). The autoinducer signalling molecules produced by LuxI are called *N*-acyl homoserine lactones (AHLs). Once a minimum ‘quorum’ is reached the AHL binds LuxR and the AHL-LuxR complex bind a *lux*-box located upstream of *luxI*. This promotes *luxI* expression and therefore increases AHL production and induces expression of *luxCDABE,* which control bioluminescence. This QS system of *V. fischeri* has now become the most well-studied QS system to date, and similar AHL-producing systems are widespread across Gram-negative bacteria. Newly discovered AHL systems are routinely named after the *V. fischeri* system, where the AHL synthase is “I” and the receptor is “R” (6). The gene encoding the LuxR-family protein is generally (but not always)

located upstream of the *luxI* gene encoding the AHL-synthase responsible for the synthesis of the cognate AHL molecule.

1.2.1 *N*-acyl homoserine lactones

AHL molecules consist of a homoserine lactone ring and an acyl chain – the chain of which can vary in length from C_4-C_{18}. The molecule may also vary by a substitution of a hydrogen, oxo or hydroxyl group at the C_3 position as well as varying degrees of fatty acid chain saturation (7, 8) (Figure 1.1). The LuxI-family AHL synthase catalyses the synthesis of AHL molecules from two substrates – S-adenosyl-L-methionine (SAM) and acyl-acyl carrier protein (acyl-ACP). The reaction gives rise to the production of AHLs, as well as two other by-products – holo-acyl carrier protein and 5'-methylthioadenosine (9, 10). While LuxI-family synthases generally synthesise AHLs capable of activating the cognate LuxR-family protein, it is also possible for one synthase to produce a range of AHL molecules, such as YtbI of *Yersinia pseudotuberculosis*, which produces at least 24 different AHLs (11).

AHL molecules vary greatly in the chain length and substitution at the third carbon position (8). Varying degrees of saturation within the fatty acid chain is not as common, with most characterised AHLs being saturated and only a few examples of unsaturated AHLs have been described. *Rhodobacter sphaeroides* produces an unsaturated 7-cis-C14-HSL molecule (12), *Sinorhizobium meliloti* produces a 9-cis-C16-HSL molecule (12, 13), and a marine *Mesorhizobium sp*. produces 5-cis-3-oxo-C12-HSL and 5-cis-C12-HSL (14). While these AHL species have been identified in nature, the mechanism of synthesis or the catalysis of double bonds in AHLs are not yet understood.

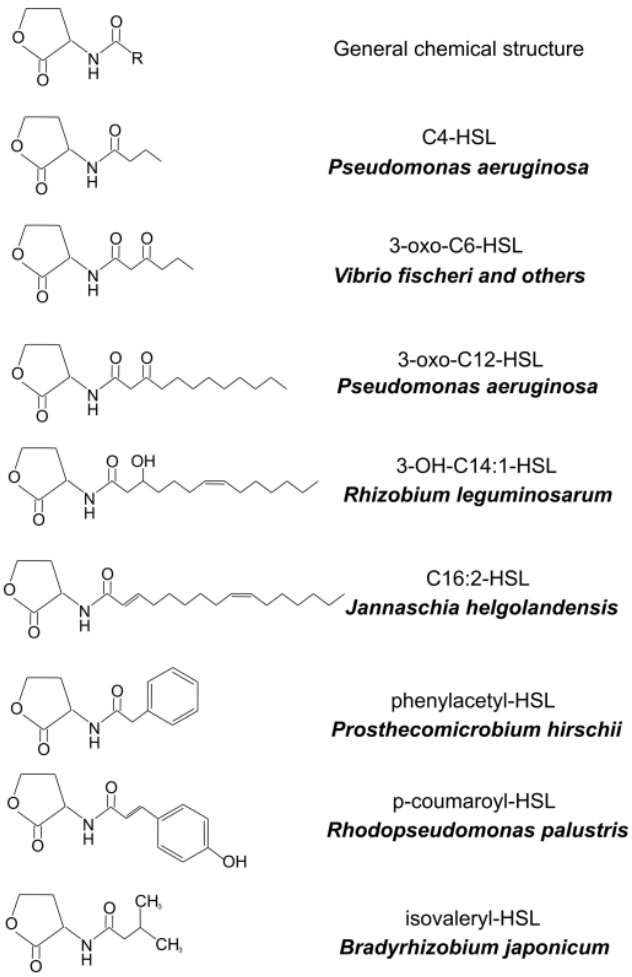


Figure 1.1 Examples of different AHL structures. The top structure depicts the generalised AHL structure, where the R group is any organic acid. The other structures are examples of naturally occurring AHL molecules produced by the indicated bacteria, each with specific chain lengths, degrees of fatty acid saturation and C_3 modifications (7).

1.3 Diffusible signal factors

A second QS system of Gram-negative bacteria – distinct from the LuxRI AHL-producing systems – uses signalling molecules called diffusible signal factors (DSFs) to control various phenotypes based on population density (15, 16). Like AHLs, DSFs can vary in chain length as well as branching, but unlike AHLs almost all characterised DSF molecules include a double bond between the second and third carbon (17). Genes for DSF synthesis are located in the *rpf* (regulation of pathogenicity factors) gene cluster. RpfF is an enzyme belonging to the crotonase superfamily and is essential for the synthesis of DSFs (15). RpfF is a uniquely bi-functional crotonase-family enzyme, acting

as both a thioesterase and a dehydratase during DSF synthesis (18, 19), acting on acyl-ACP substrates of C_8-C_{14} in length, but preferentially C_{10}-C_{12} (19). The RpfF homologue of *Burkholderia cenocepacia* – Bcam0581 – acts as a dehydratase on the acyl-ACP substrate 3-hydroxydodecanoyl-ACP and introduces the double bond at the C_2 position, creating *cis*-2-dodecenoyl-ACP. Bcam0581 then also acts as a thioesterase and releases the DSF-family molecule *cis*-2-dodecenoic acid (18).

Various sensor and response pathways are described for these DSF molecules; one utilises RpfC and RpfG as a two-part sensor and response system (20), a second utilises RpfR which acts as a sensor protein (21, 22), and a third employs RpfS – a distinct DSF-sensor protein (23). Few bacterial species are capable of producing DSF molecules, however they can reportedly be recognised by diverse bacteria and fungi (24, 25).

1.4 Quorum sensing cross-talk

Some bacteria are incapable of producing AHLs but still harbour orphan LuxR-family proteins, which can detect and respond to exogenous AHLs. *Escherichia*, *Klebsiella* and *Salmonella* spp. do not produce their own AHLs, but they still encode a functional LuxR homologue called SdiA, which is able to recognise and respond to 3-oxo-C6-HSL and 3-oxo-C8-HSL molecules of other bacteria in the environment (26, 27). Similarly, *Pseudomonas aeruginosa* PAO1 lacks the necessary gene to synthesise a distinct autoinducer molecule – AI-2. However, PAO1 virulence genes are regulated by AI-2 produced by neighbouring bacteria in the environment (28). QS not only regulates phenotypes between species but also between different kingdoms (29). For example, *Streptococcus* spp., *P. aeruginosa* and *Candida albicans* commonly coexist in the mouth. *Streptococcus* spp. produce a DSF-family molecules *(trans*-2-decenoic acid) and *P. aeruginosa* produces an AHL (3-oxo-C12-HSL), both of which supress *C. albicans* ability to form hyphae (30, 31).

1.5 Quorum sensing inhibition – quorum quenching

QS can be disrupted by a phenomenon called quorum quenching (QQ). There are several mechanisms responsible for QQ, including the inhibition of AHL synthesis, LuxR antagonism and AHL degradation. Disruption of AHL occurs either by degradation of AHL precursor molecules S-adenosyl methionine (SAM), acyl-acyl carrier protein (acyl-ACP) and acyl-CoA or inhibition of molecule binding to the AHL synthase protein (32-35). LuxR antagonism employs small ligands that bind the LuxR-family receptor and either occludes AHL binding or causes the LuxR protein to dimerise in a different formation than the LuxR-AHL complex, preventing DNA binding to the *lux*-box. While the antagonistic ligand is often a distinct molecule, it's also possible that non-cognate AHLs can act as an antagonist (36-39). The most common mechanism of QQ is degradation of the autoinducer molecule itself. Three families of enzymes are responsible for QQ degradation; acylases, lactonases and oxidoreductases. Acylases act by hydrolysing the amide bond and separating the fatty acid chain from the lactone ring, lactonases hydrolyse the ester bond in the lactone moiety, opening up the ring and oxidoreductases oxidise or reduce the third carbon in the fatty acid chain (Figure 1.2). The lactonylsis reaction is reversible by acidification, whereas the other two enzymatic reactions are non-reversible. The lactone ring of AHLs is also susceptible to hydrolysis at pH >7, which is also reversible by acidification. Lactonase enyzmes fall into a number of protein superfamilies, each with different characteristics such as differing substrate recognition, thermostability and whether or not metal cations are required for enzyme function (40).

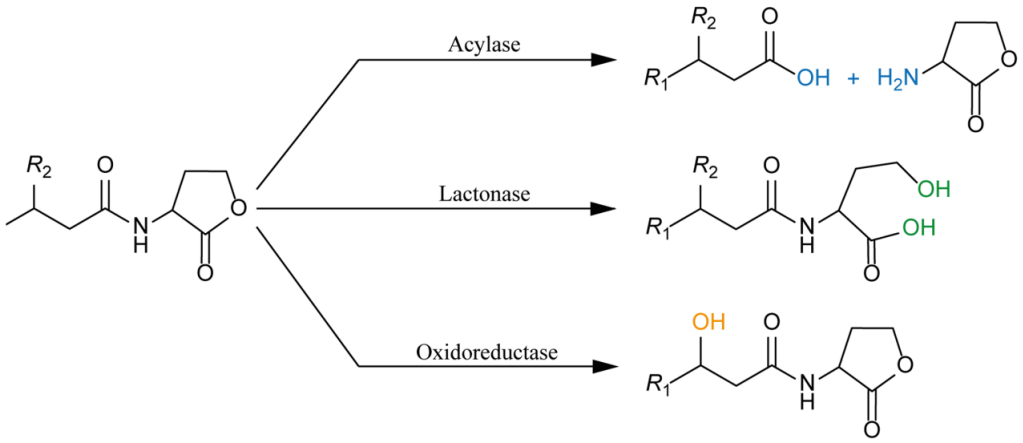


Figure 1.2 Depiction of three classes of AHL degradation enzymes. Acylases hydrolyse the amide bond to break the lactone ring from the fatty acid chain (blue), lactonases hydrolyse the ester bond to open the lactone ring (green) and oxidoreductases modify the fatty acid chain (orange) (41).

1.6 Quorum sensing in the rhizosphere

The term rhizosphere describes the nutrient-rich soil surrounding plant roots and harbours a vast range of pathogenic and beneficial microorganisms. Many bacteria capable of producing and/or responding to AHLs have been isolated from the rhizosphere, including species belonging to the genera *Serratia*, *Pseudomonas*, *Agrobacterium*, *Burkholderia*, *Rhizobia*, *Sinorhizobium*, *Mesorhizobium* and more (42, 43).

1.6.1 *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is isolated from soil and water and is a critical opportunistic human pathogen. *P. aeruginosa* expresses two AHL-based QS systems – LasR-LasI and RhlR-RhlI – as well as several lone LuxR-family proteins with no cognate LuxI, known as orphan LuxRs. LasR acts as a master regulator of the complex hierarchical QS network and binds the LasI-derived 3-oxo-C12-HSL molecule. The LasR-AHL complex further upregulates expression of *lasI*, expression of virulence genes and also induces the Rhl QS system. RhlI synthesises C4-HSL, which binds the cognate RhlR and subsequently upregulates *rhlI* expression as well as virulence genes, some of which overlap with those induced by LasR. An orphan LuxR-family protein –

QscR – is activated by LasI-derived AHLs and represses global AHL-associated QS by increasing the threshold AHL concentration at which LasR and RhlR are activated, although the mechanism is so far unclear (44, 45). While LasR is at the top of the hierarchy and regulates the Rhl system, *lasR* deletion mutants are able to circumvent this hierarchy to enable regular Rhl activity through the mutation of a LysR-family transcriptional regulator MexT (46). In addition to the AHL-based circuits, *P. aeruginosa* also encodes another QS system which utilises a signalling molecule – 2-heptyl-3-hydroxy-4-quinolone – called *Pseudomonas* quinolone signal (PQS). When bound to its receptor PqsR, PQS activates a range of genes in a similar manner to the AHL driven systems. PQS expression is positively regulated by LasR and negatively regulated by RhlR, while PQS also in turn positively regulates Rhl expression. On top of the regulatory mechanisms already described, a number of other proteins play various roles in the promotion or repression of these QS systems (47, 48), and such proves QS to be an extremely tightly controlled phenomenon.

1.6.2 *Agrobacterium tumefaciens*

Agrobacterium tumefaciens is a plant pathogen causing crown gall disease in a wide range of plant hosts. The crown gall tumours develop when *A. tumefaciens* transfers a portion of DNA called transfer DNA (T-DNA), which is carried on the tumour inducing plasmid (pTi) and translocated into the plant cell nucleus. Genes encoded by the *A. tumefaciens* derived T-DNA are then expressed in the plant leading to abnormal cell division and the development of crown gall tumours (49-51). The Ti plasmid harbours LuxR and LuxI family proteins called TraR and TraI respectively. Expression of *traR* is activated in the presence of signalling molecules called opines, which are encoded on the T-DNA and only expressed within the plant tumour environment (52). TraI synthesises 3-oxo-C8-HSL and the TraR-AHL complex induces expression of genes for conjugation and replication, allowing for QS-mediated regulation of pTi horizontal transfer and copy number (52-55).

The TraR protein is inherently unstable and subject to rapid proteolysis in the absence of its 3-oxo-C8-HSL AHL ligand. Successful folding of TraR requires the AHL binding in the ligand-binding pocket of the nascent protein, which then becomes buried within the middle of the tertiary structure. One AHL molecule binds one TraR monomer and the TraR-AHL complex subsequently acts as a homodimer to bind *tra*-boxes located upstream of TraR-regulated promoters (56, 57). The Ti plasmid also encodes TraM – an antiactivator protein that acts to bind and inactivate TraR in both the DNA bound and unbound forms (58, 59). The TraR-AHL complex binds a *tra*-box upstream of *traM*, upregulating TraM production and creating a negative feedback loop. In an extra layer of quorum sensing regulation, pTi encodes TrlR – a truncated homologue of the first 181 amino acids of TraR. TrlR is able to bind AHLs and dimerise like TraR but lacks the DNA-binding domain. TrlR readily forms dimers with TraR, creating a TraR:TrlR-AHL complex incapable of binding the *tra*-box DNA and therefore inhibiting TraR-dependent gene expression (60, 61) (Figure 1.3).

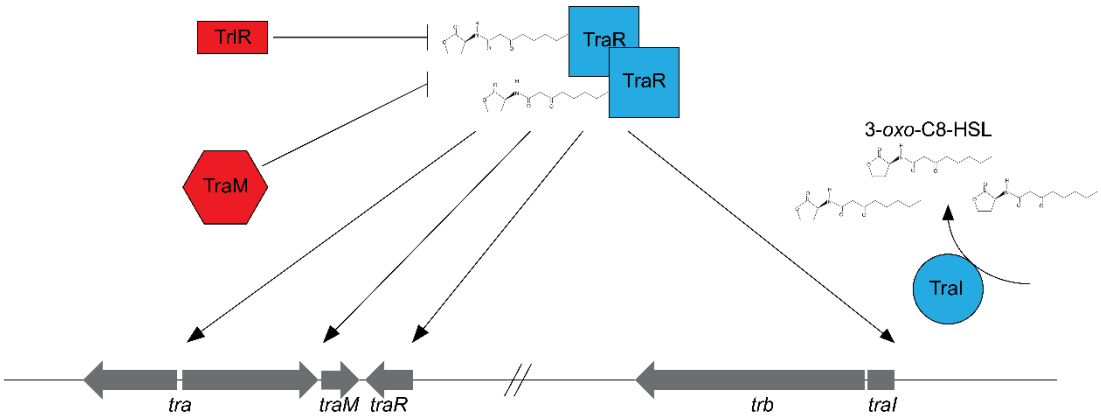


Figure 1.3. Overview of quorum sensing regulation in pTi of *Agrobacterium tumefaciens*. TraI produces the AHL species 3-oxo-C8-HSL, which binds TraR in a 1:1 ratio forming the TraR-AHL complex. Straight black arrows depict where the TraR-AHL complex binds *tra*-boxes located upstream of various promoters, inducing expression of those genes. TraR activity is inhibited by two mechanisms; the binding of antiactivator TraM, and by forming dimers with TrlR creating a complex unable to bind *tra*-box DNA. Proteins involved in QS are coloured blue, proteins involved in TraR inhibition are coloured red. TraR-regulated loci are depicted in grey block arrows. Figure adapted from (59).

1.6.3 *Rhizobium leguminosarum*

The first AHL described in rhizobia was discovered in *Rhizobium leguminosarum* and initially named 'small bacteriocin' due to the bacteriocin-like inhibitory effect on other strains of *R. leguminosarum*. The molecule has since been characterised as 3-OH-7-cis-C14-HSL (62, 63) and is known to be synthesised by the chromosomally encoded LuxI-type CinI (64). To date, four LuxI-family AHL synthases have been identified in *R. leguminosarum* – CinI, RhiI, RaiI and TraI. It was first thought that CinI-derived AHLs were essential for induction of the other QS systems. However, downstream of *cinI* is a small protein-encoding gene *cinS* which is cotranscribed with *cinI* and therefore is also cell-density dependent. Once accumulated, CinS acts independently of AHLs as an anti-repressor by binding PraR, a transcriptional repressor of *rhiR and raiR* expression (65, 66). Derepression of RhiR induces expression of *rhiI* leading to synthesis of C6-HSL, C7-HSL and C8-HSL (67) as well as expression of the *rhiABC* operon, believed to be involved in nodulation although so far, the exact function is not known (68, 69). Strains carrying mutations in *raiR* exhibit lower nitrogen fixing efficiency compared to wild-type strains, however the exact function of this transcriptional regulator is also not yet known (70).

Plasmid transfer is regulated by the TraRI QS system in response to AHLs produced by potential plasmid recipients. The symbiotic plasmid pRL1JI encodes TraRI and the orphan BisR, and CinRI is encoded on the chromosome. The CinI AHL-synthase produces 3-OH-7-cis-C14-HSL, however the plasmid-encoded BisR represses *cinI* expression so strains harbouring pRL1JI produce only low levels of 3-OH-7-cis-C14-HSL. Potential recipient strains lacking BisR produce elevated levels of 3-OH-7-cis-C14-HSL, which are perceived by BisR encoded on pRL1JI in the donor strain. AHL-induced BisR then activates TraR, which in turn autoregulates the *traI-trb* operon leading to TraI-derived 3-oxo-C8-HSL and C8-HSL and plasmid transfer. Similar to TraRI-regulated plasmid transfer of *A. tumefaciens*, *R.*

leguminosarum pRL1JI encodes the antiactivator TraM which binds and inhibits TraR, repressing *traI-trb* expression. Upon BisR-AHL activation of TraR, *traR* is upregulated and overcomes the TraM repression (71) (Figure 1.4).

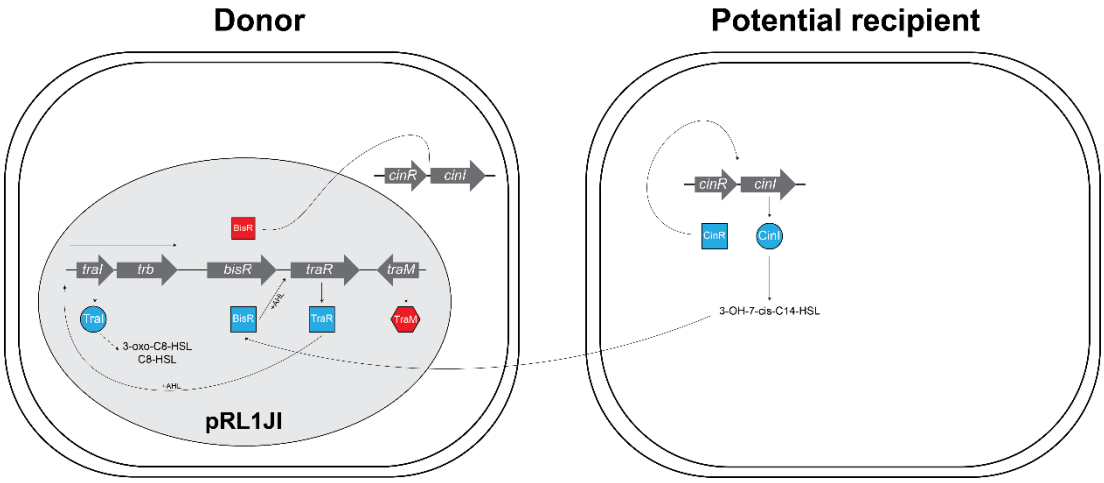


Figure 1.4 Model of recipient-induced horizontal gene transfer in *R. leguminosarum*. Proteins involved in QS are coloured blue and proteins involved in inhibiting QS are coloured red. The chromosomal *cinI* gene product produces 3-OH-7-cis-C14-HSL which in the presence of the pRL1JI plasmid (as in a donor cell) is repressed by BisR, TraM also inhibits TraR activity. Recipient cells lacking pRL1JI freely produces these AHLs which are recognised by BisR of the donor cell. BisR activates *traR* expression, resulting in *traI-trb* expression, upregulation of TraI-AHL production and transfer of pRL1JI. Figure adapted from (71).

1.6.4 *Sinorhizobium meliloti*

Sinorhizobium meliloti exists in the rhizosphere as a free-living bacterium or symbiotically with legumes of the genera *Medicago*, *Melilotus* and *Trigonella*. *S. meliloti* encodes *luxR-luxI*-type genes – *sinR* and *sinI* – as well as a gene encoding an orphan LuxR-type protein – ExpR. SinI catalyses the synthesis of several longer-chain AHLs, with 3-oxo-9-cis-C16-HSL as the main product (72). Transcription of *sinR* is influenced by nutrient depletion such as low phosphate and nitrogen/carbon starvation. SinR drives expression of *sinI* and therefore AHL production, however these AHLs do not affect SinR and instead activate the orphan ExpR. The ExpR-AHL complex positively regulates expression of *sinI* and represses sinR, while also regulating many genes

including those involved in extracellular polysaccharide and flagella production (73, 74). *S. meliloti* carries a number of other orphaned LuxR-type proteins (those without corresponding LuxI proteins), although these receptors are not known to respond to AHLs and instead respond to other environmental stimuli and host-derived signalling molecules (75). *S. meliloti* Rm41 carries a large plasmid called pRme41a which encodes QS genes homologous to *traRI* of *Agrobacterium* spp. (72, 76). The TraRI circuit of *S. meliloti* also produces 3-oxo-C8-HSL molecules and regulates conjugation, and has since been characterised in other *Sinorhizobium* species (76).

1.6.5 Mesorhizobium spp.

Mesorhizobia are mid-to-slow growing soil dwelling bacteria that often form nitrogen fixing symbiosis with specific legume host plants. Members of the *Mesorhizobium* genera encode a range of QS genes, including *mtqRI* in *M. tianshanense* which regulates growth rate and nodulation (77), *mrlI1*, *mrlI2* and *mrlI3* of *M. loti* NZP 2213 – all three of which reportedly affect nodulation (78), a marine *Mesorhizobium sp*. produces unique long chain unsaturated AHLs 3-oxo-5-cis-C12-HSL and 5-cis-C12-HSL (14) and recently *M. japonicum* MAFF 303099 was found to produces 2,4-trans-C12-HSL (79). One of the most well characterised QS system in mesorhizobia is the TraRI system which regulates horizontal gene transfer of mobile genetic elements encoding symbiosis genes (80-83).

1.7 QS-regulated horizontal gene transfer in *Mesorhizobium* spp.

Integrative and conjugative elements (ICEs) are mobile genetic elements that are transferred to recipient cells via conjugation. In mesorhizobia these ICEs can carry genes required for symbiosis and nitrogen fixation. Early work in *M. japonicum* (previous *M. loti* and *Rhizobium loti*) discovered that strains of *Mesorhizobium* that had been used as an inoculum for nitrogen fixation seven years prior had subsequently transferred a large portion of the chromosome – including symbiosis genes – to distinct non-symbiotic *Mesorhizobium* strains

indigenous to the soil (84). The transferrable symbiosis genes have since been further characterised as a 501.8 kb ICE called ICE*Ml*Sym^{R7A} (85, 86).

ICE*Ml*Sym^{R7A} horizontal gene transfer is regulated by the QS proteins TraR and TraI1. TraI1 catalyses the production of the AHL 3-oxo-C6-HSL. The *traI2* gene is likely a pseudogene but still maintains a *tra*-box in the promoter region, allowing TraR binding and activation (81). Encoded directly downstream of *traI2* are two genes; *msi172* and *msi171*. Upon induction of P*traI2*, a programmed ribosomal frameshift within msi*172*-msi*171*, occurring in approximately 4-12% of translation events, produces the protein FseA which directly induces expression of the excisionase RdfS leading to ICE*Ml*Sym^{R7A} excision (83).

Not unlike the TraRI systems of *A. tumefacien*s and *R. leguminosarum*, ICE*Ml*Sym^{R7A} encodes an antiactivator protein called QseM which inhibits TraR and FseA (81). QseM directly blocks both TraR and FseA activity via protein-protein interactions resulting in strong repression of QS and ICE*Ml*Sym^{R7A} excision. Downstream of *qseM* is a gene encoding the controller protein QseC. QseC binds two operator sites within its promoter region to regulate expression of *qseM* and its own expression (81, 87) (Figure 1.5).

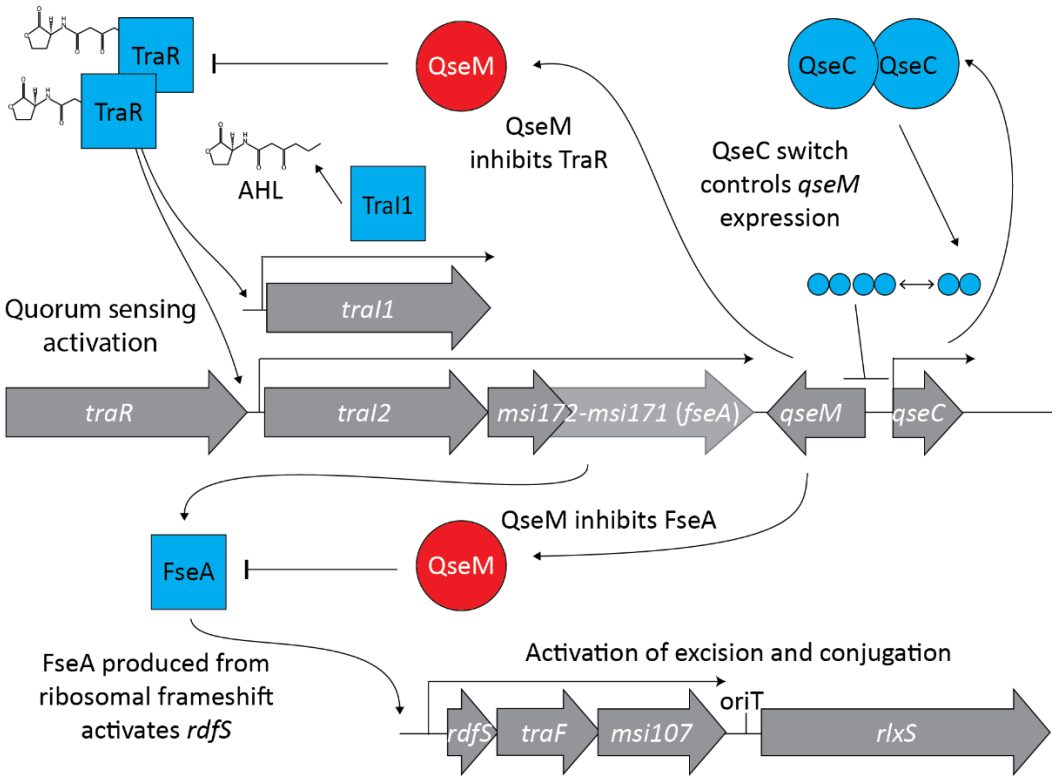


Figure 1.5 Quorum sensing, excision and transfer in R7A. Depiction of *traR-qseC* and *rdfS-rlxS* encoded on ICE*Ml*Sym^{R7A}. Proteins that positively regulate QS are coloured blue, QseM is coloured red. TraI1 produces AHLs that activate TraR, TraR-AHL then activates two *tra*-boxes upstream of *traI1* and *traI2*, upregulating AHL production and expression of *fseA* following the low frequency +1 ribosomal frameshift. FseA activates P*rdfS*, inducing excision and conjugative transfer. When QS is repressed, QseM binds TraR and FseA, inhibiting their functions in QS and excision. QseC regulates the P*qseM* and P*qseC* promoters. Figure adapted from (87).

1.8 Rhizobia and symbiosis

The term rhizobia describes a collective of soil dwelling α-proteobacteria and β-proteobacteria that are capable of forming symbiotic relationships with leguminous host plants. These symbiotic bacteria fix atmospheric nitrogen (N_2) to the biologically available form ammonia (NH_3) that can then be assimilated by the plant host (88-90). The plant releases compounds called flavonoids into the rhizosphere which are recognised by rhizobial NodD proteins. These transcriptional regulators bind *nod* boxes, activating transcription of *nod* genes and producing molecules called Nod factors (88, 91-95). Nod factors stimulate several responses for rhizobial infection in the plant host including plant cortical cell division and entrapment of the bacterium via root hair curling.

Entrapped bacteria migrate down the root hair by cell division, forming a structure called the infection thread. The rhizobia at the end of the infection thread are internalised and form the nodule primordium and eventually differentiate into nitrogen-fixing bacteroids (88, 92, 96-100). The decrease in oxygen triggers oxygen-sensing regulatory systems which subsequently induce expression of *nif* genes encoding nitrogenase, the enzyme complex required for nitrogen fixation (88, 101-104).

Members of the genus *Mesorhizobium* have largely been isolated from legume root nodules, and therefore all of these symbiotic isolates are capable of nitrogen fixation and harbour nitrogen fixation genes (105). Recently, 144 *Mesorhizobium* spp. were isolated from Western Australian soils and sequenced, with 126 of the 144 strains lacking symbiosis genes and therefore are non-symbiotic (106). The bias in isolating rhizobial strains from root nodules excludes a large portion of the natural diversity in the rhizosphere.

1.9 Aims of this thesis

Quorum sensing is a complex regulatory phenomenon employed by almost all bacteria to regulate a plethora of phenotypes including metabolism, virulence factors, antibiotic production, pigment production, horizontal gene transfer, polysaccharide production, biofilm production, motility, bioluminescence, plant nodulation, nitrogen fixation and more. Distinct bacterial species often produce the same autoinducer molecules to regulate different phenotypes in different hosts. Understanding autoinducer production, modification, degradation and regulation of target genes is critical to further our knowledge of all things microbiology, including but certainly not limited to symbiosis, human, animal and plant pathogens and synthetic biology.

The aims of this thesis were as follows;

- (Chapter 3) Characterise the AHL molecule(s) produced by chromosomally encoded QS system of *M. japonicum* R7A and *M. ciceri* WSM1271.

- (Chapter 4 & 5) Investigate the biological activity of these AHLs and identify target genes of the LuxR-family receptor.

- (Chapter 6) Further characterise the QS regulation of the ICE*Ml*Sym^{R7A} encoded TraRI QS system responsible for activating horizontal gene transfer.

- (Chapter 7) Explore the cross talk between different QS systems in the rhizosphere

Chapter 2

Materials and methods

2.1 Media and growth conditions

Escherichia coli strains were cultured in Lysogeny broth (LB) (107) at 37°C. *E. coli* ST18 (108) was grown in media supplemented with 5-aminolevulinic acid (ala). *Mesorhizobium* sp. strains were cultured in tryptone-yeast (TY) (109) or rhizobium defined media (RDM) (110). RDM was supplemented with either 0.4% glucose (G/RDM) or 5% sucrose as a carbon source. RDM media was also supplemented with nicotinate (1 µg/mL), biotin (20 ng/mL) and thiamine (1 µg/mL) except when selecting for the presence of ICE*Ml*Sym^{R7A} such as in conjugation experiments. Where appropriate, TY liquid media was made with 40 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer at a final pH of 6.5, which allowed cultures to remain at a neutral pH throughout growth. *Chromobacterium violaceum* was cultured in TY media at 28°C. Agar (Sigma-aldrich) was added at a concentration of 1.6% w/v to solidify media where necessary, or 0.4% w/v for semi-solid agar overlays. Antibiotics and supplements were added to media as required (Table 2.1).

Table 2.1. Antibiotic and supplement concentrations

Antibiotic	**Abbreviation**	**Concentration**	
		E. coli	*Mesorhizobium* sp.
Ampicillin	Ap	100 µg/mL	-
Gentamicin	Gm	50 µg/mL	25 µg/mL
Kanamycin	Km	50 µg/mL	-
Neomycin	Nm	-	200 µg/mL
Spectinomycin	Sp	50 µg/mL	100 µg/mL
Streptomycin	Sm	100 µg/mL	200 µg/mL
Tetracycline	Tc	10 µg/mL	0.5-2 µg/mL
Supplement	**Abbreviation**	**Concentration**	
		E. coli	*Mesorhizobium* sp.
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside	X-gal	50 µg/mL	50 µg/mL
5-Aminolevulinic acid	Ala	50 µg/mL	-
Arabinose	Ab	0.01%	-

Isopropyl β-d-1-thiogalactopyranoside	IPTG	200 µM	200 µM

2.2 Bacterial strains and plasmids

Strains and plasmids used in this study are described in Table 2.2.

Table 2.2. Bacterial strains and plasmids

Strain	Description	Reference
Escherichia coli		
Epi300	General cloning strain, arabinose inducible *oriV* for increased copy number	Epicentre
ST18	S17 derivative, Δ*hemA*, auxotrophic for ala	(108)
BL21(de3)pLysS	Recombinant protein expression strain. Cm^R	(111)
NiCo de3	BL21(de3) derivative, reduced nickel-binding proteins	(112)
Mesorhizobium japonicum		
R7A	Wild-type symbiont of lotus, field reisolate of ICMP3153	(84)
R7ANS	Non-symbiotic derivative of R7A. Lacks ICE*Ml*Sym^{R7A}	(85)
R7A*	QS and conjugation active variant of R7A	(87)
R7A*2	QS and conjugation active variant of R7A. Isolated independently from R7A*	(87)
R7ANS*	Derivative of R7A* cured of ICE*Ml*Sym^{R7A}	(87)
R7A*RV	Variant of R7A* exhibiting decreased AHL production after several passage events	This study, (87)
R7ANSΔ*mqsRIC*	*mqsRIC* replaced with ΩaadA cassette, Sm^R, Sp^R	Tim Haskett, unpublished
R7AΔ*mqsR*	Markerless, in-frame deletion of *mqsR* in R7A	Ronson lab, Otago, NZ, unpublished
R7ANSΔ*mqsR*	Markerless, in-frame deletion of *mqsR* in R7ANS	Ronson lab, Otago, NZ, unpublished

R7AΔ*mqsRIC*	Markerless, in-frame deletion of the *mqsRIC* genes in R7A	This study
R7AΔ*mqsI*	Markerless, in-frame deletion of *mqsI* in R7A	This study
R7A::*exoA*	Insertion mutant created in R7A, pFUS2 plasmid inserted into the chromosome to disrupt *exoA* and measure expression from the pFUS2-borne *lacZ*	Ronson lab, Otago, NZ, unpublished
R7AΔ*mqsR*::*exoA*	Insertion mutant created in R7AΔ*mqsR*, pFUS2 plasmid inserted into the chromosome to disrupt *exoA* and measure expression from the pFUS2-borne *lacZ*	This study
R7A::*exoU*	Insertion mutant created in R7A, pFUS2 plasmid inserted into the chromosome to disrupt *exoU* and measure expression from the pFUS2-borne *lacZ*	Ronson lab, Otago, NZ, unpublished
R7AΔ*mqsR*::*exoU*	Insertion mutant created in R7AΔ*mqsR*, pFUS2 plasmid inserted into the chromosome to disrupt *exoU* and measure expression from the pFUS2-borne *lacZ*	This study
R7A::*exoY*	Insertion mutant created in R7A, pFUS2 plasmid inserted into the chromosome to disrupt *exoY* and measure expression from the pFUS2-borne *lacZ*	Ronson lab, Otago, NZ, unpublished
R7AΔ*mqsR*::*exoY*	Insertion mutant created in R7AΔ*mqsR*, pFUS2 plasmid inserted into the chromosome to disrupt *exoY* and measure expression from the pFUS2-borne *lacZ*	This study
R7AΔ*qseM*	Markerless, in-frame deletion of *qseM* in R7A	(81)
R7AΔ*traR*	Markerless, in-frame deletion of *traR* in R7A	(82)
R7A*Δ*traR*	Markerless, in-frame deletion of *traR* in R7A*	(87)
R7AΔ*qseM*Δ*traR*	Markerless, in-frame deletion of *qseM* in the R7AΔ*traR* background	(81)
R7AΔ*qseC2*	Markerless, in-frame deletion of *qseC2* in R7A	(87)
R7A*Δ*qseC2*	Markerless, in-frame deletion of *qseC2* in R7A*	(87)

R7AΔP*asqseC* Δ*qseC2*	Markerless deletion of the region between the 3' end of *qseC* through to the end of *qseC2* in R7A	(87)
R7A*ΔP*asqseC* Δ*qseC2*	Markerless deletion of the region between the 3' end of *qseC* through to the end of *qseC2* in R7A*	(87)
R7AΔP*asqseC* Δ*qseC2*::P*nptII*	P*asqseC-qseC2* region of R7A replaced with the *nptII* promoter, P*nptII* in the same orientation as *qseC*	This study, (87)
R7A*ΔP*asqseC* Δ*qseC2*::P*nptII*	P*asqseC-qseC2* region of R7A* replaced with the *nptII* promoter, P*nptII* in the same orientation as *qseC*	This study, (87)
R7AΔPas*qseC* Δ*qseC2*::Pas*nptII*	P*asqseC-qseC2* region of R7A replaced with the *nptII* promoter, P*nptII* orientated antisense to *qseC*	This study, (87)
R7A*ΔPas*qseC* Δ*qseC2*::Pas*nptII*	P*asqseC-qseC2* region of R7A* replaced with the *nptII* promoter, P*nptII* orientated antisense to *qseC*	This study, (87)
R7A*qseC*(Am)	Allelic replacement of 10th and 11th codons of R7A to create N[10]* and M[11]K	This study, (87)
R7A**qseC*(Am)	Allelic replacement of 10th and 11th codons of R7A* to create N[10]* and M[11]K	This study, (87)
Mesorhizobium ciceri		
CC1192	Wild-type symbiont of *Cicer arietinum*	(113)
CC1192Δ*mqsRIC*	Markerless, in-frame deletion of the *mqsRIC* genes in CC1192	This study
WSM1271	Wild-type symbiont of B*isserula pelecinus*	(114)
***Mesorhizobium* spp.**		
WSM4303	Wild-type	(115)
WSM4306	Wild-type	(115)
Chromobacterium violaceum		
CV026	Derivative of ATCC 31532, *cvil*::mini-Tn5, AHL^{-}, C4-C8-HSL biosensor, Km^{R}	(116)
Plasmids		
pSDz	Derivative of pFAJ1700, carrying a promoterless *lacZ* and IPTG-inducible *lacI* promoter, Tc^{R}	(83)
pPmqsI	pSDz vector containing the *mqsI* promoter from R7A upstream of	This study

	lacZ, amplified using primers 1 and 2 and cloned as NsiI-XhoI fragment, Tc^R	
pPmqsIzR	pPmqsI containing the *mqsR* coding sequence from R7A amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsI1	pSDz containing P*mqsI1* upstream of *lacZ*, amplified from pPmqsI using primers 5 and 6 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsI2	pSDz containing P*mqsI2* upstream of *lacZ*, amplified from pPmqsI using primers 5 and 7 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsI3	pSDz containing P*mqsI3* upstream of *lacZ*, amplified from pPmqsI using primers 5 and 8 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsI4	pSDz containing P*mqsI4* upstream of *lacZ*, amplified from pPmqsI using primers 5 and 9 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pSDzR	pSDz containing the IPTG inducible *mqsR* of R7A, Tc^R	Ronson lab, Otago, NZ, unpublished
pPmqsRNA1	pSDz containing the *mqsRNA1* promoter upstream of *lacZ*, amplified using primers 12 and 13 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsRNA2	pSDz containing the *mqsRNA2* promoter upstream of *lacZ*, amplified using primers 14 and 15 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsRNA3	pSDz containing the *mqsRNA3* promoter upstream of *lacZ*, amplified using primers 16 and 17 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsRNA5	pSDz containing the *mqsRNA5* promoter upstream of *lacZ*, amplified using primers 18 and 19 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsRNA6	pSDz containing the *mqsRNA6* promoter upstream of *lacZ*, amplified using primers 20 and 21	This study

	and cloned as an NsiI-XhoI fragment, Tc^R	
pPmqsRNA7	pSDz containing the *mqsRNA7* promoter upstream of *lacZ*, amplified using primers 10 and 11 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPtRNA-leu	pSDz containing the *tRNA-leu* promoter upstream of *lacZ*, amplified using primers 22 and 23 and cloned as an NsiI-XhoI fragment, Tc^R	This study
pPmqsRNA1zR	pPmqsRNA1 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsRNA2zR	pPmqsRNA2 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsRNA3zR	pPmqsRNA3 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsRNA5zR	pPmqsRNA5 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsRNA6zR	pPmqsRNA6 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPmqsRNA7zR	pPmqsRNA7 containing the *mqsR* coding sequence from R7A downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPtRNA-leuzR	pPtRNA-leu containing the *mqsR* coding sequence from R7A	This study

	downstream of the *lac* promoter, amplified using primers 3 and 4 and cloned as an EcoRI-HindIII fragment, Tc^R	
pSDz-PqseC2	pSDz containing the *qseC2* promoter amplified using primers 24 and 25 and cloned downstream of *lacZ* in the NsiI site, Tc^R	This study, (87)
pSDz-PasqseC	pSDz containing the P*asqseC* promoter amplified using primers 24 and 25 and cloned downstream of *lacZ* in the NsiI site, Tc^R	This study, (87)
pSDz-PqseC2-qseC2	pSDz-PqseC2 containing the *qseC2* gene downstream of the *lac* promoter, amplified using primers 26 and 27 and cloned as an EcoRI-HindIII fragment, Tc^R	This study, (87)
pSDz-PasqseC-qseC2	pSDz-PasqseC containing the *qseC*2 gene downstream of the *lac* promoter, amplified using primers 26 and 27 and cloned as an EcoRI-HindIII fragment, Tc^R	This study, (87)
pSDz-PqseC	pSDz containing the *qseC* promoter amplified using primers 28 and 29 and cloned downstream of *lacZ* in the NsiI site, Tc^R	This study
pSDz-PqseM	pSDz containing the *qseM* promoter amplified using primers 28 and 29 and cloned downstream of *lacZ* in the NsiI site, Tc^R	This study
pSDz-PqseC-qseC	pSDz-PqseC containing the *qseC* gene downstream of the *lac* promoter, amplified using primers 30 and 31 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pSDz-PqseM-qseC	pSDz-PqseM containing the *qseC* gene downstream of the *lac* promoter, amplified using primers 30 and 31 and cloned as an EcoRI-HindIII fragment, Tc^R	This study
pPtraI	pSDz vector containing the *traI1* promoter from R7A upstream of *lacZ*, Tc^R	Tim Haskett, unpublished
pPtraIzR	pPtraI containing the *traR* coding sequence from R7A downstream of the *lac* promoter, Tc^R	Tim Haskett, unpublished
pJN105Q	pJN105 derivative, *qscR* expression under control of the arabinose-inducible *araC-pBAD* promoter, Gm^R	(117)

pSC11-Q	PA1897-*lacZ* fusion, QscR reporter, Ap^R	(118)
pJN105L	pJN105 derivative, *lasR* expression under control of the arabinose-inducible *araC-pBAD* promoter, Gm^R	(117)
pSC11	*lasI-lacZ* fusion, LasR reporter, Ap^R	(45)
pPR3	pVS1/p15a replicon, gfp reporter, *nptII* promoter, Km^R	(119)
pPROBE-GT	pVS1/p15a replicon, gfp reporter, Gm^R	(120)
pPR3G	pPROBE-GT carrying the *nptII* promoter, Gm^R	(87)
pNmqsI	pPR3 containing the *mqsI* coding sequence downstream of P*nptII*, Km^R, Nm^R	Tim Haskett, unpublished
pNmqsC	pPR3 containing the *mqsC* coding sequence, amplified using primers 32 and 33 and cloned downstream of P*nptII* as a BamHI-KpnI fragment, Km^R, Nm^R	This study
pNmqsIC	pPR3 containing the *mqsI-mqsC* region downstream of P*nptII*, Km^R, Nm^R	Tim Haskett, unpublished
pNmqsR	pPR3 containing the *mqsR* coding sequence, amplified using primers 34 and 35 and cloned downstream of P*nptII* as a BamHI-EcoRI fragment, Km^R, Nm^R	This study
pNqseM	pPR3 containing *qseM* downstream of the *nptII* promoter, Km^R, Nm^R	(81)
pNaidM	pPR3 containing *aidM*, amplified from WSm1271 gDNA using primers 36 and 37 and cloned downstream of P*nptII* as a BamHI-KpnI fragment, Km^R, Nm^R	This study
pNqseC2	pPR3G containing *qseC2* downstream of the *nptII* promoter, Gm^R	(87)
pPR3G-antiqseC	pPR3G containing the antisense qseC region, amplified from R7A using primers 38 and 39 and cloned downstream of PnptII as a KpnI-EcoRI fragment	This study, (87)
pPR3G-sacB	pPR3G containing the *sacB* cassette, amplified from pJQ200SK using primers 40 and 41 and cloned in the EcoRI site convergently to the *nptII* promoter, Gm^R	This study, (87)

pNqseC2-sacB	pNqseC2 containing the *sacB* cassette, amplified from pJQ200SK using primers 40 and 42 and cloned into the EcoRI site convergently to the *nptII* promoter, Gm^R	This study, (87)
pFAJ1700	Broad host-range plasmid, Tc^R	(121)
pFAJ1708	pFAJ1700 containing the *nptII* promoter, Tc^R	(121)
pFAJmqsI	pFAJ1708 containing *mqsI* cloned downstream of the *nptII* promoter, Tc^R	Ronson lab, Otago, NZ, unpublished
pFAJmqsIC	pFAJ1708 containing *mqsI-mqsC* cloned downstream of the *nptII* promoter, Tc^R	Ronson lab, Otago, NZ, unpublished
pFAJmqsC	pFAJ1708 containing *mqsC* cloned downstream of the *nptII* promoter, Tc^R	Ronson lab, Otago, NZ, unpublished
pFAJmqsRIC (CC1192)	pFAJ1700 containing the whole *mqsRIC* region including native promoters from CC1192, amplified using primers 43 and 44 and cloned as a HindIII-KpnI fragment, Tc^R	This study
pFAJmqsRIC (R7A)	pFAJ1700 containing the whole *mqsRIC* region including native promoters from R7A, amplified using primers 45 and 46 and cloned as a HindIII-KpnI fragment, Tc^R	This study
pETM11	N-terminally tagged 6H protein expression vector, Km^R	(122)
pETMmqsC	pETM11 vector containing the *mqsC* coding sequence amplified from a gBlock (IDT) (Table 2.3) using primers 45 and 46 and cloned as an NcoI-HindIII fragment, Km^R	This study
pETMaidM	pETM11 vector containing the *aidM* coding sequence amplified from a gBlock (IDT) (Table 2.3) using primers 45 and 46 and cloned as an NcoI-HindIII fragment, Km^R	This study
pETMaidM$_{S102G}$	pETM11 vector containing the *aidM*$_{S102G}$ coding sequence amplified from a gBlock (IDT) (Table 2.3) using primers 45 and 46 and cloned as an NcoI-HindIII fragment, Km^R	This study
pEX18Tc	Suicide vector, harbours *sacB*, Tc^R	(123)
pEX1192-RIC	pEX18Tc containing the regions flanking *mqsRIC*, amplified from CC1192 using primers 49, 50, 51 and 52 and cloned using three-way	This study

	ligation as HindIII-BamHI-KpnI fragments, Tc^R	
pJQ200SK	Suicide vector, harbours *sacB*, Gm^R	(124)
pJQR7A-RIC	pJQ200SK containing the regions flanking *mqsRIC*, amplified from R7A using primers 53, 54, 55 and 56	This study
pJQΔPasqseC ΔqseC2::PnptII	pJQ200SK containing flanking regions of P*asqseC-qseC*, amplified using primers 57, 58, 59 and 60 and cloned using three-way ligation as SacI-BamHI-PstI fragments. P*nptII* then amplified with primers 61 and 62 cloned as a BamHI fragment. P*nptII* in same orientation as *qseC*, Gm^R	This study, (87)
pJQΔPasqseC ΔqseC2::PnptII-AS	Same as above, except P*nptII* orientated antisense to *qseC*, Gm^R	This study, (87)
pJQqseC-amber	pJQ200SK containing regions flanking the 10^{th} and 11^{th} codons of qseC, amplified using primers 63, 64, 65 and 66 and cloned using three-way ligation as PstI-XbaI-SacI fragments, Gm^R	This study, (87)
pFUS2	Suicide vector, *lacZ* transcriptional reporter, Gm^R	(125)
pFUS::exoA	pFUS2 plasmid with *exoA* flanking regions cloned for insertion into chromosome. Gm^R	Ronson lab, Otago, NZ, unpublished
pFUS::exoU	pFUS2 plasmid with *exoU* flanking regions cloned for insertion into chromosome. Gm^R	Ronson lab, Otago, NZ, unpublished
pFUS::exoY	pFUS2 plasmid with *exoY* flanking regions cloned for insertion into chromosome. Gm^R	Ronson lab, Otago, NZ, unpublished
pJR201	pUC8 plasmid containing regions of attP, attB and melR, used as a standard template for ICE*Ml*Sym^{R7A} qPCR, Ap^R	(85)

2.3 Molecular and genetic techniques

All enzymes were purchase from New England Biolabs (NEB) unless otherwise stated. Primers and synthetic DNA were purchased from Integrated DNA Technologies (IDT).

2.3.1 Genomic DNA extraction from *Mesorhizobium* spp.

Genomic DNA from *Mesorhizobium* spp. was isolated using the FavorPrep genomic DNA extraction mini kit (Favorgen) with the following additional steps; cell pellets were resuspended in Tris-EDTA buffer (pH 8) containing 75 mg/mL RNase A (Sigma-Aldrich) and 5 mg/mL lysozyme (Sigma-Aldrich). The suspension was incubated at room temperature for 15 minutes and then at 70°C for two minutes and -80°C for two minutes. These freeze-thaw steps were repeated four times to assist in lysing cells. Manufacturer's instructions were followed hereafter.

2.3.2 Plasmid DNA extraction from *E. coli*

Plasmid DNA was extracted from *E. coli* using the FavorPrep Plasmid DNA Extraction mini kit (Favorgen) as per manufacturer's instructions.

2.3.3 Polymerase Chain Reaction (PCR)

PCR was used to amplify DNA for cloning using Phusion High-Fidelity DNA Polymerase. PCR screening of clones and mutants as well as amplifying regions for sequencing was performed using One*Taq* DNA Polymerase.

Phusion PCR reactions comprised 1x Phusion HF Buffer, 200 µM dNTPs, 0.4 µM of each primer, 1 unit of Phusion DNA Polymerase, 1-3 µL of template DNA and milliQ H_2O up to 50 µL. One*Taq* PCR reactions comprised 1x One*Taq* Standard Reaction Buffer, 200 µM dNTPs, 0.2 µM of each primer, 1.25 units of One*Taq* DNA Polymerase, 1-3 µL of template DNA and milliQ H_2O up to 20 µL. Due to the high GC% content, all PCR reactions of *Mesorhizobium* DNA also contained 3% DMSO to assist with denaturing. PCR cycling conditions were as per manufacturer's instructions. Primers used are listed in Table 2.3.

Table 2.3. Primers, probes and synthesised DNA used in this study†

	Name	Sequence
1	PmqsI_F_NsiI	AATTATGCATAAGCGATAAGGCGGGGCTA
2	PmqsI_R_XhoI	AATTCTCGAGAAGCATGCCGTCGTCT
3	mqsR_F_EcoRI	AATTGAATTCCGGTGGCTGCCTTGGT
4	mqsR_R_HindIII	AATTAAGCTTACGGCCGACCTAGCTAATGT
5	PmqsI_R_XhoI	TTAGCTTCCTTAGAGAGCTCATCTCGAG
6	PmqsI1_F_NsiI	AATGAAAGATGCATGATCTCCAAAAGCAAA
7	PmqsI2_F_NsiI	AATTATGCATATAAGGCGGGGCTACATTAGCTA
8	PmqsI3_F_NsiI	AATTATGCATAGAATTTCTTGCCTACGTGAT
9	PmqsI4_F_NsiI	AATTATGCATCAATTTTCTCGTTGAGGCACGG
10	Prna7_F_NsiI	AAATGCATAGGGAATCAATCGCAGATCTG
11	Prna7_R_XhoI	TTGAGCTCTTTCTTGCCGGCGGTCGC
12	Prna1_F_NsiI	AAATGCATCAAGTCTCTGGGCAAGGG
13	Prna1_R_XhoI	TTGAGCTCCGGCGGGCTTTTGGAGT
14	Prna2_F_NsiI	TTATGCATGTGGTTCGAGCCCATGAG
15	Prna2_R_XhoI	TTGAGCTCGCTTTCGCAGTGGCTACAT
16	Prna3_F_NsiI	ATATGCATAATCGCTATGCCGGCCGATT
17	Prna3_R_XhoI	ATGAGCTCTGCAGCTCTCTAGTGAAGAAC
18	Prna5_F_NsiI	TTATGCATGCTCAGGCATTTTACGACCG
19	Prna5_R_XhoI	AAGAGCTCCAGAAAAGGCGCAAGTGTCC
20	Prna6_F_NsiI	TAATGCATCTCGGCCGTGCGAATCAAA
21	Prna6_R_XhoI	AAGAGCTCCTGGAGAGTAAGCGATGGG
22	PtRNA-leu_F_NsiI	AAATGCATAGATCCGCCGTGCAAACAC
23	PtRNA-leu_R_XhoI	TTGAGCTCTTTCAGATGCCAGACACAGG
24	qseC2_op_F_NsiI	TATAATGCATCGGATTGCCGAAATCCATCTGGATGA
25	qseC2_op_R_NsiI	TATAATGCATCATCGAAGGGATTTCGGCATCAGATGAA
26	qseC2CDS_F_EcoRI	TATAGAATTCTTAAGAAGGAGATATACCATGCCGAAATCCCTTCGATCGC
27	qseC2CDS_R_HindIII	TATAAAGCTTCTATCGGTCACGCAAGCGACGCACACGTGCGAGAA
28	qseC_op_F_NsiI	TATAATGCATTCTTGCCGGTTTGGGGTTTGTTGTGAT
29	qseC_op_R_NsiI	TATAATGCATCACGCCATGACTTCCGTGAGGTCCAT

30	qseCCDS_F_EcoRI	TATAGAATTCTTAAGAAGGAGATATACCATGGACCTCACGGAAGTCATGGCGA
31	qseCCDS_R_HindIII	TATAAAGCTTACTGCGGCGGCCTAATCAACTCACAAGGATCAACGTCA
32	mqsC_F_BamHI	AATTGGATCCCAGGAGACGGCGGAATGAGATCGGAGACAGTCAT
33	mqsC_R_KpnI	AATTGGTACCAATCATGATGCTTAAGCCGCCAGCCTGGCGATA
34	mqsR_F_BamHI	AATTGGATCCATCTCAATGCATGAGGCAACAT
35	mqsR_R_EcoRI	AATTGAATTCACGGCCGACCTAGCTAATGT
36	aidM_F_BamHI	AATTGGATCCTTTGCTGCGCAACTGCA
37	aidM_R_KpnI	AATTGGTACCTCACCCACCCAACGCGAA
38	antiqseC_F_KpnI	AATTGGTACCTTGCAATCCACAAGGCAAACC
39	antiqseC_R_EcoRI	AATTGAATTCATGGACCTCACGGAAGTCATGG
40	RevSacB_F_Gib	ATCCCCGGGTACCGAGCTCGTAAATTGTCACAACGCCGCG
41	RevSacB_R_Gib	GTTAGTTAGGGAATAAGCCGTGCCAAAGAGCTACACCGAC
42	q2_RevSacB_R_Gib	GTGCGTCGCTTGCGTGACCGATAGGTGCCAAAGAGCTACACCG
43	1192_RIC_F_HindIII	AATTAAGCTTAGTCGTGCTCGCACCGGTT
44	1192_RIC_R_KpnI	AATTGGTACCAATCATGATGCTTAAGCCGCCA
45	R7A_RIC_F_HindIII	AATTAAGCTTACTCGTGTTCGCACCGGTT
46	R7A_RIC_R_KpnI	AATTGGTACCGTTGCTTAGGCCGCGA
47	Universal_gBlock_F	CGGGAACTAGTAGCGCACGCTG
48	Universal_gBlock_R	CGGGGTTGAATGCTCGACACTA
49	1192_R_5'_F_HindIII	AATTAAGCTTATAGTCGCGCAGCATGCGCT
50	1192_R_5'_R_BamHI	AATTGGATCCACCGAAATCCGATGTTATGCC
51	1192_C_3'_F_BamHI	AATTGGATCCGAGATCTCGCTTTTCCAGCA
52	1192_C_3'_R_KpnI	AATTGGTACCACCGCTCCTGGTACAAT
53	R7A_R_5'_F_SacI	AATTGAGCTCATAGTCGCGCAGCAT
54	R7A_R_5'_R_XbaI	AATTTCTAGAAAATTGACCCATCGCAACA
55	R7A_C_3'_F_XbaI	AATTTCTAGAATTGCAGGACGACACCA
56	R7A_C_3'_R_	AATTGGATCCATTGCAGATCGAGCTGGT

	BamHI	
57	asqseC_5'_F_SacI	AATTGAGCTCTCACATGCTCATGGATCGTG
58	asqseC_5'_R_BamHI	AATTGGATCCTTTGCAATCCACAAGGCAAAC
59	qseC2_3'_F_BamHI	AATTGGATCCGTGACCGATAGCTTGTGCTG
60	qseC2_3'_R_PstI	AATTCTGCAGCATCAGGCGATCGACAAGTA
61	nptII_F_BamHI	AATTGGATCCCACGCTGCCGCAAGCACT
62	nptII_R_BamHI	AATTGGATCCAGCTAGTCATCCTGTCTCTTGATCAGAT
63	qCstop_5'_F_PstI	AATTCTGCAGATGTCGGGCGGCACTCAT
64	qCstop_5'_R_XbaI	ACGCTTCTAGATCGCCATGACTTCCGT
65	qCstop_3'_F_XbaI	GCGATCTAGAAGCGTCGTTTGCGCCAT
66	qCstop_3'_R_SacI	AATTGAGCTCTGTGCGCAGCATCATCTACAT

qPCR primers and probes	
R7A_attB_F	GCGTGTCGGTGGTTCGA
R7A_attB_R	GCGACATGAAATCGGGTATGA
R7A_attP_F	GACGATGTTGGAGGCCAAAC
R7A_attP_R	CGAAAGAGACTGTCGGGAGAA
R7A_melR_F	CCAACCGCAGCAACAATTC
R7A_melR_R	AATCCGACAACGACAATTTCG
TAQATT	/56-FAM/TCCGCCTCT/ZEN/GGGCA/3IABkFQ/
TAQMR	/56-FAM/TTGACGGCA/ZEN/TGCTC/3IABkFQ/
TAQMR_V2	/56-FAM/TTGACGGCA/ZEN/TGCTCGGCCTT/3IABkFQ/

	Synthesised DNA
MqsC_gBlock	CGGGAACTAGTAGCGCACGCTGCCATGGCGAT GAACCCTCACCCTCAACCCGAACTGCGTAATCA TACTTTGATCGTTACCGTATCCTCAAACGATGGA CGCCCTGTGTTAGATCGCCGTGCGTATGAGTCA TTGGCCCGCACTTTCCATGAAGCGGCAGATAAC GATGAGGTTCGCGTCGTCGTATTGCGTGGTTTG GCAGGTTGTTTTTGTCTTGGGGGTGACTTTAGC GAATTCTTAGATGCCACGAAGCATCAGAAACTG ATTGCTGCGGTAACAGACATGTTTCGTACCCTG GCGACATTCCCGAAACCAATTCTGGCCTGTGTC

	GATGGGGACGCTGTTGGTGTCGGGTGTACCATT TTATTTCACTGCGACATGGTTATTGCATCTAACG AATCCACGTTTCGCGTGCCTTTCGTGGATTTCG GTCTTGTTCCTGATGCGGCTACGTCCATTCTTG CTCCTCAGAAGCTGGGGTATGCTGGTGCTTTCC GCTTCTTCTGCCTGGGCGACACGTTGCATGCGG AGGACGCTCGTGCCCTTGGTCTGGTGGCAGAG ATTGTGCATGACGGAGTAGAGGAAGCGACTTTG GGACGTGCACGTCAACTTGCTAAGAAACCCGTA GCAGCATTGCTTCAGACGCGTGGTCTGCTTAAA GGAAATACGGGTGCGTTATGTGACCGCATCGAC CAGGAAATTTCGCTTTTTCAGCAGGCCTTACAA GATGACACTACGTTGCGTCGCCTTCAACGCATC GCCCGTCTGGCGGCCTGAGCATGAAAGCTTTA GTGTCGAGCATTCAACCCCG
AidM_gBlock	CGGGAACTAGTAGCGCACGCTGCCATGGCGAT GACGATCAGCCAGAAAACCCTTGAAACCTCACA CGGTAAGATTGCTGTGCGTGAAACAAGCGGTCA AGGCACTGCGGTCATGCTGATTCATGGGAATAG CTCCTCGTCAGCCGTGTTCCGCAATCAGCTTGA TGGTCCGCTTGGCGAGCGCTACCATTTAATCGC TCCAGACTTACCCGGACATGGAGCGAGCGGTG ATGCCATCGATCCAGAGCGTTCGTACAGCATGG AAGGGTATGCGGACGCAATGACAGAGGTATTAG GACTTTTGGGAATTGATAAAGCGATTGTGTTCG GGTGGTCTCTGGGGGGCCACATCGGGCTTGAA ATGATCGACCGCTTCCCCGGACTGTTAGGATTG ATGATCACAGGAACACCTCCGGTATCGCCAGAG GAAGTGGGGTCGGGTTTCAAACCCTCTCCCCAC ATGCACCTTGCAGGCCAAGAGACGTTTACAGGC GCGGATGTGGAAGCCTATGCCCGTAGCACGTG TGGCGAACCCTTTGAGCCATTTCTTATTGACACA GTGGCACGTACAGATGGTCGTGCCCGTCGCCT TATGTTCGAGAAATTTGCGGCGGGAACCGGACG CAACCAACGTGAGATCGTGGCTGGTAAGACTCC GCCTATTGCCGTTTTGAATGGGATTGACGAGCC TTTTGTGAATACTGATTTTGTAAGTGCAGTAAAA TTTTCAAATCTGTGGGAAGGTAAAACTCACTTAC TTGACAAATCCGGACACGCCCCATTCTGGGACA GCCCCGACCGTTTTAATCCAGTTCTTGCTCGCT TTCTGGCTTCAGTTGATCGTGCATGAAAGCTTTA GTGTCGAGCATTCAACCCCG
AidM$_{S102G}$_gBlock	CGGGAACTAGTAGCGCACGCTGCCATGGCGAT GACGATCAGCCAGAAAACCCTTGAAACCTCACA CGGTAAGATTGCTGTGCGTGAAACAAGCGGTCA AGGCACTGCGGTCATGCTGATTCATGGGAATAG CTCCTCGTCAGCCGTGTTCCGCAATCAGCTTGA TGGTCCGCTTGGCGAGCGCTACCATTTAATCGC TCCAGACTTACCCGGACATGGAGCGAGCGGTG ATGCCATCGATCCAGAGCGTTCGTACAGCATGG

	AAGGGTATGCGGACGCAATGACAGAGGTATTAG GACTTTTGGGAATTGATAAAGCGATTGTGTTCG GGTGGGGGCTGGGGGGCCACATCGGGCTTGAA ATGATCGACCGCTTCCCCGGACTGTTAGGATTG ATGATCACAGGAACACCTCCGGTATCGCCAGAG GAAGTGGGGTCGGGTTTCAAACCCTCTCCCCAC ATGCACCTTGCAGGCCAAGAGACGTTTACAGGC GCGGATGTGGAAGCCTATGCCCGTAGCACGTG TGGCGAACCCTTTGAGCCATTTCTTATTGACACA GTGGCACGTACAGATGGTCGTGCCCGTCGCCT TATGTTCGAGAAATTTGCGGCGGGAACCGGACG CAACCAACGTGAGATCGTGGCTGGTAAGACTCC GCCTATTGCCGTTTTGAATGGGATTGACGAGCC TTTTGTGAATACTGATTTTGTAAGTGCAGTAAAA TTTTCAAATCTGTGGGAAGGTAAAACTCACTTAC TTGACAAATCCGGACACGCCCCATTCTGGGACA GCCCCGACCGTTTTAATCCAGTTCTTGCTCGCT TTCTGGCTTCAGTTGATCGTGCATGAAAGCTTTA GTGTCGAGCATTCAACCCCG

†Restriction sites are underlined

2.3.4 Quantitative PCR (qPCR)

Buffered TY cultures were inoculated with frozen inocula and incubated with shaking at 28°C. One millilitre samples were taken during log-phase and stationary-phase, and DNA was extracted using PrepMan Ultra Sample Preparation Reagent (ThermoFisher) according to manufacturer's instructions. TaqMan probe-based qPCR assays were used to measure ICE*Ml*Sym^{R7A} excision.

