

Characterization of Genetic Loci for Carbon Metabolism and Competition  
for Nodule Occupancy in *Sinorhizobium meliloti*

By

Barney A. Geddes

A Thesis Submitted to the Faculty of Graduate Studies of  
The University of Manitoba  
In Partial Fulfillment of the Requirements of the Degree of

Doctor of Philosophy

Department of Microbiology  
University of Manitoba  
Winnipeg, Manitoba, Canada

Copyright © 2014 by Barney A. Geddes

## Abstract

In agriculture nitrogen fixation by rhizobial inocula is an environmentally and economically beneficial alternative to synthetic fertilization. The effectiveness of rhizobial inocula can be limited by the inability of inoculum strains to compete with indigenous strains for nodule occupancy. *Sinorhizobium meliloti* fixes nitrogen in a complex symbiotic relationship with legume hosts including the agriculturally important forage *Medicago sativa* and the model legume *Medicago truncatula*. The ability to utilize organic compounds has emerged as an important trait for competitiveness for nodule occupancy in *S. meliloti* and other rhizobia. This thesis describes the use of bacterial genetics to characterize two carbon metabolism loci in *S. meliloti*. A genetic locus for erythritol catabolism was characterized and shown to encode an ABC transporter that is required for the catabolism of erythritol, adonitol and L-arabitol, as well as the genes for the catabolism of these three polyols. The ability to utilize erythritol was not necessary for the ability to compete for nodule occupancy in *S. meliloti*, in contrast to *Rhizobium leguminosarum*. A genetic locus that encodes components of the De Ley-Doudoroff pathway of galactose catabolism was identified and also characterized. The inability to catabolize galactose resulted in an increased ability to compete for nodule occupancy in *S. meliloti*. Evidence is presented that is consistent with the hypothesis that increased competitiveness resulted from enhanced production of the symbiotic exopolysaccharide succinoglycan. Inferences are drawn that contribute to the broader understanding of rhizobium-legume symbiosis.

## Acknowledgements

First and foremost, I would like to acknowledge my PhD supervisor Ivan Oresnik, without whom I very well may not have found my love for scientific research. His mentorship and guidance was upheld throughout the many years we spent working together spanning from my time as an undergraduate to becoming the recipient of a PhD. Moreover, he became a close friend and our professional relationship evolved to feel as though we were working together as colleagues, exploring exciting new science together. I was given every opportunity to learn and explore my interests throughout my PhD and I will be ever grateful for all that I received from him.

I would also like to thank the past and present members of the Oresnik lab, many of whom I also became close friends with. I always felt coming to work that I was excited to be there, and beyond being interested in the science, I looked forward to the friendly interactions we would have on a daily basis. I would especially recognise Damien for all that we shared during our virtually paralleled time in the Oresnik lab, from scientific and non-scientific “discussions” to friendship and many good times outside of work.

Finally, I would like to recognise my scientific colleagues, especially my committee members at the University of Manitoba, and members of the Canadian rhizobium communittee, who have always been supportive of me and bettered my science through intelligent discussion. I also acknowledge funding from the Natural Sciences and Engineering Research Council of Canada and the Government of Manitoba that allowed me to carry out my PhD work.

## **Dedication**

I would happily like to dedicate this work to my family. To my Fiancé who has never wavered in her support of me throughout my scientific career, and continues to enthusiastically support me during our move to Oxford. She was always able transform adversity from feeling as heavy as a mountain to feeling as light as a feather, and bolstered the high points and accomplishments that I had throughout my degree from feelings of subtle satisfaction to feelings of great joy and excitement.

To my Mom and Dad, Maurice and Jane Geddes, who nurtured me through the most amazing childhood I could possibly conceive for a growing young man. I count myself as truly fortunate to have been raised by the people I was, and in the environment I was, and I would not change a single thing looking back. I firmly believe that the road to any success I have had or would have in the future was, and continues to be paved by their love, support and nurturing. Further, growing up in the farm setting that I did has given me a perspective on agricultural science that I believe few have, and enriches my love of the work I have done and continue to do.

Finally, to my best friend and brother Charles, who its been incredible to grow with in life both outside of, and within academics. He has never ceased to astound me with what he has been able to accomplish. There are many aspects of his personality (especially his work ethic) that are a complete inspiration to me, and that I have tried hard to mirror in myself, as I know they would make both a better man and a better scientist.

I have been truly blessed to have the love and support of all those around me throughout my time as a student. I hope that in part with the completion of this PhD, but mostly moving forward I can prove myself worthy of it all.

## Table of Contents

<b>Abstract</b>	ii
<b>Acknowledgements</b>	iii
<b>Dedication</b>	iv
<b>List of Tables</b>	ix
<b>List of Figures</b>	x
<b>List of Abbreviations</b>	xi
<b>Chapter 1: Literature Review</b>	1
<b>1.1 Introduction to Rhizobia</b>	2
1.1.1 Nitrogen fixation	2
1.1.2 Agricultural fertilization	3
1.1.3 Rhizobia	5
1.1.4 <i>Sinorhizobium meliloti</i> Rm1021 (wild-type)	6
<b>1.2 Rhizobium-Legume Symbiosis</b>	8
1.2.1 Signal exchange in the rhizosphere	8
1.2.2 Root hair invasion	13
1.2.3 Bacteroid differentiation	17
1.2.4 Nodule development and physiology	20
<b>1.3 <i>Sinorhizobium meliloti</i> Central Carbon Metabolism</b>	23
1.3.1 Entner-Doudoroff pathway	24
1.3.2 Embden-Meyerhoff-Parnas pathway	30
1.3.3 Pentose phosphate pathway	33
1.3.4 Calvin-Benson-Bassham cycle	35
1.3.5 Tricarboxylic acid cycle	37
1.3.6 Gluconeogenesis	39
1.3.7 Anaplerotic pathways	40
<b>1.4 <i>Sinorhizobium meliloti</i> Peripheral Carbon Metabolism</b>	42
1.4.1 Disaccharide catabolism and transport	42
1.4.2 Hexose catabolism and transport	46
1.4.3 Pentose catabolism and transport	50
1.4.4 Methylpentose catabolism and transport	52
1.4.5 Sugar alcohol catabolism and transport	55
1.4.6 Regulation of peripheral metabolism	61
<b>1.5 Biological Relevance of Carbon Metabolism</b>	63
1.5.1 Free-living catabolism	63
1.5.2 Carbon catabolism and invasion	67
1.5.3 Bacteroid metabolism	71
<b>1.6 Thesis Objectives</b>	74
<b>Chapter 2: A Locus Necessary for the Transport and Catabolism of Erythritol in <i>Sinorhizobium meliloti</i></b>	76
<b>2.1 Introduction</b>	77
<b>2.2 Materials and Methods</b>	79

2.2.1 Bacterial strains, plasmids and media	79
2.2.2 DNA manipulations and constructions	79
2.2.3 Genetic manipulations	84
2.2.4 $\beta$ -galactosidase assays	85
2.2.5 RNA isolation and cDNA synthesis	85
2.2.6 Junction PCR	86
2.2.7 Quantitative RT-PCR	86
2.2.8 Transport Assay	87
2.2.9 Competition Assay	87
<b>2.3 Results</b>	87
2.3.1 Isolation of erythritol transport and catabolic mutants	87
2.3.2 Erythritol utilization genes are in different complementation groups	91
2.3.3 <i>eryA</i> and <i>eryB</i> are contained in a single transcript	93
2.3.4 Erythritol locus does not contain uncharacterized genes involved in erythritol catabolism	93
2.3.5 Components of the ABC-type transporter defined by <i>SMc01628-SMc01624</i> are necessary for transport of erythritol	97
2.3.6 EryD and <i>SMc01615</i> negatively regulate their own transcripts and affect the expression of other ery transcripts	99
2.3.7 The inability to catabolise erythritol affects gene expression	102
2.3.8 The erythritol locus contains determinants for adonitol and arabitol utilization	103
2.3.9 Lack of erythritol, adonitol, and arabitol utilization does not affect competition for nodule occupancy in <i>S. meliloti</i>	105
<b>2.4 Discussion</b>	105
<b>Chapter 3: Genetic Characterization of a Complex Locus Necessary for the Transport and Catabolism of Erythritol, Adonitol, and L-Arabitol in <i>Sinorhizobium meliloti</i></b>	112
<b>3.1 Introduction</b>	113
<b>3.2 Materials and Methods</b>	115
3.2.1 Bacterial strains, plasmids and media	115
3.2.2 Genetic techniques	115
3.2.3 DNA manipulations	119
3.2.4 RNA isolation and cDNA synthesis	119
3.2.5 Quantitative RT-PCR	120
3.2.6 Kinase assays	120
3.2.7 Transport assays	121
<b>3.3 Results</b>	121
3.3.1 Adonitol and L-arabitol compete with erythritol for transport	121
3.3.2 EryB is not necessary for adonitol or L-arabitol catabolism	123
3.3.3 EryA can use erythritol, adonitol, and L-arabitol as substrates.	128
3.3.4 Genes in the locus are uniquely necessary for adonitol and/or L-arabitol utilization	129
3.3.5 <i>xytB</i> is necessary for D-arabitol but not L-arabitol or adonitol utilization	133
3.3.6 Bioinformatic characterization of genes necessary for adonitol and L-arabitol utilization	136

3.3.7 Adonitol and L-arabitol only induce transcripts necessary for their utilization	137
<b>3.4 Discussion</b>	139
<b>Chapter 4: Phylogenetic Analysis of Erythritol Catabolic Loci within the <i>Rhizobiales</i> and Proteobacteria</b>	145
4.1 Introduction	146
4.2 Materials and Methods	148
4.2.1 Identification of erythritol loci	148
4.2.2 Phylogenetic analysis	150
4.3 Results	152
4.3.1 Phylogenetic distribution of putative erythritol loci	152
4.3.2 Genetic content of loci	152
4.3.3 The genetic structure of erythritol loci	157
4.3.4 The phylogenies of erythritol proteins do not correlate with species phylogeny	158
4.3.5 Modes of evolution for the polyol utilization loci	164
4.4 Discussion	167
<b>Chapter 5: Inability to Catabolize Galactose Leads to an Increased Ability to Compete for Nodule Occupancy in <i>Sinorhizobium meliloti</i></b>	174
5.1 Introduction	175
5.2 Materials and Methods	177
5.2.1 Bacterial strains, plasmids and media	177
5.2.2 Genetic techniques	177
5.2.3 DNA manipulations	180
5.2.4 Non-denaturing gel electrophoresis and dehydrogenase assays	181
5.2.5 Transport competition assays	182
5.2.6 RNA isolation, cDNA synthesis and quantitative RT-PCR	182
5.2.7 Competition for nodule occupancy assays	183
5.2.8 $\beta$ -galactosidase assays	184
5.3 Results	184
5.3.1 Identification of the galactose locus	184
5.3.2 Genetic characterization of the galactose locus	185
5.3.3 GalD is a galactose dehydrogenase	191
5.3.4 Identification of a galactose transporter	193
5.3.5 Galactose and glucose compete with arabinose for transport	195
5.3.6 Induction of the galactose locus	198
5.3.7 Inability to catabolize galactose affects <i>araABC</i> expression	200
5.3.8 SRmD144 is more competitive than Rm1021 for nodule occupancy	200
5.3.9 <i>araA</i> does not complement an <i>Agrobacterium chvE</i> mutant for virulence	202
5.4 Discussion	205
<b>Chapter 6: Exopolysaccharide production in response to decreased pH is correlated with an increase in competition for nodule occupancy</b>	209
6.1 Introduction	210
6.2 Materials and Methods	213
6.2.1 Bacterial strains, plasmid and media	213

6.2.2 DNA and genetic manipulations	213
6.2.3 Non-denaturing gel electrophoresis and dehydrogenase assays	215
6.2.4 Quantitative analysis of exopolysaccharide production	215
6.2.5 Exopolysaccharide composition analysis	216
6.2.6 Exopolysaccharide molecular weight distribution	217
6.2.7 RNA isolation, cDNA synthesis and quantitative RT-PCR	218
6.2.8 Fluorescent microscopy	219
6.2.9 Competition for nodule occupancy assays	220
6.2.10 Nodule kinetics assays	220
<b>6.3 Results</b>	221
6.3.1 Altered exopolysaccharide production in a galactose mutant of <i>S. meliloti</i>	221
6.3.2 Early exopolysaccharide production and acidification of culture media by SRmD304	225
6.3.3 Early exopolysaccharide produced in SRmD304 is succinoglycan	227
6.3.4 Gene expression correlates with early succinoglycan production	231
6.3.5 <i>SMc00588</i> encodes a galactose dehydrogenase that is required for galactose-dependent culture medium acidification	234
6.3.6 Early nodulation in an <i>S. meliloti</i> galactose mutant	237
6.3.7 EPS phenotypes in galactose mutants correlate with competition for nodule occupancy	237
6.3.8 Acidification of the curled colonized root hair during invasion by <i>S. meliloti</i>	241
<b>6.4 Discussion</b>	242
<b>Chapter 7: Conclusion</b>	247
7.1 Thesis Conclusions and Observations	248
<b>Literature Cited</b>	267



## List of Tables

<b>Table 1.1.</b> Genes and enzymes of central metabolism in <i>S. meliloti</i>	25
<b>Table 2.1.</b> Bacterial strains and plasmids	80
<b>Table 2.2.</b> PCR primers	82
<b>Table 2.3.</b> Erythritol catabolic genes are in separate complementation groups	92
<b>Table 2.4.</b> Complementation of mutants with erythritol catabolic genes	96
<b>Table 2.5.</b> Induction of erythritol locus $\beta$ -galactosidase gene fusions by erythritol	100
<b>Table 2.6.</b> Regulation of gene expression by <i>eryD</i> and <i>SMc01615</i> analyzed by qRT-PCR	101
<b>Table 2.7.</b> Polyol utilization phenotypes of mutations across the erythritol locus	104
<b>Table 2.8.</b> Phenotypic suppression of a <i>tpiA</i> mutant by <i>tpiB</i> induction	109
<b>Table 3.1.</b> Bacterial strains and plasmids	116
<b>Table 3.2.</b> Growth of erythritol catabolic mutants on adonitol and L-arabitol	127
<b>Table 3.3.</b> Growth of <i>R. leguminosarum</i> with the <i>S. meliloti</i> erythritol cosmid on adonitol and L-arabitol	131
<b>Table 3.4.</b> Complementation analysis of strains containing deletions	132
<b>Table 3.5.</b> Carbon phenotypes of central catabolism mutants	134
<b>Table 3.6.</b> Induction by adonitol and L-arabitol	138
<b>Table 4.1.</b> Bacterial genomes used in this study containing erythritol loci	153
<b>Table 4.2.</b> Content of putative erythritol loci	155
<b>Table 5.1.</b> Bacterial strains and plasmids	178
<b>Table 5.2.</b> Growth of strains on galactose	188
<b>Table 5.3.</b> Toxic growth phenotypes of strains grown on galactose and glycerol	190
<b>Table 5.4.</b> Growth phenotypes of mutants of galactose-induced transporters	194
<b>Table 5.5.</b> Doubling time (h) of an <i>araA</i> mutant	196
<b>Table 5.6.</b> qRT-PCR analysis of galactose locus gene expression	199
<b>Table 5.7.</b> qRT-PCR analysis of <i>araABC</i> gene expression	201
<b>Table 5.8.</b> Complementation of <i>A. tumefaciens chvE</i>	204
<b>Table 6.1.</b> Bacterial strains and plasmids	214
<b>Table 6.2.</b> Calcofluor UV-fluorescence phenotypes of a galactose mutant	223
<b>Table 6.3.</b> Colony mucoidy phenotypes of a galactose mutant in an <i>expR</i> <sup>+</sup> background	224
<b>Table 6.4.</b> Changes in gene expression in SRmD304	233