The pJR201 plasmid contains single copies of *attP*, *attB* and a region of the chromosomal *melR* from R7A (85). This plasmid was linearised with BamHI and used as a standard for ICE*Ml*Sym^{R7A} qPCR assays.

TaqMan probe-based reactions comprised 10 µL SensiFAST Probe Lo-Rox Kit (Bioline), 250 nM of probe and 900 nM each primer, 1-5 µL DNA and made up to 20 µL with TE buffer. SYBR Green-based reactions comprised 10 µL SensiFAST SYBR Lo-Rox Kit (Bioline), 500 nM each primer, 1-5 µL DNA and made up to 20 µL with TE buffer.

Primer efficiencies were determined by assaying serially diluted pJR201 or pTHqPS1 DNA and plotting each dilution against log(Ct) and applying the following equation.

$$E=10^{(\frac{1}{slope})}$$

The *attB* and *attP* data was normalised to chromosome copy number based on the *melR* reaction based on the following equation, where E_{att} is the efficiency of either *attP* or *attB* and E_{melR} is the efficiency of the *melR* reaction, $\Delta Ct_{(att)}$ is the difference between the Ct value of the standard *att* and the Ct value of the test *att* and $\Delta Ct_{(melR)}$ is the difference between the Ct value of the standard *melR* and the Ct value of the test *melR*.

$$R_{att}=\frac{(E_{att})^{\Delta Ct_{(att)}}}{(E_{melR})^{\Delta Ct_{(melR)}}}$$

2.3.5 DNA electrophoresis

DNA Samples were mixed with Gel Loading Dye (NEB) to a final concentration of 1x. DNA was separated in 1.5% (w/v) molecular biology grade agarose (Fisher Biotec) gels made with TAE buffer containing 1x GelRed stain (Biotium). The 1 kb DNA Ladder (NEB) or 1 kb plus DNA Ladder (NEB) were used as molecular weight markers. Gels were run at 4 V/cm for approximately one hour and visualised on the ChemiDoc MP Imaging System (Bio-Rad).

2.3.6 Sequencing

Plasmids and mutants constructed in this thesis were sequence confirmed using Sanger sequencing. Purified PCR or plasmid DNA was combined with an appropriate primer as per the facility service guide and sequenced by the Australian Genome Research Facility (AGRF).

2.4 Construction of plasmids

2.4.1 Restriction digestion and dephosphorylation

Plasmid DNA and PCR-amplified DNA was digested using the appropriate buffer for the restriction enzyme being used – either CutSmart Buffer or NEBuffer 1.1 – to a final concentration of 1x and incubated at 37°C for 1-4 hours. Plasmid DNA was dephosphorylated by adding Antarctic Phosphatase Reaction buffer to a final concentration of 1x and 0.25 units/µL of Antarctic Phosphatase directly to digest reactions and subsequently incubated at 37°C for a further 30 minutes.

2.4.2 DNA Purification

Plasmid DNA was electrophoresed on a 1.5% agarose gel and the corresponding band cut from the gel using a scalpel. Plasmid DNA and PCR-amplified DNA was purified using the FavorPrep Gel/PCR Purification Mini Kit (Favorgen) as per manufacturers instructions.

2.4.3 Two-way and three-way ligations

Ligation reactions were carried out using T4 DNA Ligase and T4 DNA ligase buffer in 20 µL reactions. Reactions comprised 1x T4 DNA ligase buffer, 100 µM ATP, 1 µL T4 DNA Ligase, 5 µL plasmid DNA, 5-10 µL digested PCR-amplified DNA and made up to 20 µL with H_2O. Ligation reactions were incubated at 16°C overnight and purified using the FavorPrep Gel/PCR Purification Mini Kit (Favorgen) as per manufacturers instructions.

2.4.4 DNA Assembly

Constructs were designed using the NEBuilder Assembly Tool (https://nebuilder.neb.com/). Primers were designed to include at least 20 bp overlap with the immediate upstream and downstream DNA fragments. NEBuilder HiFi DNA Assembly Master Mix was used as per manufacturers instructions, where reactions comprised 1x HiFi DNA Assembly Master Mix, 1-

5 µL plasmid DNA and 1-5 µL undigested overlapping PCR products. Reactions were incubated in a thermocycler at 50°C for two hours.

2.4.5 Transformations

Electrocompetent *E. coli* cells were prepared by first growing a seeder culture to stationary-phase in modified SOB medium at 37°C. This culture was used to inoculate 500 mL of SOB medium in a 1:200 ratio and incubated with shaking at 37°C until log-phase growth (OD_{600} 0.4-0.6). Cultures kept on ice where possible from this point onwards, and were centrifuged at 10,000 x g for 15 minutes. Cells were subsequently washed in 10% glycerol five times, each time in decreasing volumes. After the glycerol washes, the cell pellet was resuspended in a 1:500 volume of 10% glycerol i.e. a starting culture of 500 mL was resuspended in 1 mL. Suspension was divided into 40 µL aliquots and stored at -80°C.

Electrocompetent cells were thawed on ice for ten minutes. Five microlitres of purified ligation or DNA assembly reaction was added to one 40 µL aliquot of cells and gently mixed by pipetting. The DNA-cell mixture was pipetted into a pre-chilled sterile electroporation cuvette with a 1 mm gap (BTX). Electroporation was performed using a GenePulser II (BioRad) at 1.8 kV, 25 µF capacitance and 200 Ω resistance. One millilitre of LB medium (supplemented with ala for *E. coli* ST18) was immediately added to the cuvette and the mixture was transferred to a sterile tube and incubated with shaking at 37°C for two hours. The recovered cell culture was lawn inoculated on LB media containing the appropriate antibiotics for selection (and ala for *E. coli* ST18) and incubated overnight at 37°C.

2.5 Biparental matings

2.5.1 Transfer of plasmids from *E. coli* to *Mesorhizobium*

Plasmids were electroporated into *E. coli* strain ST18 (auxotrophic for ala) and grown in TY medium supplemented with ala overnight at 37°C. *Mesorhizobium* recipient strains were grown in TY medium for 64-72 hours at 28°C. One millilitre of *Mesorhizobium* culture was pelleted at 15,000 rpm for two minutes and the supernatant discarded. Cells were resuspended in 1 mL of donor *E. coli* culture, centrifuged again and supernatant discarded. Mixed cell pellets were resuspended in 30 µL of liquid TY medium and spotted on the surface of TY agar supplemented with ala and incubated overnight at 28°C. Spots were scraped off the plate using a sterile wire loop or sterile pipette tip and resuspended in 300 µL of liquid G/RDM medium, and subsequently streaked for single colonies or serially diluted and lawn inoculated on selective G/RDM agar. Plates were incubated at 28°C for five days.

2.5.2 Conjugations to measure ICE*Ml*Sym^{R7A} transfer frequency

All donor strains harboured the pFAJ1700 plasmid for tetracycline selection. *Mesorhizobium* donor and recipient strains were grown in TY medium for 64-72 hours at 28°C. Five hundred µL of recipient strains were centrifuged at 15,000 rpm for two minutes and the supernatant discarded. Pellets were resuspended in 500 µL of the donor strains culture, centrifuged at 15,000 rpm for two minutes and supernatant discarded. Five hundred µL of the donor strain culture only was centrifuged and supernatant discarded. Mixed donor-recipient pellets and donor only pellets were resuspended in 100 µL of TY medium, 30 µL was spotted onto TY agar and incubated overnight at 28°C. Spots were scraped off the plate using a sterile wire loop or sterile pipette tip, resuspended in 1 mL of liquid G/RDM medium without vitamins added, and serially diluted. Donor-recipient mix dilutions were plated on selective G/RDM agar without vitamins added. Donor only dilutions were plated on non-selective G/RDM agar without vitamins added. Plates were incubated at 28°C for five days.

2.6 Mutant construction

The suicide vector pJQ200SK carries a gentamicin resistance marker and the *sacB* gene which produces a toxic byproduct when sucrose is metabolised. A two-step allelic replacement using pJQ200SK was carried out to replace genes with mutant alleles containing in-frame markerless deletions. Constructs were designed to incorporate approximately 700 bp – 1 kb of upstream and downstream flanking regions of the target gene or region. Once plasmids were cloned and sequenced, they were first moved into *E. coli* ST18 by electroporation and subsequently mated into the *Mesorhizobium* strain as described in section 2.5.1, with mating mixtures serially diluted and lawn inoculated onto G/RDM agar containing gentamicin to select for homologous recombination and incorporation of the pJQ200SK construct. After 5 days incubation, Gm^R colonies were single-colony purified on selective G/RDM and then one colony used to inoculate a 10 mL non-selective TY culture and grown for 64 hours at 28°C. The culture was then serially diluted, lawn inoculated onto RDM agar containing 5% sucrose as a sole carbon source and incubated at 28°C for 5 days. Colonies able to grow on sucrose have theoretically lost the pJQ200SK construct and left behind either the wild-type allele or the mutant allele. Therefore, colonies growing on sucrose were PCR-screened using primers flanking the mutated region, and PCR products from isolates with the apparent deletion mutation was sequenced. Mutants were finally confirmed to have lost the pJQ200SK plasmid by testing gentamicin sensitivity.

2.7 Frozen Inocula

Mesorhizobium cultures to be assayed were inoculated using frozen inocula. Frozen inocula were prepared by growing *Mesorhizobium* strains at 28°C to stationary phase in the medium to be used in the subsequent assay (TY or G/RDM). Sterile glycerol was added to the stationary phase culture to a final concentration of 15% (v/v) and vortexed vigorously. Aliquots of 500 µL were stored at -80°C.

2.8 Growth Curves

Growth curves were carried out in 250 mL Erlenmeyer flasks. Fifty mL broth cultures of either TY or G/RDM were inoculated with 500 µL of frozen inocula of the appropriate media (see section 2.7) and incubated with shaking at 28°C. Samples were taken at various time points over 64-72 hours, serially diluted and lawn inoculated on G/RDM agar. Plates were incubated at 28°C for five days.

2.9 AHL extraction

Cultures were grown to stationary phase in 500 mL of G/RDM medium. Cells were pelleted by centrifugation at 10,000 x g for 20 minutes and the supernatant collected. Supernatants were mixed with an equal volume of ethyl acetate acidified with 0.1% acetic acid in a separating funnel. The organic layer was collected, and the inorganic layer was again mixed with an equal volume of acidified ethyl acetate, separated, and the organic layer collected. AHLs were extracted to dryness by rotary evaporation with cooling water at 15°C, water bath at 50°C and vacuum at 240 mbar. Once dry, extracts were resuspended in 50% methanol for mass spectrometry or 100% DMSO for bioassays.

2.10 Bioassays

2.10.1 *Chromobacterium violaceum* CV026 bioassays

Mesorhizobium strains to be analysed for C4-C8 AHL production were grown in liquid buffered TY cultures at 28°C for 64-72 hours. Cells were pelleted at 15,000 rpm for 5 minutes and supernatants collected.

C. violaceum CV026 was grown overnight in liquid TY medium at 28°C. Molten semi-solid (0.4%) TY agar was cooled to ~35-40°C and inoculated with the stationary-phase CV026 culture in a 1:100 ratio. The inoculated semi-solid agar was overlayed on top of a base layer of 1.6% TY agar. A cork-borer was

used to bore wells into the semi-solid layer, and 30 µL of the collected supernatants from test strains were added to each well. The bioassay plate was incubated at 28°C for 48 hours.

2.10.2 *Mesorhizobium* LacZ bioassays

Mesorhizobium strains to be analysed for AHL production were grown in liquid buffered TY cultures at 28°C for 64-72 hours. Cells were pelleted at 15,000 rpm for 5 minutes and supernatants collected.

The appropriate *Mesorhizobium* reporter strain was grown in buffered TY at 28°C for 64-72 hours. Molten semi-solid (0.4%) TY agar was cooled to approximately 35-40°C and inoculated with the stationary-phase reporter strain culture in a 1:100 ratio. The semi-solid agar was also supplemented with IPTG and Xgal. The inoculated semi-solid agar was overlayed on top of a base layer of 1.6% TY agar. A cork-borer was used to bore wells into the semi-solid layer, and 30 µL of the collected supernatants from test strains were added to each well. The bioassay plate was incubated at 28°C for 3-4 days.

2.11 β-galactosidase assays

Recipes for buffers and solutions are found in Appendix 1. TY or G/RDM broth cultures were inoculated with frozen inocula (see section 2.7) in a 1:200 ratio. Cultures were incubated at 28°C for 64-72 hours. One hundred µL of each culture was added to a 96-well plate (master plate) and frozen at -80°C for at least 12 hours. Expression from promoter-*lacZ* fusions was measured as previously described (126). Briefly, 10 µL from each well in the master plate was transferred to a new 96-well plate (assay plate). The assay plate was freeze-thawed at -80°C and 37°C three times and 100 µL of working solution was added to each well. Fluorescence was measured every minute for 40 minutes at 360 nm excitation and 460 nm emission on the Ensight Multimode plate reader (Perkin Elmer). Relative fluorescence units were normalised to

either colony forming units by plating serial dilutions of each sample on G/RDM, or OD_{600} of the master plate.

2.12 Protein purification methods

2.12.1 Cell culture and protein induction

Protein expression vectors were introduced into either *E. coli* BL21(de3)pLysS or *E. coli* NiCo21(de3). Seeder cultures were grown in liquid LB medium at 37°C with shaking overnight and used to inoculate 1 L LB cultures in 5L Erlenmeyer flasks. Cultures were incubated at 27°C with shaking until OD_{600} 0.4-0.6 or approximately 4-5 hours. Protein expression was induced with 200 µM isopropyl β- d-1-thiogalactopyranoside (IPTG) and the temperature lowered to 18°C. Cultures were left to grow with shaking at the reduced temperature overnight. Cells were pelleted at 10,000 x g for 45 minutes at 4°C.

2.12.2 Cell lysis

Cells were resuspended in wash buffer (Appendix 1) and lysed under high pressure (approximate 2000 psi) using the Continuous Flow Cell Disruptor (Constant Systems). The clarified lysate was centrifuged at 20,000 x g for 45 minutes at 4°C to separate the soluble and insoluble fractions and the supernatant was filtered with Millex 0.22 µm syringe filters (Merck).

2.12.3 Protein purification and analytical gel filtration

Recombinant proteins were purified using immobilised metal (nickel) affinity chromatography. High performance 5 mL HisTrap nickel affinity columns (Cytiva) were used to extract hexahistidine tagged protein on the ÄKTA pure chromatography system (Cytiva). Protein was eluted from the column with increasing concentrations of imidazole (elution buffer, Appendix 1). Recombinant protein was further purified with size exclusion chromatography (SEC) using the HiLoad 16/600 Superdex 200 PG SEC column (Cytiva) on the ÄKTA pure chromatography system (Cytiva). SEC buffer is described in

Appendix 1. Analytical gel filtration (analytical SEC) was carried out using the Superdex 200 Increase 10/300 GL SEC column (Cytiva) on the ÄKTA pure chromatography system (Cytiva). UV absorbance at 230, 260 and 280 nm were measured using Unicorn 5.1 software (Cytiva).

2.12.4 Sodium dodecyl-sulphate polyacrylamide gel electrophoresis

Protein samples were visualised with sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Ten microlitres of protein was mixed with 3 µL of 4x SDS loading buffer (see Appendix 1) and boiled at 95°C for 10 minutes. Samples were loaded on a 4-15% mini-PROTEAN TGX pre-cast protein gel (Bio-Rad) alongside the 10-200 kDa unstained protein standard (NEB) and electrophoresed at 10-25 V/cm in the mini-PROTEAN tetra vertical electrophoresis cell (Bio-Rad) in 1x SDS running buffer (Appendix 1). Gels were stained using 2.5 g/L Coomassie Brilliant Blue G250 in a 10% (v/v) acetic acid and 30 % (v/v) methanol solution and de-stained using a 10% (v/v) acetic acid and 40% (v/v) methanol solution. Gels were visualised on the ChemiDoc MP Imaging System (Bio-Rad).

2.12.5 Measurement of protein concentration

A NanoDrop 1000 spectrophotometer (ThermoFisher Scientific) was used to measure absorbence at 280 nm. Theoretical protein extinction coefficients and molecular weights were calculated using ProtParam (Expassy) at https://web.expasy.org/protparam/ (127).

2.12.6 Protein crystallisation trials

Crystallisation trials were carried out using vapor diffusion in a controlled temperature room of 20°C. Primary screening conditions included the commercially available matrices Crystal Screen 1, Crystal Screen 2 and Index (Hampton Research). The Phoenix liquid handling robot (Art Robbins Scientific) was used to load Intelli-Plate 96-3 low volume reservoir sitting drop crystallisation trays (Hampton Research).

2.13 Glasshouse procedures

Cicer arietinum cv. Striker was grown in sterilised free-draining sand in a temperature-controlled glasshouse at 22°C as previously described (128). Briefly, seeds were washed in 70% ethanol for one minute, 4% NaOCl for three minutes and rinsed in deionized water. This was repeated six times. Seeds were placed on the surface of 0.9% water agar and kept at room temperature in the dark for three days to begin germination. One litre pots were prepared with coarse sand, adjusted to pH 6.5 with 5 g/L $Fe_2(SO_4)_3$ and steam sterilised for four hours. Sand in the pots were washed with boiling water to leach residual nitrogen. Pots were sown with two seeds each, and once shoots emerged, thinned to one plant per pot. Mesorhizobium ciceri CC1192 strains were grown to stationary phase in liquid TY cultures and diluted to OD_{600} of 1 in saline. One mililiter of cell suspension was added on top of each sown seed and 1 mL was added on top of the whole sand surface (3 mL total per pot). Sterile H_2O was used instead of cell cultures for uninoculated control plants. A sterile PVC pipe was pushed into the sand in each pot for feeding and watering. Pots were then covered with sterile beads to protect from airborne contamination. Pots were randomised on the benches so treatments were not grouped together, to address the possibility of subtle temperature and sunlight differences. Plants were maintained by watering three times a week and feeding with a nutrient solution once a week (recipe in Appendix 1). Uninoculated, nitrogen positive plants were fed 5 mL of 10 g/L KNO_3 once per week. Plants were harvested after eight weeks.

Acetylene reduction assays were carried out as previously described (128). Briefly, plants were harvested (including roots) and placed inside 1 L schott bottles – one plant per bottle – and capped with a silicone pierceable lid. A syringe and needle was used to add acetylene gas to the bottle through the silicone lid at a final concentration of 2% (v/v) and bottles were left at room temperature. Every hour 1 mL samples were extracted through the silicone lid with a needle and syringe. Gas chromatography was carried out on the Varian

CP-3800 gas chromatograph with a Flame Ionisation Detector (FID) with a RT-U-BOND 15 m x 0.32 mm internal diameter x 10 µm column (Restek). Parameters were as follows; oven temperature 40°C, carrier gas H_2 at 1 mL/min and 5:1 split ratio injection. FID was set at 190°C, H_2 gas at 30 mL/min and 300 ml/min air. The calculations used to determine the amount of acetylene reduced to ethylene is described in reference (128).

2.14 Liquid chromatography and mass spectrometry

Liquid chromatography was carried out as previously described (79) on the 1290 Infinity II LC system (Agilent) and liquid chromatography – mass spectrometry was carried out on the 6495 triple quadrupole LC-MS system (Agilent).

Chapter 3

Characterising the *Mesorhizobium* quorum sensing (*mqs*) locus

3.1 Introduction

Mesorhizobium species harbour a variety of LuxRI-type QS systems. The most well studied are TraR and TraI proteins of *M. japonicum* R7A and *M. ciceri* bv. *biserrulae* WSM1271, which regulate excision and transfer of ICE*Ml*Sym^{R7A} and ICE*Mc*Sym^{1271}, respectively (80, 82). TraR and TraI are distant homologues of the *A. tumefaciens* pTi and *R. leguminosarum* pRL1JI encoded TraR and TraI, which control plasmid conjugative transfer. Close relatives of the *traR* and *traI1* genes are carried by numerous symbiosis ICEs related to ICE*Ml*Sym^{R7A} and ICE*Mc*Sym^{1271}, while other more distantly-related ICEs and plasmids in the genus *Mesorhizobium* carry distinct putative LuxRI-family regulators (129). Aside from these mobile LuxRI systems, relatively little work has been done to characterise chromosomally encoded LuxRI systems in the genus. In this work, we identified a single LuxRI-family system that is largely conserved throughout the genus *Mesorhizobium*, which we named the "*Mesorhizobium* quorum sensing" (MQS) system and named the *luxRI* homologues *mqsR* and *mqsI*. A third gene encoding a crotonase-family protein found to be required for the MQS system and located directly downstream of *mqsI* was named *mqsC*.

The *mqsI* gene was previously identified in a mutagenesis study of the *Lotus* symbiont *M. loti* NZP2213 under the name *mrlI1*. In this strain, three LuxI-family AHL synthases were identified and named *mrlI1*, *mrlI2* and *mrlI3* (78). Homologues of these genes were also identified in the *M. loti* MAFF303099 genome, *mlr5638*, *mlr6385* and *mlr6103* – which correspond to the *mqsI*, *traI1* and *traI2* genes of *M. japonicum* R7A (previously *M. loti* R7A). Their study revealed that the enzyme encoded by *mrlI2* produces 3-oxo-C6-HSL, C8-HSL and C10-HSL, which is somewhat consistent with the AHLs produced by the ICE*Ml*Sym^{R7A} TraI1 enzyme in *M. japonicum* R7A. The ICE*Ml*Sym^{R7A} *traI2* gene is likely a pseudogene (82, 129) and consistent with this, no AHLs were identified as produced by the *mrlI3* gene product (78). MrlI1 was found to produce C12-HSL when overexpressed in *E. coli*. Interestingly, the authors discovered that mutations in any of these three genes, including the likely pseudogene *mrlI3*, reduced efficiency of NZP 2213 nodule formation on *Lotus*

corniculatus (78). No phenotypic differences in nodulation of *L. corniculatus* have been observed for *M. japonicum* R7A strains carrying deletions of *traI1* or *traI2* (130).

Mass spectrometry (liquid chromatography coupled to mass spectrometry (LC-MS/MS)) analysis of AHLs produced by *M. japonicum* R7A has found that *traI1* is essential for production of all shorter-chain AHLs in this strain. In wild-type R7A TraI1 was essential for production of C4-HSL and 3-oxo-C6-HSL. R7A carrying an extra copy of the traR gene on a plasmid is further induced for TraI1-dependent AHLs and produce 3-oxo-C4-HSL, C6-HSL, 3-OH-C6-HSL, C8-HSL, 3-oxo-C8-HSL and 3-OH-C8-HSL (82). All strains examined also produced 3-oxo-C12-HSL and 3-OH-C10-HSL, including the ICE*Ml*Sym^{R7A}-cured derivative R7ANS, which lacks *traI1* and *traI2* (130), suggesting these longer-chain AHLs require a chromosomally-encoded AHL synthase.

Unsaturated long-chain AHLs have been reported in a *Mesorhizobium sp*. isolated from a marine sponge. LC-MS and nuclear magnetic resonance (NMR) analysis identified the novel AHLs 5-cis-3-oxo-C12-HSL and 5-cis-C12-HSL (14). The marine sponge *Mesorhizobium sp*. produced very large amounts of these unsaturated AHLs at an estimated 5 mg per litre of growth medium, however the genes responsible for production of these compounds were not identified. AHL molecules with unsaturated side chains have been characterised previously, however, they are far less common than their saturated counterparts. The *cinI* gene of the CinRI QS system of *Rhizobium leguminosarum* is required for production of 3-OH-7-cis-C12-HSL, which was originally named bacteriocin small for its ability to inhibit the growth of some strains (62, 63).

In this chapter, we aimed to identify the AHLs produced by MqsRIC. We demonstrated that the *mqsRIC* locus is indeed responsible for production of a diffusible product that is able to activate the *mqsI* promoter when supplied

exogenously and this activation was dependent on the *mqsR* gene. We also demonstrated that both *mqsI* and *mqsC* are necessary for production of the active compound, making it the first example of a LuxRI system requiring two AHL-synthase genes. We initially suspected that this compound was either 3-oxo-5-cis-C12-HSL or 5-cis-C12-HSL and indeed detected a molecule resembling 5-cis-C12-HSL in *M. japonicum* R7ANS extracts. However, this molecule was also present in strains deleted for *mqsRIC*. More recently, the *mqsRIC* locus in *M. japonicum* MAFF 303099 was interrogated by Suo et al. (79) who identified 2,4-trans-C12-HSL as the major product. We confirmed here that 2,4-trans-C12-HSL is indeed produced by R7A and is most likely the major product produced by the *mqsRIC* locus. Synthetic 2,4-trans-C12-HSL activated the *mqsI* promoter more strongly than either 5-cis-C12-HSL or C12-HSL and production of 2,4-trans-C12-HSL by R7A was dependent on the presence of both *mqsI* and *mqsC* genes.

3.2 Results

3.2.1 Identification of a conserved quorum-sensing locus

A BLASTp search of the NCBI non-redundant protein sequences database using MqsI from *M. japonicum* R7A as a query revealed a 95% and 97% amino acid identity to MrlI1 of *M. loti* NZP2213 (78) and MqsI of *M. ciceri* WSM1271 (131) respectively. A LuxR-family regulator encoded by the *mqsR* gene was identified upstream of *mqsI* in R7A, and BLASTp searches within a database of 41 closed and completed *Mesorhizobium* genome sequences described by Colombi et al. (129) revealed the MqsRI region to be conserved in 38 of the 41 genome sequences. Further investigation of these *mqsRI* genes revealed a conserved crotonase-domain-family protein was encoded by *mqsC*, located downstream of *mqsI*.

Orthologous clusters containing the *mqsRIC* genes were identified across an extended database of 214 symbiotic and non-symbiotic *Mesorhizobium* genomes (106) using Proteinortho (132) and BLASTp. Where Proteinortho did not identify orthologues in a strain, the strain was analysed using BLASTp and if present, MqsRIC-encoding genes were manually identified across multiple sequencing contigs. Of the 214 genomes analysed, 207 were found to encode MqsR and MqsI, and 205 of those were found to also encode MqsC. Interestingly, in 206 of the 207 strains containing *mqsRI*, this locus was located upstream of a convergently encoded polyphosphate kinase 2 gene (*ppk2*). This genomic arrangement is depicted in Figure 3.1A.

Multiple sequence alignments of the 207 MqsR, 207 MqsI and 205 MqsC amino acid sequences were carried out using T-coffee (133). Alignment of the MqsR homologues revealed high amino acid conservation, especially in the C-terminus (Figure 3.1B) containing the predicted helix-turn-helix DNA-binding domain (134, 135). Alignments of MqsI and MqsC homologues also exhibited high amino acid conservation throughout. In order to visualise residue changes amongst the high conservation, maximum-likelihood phylogenetic trees were

constructed using MEGA (136) for each of MqsR, MqsI and MqsC. Seven distinct representative sequences for each protein were selected, and multiple sequence alignments carried out again using T-coffee. (Figure 3.1B-D). Phylogenetic trees can be found in Appendix 2.

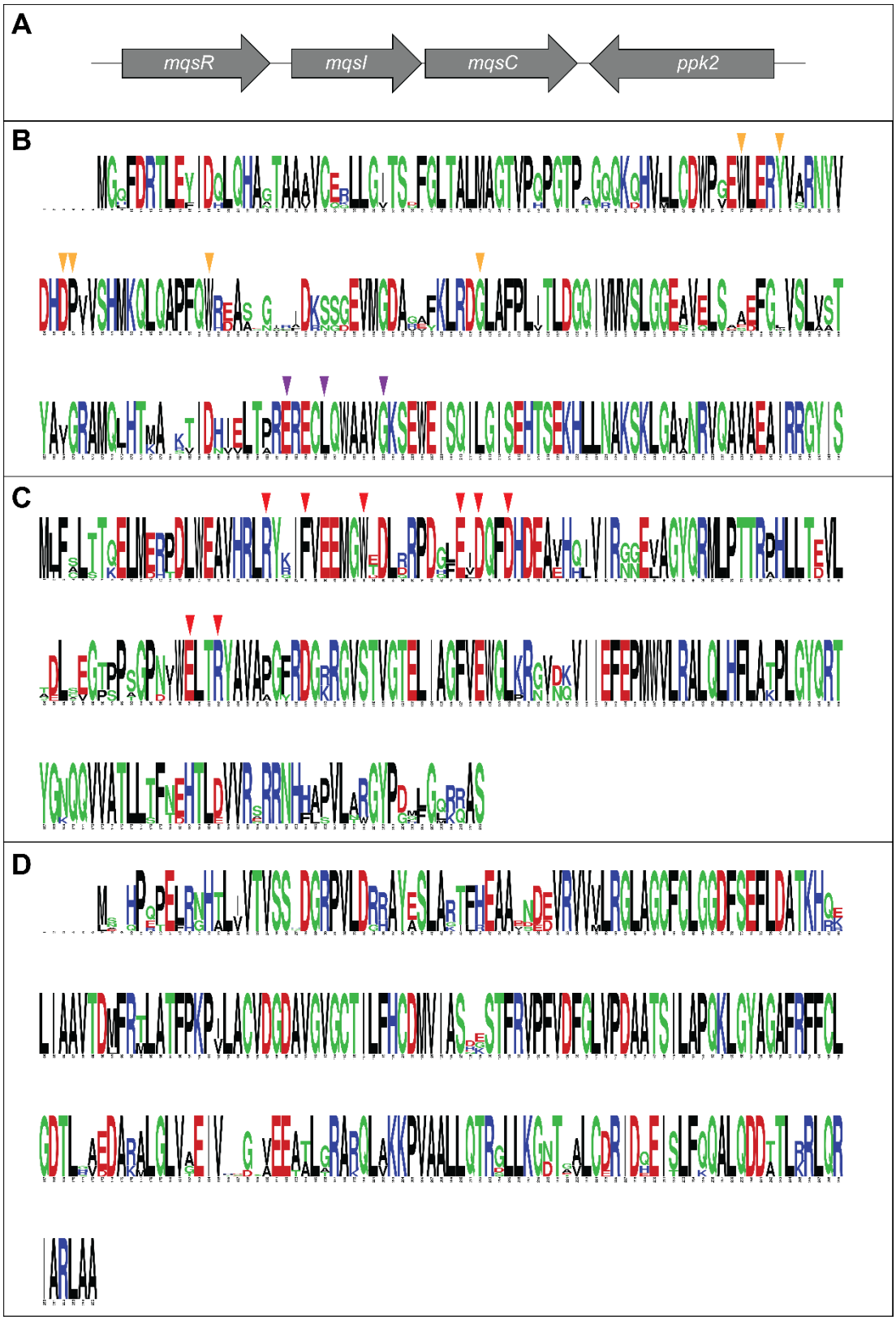


Figure 3.1 MqsRIC genome map and sequence conservation (A) Genomic arrangement of the *mqsRIC* locus alongside the conserved *ppk2* gene. (B-D) Graphical representation (137) of the multiple sequence alignment of seven distinct sequences of (B) MqsR, (C) MqsI and (D) MqsC. (B) Yellow and purple arrows denote conserved residues in the ligand-binding domain and DNA-

binding domain respectively, which are conserved in almost all LuxR-type proteins (138). (C) Red arrows denote conserved residues present in all LuxI-type AHL synthases (138).

3.2.2 MqsR is required for activation of the cognate *mqsI* promoter

LuxR-family transcriptional regulators generally require the binding of an AHL molecule produced by their cognate LuxI-family AHL-synthase for dimerisation and formation of the LuxR-AHL complex that can bind a target DNA sequence (71, 139, 140). Often, the LuxR binding site is located directly upstream of the AHL-synthase gene. To measure expression from the *mqsI* promoter, a 126-bp region upstream of *mqsI* was cloned upstream of the promoterless *lacZ* gene on pSDz (plasmid map in Appendix 3) (83). The resulting plasmid pPmqsI was introduced into R7ANS and into an R7ANS derivative carrying a markerless in-frame deletion of *mqsR*. Samples were taken at eight-hour intervals and analysed for β-galactosidase activity to quantify P*mqsI*-derived *lacZ* expression throughout growth. β-galactosidase activity from P*mqsI* was observed throughout growth in R7ANS, but only very low β-galactosidase activity was observed in the R7ANSΔ*mqsR* strain, confirming that *mqsR* was required for activation of the *mqsI* promoter (Figure 3.2A). Interestingly, while the absolute expression from P*mqsI* in R7ANS populations increased in cultures with population density, the relative expression from P*mqsI* per colony forming unit decreased during entry into stationary-phase (Figure 3.2B).

3.2.3 Construction of a biosensor strain to detect *mqsRIC*-derived AHLs in supernatants

M. japonicum R7ANS, which has been cured of ICE*Ml*Sym^{R7A} (85), encodes MqsRIC as its sole LuxR-LuxI QS system. An *mqsRIC* mutant R7ANSΔ*mqsRIC* was constructed in R7ANS. The pSDz-based vector pPmqsI, described above also carries an IPTG-inducible *lac* promoter and downstream cloning cassette that allows inducible expression of transcription factors on the same plasmid. The *mqsR* coding sequence was cloned downstream of this IPTG-inducible *lac* promoter on pPmqsI, to provide inducible *mqsR* expression. The resulting plasmid pPmqsIzR was introduced into

R7ANSΔ*mqsRIC* to create a strain able to detect the MqsR-dependent activation of the P*mqsI* promoter in the presence of exogenously supplied AHLs and IPTG. This 'reporter' strain R7ANSΔ*mqsRIC*(pPmqsIzR) was used throughout this thesis to detect AHLs able to activate MqsR.

To test the functionality of the constructed biosensor strain, supernatant extracts from the R7ANS and R7ANSΔ*mqsR* growth curves outlined in section 3.2.2 were used in bioassays with R7ANSΔ*mqsRIC*(pPmqsIzR) as the biosensor. We expected that since the R7ANS strain exhibited expression from the *mqsI* promoter, that any AHLs produced by this locus would be present in the supernatants. Since there was little activity from the *mqsI* promoter in R7ANSΔ*mqsR*, we suspected there would be an absence of AHLs present in the supernatant. The supernatant extracts were pipetted into wells made in an agar plate overlaid with soft agar containing IPTG and X-gal mixed with an overnight culture of the biosensor strain R7ANSΔ*mqsRIC*(pPmqsIzR), so that any *lacZ* expression could be visualised as a blue colour in the agar. R7ANS supernatants produced distinct zones of blue colour in bioassay plates and the diameter of the zone was larger in supernatants from stationary-phase cultures, consistent with a population-level increase in *mqsI* expression in these samples. The supernatants from R7ANSΔ*mqsR* were unable to induce any blue colour in the R7ANSΔ*mqsRIC*(pPmqsIzR) biosensor strain (Figure 3.2C).

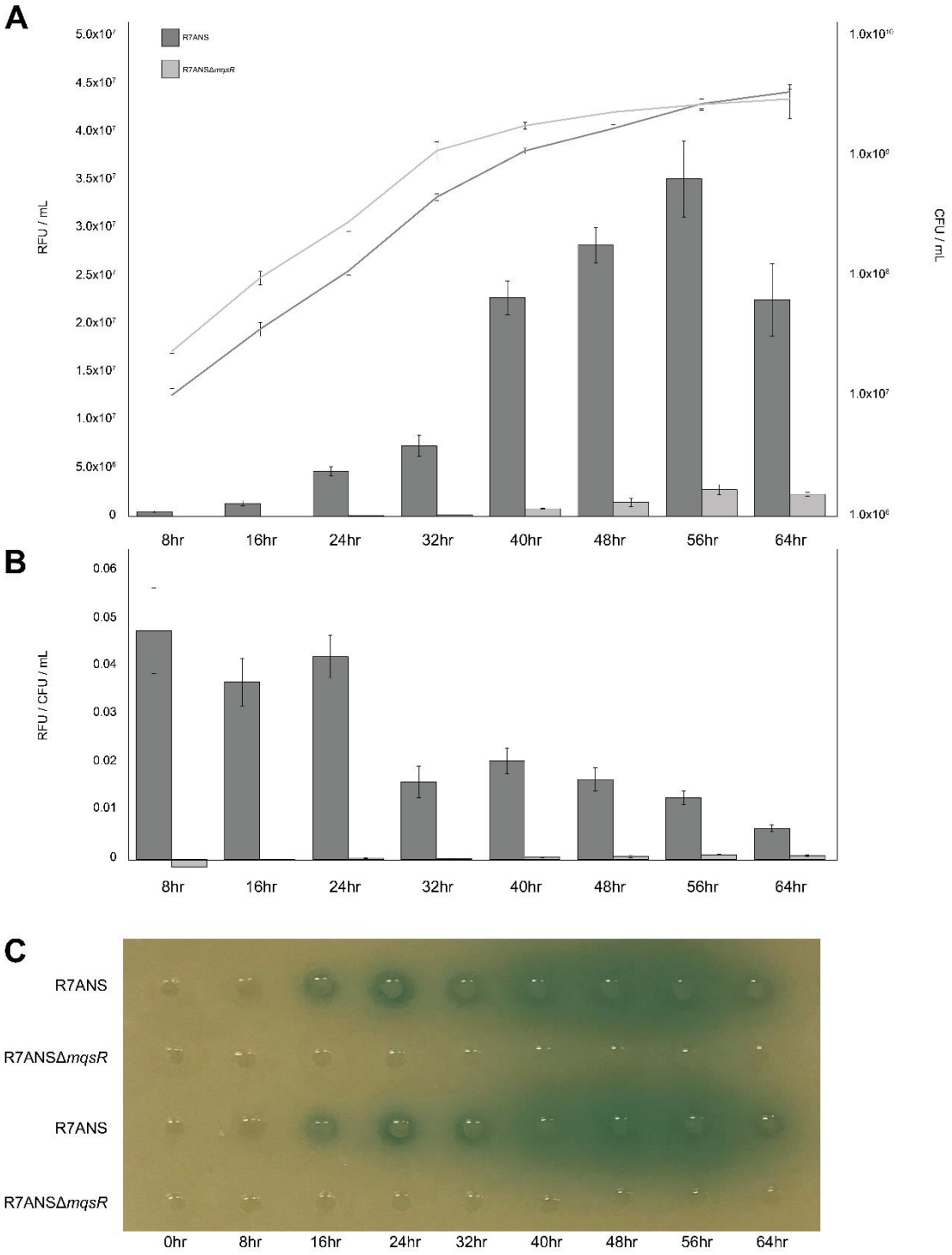


Figure 3.2 β-galactosidase activity of R7ANS and R7ANSΔ*mqsR* harbouring pPmqsI (A,B) β-galactosidase was measured as previously described (126) using the fluorescent substrate 4-methylumbelliferyl-β-D-galactopyranoside (MUG). Bars represent the mean of three biological replicates and error bars represent the standard deviation of the mean. Growth curves are shown as a continuous line and error bars represent the standard deviation of the mean. R7ANS is shown in dark grey and R7ANSΔ*mqsR* is shown in light grey. (C) The same samples as shown in A-B were also added to a reporter assay where R7ANSΔ*mqsRIC* (pPmqsIzR) was used as a biosensor to detect AHLs inducing the P*mqsI-lacZ* fusion.

3.2.4 Both MqsI and MqsC are essential for MqsR activation of P*mqsI*

To test whether the *mqsI* and *mqsC* genes were involved in producing the supernatant factor required for MqsR-dependent activation of the *mqsI* promoter, *mqsI* and *mqsC* were cloned both separately and together in distinct constructs, in which each gene or genes were placed downstream of the strong constitutive *nptII* promoter. Each of the three resulting constructs were introduced into the R7ANSΔ*mqsRIC* background. Supernatants from liquid cultures of these strains were added to a bioassay to detect activity of MqsR-based induction of P*mqsI*. Supernatants from strains constitutively expressing *mqsI* or *mqsC* alone were not capable of inducing *lacZ* expression in the reporter, however, supernatants taken from the strain constitutively expressing *mqsI* and *mqsC* together induced *lacZ* expression in the biosensor strain (Figure 3.3). This result suggested that both MqsI and MqsC were required for the production of the molecule that induces MqsR-dependent activation of the *mqsI* promoter.

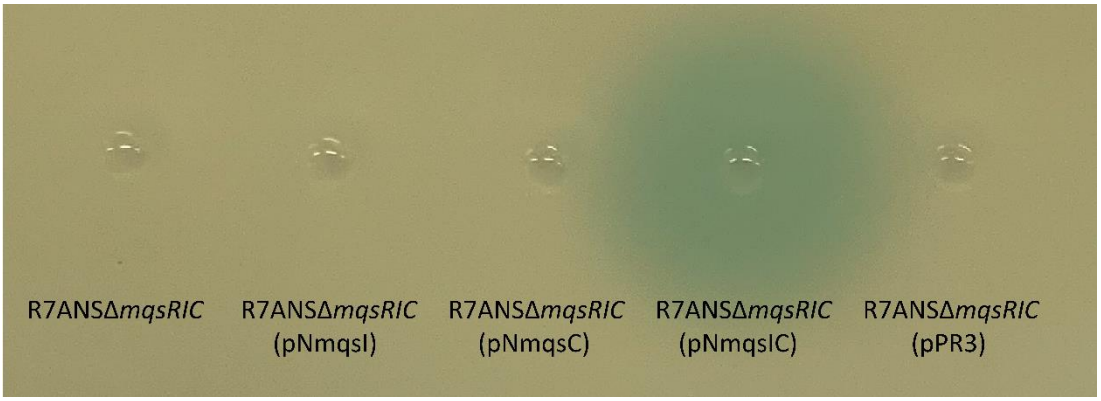


Figure 3.3 Both *mqsI* and *mqsC* are required for MqsR activation. R7ANSΔ*mqsRIC*(pPmqsIzR) was used as a biosensor in well-diffusion assays to detect AHLs able to interact with MqsR and induce P*mqsI*-driven *lacZ* expression.

3.2.5 Defining a minimal *mqsI* promoter region

The region immediately upstream of *mqsI* was taken from 207 sequences and aligned using T-coffee (133). The consensus sequence was manually scanned for any inverted repeats (IRs) which could be potential *lux*-box-like binding sites. Two IR regions were identified (Figure 3.4A). To narrow down the

essential region of the *mqsI* promoter and identify the likely '*mqs*-box' (named here after *lux* and *tra*-boxes of LuxR and TraR regulators), truncated regions upstream of *mqsI* were cloned into pSDz similarly to the construction of pPmqsI. Four plasmids were constructed with PCR-amplified fragments from 171, 122, 81 and 56 bp upstream of the mqsI translational start site to create pPmqsI1, pPmqsI2, pPmqsI3 and pPmqsI4 respectively. pPmqsI3 was truncated so that it cut off the first IR, and pPmqsI4 was truncated to cut off the first and second IR (Figure 3.4A).

The four constructs carrying the promoter fusions and the empty vector were introduced into R7A and R7AΔ*mqsR*. Expression of *lacZ* by each promoter region was measured using β-galactosidase assays. pPmqsI1-3 all exhibited high *lacZ* expression compared to the empty vector control in the R7A background, however, expression was abolished in the pPmqsI4 construct (Figure 3.4B). This suggests that the sequence upstream of P*mqsI*3 is not essential but the second IR region proximal to the *mqsI* gene is required for activation of P*mqsI*. The same experiment in the R7AΔ*mqsR* background did not show any differential *lacZ* expression by any of the tested truncated promoter regions compared to the empty vector, confirming MqsR is required for any P*mqsI* activity (Figure 3.4B).

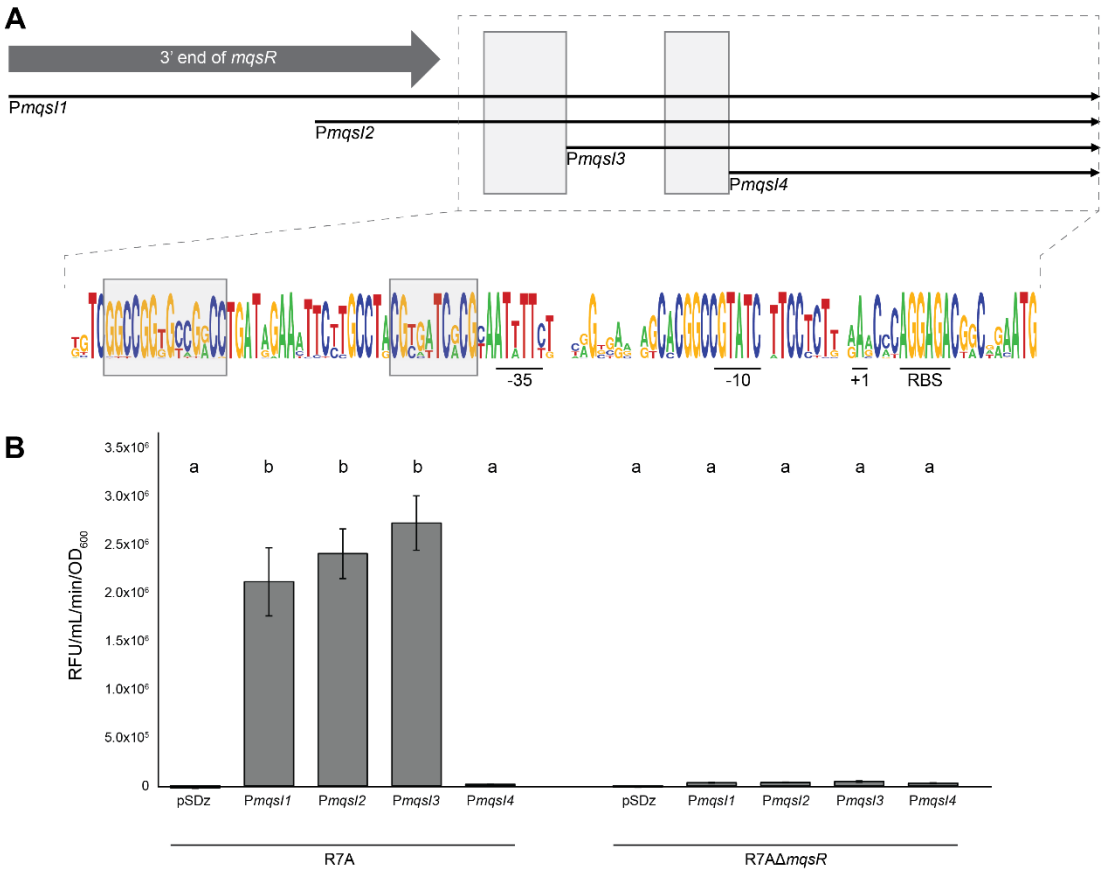


Figure 3.4 Expression from PmqsI truncations (A) Depiction of P*mqsI*1, P*mqsI*2, P*mqsI*3 and P*mqsI*4. Truncated regions upstream of *mqsI* were cloned into pSDz upstream of a promoterless *lacZ*. The 3' end of mqsR is displayed as a grey block arrow. Sequence logo shows conservation between promoter regions of 208 *mqsI* homologues. Identified inverted repeat sequences are shown in boxes. The -10 and -35 elements along with the transcriptional start site (+1) were predicted based promoters of other rhizobia (82, 141). P*mqsI*1, P*mqsI*2, P*mqsI*3 and P*mqsI*4 regions are shown as arrows above the DNA sequence. (B) β-galactosidase activity of R7A and R7AΔ*mqsR* harbouring either pSDz, pPmqsI1, pPmqsI2, pPmqsI3 or pPmqsI4. β-galactosidase was measured as previously described (126) using MUG. Bars represent the mean of three biological replicates and error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each strain. Values that are not statistically significant from each other are grouped by the same letter (a, b).

3.2.6 *M. japonicum* R7ANS supernatants contain molecules resembling 5-cis-C12-HSL and 2,4-trans-C12-HSL

Previous studies have identified long chain unsaturated AHLs in marine *Mesorhizobium* sp., namely 3-oxo-5-cis-C12-HSL and 5-cis-C12-HSL (14) and previous analysis of *M. loti* suggests C12-HSL is produced in strains lacking

TraI1 (78). More recently, Suo et al. (79) revealed *M. japonicum* MAFF 303099 produces a C12 AHL molecule with an acyl chain containing two *trans* double bonds – 2,4-trans-C12-HSL. Since *mqsRIC* is conserved across mesorhizobia, we hypothesised that MqsIC may synthesise one or more of these AHL species. AHLs were extracted from R7ANS spent supernatants in two volumes of acidified ethyl acetate. Extracts were then analysed using LC-MS/MS, targeting molecules with a product ion with a mass to charge ratio (m/z) of 102 – consistent with the homoserine lactone moiety – which is considered standard practise for identifying AHLs by mass spectrometry. This method revealed a molecule with the same m/z as 5-cis-C12-HSL was detectable at low levels. The 5-cis-C12-HSL molecule was chemically synthesised (Williams Laboratory, University of Nottingham) and analysis revealed it exhibited an identical HPLC column retention time and m/z as the identified molecule present in R7ANS extracts (Figure 3.5). To investigate if the MqsRIC system was involved in synthesis of the detected AHL molecule, extracts from R7ANSΔ*mqsRIC* were analysed by LC-MS/MS. This revealed that despite the deletion, a molecule resembling 5-cis-C12-HSL remained present and therefore this detected species is unlikely to be produced by the *mqsRIC* locus.

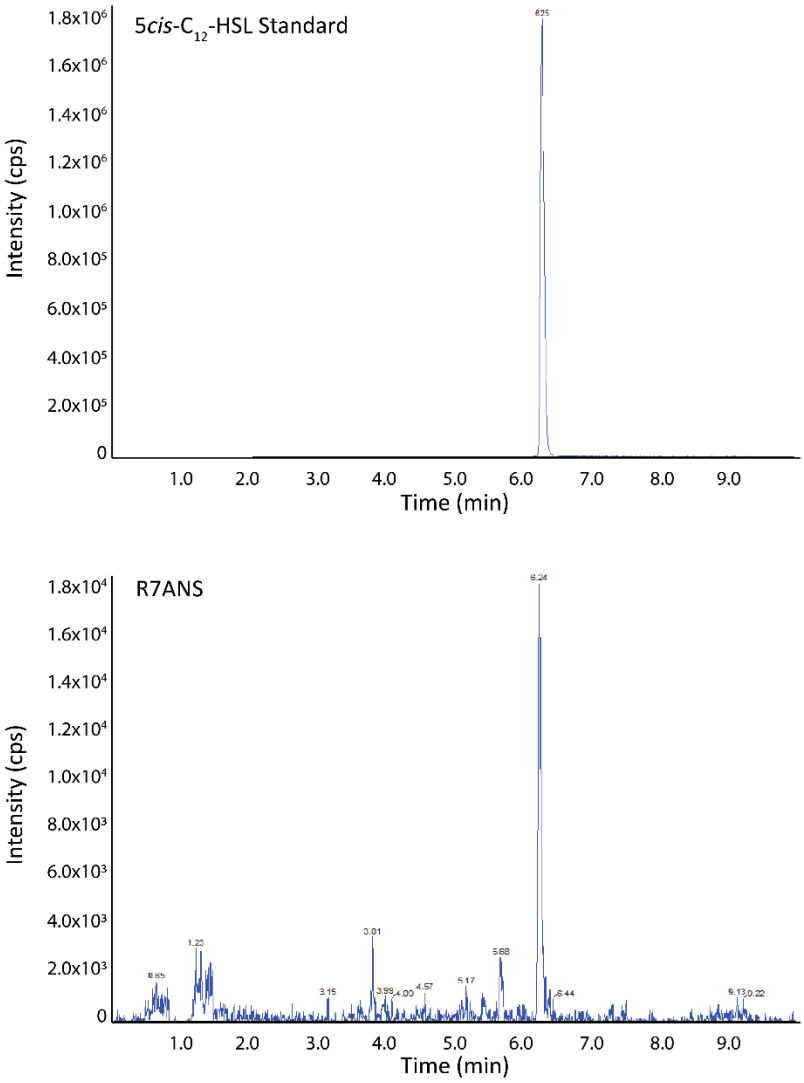


Figure 3.5 LC-MS/MS was used to analyse R7ANS supernatants for AHLs. No matter the substitutions or saturation of the fatty acid chain, AHLs generally dissociate to form an ion comprising the lactone ring moiety at m/z 102.1 (11). Extracted ion chromatograms are shown of a molecule with an identical retention time to that of 5-cis-C12-HSL synthetic standard. This molecule also produces a product ion with a m/z of 102.1.

Suo et al. (79) determined that product ion scans of 2,4-trans-C12-HSL did not detect a significant product at m/z 102 like other AHLs. Therefore, we carried out a broader scan for all molecules in R7ANS regardless of the presence of a m/z 102 product ion. By manually comparing extracts from R7ANS and R7ANSΔ*mqsRIC* we identified a product with an $[M+H]^+$ of 280.2 in the wild-type but not in the deletion mutant, which corresponds to the size of a di-unsaturated C12-HSL molecule. While LC-MS/MS alone cannot determine the position or configuration of double bonds, we hypothesised that this molecule

was likely the same molecule produced in *M. japonicum* MAFF303099. To investigate this further, 2,4-trans-C12-HSL was synthesised (Epichem, Perth, Western Australia) to use as a standard in LC-MS/MS and other assays.

The MS2 spectra of AHLs generally include a product ion at m/z 102.1 which corresponds to the lactone ring and is usually present at almost 100% abundance relative to the precursor ion, as well as product ions at m/z 84.0, 74.1 and 56.1 corresponding to certain fragments of the lactone ring (142) (Figure 3.6A). Five synthetic C12 AHL standards (listed in Table 3.1) including 2,4-trans-C12-HSL were run through LC-MS/MS to confirm their molecular weight and determine the profile of product ions for each standard. Extracted ion chromatograms (EICs) found prominent product ions at m/z 102.1, 84.0, 74.1 and 56.1 in four of the five AHL standards. Interestingly, EICs of 2,4-trans-C12-HSL did not find any of the regular AHL diagnostic product ions above background levels (Figure 3.6B). The most abundant product ion found in the MS2 spectra of 2,4-trans-C12-HSL was at m/z 179.3, which corresponds with the C12 fatty acid tail of the AHL molecule rather than the homoserine lactone moiety (Figure 3.6C), indicating that the fatty acid tail was being ionised rather than the homoserine lactone moiety.

Table 3.1. Synthetic C12 AHL standards used in LC-MS/MS experiments

AHL Standard	**$[M+H]^+$**
3OH-C12-HSL	300.22
3-oxo-C12-HSL	298.19
5-cis-C12-HSL	282.20
2,4-trans-C12-HSL	280.19
C12-HSL	284.22

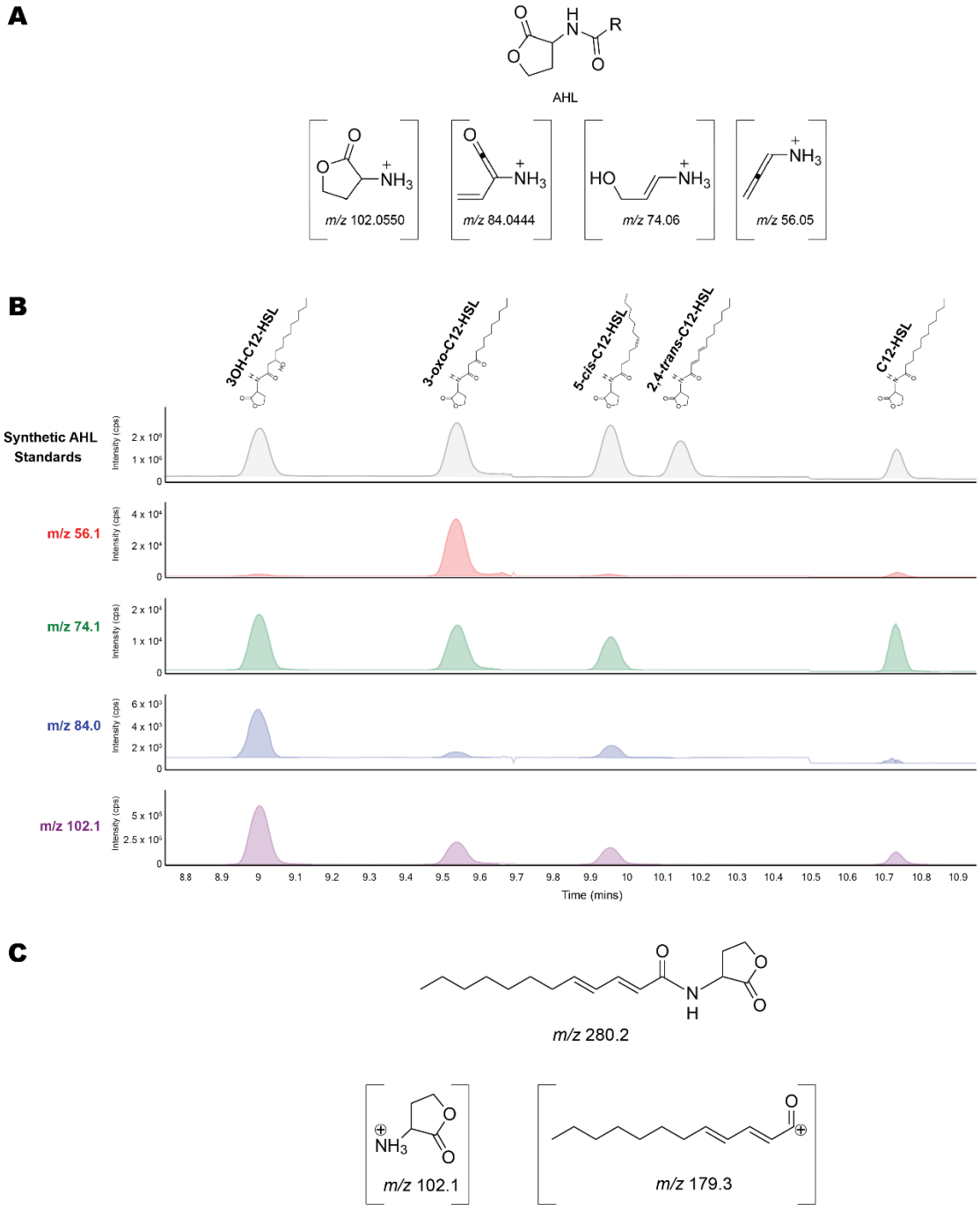


Figure 3.6. Four diagnostic product ions common to AHL molecules. (A) Depiction of an AHL molecule where R is the fatty acid chain, and four MS2 fragments commonly used to detect AHLs. Figure adapted from (143). (B) Total ion chromatogram of five C12 AHL standards is shown in grey. Each AHL is labelled above the corresponding peak with the name and structure. Extracted ion chromatograms at m/z 56.1 (red), 74.1 (green), 84.0 (blue) and 102.1 (purple) are shown below. All AHL standards except 2,4-trans-C12-HSL showed some detectable level of all four diagnostic product ions. (C) Depiction of 2,4-trans-C12-HSL, the lactone ring ion and the fatty acid chain ion.

AHLs were extracted from R7ANS, R7ANSΔ*mqs*RIC and R7ANSΔ*mqsRIC* overexpressing *mqsI*, *mqsC* and *mqsIC* and single ion monitoring (SIM) was used to look for molecules of the same molecular weight as the five different C12 AHL standards including 2,4-trans-C12-HSL (listed in Table 3.1). R7ANS wild-type extracts contained molecules of the same m/z and LC retention times as 3OH-C12-HSL and 2,4-trans-C12-HSL, as well as low levels of a molecule resembling 5-cis-C12-HSL (Figure 3.7A). Subsequent multiple reaction monitoring (MRM) of each of these molecules in R7ANS confirmed that the 280.2 precursor ion fragmented in an identical manner to synthesised 2,4-trans-C12-HSL, producing an MS2 product ion of 179. The 282.2 precursor ion in R7ANS fragmented to produce an MS2 102.1 product ion as is expected for AHL molecules, suggesting it is indeed a C12 AHL with one double bond consistent with 5-cis-C12-HSL. While we did observe a prominent molecule eluting with a similar retention time as 3OH-C12-HSL, it did not produce the same MS2 spectra as that of the 3-OH-C12-HSL standard, suggesting this peak in the LC trace was an unrelated molecule.

R7ANSΔ*mqs*RIC extracts were devoid of the 2,4-trans-C12-HSL molecule but still contained the low levels of 5-cis-C12-HSL observed in R7ANS, suggesting that the *mqsRIC* system is involved in the production of the 2,4-trans-C12-HSL molecule and not the 5-cis-C12-HSL molecule (Figure 3.7B). Overexpression of *mqsIC* in the *mqsRIC* deletion mutant restored 2,4-trans-C12-HSL production, albeit at much lower level than wild-type (Figure 3.7). We hypothesised that the strain overexpressing *mqsI* alone without *mqsC* would not produce the 2,4-trans-C12-HSL molecule but perhaps instead it could produce an intermediary AHL or fatty-acid molecule the crotonase might then modify. However, LC-MS/MS revealed very low levels of 2,4-trans-C12-HSL were present in extracts of R7ANSΔ*mqsRIC* overexpressing *mqsI* (Figure 3.7D), suggesting small amounts of 2,4-trans-C12-HSL are able to be produced in the absence of *mqsC* . R7ANSΔ*mqsRIC* overexpressing *mqsC* alone did not produce any 2,4-trans-C12-HSL (Figure 3.7E).

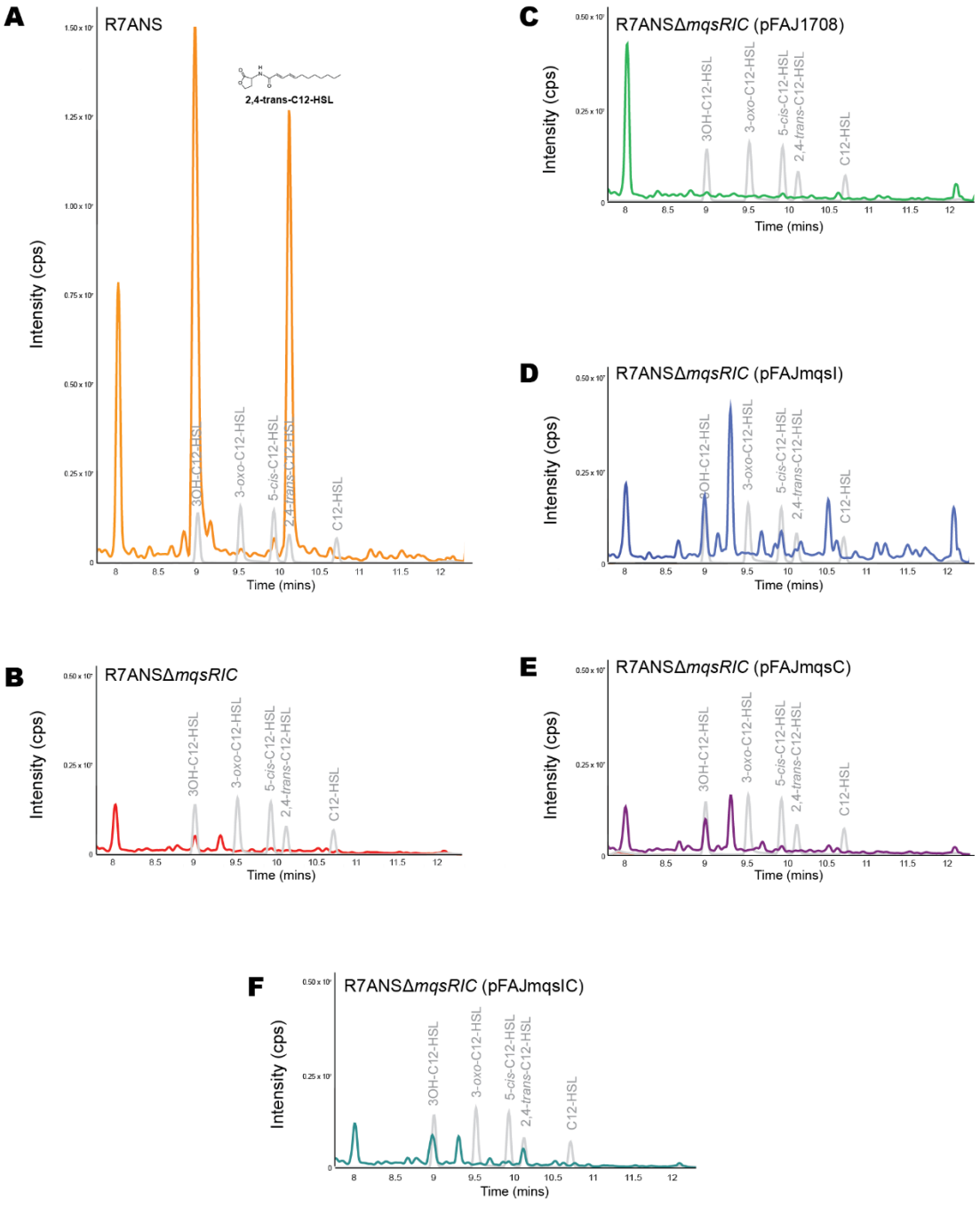


Figure 3.7 LC-MS/MS was used to analyse R7ANS supernatants for AHLs. R7ANS was deleted for *mqsRIC* and *mqsI*, *mqsC* and *mqsIC* were overexpressed. Extracted ion chromatograms are displayed here, viewing molecules with the same m/z as the synthetic C12 AHL standards; 300.2, 298.2, 282.2, 280.2 and 284.2. Each standard was prepared at 10 µg/mL and are displayed in grey in each graph.

3.2.7 Synthetic 2,4-trans-C12-HSL activates MqsR

In order to confirm whether the identified 2,4-trans-C12-HSL or other synthesised AHL molecules could activate the MqsRIC system we carried out β-galactosidase assays and well-diffusion bioassays. These assays were initially carried out using the biosensor strain described above and used in Figure 3.2 and Figure 3.3. However, we subsequently identified additional genes activated by MqsR that exhibited much stronger MqsR-dependent expression compared to the *mqsI* promoter (described in detail in Chapter 5). One of these promoters – P*mqsRNA7* – was used to construct a new biosensor plasmid with greater sensitivity and this was additionally used (Figure 3.8).

The synthetic AHL standards used in LC-MS/MS experiments were subsequently used in β-galactosidase assays and well-diffusion bioassays. The following synthetic AHLs were added to bioassay plates at 20 μM, 10 μM, 1 μM and 0.1 μM concentrations; 5-cis-C12-HSL, 3-OH-C12-HSL, C12-HSL, 3-oxo-C12-HSL and 2,4-trans-C12-HSL. 2,4-trans-C12-HSL strongly activated MqsR at all concentrations tested. Interestingly, 5-cis-C12-HSL and C12-HSL activated MqsR at >1 μM and >10 μM respectively, although it is unclear if these concentrations are biologically relevant (Figure 3.8A,B).

When carrying out liquid culture β-galactosidase assays, each biosensor strain was cultured and the synthetic AHLs and IPTG added during log-phase growth to a final concentration of 1 μM and 0.2 mM respectively. β-galactosidase assays of these cultures harvested in stationary-phase confirmed 2,4-*trans*-C12-HSL strongly induced MqsR activity and 5-cis-C12-HSL and C12-HSL also weakly induced MqsR activity. Interestingly, while the expression from both promoters (P*mqsI* and P*mqsRNA7*) in the presence of 2,4-trans-C12-HSL was similar, there was significant expression from P*mqsRNA7* in the presence of 5-cis-C12-HSL and C12-HSL but much lower expression from P*mqsI*. Other

long-chain AHLs tested had no significant effect on MqsR-mediated expression from P*mqsI* (Figure 3.8C) or P*mqsRNA7* (Figure 3.8D).

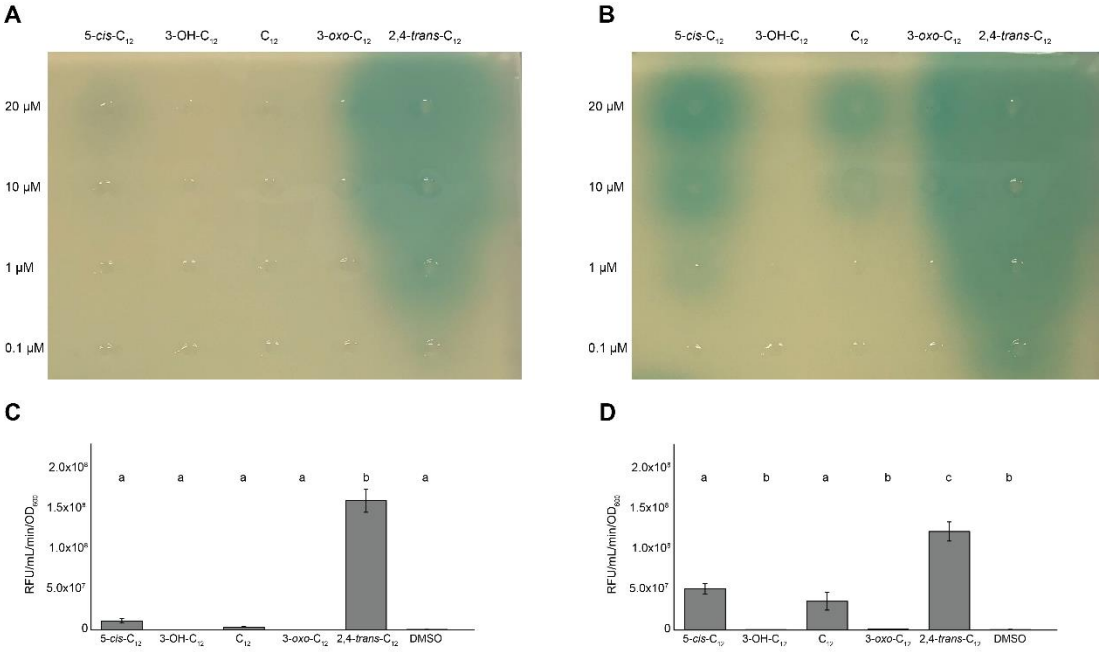


Figure 3.8 Comparisons between PmqsI and PmqsRNA7 reporters. Synthetic AHLs were added to (A, B) well-diffusion bioassays and (C, D) β-galactosidase assays with the promoters of (A,C) *mqsI* and (B,D) *mqsRNA7* used as a biosensor to detect AHLs activating MqsR and inducing the promoter-*lacZ* fusion. Bars represent the mean of three biological replicates and error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each strain. Values that are not statistically significant from each other are grouped by the same letter (a, b, c).

3.2.9 Purification of MqsC

In an attempt to gain insight into the mechanism of action of MqsC, the *mqsC* coding sequence was codon optimised for protein expression in *E. coli*. The optimised gene was then synthesised and cloned into protein expression vector pETM11. This construct incorporated a hexahistidine (6H) to the N-terminus of MqsC, allowing recombinant protein expression and purification using immobilised metal (nickel) affinity chromatography (IMAC) and size exclusion chromatography (SEC). MqsC protein was purified from *E. coli* NiCo de3 as described in Chapter 2 (section 2.12). Peak fractions were collected

from the IMAC and SEC purifications and samples run on a 4-15% SDS-PAGE to confirm the presence and purity of the recombinant MqsC protein (Figure 3.9). Purified MqsC protein was subsequently analysed using analytical gel filtration and eluted from the column at 13 minutes, corresponding to a standard protein of ~75kDa. This preliminarily molecular weight estimation suggested that MqsC, like other crotonase-family enzymes, likely forms a homotrimer in solution, however, further experiments need to be carried out to confirm the oligomerisation. Purified MqsC protein was used in initial crystallisation trials using vapour diffusion methods. Various concentrations of protein from 0.1 mg/mL to 5 mg/mL were added to 288 screening buffer conditions (see methods), however, no ideal conditions were identified and due to time constraints, this was not pursued further.

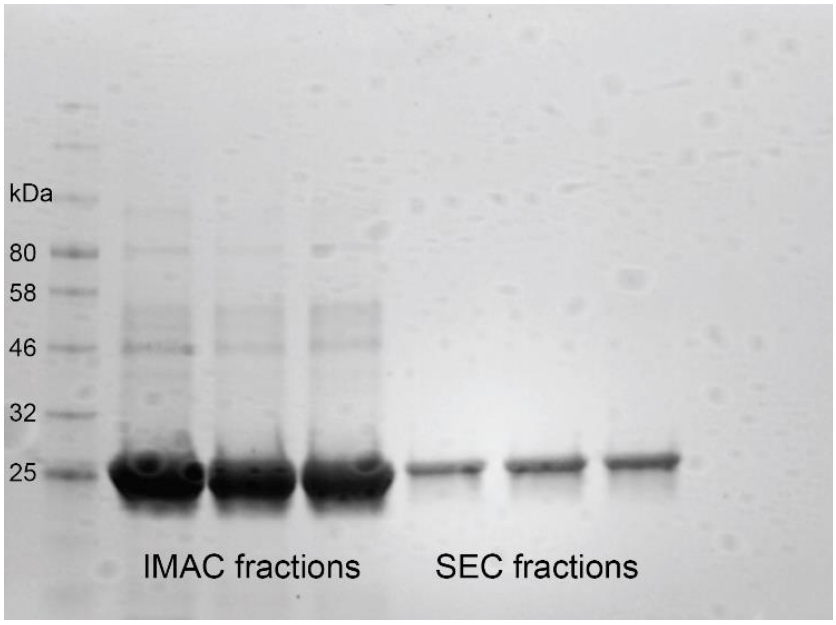


Figure 3.9 Purification of MqsC from R7A. A precast 4-15% SDS-PAGE was used to assay peak fractions at 280 nm UV from IMAC and SEC purifications of MqsC recombinant protein. MqsC should appear at 26 kDa as a monomer.

3.2.10 MqsC is unable to convert C12-HSL *in vitro* into a molecule capable of activating MqsR

It is unclear exactly how the MqsC protein is involved in the synthesis of 2,4-trans-C12-HSL. We hypothesised that like other crotonase-family dehydratases (e.g. the DSF synthase) it may modify the C12 fatty acid chain of the acyl-CoA or acyl-ACP precursor molecule. Since we had recombinantly

purified MqsC we wondered if it could perhaps convert C12 AHLs *in vitro* to 2,4-trans-C12-HSL. Purified MqsC protein was mixed with C12-HSL and 3-OH-C12-HSL and the mixture added to bioassays. Various combinations of MqsC recombinant protein at concentrations of 1 mg/mL, 2.5 mg/mL and 5 mg/mL with AHLs at 10 μM and 20 μM were created, incubated at room temperature for 30 minutes and added to bioassays. None of the MqsC-AHL mixtures were able to activate MqsR in the biosensor strain, suggesting that MqsC may not act on AHLs as a substrate, at least under the conditions tested here. The concentration of AHL in the mixtures did not surpass 20 μM, since it had already been shown that C12-HSL alone induces MqsR activity at higher concentrations (see section 3.2.7).

To further investigate whether MqsC could modify exogenously supplied C-12 AHLs into a molecule capable of activating MqsR *in vivo*, a markerless in-frame deletion of *mqsI* was constructed that contained an intact *mqsC* gene downstream of the *mqsI* promoter. The reporter plasmid pPmqsIzR was introduced into the *mqsI* mutant to create R7AΔ*mqsI*(pPmqsIzR). This strain was essentially the same as the previously described biosensor strain, however the native *mqsC* gene was still present. C12-HSL and 3-OH-C12-HSL was added to a bioassay using this new strain as a biosensor at a concentration of 20 μM. No blue coloured zones were observed, indicating that the MqsC present in the biosensor strain was not able to modify the exogenously added C12-HSL or 3-OH-C12-HSL into a molecule capable of inducing MqsR, under conditions tested.

To further examine if MqsC could modify exogenously supplied AHLs into molecules that activate MqsR, *mqsC* was overexpressed on a plasmid in the R7ANSΔ*mqsRIC* background. This strain was grown to log-phase and individual cultures were supplemented with 1 μM C12-HSL, 1 μM 3-oxo-C12-HSL or 1 μM 3-OH-C12-HSL and grown until stationary-phase. Supernatants were harvested and added to a bioassay, and again induction of MqsR by these supernatants was not observed. These data suggest that MqsC either

cannot use these specific AHL molecules as substrates to produce 2,4-trans-C12-HSL or that MqsC instead acts on fatty-acid substrates (acyl-ACP or acyl-CoA) before MqsI incorporates them into AHL molecules.

3.2.11 Strains lacking *mqsC* still have a functional *mqsI*

Two strains of *Mesorhizobium* sp. – WSM4303 and WSM4306 – encode MqsR and MqsI but naturally lack *mqsC*. This suggests they may have evolved to produce and respond to AHLs produced by mqsI in the absence of MqsC. We wondered if these strains were still capable of producing AHLs that activate the R7A MqsR protein in the presence or absence of MqsC. The MqsC overexpression construct pNmqsC was introduced into WSM4303 and WSM4306, and the strains were grown to stationary phase. Spent supernatants were added to a bioassay using R7AΔ*mqsRIC*(pPmqsIzR) as a biosensor. Overexpression of R7A-*mqsC* in these strains resulted in increased β-galactosidase in the biosensor strain, suggesting that while these strains do not carry *mqsC*, the AHL-synthase gene *mqsI* is still functional and together with *mqsC* are capable of producing a molecule that can activate R7A MqsR. Importantly, the lack of any activation of R7A MqsR by supernatants from wild-type WSM4303 or WSM4306 confirms that strains lacking *mqsC* do not produce 2,4-trans-C12-HSL detectable by the MqsR bioassay. Further studies are required to determine whether the MqsR proteins of WSM4303 and WSM4306 are functional, what AHLs they respond to in the absence of MqsC.

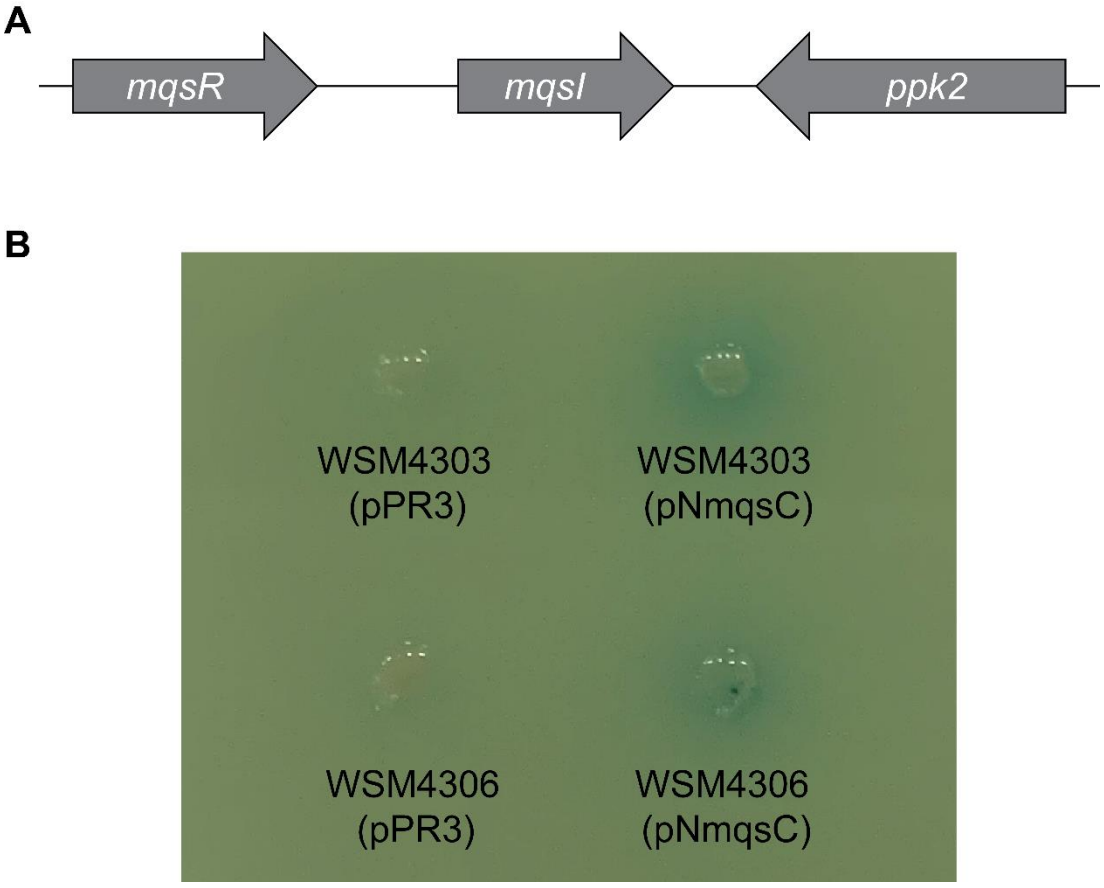


Figure 3.10 Overexpression of *mqsC* in strains that lack *mqsC*. (A) Genomic arrangement of the *mqsRI* genes in WSM4303 and WSM4306. As is the case in other *Mesorhizobium* spp., these genes are still located upstream of *ppk2*. (B) WSM4303 and WSM4306 overexpressing R7A-*mqsC* (and empty vector controls) were added to a bioassay. The AHL molecule in these strains induced R7A-MqsR activity in the biosensor strain.

3.2.12 MqsR may contain an extra helix in the ligand-binding domain

LuxR-family proteins form two domains – the ligand- or AHL-binding domain and the DNA-binding domain (134, 135). Given the uniqueness of MqsR in that it selectively responds to 2,4-trans-C12-HSL over other C12 AHL species, we wondered if there might be some evident structural features of MqsR that could indicate how this selectivity was achieved. Alphafold2-multimer (144) was used to predict homodimer structures for the most distinct MqsR representatives in the genus *Mesorhizobium* and compare the ligand-binding domains to other known LuxR proteins that recognise C12 AHLs. Eight distinct MqsR sequences were selected from the phylogenetic tree used in section 3.2.1 (Appendix 2). When these predicted structures were aligned to MqsR

from R7A using PyMOL (145) the structures aligned almost perfectly (Figure 3.11A). The predicted structures included an additional predicted helix in the ligand-binding domain compared to other LuxR homologues such as LasR and QscR (Figure 3.11). Interestingly, a comparison of the ligand-binding domain of R7A-MqsR with the predicted structure for MqsR from WSM4306, which does not carry the *mqsC* gene, revealed this variant of MqsR lacks this additional predicted helix (Figure 3.11B).

Table 3.2. RMSD values of MqsR predicted structures aligned to the predicted structure of R7A-MqsR

Strain	RMSD (Å)
Mesorhizobium sp. CA10	0.323
M. ciceri CC1192	0.647
Mesorhizobium sp. ES11	0.322
Mesorhizobium sp. M4B	0.290
M. japonicum MAFF 303099	0.833
Mesorhizobium sp. NZP 2234	0.665
M. ciceri WSM1271	0.790

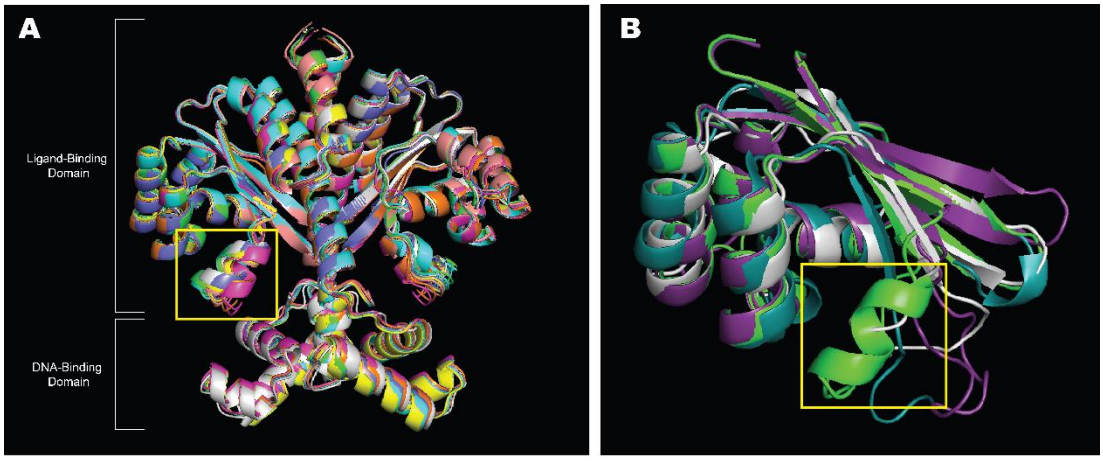


Figure 3.11 Predicted structures of MqsR homologues (A) Predicted structures of eight distinct MqsR homologues overlayed on one another; R7A (green), CA10 (blue), CC1192 (pink), ES11 (yellow), M4B (peach), MAFF 303099 (grey), NZP2234 (purple), WSM1271 (orange). (B) The ligand-binding

domain of R7A-MqsR (green), LasR (white), QscR (teal) and WSM4306-MqsR (purple). The helix mentioned in text is highlighted with a yellow box.

3.3 Discussion

In this chapter, a chromosomally encoded quorum sensing locus was identified to be conserved across a database of over 200 strains of mesorhizobia. Further investigation established that this MQS system produces an unsaturated long-chain AHL sharing identical properties to the previously identified 2,4-trans-C12-HSL. Unlike all other characterised LuxRI quorum sensing systems, MQS was found to require a second AHL-synthase – MqsC – for production of AHLs to induce MqsR-mediated activity. The inverted repeat sequence CGTGATCACG was identified in R7A to be essential for P*mqsI* expression, and likely forms the *mqs*-box, i.e., the DNA binding site for the MqsR-AHL complex. MqsR exhibited a relatively narrow range of recognition, only responding strongly to the cognate 2,4-trans-C12-HSL and to high concentrations of 5-cis-C12-HSL and C12-HSL.

LC-MS/MS is limited in that we can elucidate the chemical formula of the molecule, but not the configuration of the double bonds. NMR spectroscopy could be used to determine the configuration, however since 2,4-trans-C12-HSL has previously been identified in *M. japonicum*, we decided further characterisation of the double bonds within the MqsIC-derived AHL was unnecessary and we can reasonably assume the double bonds are in the *trans* configuration. Moreover, we showed that synthesised 2,4-trans-C12-HSL strongly activates the MqsRIC system. While we showed that the main AHL species produced by R7ANS is 2,4-trans-C12-HSL, we also observed low levels of a molecule resembling 5-cis-C12-HSL in wild-type and *mqsRIC* deletion mutants and synthesised 5-cis-C12-HSL was able to induce MqsR activity. The presence of the 5-cis-C12-HSL molecule in R7ANS and R7ANSΔ*mqsRIC* extracts suggests the presence of another AHL-synthase in the genome, however, bioinformatic searches were unsuccessful in identifying another LuxI-family encoding gene.

The MQS system is the first known LuxRI system to incorporate a second AHL-synthase and here we have shown that the crotonase-family protein MqsC is

required for activation of MqsR. In this chapter we observed MqsR to respond to a relatively narrow range of AHLs compared to other more promiscuous LuxR-family proteins, with a strong affinity to its cognate di-unsaturated AHL 2,4-trans-C12-HSL and some affinity to the unsaturated 5-cis-C12-HSL. The extreme conservation of *mqsRIC* in *Mesorhizobium* spp., the requirement for the second AHL-synthase and the preference towards somewhat unusual unsaturated AHL molecules may suggest an evolved mechanism for mesorhizobia specifically to sense and recognise only members of its own genera within the rhizosphere.

We identified only two strains of *Mesorhizobium sp*. – WSM4303 and WSM4306 – which carry *mqsR* and *mqsI* without *mqsC*. Overexpression of *mqsC* from R7A in these strains lead to production of an AHL able to activate the R7A-MqsR in the biosensor strain. These results lead us to wonder whether *mqsRI* in these strains still comprise a functional QS system with a different – perhaps saturated – main AHL product, or if a loss of *mqsC* has produced a vestigial QS system that only responds to exogenous 2,4-trans-C12-HSL. Predicted MqsR structures from eight distinct *Mesorhizobium* strains all exhibited an extra helix within the ligand-binding domain compared to other LuxR proteins known to recognise C12 AHLs. We hypothesise that this additional helix in the MqsR predicted structures may give rise to an adapted ligand-binding pocket optimised for the di-unsaturated ligand. The residues of this helix should be targeted for future mutagenesis and structural biology work to determine the exact binding site for 2,4-trans-C12-HSL. Interestingly, predicted structural models of the MqsR proteins from WSM4303 and WSM4306 did not contain this extra helix, consistent with these MqsR orthologues having evolved to perceive an AHL ligand other than 2,4-trans-C12-HSL.

LuxI-family AHL synthases generally fall into two categories, those that use acyl-ACPs as a substrate and those that use acyl-CoAs as a substrate. The *mqsI* homologue in *M. japonicum* MAFF 303099 was recently bioinformatically

classified as an acyl-CoA-dependent enzyme and Suo et al. predicts the substrate to be 2,4-*trans*-C12-CoA (79). Acyl-CoA LuxI proteins generally encode two consecutive residues that form the 'indole platform' at the CoA binding site. These residues are usually two tryptophans, however the first is sometimes replaced with phenylalanine or tyrosine, all of which are hydrophobic residues with an aromatic ring. The 'indole platform' acts by sandwiching the adenine ring of the CoA molecule between the two aromatic rings of the amino acids using π-stacking interactions (146). Most of the MqsI homologues identified in this chapter, including that of MAFF 303099, encode Met146 and Trp147 in these positions. The lack of the canonical 'indole platform' may suggest that acyl-CoA is not the optimal substrate or perhaps that the binding of CoA may occur elsewhere in within the MqsI protein.

A distinct QS system produces unsaturated fatty acid signalling molecules called diffusible signal factors (DSFs) (15). Synthesis of DSFs in *Burkholderia cenocepacia* (BDSFs) involve a dual-function crotonase-family enzyme which first acts as a dehydratase to introduce a *cis* double bond within the acyl-ACP substrate, and then acts as a thioesterase to release the final DSF product from the ACP. The *B. cenocepacia* crotonase-family DSF synthase uses 3-OH-C12-ACP as a substrate and dehydrates the 3-hydroxyl group to create the 2-*cis* double bond (18). While it is still unknown as to what MqsI uses as a substrate to create 2,4-trans-C12-HSL, we have shown that MqsC is required to produce the molecule that induces MqsR activity. We hypothesise that MqsC may act similarly to the crotonase-family DSF synthase, whereby MqsC introduces the two double bonds into the fatty-acid chain of the substrate molecule before MqsI brings together SAM and the unsaturated substrate to produce 2,4-trans-C12-HSL.

Previously, Yang et al. described the MqsI homologue of *M. loti* NZP 2213 to produce C12-HSL as its main AHL species using MS/MS, screening for product ions of mass 102 (78). As found by Suo et al. (79) and confirmed in this chapter, the 2,4-trans-C12-HSL molecule produced by MqsI does not

produce a significant 102 product ion, and therefore screening for m/z 102 ions would miss the di-unsaturated AHL. It is possible that MqsI produces different AHLs in different strains of *Mesorhizobium* depending on the fatty-acid pool of the cell, however we suspect it is more likely that *M. loti* NZP 2213 also produces 2,4-trans-C12-HSL as the main AHL product and perhaps low levels of C12-HSL as a secondary product. MqsR and MqsI of R7A shares 96% and 97% amino acid identity with MqsR and MqsI of NZP 2213 respectively. There are no differences in residues comprising the ligand-binding domain of MqsR or the likely substrate binding site of MqsI between the two strains, suggesting that the MqsI and MqsR proteins likely produce and respond to the same molecule.

In summary, the data in this chapter describe a unique QS system almost completely conserved in the *Mesorhizobium* genus which we named the Mesorhizobium Quorum Sensing loci. The MQS system incorporates a crotonase-family enzyme, essential for production of an AHL capable of inducing the cognate LuxR receptor. The main AHL species produced by this system is 2,4-trans-C12-HSL which has also been found in *M. japonicum* MAFF 303099 (79). We also describe only two strains of *Mesorhizobium sp.* that do not encode MqsC but still preserve the surrounding genomic context. The level of conservation of MQS within mesorhizobia suggest an important role, although deletion of all three genes has no obvious effect on growth under the conditions tested. Bacteria often use QS to sense other genera or species within their environment, so it is possible that mesorhizobia all encode this unique system in order to distinguish between ‘friend or foe’ within the rhizosphere.

Chapter 4

Phenotypic exploration of *mqsRIC* mutants

4.1 Introduction

Bacteria of the rhizosphere use quorum sensing to regulate a multitude of phenotypes. Cells communicate with other cells of the same species within their own population as well as neighbouring populations of different species, genera or even kingdoms. Some rhizobacteria produce AHL molecules which can be recognised by plants, such as the beneficial *Sinorhizobium meliloti* and the pathogenic *Pseudomonas aeruginosa* on their plant host *Medicago truncatula*. Upon exposure to 3-oxo-C12-HSL of *P. aeruginosa* and 3-oxo-11-cis-C16-HSL of *S. meliloti*, *M. truncatula* exhibits significant changes in its proteome (147). Rhizobacteria also use QS to regulate a plethora of bacterial symbiosis phenotypes such as polysaccharide production, horizontal gene transfer, root colonisation, nodule formation, infection thread formation, development of bacteroids and nitrogen fixation (75).

Many rhizobacteria use QS to regulate horizontal gene transfer. One example is *Rhizobium leguminosarum*, which encodes four LuxI-family AHL synthases – TraI, CinI, RhiI and RaiI. Each of these synthases have a cognate LuxR-family receptor, and the genome also encodes a number of orphan LuxR proteins (148). *R. leguminosarum* uses a complex network of different LuxR-family regulators to sense potential recipient cells and induce horizontal transfer of the conjugative plasmid pRL1JI. Genes encoding TraRI and BisR are located on this plasmid. BisR perceives 3-OH-7-cis-C14-HSL produced by the chromosomal CinI and the BisR-AHL complex then induces expression of *traR*. This in turn upregulates *traI* expression and TraI-derived AHL production, leading to plasmid conjugation. BisR also acts as a repressor of CinI, knocking down AHL production in cells harbouring pRL1JI (donor cells) (65, 66, 70, 71) The CinRI-BisR-TraRI QS network essentially allows BisR to supress CinI-derived AHL production in cells harbouring the pRL1JI plasmid, while allowing cells to sense the CinI-derived 3-OH-7-cis-C14-HSL being produced in potential recipient cells lacking pRL1JI. The outcome of these complex layers of regulation is that only the presence of 3-OH-7-cis-C14-HSL produced by plasmid-less recipient strains can induce pRLJ1 transfer from donors.

Mesorhizobium japonicum R7A and *M. ciceri* WSM1271 also use QS to regulate horizontal transfer of integrative and conjugative elements (ICEs). The LuxI-family AHL synthase TraI1 produces 3-oxo-C6-HSL as the main product. The TraR-AHL complex upregulates *traI1* expression, increasing AHL production and activates a cascade of downstream gene expression leading to excision and conjugation of the ICE (80, 82, 85). This QS system in *M. japonicum* R7A is epigenetically regulated by additional DNA-binding proteins (described in detail in Chapter 6).

Sinorhizobium meliloti harbours many AHL-based QS genes including *expR*, *sinR* and *sinI*. SinI is a LuxI-family AHL synthase and produces a range of long chain AHLs from C12-C18, with 3-oxo-11-cis-C16-HSL as the main product (13, 72). The *sinR* gene is located upstream of *sinI*, however SinR activity is not induced by AHLs. Instead, SinI-derived AHLs activate the orphan LuxR-family ExpR, which in turn upregulates *sinI* expression. After the AHL concentration reaches a certain threshold, ExpR also represses *sinR* transcription leading to a negative feedback loop for *sinI* regulation (149). One of the most well characterised phenotypes regulated by ExpR is exopolysaccharide (EPS) synthesis. *S. meliloti* produces two main exopolysaccharides – EPS I (succinoglycan) and EPS II (galactoglucan), both of which are regulated by QS (150-152). Under phosphate limiting conditions, PhoB of the Pho Regulon induces SinR activity which leads to upregulation of *sinI* expression and increased AHL production. Therefore, the QS-dependent EPS production is regulated in response to phosphate starvation (149). Mutations in this QS system and disruption in EPS production results in reduced nodulation efficiency and formation of incomplete, immature nodules (153).

In this chapter, we aimed to identify potential phenotypes controlled by the MQS system described in Chapter 3. We hypothesised that the extremely high level of conservation of the *mqsRIC* genes across the *Mesorhizobium* genus

suggests an important role in basic cell growth or perhaps symbiosis phenotypes. In contrast to work by others on the MQS genes (78), we observed no significant differences in symbiosis phenotypes between *mqsRIC* deletion mutants and their wild-type counterparts in both *M. japonicum* R7A and *M. ciceri* CC1192. We also observed no differences in ICE*Ml*Sym^{R7A} horizontal transfer frequencies when *mqsR* was deleted either in the donor, recipient or both donor and recipient. We also created promoter-*lacZ* fusions for three exopolysaccharide genes and identified small but significant changes in the transcription or translation of the *exoY* and *exoU* genes, both of which are involved in polysaccharide synthesis (154, 155).

4.2 Results

4.2.1 MqsRIC is not required for effective symbiosis

In other nitrogen-fixing rhizobia, QS can regulate phenotypes such as legume nodulation, symbiosis gene horizontal transfer and in some cases QS systems are required for optimum nitrogen fixation (75). To determine whether the MQS system plays a role in regulating nitrogen fixation or plant symbiosis, *mqsRIC* mutants and plasmid-complemented strains were used as inocula *in planta*. Both *M. ciceri* CC1192 and *M. japonicum* R7A carrying constructed deletions in *mqsRIC* were tested for symbiotic effectiveness with legumes *Cicer arietinum* cv. Striker (chickpea) and *Lotus japonicus* gifu (lotus), respectively. The *mqsRIC* genes locus including the *mqsR* promoter region were amplified from DNA from each strain and cloned into the low-copy number plasmid pFAJ1700 (which is stably maintained *in planta*) for genetic complementation of the deletion in each strain. The *mqsIC* operon was also constitutively overexpressed in the Δ*mqsRIC* backgrounds to create strains capable of producing MqsIC-derived AHLs without a cognate MqsR, in case there were plant or bacterial responses to AHLs that did not involve MqsR.

4.2.1.1 Deletion of *mqsRIC* has no effect on nitrogen fixing symbiosis in chickpea

Chickpea plants were harvested after six weeks and nitrogenase activity was measured using acetylene reduction assays (ARA). Nitrogenase is an enzyme which catalyses nitrogen fixation. Nitrogenase also reduces acetylene to ethylene, and this reaction can be measured to determine nitrogenase activity and therefore nitrogen fixing effectiveness of the bacterial inoculant (128, 156). Intact chickpea plants were harvested and acetylene reduction assays performed as previously described (128). Data from two separate experiments are presented (Figure 4.1), each was carried out using five replicates per treatment. While there appeared to be some slight differences in acetylene reduction, the small changes were not statistically significant nor were they consistent across experiments (Figure 4.1). Uninoculated plants with and

without added nitrogen were also assayed, and as expected acetylene reduction was measured as zero (data not shown).

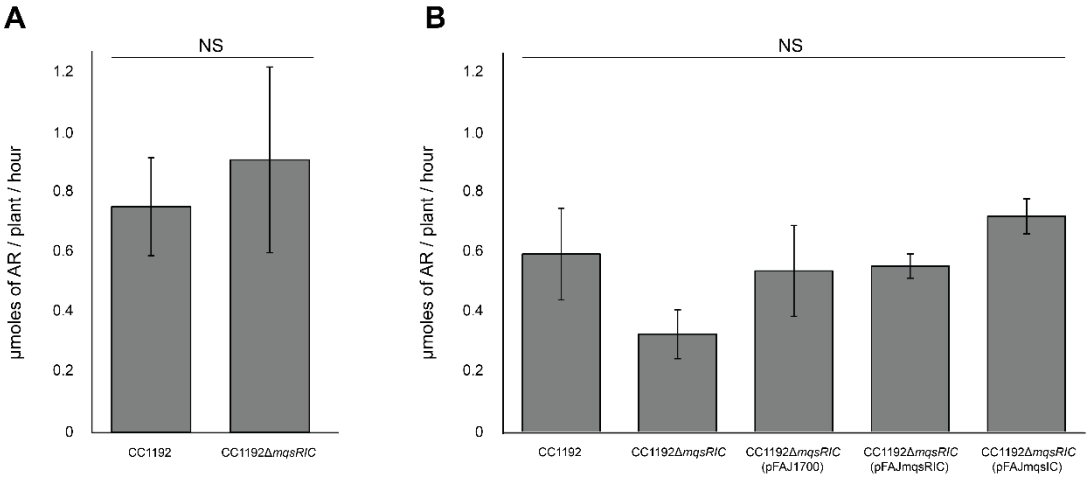


Figure 4.1 Acetylene reduction assays of chickpea plants inoculated with CC1192-derived strains either mutated for or overexpressing the *mqsRIC* genes. Bars represent the mean of three technical replicates of five biological replicates, and error bars represent the standard deviation from the mean. (A) The two treatments were compared using a student's t test, NS denotes not significant ($p>0.05$). (B) A one-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another, NS denotes not significant ($p>0.05$).

4.2.1.2 Deletion of *mqsRIC* has no effect on chickpea plant mass

Chickpea plants were inoculated with *M. ciceri* CC1192 either overexpressing or mutated for the *mqsRIC* genes. Mature plants were harvested after six weeks and the plant matter above the first cotyledon was dried and weighed. No significant differences were observed in the weight of dry plant material between any of the plants inoculated with the CC1192 variants (Figure 4.2). Surprisingly, while the uninoculated nitrogen starved plants looked less healthy than the nitrogen fed and inoculated plants (Figure 4.2C), the weights of dried plant material from nitrogen starved plants were not statistically different as would be expected, therefore we cannot draw conclusions based on plant weights alone.

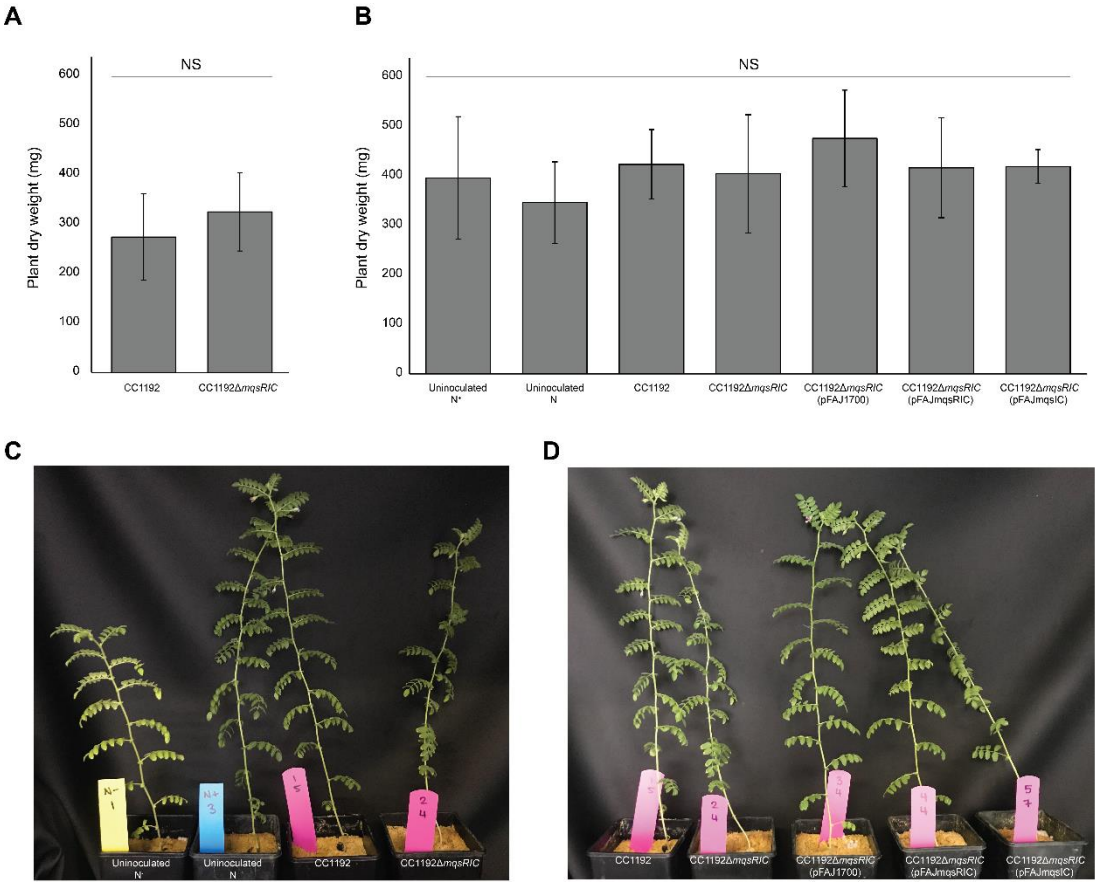


Figure 4.2 Foliage weights of chickpea plants inoculated with CC1192-derived strains either mutated for or overexpressing the *mqsRIC* genes. (A) Bars represent the mean weight of dried chickpea plant material from above the hypocotyl of five plants per treatment, error bars represent the standard deviation from the mean. The two treatments were compared using a student's t test, NS denotes not significant ($p>0.05$). (B) Bars represent the mean weight of dried chickpea plant material from above the hypocotyl of six to eight plants per treatment, error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another, NS denotes not significant ($p>0.05$). (C) One representative plant from the following treatments at six weeks post inoculation; Nitrogen starved, nitrogen fed, CC1192, CC1192Δ*mqsRIC*. (D) One representative plant from each of the CC1192-derived strain treatments at six weeks post inoculation.

4.2.1.3 Deletion of *mqsRIC* has no effect on nodule weights in chickpea

Nodules were removed from mature chickpea plant roots and dried to remove moisture. When the pilot *in planta* experiment was carried out, plants inoculated with the *mqsRIC* mutant had a marginally higher overall dry nodule weight compared to those inoculated with the wild-type strain (n=5, p=0.049). When the experiment was repeated on a larger scale with six to eight replicates

per treatment, no significant differences were observed in nodule dry weights across any of the inocula tested. As expected, nodules were not present on any of the uninoculated control plants in either experiment.

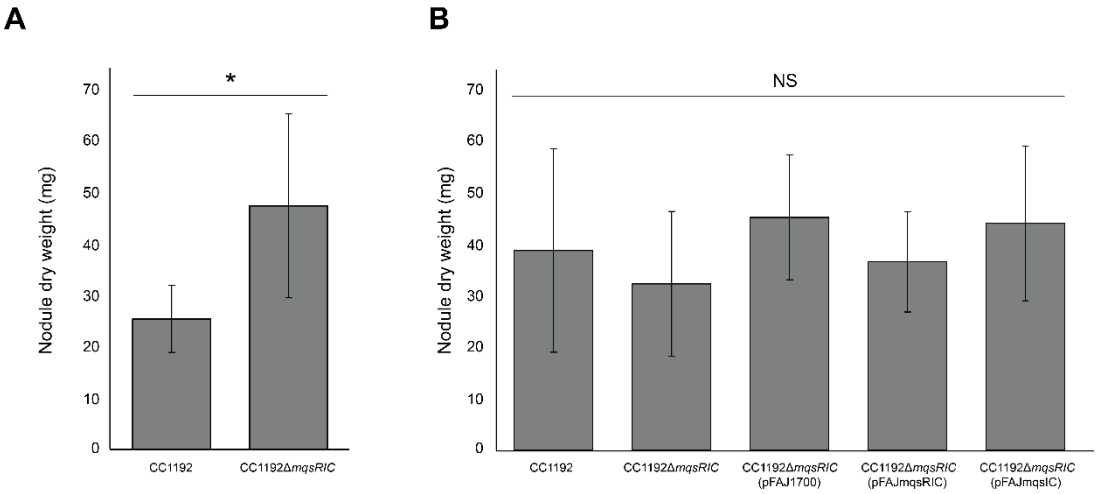


Figure 4.3 Weights of dried nodules from chickpea plants inoculated with CC1192-derived strains either mutated for or overexpressing the *mqsRIC* genes. Bars represent the mean of five (A) or six to eight (B) plants per treatment, and error bars represent the standard deviation from the mean. (A) The two treatments were compared using a student's t test, the asterisk denotes a statistical difference ($p<0.05$). (B) A one-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another, NS denotes not significant ($p>0.05$).

4.2.1.4 Deletion of *mqsRIC* has no effect on nodule number in chickpea

Nodules on chickpea plants inoculated with CC1192 variants overexpressing or mutated for *mqsRIC* genes were counted at harvest, six weeks post inoculation. During the pilot experiment, plants inoculated with the *mqsRIC* mutant exhibited a higher number of nodules per plant compared to the wild-type inoculum (n=5, p=0.025) (Figure 4.4A). However, this result was not able to be replicated, and no statistical differences in nodule counts were observed between treatments in the following chickpea experiments (Figure 4.4B). As expected, nodules were not present on uninoculated control plants in any of these experiments.

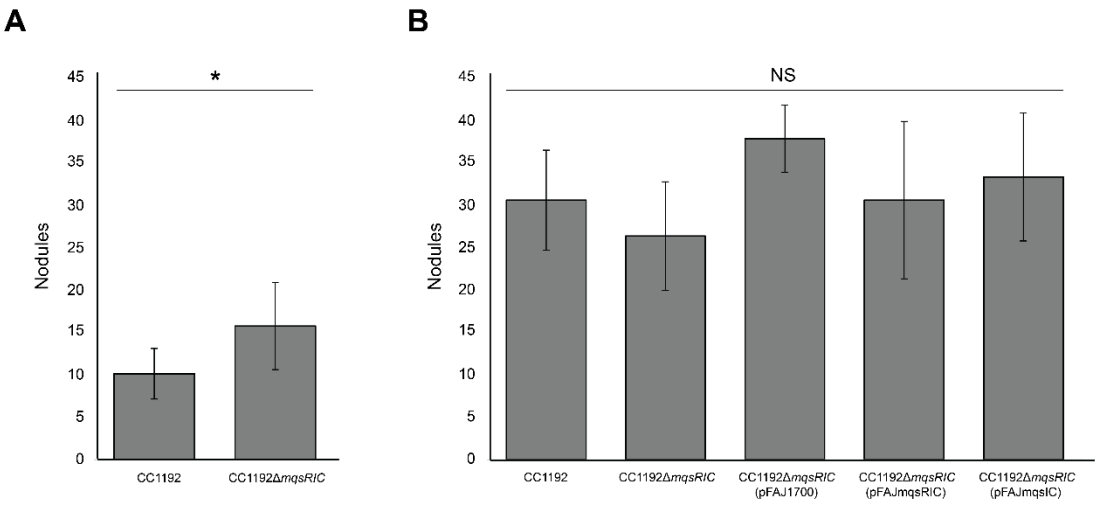


Figure 4.4 Number of nodules on chickpea plants inoculated with CC1192-derived strains either mutated for or overexpressing the *mqsRIC* genes. (A) Bars represent the mean nodule number of five chickpea plants per treatment, error bars represent the standard deviation from the mean. The two treatments were compared using the student's t test, with the asterisk denoting significance ($p<0.05$). (B) Bars represent the mean nodule number of six to eight plants per treatment, error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another, NS denotes not significant ($p>0.05$).

4.2.1.5 Deletion of *mqsRIC* has no effect on lotus plant mass

Lotus plants were inoculated with *M. japonicum* R7A strains that either overexpressed or were deleted for the *mqsRIC* genes. Plants were harvested after 30 days, and fresh plant material above the first cotyledon was weighed. As expected, the uninoculated control plants weighed significantly less than the other plants, however, no other differences were observed between plants inoculated with the various R7A strains (Figure 4.5) suggesting the MQS system does not impact the ability of R7A to increase plant weights through symbiotic nitrogen fixation.

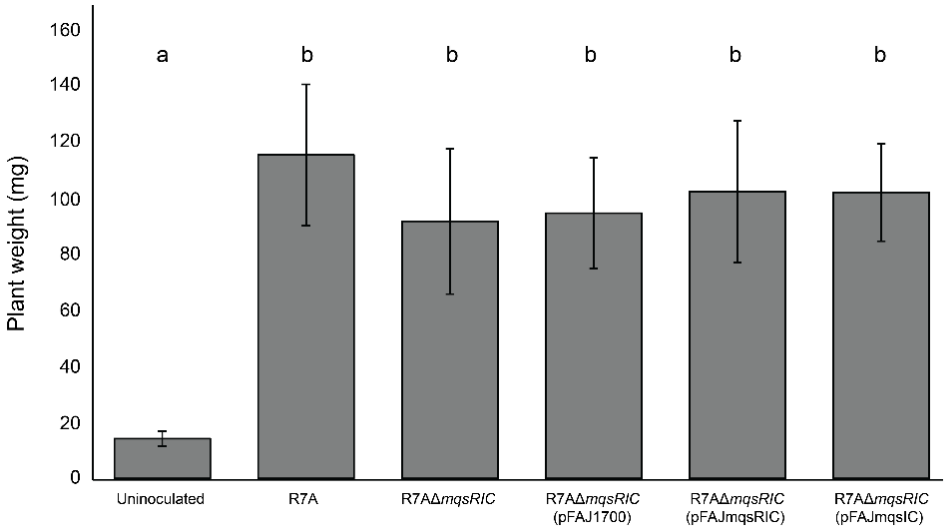


Figure 4.5 Foliage weights of lotus plants inoculated with R7A-derived strains either mutated for or overexpressing the *mqsRIC* genes. Bars represent the mean weight of lotus plant material from above the first cotyledon of 15 plants per treatment, error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another. Values that are not statistically different from each other are grouped by the same letter (a, b).

4.2.1.6 Deletion of *mqsRIC* has no effect on nodule number in lotus

Lotus plants were inoculated with R7A-derived *mqsRIC* mutant and overexpression strains and the number of nodules was monitored through glass over 30 days. While there were differences in nodule numbers observed at 24 days post inoculation (dpi) between plants inoculated with the *mqsRIC* mutant and the plasmid-complemented strain (p=0.037) (Figure 4.6A), the vector-only control R7A*mqsRIC*(pFAJ1700) strain which also lacks the MQS system exhibited nodule counts almost identical to those of R7A. Mature plants were harvested at 30 dpi and averaged 4.5 nodules per plant, with no differences in nodule numbers between inocula (Figure 4.6B).

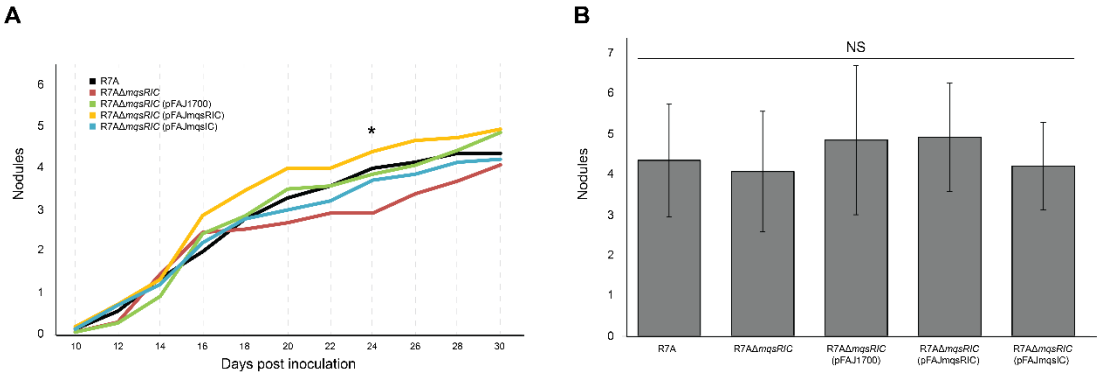


Figure 4.6 Number of nodules on lotus plants inoculated with R7A-derived strains either mutated for or overexpressing the *mqsRIC* genes. (A) Lines represent the mean nodule number of 15 lotus plants per treatment every two days 10-30 dpi. One-way ANOVA and Tukey post-hoc tests were used to compare each treatment to one another at each time point. The asterisk at 24 dpi denotes a significant difference in nodule count between treatments R7AΔ*mqsRIC* and R7AΔ*mqsRIC*(pFAJmqsRIC). (B) Bars represent the mean nodule number of 15 lotus plants per treatment at 30 dpi, error bars represent the standard deviation from the mean. A One-way ANOVA and Tukey post-hoc test was used to compare each treatment to one another, NS denotes not significant ($p>0.05$).

4.2.2 Deletion of *mqsRIC* in either donors or recipients has no effect on the frequency of ICE*Ml*Sym^{R7A} horizontal transfer

While the AHLs of the TraRI and MqsRIC systems in R7A do not appear to interact through direct AHL crosstalk (section 7.2.1), it seemed possible that these two systems might be intertwined at a regulatory level. To test if the MqsRIC system impacted ICE*Ml*Sym^{R7A} transfer, individual conjugation experiments were conducted using donors deleted for *mqsRIC*, recipients deleted for *mqsRIC* or matings in which both donors and recipients were deleted for *mqsRIC*. The pFAJ1700 (Tc^{R}) plasmid was introduced into each recipient strain to enable selection against Tc^{S} donor strains. Media lacking vitamins was used to select for exconjugants that had successfully received ICE*Ml*Sym^{R7A}, which carries vitamin synthesis genes for nicotinate, biotin and thiamine. Conjugation frequencies ranged between 4.09×10^{-6} and 5.75×10^{-6} for all mating experiments and no differences were observed for any of the strain combinations, indicating that the MqsRIC system is unlikely to greatly impact ICE*Ml*Sym^{R7A} transfer.

Table 4.1 Conjugation frequencies of ICE*Ml*Sym^{R7A} from R7A to R7ANS with or without *mqsRIC* genes

Donor	Recipient[a]	Exconjugants (per donor)	Standard deviation
R7A	R7ANS	5.75×10^{-6}	6.63×10^{-7}
R7AΔ*mqsRIC*	R7ANS	4.42×10^{-6}	9.10×10^{-7}
R7A	R7ANSΔ*mqsRIC*	4.48×10^{-6}	3.34×10^{-7}
R7AΔ*mqsRIC*	R7ANSΔ*mqsRIC*	4.09×10^{-6}	1.06×10^{-7}

[a]All recipient strains harboured the pFAJ1700 plasmid to allow for tetracycline resistance selection

4.2.3 MqsRIC may play a role in exopolysaccharide biosynthesis

Some QS systems such as the SinRI and ExpR of *Sinorhizobium* spp. play a regulatory role in polysaccharide production (73, 74). During growth of large cultures of R7A, we sporadically observed that strains lacking *mqsRIC* became extremely viscous compared to cultures grown with wild-type R7A, suggesting that there may be a difference in exopolysaccharide production. To explore this observation further, previously constructed pFUS2-based plasmid constructs were used to translationally fuse the *lacZ* gene to the 5' ends of open reading frames encoding the exopolysaccharide biosynthesis proteins ExoA, ExoU and ExoY in both R7A and R7AΔ*mqsR* backgrounds. The insertion of the pFUS2 constructs also disrupted the gene in each case. Expression from the *exoA*::*lacZ* fusion was unaffected in the absence of *mqsR*. However, *exoU* and *exoY* exhibited 1.4 times increase and 6.4 times decrease in expression, respectively, when compared to wild-type R7A (Figure 4.7). These data suggested there may be some impact on exopolysaccharide synthesis.

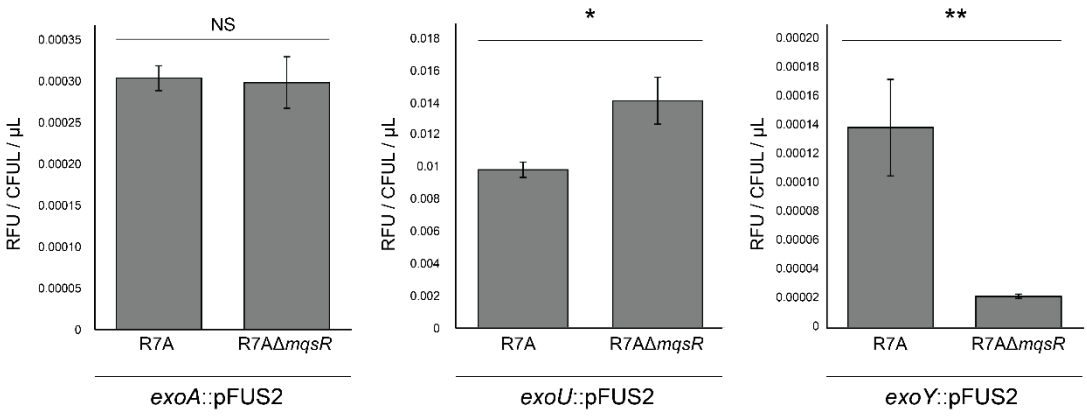


Figure 4.7 β-galactosidase activity of R7A and R7AΔ*mqsR* strains with the *lacZ*-carrying plasmid pFUS2 inserted into *exoA*, *exoU* and *exoY*. β-galactosidase was measured as previously described (126) using MUG. Bars represent the mean of three biological replicates and error bars represent the standard deviation from the mean. Student's t tests were used to compare Δ*mqsR* strains to wild-type with the same *exo* gene mutation. One asterisk denotes a statistical difference where $p<0.05$ and two asterisks denotes a statistical difference where $p<0.01$. NS denotes not significant ($p>0.05$)

4.3 Discussion

In this chapter, we aimed to identify phenotypes regulated by the MQS system. The *mqsRIC* genes are conserved across the genus *Mesorhizobium*. The level of conservation of these genes suggests an important role, however, deletion of all three genes did not appear to affect growth (Chapter 3), symbiotic phenotypes or horizontal gene transfer, at least under the laboratory conditions tested. *M. japonicum* R7A and *M. ciceri* CC1192 mutants were assayed for symbiosis phenotypes on their host legume plants lotus and chickpea, respectively. Deletions of *mqsRIC* and overexpression of *mqsIC* did not affect any of the symbiosis phenotypes tested, including nitrogen fixing efficiency, nodulation, and plant growth. Mating experiments also revealed that the MQS system had no effect on ICE*Ml*Sym^{R7A} transfer, regardless of whether the deletions were made in the donor, recipient, or both donor and recipient. Finally, *lacZ*-promoter fusions of three polysaccharide synthesis genes, *exoA*, *exoU* and *exoY*, showed that upon deletion of *mqsR*, *exoU* and *exoY* translation was increased and decreased, respectively.

The *mqsRIC* genes are highly conserved across over 200 strains of symbiotic and non-symbiotic strains of mesorhizobia (Chapter 3). The fact that these genes are present in non-symbiotic strains suggests that their conserved role has likely evolved in the absence of genes for nitrogen-fixing symbiosis and so may not be directly involved. However, recent comparative genome analyses have shown that symbiotic and non-symbiotic mesorhizobia share about 74% of their core genome (excluding ICE genes) and that most non-symbiotic strains have the capability to become a symbiont upon receival of an ICE harbouring symbiosis genes (106). Therefore, it seemed possible that the *mqsRIC* locus might impact symbiotic phenotypes indirectly. Nitrogen fixation efficiency was measured using acetylene reduction assays in chickpea plants inoculated with CC1192 wild-type and mutant strains, and no significant changes were observed. No significant or consistent differences were observed for plant mass or nodulation in chickpea or lotus plants inoculated with CC1192 strains or R7A strains, respectively. These data contradict research by Yang et al., who noted a significant nodulation deficiency in *M. loti*

NZP 2213 symbiosis with *Lotus corniculatus* when the *mqsI* homologue (there called *mrlI1*) was deleted (78). It is possible that had we left our lotus plants to grow longer, we may also have observed a difference in nodulation at time points past 30 days. While *M. japonicum* R7A and *M. loti* NZP 2213 are closely related it is also possible that phenotypes differ between these *Mesorhizobium* species. We again did not observe any consistent differences between nodulation phenotypes in chickpea plants inoculated with CC1192-derived strains carrying mutations in the MQS system. It should be noted that Yang et al. (78) observed symbiotic defects for all their constructed AHL-gene mutants, including strains deleted for homologues of *traI1* and the *traI2* pseudogene. Previous work in the Ronson laboratory has found no such symbiotic defects for R7A strains deleted for these same genes when used as inoculants for *Lotus*.