## List of Figures

<b>Figure 1.1.</b> Invasion of legumes by <i>S. meliloti</i>	9
<b>Figure 1.2.</b> Schematic of the <i>S. meliloti</i> ED pathway	27
<b>Figure 1.3.</b> Schematic of the <i>S. meliloti</i> EMP pathway	31
<b>Figure 1.4.</b> Schematic of the <i>S. meliloti</i> PP pathway	34
<b>Figure 1.5.</b> Schematic of the <i>S. meliloti</i> TCA cycle, gluconeogenesis and anaplerotic pathways	38
<b>Figure 1.6.</b> Entry-points for carbon sources into <i>S. meliloti</i> central metabolism	43
<b>Figure 1.7.</b> De Ley-Doudoroff pathway of galactose catabolism	49
<b>Figure 1.8.</b> Predicted pathway of erythritol catabolism in <i>S. meliloti</i>	59
<b>Figure 2.1.</b> Predicted ORFs and isolated mutant alleles in the erythritol locus of <i>S. meliloti</i> .	89
<b>Figure 2.2.</b> Junction PCR	94
<b>Figure 2.3.</b> Erythritol transport assays	98
<b>Figure 3.1.</b> Map of ORFs in the erythritol, adonitol and L-arabitol locus of <i>S. meliloti</i> Rm1021	122
<b>Figure 3.2.</b> Fischer projections representing stereochemistry of erythritol, adonitol, L-arabitol and D-arabitol	124
<b>Figure 3.3.</b> Competition of erythritol, adonitol and L-arabitol for transport with [ <sup>14</sup> C] erythritol in <i>S. meliloti</i> Rm1021	125
<b>Figure 3.4.</b> Predicted pathway of adonitol and L-arabitol catabolism in <i>S. meliloti</i> Rm1021	141
<b>Figure 4.1.</b> The genetic arrangement of putative erythritol loci in the proteobacteria	156
<b>Figure 4.2.</b> Phylogenetic trees of RpoD and the core erythritol genes EryA, EryB and EryD	160
<b>Figure 4.3.</b> The phylogenetic tree of erythritol proteins does not correlate with species phylogeny; evidence for horizontal gene transfer	163
<b>Figure 4.4.</b> Phylogenetic trees of erythritol transporters	165
<b>Figure 4.5.</b> <i>Mesorhizobium loti</i> contains paralogs of LalA	168
<b>Figure 4.6.</b> <i>Sinorhizobium fredii</i> encodes TpiB xenologs	169
<b>Figure 5.1.</b> Genetic map of the galactose catabolic locus of <i>S. meliloti</i>	186
<b>Figure 5.2.</b> Non denaturing PAGE of galactose-inducible enzyme activity	192
<b>Figure 5.3.</b> Competition for transport with [ <sup>3</sup> H] arabinose in <i>S. meliloti</i> Rm1021	197
<b>Figure 5.4.</b> Competition for nodule occupancy	203
<b>Figure 6.1.</b> Exopolysaccharide production by Rm1021 and SRmD304	226
<b>Figure 6.2.</b> Sugar composition of precipitated EPS.	229
<b>Figure 6.3.</b> Molecular weight profile of precipitated EPS	230
<b>Figure 6.4.</b> Molecular-weight profile of EPS synthesized by Rm1021 and SRmD304	232
<b>Figure 6.5.</b> Non denaturing PAGE of galactose-inducible enzyme activity	236
<b>Figure 6.6.</b> Effects on nodule kinetics in a galactose mutant of <i>S. meliloti</i>	238
<b>Figure 6.7.</b> Competition for nodule occupancy	240
<b>Figure 6.8.</b> The curled colonized root hair is an acidic compartment	243
<b>Figure 7.1.</b> Model for the activation of succinoglycan synthesis by acidic pH	262

## List of Abbreviations

3PG	3-Phosphoglycerate
6PG	6-Phosphogluconate
ABC	ATP-binding cassette
Ac-CoA	Acetyl- Coenzyme A
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BNF	Biological nitrogen fixation
bp	Base pairs
C	Carbon
Ca <sup>2+</sup>	Calcium ion
CCRH	Curled colonized root hair
cDNA	Complementary DNA
Cm	Centimeters
Cm	Chloramphenicol
CO <sub>2</sub>	Carbon dioxide
CPM	Counts per minute
Ct	Cycle threshold
DHAP	Dihydroxyacetone-3-phosphate
DIC	Differential interference of contrast
DNA	Deoxyribonucleic acid
e <sup>-</sup>	Electron
ED	Entner-Doudoroff
EDTA	Ethylenediaminetetraacetic acid
EPS	Exopolysaccharide
EPSI	Succinoglycan
EPSII	Galactoglucan
F6P	Fructose-6-phosphate
g	Gram
G3P	Glyceraldehyde-3-phosphate
G6P	Glucose-6-phosphate
GBq	Gigabacquerel
GFP	Green fluorescent protein
GlcNAc	<i>N</i> -acetyl-D-glucosamine
Gm	Gentamicin
GNG	Gluconeogenesis
h	hours
H	Hydrogen
H <sup>+</sup>	Hydrogen cation
H <sub>2</sub>	Dihydrogen
H <sub>2</sub> O	Water
HMW	High-molecular-weight
HPLC	High-performance liquid chromatography
IM	Inner membrane
IPTG	Isopropyl β-D-1-thiogalactopyranoside

**List of Abbreviations (continued)**

IRLC	Inverted repeat lacking clade
kb	Kilobase
KDPG	2-Keto-3-deoxy-6-phosphogluconate
KDPgal	2-Keto-3-deoxy-6-phosphogalactonate
Km	Kanamycin
L	Liter
LB	Luria-Bertani
LMW	Low-molecular-weight
LPS	Lipopolysaccharide
M	Molar
MBq	Megabacquerel
mCi	Millicurie
mg	Milligram
min	Minutes
ML	Maximum likelihood
mL	Milliliter
mM	Millimolar
mmol	Millimole
Mo	Molybdenum
mol	Mole
N	Nitrogen
N <sub>2</sub>	Dinitrogen
Na	Sodium
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NAD <sup>+</sup>	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NADP <sup>+</sup>	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NC	No change
NCR	Nodule cysteine-rich antimicrobial peptide
NF	Nod factor
NH <sub>3</sub>	Ammonia
NH <sub>4</sub> <sup>+</sup>	Ammonia
Nm	Neomycin
nmol	Nanomole
NMR	Nuclear magnetic resonance spectroscopy
NO <sub>2</sub>	Nitrogen dioxide
OM	Outer membrane
ORF	Open reading frame
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PHB	Polyhydroxybutyrate
P	Phosphate

**List of Abbreviations (continued)**

P <sub>i</sub>	Inorganic phosphate
pmol	Picomole
PP	Pentose phosphate
PQQ	Pyrroloquinoline quinone
PTS	Phosphotransferase system
qRT-PCR	Quantitative real-time PCR
R5P	Ribulose-5-phosphate
Rf	Rifampicin
RFP	Red fluorescent protein
RMM	Rhizobium minimal medium
GYM	Glutamate yeast extract mannitol
RNA	Ribonucleic acid
RT-PCR	Real-time PCR
s	Seconds
SD	Standard deviation
Sm	Streptomycin
SM	Symbiosome membrane
SMCR	Succinate mediated catabolite repression
SNF	Symbiotic nitrogen fixation
Sp	Spectinomycin
sp.	Species
spp.	Species (pl)
Tc	Tetracycline
TCA	Tricarboxylic acid
CBB	Calvin-Benson-Bassham
TE	Tris-EDTA
Tg	Teragram
Tn	Transposon
tRNA	Transfer RNA
IT	Infection thread
TY	Tryptone yeast extract
VMM	Vincent's minimal medium
U	Uniform
UV	Ultraviolet
VLCFA	Very long chain fatty acid
μCi	Microcurie
μg	Microgram
μL	Microliter
μm	Micrometer
μM	Micromolar
μM	Micromole

## **Chapter 1**

### **Literature Review**

## 1.1 Introduction to Rhizobia

### 1.1.1 Nitrogen fixation

Nitrogen is required for the biosynthesis of the basic building blocks of life. Although nitrogen exists in abundance in the earth's atmosphere, its atmospheric form dinitrogen (N<sub>2</sub>) is relatively inert. In order to be used for organic processes nitrogen must be fixed into more biologically accessible forms. There are three common forms of nitrogen fixation. These include atmospheric, biological and industrial nitrogen fixation. In total approximately 380 Tg N/year is fixed by these processes (Galloway *et al.*, 2004).

Most atmospheric nitrogen fixation occurs as a result of lightning. Energy from lightning discharge can drive reactions that form nitrogen compounds from atmospheric N<sub>2</sub> (Noxon, 1976). NO<sub>2</sub> is the most commonly measured product of these reactions and is deposited by the rainfall associated with thunderstorms (Noxon, 1976). Atmospheric nitrogen fixation is estimated to contribute approximately 5 Tg N/year into global nitrogen cycles (Galloway *et al.*, 2004).

Biological nitrogen fixation (BNF) is mediated by diazotrophic microorganisms that are capable of fixing atmospheric nitrogen using the enzyme nitrogenase. Nitrogenase catalyzes the reduction of N<sub>2</sub> to ammonia (NH<sub>3</sub>) in the energetically expensive reaction:

$$\text{N}_2 + 8 \text{H}^+ + 8 \text{e}^- + 16 \text{ATP} \rightarrow 2 \text{NH}_3 + \text{H}_2 + 16 \text{ADP} + 16 \text{P}_i$$
 Diazotrophs exist as free-living, associative or symbiotic microorganisms. Of these, symbiotic diazotrophs are the greatest contributors to BNF. BNF is the greatest contributor of fixed nitrogen to the nitrogen cycle and was estimated to contribute 230 Tg N/year (Galloway *et al.*, 2004).

Prior to industrial nitrogen fixation, the amount of N that entered the terrestrial N-cycle was limited by atmospheric and biological nitrogen fixation. The discovery of the Haber-Bosch process in the early 1900s and its subsequent industrialization has profoundly increased the quantity of nitrogen that is transformed from atmospheric  $N_2$  to  $NH_3$  (Galloway *et al.*, 2004). The Haber-Bosch process is an energy-intensive process that combines  $3H_2$  and  $N_2$  to yield  $2NH_3$ . It is the result of two reactions; the formation of  $H_2$  and  $CO_2$  from methane ( $CH_4$ ) and steam ( $H_2O$ ) using a nickel catalyst, and the conversion of  $N_2$  and  $3H_2$  to  $NH_3$  using high pressure, heat and an iron catalyst (Ertl & Jennings, 1991). In 2005, it was estimated that inorganic nitrogen was contributing 121 Tg N/year into global nitrogen cycles (Galloway *et al.*, 2008).

### ***1.1.2 Agricultural fertilization***

Nitrogen is one of the primary limiting nutrients for plant growth in agriculture. Indeed, the productivity of many ecosystems is controlled by nitrogen availability (Vitousek *et al.*, 2002). For this reason, legume crops have been used in agriculture for thousands of years in crop rotations for their ability to integrate residual nitrogen into agricultural systems. As early as 1838 it was documented that legumes could restore nitrogen to the soil and must be capable of creating it directly. More than 50 years later it was shown that the active participation of living microorganisms was necessary for the creation of nitrogen by legumes (Smil, 2004). These microorganisms are symbiotic diazotrophs, referred to as rhizobia, and have been extensively studied for their ability to fix nitrogen in a symbiotic relationship with legume crops. In modern agriculture, commercially prepared rhizobial inoculants are introduced into the soil during the



planting of legume crops to enhance SNF. Estimates suggest that currently, 40 Tg N/ year is injected into agricultural systems by SNF (Herridge *et al.*, 2008).

The primary source of nitrogen fertilization in modern agriculture is inorganic nitrogen fertilizer synthesized by the Haber-Bosch process. The development of the Haber-Bosch process sparked a “green revolution” that allowed a coincident rapid expansion in agricultural capacity for food production as well as global population following World War II (Galloway *et al.*, 2004). Unfortunately, due to the economic costs associated with inorganic fertilizer use some countries have been effectively left out of the green revolution and remain limited in their agricultural capacity. Synthetic fertilizers currently account for 121 Tg N/ year injected into agricultural systems (Galloway *et al.*, 2008), a rate that has doubled the flux of the terrestrial nitrogen cycle (Rockström *et al.*, 2009). This has resulted in a strain on the environment in the forms of waterway pollution that has resulted in eutrophication of water systems, and greenhouse gas production (Rockström *et al.*, 2009). However, the world’s population has now reached a point where its caloric requirement is greater than that which agriculture can provide without utilizing inorganic nitrogen fertilizer.

Because of the environmental and economic costs associated with the use of inorganic nitrogen as fertilizer there is growing interest in enhancing the use of SNF in agriculture to help overcome these problems, as SNF is essentially free and environmentally benign.

### 1.1.3 *Rhizobia*

Rhizobia are Gram-negative  $\alpha$ - and  $\beta$ - proteobacteria that have acquired the ability to fix atmospheric nitrogen in symbiosis with legumes (Masson-Boivin *et al.*, 2009). Rhizobia are able to elicit the formation of new organs called root nodules on the roots of host plants. Within a microoxic environment provided by the root nodules, rhizobia intracellularly fix atmospheric nitrogen to ammonia that is assimilated by the plant. Two widely distributed sets of genes encode these functions in most rhizobia: the *nod* (Nodulation) genes and the *nif* (Nitrogen Fixation) genes (Masson-Boivin *et al.*, 2009). Rhizobia tend to contain large complex genomes that often include extra replicons called megaplastids (Jumas-Bilak *et al.*, 1998). Variations to these themes exist. *Frankia* are a group of Gram-positive organisms that have acquired the ability to fix nitrogen in a symbiotic association with actinorhizal plants (Benson & Silvester, 1993). Some species of *Bradyrhizobium* that do not contain *nod* genes have been reported to be capable of eliciting nodule formation on host plants (Giraud *et al.*, 2007).

Because of their relevance to agriculture, most research has focused on rhizobia found in the order rhizobiales that nodulate crop and forage legumes. These include *Rhizobium leguminosarum* biovars *viciae* (pea), *trifolii* (clover) and *phaseoli* (kidney bean), *Rhizobium etli* (common bean), *Bradyrhizobium japonicum* (soybean), *Mesorhizobium loti* (*Lotus*), *Sinorhizobium fredii* (soybean) and *Sinorhizobium meliloti* (alfalfa). The understanding of symbiosis between rhizobia and legumes as defined in these organisms has become a paradigm of plant-microbe interaction.

#### **1.1.4 *Sinorhizobium meliloti* Rm1021 (wild-type)**

*Sinorhizobium meliloti* belongs to the Rhizobiaceae family, of the order Rhizobiales in the  $\alpha$ - proteobacteria. Along with other well-studied rhizobia, *S. meliloti* is also closely related to the plant and animal pathogens *Agrobacterium* and *Brucella*. *S. meliloti* engages in nitrogen fixing symbiosis with the agriculturally important forage *Medicago sativa* (alfalfa), a model organism for studying legume biology *Medicago truncatula* (barrel medick), as well as other legumes of the genera *Medicago*, *Melilotus* (sweet clover) and *Trigonella*. *S. meliloti* was originally isolated in New South Wales, Australia in 1937 and designated strain SU47. *S. meliloti* strain Rm1021 is a streptomycin-resistant derivative of *S. meliloti* SU47.

The genome of *S. meliloti* Rm1021 was sequenced relatively early, in 2001, and is composed of a chromosome (3,654,135 bp) and two large megaplasms called pSymA (1,354,226 bp) and pSymB (1,683,333 bp) (Barnett *et al.*, 2001; Capela *et al.*, 2001; Finan *et al.*, 2001; Galibert *et al.*, 2001). Most essential genes in *S. meliloti* are contained on the chromosome. These include genes for universal biosynthetic pathways, transcription, translation, cell division and DNA repair (Capela *et al.*, 2001). The essential genes tRNA<sup>arg</sup>, encoding the arginine tRNA, and *engA* are encoded on pSymB as well as genes for asparagine and thiamine biosynthesis (diCenzo *et al.*, 2012; Finan *et al.*, 1986). The megaplasmid pSymB also contains many gene clusters involved in the biosynthesis and export of surface polysaccharides and small molecule transport and catabolism. Based on these contents, pSymB has been suggested to play a role in the ability of *S. meliloti* to thrive during saprophytic growth in the diverse environment of the soil (Finan *et al.*, 1986; Finan *et al.*, 2001). The megaplasmid pSymA has been cured

from *S. meliloti* and therefore does not encode essential genes (Oresnik *et al.*, 2000). Based on encoding a large portion of the genes involved in symbiosis, including the genes for nodulation and nitrogen fixation, pSymA is thought of as the symbiotic plasmid of *S. meliloti* (Barnett *et al.*, 2001).

The genome of *M. truncatula* has also been sequenced recently (Young *et al.*, 2011). *M. truncatula* has been used as a model legume because it contains a small diploid genome, has a rapid generation time, prolific seed production and is amenable to genetic transformation. The *Sinorhizobium meliloti* - *Medicago truncatula* model (along with *Mesorhizobium loti* – *Lotus japonicus*) is emerging as the leading model system for studying rhizobium- legume symbiosis (Capela *et al.*, 2001). The current understanding of rhizobium-legume symbiosis with a focus on the *S. meliloti* - *Medicago* model is reviewed in the next section (1.2).