The transfer of ICE*Ml*Sym^{R7A} in *M. japonicum* R7A is well described and known to be regulated by TraRI – the LuxRI QS system encoded on the ICE (81-83, 85, 87). However, there is a precedence in other rhizobacteria for QS systems to be intertwined and form a hierarchical cascade, where separate QS systems regulate one another. *Rhizobium leguminosarum* and *R. etli* use a network of QS systems, where the chromosomally encoded CinRI proteins indirectly regulates expression of *traR*, resulting in plasmid conjugative transfer (70, 71). In this work, mating experiments revealed that deleting *mqsR* in either the R7A donor strain, the R7ANS recipient strain or both donor and recipient has no effect on the frequency of ICE*Ml*Sym^{R7A} transfer. Only the *mqsR* gene was deleted for these mating experiments and the *mqsI* and *mqsC* genes were left intact. Therefore, we can also draw the conclusion that the 2,4-trans-C12-HSL molecule produced by MqsIC does not directly induce TraR activity. These data also support those described in Chapter 7, where *in vitro* β-galactosidase assays confirmed that TraR is not induced by MqsIC AHLs and MqsR is not induced by TraI1 AHLs. These results and those in Chapter 7 strongly suggest that the MQS system does not interact with the TraRI system either through QS crosstalk or gene regulation.

After several incidental observations of what appeared to be an enhanced exopolysaccharide production phenotype in *mqsRIC* mutants, we wondered whether the MQS system plays a role in polysaccharide production. The synthesis of polysaccharides is often regulated by QS, especially in rhizosphere bacteria. In this chapter, we observed a slight significant increase in *exoU* transcription or translation, and a decrease in *exoY* transcription or translation in an *mqsR* deletion strain. ExoA, ExoY and ExoU are glycosyltransferases involved exopolysaccharide synthesis. ExoA and ExoY are involved in the first step, linking sugar residues together to form the EPS backbone. ExoU then works to link residues together forming the branching side chains (154). Previous works in *M. japonicum* R7A and have shown that mutations in early-stage synthesis genes, such as *exoA* and *exoY*, lead to delayed but successful nodulation and nitrogen fixation. Mutations in genes involved in mid- or late-stage EPS biosynthesis, such as *exoU*, lead to immature, ineffective nodules incapable of forming the infection thread necessary for successful infection and symbiosis (154, 155). In *Sinorhizobium meliloti*, c-di-GMP levels and QS regulate the production of two different EPSs – EPS I and EPS II. EPS I is synthesised by the *exo* gene products and affects cell aggregation and biofilm formation. EPS II is synthesised by proteins encoded by the *exp*/*wgx* gene clusters and generally promotes cell aggregation and biofilm formation. Gene mutations or repression of EPS I synthesis genes leads to increased production of EPS II and vice versa (157). In this work, deletion of *mqsR* resulted in downregulation of ExoY. We hypothesise that *M. japonicum* R7A may utilise a similar mechanism as *S. meliloti*, whereby a mutation in, or repression of *exoY* results in upregulated production of an alternative polysaccharide. So far, *M. japonicum* R7A is only known to produce one major EPS, synthesised by the gene products of the *exo* gene cluster (154, 155). However, R7A does also harbour an operon encoding homologues of *wgaG*, *wgaH*, *wgaI* and *wgaJ*, which produce EPS II sugar precursors in *S. meliloti* (158, 159). ExoX acts as a post-transcriptional repressor of ExoY, and the balance of *exoX* and *exoY* expression tightly regulates the production of EPS I (160, 161). The reduced transcription or

translation of *exoY* observed in the *mqsR* deletion mutant may be the result of an upregulation of *exoX*, which should also be tested using *in cis* promoter-only fusions. If R7A is indeed producing a second EPS and our *mqsR* mutation interrupted synthesis of the first EPS, the mucoid phenotype we have sporadically observed in *mqsRIC* mutants may suggest that the two EPSs act similarly to those of *S. meliloti*, where EPS I and EPS II regulate biofilm formation negatively and positively respectively.

In summary, the data in this chapter revealed that the MQS system likely does not play a regulatory role in symbiosis under the conditions tested. We also confirmed that deletion of *mqsR* does not affect ICE*Ml*Sym^{R7A} horizontal transfer. Preliminary promoter-*lacZ* fusions of three *exo* genes resulted in a decrease in transcription or translation of *exoY* upon deletion of *mqsR*. Further interrogation into EPS production in *mqsRIC* deletion mutants is warranted, as is investigations into the four *wgx* genes homologous with those in *S. meliloti*.

Chapter 5

MqsR activates a suite of novel non-coding RNAs

5.1 Introduction

Like QS, non-coding RNAs (ncRNAs) form a regulatory network to control a plethora of phenotypes in response to certain environmental stimuli. Bacterial ncRNAs are generally less than 300 bp in size and divided into two main groups based on function, the first works to block translation or degrade mRNA by base pairing to the target, and the second group of ncRNAs bind proteins (162, 163).

Sinorhizobium meliloti harbours several ncRNAs which respond to nitrogen limitation and regulates genes involved in nitrogen fixation (164-167). *Synechocystis sp*. – a marine dwelling nitrogen fixing cyanobacterium – harbours a ncRNA called NsiR4 which blocks translation of the glutamine synthase gene *gifA* in response to nitrogen stress (168). *Pseudomonas stutzeri* A1501 encodes three ncRNAs involved in nitrogen fixation phenotypes. NfiS and NfiR respond to stress signals and subsequently works to stabilise mRNA transcribed from the nitrogenase genes *nifK* and *nifD*, respectively. RsmZ positively regulates biofilm formation by sequestering the repressor protein RsmA from *pslA* and *sadC* mRNAs which encode polysaccharide and c-di-GMP, respectively (169).

There are a number of QS systems that are regulated by ncRNAs, such as that of *Vibrio* spp. which is by regulated quorum regulatory RNAs (Qrrs) (170). QS in *Vibrio harveyi* comprises three distinct systems producing three distinct autoinducers that all function in parallel to regulate five ncRNAs. The ncRNAs, Qrr1-5, are transcribed during low cell density and activate transcription of the master regulator encoding gene *aphA*, and repress transcription of the master regulator encoding gene *luxR*. During high cell density, the five ncRNAs are not transcribed, resulting in no transcription of *aphA* and derepression of *luxR* (171, 172). It was first believed that Qrr1-5 existed only to regulate QS itself, however it is now known that the ncRNAs also regulate a number of distinct targets, creating a collection of virulence factors, metabolic enzymes and other phenotypes indirectly controlled by QS (173).

The RhlRI QS system of *Pseudomonas aeruginosa* is intertwined with a number of ncRNAs. Under certain growth conditions, *rhlR* is transcribed from four different transcriptional start sites (TSSs) within the *rhlR* promoter region (174). PhrD is a 74 bp-long sRNA encoded in the *rhlR* promoter region and is highly conserved in *P. aeruginosa*. Transcripts from two of the four *rhlR* TSSs include a binding site for PhrD, and upon this ncRNA-mRNA interaction *rhlR* expression is increased (175). At least seven ncRNAs have also been found to be differentially regulated by AHLs in *P. aeruginosa*, one of which is located within the *rhlI* promoter region and named *rhlS* (176). RhlS regulates its own promoter, resulting in increased *rhlI* expression and increased QS, as well as post-transcriptionally regulating at least one unrelated gene encoding a pyoverdine receptor (176). To date, the function of the other QS-regulated ncRNAs in this system have not been functionally characterised.

In this chapter, RNA sequencing (RNAseq) was used to identify potential MqsRIC targets for further characterisation of the MqsRIC regulon. RNAseq was carried out for wild-type and *mqsRIC* deletion mutants of both *Mesorhizobium japonicum* R7A and *M. ciceri* CC1192, with the expectation that any changes observed in both species were more likely to be biologically significant. RNAseq revealed minimal changes in the mRNA abundance of protein-coding genes but did expose a network of ncRNAs that were highly expressed in wild-type strains but virtually absent in *mqsRIC* mutants. These ncRNAs were named here *mqsRNA1-mqsRNA7*. Putative MqsR binding sites were identified upstream of each ncRNA, and promoter-*lacZ* fusions confirmed them to be regulated by MqsRIC. Transcriptomics also showed a global downregulation of tRNA genes, however promoter-*lacZ* fusions of a representative tRNA indicated that this was likely not a direct result of MqsR regulation. Bioinformatic analyses of the mqsRNAs found that they were sporadically conserved in other mesorhizobia, with *mqsRNA7* being almost 100% conserved in the 205 strains that also encode *mqsRIC* (see chapter 3).

5.2 Results

5.2.1 RNAseq reveals deletion of *mqsRIC* in either *M. japonicum* R7A or *M. ciceri* CC1192 has only minor impacts on mRNA abundance for protein-coding genes

Given the lack of an obvious phenotype for the *mqsRIC* locus we carried out RNA sequencing (RNAseq) of markerless *mqsRIC* deletion mutants of *M. japonicum* R7A and *M. ciceri* CC1192 and corresponding wild-type strains. Two distinct species of mesorhizobia were analysed to enable higher confidence for changes observed in both species. RNAseq identified several differentially expressed loci, however, the genes with the largest differential expression were different between R7A and CC1192. Changes that were consistent between R7A and CC1192 included genes encoding two separate pilus assembly proteins, an operon of three genes including SagB/ThcOx family dehydrogenase and a response regulator. None of the identified differentially expressed genes had a sequence resembling the *mqs*-box upstream of the *mqsI*-promoter, suggesting that they were not likely to be directly regulated by MqsR. Interestingly, we did observe that numerous tRNA genes were downregulated in the *mqsRIC* mutants in both R7A and CC1192.

Table 5.1 Genes downregulated in *mqsRIC* mutants compared to wild-type, which were consistent between R7A and CC1192

	M. japonicum R7A				*M. ciceri* CC1192			
Protein product	**Locus tag**	**Log2 fold-change**	**Standard error**	**Adjusted p value**	**Locus tag**	**Log2 fold-change**	**Standard error**	**Adjusted p value**
Pilus assembly protein	R7A2020_24295	2.68	0.62	0.0013	A4R28_02395	2.80	0.301	3.72×10^{-18}
Pilus assembly protein	R7A2020_24290	2.93	0.41	3.93×10^{-10}	A4R28_02400	2.31	0.44	7.88×10^{-6}
SagB/ThcOx family dehydrogenase	R7A2020_05275	2.17	0.55	0.0052	A4R28_22850	2.063	0.21	1.65×10^{-19}
Hypothetical protein	R7A2020_05270	1.97	0.39	5.70×10^{-5}	A4R28_22855	1.17	0.25	0.00013
Hypothetical protein	R7A2020_05265	2.068	0.36	1.27×10^{-6}	A4R28_22860	2.13	0.32	2.61×10^{-9}
Response regulator	R7A2020_28340	2.16	0.17	2.20×10^{-32}	A4R28_29265	2.03	0.22	3.29×10^{-17}

5.2.2 MqsR activates transcription from *mqs*-box sequences upstream of non-coding RNA genes

Because of the paucity of highly differentially regulated protein-coding genes observed in either the R7A or CC1192 genomes, we decided to manually inspect the abundance of mapped RNAseq reads in non-coding regions of the genome in both strains. A 124 bp-long non-coding region was identified in R7A upstream of the *pstS* gene with an average of 17,271 reads mapped to the wild-type and 51 mapped to the *mqsRIC* mutant (Figure 5.1). The corresponding non-coding region in CC1192 had an average of 27,427 reads mapped to the wild-type and 252 mapped to the *mqsRIC* mutant (Figure 5.1C,D). Inspection of the region upstream of this putative non-coding ncRNA, later named *mqsRNA7*, revealed an inverted repeat sequence resembling the *mqs*-box, located approximately 50 bp upstream of the potential transcriptional start site estimated from the mapped RNA reads. To address the possibility that this region encodes an unannotated protein, the *mqsRNA7* sequence plus 150 bp either side was analysed with BLASTx. The search resulted in no matches, solidifying our hypothesis that this was in fact a non-coding RNA and not an unannotated protein.

The putative *mqs*-box of *mqsRNA7* was not identical to that upstream of *mqsI* and therefore we collected the sequence regions upstream of *mqsI* and the putative *mqs*-box upstream of *mqsRNA7* from both R7A and CC1192 and combined them into a weight matrix using the MEME software (177). Next we used this matrix to search for additional motifs in the R7A and CC1192 genomes using MAST software (178). With each potential hit identified, we searched for differentially expressed RNAs in the mapped RNAseq data for R7A and CC1192. Once a new differentially expressed ncRNA was identified, we added its corresponding *mqs*-box to our alignments and constructed a new MEME matrix and conducted further MAST searches with the updated motif. This process was carried out iteratively until no new regions were identified. In total, seven differentially expressed ncRNAs were identified downstream of *mqs*-box sequences. RNAseq mapping diagrams of each ncRNA can be found in Appendix 4.

Table 5.2 Non-coding RNAs identified by RNAseq to be regulated by MqsRIC

	***M. japonicum* R7A**				***M. ciceri* CC1192**			
ncRNA	**Coordinates**	**Strand**	**Fold-change**	**Adjusted p value**	**Coordinates**	**Strand**	**Fold-change**	**Adjusted p value**
mqsRNA1	4249847-4249896	-	12.47					
mqsRNA2	2910774-2910983	-	230.72	8.98×10^{-42}	1243647-1243708	-	73.52	1.53×10^{-16}
mqsRNA3	218241-218444	-	20.25	3.16×10^{-13}	1322010-1322163	-	25.28	6.27×10^{-20}
mqsRNA4	4126745-4127472	-	20.25	2.39×10^{-21}	1856750-1856832	-	69.55	9.93×10^{-63}
mqsRNA5	4688087-4688832	+	10.01	2.76×10^{-21}	756499-756694	-	0.50	0.047
mqsRNA6	2791745-2791872	-	3.14	4.10×10^{-4}	2705770-2705863	+	188.71	1.82×10^{-22}
mqsRNA7	1387207-1387330	+	374.81	6.23×10^{-108}	4010351-4010630	-	195.36	1.53×10^{-16}

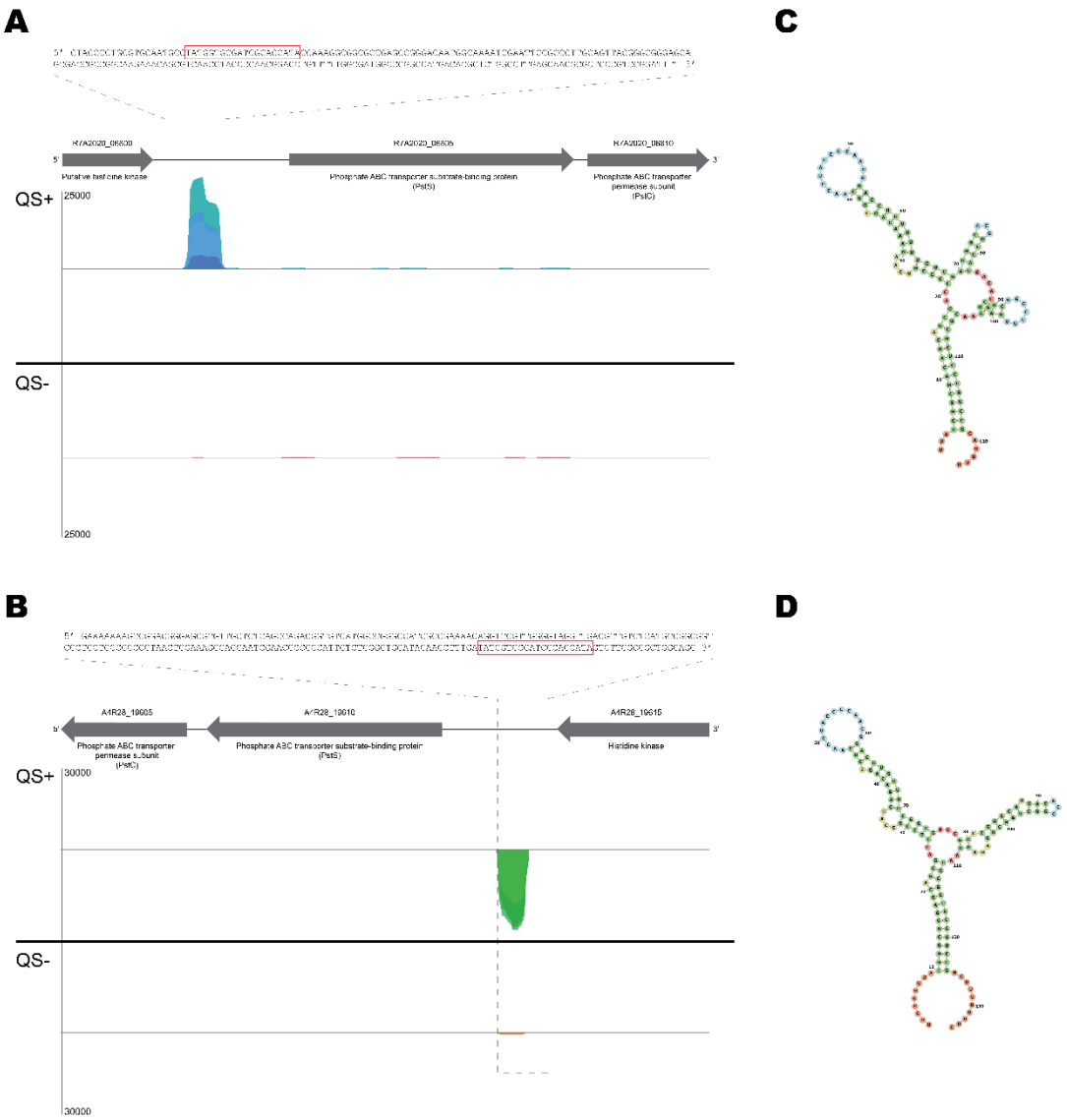


Figure 5.1 The *mqsRNA7* ncRNA is differentially regulated in *mqsRIC* deletion mutants. (A,B) RNAseq reads mapped to the (A) R7A and (B) CC1192 reference genomes. Grey block arrows represent genes (to scale), labelled with the locus ID and gene product above and below respectively. DNA sequence of the *mqsRNA7* gene is displayed above, with the putative *mqs*-box highlighted by a red box. (A) Blue and (B) green coloured reads are derived from wild-type strains, (A) red and (B) orange coloured reads are derived from *mqsRIC* mutant strains. (C,D) Predicted secondary structures of *mqsRNA7* from (C) R7A and (D) CC1192, predicted using RNAfold (179). Nucleotides are coloured green (stems), red (multiloops), yellow (interior loops), blue (hairpin loops), orange (5' and 3' unpaired region). RNAseq read maps and predicted secondary structures for the other mqsRNAs are found in Appendix 4.

5.2.3 Expression of non-coding RNAs is dependent on MqsR

To confirm the regulation of the identified ncRNAs was controlled by the MqsRIC system, promoter regions of each ncRNA from *M. japonicum* R7A were cloned into pSDz upstream of *lacZ*. Cloning of the *mqsRNA4* promoter region was unsuccessful, and due to time constrains this ncRNA was not included.

As well as the ncRNAs, the tRNA-*leu* gene promoter was also cloned into pSDz to determine whether MqsR transcriptional regulation was involved in the global downregulation of tRNA abundance observed in the *mqsRIC* mutants. These reporter plasmids were introduced into R7ANS and R7ANSΔ*mqsR* to determine if expression from the ncRNA and tRNA promoters was dependent on the presence of MqsR. Of the six mqsRNA promoters and one tRNA-*leu* promoter tested, all six mqsRNA promoters exhibited significantly reduced *lacZ* expression in the absence of MqsR. The *mqsR* gene was then cloned into the pPR3 plasmid, which harbours the constitutive *nptII* promoter. This overexpression construct was subsequently introduced back into the Δ*mqsR* strains to complement the mutation, and mqsRNA expression was restored (Figure 5.2). The representative tRNA gene promoter was not regulated by MqsR, suggesting the downregulation of tRNAs in the *mqsRIC* mutants may be an indirect effect or artifactual.

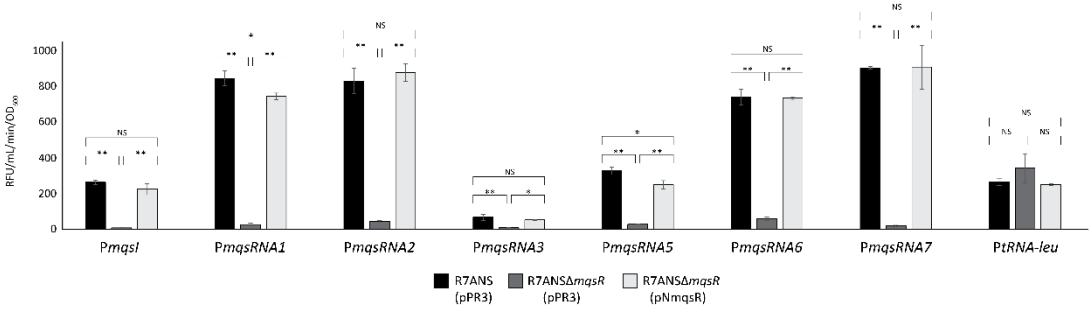


Figure 5.2 mqsRNA promoter expression is dependent on *mqsR*. β-galactosidase activity of R7ANS-derived strains with or without *mqsR* harbouring pSDz constructs with various ncRNA promoters driving *lacZ* expression. β-galactosidase was measured as previously described (126) using MUG. Bars represent the mean of three biological replicates and error

bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each background strain expressing *lacZ* from the same promoter. One asterisk denotes $p<0.05$ and two asterisks denotes $p<0.01$. NS denotes not significant ($p>0.05$)

5.2.4 Expression of non-coding RNAs is activated by MqsIC-produced AHLs

In order to further confirm that the AHL produced by MqsIC plays a role in controlling the ncRNA expression, *mqsR* was cloned into the IPTG-inducible MCS of the pSDz constructs containing the PmqsRNA-*lacZ* fusions. These constructs were then introduced into R7ANSΔ*mqsRIC* with and without a plasmid constitutively expressing *mqsIC* - pNmqsIC. The mqsRNAs identified to be differentially expressed in the presence and absence of MqsR also exhibited only low basal level expression in the absence of MqsIC, and increased expression to varying degrees when the AHL synthase genes were reintroduced (Figure 5.3). These data suggest that the small non-coding mqsRNAs are directly activated by the MqsRIC system. Again, (as in section 5.2.3), expression from the tRNA-*leu* gene promoter was not affected by the presence or absence of MqsIC.

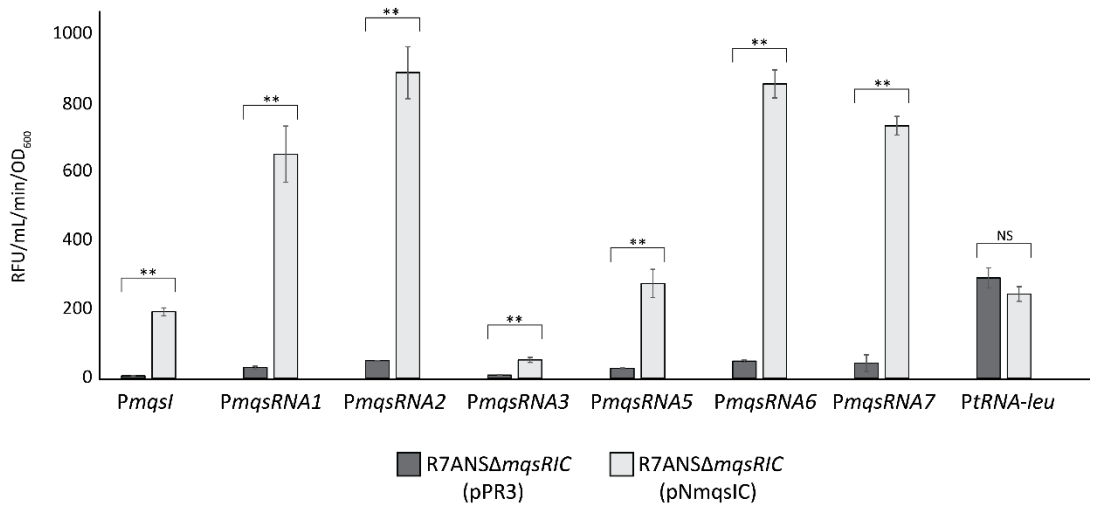


Figure 5.3 mqsRNA promoter expression is dependent on mqsIC. β-galactosidase activity of R7ANSΔ*mqsRIC* strains with or without the overexpression of *mqsIC* harbouring pSDz constructs with various ncRNA promoters driving *lacZ* expression. β-galactosidase was measured as previously described (126) using MUG. Bars represent the mean of three

biological replicates and error bars represent the standard deviation from the mean. A student's t test was used to compare strains carrying the same pSDz constructs. Two asterisks denote $p<0.01$. NS denotes not significant ($p>0.05$).

5.2.5 Refinement of the *mqs*-box sequence

Each mqsRNA identified to be regulated by MqsRIC had a *mqs*-box-like sequence located upstream of the likely transcriptional start site. We constructed an alignment of each of the newly identified *mqs*-boxes in R7A and CC1192. For each *mqs*-box, a 20 bp region centred around the 8 bp inverted repeat was aligned. This revealed that two nucleotides in the second and seventh position of the 8 bp inverted repeat frequently differed between *mqs*-boxes, while the remaining six bp were 100% conserved across all sequences (Figure 5.4). Interestingly, the mqsRNA that exhibited the highest expression was *mqsRNA7*, which had a distinct *mqs*-box sequence with extended dyad symmetry outside of the 8 bp region compared to the *mqs*-box of *mqsI*. All promoters aside from that of *mqsRNA3* exhibited stronger MqsR-dependent activation than *mqsI*, suggesting the *mqs*-box upstream of *mqsI* is sub-optimal for MqsR transcriptional activation. From this alignment we refined the mqs-box motif to be "GYGATCRC" or "VYGYGATCRCRC" with extended dyad symmetry included.

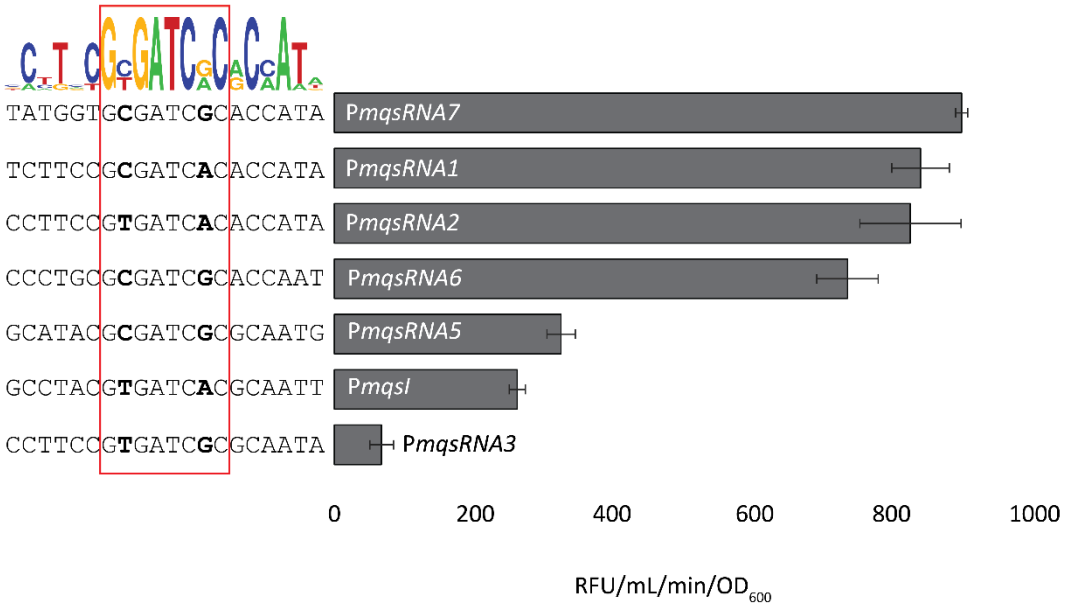


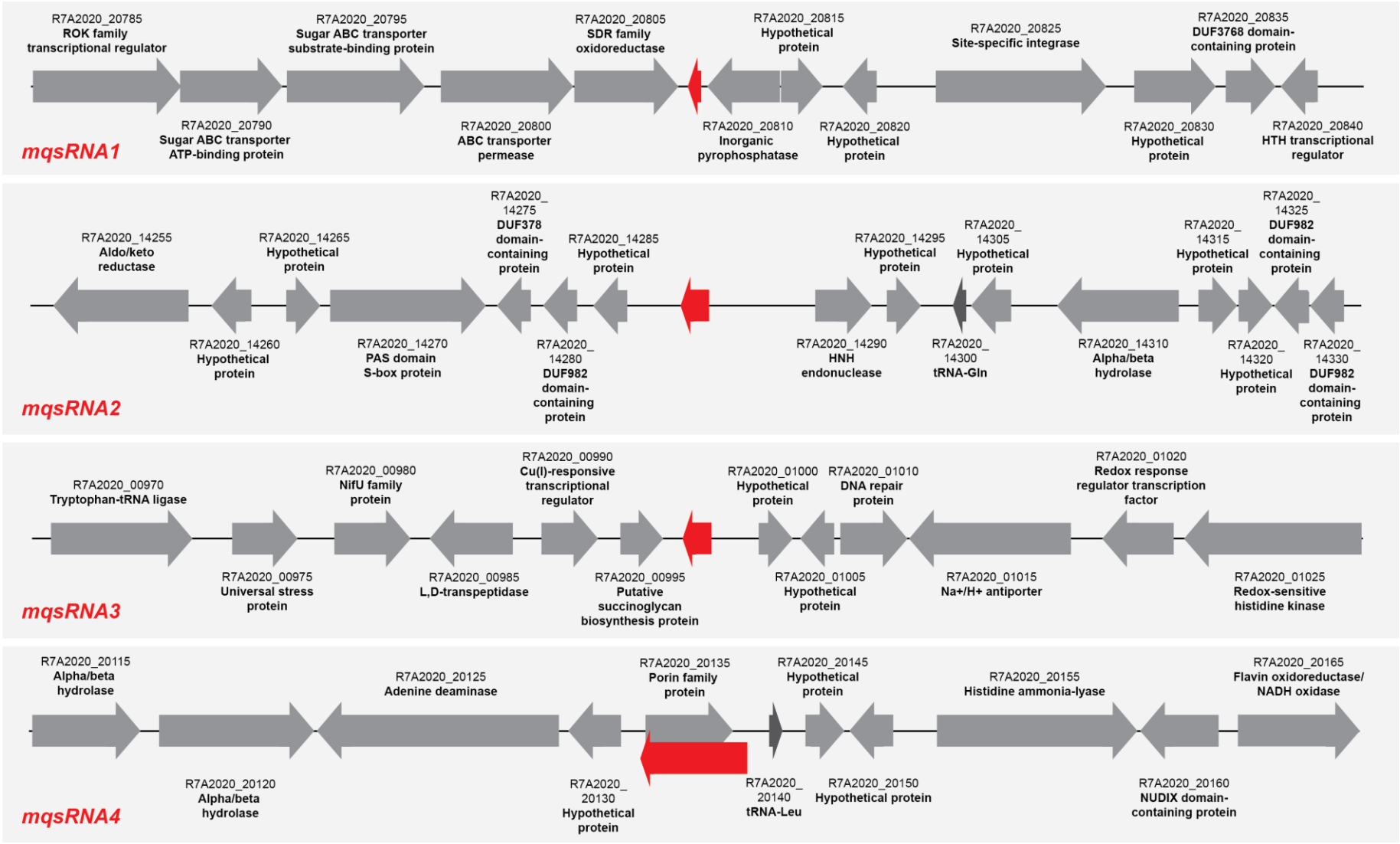
Figure 5.4 Comparison of promoter regions and *mqs*-boxes. The MqsR binding site for each mqsRNA and *mqsI* is displayed with the consensus sequence at the top, *mqs*-box sequences are highlighted in the red box. Data collected from promoter-*lacZ* fusion assays are displayed next to the corresponding *mqs*-box sequence in descending order from highest expression to lowest expression.

5.2.6 MqsRNAs are conserved in *Mesorhizobium* spp.

RNAseq analysis showed the abundance of *mqsRNA7* transcription to be approximately 300-fold higher in both the wild-type R7A and CC1192 compared to the respective *mqsRIC* mutant strains. To investigate if this ncRNA was conserved in other *Mesorhizobium* spp. and if there was a conserved genome location, the *mqsRNA7* nucleotide sequence from R7A was used as a query in a BLASTn search against the database of 214 *Mesorhizobium* spp., which included both symbiotic and non-symbiotic species and strains. Of the 214 mesorhizobia, 207 encoded *mqsRNA7* with 85-100% identity. Interestingly, the seven strains that did not encode any *mqsRNA7* homologues were the same strains that do not encode MqsRIC, thus providing corroborating evidence that the MqsRIC system and these ncRNAs are evolutionarily linked. Further investigation revealed all *mqsRNA7* regions to be located upstream of *pstS-pstC-pstA-pstB* genes encoding homologues of the well characterised phosphate transport system (Pst). BLASTn analysis of the

other mqsRNAs revealed them to be more sporadically conserved than *mqsRNA7*. BLASTn analysis of *mqsRNA1* in the same database of 214 mesorhizobia found 12 homologues, 11 of which were located downstream of a gene encoding an inorganic pyrophosphatase and upstream of a SDR family oxioreductase gene. *mqsRNA2* and *mqsRNA5* both contained conserved RNA sequence regions that were detected in 36 and 58 strains respectively and *mqsRNA3*-like sequences were identified in 24 strains. All *mqsRNA5* homologues were located upstream of *uxe-udp* which encodes UDP-xylose 4-epimerase and UPD-glucose 4-epimerase respectively. Both these enzymes are involved in biosynthesis of the arabinose-containing polysaccharide APS in *Sinorhizobium meliloti* (180).

We also found that a 40 bp region present in the 3' end of *mqsRNA6* was present 235 times within the database of 214 mesorhizobia, with up to nine copies present in a single strain. This sequence was invariably present at the 3' end of genes and contained a GC-rich inverted repeat resembling a transcriptional terminator. Interestingly, 17 strains (including CC1192) carried a version of *mqsRNA6* containing a central 100 bp deletion. Each of these shorter *mqsRNA6* sequences had an identifiable *mqs*-box located upstream. All *mqsRNA6* sequences were located between convergent genes encoding a putative saccharide export ABC transporter (ExsA) and cell division proteins MinCDE.


R7A2020_20785
ROK family transcriptional regulator
R7A2020_20790
Sugar ABC transporter ATP-binding protein
R7A2020_20795
Sugar ABC transporter substrate-binding protein
R7A2020_20800
ABC transporter permease
R7A2020_20805
SDR family oxidoreductase
R7A2020_20810
Inorganic pyrophosphatase
R7A2020_20815
Hypothetical protein
R7A2020_20820
Hypothetical protein
R7A2020_20825
Site-specific integrase
R7A2020_20830
Hypothetical protein
R7A2020_20835
DUF3768 domain-containing protein
R7A2020_20840
HTH transcriptional regulator
mqsRNA1
R7A2020_14255
Aldo/keto reductase
R7A2020_14260
Hypothetical protein
R7A2020_14265
Hypothetical protein
R7A2020_14270
PAS domain S-box protein
R7A2020_14275
DUF378 domain-containing protein
R7A2020_14280
DUF982 domain-containing protein
R7A2020_14285
Hypothetical protein
R7A2020_14290
HNH endonuclease
R7A2020_14295
Hypothetical protein
R7A2020_14300
tRNA-Gln
R7A2020_14305
Hypothetical protein
R7A2020_14310
Alpha/beta hydrolase
R7A2020_14315
Hypothetical protein
R7A2020_14320
Hypothetical protein
R7A2020_14325
DUF982 domain-containing protein
R7A2020_14330
DUF982 domain-containing protein
mqsRNA2
R7A2020_00970
Tryptophan-tRNA ligase
R7A2020_00975
Universal stress protein
R7A2020_00980
NifU family protein
R7A2020_00985
L,D-transpeptidase
R7A2020_00990
Cu(I)-responsive transcriptional regulator
R7A2020_00995
Putative succinoglycan biosynthesis protein
R7A2020_01000
Hypothetical protein
R7A2020_01005
Hypothetical protein
R7A2020_01010
DNA repair protein
R7A2020_01015
Na+/H+ antiporter
R7A2020_01020
Redox response regulator transcription factor
R7A2020_01025
Redox-sensitive histidine kinase
mqsRNA3
R7A2020_20115
Alpha/beta hydrolase
R7A2020_20120
Alpha/beta hydrolase
R7A2020_20125
Adenine deaminase
R7A2020_20130
Hypothetical protein
R7A2020_20135
Porin family protein
R7A2020_20140
tRNA-Leu
R7A2020_20145
Hypothetical protein
R7A2020_20150
Hypothetical protein
R7A2020_20155
Histidine ammonia-lyase
R7A2020_20160
NUDIX domain-containing protein
R7A2020_20165
Flavin oxidoreductase/ NADH oxidase
mqsRNA4

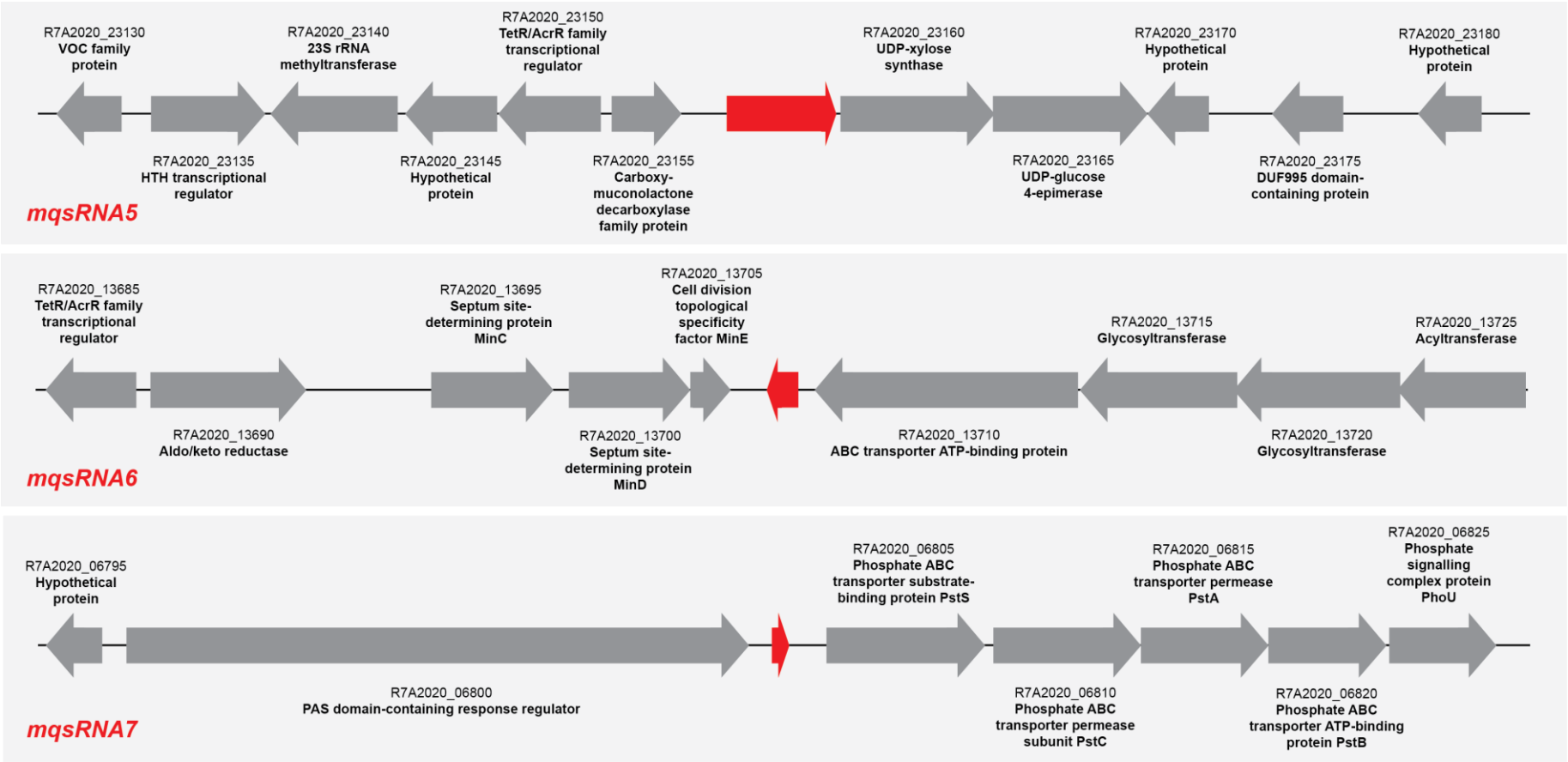


Figure 5.5 Genomic context of *mqsRNA1-7* in *M. japonicum* R7A. A 10 kb region centralised on each mqsRNA from M. japonicum R7A is displayed. Each mqsRNA shown in red, surrounding genes shown in grey and tRNA genes in dark grey. The locus tag and gene product are shown above or below each gene.

5.3 Discussion

In this chapter, RNAseq revealed a suite of ncRNAs to be directly regulated by MqsRIC. Transcription of the ncRNAs, here named mqsRNAs, was abolished in *mqsRIC* deletion mutants in both *M. japonicum* R7A and *M. ciceri* CC1192. Promoter-*lacZ* fusions using the pSDz plasmid confirmed that expression from the mqsRNA promoters was dependent on MqsR and induced by MqsIC-derived AHLs. The MqsR binding site called the *mqs*-box was refined to be "GYGATCRC" or "VYGYGATCRCRC" with extended dyad symmetry. The mqsRNAs were identified across a database of symbiotic and non-symbiotic *Mesorhizobium* strains with varying levels of conservation, and the genetic context surrounding each mqsRNA was preserved.

There are many examples of ncRNAs regulating QS however there are few known examples of QS directly regulating small or ncRNAs. The Agr QS system of *Staphylococcus aureus* comprises a complex circuit regulated by the autoinducing peptide (AIP). Upon reaching a certain threshold concentration of AIP, expression of *argC-A* is activated and expressed in a positive feedback loop. As well as increasing transcription of the phenol-soluble modulin peptide toxins, activated ArgA also induces transcription of RNAIII. RNAIII generally regulates target genes by antisense base-pairing in the 5' untranslated region inhibiting translation, resulting in either direct or indirect derepression of a range of virulence (181-187). Sequence variation of *agr* genes across staphylococci allows production of AIPs with different specificities, resulting in cross-inhibition of non-self Agr systems (181, 188-190).

A group of small RNAs (sRNAs) have been described in *Pseudomonas aeruginosa* to be regulated by the RhlRI QS system (176). The sRNA *rhlS* is transcribed from same the transcriptional start site as *rhlI* and represents a truncated isoform of *rhlI*. Most of the transcription events from the *rhlI* promoter produces *rhlS*, with only 3-15% producing the full-length *rhlI* mRNA. The *rhlS*

sRNA regulates its own promoter, as well as post-transcriptionally regulating an unrelated gene encoding a pyoverdine receptor. Six other sRNAs were also regulated by RhlRI and one was subsequently characterised to regulate a host of *P. aeruginosa* virulence factors (176). A QS system of *Yersinia enterocolitica* utilises a small ncRNA to regulate swarming motility (191).

The *mqsRNA7* is present in almost all *Mesorhizobium* species we analysed in this thesis and is always present upstream of the *pstS-pstC-pstA-pstB* operon. This highly conserved genomic context may hint at a potential involvement in the Pst phosphate transport system. It has been suggested that the phosphate stress response and the nitrogen stress response is intertwined and cross-regulated (192, 193) and a small RNA of Sinorhizobium meliloti is a part of a negative feedback loop that recognises and responds to nitrogen starvation (164, 166, 194, 195). Multiple studies indicate that transcription of the nitrogen-responsive sRNA is cell-density dependent under normal lab conditions, however none have confirmed or suggested that it may be regulated by quorum sensing to our knowledge. The mqsRNAs possibly act in response to phosphate, nitrogen, or phosphate and nitrogen limiting conditions.

We identified seven RNAs regulated by MqsRIC, and the genomic context of all seven were conserved among the strains that encoded them. It is not clear whether the seven ncRNAs are redundant, perhaps *mqsRNA7* performs the main role and the others have evolved as a fallback system, or if they all serve a different function. All *mqsRNA5* homologues in our Mesorhizobium database were located upstream of genes encoding UDP-xylose 4-epimerase and UPD-glucose 4-epimerase, both of which are involved in polysaccharide synthesis in *S. meliloti* (180). The UPD-glucose 4-epimerase downstream of *mqsRNA5* in R7A shares 99% identity with *exoB* of *M. japonicum* MAFF 303099 which supplies UDP-galactose to EPS backbone synthesis and the synthesis of other polysaccharides (196). Previous studies have shown that EPS production in *S. meliloti* and *Rhizobium leguniosarum* is influenced by phosphate limitation via the global regulator MucR/RosR (196-198). The location of *mqsRNA5*

amongst these polysaccharide genes as well as the MqsR-mediated differential expression of exoY and exoU (Chapter 4) further cement our hypothesis that the MqsRIC system is involved in polysaccharide production, perhaps in response to phosphate limiting conditions.

In summary, data in this chapter uncovered a suite of non-coding RNAs regulated by the MQS systems in *M. japonicum* R7A and *M. ciceri* CC1192. The mqsRNAs were discovered by manually examining RNAseq data and confirmed to be directly regulated by MqsR and MqsIC with promoter-*lacZ* fusions and β-galactosidase assays. The mqsRNAs were found to be conserved to varying degrees throughout our database of 214 Mesorhizobium strains with *mqsRNA7* homologues sharing almost 100% nucleotide identity in 207 strains, only strains lacking the *mqsRIC* genes also did not encode *mqsRNA7*. The locations of the identified mqsRNAs in the genome hints at possibles regulatory roles in phosphate sensing/transport or polysaccharide synthesis.

Chapter 6

Epigenetic regulation of quorum sensing and ICE*Ml*Sym^{R7A} transfer

6.1 Introduction

The most well characterised LuxRI-type QS system in *Mesorhizobium* species is the TraRI system, which regulates excision and transfer of ICE*Ml*Sym^{R7A} in *M. japonicum* R7A and ICE*Mc*Sym^{1271} in *M. ciceri* WSM1271 (80, 82). TraR and TraI of *Mesorhizobium* are homologous to TraR and TraI of *A. tumefaciens* and *R. leguminosarum* which also regulate the conjugative transfer of plasmids pTi and pRL1JI respectively (54, 71). Other symbiosis ICEs and symbiosis plasmids in the *Mesorhizobium* genus encode TraR and TraI homologues (129), suggesting a similar mode of control on horizontal gene transfer.

TraI1 of ICE*Ml*Sym^{R7A} catalyses the production of 3-oxo-C6-HSL as the main AHL species. TraR recognises and binds the AHL to form a TraR:AHL complex which upregulates TraI1 expression and stimulates excision and transfer of ICE*Ml*Sym^{R7A} through DNA binding of two *tra*-boxes upstream of *traI1* and upstream of the pseudogene *traI2* (82). Within the *traI2* operon is *msi172* and *msi171* in which a programmed +1 ribosomal frameshift occurs in approximately 4-12% of translation events and produces the protein FseA (83). FseA then activates *rdfS* transcription leading to ICE*Ml*Sym^{R7A} excision and transfer (85, 199). TraRI QS and ICE*Ml*Sym^{R7A} excision and transfer occurs in only 1-5% of an R7A population. When QS and ICE*Ml*Sym^{R7A} transfer is switched off, the antiactivator QseM binds TraR and FseA, blocking the transcriptional activators responsible for QS and ICE excision. Downstream of *qseM* is the divergently encoded gene for a DNA-binding protein QseC which regulates QseM expression (81). QseC and its promoter region resemble controller proteins (C proteins) of type II restriction modification systems. C proteins form a dimer and bind palindromic DNA regions called operators. These operators overlap promoter regions and form a switch which facilitates the delayed expression of the restriction endonuclease following entry to a new cell (200, 201). C proteins are diverse and bioinformatic studies have clustered various C protein homologues based on their operator sequences. QseC most resembles C.PvuII (motif 4). Like C proteins of RM systems, QseC dimers bind an operator site O_L and then cooperatively bind a second adjacent operator site O_R. O_L and O_R are located within the overlapping divergent promoters of

qseC and *qseM* (81, 87) (Figure 6.1). Deletion of *qseC* results in increased *qseM* expression and repression of QS and ICE*Ml*Sym^{R7A} excision and transfer (81). A gene encoding a second distinct C protein is located downstream of *qseC* called *qseC2*. QseC2 resembles the C.EcoRV family (motif 7) of C proteins and QseC2 dimers bind distinct operator sites $O2_L$ and $O2_R$. QseC2 also exhibits slightly stronger affinity for the left operator $O2_L$ (87, 202).

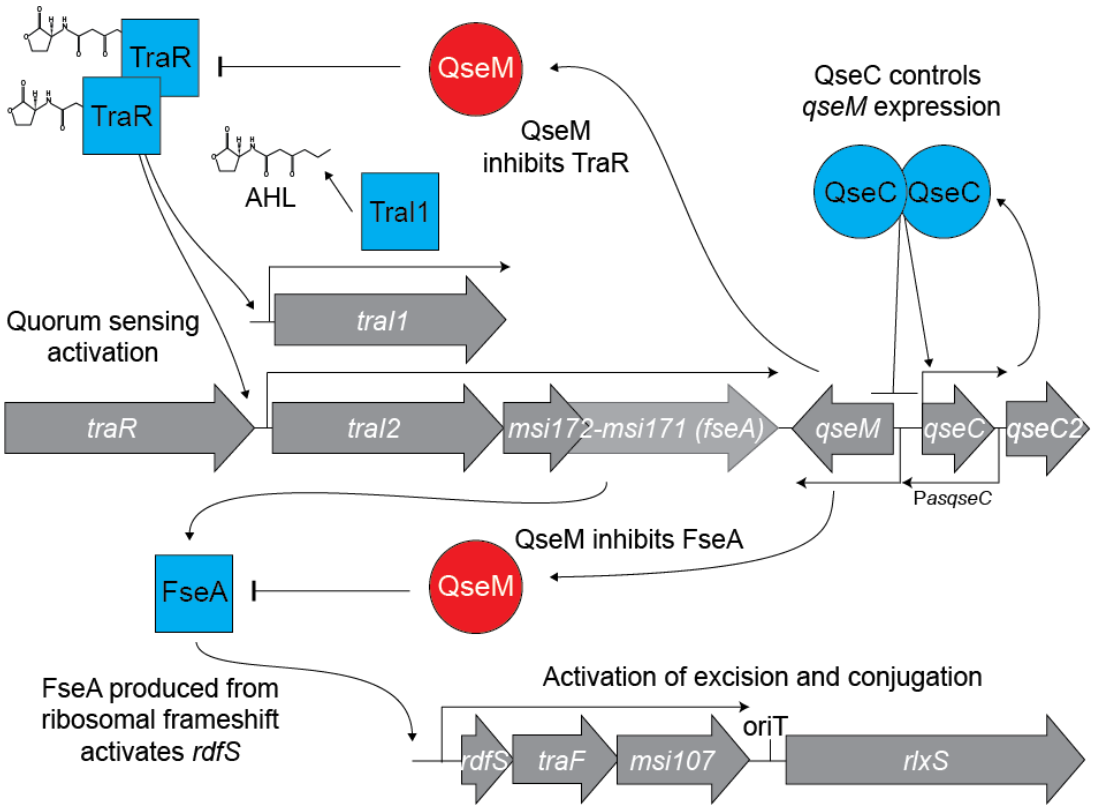


Figure 6.1. Model of QS and excision and transfer in ICE*Ml*Sym^{R7A}. Gene maps of regions spanning *traR-qseC2* and *rdfS-rlxS* of ICE*Ml*Sym^{R7A}. Proteins coloured blue positively regulate QS and ICE excision, and the antiactivator QseM is coloured red. TraI1 produces 3-oxo-C6-HSL molecules which bind and stabilises a dimerised TraR, the TraR:AHL complex binds two *tra*-boxes upstream of *traI1* and the likely pseudogene *traI2* resulting in upregulation of *traI1* and therefore increased AHL production, and upregulation from the *traI2* promoter. The *traI2* operon includes *msi172* and *msi171*, where a programmed ribosomal frameshift occurring in 4-12% of translation events causes translation of FseA (83). FseA then stimulates ICE*Ml*Sym^{R7A} excision and subsequent conjugative transfer. Figure adapted from (87).

Variants of R7A – called R7A* – can be isolated from R7A populations that are constitutively 'on' for QS and are able to induce violacein production in the C4-C8 AHL biosensor strain *Chromobacterium violaceum* CV026 (87). R7A* cells spontaneously arise in approximately 1-5% of stationary-phase R7A populations. Once R7A* cells have switched to this phenotype, they appear to remain indefinitely in this quorum-sensing active state. The R7A* state does not appear to be a result of genetic changes to the R7A genome. Two independent isolates of R7A* (R7A* and R7A*2) and R7A have been deep sequenced with both long and short-read sequencing and no causal genetic changes have been identified. RNAseq of the *traR-qseC2* region has shown that *qseM* expression is 8.7-fold higher in R7A compared to that in R7A*. RNAseq also revealed a previously undiscovered mRNA transcript antisense to the *qseC* gene to be present in R7A and not in R7A*, the transcription of which begins upstream of $O2_L$ and $O2_R$ of QseC2 (87). Transcription of this small RNA – *asqseC* – was 14.3-fold higher in R7A, however the function is so far unknown.

In this chapter, we aimed to characterise the phenotypes of R7A* and elucidate the mechanisms of the apparent epigenetic regulation of QS and ICE*Ml*Sym^{R7A} transfer. We discovered that as well as increased AHL production, R7A* also exhibits upregulated ICE*Ml*Sym^{R7A} excision and transfer to R7ANS. The experiments in this chapter are published as a co-first author paper in *Nucleic Acids Research* titled "An epigenetic switch activates bacterial quorum sensing and horizontal transfer of an integrative and conjugative element" (2022 25;50(2):975-988. doi: 10.1093/nar/gkab1217)

6.2 Results

6.2.1 Optimisation of a quantitative PCR assay for detection of ICE*Ml*Sym^{R7A} excision

A quantitative PCR (qPCR) assay previously developed by Ramsay et al. (85) is used to measure the amount of cells within a population that are in a transfer-ready state and harbouring an excised ICE*Ml*Sym^{R7A}. The established TaqMan probe-based assay amplifies the empty integration site on the chromosome (*attB*) and a region on the circularised ICE*Ml*Sym^{R7A} (*attP*). A third target is the chromosomally located *melR* gene located upstream of the ICE*Ml*Sym^{R7A} integration site and used to determine chromosome copy number for quantification of the relative *attP*/*attB* abundance (85, 130). To determine the amplification efficiencies of each individual PCR reaction, the serially diluted pJR201 plasmid – which contains a single copy of *melR*, *attB* and *attP* – is used as a template to produce standard curves. Efficiencies are then calculated from standard curves as previously described (85). Amplification efficiencies describe the hypothetical number of DNA copies derived from one individual template after each cycle. Therefore, an efficiency of two or as close to two as possible is ideal.

In this work we re-established this assay using oligonucleotide reagents purchased from IDT with slightly different chemistries. Probes and primers were designed using the same sequences as were used in the previously established assay (85) (see materials and methods). Standard curves were carried out for each PCR reaction (*attP*, *attB* and *melR*). The *attP* and *attB* reactions using ten-fold serial dilutions of pJR201 as a template exhibited a linear standard curve and calculated amplification efficiencies of 2.00 and 1.99 respectively. However, the *melR* reaction exhibited an aberrant standard curve that indicated it was not able to accurately measure templates at higher concentrations (Figure 6.2A). The probes purchased for this study were 5' FAM dye labelled and double quenched, and probes used in the previous assay were minor-groove-binding (MGB) probes. It was thought that the *melR*

probe used here may not bind its target as strongly as the MGB probe used previously. A new *melR* probe was designed to extend the region between the two quenchers by 6 bp, lengthening the probe to promote binding and compensate for the lack of MGB-conjugate. Standard curves carried out using the new probe improved the amplification efficiency, with final efficiencies for *attP*, *attB* and *melR* at 2.09, 2.03 and 2.03 respectively (Figure 6.2B).

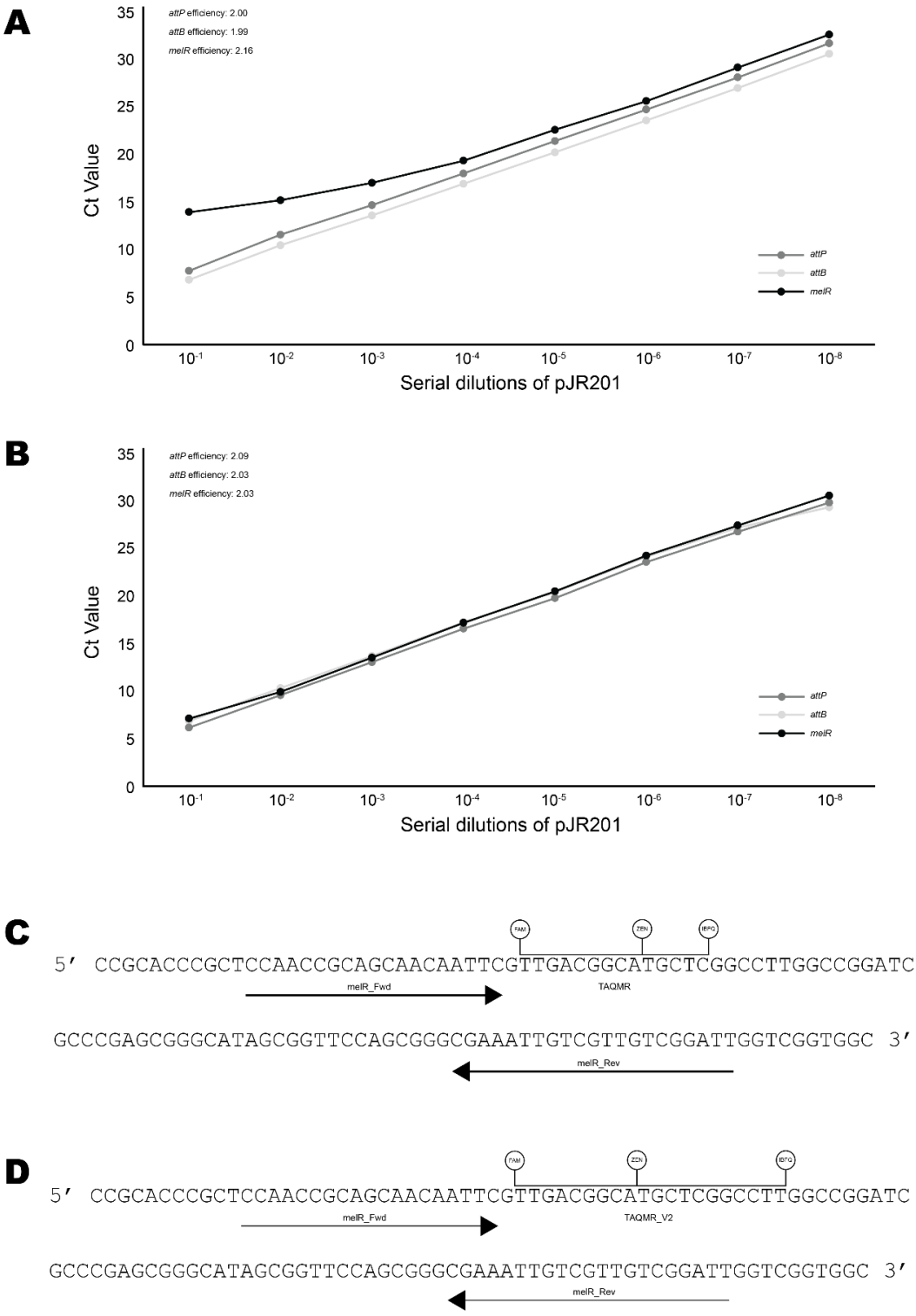


Figure 6.2 Optimisation of qPCR probes (A) Standard curves of *attP*, *attB* and *melR* PCR reactions with primers and probes of the same sequence as those used in the previously established assay. (B) Standard curves of *attP*, *attB* and *melR* PCR reactions with the new, extended *melR* probe. (C) Depiction of the first *melR* probe (TAQMR) binding site and *melR* primer binding sites (arrows). (D) Depiction of the new *melR* probe (TAQMR_V2) binding site and *melR* primer binding sites (arrows).

6.2.2 R7A* exhibits similar excision and AHL production phenotypes as a *qseM* mutant

Wild-type R7A does not produce detectable AHLs and carries an excised ICE*Ml*Sym^{R7A} in approximately 0.005% and 1% of R7A populations during log and stationary-phase respectively (Figure 6.3). Deletion of *qseM* in R7A produces a strain that constitutively produces TraI1-derived AHLs and carries excised ICE*Ml*Sym^{R7A} in 100% of cells in stationary-phase cultures (81). We repeated these results here in this work and examined AHL production and ICE*Ml*Sym^{R7A} excision frequencies in R7A*.

Chromobacterium violaceum CV026 is a biosensor strain used for detecting C4-C8 AHLs (203), and this strain was used in bioassays to detect TraI1-derived AHLs. Supernatants from R7A* and the independently isolated R7A*2 induced a purple violacein zone on the CV026 bioassay plate indicating the presence of C4-C8 AHLs. QPCR measurements of these strains revealed ICE*Ml*Sym^{R7A} was excised in approximately 3% and 72% in log and stationary-phase respectively (Figure 6.3). When *traR* was deleted from R7A*, AHL production was abolished and ICE*Ml*sym^{R7A} excision was reduced to below wild-type levels, confirming that the R7A* QS and ICE*Ml*Sym^{R7A} excision phenotypes are TraR dependent.

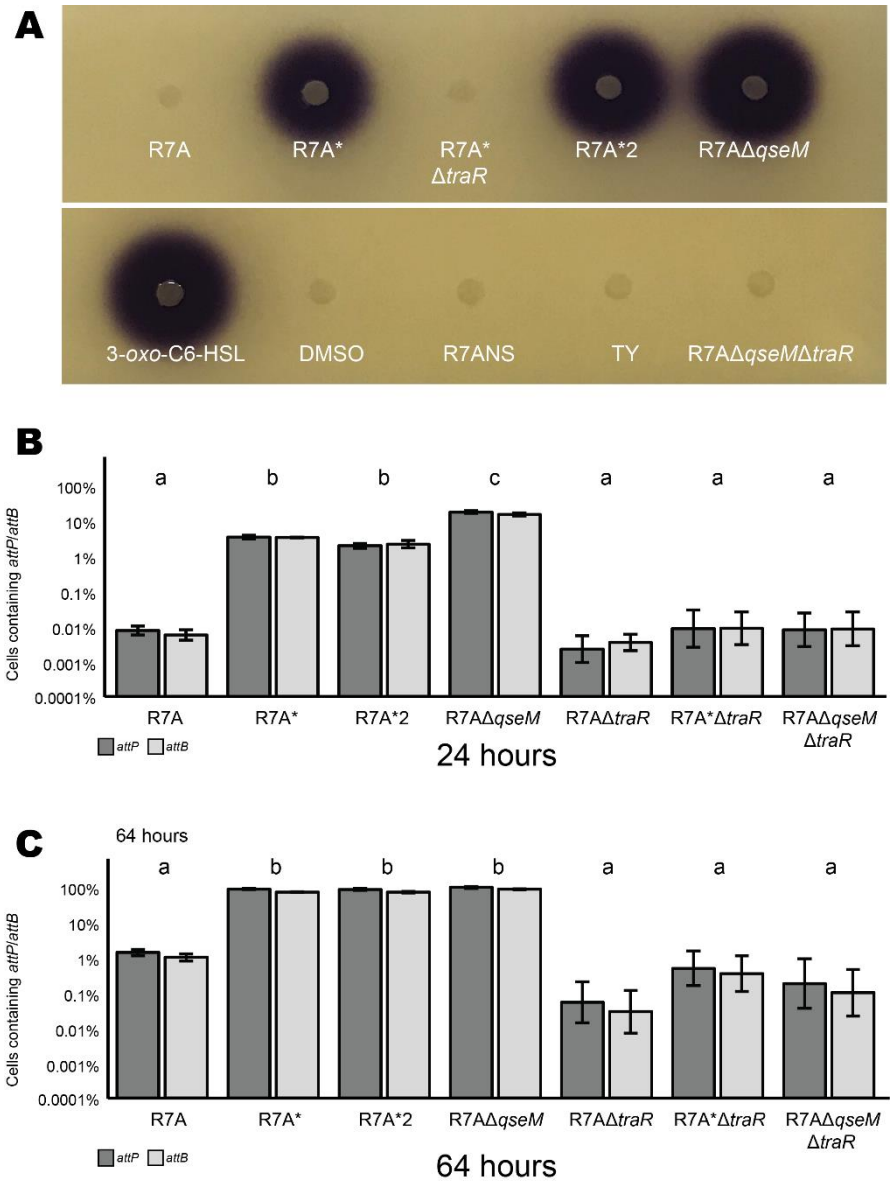


Figure 6.3 AHL production and ICE*Ml*Sym^{R7A} excision of R7A and R7A* strains (A) *C. violaceum* CV026 was used as a biosensor to detect TraI1-derived AHLs. R7A and R7A* strains were grown to stationary-phase in buffered TY and supernatants added to CV026 bioassays, to confirm that R7A* did indeed constitutively produce AHLs compared to R7A which produces no detectable AHLs. (B) qPCR was used measure ICE*Ml*Sym^{R7A} excision in R7A and R7A* strains. Bars represent the average percentage and error bars represent the standard deviation of three biological replicates. A one-way ANOVA and Tukey post-hoc test was used to compare each strain. Values that are significantly different from each other are grouped by different letters (a, b).

6.2.3 Overexpression of antiactivator QseM suppresses the R7A* phenotype

RNA sequencing of R7A* cells has previously shown that R7A* exhibits reduced *qseM* transcription compared to R7A (87). R7A* and R7AΔ*qseM* also share similarities in their AHL production and ICE*Ml*Sym^{R7A} excision phenotypes. We hypothesised that the phenotypes of R7A* were a direct result of the repression of *qseM* transcription, and that by overexpressing *qseM* in R7A* we may repress these phenotypes. A plasmid harbouring *qseM* under control of the constitutive promoter *nptII* – pNqseM – was introduced into R7A*, R7A*2 and R7AΔ*qseM*. Supernatants were harvested from stationary-phase cultures and added to CV026 bioassays, which confirmed that *qseM* overexpression abolished AHL production (Figure 6.4A).

DNA extracted from the above strains was assayed by qPCR, which confirmed *qseM* overexpression also repressed ICE*Ml*Sym^{R7A} excision in each strain. Overexpression of *qseM* in R7A reduced excision from approximately 1.4% to 0.02% during stationary-phase. Excision rates of ICE*Ml*Sym^{R7A} were approximately 81% in R7AΔ*qseM* harbouring the empty vector and overexpression of *qseM* in this strain reduced excision to 0.4%. R7A* and R7A*2 carrying the empty vector exhibited an average of 79% and 73% ICE*Ml*Sym^{R7A} excision respectively. Overexpression of *qseM* in these strains did indeed reduce ICE*Ml*Sym^{R7A} excision rates to below that of R7A at 0.05% and 0.07% respectively (Figure 6.4).

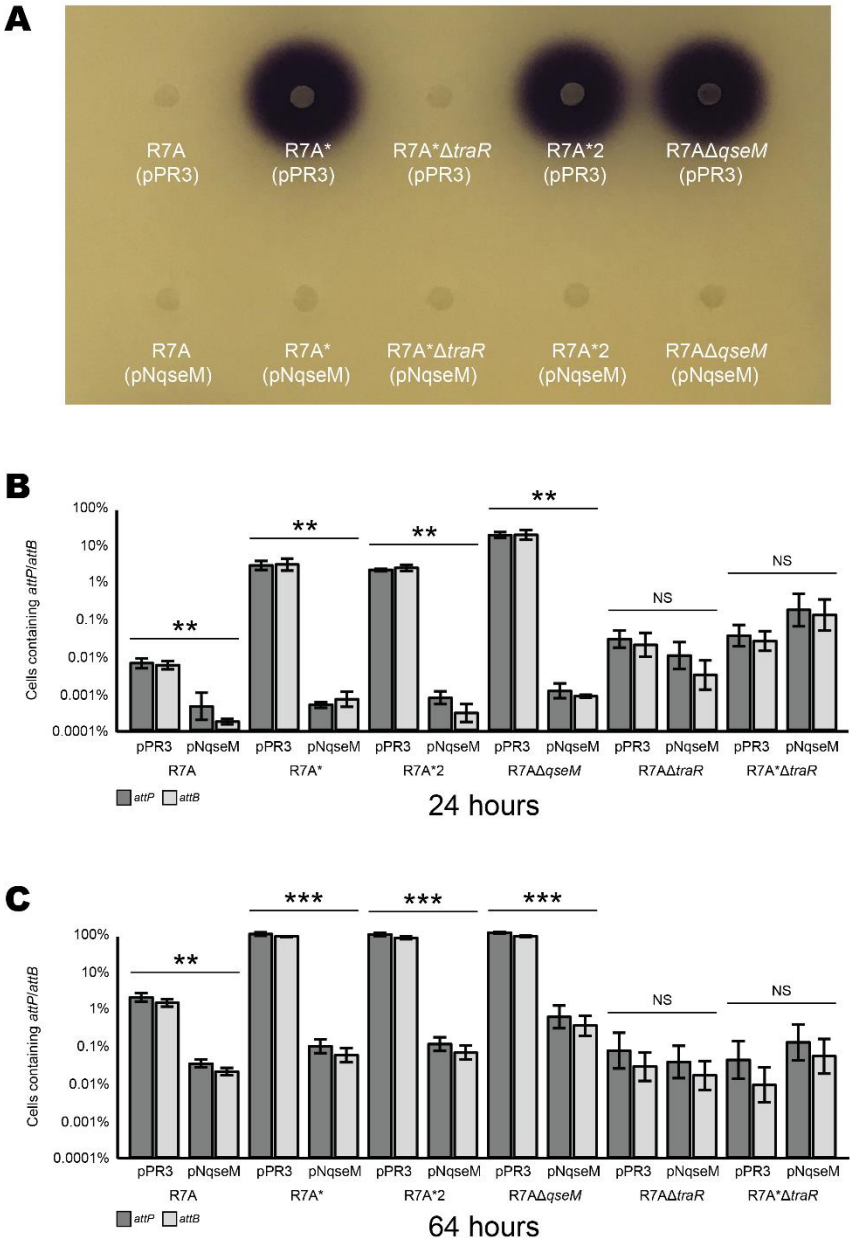


Figure 6.4 Analysis of AHL production and ICE*Ml*Sym^{R7A} excision in strains overexpressing *qseM*. (A) R7A-derived and R7A*-derived strains overexpressing *qseM* (and empty vector controls) were grown to stationary-phase in buffered TY and supernatants added to CV026 bioassays. (B-C) qPCR was used to measure ICE*Ml*Sym^{R7A} excision in R7A-derived and R7A*-derived strains overexpressing *qseM* . Samples were taken during log-phase at 24 hours, and at stationary-phase at 64 hours. Bars represent the average percentage and error bars represent the standard deviation of three biological replicates. Student's t tests were used to compare overexpression of *qseM* to empty vector controls for each background strain. ** denotes $p<0.01$, *** denotes $p<0.001$.

6.2.4 R7A* transfers ICE*Ml*Sym^{R7A} at a higher frequency than R7A and the R7A* phenotype is reset after transfer

While R7A* exhibited increased AHL production and ICE*Ml*Sym^{R7A} excision rates, we had yet to determine whether these phenotypes led to increased ICE*Ml*Sym^{R7A} horizontal transfer. R7ANS is a derivative of R7A lacking ICE*Ml*Sym^{R7A}, previously created by overexpressing *rdfS* in R7A (85). R7A* was cured of ICE*Ml*Sym^{R7A} in the same way, to create R7ANS* (87). In this study R7A, R7A*, R7AΔ*traR*, R7A*Δ*traR* and R7AΔ*qseM* were used as donors in ICE transfer experiments into R7ANS and R7ANS* recipients harbouring pFAJ1700 as a tetracycline resistant counterselectable marker. Each transfer experiment was carried out three times from three different starting cultures and the number of exconjugants was normalised to the number of donors in each experiment.

Conjugation frequencies in all experiments were almost identical between the two recipients R7ANS and R7ANS*. When ICE*Ml*Sym^{R7A} was transferred from R7A to either R7ANS or R7ANS*, frequencies were measured at an average of 1.2 x 10^{-7} exconjugants per donor. Conjugation frequencies were increased to 1.2 - 1.4 x 10^{-4} when R7AΔ*qseM* was used as a donor. R7A* and R7A*2 donors exhibited conjugation frequencies between 3.6 – 6.4 x 10^{-5}, about 400-fold higher than that of wild-type R7A and about 2.5-fold lower than R7AΔ*qseM* donors. When R7A and R7A* strains mutated for *traR* were used as donors, conjugation frequencies were reduced to below wild-type levels. Interestingly, the R7A*-derived *traR* mutant exhibited conjugation frequencies about 10-fold higher than R7AΔ*traR* suggesting that R7A* is able to maintain some level of transfer upregulation even in the absence of TraR.

Table 6.1 Conjugation frequencies of R7A* strains

Donor	Recipient†	Conjugation frequency	Standard deviation
R7A	R7ANS	1.2 x 10^{-7}	1.0 x 10^{-8}
R7A	R7ANS*	1.2 x 10^{-7}	3.8 x 10^{-8}
R7AΔ*qseM*	R7ANS	1.2 x 10^{-4}	1.1 x 10^{-5}

R7AΔ*qseM*	R7ANS*	1.4×10^{-4}	2.8×10^{-8}
R7A*	R7ANS	3.6×10^{-5}	1.4×10^{-5}
R7A*	R7ANS*	5.1×10^{-5}	1.0×10^{-6}
R7A*2	R7ANS	5.2×10^{-5}	2.5×10^{-6}
R7A*2	R7ANS*	6.4×10^{-5}	3.0×10^{-6}
R7AΔ*traR*	R7ANS	3.7×10^{-8}	1.6×10^{-9}
R7AΔ*traR*	R7ANS*	3.4×10^{-8}	2.8×10^{-9}
R7A*Δ*traR*	R7ANS	2.9×10^{-7}	1.0×10^{-8}
R7A*Δ*traR*	R7ANS*	2.8×10^{-7}	1.3×10^{-8}

†All recipients harboured pFAJ1700 for tetracycline selection

To determine whether the R7A* phenotypes were maintained after ICE*Ml*Sym^{R7A} transfer, eight exconjugants were taken from the R7A, R7A* and R7AΔ*qseM* matings and tested for AHL production using CV026 bioassays. All exconjugants from the R7AΔ*qseM* matings were positive for AHL production as expected. None of the exconjugants where R7A and R7A* were donors exhibited AHL production (Figure 6.5), indicating that the QS-active R7A* state is not transferred with ICE*Ml*Sym^{R7A} and is essentially reset in ICE*Ml*Sym^{R7A} exconjugants. This provides further evidence that the R7A* state is not a result of genetic mutations. When ICE*Ml*Sym^{R7A} was transferred from R7A* to R7ANS*, the exconjugants were essentially a reconstruction of the R7A* genome and yet these exconjugants exhibited no AHL production (Figure 6.5).

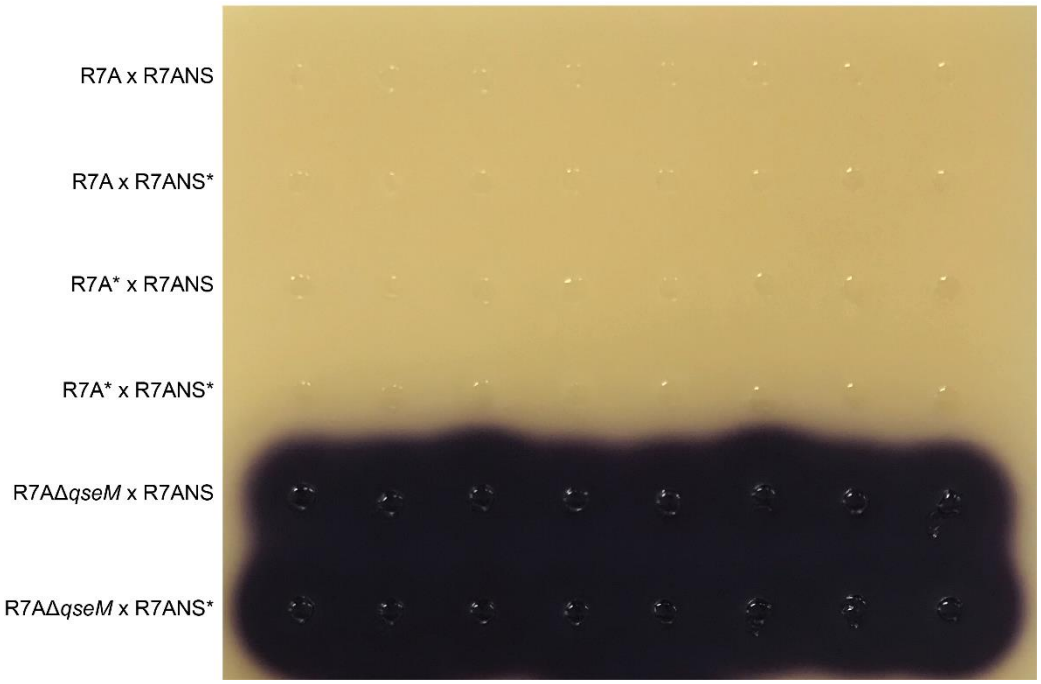


Figure 6.5 The R7A* AHL phenotype is reset following ICE*Ml*Sym^{R7A} transfer. Following conjugation experiments, eight exconjugants from each mating was grown to stationary-phase in buffered TY and supernatants added to CV026 bioassays. The transfer of ICE*Ml*Sym^{R7A} from R7A* to a non-symbiotic isogenic recipient did not carry over the AHL-producing phenotype, despite the fact that in doing so we essentially recreated the same genetically identical strain, confirming that this phenotype is epigenetically regulated. Transfer of ICE*Ml*Sym^{R7A} mutated for *qseM* created AHL-producing exconjugants as expected.

6.2.5 R7A* and R7A exhibit similar growth rates

R7A* exhibits constitutive AHL production and ICE*Ml*Sym^{R7A} excision in a transfer-ready state. We hypothesised that these phenotypes could conceivably incur a significant fitness cost for R7A* compared to wild-type R7A. R7A, R7AΔ*qseM* and R7A* were grown in the rich medium TY and the minimal medium G/RDM and samples taken over time were plated on non-selective media to determine colony forming units (CFU) per mL at given time points. The final cultures at 72 hours were tested for AHL production to confirm that R7A had not become R7A* and that R7A* had not reverted to R7A. R7A* grew almost identically to R7A and R7AΔ*qseM* in both rich and minimal media (Figure 6.6). During exponential phase R7A, R7A* and R7AΔ*qseM* exhibited a doubling time of 4.76, 4.78 and 4.85 hours respectively in TY medium, and

4.19, 4.34 and 4.28 hours respectively in G/RDM medium. These data suggest that the R7A* state does not impact growth rate.

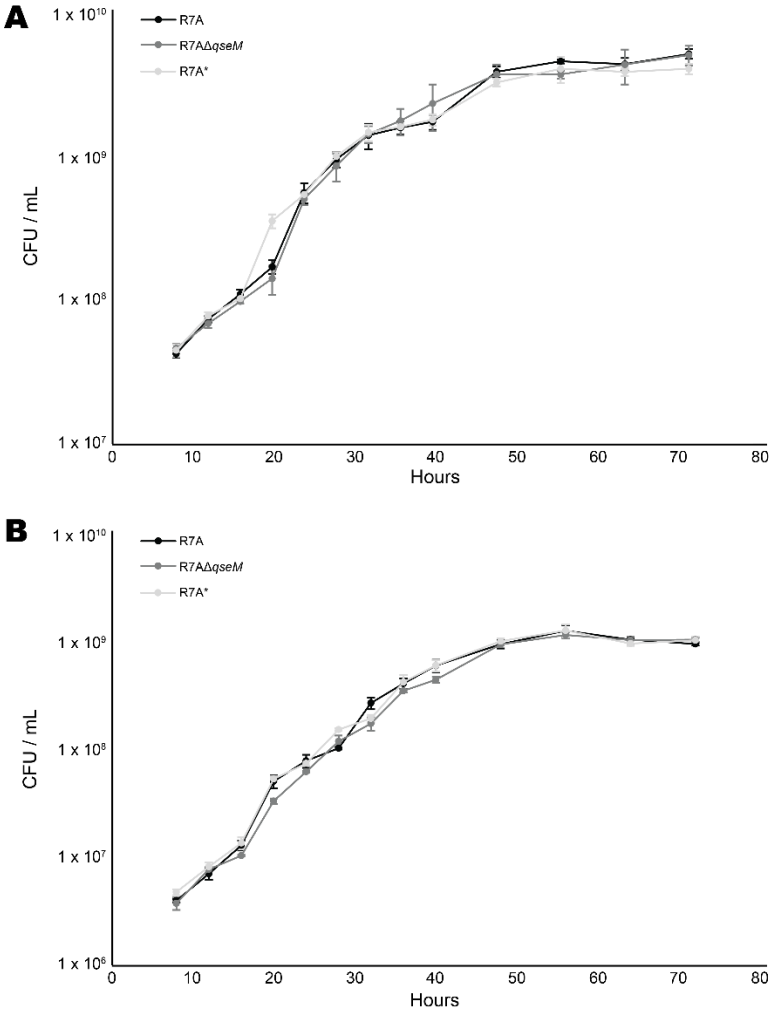


Figure 6.6 Growth of R7A, R7A* and R7AΔ*qseM* in rich and minimal media. R7A, R7A* and R7AΔ*qseM* were grown in (A) TY medium and (B) G/RDM medium for 72 hours. CFU/mL was determined at each time point by serially diluting each culture and plating on GRDM for colony counts. Graphed data represent the mean of three biological replicates and error bars depict standard deviation from the mean.

While R7A and R7A* grew at nearly identical rates in TY and G/RDM media, it seemed possible that there may be differences in carbon-source preference or utilisation. Phenotypic microarrays were carried out using two commercial 96-well microplates containing various carbon sources in each well called PM1

and PM2A (see methods). Seeder cultures were grown to stationary phase in GRDM medium, diluted to an OD_{600} of 0.05 and washed four times in RDM lacking a carbon source. Cells were resuspended in RDM containing a commercial redox dye (see methods), aliquoted into the microplate wells and incubated. As the cultures utilise each carbon source the redox dye produces a purple colour which can be measured at 540 nm absorbance along with measuring cell density at OD_{600}. PM1 and PM2A plates were grown until stationary phase and photos of the purple redox reactions taken at 72 hours (Figure 6.7A-B). PM1 plates were also grown over 72 hours with absorbance readings taken every eight hours throughout growth (Figure 6.7C). While the cell density growth curves (Figure 6.7C) suggested that R7A may be slower growing compared to R7A* and R7AΔ*qseM* in some carbon conditions, this was not a repeatable result. Ultimately, no consistent differences were observed in carbon metabolism between R7A, R7A* and R7AΔ*qseM*.

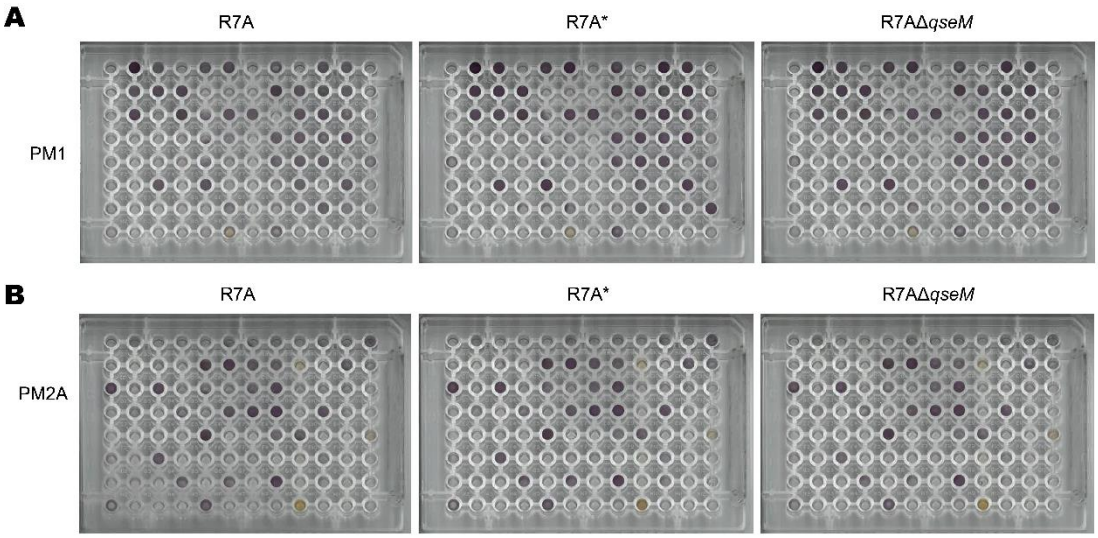
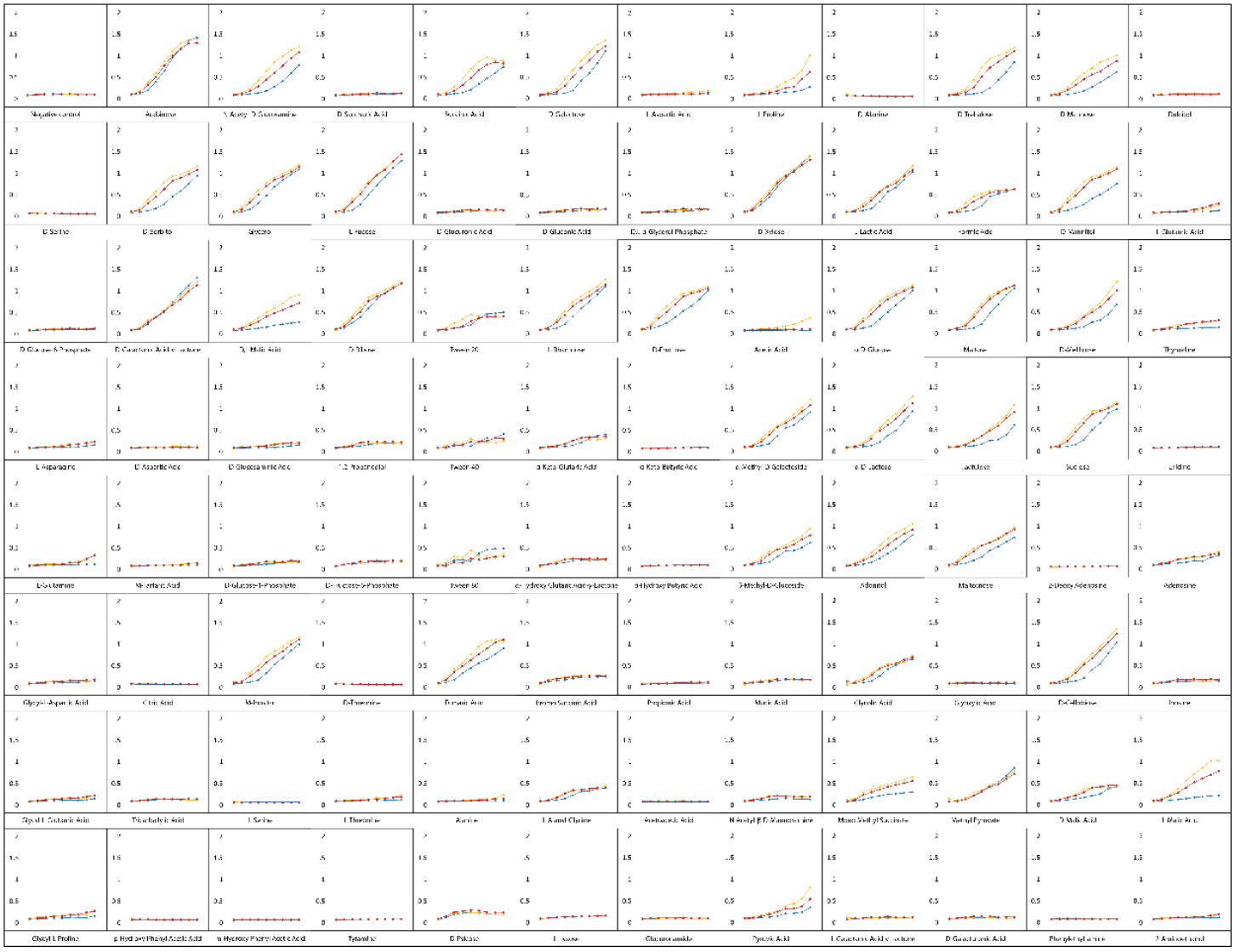


Figure 6.7 R7A, R7A* and R7AΔ*qseM* grown in phenotypic microarray plates. R7A, R7A* and R7AΔ*qseM* were grown to stationary-phase in liquid TY medium and cells were washed four times in RDM medium (no glucose). RDM supplemented with Biolog redox dye (as per manufacturers instructions) was used to resuspend cell pellets. Each cell suspension was normalised to OD_{600} 0.05 and used to inoculate PM1 and PM2A carbon source microarray plates. Plates were incubated at 28°C with a small beaker of sterile water to prevent drying out. The plates were retrieved from incubation every eight hours and measurements taken at OD_{600} (cell density) and 540 nm (redox dye absorption). Carbon metabolism was visibly observed by the purple colour of the redox dye. (A) PM1 and (B) PM2A plates at stationary phase are shown here. (C, below) Growth curves for each carbon source of the PM1 plate for R7A (blue), R7A* (yellow) and R7AΔ*qseM* (red).

c



6.2.6 R7A* reversion to an unstable R7A-like intermediary state

Throughout our work we discovered that R7A cultures could switch to R7A* at a relatively high frequency of ~2%, and that once a variant became R7A* its state was stably maintained. To determine whether it was possible for R7A* to revert back to R7A, we attempted to isolate colonies derived from R7A* populations that produced less AHLs than R7A*. Twenty-six individual colonies derived from R7A* and R7A*2 were used to inoculate buffered TY broth cultures and then supernatants from these cultures were added to a CV026 bioassay. Interestingly, the 26 supernatants exhibited varying levels of CV026 induction and two very low levels of AHLs (Figure 6.8A). One culture that exhibited almost no AHLs production (labelled as ¥ in Figure 6.8A) was again single colony purified on TY agar and 17 colonies were picked, grown to stationary-phase in buffered TY broth cultures and supernatants added to a second CV026 bioassay, this time exhibiting no AHL production. This lineage was named R7A*RV. We carried out ICE*Ml*Sym^{R7A} conjugation assays and measured ICE*Ml*Sym^{R7A} excision rates to determine whether R7A*RV had phenotypically reverted to an R7A-like state. The excision frequency of ICE*Ml*Sym^{R7A} in R7A*RV was approximately 7-9% during stationary-phase compared to ~1% in R7A and 72% in R7A* (Figure 6.8B). While excision was largely reduced in R7A*RV compared to R7A* it was still around eight times higher than R7A. Transfer of ICE*Ml*Sym^{R7A} from R7A*RV to R7ANS and R7ANS* was also approximately eight times higher than that of R7A. While AHL production, ICE*Ml*Sym^{R7A} excision and transfer were all significantly decreased compared to R7A*, R7A*RV was not phenotypically identical to R7A and instead resembled an intermediary state between R7A and R7A*. R7A*RV was frozen at -80°C for long term storage. However, upon reviving this strain from the freezer stock and testing for AHL production in a CV026 bioassay, it was discovered that R7A*RV AHL production was now at levels resembling the original R7A*, suggesting that this was not a true phenotypic revertant.

Table 6.2 Conjugation frequencies of R7A*RV

Donor	Recipient†	Conjugation frequency	Standard deviation
R7A	R7ANS	1.2×10^{-7}	1.0×10^{-8}
R7A	R7ANS*	1.2×10^{-7}	3.8×10^{-8}
R7A*	R7ANS	3.6×10^{-5}	1.4×10^{-5}
R7A*	R7ANS*	5.1×10^{-5}	1.0×10^{-6}
R7A*RV	R7ANS	9.9×10^{-7}	1.5×10^{-7}
R7A*RV	R7ANS*	9.0×10^{-7}	1.2×10^{-7}

†All recipients harboured pFAJ1700 for tetracycline selection

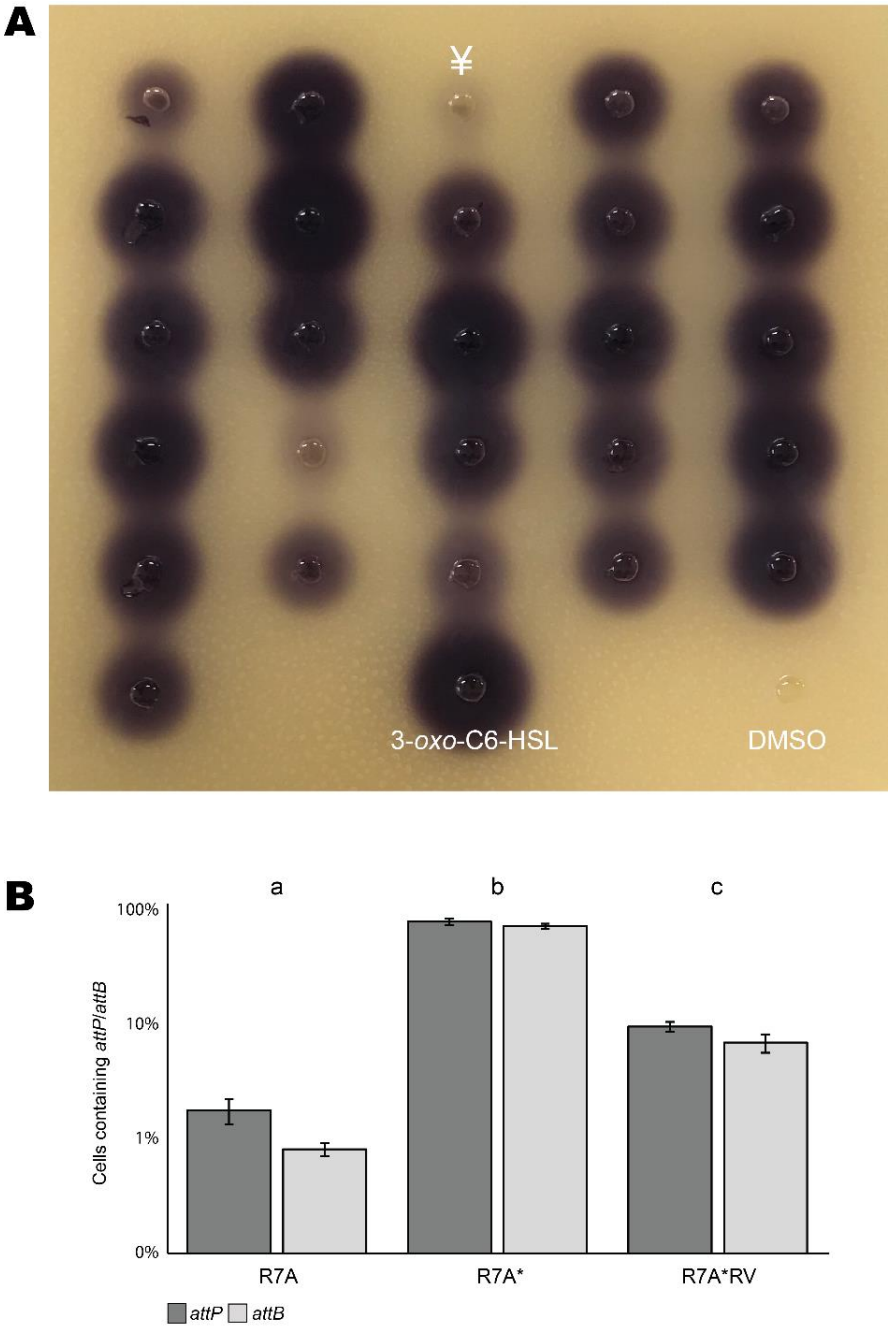


Figure 6.8 AHL and ICE*Ml*Sym^{R7A} excision phenotypes of R7A*RV (A) Single colony isolates of R7A* and R7A*2 were grown to stationary-phase in buffered TY medium and supernatants added to *C. violaceum* CV026 bioassays to detect C4-C8 AHL production. The isolate labelled ¥ was named R7A*RV and further analysed for ICE*Ml*Sym^{R7A} excision and transfer. (B) qPCR was used to measure ICE*Ml*Sym^{R7A} excision in R7A*RV compared to R7A and R7A*. Bars represent the average percentage and error bars represent the standard deviation of three biological replicates. A one-way ANOVA and Tukey post-hoc test was used to compare each strain. Values that are significantly different from each other are grouped by different letters (a, b, c).

6.2.7 QseC2 and P*asqseC* are involved in the establishment and maintenance of R7A*

R7A and R7A* exhibit changes in transcription of the P*asqseC-qseC2* region (87), implicating transcription in this region in the control of the R7A* phenotypes. In this study, *qseC2* deletion mutants in the R7A and R7A* backgrounds (Figure 6.9) were analysed for quorum sensing and ICE*Ml*Sym^{R7A} excision and transfer phenotypes. Deletion of *qseC2* in R7A had no effect on AHL production or ICE*Ml*Sym^{R7A} excision (Figure 6.10A,B), however, ICE*Ml*Sym^{R7A} transfer was reduced slightly below wild-type levels (Table 6.3). When *qseC2* was deleted in the R7A* background, AHL production was abolished and ICE*Ml*Sym^{R7A} excision and transfer rates were reduced back to R7A levels, suggesting that QseC2 is essential for R7A* phenotype maintenance. Expression of *qseC2* from the strong constitutive *nptII* promoter on the plasmid pNqseC2 restored AHL production and induced ICE*Ml*Sym^{R7A} excision (Figure 6.10A,C). Overexpression of *qseC2* in R7A and R7AΔ*qseC2* also induced AHL production and ICE*Ml*Sym^{R7A} excision in these strains to levels resembling those of R7A* (Figure 6.10A,D).

We next aimed to investigate the role *asqseC* has in regulating R7A* phenotypes. Deletions were constructed in R7A and R7A* that removed the region containing P*asqseC* and the *qseC2* gene (ΔP*asqseC*Δ*qseC2*) (Figure 6.9). This deletion in R7A resulted in activation of AHL production and increased ICE*Ml*Sym^{R7A} excision and transfer to levels similar to and above those of R7A* (Figure 6.10A,B and Table 6.3). These data suggested the P*asqseC* promoter was required to prevent R7A from becoming R7A*.

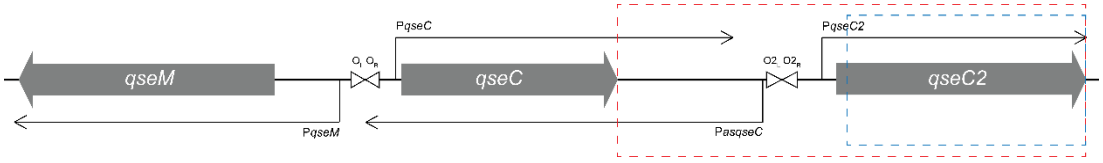


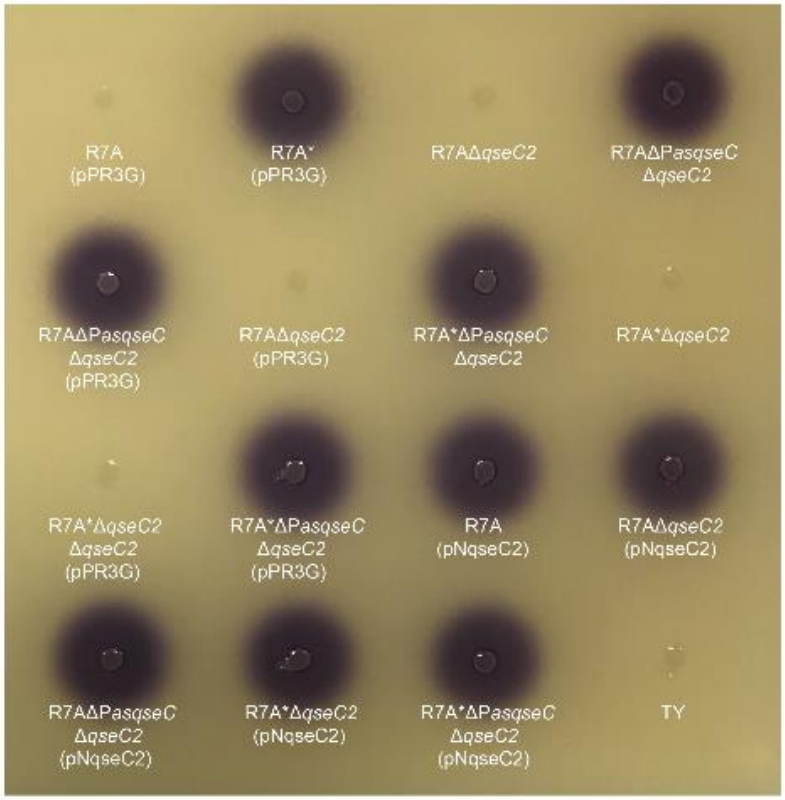
Figure 6.9 Gene map of the *qseM-qseC2* region highlighting the regions deleted in the Δ*qseC2* and ΔP*asqseC*Δ*qseC2* mutants. The Δ*qseC2* deletion is depicted by the blue dashed box and the ΔP*asqseC*Δ*qseC2* deletion is depicted by the red dashed box.

Table 6.3 Conjugation frequencies of Δ*qseC2* and ΔP*asqseC*Δ*qseC2* strains

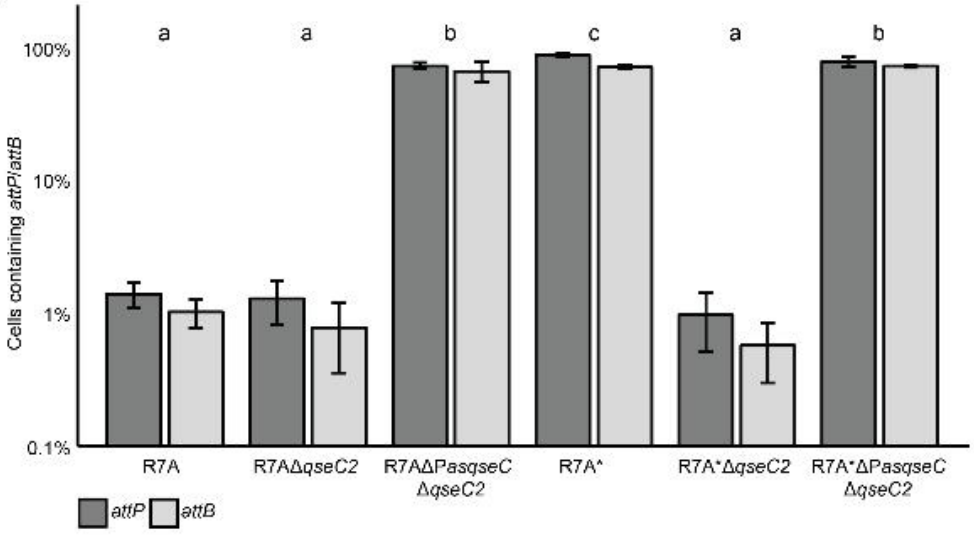
Donor	Recipient†	Conjugation frequency	Standard deviation
R7A	R7ANS	1.2×10^{-7}	1.0×10^{-8}
R7A	R7ANS*	1.2×10^{-7}	3.8×10^{-8}
R7A*	R7ANS	3.6×10^{-5}	1.4×10^{-5}
R7A*	R7ANS*	5.1×10^{-5}	1.0×10^{-6}
R7AΔ*qseC2*	R7ANS	2.0×10^{-8}	1.1×10^{-9}
R7AΔ*qseC2*	R7ANS*	3.5×10^{-8}	1.6×10^{-9}
R7A*Δ*qseC2*	R7ANS	4.2×10^{-8}	1.0×10^{-9}
R7A*Δ*qseC2*	R7ANS*	2.2×10^{-8}	3.9×10^{-9}
R7AΔP*asqseC*Δ*qseC2*	R7ANS	1.0×10^{-4}	8.8×10^{-6}
R7AΔP*asqseC*Δ*qseC2*	R7ANS*	1.0×10^{-4}	9.3×10^{-6}
R7A*ΔP*asqseC*Δ*qseC2*	R7ANS	9.7×10^{-5}	6.6×10^{-6}
R7A*ΔP*asqseC*Δ*qseC2*	R7ANS*	6.4×10^{-5}	1.1×10^{-5}

†All recipients harboured pFAJ1700 for tetracycline selection

A


R7A (pPR3G)
R7A* (pPR3G)
R7AΔqseC2
R7AΔPasqseC ΔqseC2
R7AΔPasqseC ΔqseC2 (pPR3G)
R7AΔqseC2 (pPR3G)
R7A*ΔPasqseC ΔqseC2
R7A*ΔqseC2
R7A*ΔqseC2 ΔqseC2 (pPR3G)
R7A*ΔPasqseC ΔqseC2 (pPR3G)
R7A (pNqseC2)
R7AΔqseC2 (pNqseC2)
R7AΔPasqseC ΔqseC2 (pNqseC2)
R7A*ΔqseC2 (pNqseC2)
R7A*ΔPasqseC ΔqseC2 (pNqseC2)
TY

B


a
a
b
c
a
b
100%
10%
1%
0.1%
Cells containing attP/attB
R7A
R7AΔqseC2
R7AΔPasqseC ΔqseC2
R7A*
R7A*ΔqseC2
R7A*ΔPasqseC ΔqseC2
attP
attB

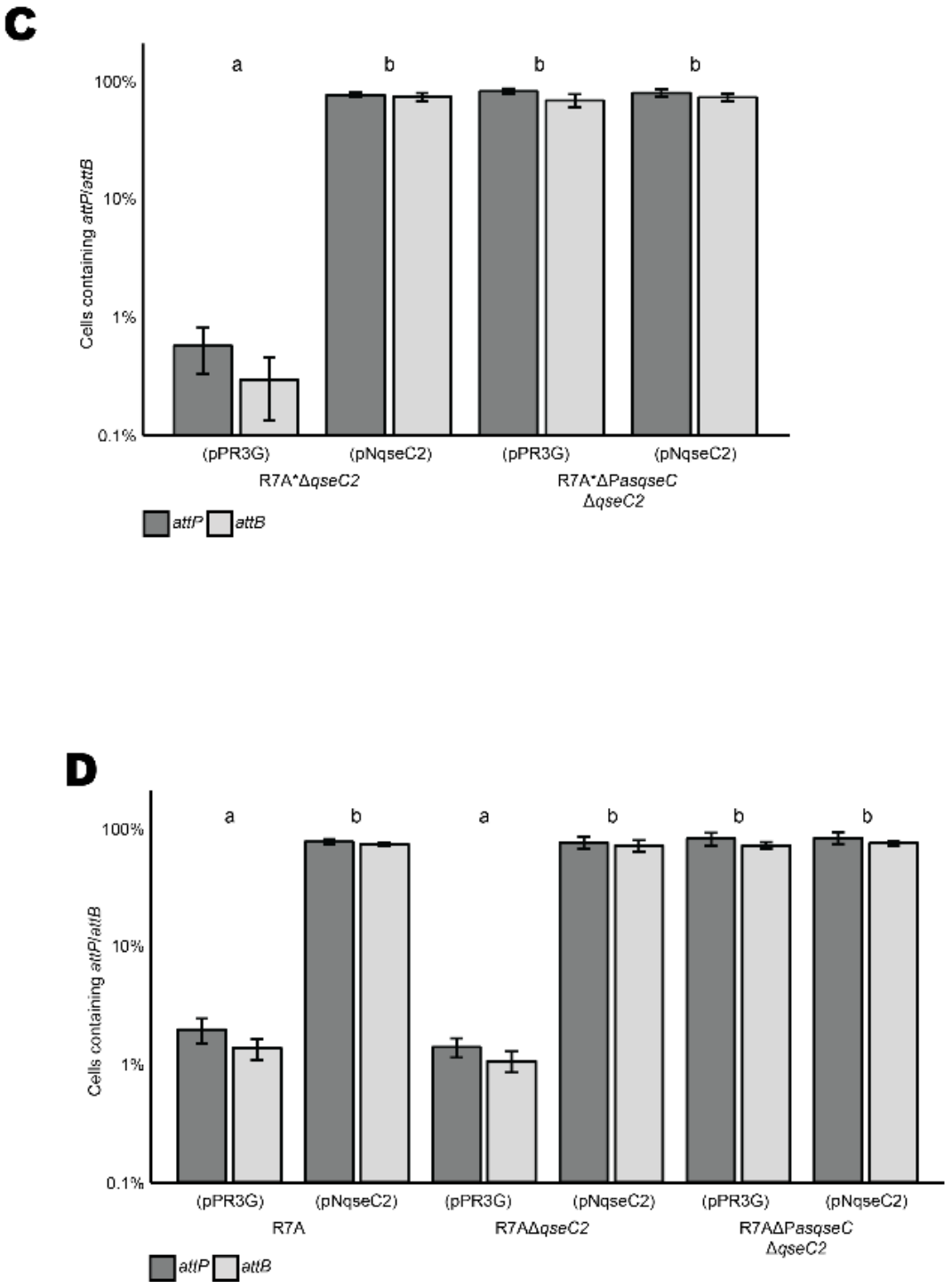


Figure 6.10 AHL production and ICE*Ml*Sym^{R7A} excision in P*asqseC* and *qseC2* mutants. (A) Strains mutated for or overexpressing P*asqseC* or *qseC2* were grown to stationary-phase in buffered TY medium and supernatants added to *C. violaceum* CV026 bioassays to detect C4-C8 AHL production. (B-D) qPCR was used to measure ICE*Ml*Sym^{R7A} excision in strains mutated for or overexpressing P*asqseC* or qseC2. Bars represent the average percentage during stationary-phase (64 hours) and error bars represent the standard deviation of three biological replicates. A one-way ANOVA and Tukey-post hoc test was used to compare each strain. Values that are not statistically different from each other are grouped by the same letter (a, b, c).

We wondered if the antisense *qseC* transcript was able to inhibit *qseC*/QseC transcription or translation when expressed *in trans* on a plasmid. The 355 bp region from the predicted P*asqseC* transcriptional start site through to the *qseC* start codon was cloned into pPR3G downstream of the constitutive promoter *nptII*. We hypothesised that if the antisense-*qseC* transcript itself was able to directly repress expression of *qseC*, then this plasmid (pPR3G-antiqseC) would repress AHL production in R7A*. However, pPR3G-antiqseC had no effect in R7A* and cells continued to produce AHLs.

Next, we wondered if we could recapitulate the functionality of the *asqseC* promoter by replacing it with the strong constitutive *nptII* promoter *in cis*. Two constructs were designed: one positioned P*nptII* anti-sense to *qseC* and the second positioned P*nptII* in the opposite direction, i.e. downstream of and in the same orientation as *qseC*. To construct allelic replacement vectors, regions flanking P*asqseC-qseC2* and P*nptII* were cloned in a three-way ligation with the suicide vector pJQ200SK. These plasmids were then used to replace P*asqseC* with P*nptII* in each orientation in the chromosomes of R7A and R7A* via allelic replacement (see methods). In the R7A* strain carrying P*nptII* positioned to drive anti-*qseC* expression, AHL production was abolished, confirming that antisense *qseC* transcription *in cis* represses the R7A* state. In contrast, R7A and R7A* carrying P*nptII* in the same orientation as *qseC* were active for AHL production (Figure 6.11). Since we were able to induce AHL production by overexpressing *qseC2* in R7A (described above) the pNqseC2 plasmid was introduced into the mutant strains harbouring P*nptII*, to test if the action of QseC2 was dependent on the P*asqseC* promoter being present. AHL production was not restored by *qseC2* overexpression in the strain carrying P*nptII* driving anti-*qseC* expression (Figure 6.11), suggesting that while QseC2 is able to induce R7A*, a QseC2 binding site within the P*asqseC* region is likely needed for QseC2-mediated induction of AHL production, further suggesting QseC2 represses P*asqseC* directly.

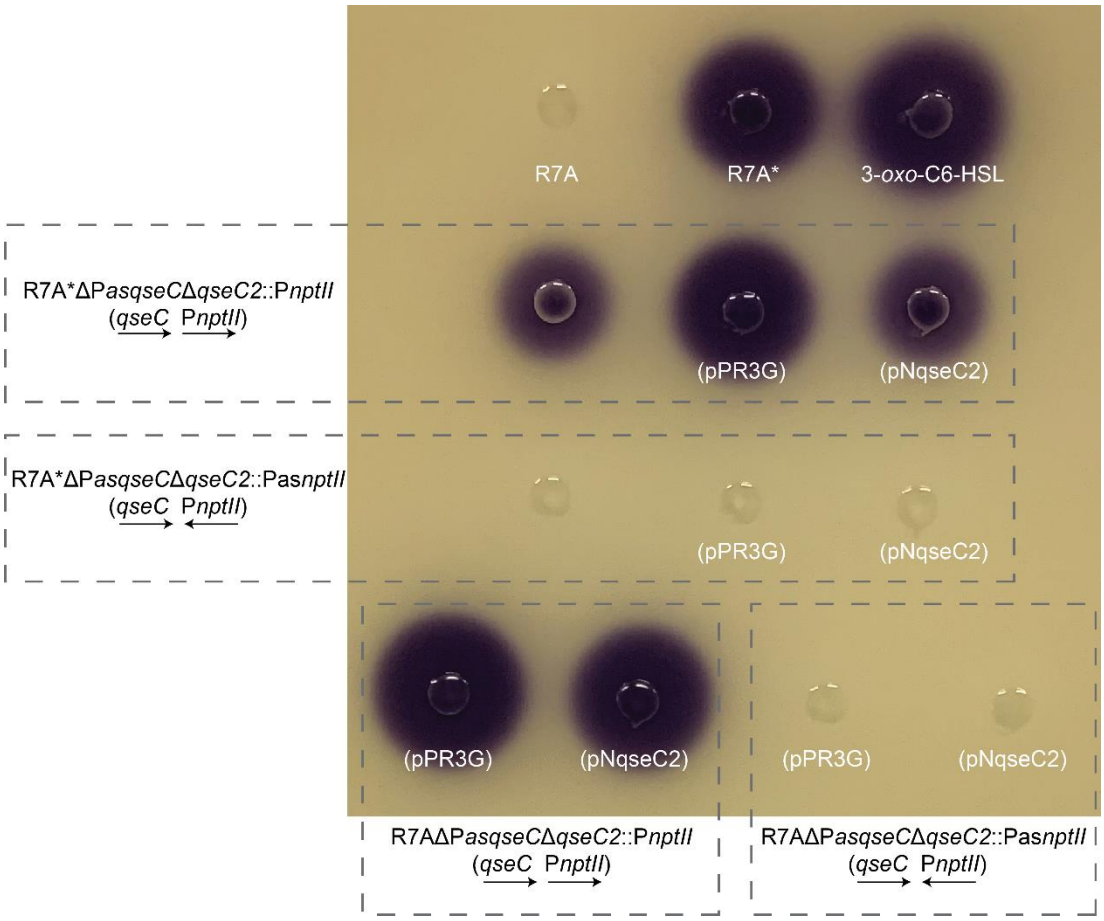


Figure 6.11 CV026 assay of strains harbouring P*nptII in cis* driving anti-*qseC* transcription. R7A and R7A* were mutated for P*asqseC-qseC2* and the region replaced with P*nptII* either driving expression of anti-*qseC* or away from *qseC*. CV026 assays showed that when P*nptII* was orientated towards anti-*qseC* AHL production was abolished. Grey dashed boxes group together samples of the same background strain which is described to the left or bottom of the bioassay plate.

6.2.8 QseC2 regulates two overlapping promoters via binding of $O2_L$ and $O2_R$ in a concentration-dependent manner

The region between *qseC* and *qseC2* includes a pair of inverted repeat sequences $O2_L$ and $O2_R$ and DNA-binding assays have shown that these operator sites are the binding sites for QseC2. Electrophoretic mobility shift assays and surface plasmon resonance have demonstrated QseC2 binds both operator sites with a slight preference for $O2_L$ (87, 202). We hypothesised that QseC2 binding these operator sites allowed QseC2-dependent regulation of the *qseC2* and *asqseC* promoters. To test this, the pSDz plasmid was used, which contains a cloning cassette positioned upstream of a promoterless *lacZ*

gene and a separate cloning cassette positioned downstream of the IPTG-inducible *lac* promoter. Previous studies have shown that the *lac* promoter on pSDz exhibits 'leaky' low-level expression in TY medium without IPTG and expression is further increased with IPTG induction (80). We utilised this to create strains with varying levels of *qseC2* expression and to measure potential QseC2 concentration-dependent regulation of the overlapping promoters P*asqseC* and P*qseC2*. A 214 bp region between *qseC* and *qseC2* (depicted in the grey dashed box in Figure 6.12A) was cloned in both orientations in the cloning cassette upstream of *lacZ* to create one plasmid to measure P*asqseC* activity (pSDz-PasqseC) and one to measure P*qseC2* activity (pSDz-qseC2). The qseC2 gene was then cloned into both plasmids downstream of the IPTG-inducible *lac* promoter. All reporter plasmids were introduced into R7ANS as this strain lacks ICE*Ml*Sym^{R7A}.

β-galactosidase assays were carried out in the P*qseC2* and P*asqseC* reporter strains, utilising the varying levels of *qseC2* expression with and without the addition of IPTG. High β-galactosidase activity was observed in the absence of QseC2. Uninduced/leaky expression of *qseC2* also had no effect on P*qseC2* expression. When *qseC2* expression was induced with IPTG, P*qseC2* expression was reduced to that of the empty vector control strain (Figure 6.12B), confirming QseC2 represses its own gene promoter. P*asqseC* was expressed at similar levels to P*qseC2* in the absence of QseC2. However, in the presence of leaky, uninduced *qseC2* expression, or in the presence of IPTG, P*asqseC* expression was reduced to levels similar to that of the empty vector control (Figure 6.12C). These data together suggested that there is a concentration-dependent regulation of the two overlapping promoters controlled by QseC2 binding in this region, where low levels of QseC2 repress P*asqseC* expression and higher concentrations of QseC2 repress P*qseC2*.

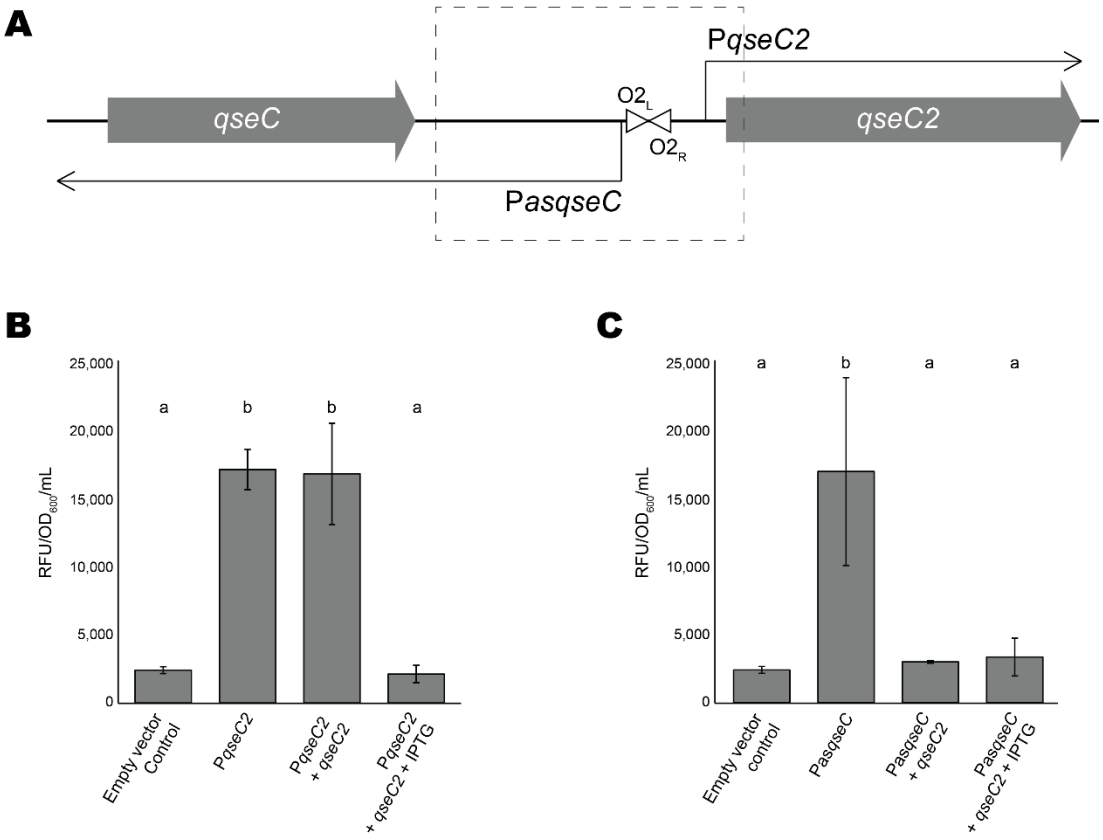


Figure 6.12 QseC2-dependent regulation of P*qseC2* and P*asqseC* (A) Map of the *qseC-qseC2* region (to scale). $O2_L$ and $O2_R$ are depicted by white triangles and the grey dashed box represents the region cloned into pSDz upstream of *lacZ*. Promoters are depicted by arrows and were previously estimated based on transcriptomics (87). (B,C) β-galactosidase assays measuring *lacZ* expression driven by P*qseC2* or P*asqseC* in the presence or absence of *qseC2* and with or without the addition of 1 mM IPTG. Bars represent the average of three biological replicates and error bars represent standard deviation from the mean. A one-way ANOVA and Tukey-post hoc test was used to compare each reporter strain. Values that are not statistically different from each other are grouped by the same letter (a, b).

6.2.9 QseC regulates two overlapping promoters in a concentration dependent manner

Previous work has shown that QseC binds two operator sites O_L and O_R located between *qseM* and *qseC* and that QseC first binds O_L and then cooperatively binds O_R at higher concentrations (81). The pSDz plasmid was again used to measure expression of the overlapping promoters P*qseM* and P*qseC* dependent on leaky or induced expression of *qseC*. A 141 bp region between *qseM* and *qseC* (depicted in the grey dashed box in Figure 6.13A) was cloned in both orientations in the cloning cassette upstream of *lacZ* in

pSDz to create two separate reporter assays to measure P*qseM* expression and P*qseC* expression. The *qseC* gene was cloned into both reporter constructs downstream of the IPTG-inducible *lac* promoter and all four reporter plasmids were introduced into R7ANS.

β-galactosidase assays revealed P*qseC* expression in the absence of QseC to be similar to that of the empty vector control strain. Low-level expression of *qseC* from the leaky *lac* promoter increased P*qseC* expression. When induced with IPTG, high level expression of *qseC* resulted in the reduction of P*qseC* expression to below those observed in the absence of the *qseC* gene (Figure 6.13B). Expression of P*qseM* in the absence of *qseC* was approximately six times higher than that of P*qseC* in the absence of *qseC*. However, even leaky expression of *qseC* without IPTG induction abolished expression from P*qseM* (Figure 6.13C). Taken together, these data suggest that in the absence of QseC, P*qseM* is highly expressed. A low concentration of QseC protein activates P*qseC* and simultaneously abolishes expression from P*qseM*. As the concentration of QseC increases and occupies both O_L and O_R sites, P*qseC* expression is repressed.

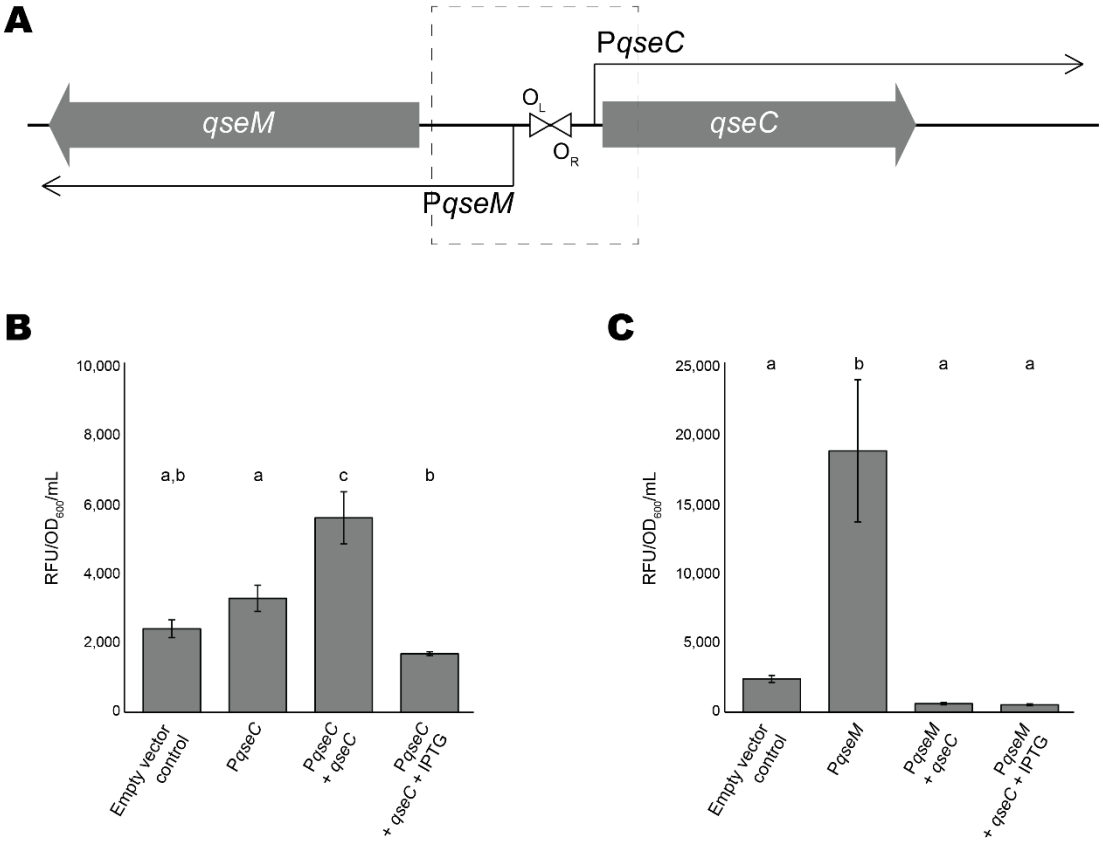


Figure 6.13 QseC-dependent regulation of P*qseC* and P*qseM* (A) Map of the *qseM-qseC* region (to scale). O_L and O_R are depicted by white triangles and the grey dashed box represents the region cloned into pSDz upstream of *lacZ*. Promoters are depicted by arrows and were previously mapped using 5'RACE (81). (B,C) β-galactosidase assays measuring *lacZ* expression driven by P*qseM* or P*qseC* in the presence or absence of *qseC* and with or without the addition of 1 mM IPTG. Bars represent the average of three biological replicates and error bars represent standard deviation from the mean. A one-way ANOVA and Tukey-post hoc test was used to compare each reporter strain. Values that are not statistically different from each other are grouped by the same letter (a, b, c).