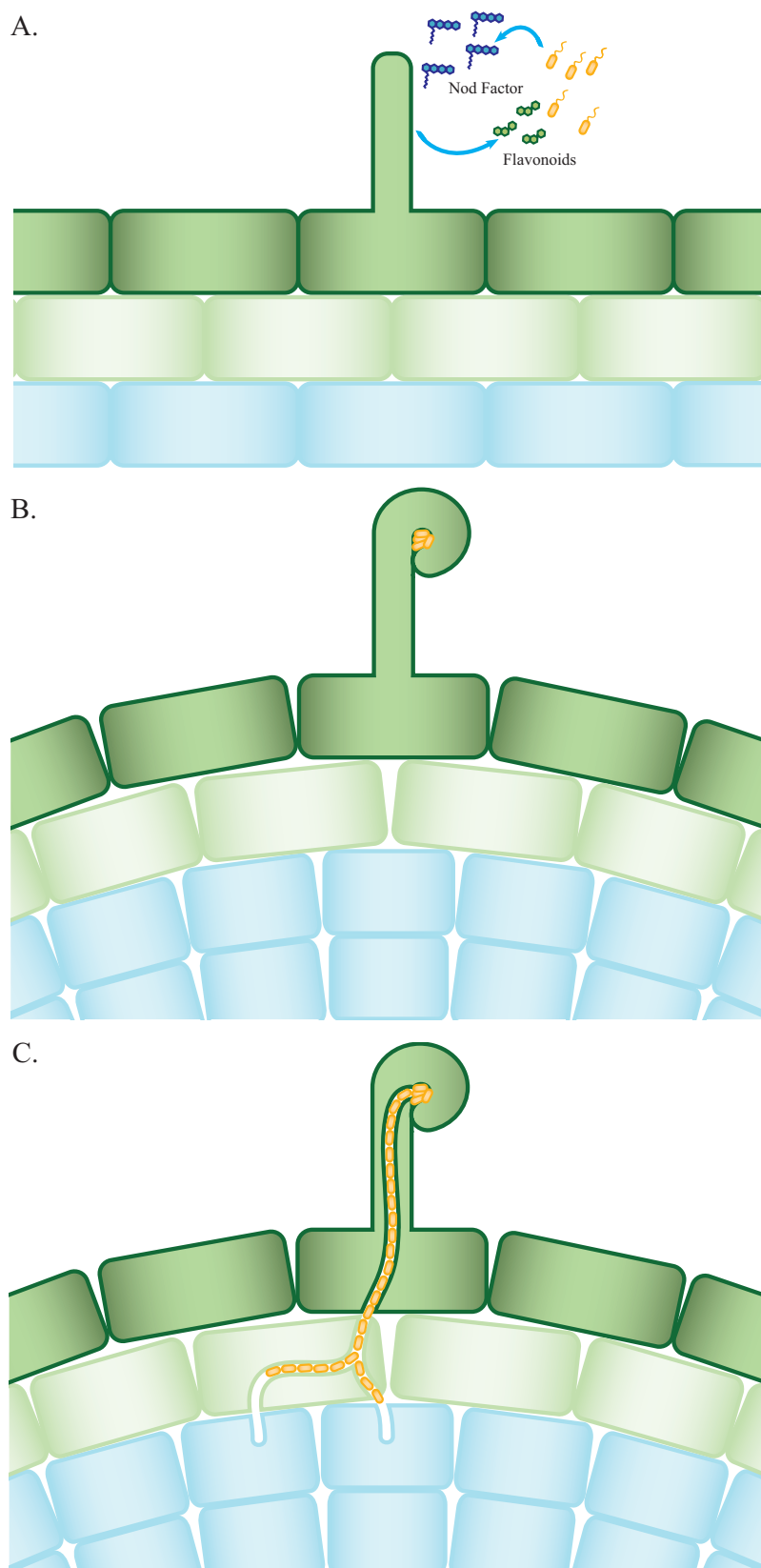
## 1.2 Rhizobium-Legume Symbiosis

The rhizobium-legume symbiosis is an elaborate process that culminates in the development of root nodules, wherein rhizobia intracellularly fix atmospheric nitrogen that is assimilated by the host plant. To achieve this end result, rhizobia must first infect the legumes through root hair cells on the root surface and traverse intracellular tubules called infection threads (IT) before reaching root inner cortical cells that form the nodule primordium, where they are endocytosed and undergo a dramatic developmental differentiation into nitrogen fixing forms referred to as bacteroids (Jones *et al.*, 2007; Oldroyd *et al.*, 2011). This process has been studied extensively in rhizobia. The current understanding of the invasion of *Medicago* species by *S. meliloti* is reviewed in this section.

### 1.2.1 Signal exchange in the rhizosphere

Invasion of the legume by rhizobia begins with a signal exchange that occurs between the legume and saprophytic, free living rhizobia in the soil environment surrounding the plant root, referred to as the rhizosphere (Figure 1.1 A). The signal exchange begins with the secretion of inducing molecules such as flavonones and betains by legumes in their root exudate (Gage, 2004). The type of inducing molecule is variable among legume species and unique to different rhizobia (Allan Downie, 1994). Flavonoids are recognized by rhizobial NodD proteins: LysR family regulators that induce the transcription of downstream *nod* genes. In *S. meliloti* the flavonoid luteolin, secreted by *M. sativa*, was shown to be responsible for the induction of *nod* genes (Peters et al., 1986). *S. meliloti* NodD1 specifically binds luteolin and induces transcription of a subset

**Figure 1.1.** Invasion of legumes by *S. meliloti*



**A.** Rhizobium-legume symbiosis begins with signal exchange in the rhizosphere. Legumes secrete phenolic compounds (flavonoids) that are recognised by rhizobial NodD proteins and trigger the production of the lipochitooligosaccharide, Nod Factor, by rhizobia. Nod Factor is recognised by legume receptors on the surface of the root hair.

**B.** Nod Factor recognition triggers a signaling cascade that leads to morphological changes in the plant, including root hair curling and inner cortical cell division. Root hair curling results in the trapping of rhizobia that have colonized the root hair surface in an apoplastic compartment referred to as a curled colonized root hair. Inner cortical cell division gives rise to the nodule primordium.

**C.** Invasion of the legume proceeds by the formation of an intracellular infection thread as a result of a reversal of polar cell growth in the root hair. The infection thread traverses the root hair and multiple cell layers and ramifies before reaching the destination inner cortical cells. There, rhizobia are endocytosed and differentiate into nitrogen fixing bacteroids.

of genes that contain a specific nucleic acid motif in their promoter called a nod-box, including the *nod* genes (Fisher & Long, 1993). The *nod* genes encode approximately 25 proteins that are involved in the synthesis and export of a lipochitooligosaccharide signaling molecule called Nod Factor (NF) (Gage, 2004). *S. meliloti* possesses two other NodD variants that are capable of activating *nod* gene expression, NodD2 and NodD3. NodD2 responds to a yet unidentified plant compound (Honma *et al.*, 1990). NodD3 is capable of activating *nod* gene expression in the absence of inducing molecules, and is integrated into a complex regulatory circuit with the positive regulator SyrM (Kondorosi *et al.*, 1991; Swanson *et al.*, 1993).

NF consists of a four to five residue chitin backbone of  $\beta$ -1,4-linked *N*-acetyl-D-glucosamine (GlcNAc) subunits with an N-linked acyl tail attached to the non-reducing end (Gage, 2004). Further decorations of the NF backbone as well as the modifications to the lipid tail are variable among rhizobial species, and even individual species are capable of synthesizing an array of different NFs (Perret *et al.*, 2000). NF synthesized by *S. meliloti* has been purified and shown to be an acetylated, sulfated  $\beta$ -1,4-linked tetrasaccharide of *N*-acetyl-D-glucosamine with a C16 acyl tail (Lerouge *et al.*, 1990). Enzymes for the synthesis of the chitin backbone and attachment of the lipid tail are encoded by the *nodABC* operon in *S. meliloti*, divergently transcribed from *nodDI* (Atkinson *et al.*, 1994; Egelhoff *et al.*, 1985; Geremia *et al.*, 1994; John *et al.*, 1993; Spaink *et al.*, 1994). Enzymes required for acetylation and sulfation of NF are encoded by *nodL* and *nodHPQ*, respectively (Ardourel *et al.*, 1995; Roche *et al.*, 1991). The *nod* genes *nodE* and *nodF* are involved in synthesis of the acyl chain (Demont *et al.*, 1993).



Following its synthesis, NF is perceived by receptor-like kinases on the plasma membrane of the host plant that contain extracellular chitooligosaccharide binding LysM domains (Limpens *et al.*, 2003; Madsen *et al.*, 2003; Radutoiu *et al.*, 2003). In *M. truncatula* these are MtNFP and MtLYK3 (Amor *et al.*, 2003; Limpens *et al.*, 2003). These LysM receptor-like kinases are involved in NF specificity as transfer of the LysM receptor-like kinases from *L. japonicus* (*NFRI* and *NFR5*) to *M. truncatula* allowed it to interact with the *L. japonicus* symbiont *M. loti* (Radutoiu *et al.*, 2007). Recognition of NF by LysM receptor-like kinases triggers a calcium ( $\text{Ca}^{2+}$ ) spiking response in the nucleus. Ehrhardt *et al.* showed that *S. meliloti* NF alone was sufficient to trigger this response in *M. sativa* (Ehrhardt *et al.*, 1996).

Nuclear  $\text{Ca}^{2+}$  spiking is central to a signal transduction pathway that integrates NF perception into physiological responses from the plant. Several components of this pathway in *M. truncatula* were identified using a large mutagenesis screen for mutants that were defective in nodule formation. Mutants of *DMI1*, *DMI2* and *DMI3* (Does Not Make Infections) and *NSP1* were shown to be defective in nodule formation in response to NF (Catoira *et al.*, 2000). It was later shown that the nuclear  $\text{Ca}^{2+}$  spiking response depends on the receptor-like kinase DMI2 and a ligand gated cation channel on the nuclear envelope DMI1 (Ané *et al.*, 2004; Endre *et al.*, 2002).  $\text{Ca}^{2+}$  spiking in the nucleus is decoded by the  $\text{Ca}^{2+}$ /calmodulin-dependent protein kinase DMI3 (CCaMK) (Lévy *et al.*, 2004). A constitutively active form of CCaMK was able to induce nodule formation in *M. truncatula* in the absence of bacterial elicitation (Gleason *et al.*, 2006). CCaMK associates with and phosphorylates IPD3 (CYCLOPS) (Horváth *et al.*, 2011), and results in the activation of two GRAS domain transcription factors NSP1 and NSP2.

Activated NSP1 and NSP2 form a complex that activates the transcription of early nodulation genes (Hirsch *et al.*, 2009), including the transcription factors NIN and ERN1 (Marsh *et al.*, 2007; Middleton *et al.*, 2007). This complex signal transduction results in dramatic physiological responses in the plant including root hair curling and the division of inner cortical cells that form the nodule primordium.

### ***1.2.2 Root hair invasion***

Asymmetric growth at the root hair tip in response to NF results in root hair curling that traps rhizobia that are colonizing the root hair surface in an apoplastic compartment between two cell walls of the curled colonized root hair (CCRH) (Figure 1.1 B). Rhizobia trapped within the CCRH continue to divide forming colonies called infection foci.

Infection is initiated by localized cell wall degradation at the site of contact between bacteria in the infection foci and the plant cell wall. It is unclear whether the rhizobia or plant are the source of the enzymes responsible for cell wall degradation. Induction of polygalacturonase in *M. sativa* was observed during infection by *S. meliloti* (Muñoz *et al.*, 1998). Bacterial cellulase mutants in *R. leguminosarum* bv. *trifolii* were deficient in the nodulation of clover (Robledo *et al.*, 2008).

Infection proceeds by a reverse of polar growth in the root hair that results in an invagination of the plasma membrane at the site of cell wall degradation. A NF induced  $\text{Ca}^{2+}$  influx at the root hair tip that is independent of nuclear calcium spiking has been observed that may be involved in the signaling that leads to the reversal of polar cell growth in *M. truncatula* that results in membrane invagination (Morieiri *et al.*, 2013).

Root hair  $\text{Ca}^{2+}$  influx may involve unique NF receptors as *nodL* mutants of *S. meliloti* that lack the NF acetyl group were impaired for root hair  $\text{Ca}^{2+}$  influx, but not nuclear  $\text{Ca}^{2+}$  spiking (Moriei *et al.*, 2013).

The invaginating plasma membrane is lined with new cell wall as it grows inward, forming a hollow tube called an infection thread (IT) (Figure 1.1 C). The IT is continually colonized by rhizobia as it traverses the root hair cell. Intracellular IT progression is accompanied by dynamic cytoskeletal rearrangements in the root hair and migration of the nucleus to the growing IT tip (Oldroyd *et al.*, 2011). Upon reaching the base of the root hair cell, localized cell wall degradation allows the IT to continue into the next cell layer. The IT continues to grow and ramify through multiple cell layers until it reaches the inner cortical cells of the root. Gage *et al.* (1996, 2002) used green fluorescent protein (GFP) and red fluorescent protein (RFP) labeled *S. meliloti* to visualize the early events of symbiosis between *S. meliloti* and *M. sativa*. They demonstrated that active growth of bacteria occurred only at the growing tip of the IT. This resulted in clonal expansion at the tip of the infection thread when inoculated with a mixed culture of *S. meliloti* expressing GFP and RFP such that only a single type of bacterium would enter the nodule (Gage, 2002; Gage *et al.*, 1996).

Successful penetration of the IT requires the continued synthesis of NF by *S. meliloti*. A *nodFL* mutant of *S. meliloti* that produces NF that lacks the acetyl group and has a modified acyl tail showed reduced and aberrant IT formation. ITs that did form were aborted before reaching the base of the root hair (Limpens *et al.*, 2003). A double mutant of *nodFE* that produces NF with a modified acyl tail showed aberrant IT formation in symbiosis with *M. truncatula* with partially depleted *LYK3* (Limpens *et al.*,

2003). Consistent with the role of NF in signaling during IT penetration, plant mutants of various components in the NF-induced  $\text{Ca}^{2+}$  spiking signaling cascade are defective in the formation of infection foci (*NIN*, *NSP1*, *NSP2*), or infection threads (*CYCLOPS*, *ERN1*) (Oldroyd *et al.*, 2011).

Another intriguing class of molecules that must be synthesized by *S. meliloti* in order to penetrate the IT are exopolysaccharides (EPS). *S. meliloti* is capable of producing two main exopolysaccharides: succinoglycan (EPSI) and galactoglucan (EPSII). However, under normal conditions *S. meliloti* strain Rm1021 only produces succinoglycan. Succinoglycan is a polymer of octosaccharide repeating units composed of one galactose residue and seven glucose residues, with pyruvyl, succinyl and acetyl modifications. The steps involved in its biosynthesis have been extensively characterized (Reuber & Walker, 1993). The repeating unit of galactoglucan is a disaccharide of glucose and galactose, decorated with pyruvyl and acetyl modifications. The genes for succinoglycan and galactoglucan biosynthesis are encoded on pSymB by the *exo* and *exp* loci, respectively (Finan *et al.*, 2001).

A mutant of *exoY*, which encodes the glycosyltransferase responsible for the first step in succinoglycan biosynthesis, is unable to initiate infection thread formation in *M. sativa* (Cheng & Walker, 1998b; Dickstein *et al.*, 1988). A mutant of *exoH*, which encodes the enzyme responsible for the addition of the succinyl modification, forms aberrant ITs that abort before reaching the base of the root hair (Cheng & Walker, 1998b). It was hypothesized that low molecular weight (LMW) forms of succinoglycan might be specifically important for invasion because EPSI synthesized in the *exoH* mutant is refractory to cleavage by the endoglycanases ExoK and ExsH into LMW forms

(York & Walker, 1998). Strains of *S. meliloti* with a functional SinI/ExpR quorum sensing system are capable of synthesis of EPSII under normal conditions. In this background, synthesis of EPSII was able to complement an *exoY* mutant for invasion of *M. sativa*, but not *M. truncatula* (Glazebrook & Walker, 1989).

It is currently unclear how EPS participates in invasion by *S. meliloti*. A microarray experiment showed increased expression of many genes associated with the plant immune response in *M. truncatula* when inoculated with an *exoY* mutant compared to wild-type *S. meliloti*. (Jones *et al.*, 2008). One hypothesis is that EPS plays a role in dampening the plant immune response, possibly as a signal that is recognized by the plant. EPS has also been shown to play a role in protection from a variety of stresses, including oxidative stress (Lehman & Long, 2013). An oxidative burst is associated with the early stages of *M. sativa* infection by *S. meliloti* (Santos *et al.*, 2001). Therefore, EPS may also play a role in tolerating different stresses experienced during infection.

A third class of macromolecule that has been implicated in invasion is the cyclic  $\beta$ -glucans. These are cyclized chains of 17-25 glucose residues, connected by  $\beta$ -1,2 linkages (Spaink, 2000). Enzymes that mediate cyclic  $\beta$ -glucan export and synthesis are chromosomally encoded by *ndvA* and *ndvB* (Nodule development) (Galibert *et al.*, 2001). Rhizobial *ndv* mutants are impaired in their ability to nodulate *M. sativa*, and form small empty pseudonodules (Dickstein *et al.*, 1988; Geremia *et al.*, 1987). They have also been reported to be defective in attachment to plant cells, and ITs formed by these mutants abort at early stage (Dylan *et al.*, 1990b). Cyclic  $\beta$ -glucans function in adaption to hypoosmotic conditions in many bacteria, including *S. meliloti*, suggesting that they may play a role in tolerating osmotic stress experienced during invasion (Dylan *et al.*, 1990a;

Miller & Wood, 1996). However, pseudorevertants in *S. meliloti ndv* mutant backgrounds that regained the ability to effectively nodulate alfalfa remained sensitive to hypoosmotic growth conditions (Dylan *et al.*, 1990b). Therefore, the role of cyclic  $\beta$ -glucans in *S. meliloti* symbiosis remains unclear.

### ***1.2.3 Bacteroid differentiation***

Rhizobia dividing at the tip of the IT enter inner cortical cells in the nodule primordium by endocytosis. This results in the acquisition of a plant-derived membrane surrounding the rhizobial cell called the symbiosome membrane (SM). The resulting organelle-like structure consisting of a rhizobium cell surrounded by a SM is referred to as a symbiosome. *M. truncatula* DMI2 is localized to the infection thread and symbiotic membranes, and knockdown studies showed that *M. truncatula* with reduced levels of DMI2 failed to release bacteria into symbiosomes (Limpens *et al.*, 2005). The space between the SM and the *S. meliloti* cell wall is referred to as the peribacteroid space. It has been shown that the peribacteroid space in *M. truncatula* is an acidic compartment (Pierre *et al.*, 2013).