6.2.10 Transient QseC2 expression triggers entry of cells into the R7A* state and both QseC and QseC2 are required to maintain the R7A* state

Results described above led to the hypothesis that the concentration of QseC2 and/or QseC may influence whether a cell enters and remains in the R7A* state. Moreover, the stable nature of R7A* cells suggest that once QseC2 concentrations achieve a certain threshold, cells maintain the QseC2 concentrations above this threshold thereafter. To test these hypotheses, we constructed a *qseC2* overexpression vector that was able to be selectively cured from cells, so that we could induce the R7A* state and test if the state

was maintained after the *qseC2* vector was removed. To create this vector, the *sacB* gene was cloned into pNqseC2 and into the empty vector pPR3G. The *sacB* gene encodes levansucrase which converts sucrose to the toxic levans, allowing for counterselection on media containing sucrose (204). Therefore having *sacB* would allow us to introduce the plasmid into a strain and subsequently cure it. The pNqseC2-sacB and pPR3G-sacB plasmids were introduced into R7A in five separate mating experiments. In all cases, the pNqseC2-sacB plasmid stimulated the AHL production while pPR3G-sacB did not. The cultures used in the CV026 assays were plated onto media containing sucrose (S/RDM) in order to select for strains that had lost the *sacB* carrying plasmids. Single colonies from each of the five independent matings were single colony purified once on S/RDM and then once on G/RDM (isolated colonies were also tested for gentamicin sensitivity to confirm loss of the plasmid). Single colonies from the G/RDM plates were then grown in buffered TY and their supernatants were tested on CV026 assays for AHL production. Each of the strains that previously harboured pNqseC2-sacB remained active for AHL production (Figure 6.14A). QPCR assays confirmed that these strains also exhibited ICE*Ml*Sym^{R7A} excision at the same frequency as R7A* (Figure 6.14B-C). Therefore transient overexpression of *qseC2* was able to induce cells to enter the R7A* state.

Transient overexpression of *qseC2* stimulated R7A to become R7A* and those cells remained in the R7A* state even after they were cured of the pNqseC2-sacB plasmid. We hypothesised that the chromosomal *qseC2* and/or *qseC* were likely responsible for the maintenance of R7A* phenotypes after they had been induced by QseC2. The pNqseC2-sacB and pPR3G-sacB plasmids were introduced into R7AΔ*qseC2* in five separate mating experiments. Strains harbouring the *qseC2* overexpression plasmid were activated for AHL production as expected, and strains that had received the empty vector did not produce AHLs. After R7AΔ*qseC2* was cured of the *qseC2* plasmid, these strains ceased producing AHLs (Figure 6.14A), confirming that *qseC2* was essential for maintainance of the R7A* state once induced.

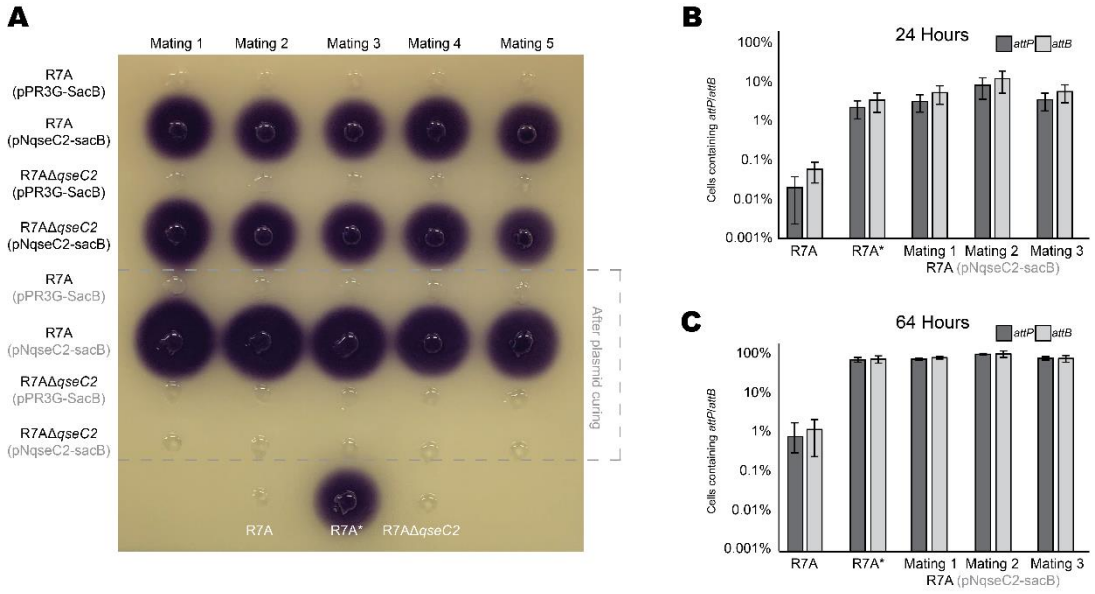


Figure 6.14 Analysis of AHL production and ICE*Ml*Sym^{R7A} excision following transient overexpression of *qseC2*. (A) R7A and R7AΔ*qseC2* either harbouring the curable *qseC2* overexpression plasmid or cured of the *qseC2* overexpression plasmid from five separate matings were grown to stationary phase in buffered TY medium and supernatants added to CV026 bioassays. Strains harbouring qseC2 as well as those that were cured of qseC2 exhibited AHL production, confirming that the QseC2 concentrations within the cell triggers and maintains the R7A* phenotypes. R7AΔ*qseC2* did not maintain AHL production after curing of the *qseC2* plasmid, indicating that the *in cis qseC2* gene is essential for maintaining AHL production. (B-C) qPCR was used to measure ICE*Ml*Sym^{R7A} excision, confirming that R7A strains previously overexpressing *qseC2* exhibited R7A*-like excision frequencies. Bars represent the average percentage and error bars represent the standard deviation of three biological replicates. (A, B, C) Grey text represents strains that have been cured of the plasmid.

While we had confirmed that QseC2 was essential induction of the R7A* state and its maintenance, it was still unclear as to whether QseC2-mediated activation of the R7A* state was due to repression of antisense *qseC* transcription and reduced production of QseC. It seemed plausible that P*asqseC* could be more directly controlling *qseM* expression and that QseC2 merely repressed such expression (Figure 6.15). We therefore designed experiments to test whether the translation of QseC was essential for the repression of *qseM* in cells overexpressing *qseC2*. For these experiments we had to disrupt QseC translation without disrupting transcription from P*asqseC*,

so instead we designed an allelic replacement vector to introduce a premature stop codon in the *qseC* gene (amber mutation), so as to disrupt QseC translation without disrupting any of the surrounding genes and promoters. Three base pairs were altered to change the tenth codon to a stop codon and disrupt the downstream ATG. The design also incorporated an XbaI restriction endonuclease site allowing for screening of potential mutants by PCR and restriction digest (Figure 6.16A). The mutation was introduced into R7A and R7A* by allelic replacement (see methods) creating R7A*qseC*(Am) and R7A**qseC*(Am). R7A*qseC*(Am) exhibited no AHL production as expected, however R7A**qseC*(Am) was also unable to produce AHLs. The pNqseC2 plasmid was then introduced to these mutants in five separate conjugation events and neither R7A*qseC*(Am) nor R7A**qseC*(Am) exhibited any AHL production (Figure 6.16B), confirming that a functional *qseC* gene is essential for entry into the R7A* state and that the control exerted by QseC2 depends on QseC.

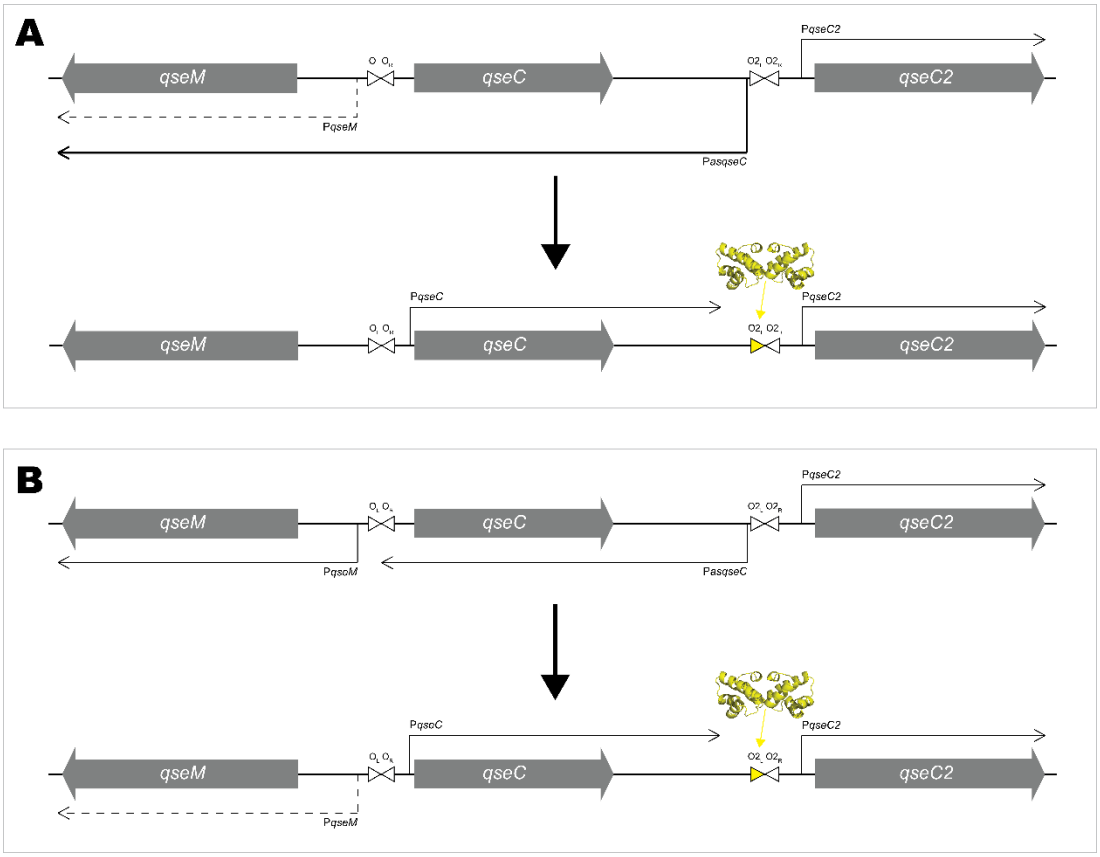


Figure 6.15 Two possible mechanisms of P*qseM* repression by QseC2 binding of O2L. QseC translation was disrupted to determine whether QseC2 binding of O2L (A) represses transcription from P*asqseC* and repression of read through transcription of *qseM*, independently of P*qseC* transcription. Or (B) represses transcription from P*asqseC*, upregulating P*qseC* and subsequent QseC-dependent repression of P*qseM*.

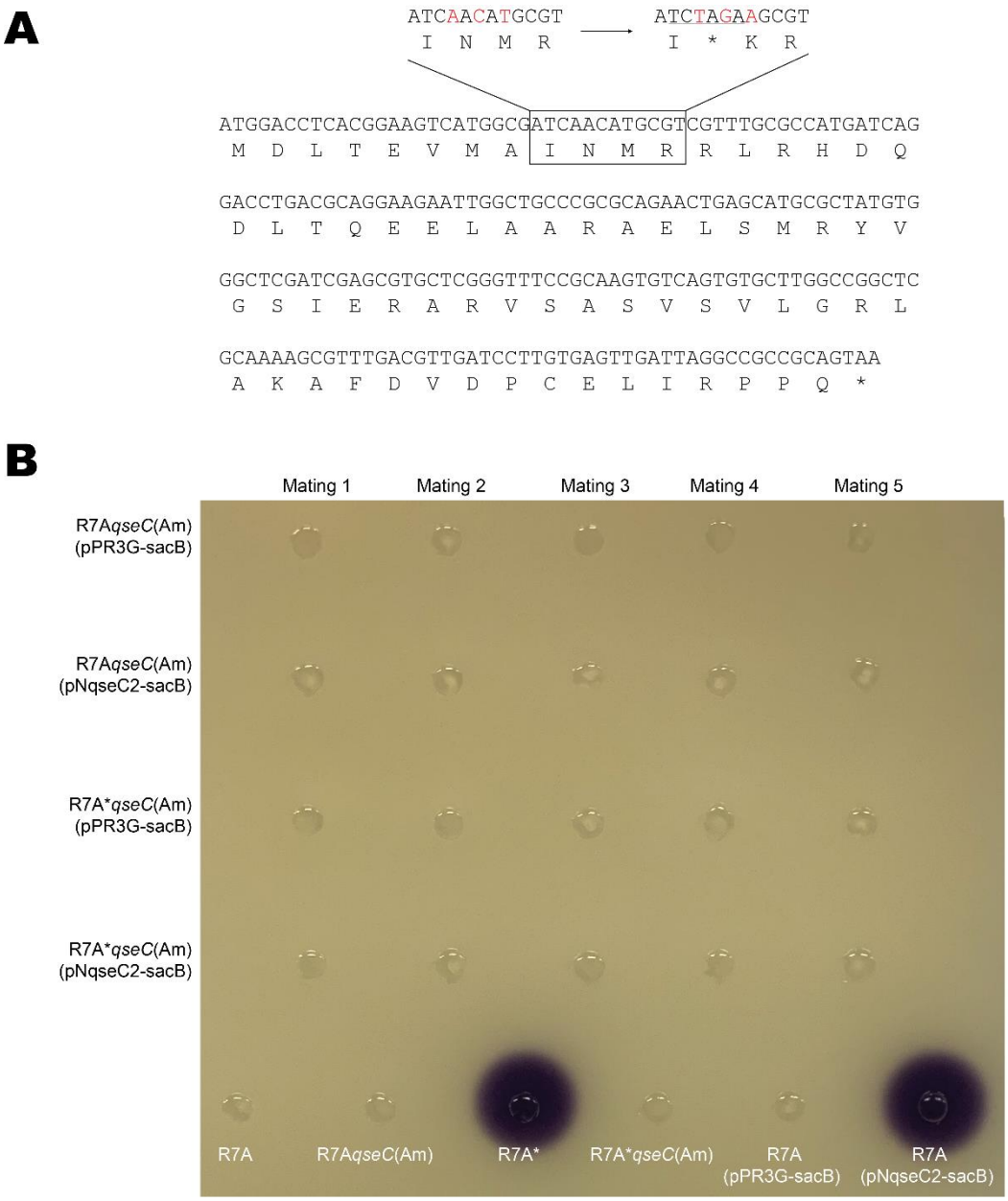


Figure 6.16 QseC translation is essential for R7A*. (A) A stop codon was introduced in the *qseC* coding sequence in the chromosome. Altered nucleotides are coloured red and the resulting residue changes are displayed below. An XbaI restriction site was also introduced to aid in mutant screening. (B) R7A and R7A* derived *qseC* mutants were grown to stationary-phase and supernatants added to CV026 bioassays. All strains with the mutated *qseC* gene lost the ability to produce AHLs, and overexpression of *qseC2* failed to trigger AHL production.

6.3 Discussion

In this chapter, we further characterise a phenotypic *M. japonicum* R7A variant termed R7A*. Quantitative PCR and *Chromobacterium violaceum* CV026 bioassays confirmed that R7A* resembles R7AΔ*qseM*, in that TraI1-derived AHLs are constitutively produced and ICE*Ml*Sym^{R7A} is excised in almost 100% of cells in the population (81). Overexpression of *qseM* in the R7A* background reduced AHL production and ICE*Ml*Sym^{R7A} excision to below R7A levels, confirming that the R7A* phenotypes are a result of *qseM* repression. We also discovered that conjugation of ICE*Ml*Sym^{R7A} occurred at higher frequencies in R7A* compared to R7A, and that the exconjugants were all reset to the R7A phenotype, confirming that it is an epigenetic factor of the donor cells regulating the switch rather than a genetic mutation. We observed no differences in growth rates in rich and minimal media, nor any differences in carbon utilisation between R7A and R7A* cells. Independent R7A* cultures exhibited varying levels of AHLs in CV026 bioassays, however we were unable to isolate a revertant variant that switched from R7A* to R7A, indicating that once cells entered the R7A* state, they were stably maintained in that state. A series of promoter-*lacZ* fusions allowed us to determine that QseC and QseC2 regulated overlapping promoters by binding operator sites within their respective promoters in a concentration dependent manner. We also determined that the anti-*qseC* transcript was essential in preventing cells from become R7A* and must be expressed *in cis*. And finally, data presented here confirmed that QseC and QseC2 were essential for establishing and maintaining the epigenetic R7A* state.

Previously, our research group had proposed that QseC and its operator sites O_L and O_R constituted a molecular bimodal switch, whereby binding of QseC to the operator sites regulated its own expression as well as repressing *qseM* expression (83). Data presented here and in (87) confirms that QseC does indeed bind O_L and O_R to regulate P*qseC* and P*qseM*, and we additionally found that the transcription antisense to *qseC* is critical in preventing R7A cells from entering into the R7A* state. Overexpression of P*asqseC* on a plasmid did not complement the P*asqseC* deletion, however, replacement of P*asqseC*

with the *nptII* promoter in the chromosome did inhibit R7A* establishment, indicating that the antisense RNA functions only *in cis*. RNAseq shows that the abundance of *qseC* mRNA transcripts similar in R7A and R7A* (which lacks expression from P*asqseC*) (87), suggesting that P*asqseC*-facilitated repression of *qseC* is likely not due to transcription inhibition nor due to dsRNA-mediated mRNA degradation. It is unclear exactly how P*asqseC* might repress QseC. The lack of a change in *qseC* mRNA levels suggest *asqseC* RNA could inhibit QseC translation, however, it is unclear why such repression would not be functional when the RNA was expressed from a plasmid. Most characterised mechanisms of RNA-mediated translation inhibition involve binding to Shine-Dalgarno sequences in the 5' untranslated region. However, QseC is translated from a leaderless mRNA and lacks a Shine-Dalgarno sequence. Mechanisms such as ppGpp-mediated stress responses or abundance of translation factors regulates leaderless mRNA translation on a global scale, however, as far as we are aware there has been little investigation into the antisense regulation of individual leaderless mRNAs (205, 206).

The bimodal switch described in this chapter comprises two tandem controller proteins (C proteins) and their respective operator sites. Clusters of two or more QseC-like protein encoding genes have been found on distinct ICEs and each *qseC* can be clustered into six groups, each with distinct operator sites within the promoter region (81, 87, 129, 130). Some of these *qseC* genes have undergone mutations, duplications, or rearrangements and two or more *qseC* homologues are tandemly encoded with *qseM* and other QS gene homologues, for example *M. loti* NZP2037 encodes *qseC-qseC3-qseC2-qseC6* downstream of *qseM*. While it is yet to be experimentally tested, we suspect that the presence of these tandemly encoded *qseC* genes in diverse ICEs suggest there are multiple evolved molecular switches like that on ICE*Ml*Sym^{R7A}, and perhaps additional *qseC* homologues leads to even tighter regulation and stability of the switch.

In summary, data presented in this chapter provide further details on how the R7A* state is epigenetically activated and maintained to control QS and ICE*Ml*Sym^{R7A} transfer. Our model for R7A* establishment begins with the accumulation of QseC2 and QseC proteins, and at a certain concentration threshold for each protein they bind their respective O_L sites. Binding of QseC2 to $O2_L$ leads to repression of P*asqseC*, and binding of QseC to O_L leads to autoregulation of P*qseC* and repression of P*qseM* and consequently derepression of QS and ICE excision genes. As QseC and QseC2 protein concentrations increase, the O_R and $O2_R$ sites become occupied and repress P*qseC* and P*qseC2* respectively (Figure 6.17). Following transfer of ICE*Ml*Sym^{R7A} to a new host, QseC and QseC2 proteins are lost, and switch is reset.

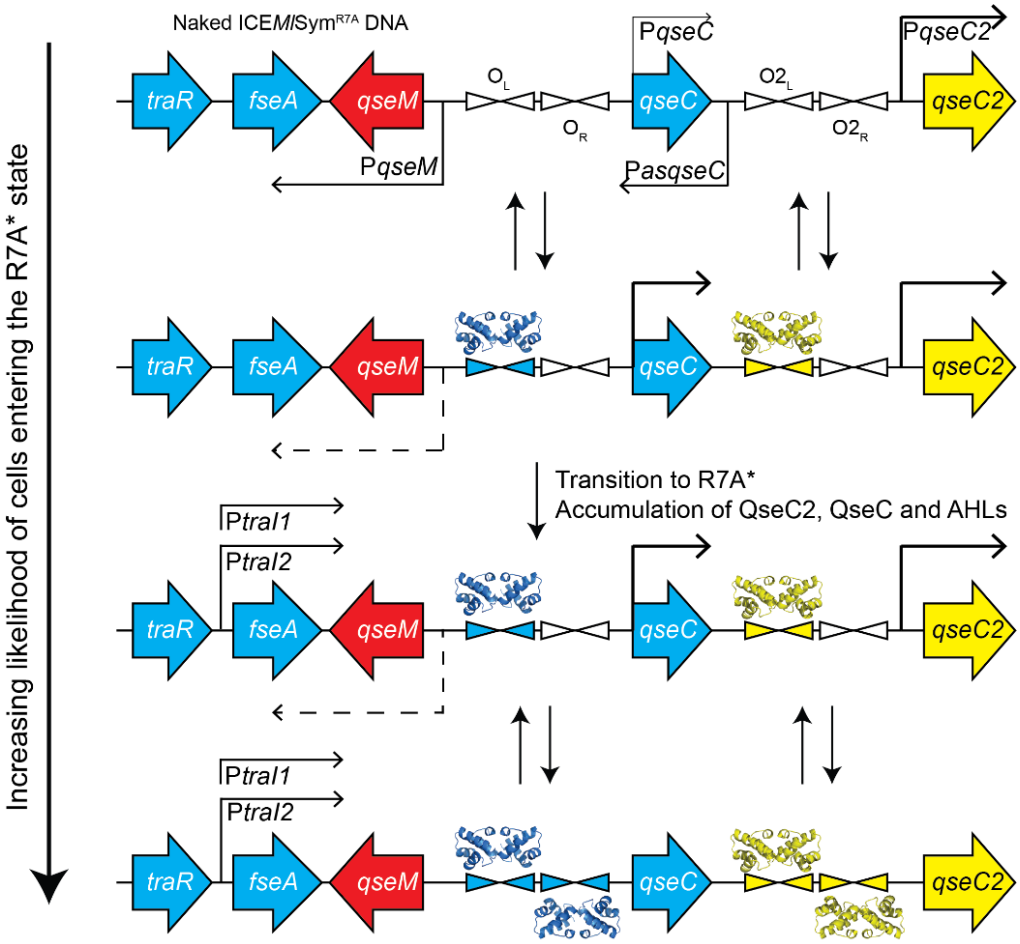


Figure 6.17 Overall model of QS and ICE*Ml*Sym^{R7A} excision and transfer in R7A and R7A*. In the absence of QseC or QseC2 protein and unoccupied

operator sites, the *qseM*, *asqseC* and *qseC2* promoters are all activated. Accumulation of QseC2 protein and binding of $O2_L$ leads to repression of P*asqseC* and subsequent derepression of P*qseC*. As QseC protein starts to accumulate and bind O_L, P*qseC* is upregulated resulting in cooperative binding of QseC to O_R, repressing *qseM* expression. Repression of *qseM* allows TraR and FseA to carry out their roles in AHL production and ICE*Ml*Sym^{R7A} excision and transfer. Figure adapted from (87).

Chapter 7

Quorum sensing cross-talk and quorum quenching

7.1 Introduction

Bacteria utilise quorum sensing to regulate various phenotypes on a population-density basis. LuxR-family receptors vary in specificity. Some are induced only by AHLs synthesised by the cognate LuxI, and others have a broader range of specificity and can perceive many AHLs of different chain lengths and C3 substitutions. As well as recognising self-produced AHLs some LuxR-type receptors can recognise AHLs synthesised by diverse neighbouring species within the environment, resulting in cross-talk between distinct QS systems of distinct organisms. For instance, LasR and QscR of *Pseudomonas aeruginosa* are widely promiscuous in the molecules they perceive (207-209). These characteristics essentially allow *P. aeruginosa* to sense and communicate with neighbouring bacteria in the environment.

Some bacteria also have the ability to inhibit quorum sensing, in a phenomenon called quorum quenching (QQ). QQ encompasses several mechanisms, including enzymatic degradation of the AHLs, LuxR antagonism and inhibition of AHL synthesis. AHL degradation is carried out by one of three enzymes families; oxidoreductases, acylases and lactonases. AHL-oxidoreductases act by oxidising or reducing the group at C3, changing the structure of the fatty acid chain, decreasing or eliminating its affinity to the LuxR-type receptor. AHL-acylases hydrolyse the amide bond between the homoserine lactone ring and the fatty acid chain, and AHL-lactonases hydrolyse the ester bond within the lactone moiety, opening the ring. Of the three AHL degradation enzymes, only reactions catalysed by lactonases are reversible, whereby the lactone ring is reformed under acidic conditions (210, 211). AHLs also undergo nonenzymatic lactonolysis under alkaline conditions which, like the lactonase reaction, is a reversible reaction under acidic conditions. Most AHL lactonases require metal ions for lactone binding and include a conserved metal-binding motif (115), however AidH of *Ochrobactrum sp*. functions with no metal ion and does not encode the canonical metal-binding motif of other lactonases (212).

Bacillus sp. encode an AHL lactonase AiiA, despite the fact that gram positive bacteria don't produce AHLs. *Pseudomonas aeruginosa* produces 3-oxo-C12-HSL which is a precursor to a bactericidal tetramic acid molecule, active against gram positive bacteria (213) and strains of *Bacillus* expressing AiiA are resistant to tetramic acid activity due to the inactivation of 3-oxo-C12-HSL molecules. AiiA of *Bacillus sp*. also provides protection to the plant hosts against QS-mediated virulence in neighbouring pathogenic bacteria (214).

In this chapter, we aimed to explore the specificity of *Mesorhizobium* QS systems in comparison to the well characterised QS systems of *P. aeruginosa*, and investigate an identified AHL-degradation protein, AidM. We demonstrated that TraR and MqsR recognise a narrow range of AHL molecules compared to the more indiscriminate LasR and QscR of *P. aeruginosa*. We confirmed that AidM functions as a lactonase and resembles AidH in that it does not require a metal ion to catalyse enzyme activity, nor does it include the canonical metal-binding motif. We also observed that AidM degrades *Mesorhizobium* AHLs with different affinities.

7.2 Results

7.2.1 Investigation of cross-talk between LuxR-type AHL receptors in *Mesorhizobium japonicum* R7A

The *M. japonicum* R7A TraI1 protein produces 3-oxo-C6-HSL QS molecules as the main AHL species (82) and MqsIC produces 2,4-trans-C12-HSL as the main AHL species (Chapter 3). β-Galactosidase assays were used to test whether MqsR is able to perceive 3-oxo-C6-HSL, and whether TraR is able to perceive 2,4-trans-C12-HSL and other C_{12} AHLs. A reporter plasmid was used to measure TraRI activity, where the *traI1* promoter is cloned into the promoterless *lacZ* cassette and the *traR* coding sequence is cloned into the IPTG inducible cassette of pSDz. This plasmid – pPtraI1zR – was introduced into R7ANSΔ*mqsRIC* to create a biosensor strain for measuring TraR-mediated expression of *traI1*. The P*mqsRNA7*/*mqsR* and P*traI*/*traR* biosensor strains were cultured with the synthetic AHLs and IPTG added during log-phase growth to a final concentration of 1 μM and 0.2 mM respectively. As described in section 3.2.7, MqsR was induced strongly by 2,4-trans-C12-HSL and weakly by 5-cis-C12-HSL and C12-HSL. However, 3-oxo-C6-HSL and C6-HSL were unable to induce the MQS system to a level significantly above that of the solvent-only control (Figure 7.1A). Likewise, TraR was activated by the known cognate TraI1 molecules – 3-oxo-C6-HSL and C6-HSL – but not by 2,4-trans-C12-HSL, 5-cis-C12-HSL or C12-HSL (Figure 7.1B). These results suggest that there is likely little to no cross-talk between the two QS systems of *M. japonicum* R7A, as the two LuxR receptors are unable to recognise the AHLs produced and perceived by the other system.

To determine whether the MqsI AHLs are perceived by other LuxR-family receptors outside of the *Mesorhizobium* genus, the synthetic AHLs were assayed with LasR and QscR from *Pseudomonas aeruginosa*, both known to respond to 3-oxo-C12-HSL. The LasR biosensor strain was *E. coli* DH5α harbouring two plasmids, pJN105L and pSC11-L, which carry an arabinose inducible *lasR* and a P*lasI*-*lacZ* fusion respectively. The QscR biosensor strain

was *E. coli* DH5α harbouring pJN105Q and pSC11-Q, which carry an arabinose inducible *qscR* and a P*PA1897-lacZ* fusion respectively. Both biosensor strains were cultured with the synthetic AHLs and arabinose added during log-phase to a final concentration of 1 μM and 0.1% respectively. LasR and QscR were both induced by all C12 AHLs tested to varying levels, with the exception of 2,4-trans-C12-HSL which was perceived by QscR but not LasR (Figure 7.1C-D).

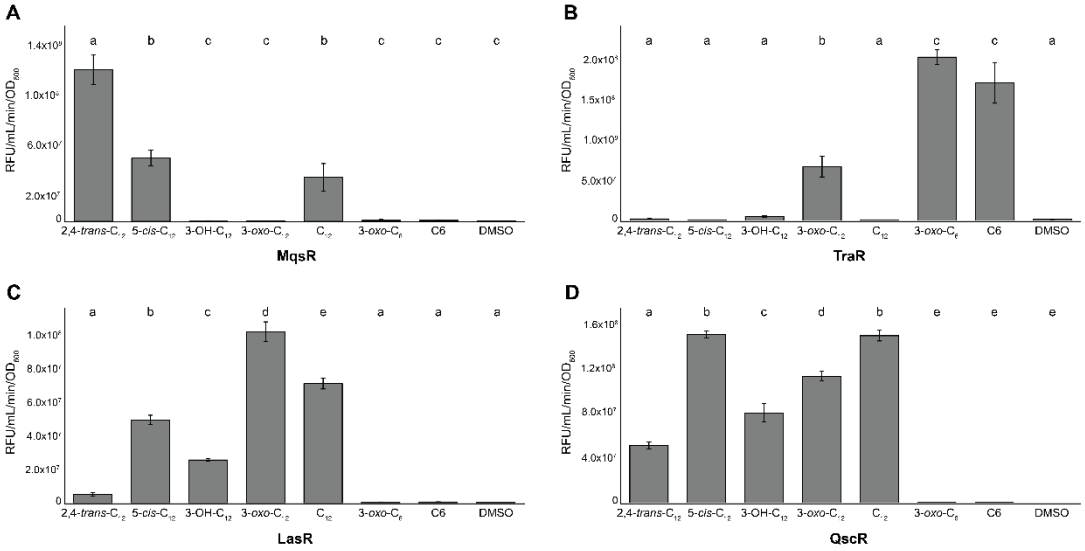


Figure 7.1 Substrate specificity of LuxR receptors. β-Galactosidase assays were used to determine the activity of (A) MqsR, (B) TraR, (C) LasR and (D) QscR in the presence of synthetic C6 and C12 AHLs. Biosensor strains were grown to stationary-phase in liquid cultures with AHLs and IPTG or arabinose added during log-phase to a final concentration of 1 μM and 0.2 mM or 0.1% respectively. Bars represent the mean of three biological replicates and error bars represent the standard deviation from the mean. A one-way ANOVA and Tukey post-hoc test was used to compare each strain. Values that are not statistically significant from each other are grouped by the same letter (a, b, c, d, e).

7.2.2 Identification and purification of an AHL degradation protein

As well as encoding a number of QS systems, some members of the *Mesorhizobium* genus also encode a quorum quenching (QQ) AHL-degradation enzyme, named here AidM. Our research group had recently bioinformatically identified the putative α/β-fold hydrolase in at least 20 *Mesorhizobium* strains, including *M. ciceri* WSM1271 and CC1192 with 81%

and 80% amino acid similarity to AidH and AiiO respectively, both of *Ochrobactrum* spp.. AidH functions as a lactonase (212) and there is conflicting information as to whether AiiO functions as an acylase or a lactonase (215, 216). Bioinformatic analysis alone could not ascertain the enzymatic function of the *Mesorhizobium* AHL-degradation protein. Here, multiple sequence alignments were carried out of the amino acid sequences of AidM (*M. ciceri* WSM1271), AidH (*Ochrobactrum sp*. T63) and AiiO (*Ochrobactrum sp*. A44) using T-coffee (133). The canonical catalytic triad Ser102-His248-Glu219 common to α/β-fold hydrolases (212) was present in AidM (Figure 7.2).

```
          ***    ****** *****  * *     ******* * *  ** *  *     ********* * **
AidM      MTISQKTLETSHGKIAVRETSGQGTAVMLIHGNSSSSAVFRNQLDGPLGERYHLIAPDLPGHGASGDA
AidH      MTINYHELETSHGRIAVRESEGEGAPLLMIHGNSSSGAIFAPQLEGEIGKKWRVIAPDLPGHGKSTDA
AiiO      MTINYHELETSHGRIAVRESEGEGASLLMIHGNSSSGAIFAPQLEGEIGKKWRVIAPDLPGHGKSSDA

          *** ***************   ***  * **********  *** * *   **********  ****
AidM      IDPERSYSMEGYADAMTEVLGLLGIDKAIVFGWSLGGHIGLEMIDRFPGLLGLMITGTPPVSPEEVGS
AidH      IDPDRSYSMEGYADAMTEVMQQLGIADAVVFGWSLGGHIGIEMIARYPEMRGLMITGTPPVAREEVGQ
AiiO      IDPDRSYSMEGYADAMTEVMQKLGIADAVVFGWSLGGHIGIEMIARYPAMRGLMITGTPPVAREEVGQ

          ***  * * ***** *   *** ***********  * * ********** *****  *** *** **
AidM      GFKPSPHMHLAGQETFTGADVEAYARSTCGEPFEPFLIDTVARTDGRARRLMFEKFAAGTGRNQREIV
AidH      GFKSGPDMALAGQEIFSERDVESYARSTCGEPFEASLLDIVARTDGRARRIMFEKFGSGTGGNQRDIV
AiiO      GFKSGPDMALAGQEIFSERDVESYARSTCGEPFETSLLDIVARTDGRARRIMFEKFGNGTGGNQRDIV

          *    **** ** ***** ****  **** *** *** ****  *  ***** *  *  ****
AidM      AGKTPPIAVLNGIDEPFVNTDFVSAVKFSNLWEGKTHLLDKSGHAPFWDSPDRFNPVLARFLASVDRA
AidH      AEAQLPIAVVNGRDEPFVELDFVSKVKFGNLWEGKTHVIDNAGHAPFREAPAEFDAYLARFIRDCTQ-
AiiO      AEAKLPIAVVNGRDEPFVELDFVSKVKFGNLWDGKTHVIDNSGHAPFREAPAEFDAYLARFIGDCTK-
```

Figure 7.2 Multiple sequence alignment of AidM, AidH and AidO. Alignments of AidM from *M. ciceri* WSM1271, AidH from *Ochrobactrum sp*. T63 and AiiO from *Ochrobactrum sp*. A44. The three resides of the catalytic triad are coloured red. The tyrosine residue coloured blue sits within the cap domain, interacts with the bound substrate and mutation of this residue was shown to completely abolish lactonase in AidH (212). Asterisks denote 100% identity.

To gain insight into the enzymatic mechanism and substrate specificity of AidM, constructs were designed for recombinant protein expression. The *aidM* coding sequence was codon optimised for protein expression in *E. coli*. The optimised gene was synthesised and cloned into the protein expression vector

pETM11 which incorporates a hexahistidine (6H) to the N-terminus of the protein. The same sequence was also synthesised and cloned with a mutation in one of the residues of the catalytic triad – S102G. AidM and AidM$_{S102G}$ were both purified from *E. coli* BL21(de3)pLysS as described in Chapter 2 (section 2.12). Peak fractions were collected from the IMAC and SEC purifications and samples run on a 4-15% SDS-PAGE to confirm the presence and purity of the recombinant AidM and AidM$_{S102G}$ proteins (Figure 7.3). Purified AidM protein was used in initial crystallisation trials using vapour diffusion. Various concentrations of protein from 0.5 mg/mL to 5 mg/mL were added to 288 screening buffer conditions (see methods), however no ideal conditions were identified and due to time constraints, this was not pursued further.

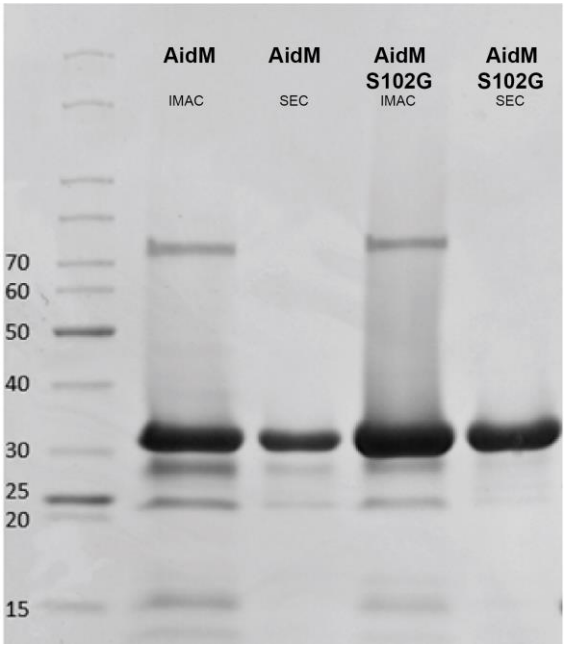


Figure 7.3 Purification of AidM and AidM$_{S102G}$ of WSM1271. A precast 4-15% SDS-PAGE was used to assay peak fractions at 280 nM UV from IMAC purifications of AidM and AidM$_{S102G}$ recombinant protein. AidM is expected to be 30 kDa as a monomer.

7.2.3 AidM is a lactonase and active against 3-oxo-C6-HSL *in vitro*

The AidM and AidM$_{S102G}$ recombinant proteins purified in section 7.2.2 were tested for their activity against 3-oxo-C6-HSL *in vitro*. We hypothesised that if AidM functions similarly to AidH of *Ochrobactrum* sp., we would observe

degradation of AHLs catalysed by the wild-type protein and a complete loss of degradation activity in the S102G mutant. Protein-AHL mixtures were prepared with either AidM or AidM$_{S102G}$ at 2.5 µM with 3-oxo-C6-HSL at 2.5 µM and incubated at room temperature for 30 minutes. *Chromobacterium violaceum* CV026 is a biosensor strain that produces a purple pigment in the presence of C4-C8 AHLs. Protein-AHL mixes were added to wells of CV026 bioassays, alongside a protein buffer control and 3-oxo-C6-HSL control. The CV026 bioassays revealed the recombinant AidM protein did indeed degrade the synthetic 3-oxo-C6-HSL, whereas the AidM$_{S102G}$ mutant resembled that of the buffer only control (Figure 7.4A), suggesting that the mutation in one of the putative catalytic triad residues is enough to abolish enzyme activity, and that AidM likely functions similarly to AidH.

Of the three families of AHL-degrading enzymes, lactonases are the only family where the AHL degradation can be reversed by acidification. To confirm that AidM is in fact a lactonase like AidH, AidM-3-oxo-C6-HSL mixes were filtered with a 0.22 µm syringe filter, adjusted to pH 2.0 (or left at pH 7.5) and incubated at room temperature for 30 minutes. CV026 assays revealed detectable AHLs in the acidified samples, suggesting reversal of lactonolysis (Figure 7.4B).

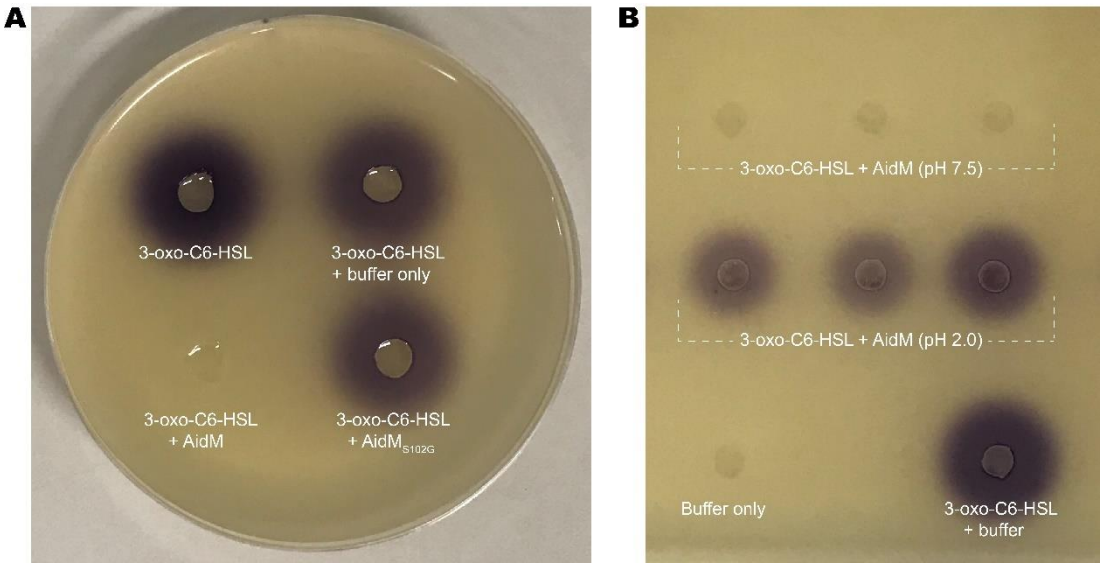


Figure 7.4 AidM is a lactonase and degrades 3-oxo-C6-HSL. *Chromobacterium violaceum* CV026 was used as a biosensor strain to analyse *in vitro* AidM activity against synthetic 3-oxo-C6-HSL. (A) Synthetic 3-oxo-C6-HSL (2.5 µM) was mixed with purified AidM protein (2.5 µM), purified AidM$_{S102G}$ protein (2.5 µM) and protein elution buffer only. CV026 bioassays revealed that the recombinant AidM protein was indeed active towards the synthetic AHLs, and the AidM$_{S102G}$ mutant was non-functional. (B) The synthetic 3-oxo-C6-HSL degraded by AidM was pH adjusted to 2.0 and added to CV026 bioassays. Production of purple violacein indicated intact AHLs in the acidified samples compared to the unadjusted samples at pH 7.5, confirming that AidM-mediated AHL degradation is reversible and therefore AidM acts as a lactonase.

7.2.4 AidM exhibits varying activity *in vivo*

After confirming that AidM is active on synthetic 3-oxo-C6-HSL *in vitro* (section 7.2.3), we next wanted to investigate AidM activity *in vivo*. BLASTp searches confirmed that *M. japonicum* R7A does not naturally encode this AHL degradation enzyme, so we used R7A as a background to avoid interference from other lactonases. Specifically, R7A* (described in depth in Chapter 6) was used for *in vivo* lactonolysis assays as this strain constitutively produces 3-oxo-C6-HSL AHLs. The *aidM* gene from *M. ciceri* WSM1271 was cloned into the pPR3 plasmid downstream of the constitutive *nptII* promoter to create pNaidM. The *aidM* overexpression plasmid as well as the empty vector was introduced into R7A* and grown to stationary phase in buffered TY media. Supernatants harvested from these cultures were added to bioassays using two separate biosensor strains – CV026 to detect C4-C8 AHLs synthesised by

TraI, and the MqsR biosensor to detect MqsIC-derived AHLs. Overexpression of *aidM* in R7A* resulted in the shorter chain AHLs being completely degraded to below detectable levels (Figure 7.5A). However, overexpression of *aidM* appears to have only partially degraded the longer chain MqsIC AHLs (Figure 7.5B).

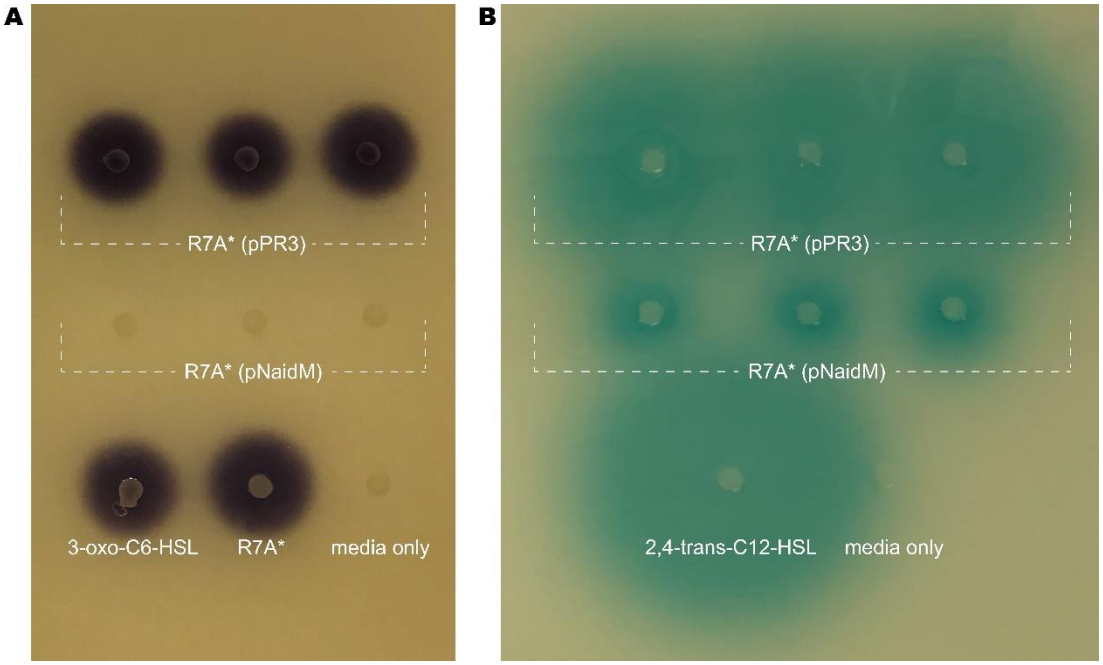


Figure 7.5 Expression of *aidM* in vivo exhibits substrate discrimination. The *aidM* gene was cloned downstream of PnptII in pPR3 and introduced into R7A*. R7A* strains overexpressing aidM were grown in buffered TY to stationary-phase and supernatants added to (A) CV026 bioassays and (B) MqsR bioassays. CV026 assays revealed that overexpression of *aidM* led to degradation of all detectable TraI1-derived AHLs. MqsR bioassays revealed that overexpression of *aidM* did not yield complete degradation of longer-chain AHLs.

7.2.5 AidM and ICE*Ml*Sym^{R7A} transfer

Horizontal transfer of ICE*Ml*Sym^{R7A} and ICE*Mc*Sym^{1271} are regulated by TraRI QS in *M. japonicum* R7A and *M. ciceri* WSM1271 respectively. We expected that overexpression of *aidM* in donor strains would decrease transfer rates, as QS would be disrupted resulting in a repression of ICE excision and conjugation. However, we wondered whether transfer rates would also be influenced by QQ in the recipient strains such as in *Rhizobium leguminosarum*, whereby sensing of AHLs produced by chromosomal AHL-synthases in the

recipient induces transfer. The pSaidM plasmid harbours *aidM* cloned from WSM127 under control of an IPTG inducible promoter. This plasmid alongside the empty vector pSacB were introduced into R7A, R7A* and R7ANS and the resulting strains were used in conjugation assays. Induced expression of *aidM* in R7A or R7A* donors resulted in an approximately 10-100-fold decrease in ICE*Ml*Sym^{R7A} transfer compared to the empty vector control donor strains. When *aidM* expression was induced in the recipient strains, no significant differences in ICE*Ml*Sym^{R7A} transfer were observed.

Table 7.1 Conjugation frequencies of ICE*Ml*Sym^{R7A} from R7A to R7ANS with or without *aidM*

Donor	Recipient[a]	Exconjugants (per donor)	Standard deviation
R7A (pSacB)	R7ANS	2.13×10^{-6}	9.26×10^{-7}
R7A (pSaidM)	R7ANS	1.36×10^{-8}	9.75×10^{-10}
R7A	R7ANS (pSacB)	2.91×10^{-6}	1.11×10^{-7}
R7A	R7ANS (pSaidM)	4.59×10^{-6}	3.40×10^{-7}
R7A* (pSacB)	R7ANS	3.52×10^{-5}	3.97×10^{-6}
R7A* (pSaidM)	R7ANS	2.73×10^{-6}	5.29×10^{-7}
R7A*	R7ANS (pSacB)	2.65×10^{-5}	4.25×10^{-6}
R7A*	R7ANS (pSaidM)	2.19×10^{-5}	6.73×10^{-6}

[a]All recipient strains harboured the pFAJ1700 plasmid to allow for tetracycline resistance selection

7.3 Discussion

In this chapter, we aimed to investigate cross-talk between QS systems of mesorhizobia and characterise the AHL degradation protein found in a large number of mesorhizobia. Data presented here confirms that this enzyme, AidM, acts as a lactonase to open the lactone ring of the AHL and the reaction is reversible by acidification. Mutation of the Ser102 which lies within the canonical catalytic triad inactivated the enzyme, consistent with the lactonase AidH of *Ochrobactrum sp*.. Interestingly, *in vivo* expression of aidM in R7A resulted in total abolishment of TraRI-derived AHLs (3-oxo-C6-HSL) but only partial degradation of MqsRIC AHLs (2,4-trans-C12-HSL). As expected, expression of *aidM* in donor strains resulted in decreased ICE*Ml*Sym^{R7A} transfer, and expression of *aidM* in the recipient strain had no effect on transfer frequencies.

Data presented in this chapter confirmed that there is little to no cross-talk between the TraRI and MqsRIC systems of *M. japonicum* R7A. Of the AHLs tested, MqsR responded strongly to the cognate 2,4-trans-C12-HSL molecule and to 5-cis-C12-HSL and C12-HSL with less affinity, whereas TraR responded strongly to the cognate 3-oxo-C6-HSL and C6-HSL and to 3-oxo-C12-HSL with less affinity. These two QS systems of mesorhizobia exhibit a relatively narrow spectrum of recognised AHLs compared to LasR and QscR of *Pseudomonas aeruginosa*, both of which recognise almost all C12 AHLs tested. The lack of cross-talk observed here complements the conjugation experiments in chapter four, in which it was demonstrated that the presence or absence of the *mqsRIC* genes had no effect on TraRI-regulated conjugative transfer.

Bacteria use quorum sensing to coordinate gene expression and various phenotypes based on population density. QS itself is tightly regulated, and individual QS systems are often a part of complex networks of sensor and response regulators, DNA-binding proteins and other QS systems within the same cell. AHL production and ICE excision and transfer is tightly regulated in

mesorhizobia (Chapter six, (80, 81, 85, 87, 217)), therefore the presence of *aidM* in some *Mesorhizobium* species is not surprising and likely adds an extra layer of regulation of these phenotypes. During previous work in our research group (131) it was attempted to create a markerless deletion of *aidM* in WSM1271 but was unsuccessful. A mutation was then made in the *traI1* gene, and a subsequent *aidM* deletion was successful. This suggests production and/or stability of TraI1-derived AHLs is lethal or at least causes a significant growth defect. A similar observation is made when the excisionase-encoding *rdfS* is overexpressed in R7A stimulating constitutive ICE*Ml*Sym^{R7A} excision, resulting in a severe growth defect close to lethality (85).

Just as QS is used as a means of cross-talk between different species, QQ is also used as a defence mechanism against competitors in the surrounding environment. The first AHL inactivating enzyme described was a lactonase in a soil-dwelling *Bacillus sp*., a gram-positive bacterium that does not encode any AHL-based QS genes. When expressed *in planta*, this lactonase inactivates AHLs, reducing QS-regulated virulence factors in the pathogenic *Erwinia carotovora* resulting in a significant decrease in host plant disease (214, 218). It is possible that the AidM lactonase of *Mesorhizobium* spp. may act in self-preservation as well as protecting the host plant from competing or pathogenic bacteria in the rhizosphere. Interestingly, data in this chapter revealed that expression of the *Mesorhizobium* lactonase *aidM* in R7A lead to complete degradation of TraI1-derived AHLs, but only partial degradation of MqsIC-derived AHLs. Further work is required to determine whether this is due to protein concentration imbalance or if our plasmid-borne *aidM* may not represent true biological expression, or whether AidM exhibits a narrow range of specificity and is indeed unable to degrade certain long-chain unsaturated AHLs. Further investigations should also be made as to what triggers lactonase expression. Some studies have claimed lactonase expression to be regulated by QS, however our RNAseq experiments carried out in chapter 5 did not find any significant differences in transcription of *aidM* between wild-type CC1192 and *mqsRIC* deletion mutants (fold change= 0.86, adjusted p

value=0.86), indicating that transcription of *aidM* is not directly regulated by MqsRIC.

In summary, we have determined that the QS systems in mesorhizobia exhibit minimal cross-talk between each other as well as LasR and QscR of *P. aeruginosa*. Some *Mesorhizobium* species encode an AHL lactonase which seems to completely degrade short-chain TraI1-derived AHLs but only partially degrade the longer-chain MqsIC-derived AHLs. There are many possibilities as to the inefficient inactivation of longer-chain AHLs, and this should be explored further to determine whether AidM has evolved to selectively target non-MqsRIC AHLs.

Chapter 8

Concluding remarks

Over the course of this study, we characterised two quorum sensing systems in the soil dwelling, nitrogen fixing *Mesorhizobium*. Mesorhizobia are generally known for forming symbiotic partnerships with leguminous plants, where they convert atmospheric nitrogen (N_2) to the bioavailable form ammonia (NH_3), in exchange for a carbon source. One QS system – MqsRIC – is encoded on the chromosome of almost all *Mesorhizobium* strains and uniquely incorporates a second AHL synthase to produce an unusual di-unsaturated AHL molecule to activate the cognate MqsR. Here, we also discovered that MqsRIC directly regulates a network of ncRNAs. The second QS system described in this thesis is TraRI of *M. japonicum* R7A which regulates excision and conjugative transfer of ICE*Ml*Sym^{R7A}. In this work we investigated the complex epigenetic regulatory mechanisms of QS and ICE transfer in an R7A variant called R7A*.

8.1 The MqsRIC QS system

Data presented in Chapters 3, 4 and 5 of this thesis describes the highly conserved quorum sensing system of *Mesorhizobium* – MqsRIC. The MQS system comprises the paradigm LuxR-family transcriptional activator and LuxI-family AHL synthase, MqsR and MqsI respectively. However, this system also uniquely encodes a second AHL synthase belonging to the crotonase enzyme family, MqsC. While this requirement for two AHL synthase genes seems to be unique in AHL-based QS systems, it is vaguely reminiscent of DSF-based QS systems whereby a crotonase-family enzyme dehydrates the fatty acid chain of the precursor molecule to produce an unsaturated autoinducer (15, 18, 19).

The presence of the crotonase-family protein encoding gene *mqsC* downstream of *mqsI* in what appeared to be an operon first led us to believe that the molecule produced was possibly unsaturated. To date, various unsaturated AHLs had been discovered in rhizobia, including 3-oxo-5-cis-C12-HSL and 5-cis-C12-HSL in a marine *Mesorhizobium* sp. (14). Using LC-MS/MS, we identified low levels of a molecule resembling 5-cis-C12-HSL in R7ANS supernatants, however the same molecule was also present in

mqsRIC mutants indicating that it is not produced by the MQS system. While outside the scope of this study, it is worth investigating this 5-cis-C12-HSL molecule in *Mesorhizobium* spp.. Preliminary bioinformatic searches in the R7A genome failed to uncover any uncharacterised *luxI*-type genes, therefore it's possible that this AHL is produced by a distinct autoinducer synthase, or perhaps even modified from another AHL.

Recently Suo et al. uncovered the di-unsaturated 2,4-trans-C12-HSL in *M. japonicum* MAFF 303099, produced by MqsIC homologues. This molecule was identified with LC-MS/MS methods scanning for all molecules, rather than using the 'gold standard' AHL identification method of scanning for product ions of m/z 102 (79). Further LC-MS/MS experiments on R7ANS strains, this time without discriminating based on product ions, revealed high amounts of 2,4-trans-C12-HSL in R7ANS wild-type AHL extracts, and the molecule was absent in AHL extracts from *mqsRIC* mutant strains. Our experience with this revelation relatively late into the study begs the question; how many other AHLs are naturally produced that are missed by the commonly used narrow-spectrum LC-MS/MS methods?

We hypothesised that MqsI likely acts like a LuxI-family AHL synthase, and the MqsC crotonase acts as a dehydratase to incorporate double bonds into the fatty acid chain. Whether this occurs before or after MqsI has carried out its function we do not know, although in this work MqsC was unable to convert exogenously added C12 AHLs, indicating that it may act on the fatty acid chain of the precursor, rather than the chain of AHL substrate. We expected LC-MS/MS analysis of AHL extracts from *mqsRIC* mutants overexpressing *mqsI* alone to provide insight into the intermediary molecule in the absence of MqsC. Extracts from this *mqsI* overexpression strain did exhibit a significant peak at approximately 9.3 minutes retention time which was not present in the wild-type extracts. However, the same molecule was also present in extracts from the *mqsC* overexpression strain and the empty vector control strain, suggesting it is unrelated, and no obvious candidate molecules were identified.

Determining the appropriate substrate molecules for MqsI and MqsC activity should be a focus of future works, including investigating the mechanism of action of MqsC and at which point in the 2,4-trans-C12-HSL synthesis it intervenes.

8.2 Possible roles of the MQS system

Unlike many other QS systems in rhizobia, data presented in this thesis suggests that MQS likely does not regulate symbiosis phenotypes. It has recently been reported that the *Mesorhizobium* genus includes a vast number of non-symbiotic strains (106) and this study determined that the non-symbiotic strains also encode *mqsRIC*, indicating a role outside of those involved in nitrogen fixation or symbiosis.

Proteins with related functions are often encoded together in the genome and the genomic context of the *mqsRIC* genes as well as the most conserved ncRNA – *mqsRNA7* – hint at a possible role in phosphate sensing, phosphate transport or the phosphate stress response (Figure 8.1). In every strain encoding *mqsRIC*, a convergent *ppk2* gene was found downstream of *mqsC*. Polyphosphate kinases exist in two distinct families, both of which are capable of catalysing inorganic polyphosphate (polyP) synthesis or polyP degradation, however PPK1 proteins preferentially synthesise polyP and PPK2 preferentially breaks down polyP. It has been suggested that PPK2 breaks down polyP to release stored energy and promiscuously phosphorylate nucleotides to regenerate the nucleotide pool as a stress response (219-221). Downstream of *mqsRIC* and *ppk2* are *phoR* and *phoU* of the Pho regulon. Elsewhere in the genome (in R7A, approximately 5,000 kb apart) *mqsRNA7* is consistently encoded upstream of *pstS*, *pstC*, *pstA*, *pstB*, *phoU* and *phoB*, all belonging to the Pho regulon. Under phosphate limiting conditions, the sensor kinase PhoR phosphorylates the response regulator PhoB and the phosphorylated PhoB binds and activates certain stress-response genes. PstSCAB form an ABC-type transporter complex, actively transporting phosphate into the cell during periods of phosphate starvation. Upon

phosphate repletion, PhoU is required to dephosphorylate PhoB and regulate the PstSCAB complex to prevent toxicity of uncontrolled phosphate uptake (222-225). The Pho regulon has been widely investigated over the last three decades (226) and it is well known that PhoR acts as a bifunctional kinase/phosphatase in response to extracellular inorganic phosphate levels. However, the exact mechanism of this phosphate sensing is still not well understood. PhoR does include a Per-Arnt-Sim (PAS) domain and it has been suggested that this PAS domain senses an unknown cytoplasmic signal molecule based on extracellular phosphate levels (227, 228). Although the Pho regulon exists ubiquitously in bacteria and the MqsRIC QS system is essentially restricted to the *Mesorhizobium* genus, it seems possible that the QS system and mqsRNAs may be involved in sensing extracellular phosphate levels, and perhaps even the activation of PhoR in mesorhizobia. RpfC of *Xanthomonas campestris* is a PAS domain-containing sensor kinase that recognises the distinct unsaturated DSF-family QS molecules to regulate expression of extracellular enzymes (23). RNAseq experiments carried out in this thesis did not reveal any transcriptional changes in any of Pho regulon genes, suggesting that if they are in some way regulated by MqsRIC and the mqsRNAs, it is likely via a post transcriptional mechanism or perhaps only occurs under certain conditions that were not tested in these experiments, such as phosphate limiting conditions.

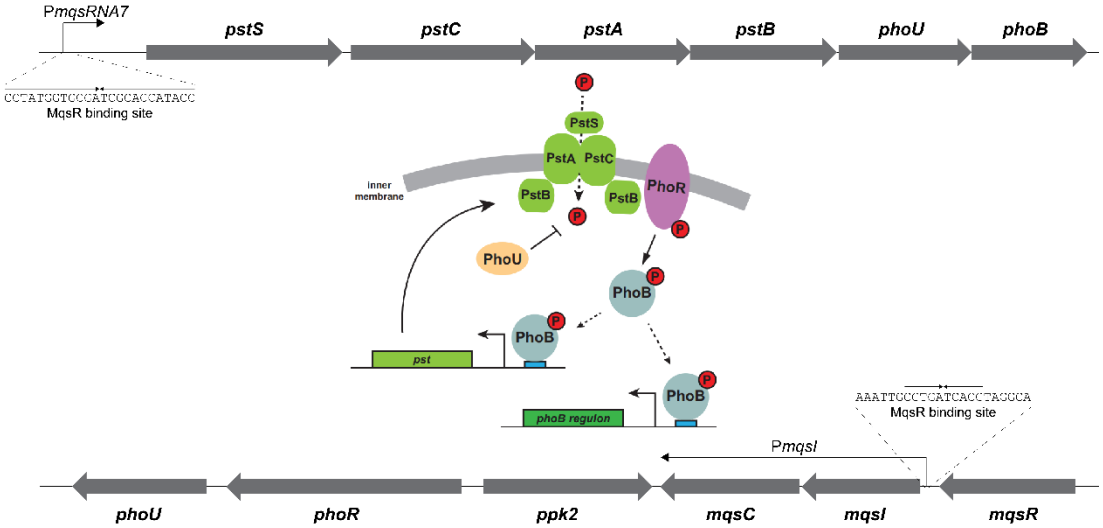


Figure 8.1 Gene map of *pho*, *mqs* and *ppk2* genes and depiction of the Pho regulon. Top: Map of genes downstream and including *mqsRNA7*. MqsR binding site within the *mqsRNA7* promoter region is shown. Bottom: Map of genes downstream and including mqsRIC. MqsR binding site within the *mqsI* promoter region is shown. Middle: Model of Pho regulon gene products. PstSCAB forms an ABC-like transporter complex, where PstS is the substrate-binding component, PstB is an ATP-binding protein which hydrolyses ATP to provide energy for the transport event and PstA and PstC are membrane permease proteins. The PstSCAB complex imports inorganic phosphate (red P). PhoR phosphorylates PhoB and phosphorylated PhoB activates target genes. PhoU is required for dephosphorylation by unknown mechanisms. Model of Pho proteins is from (229).

8.3 Epigenetic regulation of ICE*Ml*Sym^{R7A}

QS regulation of ICE*Ml*Sym^{R7A} has been well characterised, where TraI1 produces 3-oxo-C6-HSL to activate TraR, subsequently inducing expression from *tra*-boxes upstream of *traI1* and *traI2*, leading to upregulation of AHL synthesis and production of FseA via a programmed ribosomal frameshift in the *traI2-msi172-msi171* operon. This cascade of events results in *rdfS* expression triggering ICE*Ml*Sym^{R7A} excision from the chromosome and transfer to a recipient cell (82, 83, 85). However, it seems that almost every step of this process is inhibited in some manner. The antiactivator QseM represses TraR in the presence of AHLs and inactivates FseA-dependent transcriptional activation of P*rdfS*. Additionally, the ribosomal frameshift

producing FseA only occurs in 4-12% of translation events and overexpression of *traR*, *fseA* or *rdfS* all result in a significant growth defect (81-83, 85, 230). These regulatory factors suggest that ICE*Ml*Sym^{R7A} has likely evolved to remain in an integrated state, only excising for transfer under very specific circumstances.

Data presented in this thesis and in (87) describes further complex regulation of the excision and horizontal transfer of ICE*Ml*Sym^{R7A} by epigenetic mechanisms. Encoded downstream of *qseM* are two controller proteins QseC and QseC2 which bind operator sites within overlapping promoters in a concentration-dependent manner to regulate expression of P*qseM*, P*qseC*, P*qseC2* and the promoter of an RNA antisense to *qseC*. We confirmed that expression of *qseC2* triggers the QS “on” state in the R7A variant R7A*, and that the chromosomal *qseC2* gene is required for R7A* phenotype maintenance. In addition to QseM-mediated repression of ICE*Ml*Sym^{R7A} excision and transfer, the RNA antisense to *qseC* was confirmed here to inhibit cells from entering into the R7A* state in a so far unknown, but QseC-dependent mechanism. QseC is translated from a leaderless mRNA transcript (126) which lack the canonical Shine-Dalgarno sequence in the 5’ untranslated region, and instead generally employ direct binding of the 70s ribosomal particle to the start codon of the mRNA transcript (231, 232). Non-coding RNAs that act as inhibitors predominantly bind the 5’ untranslated region, occluding ribosome binding at the Sine-Dalgarno sequence (163). However, it has been shown that a sRNA in a *Salmonella sp*. binds within the first five codons of the coding region of its target gene (233), indicating a possibility for non-coding RNAs to inhibit leaderless mRNAs with a similar method. The complex regulation of TraRI QS in *M. japonicum* described here suggests ICE*Ml*Sym^{R7A} has evolved to favour integration in a host genome, only becoming primed for transfer under very specific conditions.

8.4 Conclusion

Work carried out in this thesis provides insight into the complexity of quorum sensing in gram negative bacteria. It is often found that different QS systems within a population are interconnected, with some forming hierarchical regulatory networks, whereby one system controls another. However, here we show that the MqsRIC and TraRI QS systems of mesorhizobia do not work in concert and that the molecules produced are distinct from one another, varying by chain length, C_3 substitution and fatty acid saturation. TraRI-based QS faces several suppressing mechanisms including the dual antiactivator activity of QseM, the low frequency of ribosomal frameshift events leading to expression of FseA, AHL degradation in strains encoding the AidM lactonase, growth defects when too much TraR, FseA or RdfS protein is produced and the precise regulation cascade of the QseC2-QseC genetic switch. On the contrary, MqsRIC-based QS does not appear to encounter such obstacles. We observe strong, unhindered MqsR activation from early exponential phase growth and evasion of AHL degradation by AidM. These observations suggest that the TraRI QS system has evolved to repress ICE excision and transfer unless under near-perfect conditions, but that the MqsRIC QS system is possibly more loosely regulated, perhaps involved in a core metabolism pathway, or phenotypes more critical to the growth and persistence of the population.

Further studies should focus on the network of small RNAs regulated by MqsR and their potential role in phosphate metabolism and polysaccharide production. Studies should also continue to investigate the mechanism of action of the antisense RNA *asqseC* – which we show to be crucial in the repression of TraRI QS – and investigate the triggering conditions for the R7A* switch to occur.

The work in this thesis brings to light the intricacies of bacterial communication, the sophisticated and sometimes tightly controlled regulation of certain phenotypes by different QS systems and the ability to participate in or block

chemical cross-talk. Understanding mechanisms such as those described in this thesis broadens our collective knowledge not only of QS in the rhizosphere, but also QS regulation of phenotypes including pathogenicity, horizontal gene transfer, virulence, motility and more across many bacteria.

Chapter 9

References

1. Fuqua WC, Winans SC, Greenberg EP. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. J Bacteriol. 1994;176(2):269-75.
2. Nealson KH. Autoinduction of bacterial luciferase. Occurrence, mechanism and significance. Arch Microbiol. 1977;112(1):73-9.
3. Nealson KH, Hastings JW. Bacterial bioluminescence: its control and ecological significance. Microbiol Rev. 1979;43(4):496-518.
4. Nealson KH, Platt T, Hastings JW. Cellular control of the synthesis and activity of the bacterial luminescent system. J Bacteriol. 1970;104(1):313-22.
5. Engebrecht J, Nealson K, Silverman M. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. Cell. 1983;32(3):773-81.
6. Engebrecht J, Silverman M. Identification of genes and gene products necessary for bacterial bioluminescence. Proc Natl Acad Sci U S A. 1984;81(13):4154-8.
7. Soto-Aceves MP, Diggle SP, Greenberg EP. Microbial Primer: LuxR-LuxI Quorum Sensing. Microbiology (Reading). 2023;169(9).
8. Watson WT, Minogue TD, Val DL, von Bodman SB, Churchill ME. Structural basis and specificity of acyl-homoserine lactone signal production in bacterial quorum sensing. Mol Cell. 2002;9(3):685-94.
9. Churchill ME, Chen L. Structural basis of acyl-homoserine lactone-dependent signaling. Chem Rev. 2011;111(1):68-85.
10. Val DL, Cronan JE, Jr. In vivo evidence that *S*-adenosylmethionine and fatty acid synthesis intermediates are the substrates for the LuxI family of autoinducer synthases. J Bacteriol. 1998;180(10):2644-51.
11. Ortori CA, Atkinson S, Chhabra SR, Camara M, Williams P, Barrett DA. Comprehensive profiling of *N*-acylhomoserine lactones produced by *Yersinia pseudotuberculosis* using liquid chromatography coupled to hybrid quadrupole-linear ion trap mass spectrometry. Anal Bioanal Chem. 2007;387(2):497-511.
12. Wagner-Dobler I, Thiel V, Eberl L, Allgaier M, Bodor A, Meyer S, et al. Discovery of complex mixtures of novel long-chain quorum sensing signals in free-living and host-associated marine alphaproteobacteria. Chembiochem. 2005;6(12):2195-206.
13. Teplitski M, Eberhard A, Gronquist MR, Gao M, Robinson JB, Bauer WD. Chemical identification of *N*-acyl homoserine lactone quorum-sensing signals produced by *Sinorhizobium meliloti* strains in defined medium. Arch Microbiol. 2003;180(6):494-7.

14. Krick A, Kehraus S, Eberl L, Riedel K, Anke H, Kaesler I, et al. A marine *Mesorhizobium sp*. produces structurally novel long-chain *N*-acyl-*L*-homoserine lactones. Appl Environ Microbiol. 2007;73(11):3587-94.
15. Barber CE, Tang JL, Feng JX, Pan MQ, Wilson TJ, Slater H, et al. A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. Mol Microbiol. 1997;24(3):555-66.
16. Tang JL, Liu YN, Barber CE, Dow JM, Wootton JC, Daniels MJ. Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. Mol Gen Genet. 1991;226(3):409-17.
17. Zhou L, Zhang LH, Camara M, He YW. The DSF Family of Quorum Sensing Signals: Diversity, Biosynthesis, and Turnover. Trends Microbiol. 2017;25(4):293-303.
18. Bi H, Christensen QH, Feng Y, Wang H, Cronan JE. The *Burkholderia cenocepacia* BDSF quorum sensing fatty acid is synthesized by a bifunctional crotonase homologue having both dehydratase and thioesterase activities. Mol Microbiol. 2012;83(4):840-55.
19. Zhou L, Yu Y, Chen X, Diab AA, Ruan L, He J, et al. The Multiple DSF-family QS Signals are Synthesized from Carbohydrate and Branched-chain Amino Acids via the FAS Elongation Cycle. Sci Rep. 2015;5:13294.
20. Slater H, Alvarez-Morales A, Barber CE, Daniels MJ, Dow JM. A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. Mol Microbiol. 2000;38(5):986-1003.
21. Deng Y, Schmid N, Wang C, Wang J, Pessi G, Wu D, et al. Cis-2-dodecenoic acid receptor RpfR links quorum-sensing signal perception with regulation of virulence through cyclic dimeric guanosine monophosphate turnover. Proc Natl Acad Sci U S A. 2012;109(38):15479-84.
22. Suppiger A, Eshwar AK, Stephan R, Kaever V, Eberl L, Lehner A. The DSF type quorum sensing signalling system RpfF/R regulates diverse phenotypes in the opportunistic pathogen *Cronobacter*. Sci Rep. 2016;6:18753.
23. An SQ, Allan JH, McCarthy Y, Febrer M, Dow JM, Ryan RP. The PAS domain-containing histidine kinase RpfS is a second sensor for the diffusible signal factor of *Xanthomonas campestris*. Mol Microbiol. 2014;92(3):586-97.
24. Deng Y, Wu J, Tao F, Zhang LH. Listening to a new language: DSF-based quorum sensing in Gram-negative bacteria. Chem Rev. 2011;111(1):160-73.
25. Ryan RP, Dow JM. Diffusible signals and interspecies communication in bacteria. Microbiology (Reading). 2008;154(Pt 7):1845-58.
26. Michael B, Smith JN, Swift S, Heffron F, Ahmer BM. SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. J Bacteriol. 2001;183(19):5733-42.
27. Walters M, Sperandio V. Quorum sensing in *Escherichia coli* and *Salmonella*. Int J Med Microbiol. 2006;296(2-3):125-31.
28. Duan K, Dammel C, Stein J, Rabin H, Surette MG. Modulation of *Pseudomonas aeruginosa* gene expression by host microflora through interspecies communication. Mol Microbiol. 2003;50(5):1477-91.

29. Williams P. Quorum sensing, communication and cross-kingdom signalling in the bacterial world. Microbiology (Reading). 2007;153(Pt 12):3923-38.
30. Hogan DA, Vik A, Kolter R. A *Pseudomonas aeruginosa* quorum-sensing molecule influences *Candida albicans* morphology. Mol Microbiol. 2004;54(5):1212-23.
31. Vilchez R, Lemme A, Ballhausen B, Thiel V, Schulz S, Jansen R, et al. *Streptococcus mutans* inhibits *Candida albicans* hyphal formation by the fatty acid signaling molecule trans-2-decenoic acid (SDSF). Chembiochem. 2010;11(11):1552-62.
32. Chan KG, Liu YC, Chang CY. Inhibiting *N*-acyl-homoserine lactone synthesis and quenching *Pseudomonas* quinolone quorum sensing to attenuate virulence. Front Microbiol. 2015;6:1173.
33. Chang CY, Krishnan T, Wang H, Chen Y, Yin WF, Chong YM, et al. Non-antibiotic quorum sensing inhibitors acting against *N*-acyl homoserine lactone synthase as druggable target. Sci Rep. 2014;4:7245.
34. Koh CL, Sam CK, Yin WF, Tan LY, Krishnan T, Chong YM, Chan KG. Plant-derived natural products as sources of anti-quorum sensing compounds. Sensors (Basel). 2013;13(5):6217-28.
35. Lade H, Paul D, Kweon JH. Quorum quenching mediated approaches for control of membrane biofouling. Int J Biol Sci. 2014;10(5):550-65.
36. Chbib C. Impact of the structure-activity relationship of AHL analogues on quorum sensing in Gram-negative bacteria. Bioorg Med Chem. 2020;28(3):115282.
37. Shukla A, Parmar P, Rao P, Goswami D, Saraf M. Twin Peaks: Presenting the Antagonistic Molecular Interplay of Curcumin with LasR and LuxR Quorum Sensing Pathways. Curr Microbiol. 2020;77(8):1800-10.
38. Swem LR, Swem DL, O'Loughlin CT, Gatmaitan R, Zhao B, Ulrich SM, Bassler BL. A quorum-sensing antagonist targets both membrane-bound and cytoplasmic receptors and controls bacterial pathogenicity. Mol Cell. 2009;35(2):143-53.
39. Wysoczynski-Horita CL, Boursier ME, Hill R, Hansen K, Blackwell HE, Churchill MEA. Mechanism of agonism and antagonism of the *Pseudomonas aeruginosa* quorum sensing regulator QscR with non-native ligands. Mol Microbiol. 2018;108(3):240-57.
40. Sikdar R, Elias M. Quorum quenching enzymes and their effects on virulence, biofilm, and microbiomes: a review of recent advances. Expert Rev Anti Infect Ther. 2020;18(12):1221-33.
41. Utari PD, Vogel J, Quax WJ. Deciphering Physiological Functions of AHL Quorum Quenching Acylases. Front Microbiol. 2017;8:1123.
42. Venturi V, Keel C. Signaling in the Rhizosphere. Trends Plant Sci. 2016;21(3):187-98.
43. Ferluga S, Steinder, L., Venturi, V. *N*-Acyl Homoserine Lactone Quorum Sensing in Gram-Negative Rhizobacteria. Secondary Metabolites in Soil Ecology. 14. Berlin: Springer; 2008. p. 69-90.
44. Nyffeler KE, Santa EE, Blackwell HE. Small Molecules with Either Receptor-Selective or Pan-Receptor Activity in the Three LuxR-Type Receptors that Regulate Quorum Sensing in *Pseudomonas aeruginosa*. ACS Chem Biol. 2022;17(11):2979-85.

45. Chugani SA, Whiteley M, Lee KM, D'Argenio D, Manoil C, Greenberg EP. QscR, a modulator of quorum-sensing signal synthesis and virulence in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A. 2001;98(5):2752-7.
46. Kostylev M, Kim DY, Smalley NE, Salukhe I, Greenberg EP, Dandekar AA. Evolution of the *Pseudomonas aeruginosa* quorum-sensing hierarchy. Proc Natl Acad Sci U S A. 2019;116(14):7027-32.
47. Qin S, Xiao W, Zhou C, Pu Q, Deng X, Lan L, et al. *Pseudomonas aeruginosa*: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. Signal Transduct Target Ther. 2022;7(1):199.
48. LaSarre B, Federle MJ. Exploiting quorum sensing to confuse bacterial pathogens. Microbiol Mol Biol Rev. 2013;77(1):73-111.
49. Gelvin SB. *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev. 2003;67(1):16-37, table of contents.
50. Chilton MD, Saiki RK, Yadav N, Gordon MP, Quetier F. T-DNA from *Agrobacterium* Ti plasmid is in the nuclear DNA fraction of crown gall tumor cells. Proc Natl Acad Sci U S A. 1980;77(7):4060-4.
51. Lee CW, Efetova M, Engelmann JC, Kramell R, Wasternack C, Ludwig-Muller J, et al. *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. Plant Cell. 2009;21(9):2948-62.
52. Fuqua WC, Winans SC. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. J Bacteriol. 1994;176(10):2796-806.
53. Zhang L, Murphy PJ, Kerr A, Tate ME. *Agrobacterium* conjugation and gene regulation by *N*-acyl-L-homoserine lactones. Nature. 1993;362(6419):446-8.
54. Hwang I, Li PL, Zhang L, Piper KR, Cook DM, Tate ME, Farrand SK. TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid *N*-acylhomoserine lactone autoinducer. Proc Natl Acad Sci U S A. 1994;91(11):4639-43.
55. Pappas KM, Winans SC. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. Mol Microbiol. 2003;48(4):1059-73.
56. Zhu J, Winans SC. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. Proc Natl Acad Sci U S A. 1999;96(9):4832-7.
57. Zhu J, Winans SC. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proc Natl Acad Sci U S A. 2001;98(4):1507-12.
58. Fuqua C, Burbea M, Winans SC. Activity of the *Agrobacterium* Ti plasmid conjugal transfer regulator TraR is inhibited by the product of the *traM* gene. J Bacteriol. 1995;177(5):1367-73.
59. Swiderska A, Berndtson AK, Cha MR, Li L, Beaudoin GM, 3rd, Zhu J, Fuqua C. Inhibition of the *Agrobacterium tumefaciens* TraR quorum-sensing regulator. Interactions with the TraM anti-activator. J Biol Chem. 2001;276(52):49449-58.

60. Chai Y, Zhu J, Winans SC. TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function in vitro by forming inactive TrlR:TraR dimers. Mol Microbiol. 2001;40(2):414-21.
61. Zhu J, Winans SC. Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated, dominant defective TraR-like protein. Mol Microbiol. 1998;27(2):289-97.
62. Gray KM, Pearson JP, Downie JA, Boboye BE, Greenberg EP. Cell-to-cell signaling in the symbiotic nitrogen-fixing bacterium *Rhizobium leguminosarum*: autoinduction of a stationary phase and rhizosphere-expressed genes. J Bacteriol. 1996;178(2):372-6.
63. Schripsema J, de Rudder KE, van Vliet TB, Lankhorst PP, de Vroom E, Kijne JW, van Brussel AA. Bacteriocin small of *Rhizobium leguminosarum* belongs to the class of *N*-acyl-L-homoserine lactone molecules, known as autoinducers and as quorum sensing co-transcription factors. J Bacteriol. 1996;178(2):366-71.
64. Lithgow JK, Wilkinson A, Hardman A, Rodelas B, Wisniewski-Dye F, Williams P, Downie JA. The regulatory locus *cinRI* in *Rhizobium leguminosarum* controls a network of quorum-sensing loci. Mol Microbiol. 2000;37(1):81-97.
65. Frederix M, Edwards A, McAnulla C, Downie JA. Co-ordination of quorum-sensing regulation in *Rhizobium leguminosarum* by induction of an anti-repressor. Mol Microbiol. 2011;81(4):994-1007.
66. Edwards A, Frederix M, Wisniewski-Dye F, Jones J, Zorreguieta A, Downie JA. The *cin* and *rai* quorum-sensing regulatory systems in *Rhizobium leguminosarum* are coordinated by ExpR and CinS, a small regulatory protein coexpressed with CinI. J Bacteriol. 2009;191(9):3059-67.
67. Rodelas B, Lithgow JK, Wisniewski-Dye F, Hardman A, Wilkinson A, Economou A, et al. Analysis of quorum-sensing-dependent control of rhizosphere-expressed (*rhi*) genes in *Rhizobium leguminosarum* bv. *viciae*. J Bacteriol. 1999;181(12):3816-23.
68. Cubo MT, Economou A, Murphy G, Johnston AW, Downie JA. Molecular characterization and regulation of the rhizosphere-expressed genes *rhiABCR* that can influence nodulation by *Rhizobium leguminosarum* biovar *viciae*. J Bacteriol. 1992;174(12):4026-35.
69. Sanchez-Contreras M, Bauer WD, Gao M, Robinson JB, Allan Downie J. Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. Philos Trans R Soc Lond B Biol Sci. 2007;362(1483):1149-63.
70. Zheng H, Mao Y, Zhu Q, Ling J, Zhang N, Naseer N, et al. The quorum sensing regulator CinR hierarchically regulates two other quorum sensing pathways in ligand-dependent and -independent fashions in *Rhizobium etli*. J Bacteriol. 2015;197(9):1573-81.
71. Danino VE, Wilkinson A, Edwards A, Downie JA. Recipient-induced transfer of the symbiotic plasmid pRL1JI in *Rhizobium leguminosarum* bv. *viciae* is regulated by a quorum-sensing relay. Mol Microbiol. 2003;50(2):511-25.
72. Marketon MM, Gronquist MR, Eberhard A, Gonzalez JE. Characterization of the *Sinorhizobium meliloti sinR*/*sinI* locus and the production of novel *N*-acyl homoserine lactones. J Bacteriol. 2002;184(20):5686-95.

73. Bahlawane C, McIntosh M, Krol E, Becker A. *Sinorhizobium meliloti* regulator MucR couples exopolysaccharide synthesis and motility. Mol Plant Microbe Interact. 2008;21(11):1498-509.
74. Pellock BJ, Teplitski M, Boinay RP, Bauer WD, Walker GC. A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. J Bacteriol. 2002;184(18):5067-76.
75. Calatrava-Morales N, McIntosh M, Soto MJ. Regulation Mediated by *N*-Acyl Homoserine Lactone Quorum Sensing Signals in the *Rhizobium*-Legume Symbiosis. Genes (Basel). 2018;9(5).
76. Marketon MM, Gonzalez JE. Identification of two quorum-sensing systems in *Sinorhizobium meliloti*. J Bacteriol. 2002;184(13):3466-75.
77. Cao H, Yang M, Zheng H, Zhang J, Zhong Z, Zhu J. Complex quorum-sensing regulatory systems regulate bacterial growth and symbiotic nodulation in *Mesorhizobium tianshanense*. Arch Microbiol. 2009;191(3):283-9.
78. Yang M, Sun K, Zhou L, Yang R, Zhong Z, Zhu J. Functional analysis of three AHL autoinducer synthase genes in *Mesorhizobium loti* reveals the important role of quorum sensing in symbiotic nodulation. Can J Microbiol. 2009;55(2):210-4.
79. Suo Z, Cummings DA, Jr., Puri AW, Schaefer AL, Greenberg EP. A *Mesorhizobium japonicum* quorum sensing circuit that involves three linked genes and an unusual acyl-homoserine lactone signal. mBio. 2023:e0101023.
80. Haskett TL, Terpolilli JJ, Ramachandran VK, Verdonk CJ, Poole PS, O'Hara GW, Ramsay JP. Sequential induction of three recombination directionality factors directs assembly of tripartite integrative and conjugative elements. PLoS Genet. 2018;14(3):e1007292.
81. Ramsay JP, Major AS, Komarovsky VM, Sullivan JT, Dy RL, Hynes MF, et al. A widely conserved molecular switch controls quorum sensing and symbiosis island transfer in *Mesorhizobium loti* through expression of a novel antiactivator. Mol Microbiol. 2013;87(1):1-13.
82. Ramsay JP, Sullivan JT, Jambari N, Ortori CA, Heeb S, Williams P, et al. A LuxRI-family regulatory system controls excision and transfer of the *Mesorhizobium loti* strain R7A symbiosis island by activating expression of two conserved hypothetical genes. Mol Microbiol. 2009;73(6):1141-55.
83. Ramsay JP, Tester LG, Major AS, Sullivan JT, Edgar CD, Kleffmann T, et al. Ribosomal frameshifting and dual-target antiactivation restrict quorum-sensing-activated transfer of a mobile genetic element. Proc Natl Acad Sci U S A. 2015;112(13):4104-9.
84. Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc Natl Acad Sci U S A. 1995;92(19):8985-9.
85. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW. Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. Mol Microbiol. 2006;62(3):723-34.

86. Sullivan JT, Ronson CW. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc Natl Acad Sci U S A. 1998;95(9):5145-9.
87. Ramsay JP, Bastholm TR, Verdonk CJ, Tambalo DD, Sullivan JT, Harold LK, et al. An epigenetic switch activates bacterial quorum sensing and horizontal transfer of an integrative and conjugative element. Nucleic Acids Res. 2022;50(2):975-88.
88. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat Rev Microbiol. 2007;5(8):619-33.
89. Oldroyd GE, Murray JD, Poole PS, Downie JA. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet. 2011;45:119-44.
90. Sprent JI. 60Ma of legume nodulation. What's new? What's changing? J Exp Bot. 2008;59(5):1081-4.
91. Geurts R, Bisseling T. *Rhizobium* nod factor perception and signalling. Plant Cell. 2002;14 Suppl(Suppl):S239-49.
92. Wais RJ, Keating DH, Long SR. Structure-function analysis of nod factor-induced root hair calcium spiking in *Rhizobium*-legume symbiosis. Plant Physiol. 2002;129(1):211-24.
93. Kamboj DV, Bhatia R, Pathak DV, Sharma PK. Role of *nodD* gene product and flavonoid interactions in induction of nodulation genes in *Mesorhizobium ciceri*. Physiol Mol Biol Plants. 2010;16(1):69-77.
94. Spaink HP, Wijffelman CA, Pees E, Okker RJH, Lugtenberg BJJ. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature. 1987;328(6128):337-40.
95. Perret X, Staehelin C, Broughton WJ. Molecular basis of symbiotic promiscuity. Microbiol Mol Biol Rev. 2000;64(1):180-201.
96. Alunni B, Gourion B. Terminal bacteroid differentiation in the legume-*rhizobium* symbiosis: nodule-specific cysteine-rich peptides and beyond. New Phytol. 2016;211(2):411-7.
97. Fournier J, Timmers AC, Sieberer BJ, Jauneau A, Chabaud M, Barker DG. Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. Plant Physiol. 2008;148(4):1985-95.
98. Gage DJ. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev. 2004;68(2):280-300.
99. Esseling JJ, Lhuissier FG, Emons AM. Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. Plant Physiol. 2003;132(4):1982-8.
100. Murray JD. Invasion by invitation: rhizobial infection in legumes. Mol Plant Microbe Interact. 2011;24(6):631-9.
101. Robertson JG, Lyttleton P. Division of peribacteroid membranes in root nodules of white clover. J Cell Sci. 1984;69:147-57.
102. Sciotti MA, Chanfon A, Hennecke H, Fischer HM. Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*. J Bacteriol. 2003;185(18):5639-42.

103. Tsoy OV, Ravcheev DA, Cuklina J, Gelfand MS. Nitrogen Fixation and Molecular Oxygen: Comparative Genomic Reconstruction of Transcription Regulation in Alphaproteobacteria. Front Microbiol. 2016;7:1343.
104. Udvardi M, Poole PS. Transport and metabolism in legume-rhizobia symbioses. Annu Rev Plant Biol. 2013;64:781-805.
105. Laranjo M, Alexandre A, Oliveira S. Legume growth-promoting rhizobia: an overview on the *Mesorhizobium* genus. Microbiol Res. 2014;169(1):2-17.
106. Colombi E, Hill Y, Lines R, Sullivan JT, Kohlmeier MG, Christophersen CT, et al. Population genomics of Australian indigenous *Mesorhizobium* reveals diverse nonsymbiotic genospecies capable of nitrogen-fixing symbioses following horizontal gene transfer. Microb Genom. 2023;9(1).
107. Miller JH. Experiments in Molecular Genetics. New York: Cold Spring Harbor Laboratory Press; 1972.
108. Thoma S, Schobert M. An improved *Escherichia coli* donor strain for diparental mating. FEMS Microbiol Lett. 2009;294(2):127-32.
109. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol. 1974;84(1):188-98.
110. Ronson CW, Nixon BT, Albright LM, Ausubel FM. *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J Bacteriol. 1987;169(6):2424-31.
111. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 1986;189(1):113-30.
112. Robichon C, Luo J, Causey TB, Benner JS, Samuelson JC. Engineering *Escherichia coli* BL21(DE3) derivative strains to minimize *E. coli* protein contamination after purification by immobilized metal affinity chromatography. Appl Environ Microbiol. 2011;77(13):4634-46.
113. Corbin E, J., Brockwell, J., Gault, R., R. Nodulation studies on chickpea (*Cicer arietinum*). Australian Journal of Experimental Agriculture. 1977;17.
114. Nandasena K, Yates R, Tiwari R, O'Hara G, Howieson J, Ninawi M, et al. Complete genome sequence of *Mesorhizobium ciceri* bv. *biserrulae* type strain (WSM1271(T)). Stand Genomic Sci. 2014;9(3):462-72.
115. Elias NV, Herridge DF. Naturalised populations of mesorhizobia in chickpea (*Cicer arietinum* L.) cropping soils: effects on nodule occupancy and productivity of commercial chickpea. Plant and Soil. 2015;387(1):233-49.
116. McClean KH, Winson MK, Fish L, Taylor A, Chhabra SR, Camara M, et al. Quorum sensing and Chromobacterium violaceum: exploitation of violacein production and inhibition for the detection of N-acylhomoserine lactones. Microbiology (Reading). 1997;143 (Pt 12):3703-11.
117. Lee JH, Lequette Y, Greenberg EP. Activity of purified QscR, a *Pseudomonas aeruginosa* orphan quorum-sensing transcription factor. Mol Microbiol. 2006;59(2):602-9.
118. Eibergen NR, Moore JD, Mattmann ME, Blackwell HE. Potent and Selective Modulation of the RhlR Quorum Sensing Receptor by Using Non-native Ligands: An Emerging Target for Virulence Control in *Pseudomonas aeruginosa*. Chembiochem. 2015;16(16):2348-56.
119. Rodpothong P, Sullivan JT, Songsrirote K, Sumpton D, Cheung KW, Thomas-Oates J, et al. Nodulation gene mutants of *Mesorhizobium loti* R7A-

nodZ and *nolL* mutants have host-specific phenotypes on *Lotus* spp. Mol Plant Microbe Interact. 2009;22(12):1546-54.
120. Miller WG, Leveau JH, Lindow SE. Improved *gfp* and *inaZ* broad-host-range promoter-probe vectors. Mol Plant Microbe Interact. 2000;13(11):1243-50.
121. Dombrecht B, Vanderleyden J, Michiels J. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol Plant Microbe Interact. 2001;14(3):426-30.
122. Dummler A, Lawrence AM, de Marco A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. Microb Cell Fact. 2005;4:34.
123. Hoang TT, Karkhoff-Schweizer RR, Kutchma AJ, Schweizer HP. A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. Gene. 1998;212(1):77-86.
124. Quandt J, Hynes MF. Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene. 1993;127(1):15-21.
125. Antoine R, Alonso S, Raze D, Coutte L, Lesjean S, Willery E, et al. New virulence-activated and virulence-repressed genes identified by systematic gene inactivation and generation of transcriptional fusions in *Bordetella pertussis*. J Bacteriol. 2000;182(20):5902-5.
126. Ramsay JP. High-throughput β-galactosidase and β-glucuronidase Assays Using Fluorogenic Substrates. Bio-protocol. 2013;3(14).
127. Gasteiger E. HC, Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. The Proteomics Protocols Handbook, Humana Press: Humana Press; 2005.
128. Howieson JGaD, M.J. Working with rhizobia. Canberra: Australian Centre for International Agricultural Research 2016.
129. Colombi E, Perry BJ, Sullivan JT, Bekuma AA, Terpolilli JJ, Ronson CW, Ramsay JP. Comparative analysis of integrative and conjugative mobile genetic elements in the genus *Mesorhizobium*. Microb Genom. 2021;7(10).
130. Ramsay JP. Regulation of Excision and Transfer of the *Mesorhizobium loti* R7A Symbiosis Island. Dunedin, New Zealand: University of Otago; 2008.
131. Haskett TL. Discovery and characterisation of tripartite Integrative & Conjugative Elements. Perth, Western Australia: Murdoch University; 2018.
132. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics. 2011;12:124.
133. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000;302(1):205-17.
134. Choi SH, Greenberg EP. The C-terminal region of the *Vibrio fischeri* LuxR protein contains an inducer-independent *lux* gene activating domain. Proc Natl Acad Sci U S A. 1991;88(24):11115-9.
135. Hanzelka BL, Greenberg EP. Evidence that the N-terminal region of the *Vibrio fischeri* LuxR protein constitutes an autoinducer-binding domain. J Bacteriol. 1995;177(3):815-7.

136. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021;38(7):3022-7.
137. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14(6):1188-90.
138. Fuqua C, Greenberg EP. Listening in on bacteria: acyl-homoserine lactone signalling. Nat Rev Mol Cell Biol. 2002;3(9):685-95.
139. Patankar AV, Gonzalez JE. Orphan LuxR regulators of quorum sensing. FEMS Microbiol Rev. 2009;33(4):739-56.
140. Poulter S, Carlton TM, Spring DR, Salmond GP. The *Serratia* LuxR family regulator CarR 39006 activates transcription independently of cognate quorum sensing signals. Mol Microbiol. 2011;80(4):1120-31.
141. MacLellan SR, MacLean AM, Finan TM. Promoter prediction in the rhizobia. Microbiology (Reading). 2006;152(Pt 6):1751-63.
142. Patel NM, Moore JD, Blackwell HE, Amador-Noguez D. Identification of Unanticipated and Novel *N*-Acyl L-Homoserine Lactones (AHLs) Using a Sensitive Non-Targeted LC-MS/MS Method. PLoS One. 2016;11(10):e0163469.
143. Rodrigues AMS, Lami R, Escoubeyrou K, Intertaglia L, Mazurek C, Doberva M, et al. Straightforward *N*-Acyl Homoserine Lactone Discovery and Annotation by LC-MS/MS-based Molecular Networking. J Proteome Res. 2022;21(3):635-42.
144. Richard Evans MON, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Zidek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper and Demis Hassabis. Protein complex prediction with AlphaFold-Multimer. bioRxiv. 2021.
145. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.
146. Dong SH, Frane ND, Christensen QH, Greenberg EP, Nagarajan R, Nair SK. Molecular basis for the substrate specificity of quorum signal synthases. Proc Natl Acad Sci U S A. 2017;114(34):9092-7.
147. Mathesius U, Mulders S, Gao M, Teplitski M, Caetano-Anolles G, Rolfe BG, Bauer WD. Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. Proc Natl Acad Sci U S A. 2003;100(3):1444-9.
148. Wisniewski-Dye F, Downie JA. Quorum-sensing in *Rhizobium*. Antonie Van Leeuwenhoek. 2002;81(1-4):397-407.
149. McIntosh M, Meyer S, Becker A. Novel *Sinorhizobium meliloti* quorum sensing positive and negative regulatory feedback mechanisms respond to phosphate availability. Mol Microbiol. 2009;74(5):1238-56.
150. Cheng HP, Walker GC. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol. 1998;180(19):5183-91.
151. Glenn SA, Gurich N, Feeney MA, Gonzalez JE. The ExpR/Sin quorum-sensing system controls succinoglycan production in *Sinorhizobium meliloti*. J Bacteriol. 2007;189(19):7077-88.
152. Marketon MM, Glenn SA, Eberhard A, Gonzalez JE. Quorum sensing controls exopolysaccharide production in *Sinorhizobium meliloti*. J Bacteriol. 2003;185(1):325-31.

153. Gonzalez JE, Marketon MM. Quorum sensing in nitrogen-fixing rhizobia. Microbiol Mol Biol Rev. 2003;67(4):574-92.
154. Kelly SJ, Muszynski A, Kawaharada Y, Hubber AM, Sullivan JT, Sandal N, et al. Conditional requirement for exopolysaccharide in the *Mesorhizobium-Lotus* symbiosis. Mol Plant Microbe Interact. 2013;26(3):319-29.
155. Muszynski A, Heiss C, Hjuler CT, Sullivan JT, Kelly SJ, Thygesen MB, et al. Structures of Exopolysaccharides Involved in Receptor-mediated Perception of *Mesorhizobium loti* by *Lotus japonicus*. J Biol Chem. 2016;291(40):20946-61.
156. Dilworth MJ. Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum*. Biochim Biophys Acta. 1966;127(2):285-94.
157. Schaper S, Krol E, Skotnicka D, Kaever V, Hilker R, Sogaard-Andersen L, Becker A. Cyclic Di-GMP Regulates Multiple Cellular Functions in the Symbiotic Alphaproteobacterium *Sinorhizobium meliloti*. J Bacteriol. 2016;198(3):521-35.
158. Bahlawane C, Baumgarth B, Serrania J, Ruberg S, Becker A. Fine-tuning of galactoglucan biosynthesis in *Sinorhizobium meliloti* by differential WggR (ExpG)-, PhoB-, and MucR-dependent regulation of two promoters. J Bacteriol. 2008;190(10):3456-66.
159. Becker A, Ruberg S, Kuster H, Roxlau AA, Keller M, Ivashina T, et al. The 32-kilobase exp gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. J Bacteriol. 1997;179(4):1375-84.
160. Reed JW, Capage M, Walker GC. *Rhizobium meliloti exoG* and *exoJ* mutations affect the *exoX-exoY* system for modulation of exopolysaccharide production. J Bacteriol. 1991;173(12):3776-88.
161. Zhan HJ, Leigh JA. Two genes that regulate exopolysaccharide production in *Rhizobium meliloti*. J Bacteriol. 1990;172(9):5254-9.
162. Han Y, Li C, Yan Y, Lin M, Ke X, Zhang Y, Zhan Y. Post-transcriptional control of bacterial nitrogen metabolism by regulatory noncoding RNAs. World J Microbiol Biotechnol. 2022;38(7):126.
163. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell. 2011;43(6):880-91.
164. Ceizel Borella G, Lagares A, Jr., Valverde C. Expression of the *Sinorhizobium meliloti* small RNA gene *mmgR* is controlled by the nitrogen source. FEMS Microbiol Lett. 2016;363(9).
165. Ceizel Borella G, Lagares A, Jr., Valverde C. Expression of the small regulatory RNA gene *mmgR* is regulated negatively by AniA and positively by NtrC in *Sinorhizobium meliloti* 2011. Microbiology (Reading). 2018;164(1):88-98.
166. Lagares A, Jr., Borella GC, Linne U, Becker A, Valverde C. Regulation of Polyhydroxybutyrate Accumulation in *Sinorhizobium meliloti* by the Trans-Encoded Small RNA MmgR. J Bacteriol. 2017;199(8).
167. Muriel-Millan LF, Castellanos M, Hernandez-Eligio JA, Moreno S, Espin G. Posttranscriptional regulation of PhbR, the transcriptional activator of polyhydroxybutyrate synthesis, by iron and the sRNA ArrF in *Azotobacter vinelandii*. Appl Microbiol Biotechnol. 2014;98(5):2173-82.
168. Klahn S, Schaal C, Georg J, Baumgartner D, Knippen G, Hagemann M, et al. The sRNA NsiR4 is involved in nitrogen assimilation control in

cyanobacteria by targeting glutamine synthetase inactivating factor IF7. Proc Natl Acad Sci U S A. 2015;112(45):E6243-52.
169. Shang L, Yan Y, Zhan Y, Ke X, Shao Y, Liu Y, et al. A regulatory network involving Rpo, Gac and Rsm for nitrogen-fixing biofilm formation by *Pseudomonas stutzeri*. NPJ Biofilms Microbiomes. 2021;7(1):54.
170. Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell. 2004;118(1):69-82.
171. Rutherford ST, van Kessel JC, Shao Y, Bassler BL. AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. Genes Dev. 2011;25(4):397-408.
172. Tu KC, Bassler BL. Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi*. Genes Dev. 2007;21(2):221-33.
173. Shao Y, Feng L, Rutherford ST, Papenfort K, Bassler BL. Functional determinants of the quorum-sensing non-coding RNAs and their roles in target regulation. EMBO J. 2013;32(15):2158-71.
174. Medina G, Juarez K, Diaz R, Soberon-Chavez G. Transcriptional regulation of *Pseudomonas aeruginosa rhlR*, encoding a quorum-sensing regulatory protein. Microbiology (Reading). 2003;149(Pt 11):3073-81.
175. Malgaonkar A, Nair M. Quorum sensing in *Pseudomonas aeruginosa* mediated by RhlR is regulated by a small RNA PhrD. Sci Rep. 2019;9(1):432.
176. Thomason MK, Voichek M, Dar D, Addis V, Fitzgerald D, Gottesman S, et al. A *rhlI* 5' UTR-Derived sRNA Regulates RhlR-Dependent Quorum Sensing in *Pseudomonas aeruginosa*. mBio. 2019;10(5).
177. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 1994;2:28-36.
178. Bailey TL, Noble WS. Searching for statistically significant regulatory modules. Bioinformatics. 2003;19 Suppl 2:ii16-25.
179. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6:26.
180. Schaper S, Wendt H, Bamberger J, Sieber V, Schmid J, Becker A. A Bifunctional UDP-Sugar 4-Epimerase Supports Biosynthesis of Multiple Cell Surface Polysaccharides in *Sinorhizobium meliloti*. J Bacteriol. 2019;201(10).
181. Le KY, Otto M. Quorum-sensing regulation in staphylococci-an overview. Front Microbiol. 2015;6:1174.
182. Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, Sturdevant DE, et al. RNAIII-independent target gene control by the *agr* quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. Mol Cell. 2008;32(1):150-8.
183. Novick RP, Projan SJ, Kornblum J, Ross HF, Ji G, Kreiswirth B, et al. The agr P2 operon: an autocatalytic sensory transduction system in *Staphylococcus aureus*. Mol Gen Genet. 1995;248(4):446-58.
184. Koenig RL, Ray JL, Maleki SJ, Smeltzer MS, Hurlburt BK. *Staphylococcus aureus* AgrA binding to the RNAIII-*agr* regulatory region. J Bacteriol. 2004;186(22):7549-55.

185. Ji G, Beavis RC, Novick RP. Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. Proc Natl Acad Sci U S A. 1995;92(26):12055-9.
186. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, et al. *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes Dev. 2007;21(11):1353-66.
187. Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S. Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J. 1993;12(10):3967-75.
188. Olson ME, Todd DA, Schaeffer CR, Paharik AE, Van Dyke MJ, Buttner H, et al. *Staphylococcus epidermidis agr* quorum-sensing system: signal identification, cross talk, and importance in colonization. J Bacteriol. 2014;196(19):3482-93.
189. Otto M, Echner H, Voelter W, Gotz F. Pheromone cross-inhibition between *Staphylococcus aureus* and *Staphylococcus epidermidis*. Infect Immun. 2001;69(3):1957-60.
190. Ji G, Beavis R, Novick RP. Bacterial interference caused by autoinducing peptide variants. Science. 1997;276(5321):2027-30.
191. Tsai CS, Winans SC. The quorum-hindered transcription factor YenR of *Yersinia enterocolitica* inhibits pheromone production and promotes motility via a small non-coding RNA. Mol Microbiol. 2011;80(2):556-71.
192. Hagberg KL, Price JP, Yurgel SN, Kahn ML. The *Sinorhizobium meliloti* Nitrogen Stress Response Changes Radically in the Face of Concurrent Phosphate Stress. Front Microbiol. 2022;13:800146.
193. Choudhary KS, Kleinmanns JA, Decker K, Sastry AV, Gao Y, Szubin R, et al. Elucidation of Regulatory Modes for Five Two-Component Systems in *Escherichia coli* Reveals Novel Relationships. mSystems. 2020;5(6).
194. Garcia-Tomsig NI, Garcia-Rodriguez FM, Guedes-Garcia SK, Millan V, Becker A, Robledo M, Jimenez-Zurdo JI. A double-negative feedback loop between NtrBC and a small RNA rewires nitrogen metabolism in legume symbionts. mBio. 2023:e0200323.
195. Prasse D, Schmitz RA. Small RNAs Involved in Regulation of Nitrogen Metabolism. Microbiol Spectr. 2018;6(4).
196. Acosta-Jurado S, Fuentes-Romero F, Ruiz-Sainz JE, Janczarek M, Vinardell JM. Rhizobial Exopolysaccharides: Genetic Regulation of Their Synthesis and Relevance in Symbiosis with Legumes. Int J Mol Sci. 2021;22(12).
197. Janczarek M, Skorupska A. Modulation of *rosR* expression and exopolysaccharide production in *Rhizobium leguminosarum* bv. *trifolii* by phosphate and clover root exudates. Int J Mol Sci. 2011;12(6):4132-55.
198. Mueller K, Gonzalez JE. Complex regulation of symbiotic functions is coordinated by MucR and quorum sensing in *Sinorhizobium meliloti*. J Bacteriol. 2011;193(2):485-96.
199. Verdonk CJ, Sullivan JT, Williman KM, Nicholson L, Bastholm TR, Hynes MF, et al. Delineation of the integrase-attachment and origin-of-transfer regions of the symbiosis island ICE*Ml*Sym$^{(R7A)}$. Plasmid. 2019;104:102416.
200. Sorokin V, Severinov K, Gelfand MS. Systematic prediction of control proteins and their DNA binding sites. Nucleic Acids Res. 2009;37(2):441-51.

201. Williams K, Savageau MA, Blumenthal RM. A bistable hysteretic switch in an activator-repressor regulated restriction-modification system. Nucleic Acids Res. 2013;41(12):6045-57.
202. Verdonk C. Investigating the structure and function of RdfS, a recombination directionality factor from an Integrative and Conjugative Element. Perth, Australia: University of Western Australia; 2023.
203. McClean KH, Winson MK, Fish L, Taylor A, Chhabra SR, Camara M, et al. Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of *N*-acylhomoserine lactones. Microbiology. 1997;143 (Pt 12):3703-11.
204. Gay P, Le Coq D, Steinmetz M, Berkelman T, Kado CI. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol. 1985;164(2):918-21.
205. Bharmal MM, Gega A, Schrader JM. A combination of mRNA features influence the efficiency of leaderless mRNA translation initiation. NAR Genom Bioinform. 2021;3(3):lqab081.
206. Leiva LE, Katz A. Regulation of Leaderless mRNA Translation in Bacteria. Microorganisms. 2022;10(4).
207. Gerdt JP, Wittenwyler DM, Combs JB, Boursier ME, Brummond JW, Xu H, Blackwell HE. Chemical Interrogation of LuxR-type Quorum Sensing Receptors Reveals New Insights into Receptor Selectivity and the Potential for Interspecies Bacterial Signaling. ACS Chem Biol. 2017;12(9):2457-64.
208. Mattmann ME, Geske GD, Worzalla GA, Chandler JR, Sappington KJ, Greenberg EP, Blackwell HE. Synthetic ligands that activate and inhibit a quorum-sensing regulator in *Pseudomonas aeruginosa*. Bioorg Med Chem Lett. 2008;18(10):3072-5.
209. Mattmann ME, Shipway PM, Heth NJ, Blackwell HE. Potent and selective synthetic modulators of a quorum sensing repressor in *Pseudomonas aeruginosa* identified from second-generation libraries of *N*-acylated *L*-homoserine lactones. Chembiochem. 2011;12(6):942-9.
210. Byers JT, Lucas C, Salmond GP, Welch M. Nonenzymatic turnover of an *Erwinia carotovora* quorum-sensing signaling molecule. J Bacteriol. 2002;184(4):1163-71.
211. Yates EA, Philipp B, Buckley C, Atkinson S, Chhabra SR, Sockett RE, et al. *N*-acylhomoserine lactones undergo lactonolysis in a pH-, temperature-, and acyl chain length-dependent manner during growth of *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa*. Infect Immun. 2002;70(10):5635-46.
212. Gao A, Mei GY, Liu S, Wang P, Tang Q, Liu YP, et al. High-resolution structures of AidH complexes provide insights into a novel catalytic mechanism for *N*-acyl homoserine lactonase. Acta Crystallogr D Biol Crystallogr. 2013;69(Pt 1):82-91.
213. Kaufmann GF, Sartorio R, Lee SH, Rogers CJ, Meijler MM, Moss JA, et al. Revisiting quorum sensing: Discovery of additional chemical and biological functions for 3-oxo-*N*-acylhomoserine lactones. Proc Natl Acad Sci U S A. 2005;102(2):309-14.
214. Dong YH, Xu JL, Li XZ, Zhang LH. AiiA, an enzyme that inactivates the acylhomoserine lactone quorum-sensing signal and attenuates the virulence of *Erwinia carotovora*. Proc Natl Acad Sci U S A. 2000;97(7):3526-31.

215. Czajkowski R, Krzyzanowska D, Karczewska J, Atkinson S, Przysowa J, Lojkowska E, et al. Inactivation of AHLs by *Ochrobactrum sp.* A44 depends on the activity of a novel class of AHL acylase. Environ Microbiol Rep. 2011;3(1):59-68.
216. Xia R, Yang Y, Pan X, Gao C, Yao Y, Liu X, et al. Improving the production of AHL lactonase AiiO-AIO6 from *Ochrobactrum sp.* M231 in intracellular protease-deficient *Bacillus subtilis*. AMB Express. 2020;10(1):138.
217. Haskett TL, Terpolilli JJ, Bekuma A, O'Hara GW, Sullivan JT, Wang P, et al. Assembly and transfer of tripartite integrative and conjugative genetic elements. Proc Natl Acad Sci U S A. 2016;113(43):12268-73.
218. Dong YH, Wang LH, Xu JL, Zhang HB, Zhang XF, Zhang LH. Quenching quorum-sensing-dependent bacterial infection by an *N*-acyl homoserine lactonase. Nature. 2001;411(6839):813-7.
219. Neville N, Roberge N, Jia Z. Polyphosphate Kinase 2 (PPK2) Enzymes: Structure, Function, and Roles in Bacterial Physiology and Virulence. Int J Mol Sci. 2022;23(2).
220. Nocek B, Kochinyan S, Proudfoot M, Brown G, Evdokimova E, Osipiuk J, et al. Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria. Proc Natl Acad Sci U S A. 2008;105(46):17730-5.
221. Frank C, Teleki A, Jendrossek D. Characterization of *Agrobacterium tumefaciens* PPKs reveals the formation of oligophosphorylated products up to nucleoside nona-phosphates. Appl Microbiol Biotechnol. 2020;104(22):9683-92.
222. Rice CD, Pollard JE, Lewis ZT, McCleary WR. Employment of a promoter-swapping technique shows that PhoU modulates the activity of the PstSCAB2 ABC transporter in *Escherichia coli*. Appl Environ Microbiol. 2009;75(3):573-82.
223. Carmany DO, Hollingsworth K, McCleary WR. Genetic and biochemical studies of phosphatase activity of PhoR. J Bacteriol. 2003;185(3):1112-5.
224. Muda M, Rao NN, Torriani A. Role of PhoU in phosphate transport and alkaline phosphatase regulation. J Bacteriol. 1992;174(24):8057-64.
225. Aguena M, Ferreira GM, Spira B. Stability of the pstS transcript of *Escherichia coli*. Arch Microbiol. 2009;191(2):105-12.
226. Wanner BL, Chang BD. The *phoBR* operon in Escherichia coli K-12. J Bacteriol. 1987;169(12):5569-74.
227. Santos-Beneit F. The Pho regulon: a huge regulatory network in bacteria. Front Microbiol. 2015;6:402.
228. Gardner SG, McCleary WR. Control of the *phoBR* Regulon in *Escherichia coli*. EcoSal Plus. 2019;8(2).
229. Lubin EA, Henry JT, Fiebig A, Crosson S, Laub MT. Identification of the PhoB Regulon and Role of PhoU in the Phosphate Starvation Response of Caulobacter crescentus. J Bacteriol. 2016;198(1):187-200.
230. Jowsey WJ, Morris CRP, Hall DA, Sullivan JT, Fagerlund RD, Eto KY, et al. DUF2285 is a novel helix-turn-helix domain variant that orchestrates both activation and antiactivation of conjugative element transfer in proteobacteria. Nucleic Acids Res. 2023;51(13):6841-56.

231. O'Donnell SM, Janssen GR. Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in *Escherichia coli*. J Bacteriol. 2002;184(23):6730-3.
232. Udagawa T, Shimizu Y, Ueda T. Evidence for the translation initiation of leaderless mRNAs by the intact 70 S ribosome without its dissociation into subunits in eubacteria. J Biol Chem. 2004;279(10):8539-46.
233. Bouvier M, Sharma CM, Mika F, Nierhaus KH, Vogel J. Small RNA binding to 5' mRNA coding region inhibits translational initiation. Mol Cell. 2008;32(6):827-37.

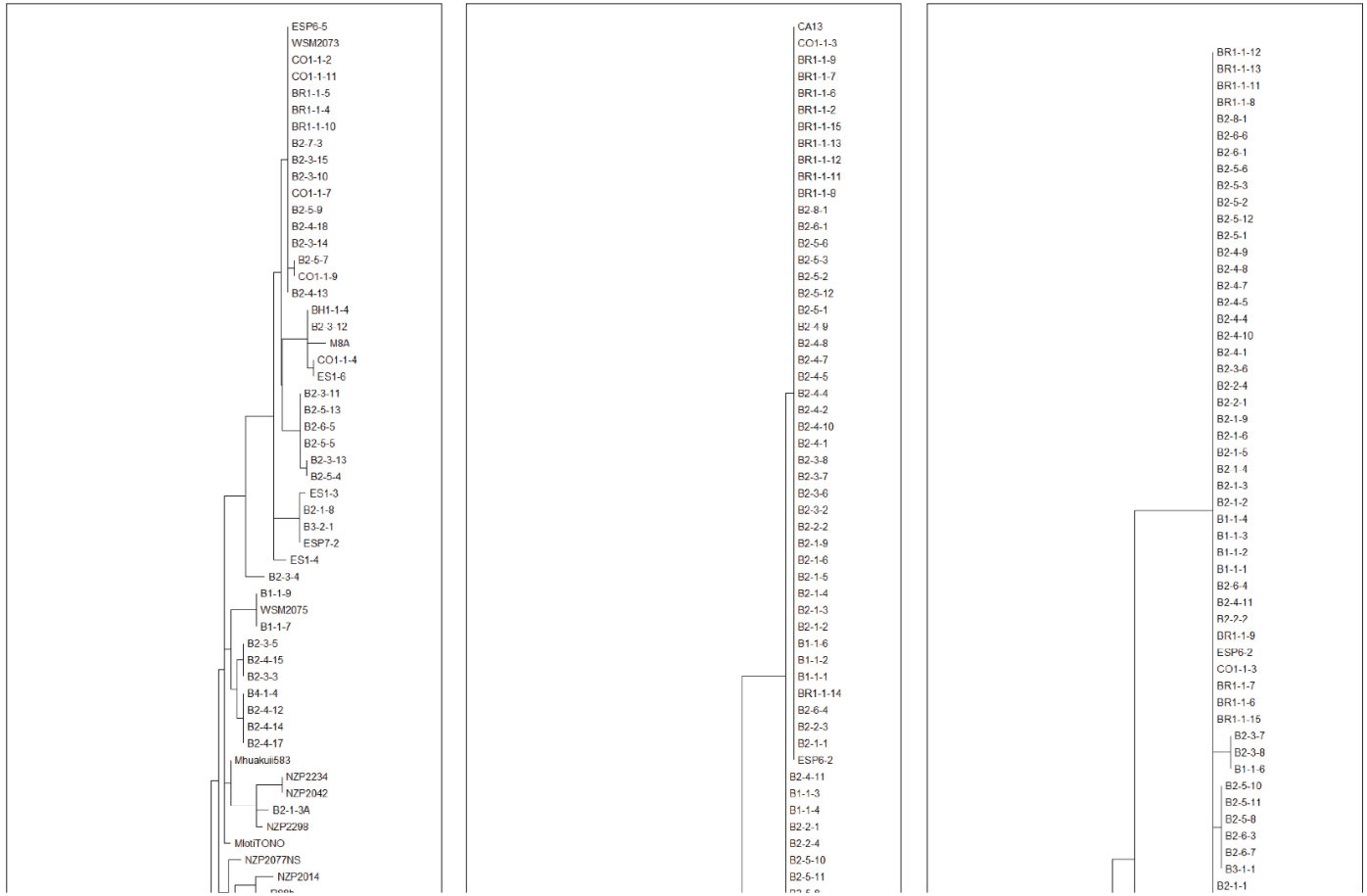
Appendix 1

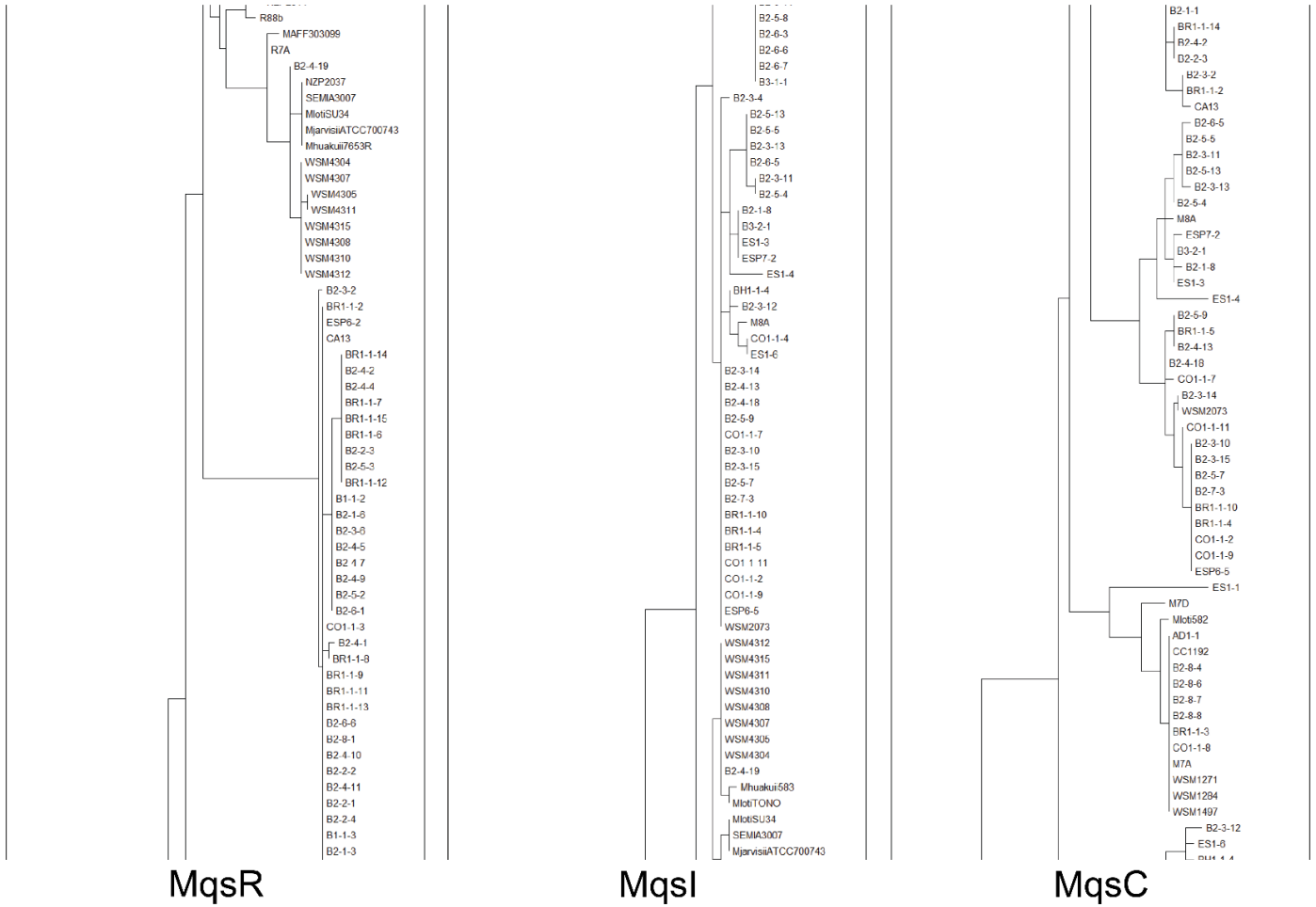
Buffer and solution recipes

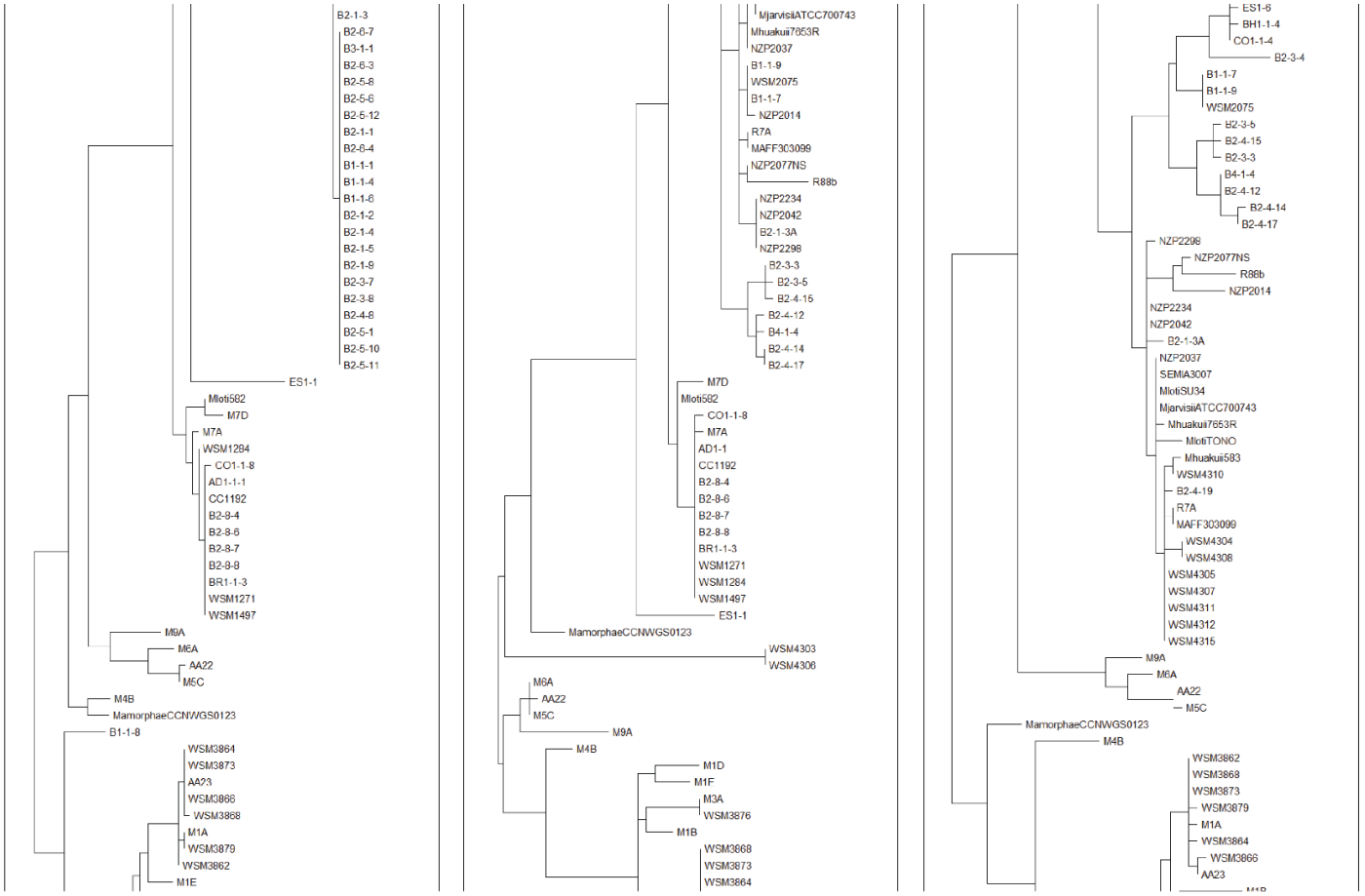
β-galactosidase assays	
Phosphate-buffered saline (PBS)	8 g/L NaCl, 1.44 g/L Na_2HPO_4, 0.2 g/L KCl, 0.24 g/L KH_2PO_4, pH adjusted to 7.2 with HCl, made up to volume with H_2O and autoclaved.
200x Stock solution	50 mg/mL 4-Methylumbelliferyl β-D-galactopyranoside (Sigma-Aldrich) dissolved in DMSO
Working solution	1x Stock solution diluted in PBS buffer, 2 mg/mL lysozyme (Sigma-Aldrich)
MqsC protein purification	
Protein wash buffer	50 mM NaH_2PO_4, 300 mM NaCl, 50 mM imidazole, 1mM DTT, 5% (v/v) glycerol, pH 8.0
Protein elution buffer	50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, 1mM DTT, 5% (v/v) glycerol, pH 8.0
SEC buffer	50 mM NaH_2PO_4, 150 mM NaCl, 1 mM DTT, pH 8.0
AidM and AidM$_{S102G}$ protein purification	
Protein wash buffer	50 mM NaH_2PO_4, 400 mM NaCl, 30 mM imidazole, 5% (v/v) glycerol, pH 7.0
Protein elution buffer	50 mM NaH_2PO_4, 400 mM NaCl, 300 mM imidazole, 5% (v/v) glycerol, pH 7.0
SEC buffer	50 mM NaH_2PO_4, 200 mM NaCl, pH 7.0
SDS-PAGE gels	
4x SDS loading buffer	50 mM Tris (pH 6.8), 100 mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 30% (v/v) glycerol
SDS running buffer	25 mM Tris (pH 8.3), 192 mM glycine, 0.1% (w/v) SDS
Glasshouse experiments	
Nutrient solution	12.3 g/L $MgSO_4•7H_2O$, 6.8 g/L KH_2PO_4, 17.5 g/L K_2SO_4, 2.5 g/L Fe-EDTA, 0.464 g/L H_3BO_3, 0.018 g/L $Na_2MoO_4•2H_2O$, 0.539 g/L $ZnSO_4•7H_2O$, 0.042 g/L $MnSO_4•4H_2O$, 0.141 g/L $CoSO_4•7H_2O$, 0.125 g/L $CuSO_4•5H_2O$, 2.04 g/L $CaSO_4$.