Rhizobial lipopolysaccharide (LPS) may be important for interaction with the SM or tolerating conditions of the peribacteroid space. *S. meliloti* mutants of *bacA* and *lpsB* that express an altered LPS on the cell surface fail to survive the symbiosome and senesce following endocytosis (Campbell *et al.*, 2002; Glazebrook *et al.*, 1993). LPS consists of an O-antigen repeating unit, attached to a polysaccharide core that is anchored to the membrane by lipid A. Mutants of *lpsB* produce a modified LPS core with altered

polysaccharide composition (Campbell *et al.*, 2002). Mutants of *bacA* lack the ability to modify lipid A with a very long chain fatty acid moiety (Ferguson *et al.*, 2004).

Rhizobia grow and divide along with the SM until the cytoplasm of the infected host cell is packed with thousands of symbiosomes (Udvardi & Poole, 2013). Within the symbiosome, *S. meliloti* undergo a dramatic differentiation into their nitrogen fixing bacteroid form. *S. meliloti*, along with their host plant cells undergo several rounds of endoreduplication, yielding chromosome counts of 24 in *S. meliloti* bacteroids.

Differentiated bacteroids appear swollen and pleomorphic in shape and show membrane permeability. In *S. meliloti* this differentiation is said to be terminal; terminally differentiated bacteroids cannot be cultured from nodules (Oldroyd *et al.*, 2011).

Terminal differentiation is not universal among legumes, and only takes place in legumes of the inverted repeat lacking clade (IRLC), including *Medicago*, *Pisium sativum* (pea), and *Viciae faba* (faba bean). In other legumes such as *L. japonicas*, bacteroids do not undergo endoreduplication and remain culturable from nodules (Oldroyd *et al.*, 2011).

Terminal differentiation is thought to be a plant-dependent trait. This was demonstrated in an experiment where *R. leguminosarum* bv. *viciae* ( which undergoes terminal differentiation in symbiosis with its host *P. sativum*) was modified so that it could colonize *L. japonicas* nodules. During nodulation of *L. japonicus*, *R. leguminosarum* bv. *viciae* bacteroids did not undergo endoreduplication and maintained their normal size and shape. Conversely, in another experiment *R. leguminosarum* bv. *phaseoli* (not terminally differentiated in symbiosis with its host *Phaseolus vulgaris*) was modified to infect *P. sativum* and showed large, branched bacteroids and endoreduplication. Therefore, some

legumes are able to take control of the bacterial cell cycle and impose terminal differentiation upon their rhizobial symbionts (Mergaert *et al.*, 2006).

Terminal differentiation is imposed on rhizobia by small plant peptides that co-localize with bacteroids in the nodule, referred to as nodule cysteine-rich antimicrobial peptides (NCRs). NCRs are short 60-90 amino acid peptides with conserved cysteine-rich motifs. Treatment of free-living rhizobia with NCR peptides induced phenotypes consistent with those observed during terminal differentiation such as membrane permeabilization, endoreduplication and loss of viability (Van de Velde *et al.*, 2010). Microarray analysis showed that over 300 NCR peptides that are absent from *Lotus japonicus* are induced in the nodules of *M. truncatula* (Mergaert *et al.*, 2003). A signal peptidase, DNF1 is required to target NCRs to the SM, where they induce the changes responsible for terminal differentiation (Wang *et al.*, 2010). Interestingly, *bacA* mutants show resistance to antimicrobial peptides (Marlow *et al.*, 2009). BacA is a membrane component of the ABC (ATP binding cassette) transporter family, encoded on pSymB. Although BacA is involved in the transport of VLCFA, VLCFA biosynthetic mutants show unaltered resistance to antimicrobial peptides and form a successful symbiosis, suggesting an alternate role for BacA. (Oldroyd *et al.*, 2011). It has been hypothesized that BacA plays a role in the action of NCRs on *S. meliloti* (Marlow *et al.*, 2009). In the broad host range *Sinorhizobium fredii* NGR234, BacA is not required for nodule formation on legume hosts outside of the IRLC. This evidence supports the role of BacA in the terminal differentiation of IRLC legumes, perhaps through interactions with NCRs (Ardisson *et al.*, 2011).



#### ***1.2.4 Nodule development and physiology***

Nodule organogenesis is mediated by complex hormone signaling, involving the activation of cytokinin and the suppression of polar auxin transport in the root cortex (Oldroyd *et al.*, 2011). Mature nodules are composed of a central infected tissue that contains a mixture of infected, and uninfected cells, surrounded by uninfected tissues that connect to the root vascular system (Udvardi & Poole, 2013). *Medicago* species form indeterminate nodules that are elliptical in shape and contain a persistent meristem. Indeterminate nodules are organized into several zones: the meristem at the growing tip, an invasion zone that contains undifferentiated rhizobia and is the site of IT penetration into the nodule, an interzone where rhizobia undergo differentiation, a nitrogen fixation zone that contains mature nitrogen-fixing bacteroids, and a senescence zone that is absent of nitrogen fixation where bacteroids have degraded (Udvardi & Poole, 2013).

Although rhizobia are obligate aerobic organisms, bacterial nitrogenase is oxygen labile. In order to permit nitrogen fixation, nodule tissue forms a barrier to gaseous diffusion that limits oxygen entry into the nodule, forming a microaerobic environment within. Oxygen concentrations within the nodule are in the tens of nanomolar range, approximately four orders of magnitude lower than equilibrium in water. Continued cellular respiration by bacteroids under these conditions is facilitated by both plant and bacterial factors. Legumes synthesize nodule specific proteins called legume hemoglobins (leghemoglobins) that are critical to maintaining the microaerobic environment. The synthesis of these proteins results in a pink coloration of nodule tissue. Leghemoglobins are able to bind free oxygen and deliver it to bacteroids as well as plant mitochondria (Udvardi & Poole, 2013). RNA interference with the gene encoding

leghemoglobin in *L. japonicus* resulted in higher levels of free oxygen, and a complete absence of nitrogenase activity (Ott *et al.*, 2009). Bacteroids also synthesize new enzymes for microaerobic respiration, including a new, high-affinity terminal oxidase that is critical for microaerobic respiration. These enzymes as well as the genes encoding nitrogenase are under the control of a complex regulatory circuit that is induced by low oxygen concentrations (Jones *et al.*, 2007).

In *S. meliloti* this regulatory circuit is controlled by a two-component system FixL/FixJ. FixL is a sensor kinase that senses oxygen concentrations and phosphorylates the response regulator FixJ in the absence of oxygen (Gilles-Gonzalez *et al.*, 1991; Lois *et al.*, 1993). Phosphorylated FixJ activates the transcription of genes that encode nitrogenase synthesis and microaerobic respiration, including two key regulators. These include *nifA* which encodes a regulator that controls the transcription of the nitrogenase genes, and *fixK* which encodes a regulator that controls the transcription of the genes that encode the high-affinity terminal oxidase, *fixNOPQ* (Foussard *et al.*, 1997) (de Philip *et al.*, 1990).

*S. meliloti* fixes nitrogen using molybdenum (Mo)-nitrogenase. The mature holoenzyme is composed of two subunits: dinitrogen reductase and nitrogenase. Dinitrogen reductase contains the iron-molybdenum cofactor and is encoded by *nifH*, and nitrogenase is composed of  $\alpha$  and  $\beta$  subunits encoded by *nifD* and *nifK*. Several other *nif* and *fix* genes are required for the synthesis of the iron-molybdenum cofactor and the proper maturation of nitrogenase (Masson-Boivin *et al.*, 2009). Electrons flow from a donor to dinitrogen reductase, to nitrogenase where they reduce  $N_2$  to  $NH_3$ . Collectively this process requires 8 electrons and 16 molecules of ATP to reduce one molecule of  $N_2$

to 2 molecules of  $\text{NH}_3$  (ammonia). The energy for nitrogen fixation is dicarboxylic acids derived from sucrose photosynthate that are supplied to the rhizobia by the plant (See section 1.5.3). The synthesized ammonia is thought to passively diffuse into the acidic peribacteroid space where it is trapped by protonation to  $\text{NH}_4^+$  (ammonium) (Day *et al.*, 2001). Ammonium is thought to be further transported across the SM by  $\text{NH}_3$ , ion, or aquaporin-like channels, although the molecular identity of an  $\text{NH}_4^+$  channel in the SM has yet to be identified (Udvardi & Poole, 2013). Ammonium is primarily assimilated by the plant using asparagine and glutamine synthetases to convert ammonia into amino acids (Barsch *et al.*, 2006) (Cordoba *et al.*, 2003).

### 1.3 *Sinorhizobium meliloti* Central Carbon Metabolism

*S. meliloti* is able to engage in a complex symbiotic relationship with its legume hosts as discussed in the previous section. Carbon metabolism is critical for all aspects of the nitrogen fixing symbiosis. The soil is a diverse environment, and the ability to metabolize a diverse array of carbon compounds can provide a competitive advantage to rhizobia (Triplett & Sadowsky, 1992). Throughout invasion, rhizobia must utilize energy derived from catabolism to continue to actively divide, as well as to biosynthesize the macromolecules required for invasion. Finally, in the bacteroid, the high energy demand for nitrogen fixation is met through the catabolism of dicarboxylic acids.

For these reasons, carbon metabolism in *S. meliloti* has been studied for many years. An emerging picture of *S. meliloti* catabolism contains unique paradigms of metabolism compared to those established in the model organism *Escherichia coli*. The model of *S. meliloti* metabolism may more accurately reflect common metabolic schemes in the  $\alpha$ -proteobacteria. Carbon metabolism in *S. meliloti* is discussed in the following sections, including central metabolism (1.3), and peripheral sugar and sugar alcohol catabolism (1.4). The biological relevance of carbon metabolism to symbiosis is further discussed in section 1.6.

Glycolysis in *S. meliloti* proceeds through a cyclic Entner-Doudoroff (ED) pathway that utilizes the upper Embden-Meyerhoff-Parnas (EMP) pathway in a gluconeogenic manner. *S. meliloti* also contains a complete pentose phosphate (PP) pathway and tricarboxylic acid (TCA) cycle. *S. meliloti* Rm1021 is one of the few rhizobia to contain an operational Calvin-Benson-Bassham (CBB) cycle that allows formate dependent autotrophic growth. The genetic and biochemical evidence for the

functions of these pathways is described below. A summary of the predicted and verified genes that encode the enzymes in these pathways is presented in Table 1.1.

Metabolically, rhizobia have been sub-divided into fast-growing and slow-growing groups based on their growth rates. Particular metabolic schemes that correlate with either group have been identified. The slow-growing rhizobia include rhizobia isolated from legumes of tropical origin, such as *Bradyrhizobium* spp. The fast growing rhizobia include rhizobia isolated from legumes of temperate origin such as *Sinorhizobium* spp. and *Rhizobium* spp. (Stowers, 1985). In general the pathways described in this section from *S. meliloti* are representative of members of the fast-growing clade. Conversely, findings in other members of the fast growing rhizobia will be used to augment our understanding of *S. meliloti* catabolism.

### ***1.3.1 Entner-Doudoroff pathway***

It is well established that rhizobia possess an ED pathway for hexose catabolism (Stowers, 1985). The ED pathway is a widely distributed, alternative pathway of glycolysis to the EMP pathway used in *E. coli*. ED catabolism proceeds via oxidation of glucose-6-phosphate (G6P) to 6-phosphogluconolactone using G6P dehydrogenase. 6-phosphogluconolactone is resolved to 6-phosphogluconate (6PG) by 6-phosphogluconolactonase. 6PG is dehydrated using 6PG dehydratase to 2-keto-3-deoxy-6-phosphogluconate (KDPG), which is split by KDPG aldolase into glyceraldehyde-3-phosphate (G3P) and pyruvate that enter the lower EMP pathway (Figure 1.2) (Conway, 1992). 6PG dehydratase and KDPG aldolase are the two key

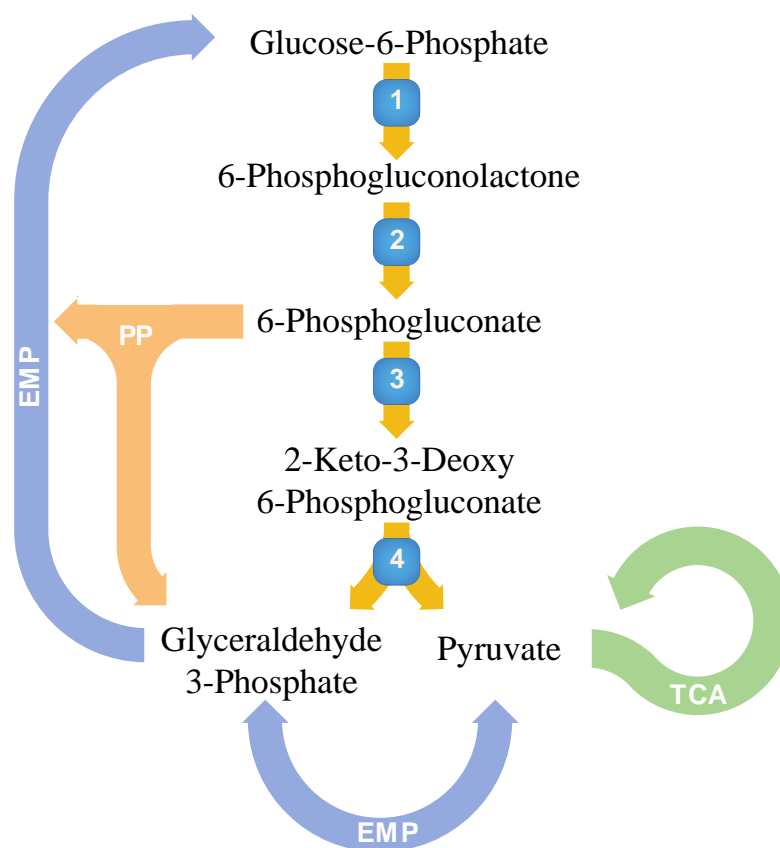
**Table 1.1.** Genes and enzymes of central metabolism in *S. meliloti*

Enzyme	Gene	Pathway	Reference(s)
Phosphoglucose isomerase (5.3.1.9)	<i>pgi</i> <sup>P</sup>	EMP; GNG	Arias <i>et al.</i> 1979
Fructose bisphosphatase (3.1.3.11)	<i>cbbF</i> <sup>P</sup>	GNG; CBB	Arias <i>et al.</i> 1979; Pickering and Oresnik, 2008
Fructose-bisphosphate aldolase (4.1.2.13)	<i>cbbA</i> <sup>P</sup> <i>cbbA2</i> <sup>P</sup> <i>fbaB</i> <sup>P</sup>	EMP; GNG; CBB	Pickering and Oresnik, 2008
Triose-phosphate isomerase (5.3.1.1)	<i>tpiA</i> <sup>D</sup> <i>tpiB</i> <sup>D</sup>	EMP; GNG; CBB	Poysti and Oresnik, 2007
Glyceraldehyde-3-phosphate dehydrogenase (1.2.1.12)	<i>gap</i> <sup>D</sup>	EMP; GNG; CBB	Finan <i>et al.</i> 1988, 1991
Phosphoglycerate kinase (2.7.2.3)	<i>pgk</i> <sup>D</sup>	EMP; GNG; CBB	Finan <i>et al.</i> 1991
Phosphoglycerate mutase (5.4.2.1)	<i>gpmA</i> <sup>P</sup> <i>gpmB</i> <sup>P</sup>	EMP; GNG; CBB	
Enolase (4.2.1.11)	<i>eno</i> <sup>D</sup>	EMP; GNG	Finan <i>et al.</i> 1988, 1991
Pyruvate kinase (2.7.1.40)	<i>pykA</i> <sup>P</sup>	EMP; GNG	
Pyruvate dehydrogenase complex (1.2.4.1), (2.3.1.12) (1.8.1.4)	<i>pdhA</i> <sup>D</sup> <i>pdhB</i> <sup>D</sup> <i>pdhC</i> <sup>D</sup> <i>lpdA1</i> <sup>P</sup>		Cabanes <i>et al.</i> 2000; Soto <i>et al.</i> 2001
Glucose-6-phosphate dehydrogenase (1.1.1.49)	<i>zwf</i> <sup>D</sup>	ED; PP	Cerveňanský and Arias 1984; Barra <i>et al.</i> 2003; Willis and Walker, 1999
6-Phosphogluconolactonase (3.1.1.31)	<i>pgl</i> <sup>P</sup>	ED; PP	Willis and Walker, 1999
6-Phosphogluconate dehydratase (4.2.1.12)	<i>edd</i> <sup>P</sup>	ED	Willis and Walker, 1999
2-Keto-3-deoxy-6-phosphogluconate aldolase (4.1.3.16/4.1.2.14), 6-Phosphogluconate dehydrogenase (1.1.1.44)	<i>eda1</i> <sup>P</sup> <i>eda2</i> <sup>P</sup>  <i>gnd</i> <sup>P</sup>	ED  PP	Martinez-De Drets and Arias, 1972, Irigoyen <i>et al.</i> 1990.
Ribose-5-phosphate isomerase (5.3.1.6)	<i>rpiA</i> <sup>P</sup> <i>rpiB</i> <sup>P</sup>	PP	Poysti and Oresnik, 2007
Ribulose-5-phosphate epimerase (5.1.3.1)	<i>rpe</i> <sup>D</sup> <i>ppe</i> <sup>P</sup>	PP	Pickering and Oresnik, 2008
Transketolase (2.2.1.1)	<i>tkt1</i> <sup>P</sup> <i>tkt2</i> <sup>P</sup> <i>cbbT</i> <sup>P</sup>	PP; CBB	

Transaldolase (2.2.1.2)	<i>tal</i> <sup>P</sup>	PP	
Phosphoribulokinase (2.7.1.19)	<i>cbbP</i> <sup>P</sup>	CBB	Pickering and Oresnik, 2008
Ribulose-1,5-bisphosphate carboxylase (4.1.1.39)	<i>cbbL</i> <sup>P</sup> <i>cbbS</i> <sup>P</sup>	CBB	Pickering and Oresnik, 2008
Citrate synthase (2.3.3.1)	<i>gltA</i> <sup>D</sup>	TCA	Mortimer <i>et al.</i> 1999
Aconitase (4.2.1.3)	<i>acnA</i> <sup>D</sup>	TCA	Koziol <i>et al.</i> 2009
Isocitrate dehydrogenase (1.1.1.42)	<i>icd</i> <sup>D</sup>	TCA	McDermott and Kahn, 1992
$\alpha$ -Ketoglutarate dehydrogenase complex (1.2.4.1), (2.3.1.61), (1.8.1.4)	<i>sucA</i> <sup>D</sup> <i>sucB</i> <sup>D</sup> <i>lpdA2</i> <sup>P</sup>	TCA	Duncan and Fraenkel, 1979; Dymov <i>et al.</i> 2004, Soto <i>et al.</i> 2001
Succinyl-CoA synthetase complex (6.2.1.5), Succinic dehydrogenase complex (1.3.99.1),	<i>sucC</i> <sup>D</sup> <i>sucD</i> <sup>D</sup> <i>sdhA</i> <sup>P</sup> <i>sdhB</i> <sup>P</sup> <i>sdhC</i> <sup>P</sup> <i>sdhD</i> <sup>P</sup>	TCA TCA	Dymov <i>et al.</i> 2004. Gardiol <i>et al.</i> 1982
Fumarase (4.2.1.2)	<i>fumC</i> <sup>P</sup>	TCA	
Malate dehydrogenase (1.1.1.37)	<i>mdh</i> <sup>D</sup>	TCA	Dymov <i>et al.</i> 2004
Phosphoenolpyruvate Carboxykinase (4.1.1.49)	<i>pckA</i> <sup>D</sup>	GNG	Østerås <i>et al.</i> 1995, 1997
NAD <sup>+</sup> Malic enzyme (1.1.1.39)	<i>dme</i> <sup>D</sup>	GNG	Driscoll and Finan, 1993, 1997
NADP <sup>+</sup> Malic enzyme (1.1.1.40)	<i>tme</i> <sup>D</sup>	GNG	Driscoll and Finan, 1996, 1997
Pyruvate orthophosphate dikinase (2.7.9.1)	<i>ppdK</i> ( <i>pod</i> ) <sup>D</sup>	GNG	Driscoll and Finan 1997, Østerås <i>et al.</i> 1997
Pyruvate carboxylase (6.4.1.1)	<i>pyc</i> <sup>D</sup>	TCA	Dunn <i>et al.</i> 2001
Isocitrate lyase (4.1.3.1)	<i>aceA</i> <sup>D</sup>	TCA	Duncan and Fraenkel, 1979; Ramírez-Trujillo <i>et al.</i> 2007
Malate synthase (2.3.3.9)	<i>glcB</i> <sup>D</sup>	TCA	Duncan and Fraenkel, 1979; Ramírez-Trujillo <i>et al.</i> 2007

Enzymes are presented that carry out the following key pathways of central metabolism in *S. meliloti*: Embden-Meyerhof-Pernas pathway (EMP), Entner Doudoroff pathway (ED), Calvin-Benson-Bassham pathway (CBB), Gluconeogenesis (GNG), Tricarboxylic acid cycle (TCA). Genes either predicted to encode these enzymes based on sequence homology (<sup>P</sup>), or that have been shown to encode them based on carbon utilization, and or enzyme activity assays (<sup>D</sup>) are included with systematic identifier numbers (Sys Id Num).

**Figure 1.2.** Schematic of the *S. meliloti* ED pathway



Yellow arrows correspond to the enzymatic reactions of the ED pathway. Enzymes are represented by circled numbers that correspond to: 1) glucose-6-phosphate dehydrogenase; 2) 6-phosphogluconolactonase; 3) 6-phosphogluconate dehydratase; 4) 2-keto-3-deoxy-6-phosphogluconate aldolase. Bold auxiliary arrows are used to represent interactions with the EMP pathway (blue), PP pathway (orange) and the TCA cycle (green). Arrows indicate the direction of carbon flow as described by the literature. See text for details.



enzymes of the ED pathway that direct carbon from hexoses into assimilation via pyruvate, rather than through the pentose phosphate pathway (see section 1.3.3).

Early studies demonstrated the presence of enzymes for the ED pathway in a broad range of rhizobial species. ED enzymes were found at particularly highly levels in fast-growing rhizobia (Martinez de Drets *et al.*, 1974). Enzymes of the ED pathway are highly induced in *S. meliloti* during growth with glucose compared to succinate (Finan *et al.*, 1988; Irigoyen *et al.*, 1990). Consistent with its role as the main pathway of hexose assimilation, the ED genes have been shown to be among the most highly expressed genes during growth with glucose in *S. meliloti* (Barnett *et al.*, 2004).

Mutants have been isolated in *S. meliloti* that lacked both NAD<sup>+</sup> and NADP<sup>+</sup> linked G6P dehydrogenase activities (Cerveñansky & Arias, 1984). These mutants were unable to grow using glucose, fructose, ribose, xylose, mannitol, or sorbitol as sole carbon sources, but were able to grow using gluconeogenic carbon sources such as succinate (Cerveñansky & Arias, 1984). Notably, they were also able to grow using galactose or L-arabinose as a sole carbon source, supporting alternate modes of assimilation for these sugars (see sections 1.4.2.2 and 1.4.3.1). The inability of mutants to grow using most hexoses and pentoses supports the role of the ED pathway as the primary mode of glycolysis. Some mutants that lacked G6P dehydrogenase activity also lacked 6PG dehydratase /KDPG aldolase activity, suggesting that the dedicated ED enzymes may be encoded in the same genetic locus as G6P dehydrogenase (Cerveñansky & Arias, 1984). Such a locus was identified adjacent to the genes for  $\alpha$ -glucoside utilization (Willis & Walker, 1999). The locus contains *edd*, predicted to encode 6PG dehydratase, *pgl*, predicted to encode 6-phosphogluconolactonase and *zwf*, predicted to

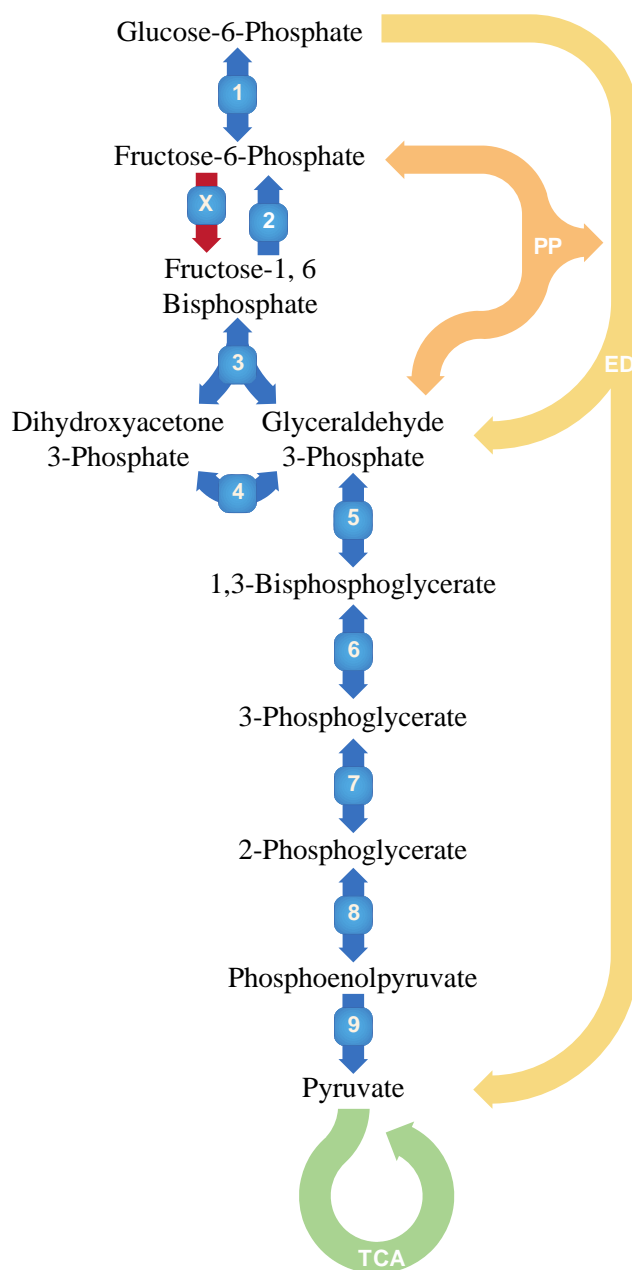
encode G6P dehydrogenase (Table 1.1). A strain carrying a *zwf* mutation was later shown to lack G6P dehydrogenase activity, reinforcing this region as an ED locus in *S. meliloti* (Barra *et al.*, 2003). A gene encoding KDPG aldolase has not been experimentally identified, but two putative KDPG aldolase genes are encoded on the chromosome (*eda1*, *eda2*) (Table 1.1) (Galibert *et al.*, 2001).

Several lines of evidence point to the functioning of a cyclic ED pathway in *S. meliloti*, where G3P is cycled through the upper EMP into hexoses that can be used for biosynthesis or recycled through the ED pathway. The cyclic ED pathway is in contrast to the linear ED pathway of *E. coli* (mainly for the catabolism of gluconate), where the end product G3P is catabolized to pyruvate through the lower EMP pathway (Conway, 1992). Experiments by two different groups using chromatography-mass spectrometry and nuclear magnetic resonance with labeled carbon compounds clearly demonstrated that some of the G3P produced by ED is converted back into high-molecular weight compounds through the upper EMP pathway (Fuhrer *et al.*, 2005; Portais *et al.*, 1999). These studies also confirmed that glucose was degraded primarily through the ED pathway, and that glycolysis through the EMP pathway was absent (Fuhrer *et al.*, 2005; Portais *et al.*, 1999). A survey of carbon flux in a number of microorganisms identified a cyclic ED pathway in *S. meliloti*, as well as pseudomonads. Further, the ED pathway was the predominant route of glucose catabolism in most microorganisms that possessed both ED and EMP pathways (Fuhrer *et al.*, 2005). Early studies also identified a cyclic ED pathway in slow-growing rhizobia (Stowers, 1985).

### 1.3.2 Embden-Meyerhoff-Parnas pathway

Evidence is consistent with the EMP pathway being used in a gluconeogenic role rather than a glycolytic role in *S. meliloti* (Figure 1.3). The dedicated glycolytic enzyme of the EMP, phosphofructokinase, has not been detected in cell-free extracts of *S. meliloti* (Arias *et al.*, 1979; Irigoyen *et al.*, 1990). Consistent with the absence of detectable phosphofructokinase activity, the genome is not predicted to contain an ATP-dependent *pfk* gene (Capela *et al.*, 2001). The remaining enzymes of EMP that are sufficient for gluconeogenesis are predicted to be encoded by the genome (Table 1.1) (Galibert *et al.*, 2001). The activities of many of these enzymes have been detected, including phosphoglucose isomerase, fructose-bisphosphate aldolase, triose-phosphate isomerase, G3P dehydrogenase, 3-phosphoglycerate kinase, phosphoglycerate mutase and enolase (Arias *et al.*, 1979; Finan *et al.*, 1988; Irigoyen *et al.*, 1990). In contrast to the enzymes of the ED pathway, which showed substantial induction during growth with glucose compared to succinate, the EMP pathway enzymes showed similar levels of expression during growth with either carbon source (Arias *et al.*, 1979; Finan *et al.*, 1988; Irigoyen *et al.*, 1990).

Mutants of genes in the upper and lower EMP pathway have been isolated. Mutants that were deficient in phosphoglucose isomerase were unable to grow using carbon sources that enter central metabolism either directly through fructose-6-phosphate (F6P), including fructose, mannose, sorbitol and mannitol, or indirectly through the PP pathway, including ribose and xylose (see section 1.4) (Arias *et al.*, 1979). Growth phenotypes on these carbon sources are consistent with their catabolism through the cyclic ED pathway following conversion of F6P to G6P by phosphoglucose isomerase.

**Figure 1.3.** Schematic of the *S. meliloti* EMP pathway

Blue arrows correspond to the enzymatic reactions of the EMP pathway. Enzymes are represented by circled numbers that correspond to: 1) phosphoglucose isomerase; 2) fructose biphosphatase; 3) fructose-bisphosphate aldolase; 4) triose-phosphate isomerase; 5) glyceraldehyde-3-phosphate dehydrogenase; 6) phosphoglycerate kinase; 7) phosphoglycerate mutase; 8) enolase; 9) pyruvate kinase. The red arrow represents the absence of phosphofructokinase in *S. meliloti*. Bold auxiliary arrows are used to represent interactions with the ED pathway (yellow), PP pathway (orange) and the TCA cycle (green). Arrows indicate the direction of carbon flow as described by the literature. See text for details.

Several mutants have been isolated in genes encoding the enzymes of the lower EMP. Carbon utilization phenotypes in these mutants demonstrate the importance of the lower EMP in the gluconeogenesis of carbon sources that enter central metabolism through the TCA cycle. Mutants of enolase, G3P dehydrogenase and 3-phosphoglycerate kinase were isolated based on a genetic screen for the inability to utilize succinate as a sole carbon source. These mutants were unable to grow using TCA cycle intermediates or pyruvate as sole carbon sources (Finan *et al.*, 1988). Mutants were capable of growth using glucose as a sole carbon source, consistent with glycolysis occurring through the ED pathway. However, growth using glucose as a sole carbon source was slower in these mutants than in the wild type, suggesting that although the EMP is not required for glycolysis, it may be advantageous to catabolize some of the G3P synthesized by the ED pathway through the lower EMP pathway. Based on the absence of 3-phosphoglycerate kinase activity in some G3P dehydrogenase mutants it was suggested that *gap* and *pgk* may be located in the same genetic locus (Finan *et al.*, 1988). The genome sequence of *S. meliloti* Rm1021 supports this hypothesis (Table 1.1) (Galibert *et al.*, 2001).

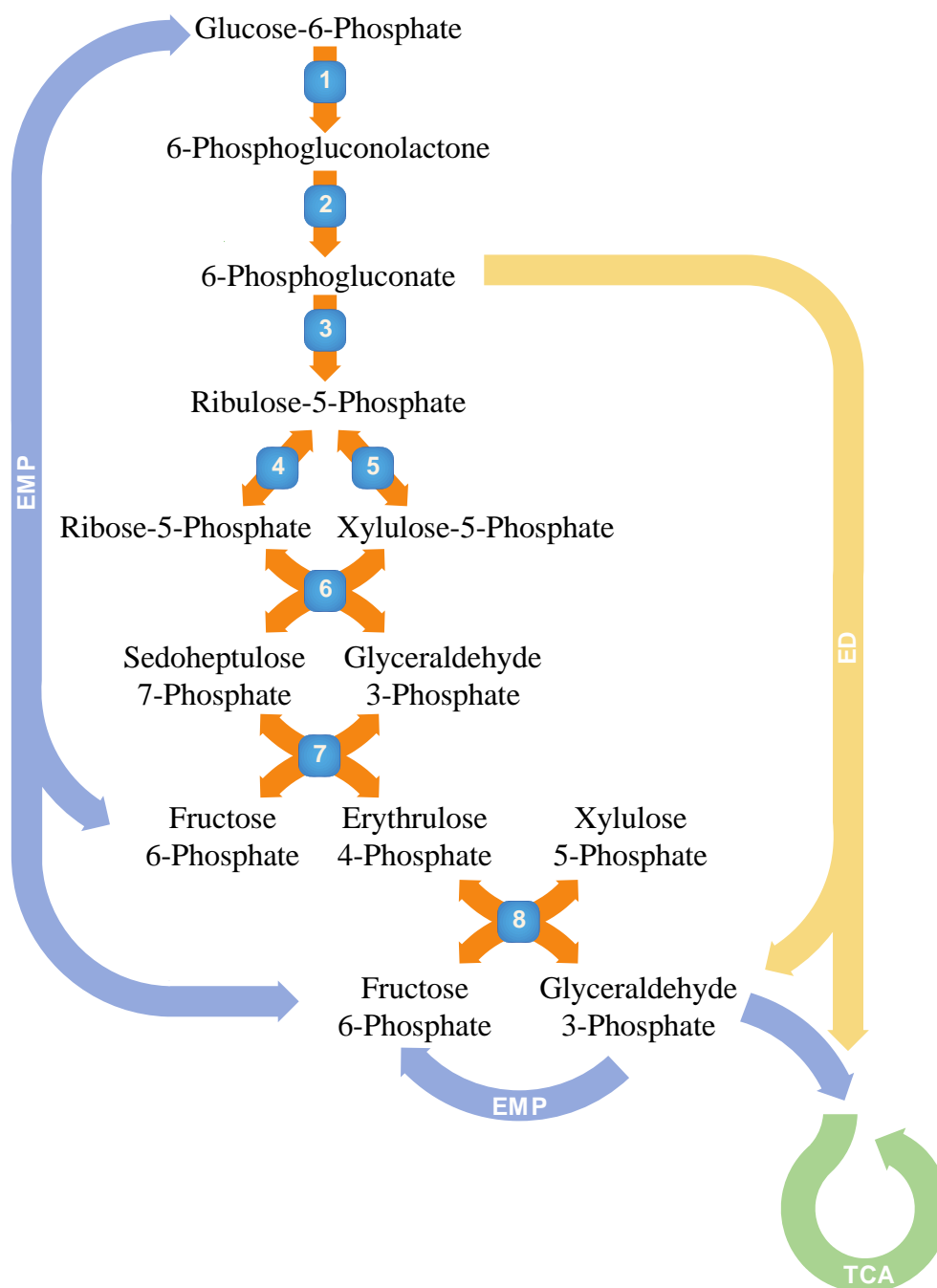
Unlike *E. coli* and *B. subtilis*, *S. meliloti* and many fast-growing rhizobia contain two triose-phosphate isomerase genes. Mutants in both of the triose-phosphate isomerase genes encoded in the genome have been isolated (*tpiA* and *tpiB*) (Poysti & Oresnik, 2007). Both genes were shown to encode functional triose-phosphate isomerases that were required for the catabolism of carbon sources that enter central metabolism through dihydroxyacetone-3-phosphate (DHAP). TpiA was required for growth using glycerol and rhamnose as sole carbon sources, whereas TpiB was specifically required for

erythritol catabolism (see sections 1.4.4.1, and 1.4.5.4 and 1.4.5.5). A double mutant was unable to grow on gluconeogenic carbon sources such as succinate, demonstrating that they are required for gluconeogenesis (Poysti & Oresnik, 2007). Labeled carbon experiments also demonstrated the cycling of G3P synthesized during the catabolism of glucose by the cyclic ED pathway through triose-phosphate isomerase and fructose-bisphosphate aldolase (Fuhrer *et al.*, 2005; Portais *et al.*, 1999).