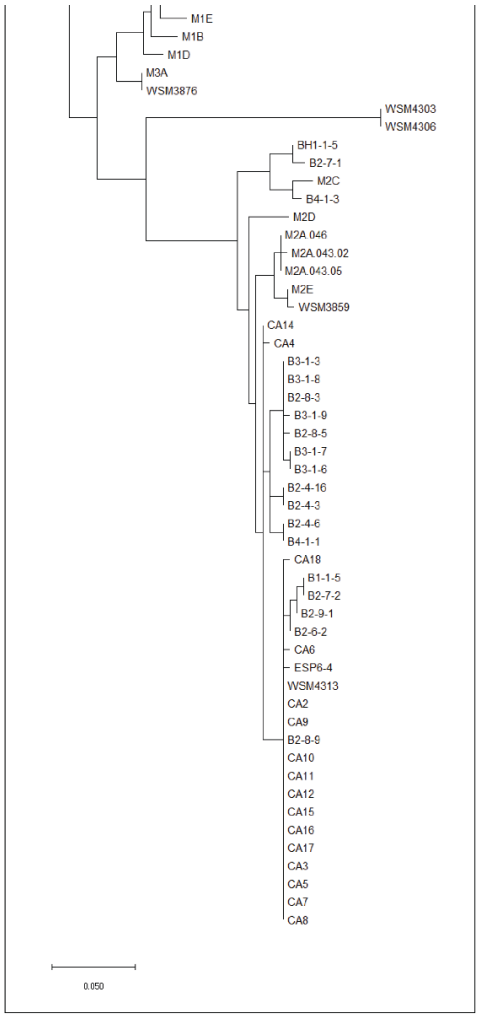
Appendix 2

Maximum-likelihood phylogenetic trees of 207 MqsR, 207 MqsI and 205 MqsC amino acid sequences.

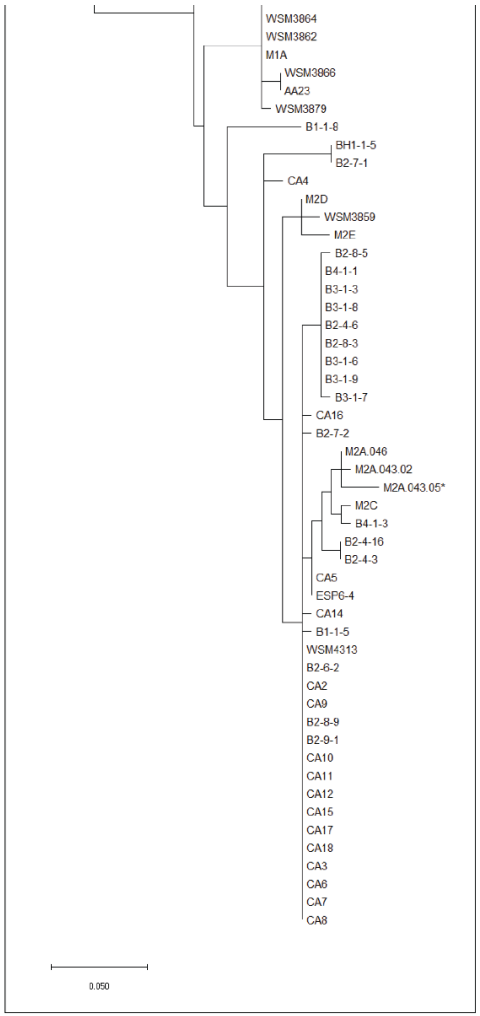



R88b
MAFF303099
R7A
B2-4-19
NZP2037
SEMIA3007
MlotiSU34
MjarvisiiATCC700743
Mhuakui7653R
WSM4304
WSM4307
WSM4305
WSM4311
WSM4315
WSM4308
WSM4310
WSM4312
B2-3-2
BR1-1-2
ESP6-2
CA13
BR1-1-14
B2-4-2
B2-4-4
BR1-1-7
BR1-1-15
BR1-1-6
B2-2-3
B2-5-3
BR1-1-12
B1-1-2
B2-1-8
B2-3-8
B2-4-5
B2-4-7
B2-4-9
B2-5-2
B2-6-1
CO1-1-3
B2-4-1
BR1-1-8
BR1-1-9
BR1-1-11
BR1-1-13
B2-6-6
B2-8-1
B2-4-10
B2-2-2
B2-4-11
B2-2-1
B2-2-4
B1-1-3
B2-1-3
MqsR
B2-5-8
B2-6-3
B2-6-6
B2-6-7
B3-1-1
B2-3-4
B2-5-13
B2-6-5
B2-3-13
B2-6-5
B2-3-11
B2-5-4
B2-1-8
B3-2-1
ES1-3
ESP7-2
ES1-4
BH1-1-4
B2-3-12
M8A
CO1-1-4
ES1-6
B2-3-14
B2-4-13
B2-4-18
B2-5-9
CO1-1-7
B2-3-10
B2-3-15
B2-5-7
B2-7-3
BR1-1-10
BR1-1-4
BR1-1-5
CO1-1-11
CO1-1-2
CO1-1-9
ESP6-5
WSM2073
WSM4312
WSM4315
WSM4311
WSM4310
WSM4308
WSM4307
WSM4305
WSM4304
B2-4-19
Mhuakui583
MlotiTONO
MlotiSU34
SEMIA3007
MjarvisiiATCC700743
MqsI
B2-1-1
BR1-1-14
R2-4-2
D2-2-3
B2-3-2
BR1-1-2
CA13
B2-6-5
B2-5-5
B2-3-11
B2-5-13
B2-3-13
B2-5-4
M8A
ESP7-2
B3-2-1
B2-1-8
ES1-3
ES1-4
B2-5-9
BR1-1-5
B2-4-13
B2-4-18
CO1-1-7
B2-3-14
WSM2073
CO1-1-11
B2-3-10
B2-3-15
B2-5-7
B2-7-3
BR1-1-10
BR1-1-4
CO1-1-2
CO1-1-9
ESP6-5
ES1-1
M7D
Mloti582
AD1-1
CC1192
B2-8-4
B2-8-6
B2-8-7
B2-8-8
BR1-1-3
CO1-1-8
M7A
WSM1271
WSM1284
WSM1497
B2-3-12
ES1-6
MqsC

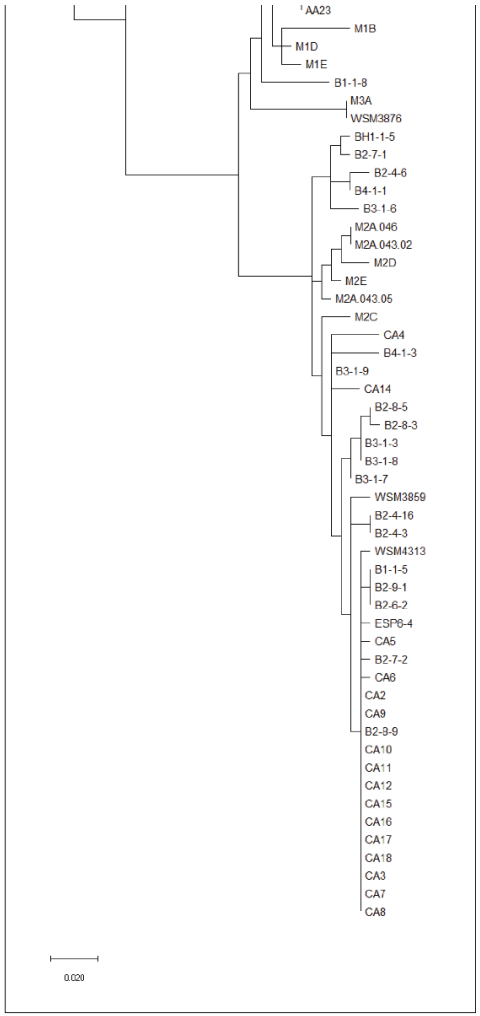

B2-1-3
B2-6-7
B3-1-1
B2-6-3
B2-5-8
B2-5-6
B2-5-12
B2-1-1
B2-6-4
B1-1-1
B1-1-4
B1-1-6
B2-1-2
B2-1-4
B2-1-5
B2-1-9
B2-3-7
B2-3-8
B2-4-8
B2-5-1
B2-5-10
B2-5-11
ES1-1
Mloti582
M7D
M7A
WSM1284
CO1-1-8
AD1-1-1
CC1192
B2-8-4
B2-8-6
B2-8-7
B2-8-8
BR1-1-3
WSM1271
WSM1497
M9A
M6A
AA22
M5C
M4B
MamorphaeCCNWGS0123
B1-1-8
WSM3864
WSM3873
AA23
WSM3866
WSM3868
M1A
WSM3879
WSM3862
M1E
MqsR
MjarvisiiATCC700743
Mhuakui7653R
NZP2037
B1-1-9
WSM2075
B1-1-7
NZP2014
R7A
MAFF303099
NZP2077NS
R88b
NZP2234
NZP2042
B2-1-3A
NZP2298
B2-3-3
B2-3-5
B2-4-15
B2-4-12
B4-1-4
B2-4-14
B2-4-17
M7D
Mloti582
CO1-1-8
M7A
AD1-1
CC1192
B2-8-4
B2-8-6
B2-8-7
B2-8-8
BR1-1-3
WSM1271
WSM1284
WSM1497
ES1-1
MamorphaeCCNWGS0123
WSM4303
WSM4306
M6A
AA22
M5C
M9A
M4B
M1D
M1F
M3A
WSM3876
M1B
WSM3868
WSM3873
WSM3864
MqsI
ES1-6
BH1-1-4
CO1-1-4
B2-3-4
B1-1-7
B1-1-9
WSM2075
B2-3-6
B2-4-15
B2-3-3
B4-1-4
B2-4-12
B2-4-14
B2-4-17
NZP2298
NZP2077NS
R88b
NZP2014
NZP2234
NZP2042
B2-1-3A
NZP2037
SEMIA3007
MlotiSU34
MjarvisiiATCC700743
Mhuakui7653R
MlotiTONO
Mhuakui583
WSM4310
B2-4-19
R7A
MAFF303099
WSM4304
WSM4308
WSM4305
WSM4307
WSM4311
WSM4312
WSM4315
M9A
M6A
AA22
M5C
MamorphaeCCNWGS0123
M4B
WSM3862
WSM3868
WSM3873
WSM3879
M1A
WSM3864
WSM3866
AA23
MqsC



MqsR



MqsI



MqsC

Appendix 3

Plasmid map of pSDz, commonly used throughout this thesis for promoter-fusion assays.

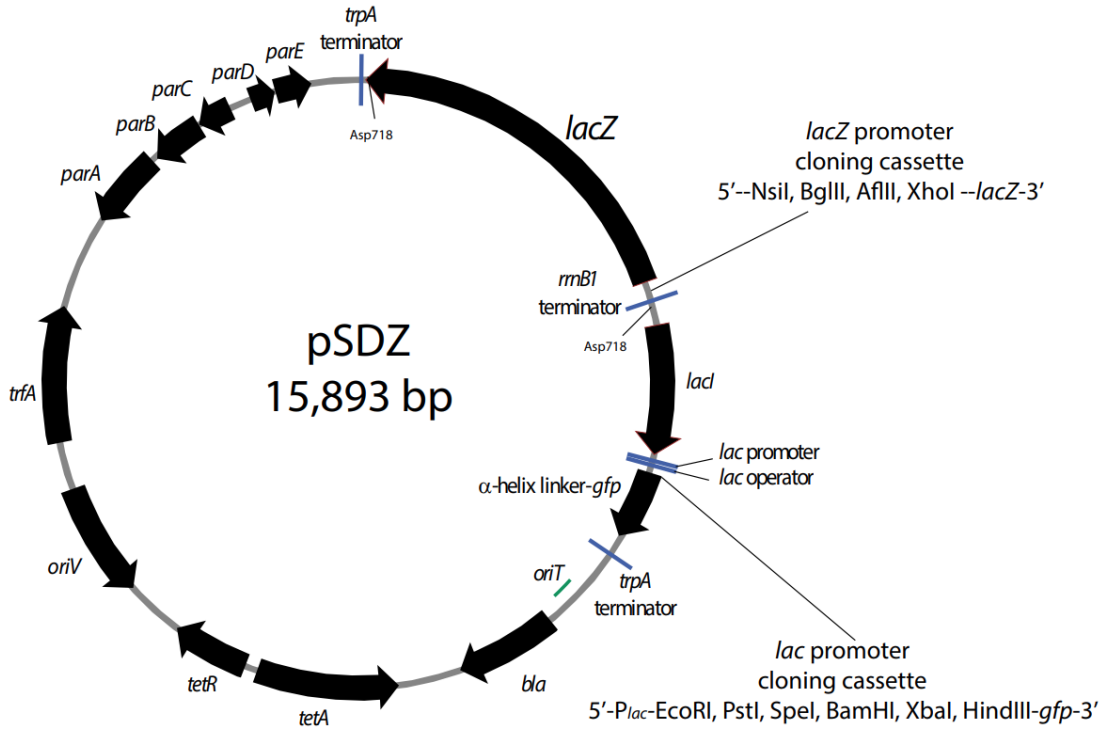


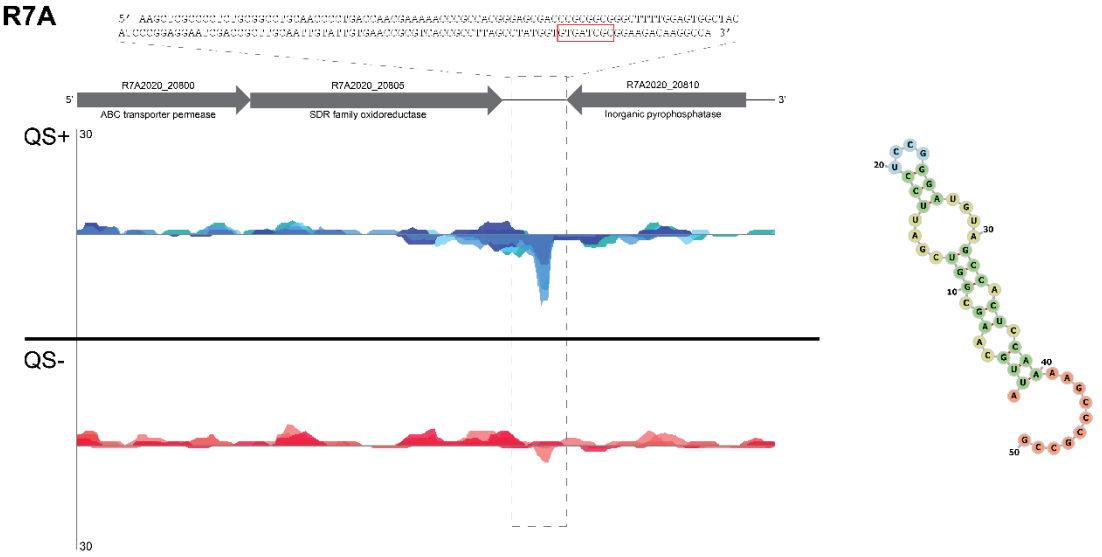
Figure from reference (83).

Appendix 4

Mapped RNAseq reads (wild type in blue and *mqsRIC* mutant in red), genomic context and predicted structure for each mqsRNA described in Chapter 5.

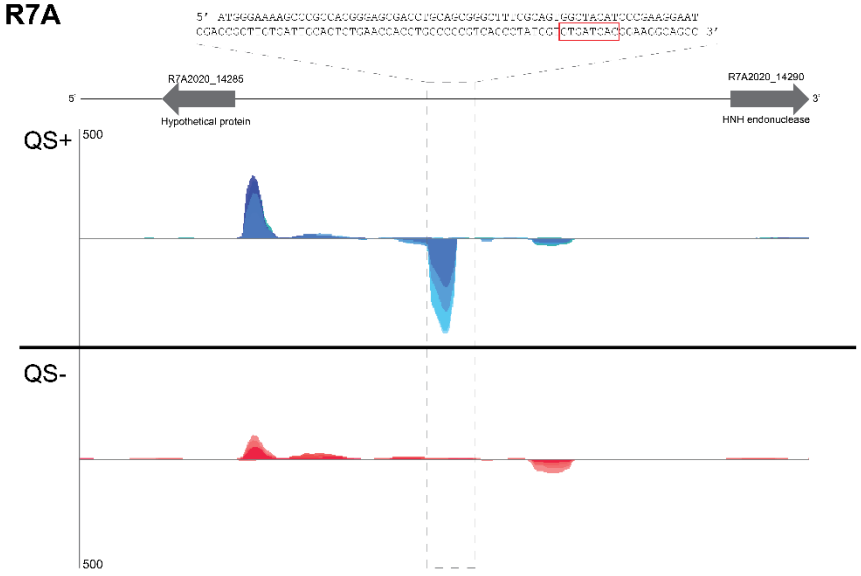
mqsRNA1

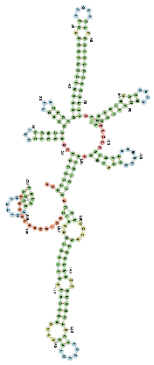
R7A



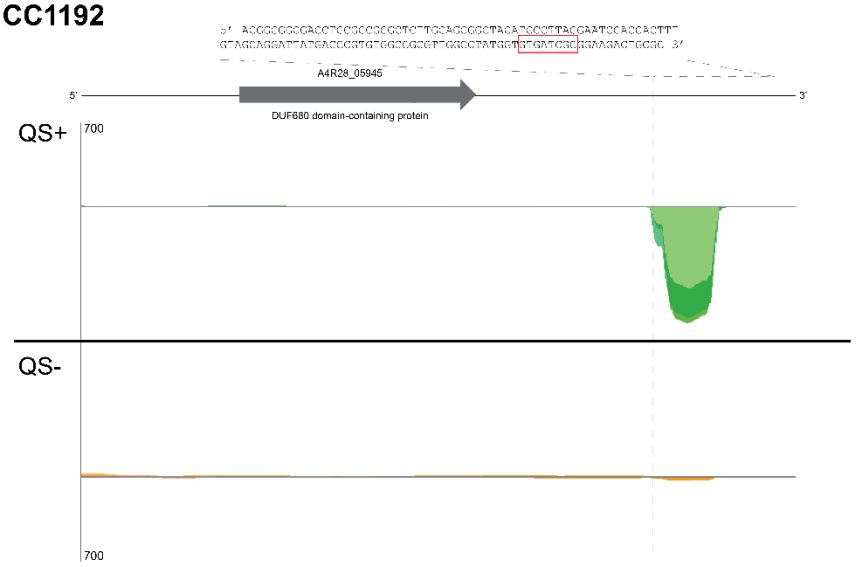
mqsRNA2

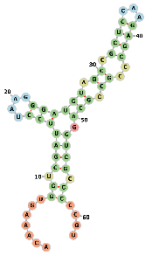
R7A





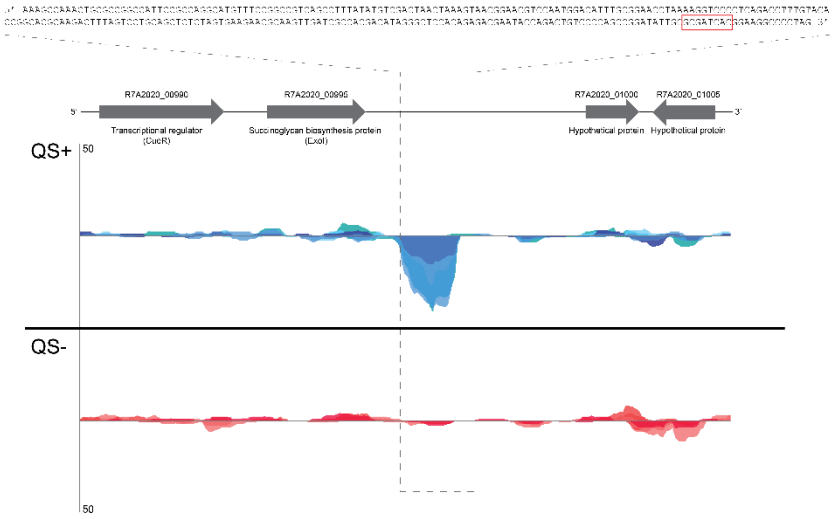
CC1192

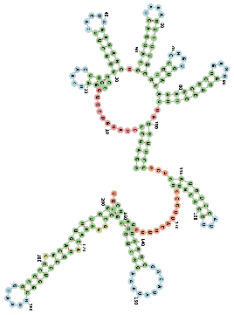




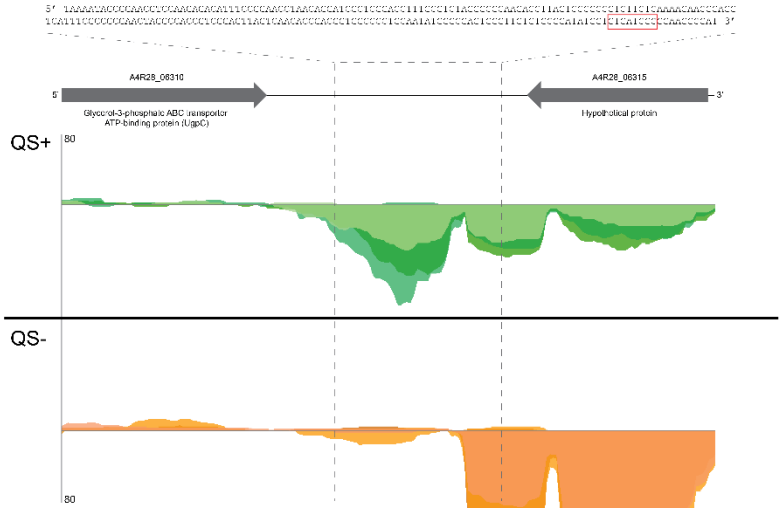
mqsRNA3

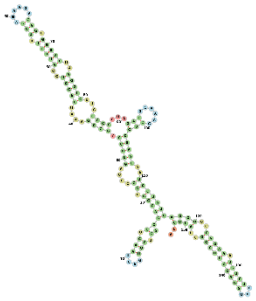
R7A





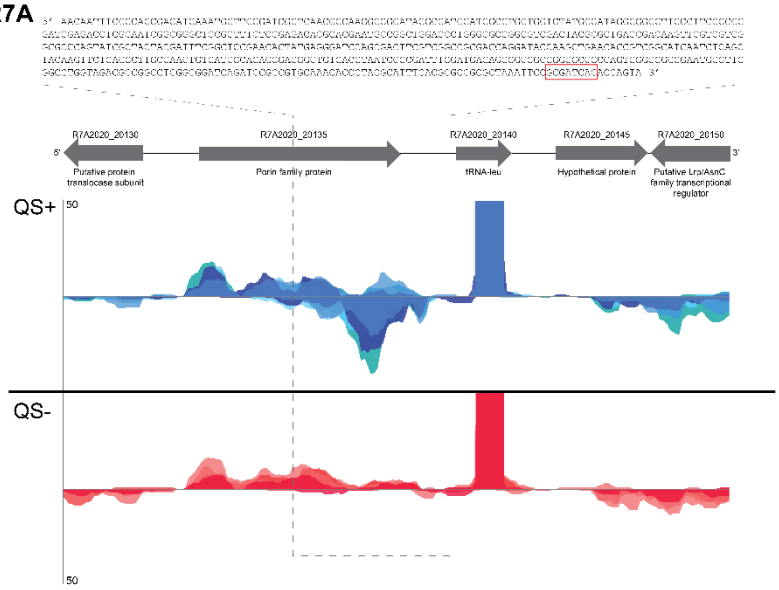
CC1192

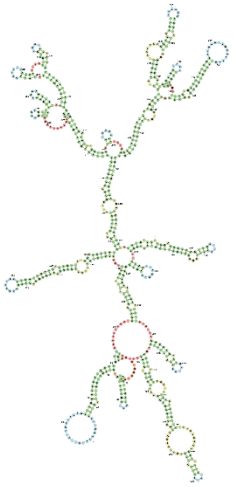




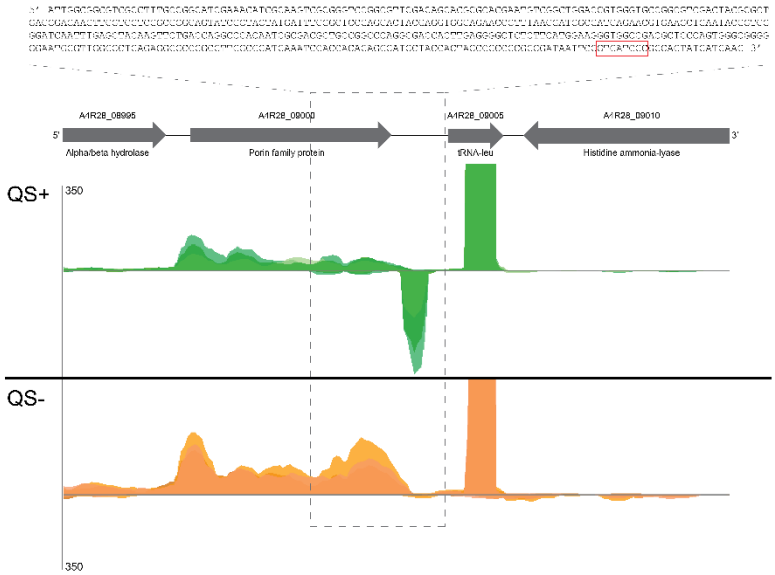
mqsRNA4

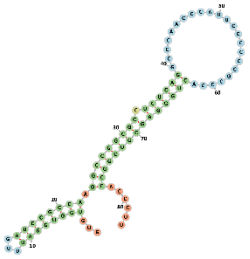
R7A





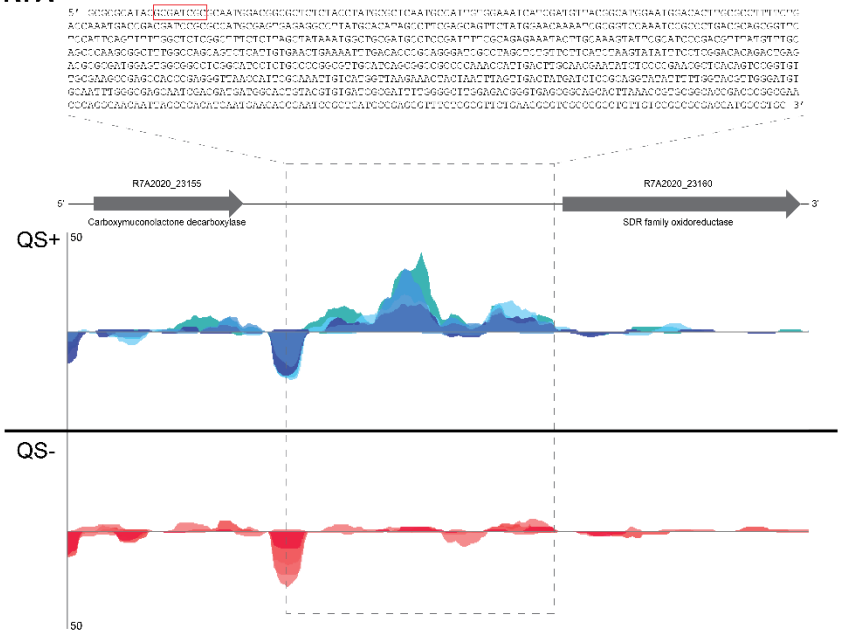
CC1192

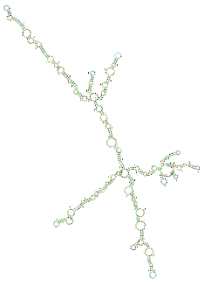




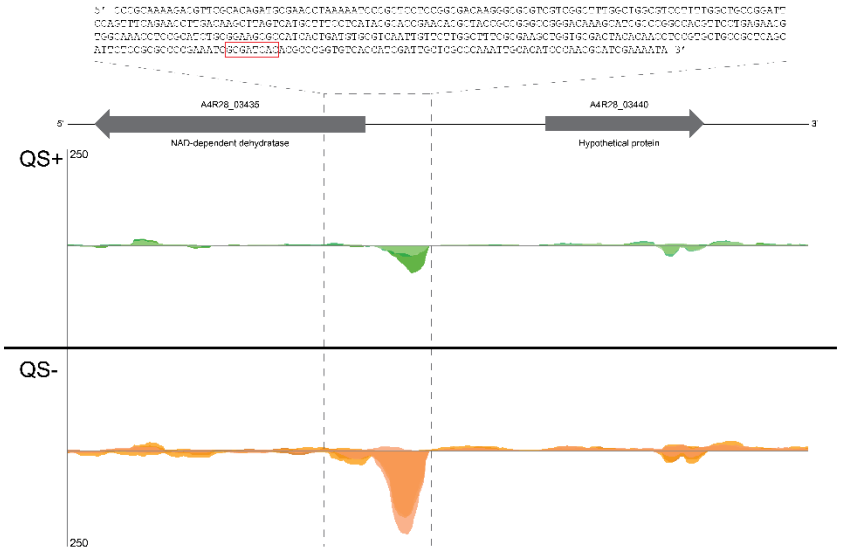
mqsRNA5

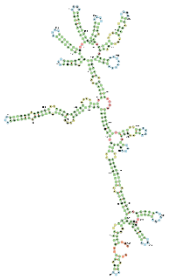
R7A





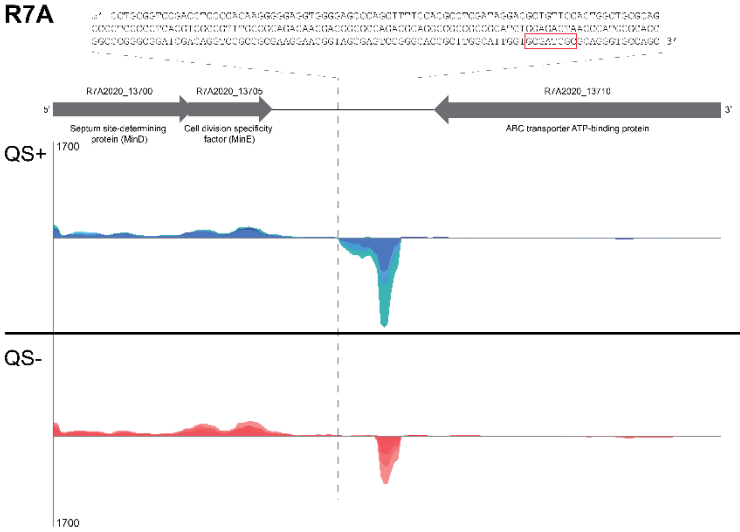
CC1192

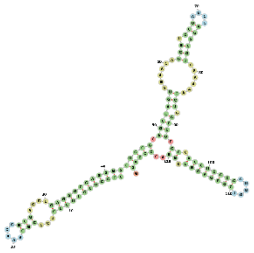




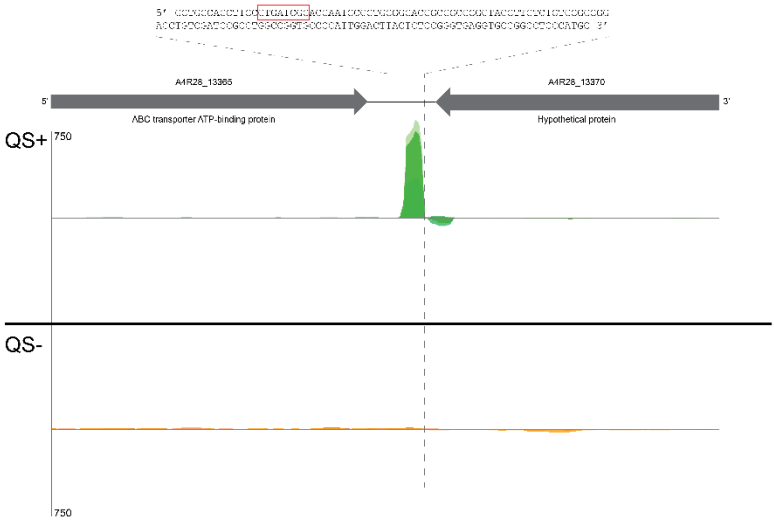
mqsRNA6

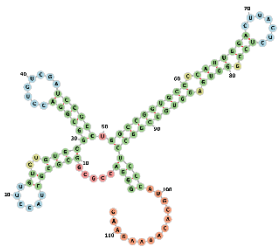
R7A





CC1192





Appendix 5

Differentially expressed genes in R7A wild type compared to R7AΔ*mqsRIC*

Locus Tag	Gene product	Log2 Fold change	Adjusted p value
R7A2020_23705	porin	-3.83313	9.63E-18
R7A2020_13690	aldo/keto reductase	-1.80818	0.000573
R7A2020_30725	hypothetical protein	-1.54846	0.007371
R7A2020_30720	carboxylating nicotinate-nucleotide diphosphorylase	-1.53221	0.011108
R7A2020_30750	beta-ketoacyl-ACP synthase III	-1.44192	0.026619
R7A2020_30730	biotin synthase BioB	-1.3826	0.004571
R7A2020_30710	quinolinate synthase NadA	-1.38257	0.043127
R7A2020_11080	cold-shock protein	-1.38195	0.014702
R7A2020_22170	chaperonin GroEL	-1.35806	0.004373
R7A2020_01045	ABC transporter substrate-binding protein	-1.35547	0.005348
R7A2020_22110	substrate-binding domain-containing protein	-1.31067	2.90E-05
R7A2020_22165	co-chaperone GroES	-1.2995	0.00495
R7A2020_19620	30S ribosomal protein S7	-1.28598	0.026028
R7A2020_30715	L-aspartate oxidase	-1.2833	0.040057
R7A2020_23295	30S ribosomal protein S18	-1.23881	0.020092
R7A2020_00135	helix-turn-helix transcriptional regulator	-1.21946	0.000129
R7A2020_30765	co-chaperone GroES	-1.13594	0.011014
R7A2020_06725	peroxiredoxin	-1.13578	1.81E-06
R7A2020_15565	precorrin-6A synthase (deacetylating)	-1.1032	0.048461
R7A2020_19530	30S ribosomal protein S8	-1.09539	0.046744
R7A2020_17605	DUF3572 domain-containing protein	-1.07835	0.049636
R7A2020_30770	chaperonin GroEL	-1.03187	0.00115
R7A2020_31230	sugar ABC transporter substrate-binding protein	-1.01401	1.72E-05
R7A2020_15535	molybdopterin-binding protein	-0.95159	0.005565
R7A2020_23485	VOC family protein	-0.94262	0.002076
R7A2020_15770	BA14K family protein	-0.94056	9.19E-05
R7A2020_04730	cupin domain-containing protein	-0.93536	0.002267
R7A2020_06265	amino acid ABC transporter substrate-binding protein	-0.93067	0.017147

R7A2020_16230	nitroreductase	-0.90785	0.02158
R7A2020_19670	50S ribosomal protein L1	-0.89251	0.018742
R7A2020_23720	pyridoxal phosphate-dependent aminotransferase	-0.86005	0.000737
R7A2020_06840	D-xylose ABC transporter substrate-binding protein	-0.83008	0.006198
R7A2020_06730	MBL fold metallo-hydrolase	-0.8224	2.21E-05
R7A2020_09900	PotD/PotF family extracellular solute-binding protein	-0.82084	0.028853
R7A2020_14640	sugar-binding protein	-0.81347	0.00344
R7A2020_14920	ABC transporter substrate-binding protein	-0.78242	0.002597
R7A2020_01930	PTS sugar transporter subunit IIA	-0.77472	0.039239
R7A2020_24345	sugar ABC transporter substrate-binding protein	-0.7591	0.002262
R7A2020_25505	substrate-binding domain-containing protein	-0.75136	2.10E-05
R7A2020_12955	substrate-binding domain-containing protein	-0.731	0.012815
R7A2020_05870	DUF2312 domain-containing protein	-0.72307	0.047616
R7A2020_14965	energy-dependent translational throttle protein EttA	-0.70836	0.024511
R7A2020_31050	rhamnose ABC transporter substrate-binding protein	-0.70636	0.007371
R7A2020_28665	translation initiation factor IF-1	-0.70515	0.044372
R7A2020_23005	MucR family transcriptional regulator	-0.70171	0.00145
R7A2020_28520	type II and III secretion system protein family protein	-0.6851	0.038103
R7A2020_31375	acyl-CoA dehydrogenase	-0.62551	0.02672
R7A2020_09520	ABC transporter substrate-binding protein	-0.61996	0.018115
R7A2020_23350	amidophosphoribosyltransferase	-0.60885	0.012891
R7A2020_21030	sugar ABC transporter substrate-binding protein	-0.59982	0.001959
R7A2020_01850	ABC transporter substrate-binding protein	-0.55559	0.024256
R7A2020_18480	acyl-ACP--UDP-N-acetylglucosamine O-acyltransferase	-0.50981	0.037868
R7A2020_06185	phosphoglucosamine mutase	0.493913	0.0428
R7A2020_01265	sugar transferase	0.511255	0.014702
R7A2020_01160	glycosyltransferase	0.544839	0.048461
R7A2020_27700	AraC family transcriptional regulator	0.572325	0.027669
R7A2020_23725	MarR family winged helix-turn-helix transcriptional regulator	0.605949	0.005348
R7A2020_01180	lipopolysaccharide biosynthesis protein	0.606968	0.017954
R7A2020_29280	Mov34/MPN/PAD-1 family protein	0.642483	0.005002
R7A2020_03105	hypothetical protein	0.669444	0.035974

R7A2020_20880	nucleotidyltransferase	0.672797	0.028478
R7A2020_23875	bifunctional [glutamine synthetase] adenylyltransferase/[glutamine synthetase]-adenylyl-L-tyrosine phosphorylase	0.675499	0.026619
R7A2020_27760	response regulator	0.689332	0.026619
R7A2020_21620	exopolyphosphatase	0.69825	0.047922
R7A2020_20215	PilZ domain-containing protein	0.698499	0.044212
R7A2020_26905	DNA-3-methyladenine glycosylase I	0.701504	0.049219
R7A2020_01195	hypothetical protein	0.707424	0.02672
R7A2020_05995	DUF3857 domain-containing transglutaminase family protein	0.708081	0.011014
R7A2020_20920	His-Xaa-Ser system radical SAM maturase HxsC	0.723292	0.049219
R7A2020_21345	ABC transporter substrate-binding protein	0.72354	0.001554
R7A2020_29635	DUF2285 domain-containing protein	0.753915	0.032781
R7A2020_29275	hypothetical protein	0.7598	0.000306
R7A2020_05935	YmdB family metallophosphoesterase	0.78251	0.006198
R7A2020_19420	DUF2735 domain-containing protein	0.793285	0.021063
R7A2020_27165	SURF1 family protein	0.812279	0.015237
R7A2020_11765	PadR family transcriptional regulator	0.81667	0.036474
R7A2020_13185	exopolysaccharide transport family protein	0.817333	0.001147
R7A2020_04900	LTA synthase family protein	0.825453	0.002076
R7A2020_28755	hypothetical protein	0.835461	0.0056
R7A2020_27755	YaiO family outer membrane beta-barrel protein	0.836373	0.001887
R7A2020_17255	hypothetical protein	0.840971	0.010981
R7A2020_04205	hypothetical protein	0.86328	0.001633
R7A2020_15405	glucose/quinate/shikimate family membrane-bound PQQ-dependent dehydrogenase	0.86638	0.000735
R7A2020_14400	glycosyltransferase family 39 protein	0.876537	0.042791
R7A2020_17245	DUF982 domain-containing protein	0.878019	0.002267
R7A2020_21675	DUF58 domain-containing protein	0.87952	0.028924
R7A2020_13190	VpsF family polysaccharide biosynthesis protein	0.879649	0.005958
R7A2020_20875	SAVED domain-containing protein	0.891615	0.005736
R7A2020_17595	hypothetical protein	0.89673	0.000209
R7A2020_20945	DUF4238 domain-containing protein	0.94037	0.007389
R7A2020_00435	DUF1800 family protein	0.948007	0.007389

R7A2020_10535	hypothetical protein Tetratricopeptide (TPR) repeat C-terminus	0.954267	6.38E-07
R7A2020_27735	DNA ligase	0.960889	0.006198
R7A2020_27745	glycosyltransferase family 2 protein	0.967549	0.000617
R7A2020_22140	helix-turn-helix domain-containing GNAT family N-acetyltransferase	0.976657	0.036474
R7A2020_27750	HEAT repeat domain-containing protein	0.99558	0.011108
R7A2020_10420	tRNA-Met	1.012087	0.026619
R7A2020_10280	ABC transporter substrate-binding protein	1.030728	0.007371
R7A2020_17785	hypothetical protein	1.042722	0.010538
R7A2020_11025	hypothetical protein	1.056238	0.000974
R7A2020_16135	bifunctional adenosylcobinamide kinase/adenosylcobinamide-phosphate guanylyltransferase	1.078764	0.00401
R7A2020_04780	DUF2867 domain-containing protein	1.129541	3.05E-06
R7A2020_03530	hypothetical protein	1.130399	0.03061
R7A2020_22710	hypothetical protein	1.137531	0.003935
R7A2020_21665	DUF4350 domain-containing protein	1.144392	0.03061
R7A2020_27705	EAL domain-containing protein	1.187269	6.80E-10
R7A2020_03365	YifB family Mg chelatase-like AAA ATPase	1.191895	0.001753
R7A2020_27625	helix-turn-helix domain-containing protein	1.212183	2.06E-06
R7A2020_27595	basic amino acid/polyamine antiporter	1.221985	0.010981
R7A2020_15615	GGDEF domain-containing protein	1.276136	0.015169
R7A2020_05760	nucleoside diphosphate kinase regulator	1.276426	0.036331
R7A2020_04895	hypothetical protein	1.288479	0.001753
R7A2020_19700	hypothetical protein crotonase superfamily domain COG3904	1.303665	3.86E-07
R7A2020_22235	cold-shock protein	1.307411	0.001022
R7A2020_22340	hypothetical protein	1.341948	0.000144
R7A2020_04475	DUF1236 domain-containing protein	1.377472	0.001952
R7A2020_20970	hypothetical protein	1.41094	1.01E-05
R7A2020_20545	multidrug effflux MFS transporter	1.443892	2.01E-05
R7A2020_09305	hypothetical protein	1.449364	0.003935
R7A2020_00420	hypothetical protein	1.458781	0.005154
R7A2020_06010	hypothetical protein	1.467423	0.018329
R7A2020_21340	ABC transporter permease	1.48424	2.61E-07

R7A2020_21035	sugar ABC transporter permease	1.525791	8.69E-06
R7A2020_18210	BA14K family protein	1.566181	0.026619
R7A2020_11185	DUF2314 domain-containing protein	1.669598	0.007162
R7A2020_06530	DUF4034 domain-containing protein	1.906954	5.41E-08
R7A2020_05270	hypothetical protein downstream of SagB/ThcOx family dehydrogenase	1.97199	5.70E-05
R7A2020_28630	tRNA-Thr	1.986566	7.59E-09
R7A2020_05265	hypothetical protein downstream of SagB/ThcOx family dehydrogenase	2.06813	1.27E-06
R7A2020_28340	response regulator	2.160544	2.20E-32
R7A2020_05275	SagB/ThcOx family dehydrogenase	2.173617	0.005161
R7A2020_24295	hypothetical protein	2.676231	0.001329
R7A2020_24290	pilus assembly protein	2.936767	3.93E-10

Differentially expressed genes in CC1192 wild type compared to CC1192Δ*mqsRIC*

Locus Tag	Gene product	Log2 Fold change	Adjusted p value
A4R28_34150	lasso peptide	-4.29462	4.31E-85
A4R28_27165	PqqD family protein	-3.08787	4.28E-11
A4R28_02575	MucR family transcriptional regulator	-1.69265	3.38E-09
A4R28_03435	SDR family oxidoreductase	-1.67201	0.000241
A4R28_03430	UDP-glucose 4-epimerase GalE	-1.42711	0.003004
A4R28_06315	hypothetical protein	-1.41733	7.03E-06
A4R28_04680	type II toxin-antitoxin system RelE/ParE family toxin	-1.38352	0.025899
A4R28_03840	cupin domain-containing protein	-1.30792	0.006206
A4R28_08735	extracellular solute-binding protein	-1.29663	0.000232
A4R28_08020	cytochrome o ubiquinol oxidase subunit IV	-1.21972	5.05E-05
A4R28_28940	extracellular solute-binding protein	-1.21278	1.45E-06
A4R28_05580	hypothetical protein	-1.19941	0.008882
A4R28_27870	VOC family protein	-1.19286	3.37E-06
A4R28_26095	type II toxin-antitoxin system VapB family antitoxin	-1.18856	0.04934
A4R28_00975	ABC transporter substrate-binding protein	-1.18009	5.16E-05
A4R28_28260	helix-turn-helix domain-containing protein	-1.15964	0.000115
A4R28_03300	pilus assembly protein	-1.15643	0.023484
A4R28_00330	calcium-binding protein	-1.13318	0.00285

A4R28_24265	DUF559 domain-containing protein	-1.13297	0.016615
A4R28_26460	ABC transporter substrate-binding protein	-1.12547	9.70E-06
A4R28_26390	DUF680 domain-containing protein	-1.11324	0.009072
A4R28_28945	sugar ABC transporter permease	-1.10932	0.000285
A4R28_11745	helix-turn-helix transcriptional regulator	-1.1086	0.005967
A4R28_07555	heat shock protein HspQ	-1.08696	0.004375
A4R28_07050	ATP-dependent Clp protease adapter ClpS	-1.07786	6.27E-06
A4R28_12150	hypothetical protein	-1.06503	1.56E-06
A4R28_28145	ABC transporter ATP-binding protein	-1.06449	0.008851
A4R28_15300	AIPR family protein	-1.05403	1.84E-05
A4R28_19240	sugar phosphate isomerase/epimerase	-1.01874	0.000247
A4R28_27840	RidA family protein	-1.00221	0.008504
A4R28_09940	hypothetical protein	-0.99423	0.032661
A4R28_28950	carbohydrate ABC transporter permease	-0.99337	0.040038
A4R28_12135	DUF559 domain-containing protein	-0.9924	0.000157
A4R28_07290	ABC transporter substrate-binding protein	-0.97553	5.12E-06
A4R28_05810	hypothetical protein	-0.97489	0.028225
A4R28_12140	ABC transporter ATP-binding protein	-0.96824	1.40E-05
A4R28_06320	50S ribosomal protein L33	-0.96012	0.007709
A4R28_07300	ABC transporter permease	-0.95768	0.029335
A4R28_21500	amino acid ABC transporter ATP-binding protein	-0.89759	0.003812
A4R28_28955	sn-glycerol-3-phosphate ABC transporter ATP-binding protein UgpC	-0.88558	0.020497
A4R28_21395	DUF4231 domain-containing protein	-0.84347	0.036415
A4R28_01110	sugar ABC transporter substrate-binding protein	-0.84194	0.023366
A4R28_11985	PIN-like domain-containing protein	-0.84132	0.008882
A4R28_00960	TIGR04076 family protein	-0.84058	0.043836
A4R28_12120	ABC transporter substrate-binding protein	-0.84055	0.000345
A4R28_03560	MucR family transcriptional regulator	-0.83817	0.013216
A4R28_27420	NIPSNAP family protein	-0.83554	0.002235
A4R28_27960	DUF2158 domain-containing protein	-0.83324	0.004307
A4R28_20675	5-demethoxyubiquinol-8 5-hydroxylase UbiM	-0.82906	0.021331
A4R28_27140	exopolysaccharide production repressor ExoX protein	-0.81448	0.001188

A4R28_16690	AraC family transcriptional regulator	-0.81105	0.045682
A4R28_07700	RidA family protein	-0.80519	0.004776
A4R28_08005	ubiquinol oxidase subunit II	-0.79709	0.028676
A4R28_29885	ABC transporter substrate-binding protein	-0.79656	0.002206
A4R28_08765	GFA family protein	-0.78945	0.017566
A4R28_01450	ABC transporter substrate-binding protein	-0.78614	0.001605
A4R28_05045	sugar ABC transporter substrate-binding protein	-0.77943	0.003437
A4R28_05290	hypothetical protein	-0.76558	0.013822
A4R28_14605	hypothetical protein	-0.76113	0.045281
A4R28_13125	AraC family transcriptional regulator	-0.75297	0.019241
A4R28_15395	cold-shock protein	-0.75007	3.81E-05
A4R28_16785	extracellular solute-binding protein	-0.74065	0.00285
A4R28_14330	sugar ABC transporter ATP-binding protein	-0.73281	0.021331
A4R28_19620	hypothetical protein	-0.72135	0.031036
A4R28_28135	ABC transporter permease	-0.70614	0.014125
A4R28_07995	ATP-binding protein	-0.69852	0.046199
A4R28_12155	dihydropyrimidinase	-0.69633	0.000232
A4R28_19860	YdcH family protein	-0.69616	0.010493
A4R28_26465	ABC transporter ATP-binding protein	-0.688	0.024078
A4R28_09390	cupin domain-containing protein	-0.68653	0.044238
A4R28_08910	transglutaminase-like cysteine peptidase	-0.6628	0.001852
A4R28_08135	septum formation initiator family protein	-0.6582	0.031955
A4R28_15820	YbjN domain-containing protein	-0.65245	0.013734
A4R28_09245	DUF3037 domain-containing protein	-0.64879	0.021812
A4R28_21795	IucA/IucC family siderophore biosynthesis protein	-0.64806	0.011486
A4R28_08890	CBS domain-containing protein	-0.64548	0.013746
A4R28_07045	phasin family protein	-0.64129	0.00285
A4R28_11285	hypothetical protein	-0.6407	0.027444
A4R28_33270	hypothetical protein	-0.63338	0.048187
A4R28_07410	ABC transporter substrate-binding protein	-0.62656	0.004776
A4R28_27745	aromatic ring-hydroxylating dioxygenase subunit alpha	-0.61239	0.015483
A4R28_10985	UDP-3-O-acyl-N-acetylglucosamine deacetylase	-0.61137	0.033654

A4R28_02535	branched-chain amino acid aminotransferase	-0.60461	0.009436
A4R28_29290	polyamine ABC transporter substrate-binding protein	-0.60071	0.034252
A4R28_02870	tetratricopeptide repeat-containing sulfotransferase family protein	-0.59893	0.021205
A4R28_13390	DUF4406 domain-containing protein	-0.59161	0.027518
A4R28_27865	bifunctional salicylyl-CoA 5-hydroxylase/oxidoreductase	-0.58037	0.035549
A4R28_13395	ATP-binding protein	-0.56439	0.046115
A4R28_18865	iron ABC transporter substrate-binding protein	-0.54323	0.021331
A4R28_27895	ABC transporter substrate-binding protein	-0.50456	0.030365
A4R28_05000	sugar ABC transporter substrate-binding protein	-0.50277	0.022076
A4R28_26480	ABC transporter substrate-binding protein	-0.45386	0.04934
A4R28_13380	septum site-determining protein MinD	0.451021	0.038876
A4R28_01650	F0F1 ATP synthase subunit B	0.485613	0.032964
A4R28_06250	hypothetical protein	0.500356	0.043067
A4R28_12310	alpha-2-macroglobulin family protein	0.503693	0.028355
A4R28_26805	30S ribosomal protein S1	0.515738	0.038966
A4R28_16725	aldehyde dehydrogenase family protein	0.520519	0.029204
A4R28_23600	RNA 2',3'-cyclic phosphodiesterase	0.522998	0.036457
A4R28_08200	alanine dehydrogenase	0.528135	0.027929
A4R28_09195	TolC family outer membrane protein	0.528494	0.03624
A4R28_20170	phosphoglucosamine mutase	0.529719	0.020169
A4R28_23985	dihydrolipoyl dehydrogenase	0.540034	0.027518
A4R28_03950	chaperonin GroEL	0.546315	0.029828
A4R28_27520	dUTP diphosphatase	0.547454	0.04128
A4R28_21170	MucR family transcriptional regulator	0.548387	0.049967
A4R28_19970	efflux RND transporter permease subunit	0.553087	0.032629
A4R28_06635	cytochrome c	0.554194	0.015647
A4R28_13495	xanthine dehydrogenase family protein molybdopterin-binding subunit	0.558386	0.041485
A4R28_24840	SIS domain-containing protein	0.56195	0.047716
A4R28_03260	DNA repair protein RadA	0.566021	0.038245
A4R28_20830	Mov34/MPN/PAD-1 family protein	0.584049	0.009072
A4R28_26075	double-strand break repair protein AddB	0.593035	0.04699
A4R28_25540	amino acid ABC transporter permease	0.594609	0.005014

A4R28_20835	E2 ligase fold family C protein	0.597426	0.00285
A4R28_23725	aspartate-semialdehyde dehydrogenase	0.599394	0.021781
A4R28_19555	lipoprotein	0.634372	0.018736
A4R28_11060	16S rRNA (cytosine(1402)-N(4))-methyltransferase RsmH	0.641659	0.006144
A4R28_11800	hypothetical protein	0.650536	0.011486
A4R28_04790	outer membrane beta-barrel protein	0.652028	0.01123
A4R28_06640	GMC family oxidoreductase	0.661343	0.001066
A4R28_16665	invasion associated locus B family protein	0.662423	0.004041
A4R28_04480	Ppx/GppA family phosphatase	0.671205	0.025605
A4R28_03575	NAD(P)/FAD-dependent oxidoreductase	0.674398	0.028169
A4R28_16130	formate--tetrahydrofolate ligase	0.683006	0.031417
A4R28_18770	hypothetical protein	0.685695	0.021358
A4R28_08445	50S ribosomal protein L2	0.687032	0.006601
A4R28_03290	30S ribosomal protein S18	0.690176	0.002235
A4R28_10000	HlyD family type I secretion periplasmic adaptor subunit	0.690984	0.014414
A4R28_08570	hypothetical protein	0.693093	0.021331
A4R28_13025	EAL domain-containing protein	0.694332	0.009364
A4R28_23890	hypothetical protein	0.695874	0.000124
A4R28_29190	Wzz/FepE/Etk N-terminal domain-containing protein	0.70395	0.009072
A4R28_20175	hypothetical protein	0.704245	0.000478
A4R28_08030	aspartate/tyrosine/aromatic aminotransferase	0.709694	0.031735
A4R28_12945	D-alanine--D-alanine ligase	0.716215	0.036807
A4R28_07105	translation elongation factor Ts	0.71933	0.011236
A4R28_13285	sulfatase	0.719971	0.022143
A4R28_03295	30S ribosomal protein S6	0.724149	0.002337
A4R28_13700	exopolysaccharide transport family protein	0.72426	0.00079
A4R28_23665	lytic transglycosylase domain-containing protein	0.72891	0.000911
A4R28_07245	triose-phosphate isomerase	0.733695	0.004776
A4R28_28825	pyridoxal phosphate-dependent aminotransferase	0.734907	0.006601
A4R28_17040	response regulator transcription factor	0.736445	0.048145
A4R28_10815	glucose/quinate/shikimate family membrane-bound PQQ-dependent dehydrogenase	0.737599	0.002577
A4R28_17620	DUF992 domain-containing protein	0.74568	0.009445

A4R28_12315	penicillin-binding protein 1C	0.746952	0.005224
A4R28_00115	LysR family transcriptional regulator	0.757136	0.021149
A4R28_24080	dephospho-CoA kinase	0.764924	0.006956
A4R28_01640	F0F1 ATP synthase subunit C	0.768675	0.000451
A4R28_04005	iron ABC transporter permease	0.778602	0.003342
A4R28_24720	YifB family Mg chelatase-like AAA ATPase	0.797987	0.005149
A4R28_17490	efflux RND transporter periplasmic adaptor subunit	0.798714	0.020808
A4R28_08535	50S ribosomal protein L1	0.801419	0.001527
A4R28_13705	hypothetical protein	0.809733	0.012352
A4R28_01185	transporter substrate-binding domain-containing protein	0.818723	0.027518
A4R28_15030	polynucleotide kinase-phosphatase	0.8224	0.001029
A4R28_09995	type I secretion system permease/ATPase	0.828227	0.003252
A4R28_16140	helix-turn-helix domain-containing protein	0.828567	0.010441
A4R28_06645	gluconate 2-dehydrogenase subunit 3 family protein	0.859602	0.000352
A4R28_18840	electron transfer flavoprotein subunit alpha/FixB family protein	0.86632	6.73E-05
A4R28_15615	autotransporter domain-containing protein	0.867404	0.018291
A4R28_03660	phosphoribosylglycinamide formyltransferase	0.870263	0.021781
A4R28_15460	Lrp/AsnC family transcriptional regulator	0.874275	0.011024
A4R28_02410	ABC transporter permease subunit	0.882235	1.14E-05
A4R28_00060	nuclear transport factor 2 family protein	0.886704	0.018736
A4R28_06795	DMT family transporter	0.890147	0.003844
A4R28_23180	tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex transferase subunit TsaD	0.894686	0.006486
A4R28_26290	tetratricopeptide repeat protein	0.906895	0.012189
A4R28_27470	DNA mismatch repair protein MutS	0.924949	0.000139
A4R28_09990	hypothetical protein	0.925124	0.000531
A4R28_13685	DUF1972 domain-containing protein	0.929344	0.00334
A4R28_19625	hypothetical protein	0.948774	0.008938
A4R28_13885	GNAT family N-acetyltransferase	0.953586	7.58E-05
A4R28_25705	hypothetical protein	0.96563	0.031417
A4R28_17690	hypothetical protein	0.968466	0.007262
A4R28_10565	hypothetical protein	0.971592	0.003942

A4R28_13710	FAD-dependent monooxygenase	0.978766	0.000232
A4R28_16685	YadA-like family protein	0.978927	2.69E-07
A4R28_14420	ABC transporter substrate-binding protein	0.980734	0.000648
A4R28_15485	LysE family translocator	0.981271	0.033939
A4R28_18215	copper chaperone PCu(A)C	0.986808	0.01123
A4R28_17370	enoyl-CoA hydratase/isomerase family protein	0.988091	0.032629
A4R28_14160	hypothetical protein	0.994329	0.045332
A4R28_06585	MFS transporter	1.017709	0.001635
A4R28_18920	diaminopimelate decarboxylase	1.019928	0.00011
A4R28_15780	hypothetical protein	1.025241	1.26E-06
A4R28_19170	3-dehydroquinate synthase	1.040608	0.000202
A4R28_19425	L,D-transpeptidase	1.050333	0.004882
A4R28_03240	hypothetical protein	1.053534	0.000567
A4R28_23870	hypothetical protein	1.05491	0.000419
A4R28_29900	MIP family channel protein	1.058726	4.28E-09
A4R28_07265	DUF1697 domain-containing protein	1.059735	2.18E-05
A4R28_05210	hypothetical protein	1.080194	0.000247
A4R28_19630	DUF2865 domain-containing protein	1.101077	1.82E-05
A4R28_04010	phosphatidate cytidylyltransferase	1.108943	0.039326
A4R28_10890	SH3 domain-containing protein	1.122301	6.94E-06
A4R28_29415	carbohydrate ABC transporter permease	1.134691	0.023266
A4R28_09985	hypothetical protein	1.134714	0.002913
A4R28_29480	Flp family type IVb pilin	1.141426	0.027444
A4R28_18085	siroheme synthase CysG	1.155282	0.042461
A4R28_22855	hypothetical protein	1.1716	0.000126
A4R28_25065	ABC transporter permease	1.172652	3.28E-05
A4R28_29905	RsmB/NOP family class I SAM-dependent RNA methyltransferase	1.192659	0.000179
A4R28_24525	RNA polymerase sigma factor	1.246227	0.027817
A4R28_26670	CGNR zinc finger domain-containing protein	1.250771	0.00086
A4R28_07640	autotransporter domain-containing protein	1.264633	3.83E-05
A4R28_35465	RHS repeat-associated core domain-containing protein	1.299984	0.00439
A4R28_12255	hypothetical protein	1.327969	4.49E-05

A4R28_13275	AraC family transcriptional regulator	1.363444	0.026461
A4R28_10435	ABC transporter ATP-binding protein	1.377288	0.003812
A4R28_15755	peptidase	1.39294	1.52E-10
A4R28_29485	CpaF family protein	1.398712	0.008927
A4R28_09565	bifunctional proline dehydrogenase/L-glutamate gamma-semialdehyde dehydrogenase PutA	1.411328	4.30E-06
A4R28_29475	Flp pilus assembly protein CpaB	1.471717	0.000309
A4R28_22415	hypothetical protein	1.517904	1.29E-07
A4R28_29385	HAD family phosphatase	1.527582	1.40E-05
A4R28_13215	xanthine dehydrogenase family protein subunit M	1.537756	0.002206
A4R28_22865	hypothetical protein	1.545798	1.45E-05
A4R28_01190	cation:dicarboxylase symporter family transporter	1.606263	2.10E-12
A4R28_18900	EAL domain-containing protein	1.640768	4.06E-08
A4R28_17340	lysozyme inhibitor LprI family protein	1.653095	2.27E-09
A4R28_17335	hypothetical protein	1.691975	2.71E-12
A4R28_06460	hypothetical protein	1.709958	1.16E-10
A4R28_08690	DUF2569 family protein	1.732912	8.22E-05
A4R28_28620	enoyl-CoA hydratase-related protein	1.753113	1.56E-14
A4R28_03475	hypothetical protein	1.783183	1.24E-15
A4R28_10920	hypothetical protein	1.80923	2.48E-08
A4R28_33975	DUF3011 domain-containing protein	1.880897	5.83E-10
A4R28_13335	hypothetical protein	1.975863	7.64E-08
A4R28_29265	response regulator	2.030264	3.29E-17
A4R28_35450	hypothetical protein	2.036195	1.57E-07
A4R28_22850	SagB/ThcOx family dehydrogenase	2.062519	1.65E-19
A4R28_10915	hypothetical protein	2.08103	2.35E-19
A4R28_26750	hypothetical protein	2.112305	4.48E-21
A4R28_22860	hypothetical protein	2.12896	2.61E-09
A4R28_07985	sigma-70 family RNA polymerase sigma factor	2.170344	4.19E-10
A4R28_02400	pilus assembly protein	2.314555	7.88E-06
A4R28_02395	hypothetical protein	2.798682	3.72E-18