Taken together, the EMP pathway performs several roles in central metabolism. It is required for gluconeogenesis of carbon sources that enter central metabolism through the TCA cycle. The EMP pathway also participates in the cycling of G3P synthesized by the ED pathway during growth on hexoses, as well the conversion of pentoses and hexoses that enter central metabolism through F6P into G6P for glycolysis by the ED pathway.

### ***1.3.3 Pentose phosphate pathway***

The pentose phosphate (PP) pathway is comprised of oxidative and non-oxidative branches. The first two steps of the oxidative branch are shared with the ED pathway. G6P dehydrogenase and 6-phosphogluconolactonase synthesize 6PG from G6P. The dedicated enzyme of the PP pathway is 6PG dehydrogenase. 6PG dehydrogenase catalyzes the irreversible decarboxylation of 6PG to ribulose-5-phosphate (R5P) (Figure 1.4). The presence or absence of 6PG dehydrogenase activity has been used to classify rhizobia as fast- or slow-growers. Fast-growing rhizobia, including *S. meliloti* contain a complete PP pathway as indicated by the presence of 6PG dehydrogenase, while most slow-growing rhizobia do not (Martínez de Drets & Arias, 1972).

**Figure 1.4.** Schematic of the *S. meliloti* PP pathway

Orange arrows correspond to the enzymatic reactions of the PP pathway. Enzymes are represented by circled numbers that correspond to: 1) glucose-6-phosphate dehydrogenase; 2) 6-phosphogluconolactonase; 3) 6-phosphogluconate dehydrogenase; 4) ribose-5-phosphate isomerase; 5) ribulose-5-phosphate epimerase; 6) transketolase; 7) transaldolase; 8) transketolase. Bold auxiliary arrows are used to represent interactions with the ED pathway (yellow), EMP pathway (blue) and the TCA cycle (green). Arrows indicate the direction of carbon flow as described by the literature. See text for details.

The non-oxidative branch of the PP pathway is a series of reversible reactions that interchange R5P with F6P and G3P (Figure 1.2). The activities of transketolase and transaldolase have been shown to be present in cell-free extracts of *S. meliloti* (Cerveñansky & Arias, 1984). A proteomic study found a probable transketolase and ribose-5-phosphate epimerase to be expressed under free living conditions as well as in the bacteroid (Djordjevic, 2004).

The active cycling of the PP pathway has been demonstrated in radiolabeled carbon experiments (Fuhrer *et al.*, 2005; Gosselin *et al.*, 2001; Portais *et al.*, 1999). Complicated labeling patterns were observed that were consistent with the reversible action of the non-oxidative branch of the PP pathway using  $^{13}\text{C}$ -NMR with glucose or fructose (Gosselin *et al.*, 2001; Portais *et al.*, 1999). Flux through the PP pathway to supply PP intermediates for biosynthesis was also observed using gas chromatography with radiolabelled carbon (Fuhrer *et al.*, 2005).

Overall, the study of the PP pathway is limited in *S. meliloti*. The genes encoding the PP pathway have not been characterized, and the regulation of the PP pathway is unclear. Possible candidates for the genes encoding all components of the PP pathway can be identified in the genome sequence and are presented in Table 1.1.

#### **1.3.4 Calvin-Benson-Bassham cycle**

The ability to fix carbon dioxide autotrophically is generally associated with the presence of ribulose-1,5-bisphosphate carboxylase activity. Across the Rhizobiaceae only *Bradyrhizobium* and *Sinorhizobium* have been shown to have this activity (Hanus *et al.*, 1979; Manian & O'Gara, 1982b; Pickering & Oresnik, 2008). Whereas *B. japonicum* is



capable of chemolithoautotrophic growth (Hanus *et al.*, 1979), *S. meliloti* is capable of fixing carbon dioxide is contingent upon the presence of another form of reduced carbon (Manian & O'Gara, 1982a, b; Pickering & Oresnik, 2008). The genes *cbbFPTALSX* and *ppe* predicted to encode determinants of the CBB pathway are found on pSymB along with a presumed LysR type positive regulator, *cbbR*, that is divergently transcribed. Together these genes could encode all the enzymes necessary for a traditional CBB pathway except for those that are shared with lower portion of the EMP pathway or PP pathway. In support of this hypothesis, growth in a medium containing formate and bicarbonate was shown to be dependent upon the *cbb* operon (Pickering & Oresnik, 2008). Formate mediated autotrophic growth was also dependent on one of the two triose-phosphate isomerase genes (*tpiA* or *tpiB*) found in the *S. meliloti* genome, suggesting that fixed carbon is removed from the cycle by the conversion of GAP to DHAP and condensing this to form fructose-1,6-bisphosphate (Pickering & Oresnik, 2008).

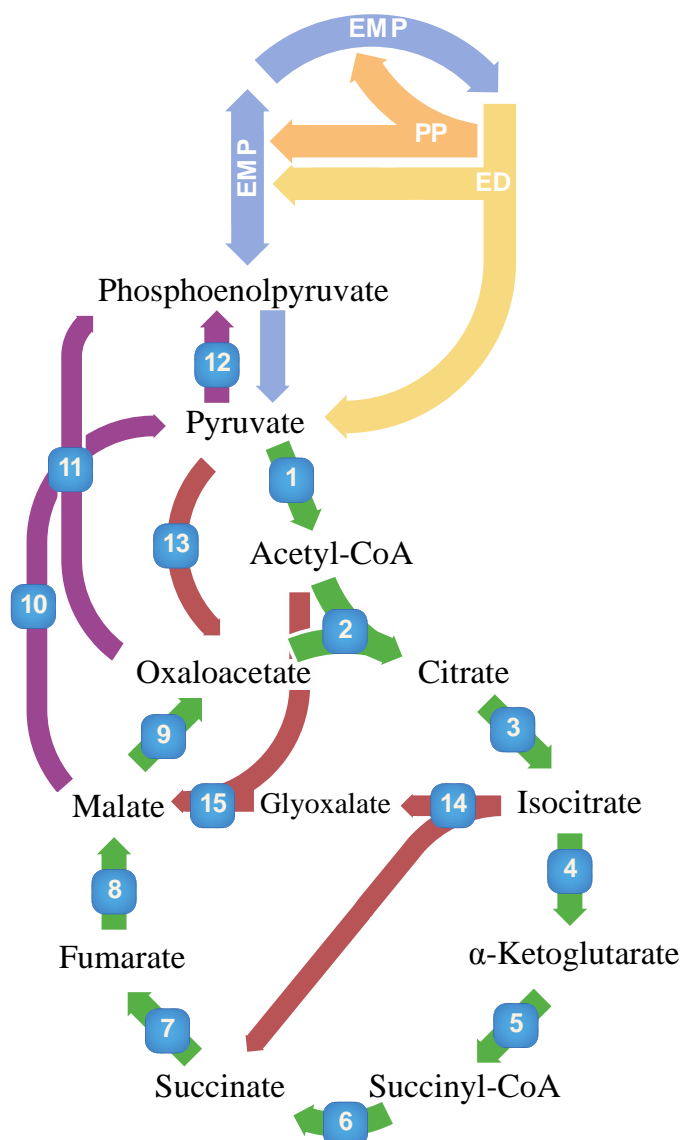
The annotation of the Rm1021 genome suggests that three formate dehydrogenases are present (*fdoGHI*, *fdsABCDG* and *SMa0487*) (Barnett *et al.*, 2001). It was demonstrated that the *fdsABCDG* locus was absolutely necessary for formate-dependent autotrophic growth, whereas a mutation that affected the pSymA encoded *fdoGHI* locus reduced the amount of carbon fixed to about 40% that of the wild type (Pickering & Oresnik, 2008). The third annotated formate dehydrogenase encoding gene, *SMa0487*, did not appear to have formate dehydrogenase activity, and did not effect autotrophic growth (Pickering & Oresnik, 2008).

### 1.3.5 Tricarboxylic acid cycle

*S. meliloti* is known to contain a complete tricarboxylic acid (TCA) cycle (Figure 1.5) (Dunn, 1998; Stowers, 1985). All the enzymes of the TCA cycle have been shown to be expressed under lab growth conditions, and in symbiotic association (Djordjevic, 2004). Mutants that were deficient in  $\alpha$ -ketoglutarate dehydrogenase (Duncan & Fraenkel, 1979) and succinate dehydrogenase (Gardiol *et al.*, 1982) activities have been previously isolated. More recently the genes encoding many of the TCA cycle components in *S. meliloti* have been identified (Table 1.1). Citrate synthase, aconitase and isocitrate dehydrogenase are encoded by *gltA*, *acnA* and *icd* respectively (Koziol *et al.*, 2009; McDermott & Kahn, 1992; Mortimer *et al.*, 1999). Knockouts of each of these genes have been constructed and shown to lack the associated enzyme activity, suggesting that they uniquely encode each of these enzymes in *S. meliloti* (Koziol *et al.*, 2009; McDermott & Kahn, 1992; Mortimer *et al.*, 1999). These three genes appear scattered throughout the chromosome as ORFan genes in *S. meliloti* (Galibert *et al.*, 2001).

Random mutagenesis identified a locus that contains many of the remaining TCA cycle genes. (Dymov *et al.*, 2004). In this study a Tn5-*tacI* mutant was isolated in *mdh* which was shown to encode malate dehydrogenase. The genes encoding succinyl-CoA synthetase (*sucCD*) and the E<sub>1</sub>/E<sub>2</sub> components of  $\alpha$ -ketoglutarate dehydrogenase (*sucAB*) complexes were shown to be encoded immediately downstream of *mdh* (Dymov *et al.*, 2004). Although not co-transcribed, also nearby in the region are the genes predicted to encode succinic dehydrogenase complex (*sdhABCD*) and the E<sub>3</sub> component of

**Figure 1.5.** Schematic of the *S. meliloti* TCA cycle, gluconeogenesis and anaplerotic pathways



Green arrows correspond to the enzymatic reactions of the TCA cycle. Purple arrows correspond to key enzymatic reactions of gluconeogenesis. Red arrows correspond to enzymatic reactions in anaplerotic pathways. Enzymes are represented by circled numbers that correspond to: 1) pyruvate dehydrogenase; 2) citrate synthase; 3) aconitase; 4) isocitrate dehydrogenase; 5)  $\alpha$ -ketoglutarate dehydrogenase; 6) succinyl-CoA synthetase; 7) succinate dehydrogenase; 8) fumarase; 9) malate dehydrogenase; 10) malic enzyme; 11) phosphoenolpyruvate carboxykinase; 12) pyruvate orthophosphate dikinase; 13) pyruvate carboxylase; 14) isocitrate lyase; 15) malate synthase. Bold auxiliary arrows are used to represent interactions with the ED pathway (yellow), EMP pathway (blue) and the PP pathway (orange). Arrows indicate the direction of carbon flow as described by the literature. See text for details.

$\alpha$ -ketoglutarate dehydrogenase: dihydrolipoyl dehydrogenase (*lpdA2*) are also in close proximity. The genes encoding the pyruvate dehydrogenase complex (*pdhABC*) are co-transcribed elsewhere on the chromosome (Cabanes *et al.*, 2000). Although dihydrolipoyl dehydrogenase can be shared between the pyruvate dehydrogenase and  $\alpha$ -ketoglutarate dehydrogenase complexes (Mattevi *et al.*, 1992), a second dihydrolipoyl dehydrogenase (*lpdA1*) is encoded near *pdhABC* and a mutation that was polar on its transcription abolished pyruvate dehydrogenase activity (Soto *et al.*, 2001).

### 1.3.6 Gluconeogenesis

Growth on TCA cycle intermediates, or carbon sources that enter central metabolism via the TCA cycle, is carried out by gluconeogenesis. Gluconeogenesis requires the EMP pathway as described in section 1.3.2, as well as a dedicated gluconeogenic enzyme to synthesize intermediates in the lower EMP pathway from TCA cycle intermediates. One such enzyme that has been identified in *S. meliloti* and other rhizobia is phosphoenolpyruvate (PEP) carboxykinase (Dunn, 1998). PEP carboxykinase catalyzes the conversion of oxaloacetate to PEP (Figure 1.5). PEP carboxykinase activity has been demonstrated in *S. meliloti*, and mutants that were deficient in PEP carboxykinase activity grew very slowly using gluconeogenic carbon sources (Finan *et al.*, 1988). PEP carboxykinase was shown to be encoded by *pckA* in *S. meliloti* (Table 1.1) (Østerås *et al.*, 1995).

An alternative gluconeogenic route is through the cooperative action of pyruvate orthophosphate dikinase and malic enzyme (Figure 1.5). Pyruvate orthophosphate dikinase catalyzes the ATP-dependent conversion of pyruvate to PEP. Increased

pyruvate orthophosphate dikinase activity was associated with suppressors to *pckA* mutants for growth on succinate (Østerås *et al.*, 1997). Pyruvate orthophosphate dikinase is encoded by *pod* (Table 1.1). Although a *pod* mutant alone showed no significant phenotype, a double mutant of *pckA* and *pod* was completely unable to utilize gluconeogenic carbon sources (Østerås *et al.*, 1997). Malic enzyme catalyzes the decarboxylation of malate to form pyruvate. *S. meliloti* contains both an NAD<sup>+</sup> dependent and an NADP<sup>+</sup> dependent malic enzyme (Driscoll & Finan, 1993, 1996, 1997). These are encoded by *dme* and *tme*, respectively (Table 1.1) (Driscoll & Finan, 1993, 1996). The suppression of a *pckA* mutant by PEP carboxykinase was dependent on the presence of *dme* (Østerås *et al.*, 1997). However, *dme* and *tme* mutants alone show no significant carbon utilization phenotype suggesting that PEP carboxykinase is the predominant gluconeogenic route in free-living *S. meliloti*. Interestingly, evidence is consistent with the use of malic enzyme rather than PEP carboxykinase for gluconeogenesis in the bacteroid (see section 1.6.3) (Driscoll & Finan, 1993, 1997; Finan *et al.*, 1988).

### 1.3.7 Anaplerotic pathways

Glycolytic growth requires an anaplerotic enzyme to resupply the TCA cycle with intermediates that are removed during amino acid biosynthesis or gluconeogenesis. Pyruvate carboxylase fulfills this role in *S. meliloti* (Dunn *et al.*, 2001). Pyruvate carboxylase catalyzes the carboxylation of pyruvate to form oxaloacetate (Figure 1.5). Pyruvate carboxylase is encoded by *pyc* in *S. meliloti*. Consistent with its role as an anaplerotic enzyme, *pyc* mutants were unable to grow using glucose or pyruvate as sole

carbon sources, but were capable of growth using succinate as a sole carbon source (Dunn *et al.*, 2001).

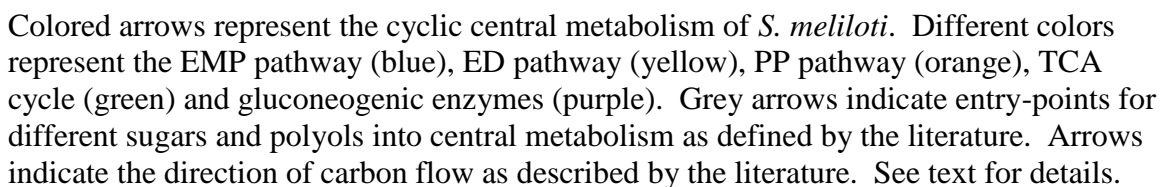
Growth with acetate as a carbon source also requires an anaplerotic pathway in order to bypass the decarboxylating steps of the TCA cycle to allow a net accumulation of carbon. In *S. meliloti* this is accomplished by the glyoxylate shunt (Figure 1.5). The glyoxylate shunt involves the cleavage of isocitrate to form succinate and glyoxylate by isocitrate lyase, followed by the synthesis of malate from glyoxylate and acetyl-CoA by malate synthase (Dunn, 1998). Both of these enzymes have been detected and are expressed during growth with acetate (Duncan & Fraenkel, 1979). These enzymes are encoded by *aceA* and *glcB*, respectively, and mutations in these genes abolished isocitrate lyase and malate synthase activities in *S. meliloti* (Ramírez-Trujillo *et al.*, 2007).

## 1.4 *Sinorhizobium meliloti* Peripheral Carbon Metabolism

The ability to utilize a broad range of carbon sources is important for survival in a diverse array of growth conditions. Rhizobia must be able to survive in the complex environment of the soil, intracellularly during infection of the plant and throughout the symbiosis. It is therefore unsurprising that a large portion of its genome is dedicated to carbon catabolism and transport (Galibert *et al.*, 2001; Mauchline *et al.*, 2006). This section discusses the peripheral pathways of metabolism of some of these sugars and sugar alcohols, as defined in *S. meliloti*. The predicted entry-points of these sugars and sugar alcohols into the pathways of central metabolism described in the previous section are presented in Figure 1.6. The biological relevance of the ability to catabolize these substrates is further discussed in section 1.5.

### 1.4.1 Disaccharide catabolism and transport

An early feature that differentiated fast-growing rhizobia from those that were slow-growing was the capacity to utilize disaccharides. Studies showed that fast growing rhizobia including *Sinorhizobium* and *Rhizobium* spp. were capable of utilizing lactose, sucrose, maltose, trehalose and cellobiose as sole carbon sources, whereas slow-growers such as *Bradyrhizobia* were not (Glenn & Dilworth, 1981; Jordan, 1984; Martinez de Drets *et al.*, 1974; Martínez de Drets & Arias, 1972; Stowers, 1985). More recently targeted and genomic approaches have identified many of the loci involved in  $\beta$ - and  $\alpha$ -galactoside and glucoside catabolism in *S. meliloti*.





#### ***1.4.1.1 $\beta$ -galactosides (lactose)***

Two different  $\beta$ -galactosidase activities have been demonstrated in *S. meliloti*, however only one of the activities was inducible by lactose. The inducible  $\beta$ -galactosidase was shown to be required for growth using lactose as a sole carbon source (Niel *et al.*, 1977). Lactose uptake has been shown to be inducible in *S. meliloti* and other fast-growing rhizobia (Ucker & Signer, 1978). Uptake of lactose was not inhibited by sucrose, maltose, trehalose or cellobiose suggesting that it uses a unique transporter (Glenn & Dilworth, 1981). The genes for lactose catabolism are carried on the megaplasmid pSymB and are adjacent to the locus for  $\alpha$ -galactoside utilization (Charles & Finan, 1991; Charles *et al.*, 1990). The lactose locus contains the  $\beta$ -galactosidase *lacZ*, genes that presumably encode a lactose ABC transporter, *lacEFGK*, and a LacI type negative regulator, *lacR* (Jelesko & Leigh, 1994). Lactose regulation proceeds differently than *E. coli* in *S. meliloti* as  $\beta$ -galactosidase activity is not inducible by IPTG (Ucker & Signer, 1978).

#### ***1.4.1.2 $\alpha$ -galactosides (melibiose and raffinose)***

The locus required for the utilization of  $\alpha$ -galactosides including melibiose and raffinose is located on pSymB adjacent to the  $\beta$ -galactoside locus (Charles & Finan, 1991; Charles *et al.*, 1990; Gage & Long, 1998). The locus contains an ABC transporter (*agpABCD*) and two putative  $\alpha$ -galactosidases, *melA* and *agaL2*. The ABC transporter was required for growth on melibiose and raffinose, but not galactose or glucose. Further, a mutant of the gene encoding the periplasmic binding protein *agpA* was shown to be unable to transport raffinose. Transcription of *agpA* was downregulated by SyrA, and

upregulated by galactose and  $\alpha$ -galactosides (Gage & Long, 1998). Induction of genes in the locus by galactose and  $\alpha$ -galactosides is mediated by an AraC-type transcriptional regulator encoded by the upstream gene *agpT*. Mutants of *agpT* were unable to grow using melibiose or raffinose as sole carbon sources (Bringhurst & Gage, 2000).

#### ***1.4.1.3 $\alpha$ - and $\beta$ -glucosides (sucrose, maltose, trehalose and cellobiose)***

Early studies of disaccharide uptake and hydrolysis provided evidence that the uptake of sucrose and maltose was facilitated by an active process, and that sucrose invertase activity was inducible in *S. meliloti* (Glenn & Dilworth, 1981). The  $\alpha$ -glucosides sucrose, maltose and trehalose were able to compete for active uptake with sucrose and maltose, whereas lactose or the  $\beta$ -glucoside cellobiose were not (Glenn & Dilworth, 1981). This early work suggested that sucrose, maltose and trehalose share a common uptake system, whereas cellobiose may be transported by a separate system. A locus that encoded  $\alpha$ -glucoside transport and catabolism was isolated on an *S. meliloti* cosmid that conferred the ability to utilize sucrose, maltose and trehalose to *Ralstonia eutropha* (Willis & Walker, 1998). The locus consisted of a set of genes that map to the chromosome of *S. meliloti* annotated as *aglEFGAK*, where *aglEFGK* and *aglA* are predicted to encode an ABC transport system and  $\alpha$ -glucosidase respectively. However, *S. meliloti* strains carrying mutations in the *agl* region were not impaired in  $\alpha$ -glucoside utilization, even in backgrounds that also carried mutations in the *agp*  $\alpha$ -galactoside region (Willis & Walker, 1999).

Evidence for redundancy in  $\alpha$ -glucoside transport systems has more recently been presented; a locus was isolated by Tn5 mutagenesis based on an impaired ability to utilize

trehalose as a sole carbon source. The locus containing the Tn5 mapped to pSymB and contained four putative ABC transport genes *thuEFGK* (Jensen *et al.*, 2002). Uptake experiments in a *thuE* mutant background showed reduced uptake of radiolabelled sucrose, maltose and trehalose. Furthermore, a double mutant of *thuE*, *aglE* was abolished for sucrose, trehalose and maltose utilization showing that *aglEFGK* and *thuEFGK* together encode the two major  $\alpha$ -glucoside transport systems in *S. meliloti* (Jensen *et al.*, 2002).

Further work characterizing trehalose catabolism showed that mutants of two putative catabolic genes in the locus, *thuAB* were strongly impaired in their ability to utilize trehalose as a sole carbon source (Jensen *et al.*, 2005). These genes had little similarity to genes encoding known enzymes of trehalose catabolism and were later shown to encode enzymes for a novel pathway of trehalose catabolism through a 3-ketotrehalose intermediate (Ampomah *et al.*, 2013; Avetisyan *et al.*, 2013). The transport genes *thuEFGK* and catabolic genes *thuAB* were further shown to be involved in the transport and catabolism of maltitol and the sucrose isomers leucrose, palatinose and trehalulose (Ampomah *et al.*, 2013).

#### ***1.4.2 Hexose catabolism and transport***

The catabolism of most hexoses requires a functional cyclic ED pathway. This is illustrated by mutations in G6P dehydrogenase that are unable to grow using fructose, mannose or glucose as sole carbon sources. Galactose catabolism does not follow this scheme, nor does it enter glycolysis through the Leloir pathway as in *E. coli*. Rather it is

catabolized through an analogous pathway to the ED pathway: the De Ley-Doudoroff pathway (Arias & Cerveñansky, 1986).

#### **1.4.2.1 Glucose**

The most direct point of entry for glucose into the ED pathway is phosphorylation by hexose kinase (Figure 1.6). However, glucose can also be oxidized by a gluconate bypass which consists of a periplasmic pyrroloquinoline quinone-dependent glucose dehydrogenase and gluconate kinase (Gosselin *et al.*, 2001; Portais *et al.*, 1999).

Although this pathway exists, glucose dehydrogenase mutants did not have a slower growth rate than wild type when grown with glucose as a sole carbon source (Gosselin *et al.*, 2001). Growth experiments in chemostat cultures containing glucose and succinate indicated that glucose is preferentially oxidized while gluconate accumulates in the culture (Bernardelli *et al.*, 2001). A genetic locus that encodes an ATP-independent periplasmic transporter that is required for gluconate utilization has been identified on pSymA (*gntABC*) (Steele *et al.*, 2009).

#### **1.4.2.2 Galactose**

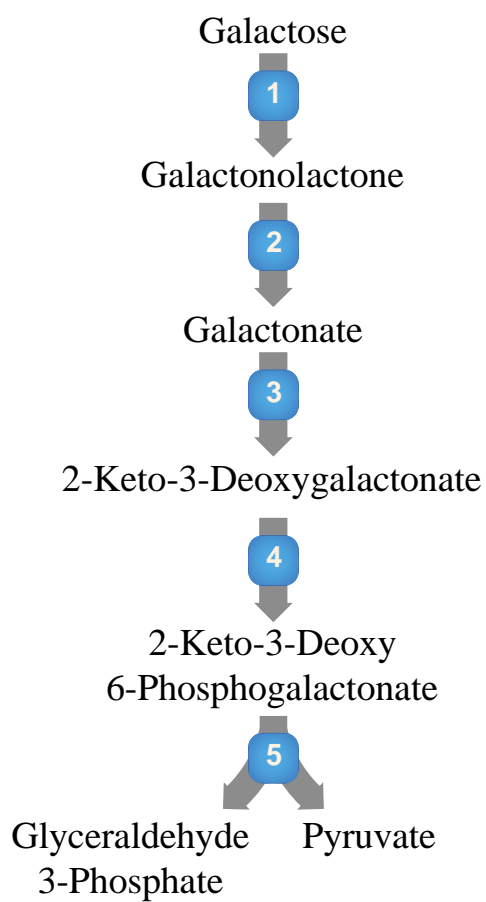
Based on the observation that a G6P dehydrogenase mutant was unable to grow with most hexoses, but was able to grow with galactose, it was hypothesized that the De Ley-Doudoroff pathway was used in *S. meliloti* (Cerveñansky & Arias, 1984). The De Ley-Doudoroff pathway of galactose catabolism was first discovered in *Pseudomonas saccharophila* (De Ley & Doudoroff, 1957; Lessie & Phibbs, 1984). Galactose catabolism proceeds through several non-phosphorylated intermediates in analogy to the

ED pathway. Briefly, galactose is oxidized to galactono- $\gamma$ -lactone, and resolved to galactonate by a lactonase. Galactonate is dehydrated to 2-keto-3-deoxygalactonate. 2-keto-3-deoxygalactonate is phosphorylated forming 2-keto-3-deoxy-6-phosphogalactonate (KDPGal), which is split by an aldolase reaction into G3P and pyruvate (Figure 1.7). Cell free extracts of *S. meliloti* L5-30 were assayed for the presence of four out of five of the enzymes required for the De Ley-Doudoroff pathway (Arias & Cerveñansky, 1986). Activities of each of these enzymes were demonstrated, consistent with De Ley-Doudoroff pathway operation. A chemically induced mutant unable to grow using galactose as a sole carbon source was found to be missing one KDPGal aldolase activity (Arias & Cerveñansky, 1986).

#### ***1.4.2.3 Fructose and mannose***

Fructose enters central metabolism through phosphoglucose isomerase (Arias *et al.*, 1979), following phosphorylation by fructose kinase (Figure 1.6). Fructose kinase activity has been demonstrated in *S. meliloti* and mutants deficient in fructose kinase activity have been isolated and shown to be unable to grow using fructose as a sole carbon source (Gardiol *et al.*, 1980). The pathway for D-mannose catabolism has been characterized in *S. meliloti* L5-30. Both mannose kinase and mannose-phosphate isomerase activities were detected and mutants that lacked these activities were isolated (Arias & Cerveñansky, 1982). These activities are consistent with the conversion of mannose to mannose-6-phosphate, and subsequently fructose-6-phosphate where it is able to enter central metabolism (Figure 1.6)

**Figure 1.7.** De Ley-Doudoroff pathway of galactose catabolism



Arrows represent enzymatic reactions as defined in *P. saccharophila*. Circled numbers represent the following enzymes: 1) galactose dehydrogenase; 2) galactonolactonase; 3) galactonate dehydratase; 4) 2-keto-3-deoxygalactonokinase; 5) KDPGal aldolase.

The transport of both mannose and fructose has been shown to be an active process (Arias & Cerveñansky, 1982; Gardiol *et al.*, 1980). A genetic locus encoding a high affinity fructose ABC transporter has been identified in *S. meliloti* following a Tn5 mutagenesis by screening for mutants with the inability to utilize fructose as a sole carbon source (Lambert *et al.*, 2001). The ABC transporter is encoded on the chromosome by *frcBCA*. Both ribose and mannose competed for binding to the fructose sugar binding protein FrcB, however mutants in the transporter were still capable of growth using these sugars as sole carbon sources whereas they were unable to grow using fructose (Lambert *et al.*, 2001). The locus also contains a putative fructose kinase *frcK*, however FrcK is not homologous to the demonstrated fructose kinase of *R. leguminosarum* (Fennington & Hughes, 1996). The gene encoding phosphoglucose isomerase (*pgi*) may also be encoded nearby (Galibert *et al.*, 2001). However, the functions of *pgi*, or the putative fructose catabolism genes have not been experimentally verified.

#### ***1.4.3 Pentose catabolism and transport***

The inability of strains carrying mutations in G6P dehydrogenase and phosphoglucose isomerase to grow using ribose and xylose is consistent with catabolism through the cyclic ED pathway via the pentose phosphate pathway. L-arabinose catabolism has been characterized for its unique entry into central metabolism in *S. meliloti* through the TCA cycle rather than the PP pathway, or glycolysis as in *E. coli* (Duncan & Fraenkel, 1979).

#### 1.4.3.1 Arabinose

In the slow-growing rhizobia such as *B. japonicum*, L-arabinose is catabolized through non-phosphorylated intermediates leading to the production of pyruvate and glycolaldehyde (Pedrosa & Zancan, 1974). In this pathway, L-arabinose catabolism begins in analogy to galactose with oxidation to arabinolactone, followed by hydration forming L-arabonate and dehydration to 2-keto-3-deoxy-L-arabonate which is split by an aldolase reaction to pyruvate and glycolaldehyde (Pedrosa & Zancan, 1974). In contrast, *S. meliloti* L5-30 has been shown to grow gluconeogenically on L-arabinose, where it enters central metabolism through  $\alpha$ -ketoglutarate (Duncan & Fraenkel, 1979). Here, the first steps of arabinose degradation are conserved leading to the production of 2-keto-3-deoxy-L-arabonate. At this point,  $\alpha$ -ketoglutarate semialdehyde dehydrogenase oxidizes 2-keto-3-deoxy-L-arabonate to  $\alpha$ -ketoglutarate through which it enters central metabolism (Figure 1.6) (Duncan & Fraenkel, 1979). Surveys of several other fast growing rhizobia showed that they all contained  $\alpha$ -ketoglutarate semialdehyde dehydrogenase rather than 2-keto-3-deoxy-L-arabonate aldolase, suggesting that gluconeogenic growth on L-arabinose is a common theme among fast growing rhizobia (Duncan & Fraenkel, 1979).

A genetic locus has been identified on pSymB that is required for L-arabinose catabolism in *S. meliloti* Rm1021. The locus contains genes for both arabinose catabolism and transport (Poysti *et al.*, 2007). Genes predicted to encode a hydratase, dehydratase and  $\alpha$ -ketoglutarate semialdehyde dehydrogenase are all contained within the locus but a gene encoding L-arabinose dehydrogenase was notably absent (Poysti *et al.*, 2007). The



ABC transporter encoded by the locus AraABC was shown to be required for intracellular accumulation of arabinose (Poysti *et al.*, 2007).

#### ***1.4.3.2 Ribose and xylose***

Ribose kinase activity has been demonstrated in cell-free extracts of *S. meliloti* L5-30, and a mutant was isolated that lacked ribose kinase activity (Duncan, 1981). However, the location of the gene encoding ribose kinase in *S. meliloti* has not been described. In many organisms the catabolism of xylose proceeds by isomerization to xylulose followed by a phosphorylation to yield xylulose-5-phosphate. Xylose isomerase activity has been demonstrated in *S. meliloti* and a mutant was isolated that lacked this enzyme activity and was unable to grow using xylose as a sole carbon source (Duncan, 1981). These data are consistent with the hypothesis that the primary route of ribose and xylose catabolism is through the PP pathway, however there is evidence that there may be a secondary “metabolic bypass” for ribose and xylose catabolism (Duncan, 1981). This was suggested based on the ability of ribose kinase and xylose isomerase mutants to continuously incorporate  $^{14}\text{C}$  labeled ribose and xylose at rates consistent with the wild-type over a 22 hour period (Duncan, 1981).

#### ***1.4.4 Methylpentose catabolism and transport***

The biochemical pathways for the methylpentoses fucose and rhamnose have been determined in *E. coli* and *Salmonella* (Lin, 1996). Although these sugars are very similar in structure, their catabolism is carried out by two parallel pathways in these organisms. Essentially the sugars are isomerized to their keto derivatives, phosphorylated, and

broken into two 3 carbon compounds by sugar-specific aldolase yielding DHAP and lactaldehyde, the lactaldehyde is converted to pyruvate.

#### **1.4.4.1 Rhamnose**

Rhamnose transport and catabolism has been extensively characterized in *R. leguminosarum* bv. *trifolii* (Oresnik *et al.*, 1998). The *R. leguminosarum* locus consists of 9 genes that are arranged in two divergent operons. The genes at this locus consist of a negative regulator (*rhaR*), the components of an ABC transporter (*rhaSTPQ*), a mutarotase (*rhaU*), as well as an isomerase (*rhaI*), kinase (*rhaK*) and dehydrogenase/aldolase (*rhaD*) (Richardson *et al.*, 2004). It was shown that in *R. leguminosarum*, transport of rhamnose is dependent on the ABC transporter (Richardson *et al.*, 2004). Proper functioning of the transporter is dependent on the sugar kinase RhaK (Richardson & Oresnik, 2007). Utilizing a linker scanning mutagenesis the ability of RhaK to affect transport was shown to be distinct from its ability to act as a kinase (Rivers & Oresnik, 2013). Genetic and biochemical evidence support the hypothesis that RhaK could directly phosphorylate rhamnose (Richardson & Oresnik, 2007; Richardson *et al.*, 2004). The complement of enzymes in the locus is consistent with the capacity to catabolize rhamnose in a similar manner to *E. coli*, however the direct phosphorylation of rhamnose by RhaK is not. Further work should be carried out to resolve these differences and determine whether the pathway of rhamnose catabolism in rhizobia may be unique.

The evidence for rhamnose catabolism is more disparate in *S. meliloti*. The genes for rhamnose catabolism were first suggested based on similarity when the *R. leguminosarum* locus was sequenced (Richardson *et al.*, 2004). The locus was

subsequently shown to be necessary for rhamnose utilization and induced in response to rhamnose as a proof of principle on a genome-wide attempt to provide functionality to ABC transporters (Mauchline *et al.*, 2006). A triose-phosphate isomerase mutant (*tpiA*) was unable to utilize rhamnose as a sole carbon source, suggesting that it is catabolized to DHAP and lactaldehyde as in *E. coli* (Figure 1.6) (Poysti & Oresnik, 2007).

#### **1.4.4.2 Fucose**

The locus used for fucose transport and utilization is found on pSymB (*SMb21103-SMb21113*). The evidence for this is as follows: *SMb21103-SMb21106* encode a putative ABC transporter that has been shown to be induced by D and L-fucose (Mauchline *et al.*, 2006); a mutant of the putative short-chain dehydrogenase *SMb21111* has been shown to be unable to use either D- or L-fucose as sole carbon sources (Jacob *et al.*, 2008); deletions spanning these regions were shown to result in an inability to utilize D-fucose as a sole carbon source (Milunovic *et al.*, 2014); and *SMb21108* encodes a protein that is highly similar to RhaU which has been crystallized and shown to be a rhamnose mutarotase (Richardson *et al.*, 2008).

The locus contains a putative racemase (*SMb21107 /manR*), a mutarotase (*SMb21108*), two dehydrogenases (*SMb21109* and *SMb21111*), two putative dehydratases (*SMb21110* and *SMb21113*), and a gene that encodes an enzyme capable of cleaving a carbon-carbon bond (*SMb21112 /hpaG*). Based on these annotations, it is possible that fucose is broken down through non-phosphorylated intermediates since a carbohydrate kinase is not present. Precedent for fucose breakdown using a non-phosphorylated pathway has previously been shown in pseudomonads where fucose is

broken down to pyruvate and lactaldehyde (Dahms & Anderson, 1972a, b, c, d). The ability of *tpiA* and *tpiB* mutants to grow using fucose as a sole carbon source is consistent with such an entry-point, rather than DHAP as in *E. coli* (Poysti & Oresnik, 2007).

#### ***1.4.5 Sugar alcohol catabolism and transport***

Rhizobia have been historically known for their ability to grow efficiently using sugar alcohols for catabolism. Traditional media used for isolating rhizobia from nature contains high concentrations of mannitol to help enrich for rhizobia (Vincent, 1970). Sugar alcohol catabolism in *S. meliloti* has been studied for many years and the catabolism of several different sugar alcohols have been characterized genetically and biochemically.

##### ***1.4.5.1 Inositol***

The cyclic six carbon polyol inositol has been suggested to be prevalent in soils and an important substrate for *S. meliloti* based on the detection of inositol dehydrogenase activity in *S. meliloti* and *S. fredii* strains grown in solutions extracted from a wide range of soils (Wood & Stanway, 2001). The metabolism of *myo*-inositol and its isomers *scyllo*- and *chiro*-inositol has been shown to be encoded by three different loci in *S. meliloti*; *idhA*, encoded on pSymB, and the chromosomally encoded *iolA* and *SMc01163-iolRCDEB* loci (Kohler *et al.*, 2010). The metabolism of inositol in *S. meliloti* is thought to proceed through a similar pathway to that characterized in *Bacillus subtilis* (as first identified in *Klebsiella aerogenes*) wherein the end-products of its catabolism are DHAP and acetyl-CoA (Figure 1.6) (Kohler *et al.*, 2010; Yoshida *et al.*, 2008). Activities

of the enzymes encoding the first two steps of the pathway (*myo*-inositol dehydrogenase and 2-keto *myo*-inositol dehydratase) have been demonstrated in *Rhizobium leguminosarum* bv. *viciae* (Poole *et al.*, 1994). The inositol dehydrogenase that carries out the oxidation of *myo*- and *chiro*-inositol in *S. meliloti* is encoded by *idhA* (Galbraith *et al.*, 1998; Kohler *et al.*, 2010). The remaining enzymes catalyzing inositol catabolism breakdown in *S. meliloti* are thought to be encoded by *iolE*, *iolD*, *iolB*, *iolC* and *iolA* (Kohler *et al.*, 2010). *SMc01163* (*iolY*) is required for the catabolism of *scyllo*-inositol but not *myo*- or *chiro*-inositol, and is predicted to encode a *scyllo*-inositol dehydrogenase that can oxidize *scyllo*-inositol to 2-keto-*myo*-inositol which then proceeds through a common pathway with *myo*- and *chiro*- inositol (Kohler *et al.*, 2010). Regulation of inositol loci is carried out by IolR which negatively regulates the transcription of the *idhA* and *SMc01163-iolREDBCA* loci but not *iolA* (Kohler *et al.*, 2011). Evidence for induction of inositol genes by an end product of *myo*-inositol catabolism has been presented in *R. leguminosarum* bv. *viciae* (Fry *et al.*, 2001). A *myo*-inositol ABC transporter has been identified in *R. leguminosarum*, bv. *viciae* and designated IntABC (Fry *et al.*, 2001). In *S. meliloti*, two ATP-dependent transport systems were shown to be inducible by *myo*-inositol. One of these systems, encoded by *ibpA-iatA-iatP* (*SMb20712-SMb20714*) shares high identity with IntABC of *R. leguminosarum* (Mauchline *et al.*, 2006).

#### **1.4.5.2 Sorbitol, mannitol and D-arabitol**

Mannitol is widely considered to be a preferred carbon source for fast-growing rhizobia (Jordan, 1984; Vincent, 1970). An early study using extracts of mannitol-grown *S. meliloti* identified two different inducible polyol dehydrogenase activities. One of

these acted specifically on sorbitol, while the other acted on mannitol and D-arabitol (Martinez de Drets & Arias, 1970). A mutation of a putative sorbitol dehydrogenase on the chromosome, *SMc01500* (*smoS*), resulted in the inability to utilize D-sorbitol, D-mannitol or D-arabitol (Jacob *et al.*, 2008). Additionally, a putative mannitol dehydrogenase, *mtlK* is encoded immediately downstream of *smoC* (Galibert *et al.*, 2001). Genes encoding an ABC transporter (*smoEFGK*) immediately upstream of *smoS* and *mtlK* have been shown to be induced by sorbitol and mannitol (Mauchline *et al.*, 2006). Therefore, it seems likely that this locus may encode both the transport and catabolism of these polyols. Sorbitol and mannitol are oxidized to fructose in *S. meliloti* and D-arabitol is oxidized to D-xylulose (Martinez de Drets & Arias, 1970). A mutant deficient in fructose kinase activity in *R. meliloti* L5-30 was unable to grow using mannitol or sorbitol as sole carbon sources, consistent with these polyols entering central metabolism through fructose (Figure 1.6) (Gardiol *et al.*, 1980). D-xylulose derived from D-arabitol oxidation presumably enters central metabolism through the PP pathway as in xylose catabolism (Figure 1.6).

#### ***1.4.5.3 Dulcitol***

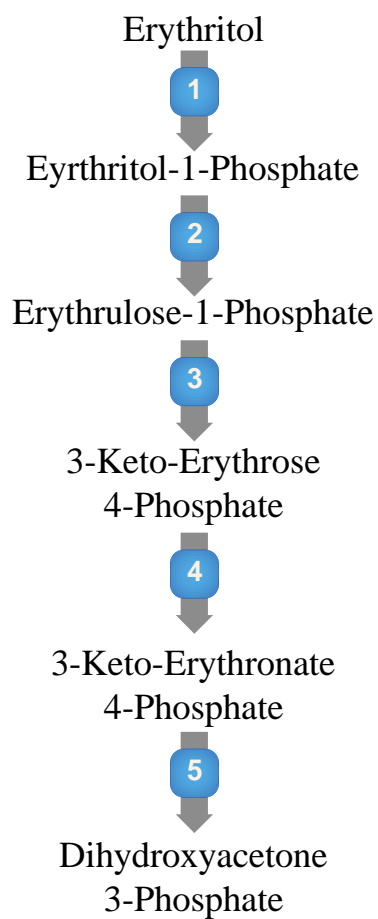
A region of pSymB has been shown to be required for the utilization of dulcitol as a sole carbon source (Charles & Finan, 1991; Milunovic *et al.*, 2014). The most likely candidate genes for dulcitol in this region are *SMb21375-SMb21377*, encoding a putative ABC transporter. Transcriptional fusions to *SMb21377* showed significant induction by dulcitol (Mauchline *et al.*, 2006). Although dulcitol dehydrogenase is present in *R. leguminosarum* bv. *trifolii* (Primrose & Ronson, 1980), an NAD<sup>+</sup> or NADP<sup>+</sup> dependent

dulcitol dehydrogenase was not detected in the extracts of dulcitol grown cells (Martinez de Drets & Arias, 1970). Therefore, dulcitol may be catabolized by a unique route in *S. meliloti*.

#### **1.4.5.4 Erythritol**

The catabolism of the four carbon polyol erythritol has been characterized in *Brucella abortus* and *R. leguminosarum* bv. *viciae* (Sangari *et al.*, 2000; Sperry & Robertson, 1975a, b; Yost *et al.*, 2006). The biochemical pathway of erythritol catabolism was elucidated in *B. abortus*. Erythritol catabolism proceeds through phosphorylation to erythritol-1-phosphate, followed by oxidation to erythrulose-1-phosphate. Further oxidations generating 3-keto-erythrose-4-phosphate and 3-keto-erythrone-4-phosphate, followed by decarboxylation yield DHAP where it enters central metabolism (Figure 1.8) (Sperry & Robertson, 1975a, b). A genetic locus was identified that encoded erythritol kinase (*eryA*), erythritol-1-phosphate dehydrogenase (*eryB*) and erythrulose-1-phosphate dehydrogenase (*eryC*) organized as an inducible operon with a putative negative regulator (*eryD*) (Sangari *et al.*, 2000). A homologous operon was later identified in *R. leguminosarum* and shown to be necessary for erythritol utilization (Yost *et al.*, 2006). The *R. leguminosarum* erythritol locus was further shown to contain a divergently transcribed ABC-transporter (*eryEFG*) that was necessary for growth using erythritol as a sole carbon source (Yost *et al.*, 2006).

**Figure 1.8.** Predicted pathway of erythritol catabolism in *S. meliloti*



Arrows represent enzymatic reactions as defined in *B. abortus*. Circled numbers represent the following enzymes: 1) erythritol kinase (EryA); 2) erythritol-1-phosphate dehydrogenase (EryB); 3) erythrulose-1-phosphate dehydrogenase (EryC); 4) 3-keto-erythrose-4-phosphate dehydrogenase; 5) 3-keto-erythronate-4-phosphate decarboxylase.



The erythritol locus of *S. meliloti* was identified during the characterization of the triose-phosphate isomerase genes *tpiA* and *tpiB*. A putative locus of erythritol catabolism that contained orthologs to *eryA*, *eryB*, *eryC* and *eryD* was identified adjacent to *tpiB* in the genome (Poysti & Oresnik, 2007). Interestingly, it was shown that *tpiB* was uniquely required for erythritol catabolism; overexpressed *tpiB* was able to complement *tpiA* mutants for growth using glycerol or rhamnose as sole carbon sources, whereas overexpressed *tpiA* was unable to complement *tpiB* mutants for growth using erythritol as a sole carbon source (Poysti & Oresnik, 2007). The inability of a triose-phosphate isomerase mutant to grow using erythritol as a sole carbon source is consistent with a DHAP entry-point into central metabolism as characterized *B. abortus* (Figure 1.6).

#### 1.4.5.5 Glycerol

Glycerol is the most widely used carbon source by rhizobia (Stowers, 1985). Glycerol catabolism has been studied in *B. japonicum* and *R. leguminosarum* bv. *viciae*, where it was found that glycerol is catabolized by glycerol kinase and  $\alpha$ -glycerolphosphate dehydrogenase to DHAP (Arias & Martinez de Drets, 1976). As well, in *R. leguminosarum* bv. *viciae*, an operon containing an ABC transporter for glycerol (*glpSTPQUV*), the glycerol kinase *glpK*, and the glycerol-3-phosphate dehydrogenase, *glpD* have been identified and were all shown to be necessary for the ability to utilize glycerol as a sole carbon source (Ding *et al.*, 2012). This operon is conserved across many of the  $\alpha$  proteobacteria.

In *S. meliloti* glycerol catabolism has not been well studied. *S. meliloti* contains an orthologous locus to the *R. leguminosarum* locus, however the kinase *glpK* is not present

(Ding *et al.*, 2012). The gene encoding glycerol kinase is divergently transcribed from *bdhA* and is found on pSymB (Aneja & Charles, 1999). The induction of both loci has been demonstrated during growth of *S. meliloti* using glycerol as a sole carbon source (White *et al.*, 2012). Taken together this suggests that the catabolism of glycerol occurs in an identical manner to that of *R. leguminosarum* where it enters central metabolism through DHAP which is subsequently converted to G3P (Figure 1.6). Consistent with this, *S. meliloti* strains carrying *tpiA* mutations are unable to grow using glycerol as a sole carbon source (Poysti & Oresnik, 2007). Presumably some carbon from glycerol is then cycled through the ED pathway, while the rest is catabolized to pyruvate. This is supported by the observation that strains carrying mutations in *eno* and *gap* are able to grow slowly with glycerol as a sole carbon source, indicating that the lower half of EMP is not required but is beneficial for growth (Finan *et al.*, 1988).

#### ***1.4.6 Regulation of peripheral metabolism***

Succinate and other C<sub>4</sub> dicarboxylic acids may be the preferred carbon sources for *S. meliloti* and *R. leguminosarum*. The regulation of the genes involved in the catabolism and transport of disaccharides, hexoses and sugar alcohols supports this, where many are subject to succinate-mediated catabolite repression (SMCR) (de Vries *et al.*, 1982; Gage & Long, 1998; Poole *et al.*, 1994; Ucker & Signer, 1978). In *R. leguminosarum* glucose mediated catabolite repression-like phenomena have also been reported during growth with sugar alcohols (Poole *et al.*, 1994; Ronson & Primrose, 1979). Glucose mediated catabolite repression of the  $\alpha$ -galactoside locus has been demonstrated in *S. meliloti* (Gage & Long, 1998).

SMCR is manifested as diauxic growth in *S. meliloti* during growth with lactose and is independent of cyclic nucleotides (Ucker & Signer, 1978). SMCR has been proposed to function by an inducer exclusion mechanism in *S. meliloti* where the presence of succinate in the cell transduces a signal that prevents the accumulation of inducing molecules that result in the expression of sugar transport and uptake systems (Bringinghurst & Gage, 2002). Although *S. meliloti* does not contain a complete PTS system for sugar transport, some components are maintained including Hpr and EIIA (Pinedo *et al.*, 2008). Phosphorylation and dephosphorylation of Hpr by the kinase/phosphatase HprK has been shown to influence SMCR (Pinedo & Gage, 2009). A model has been proposed where these components of the PTS system transduce a signal that corresponds to the presence of succinate in the cell, and result in the inducer exclusion mechanism by SMCR (Pinedo & Gage, 2009). Notably, components of the PTS<sup>NTR</sup> system in *R. leguminosarum* have been shown to globally regulate ATP transporters (Prell *et al.*, 2012)

## 1.5 Biological Relevance of Carbon Metabolism

Metabolism is fundamental to the physiology of any organism. *S. meliloti* can be found as a saprophytic organism, living in a plant rhizosphere, colonizing a root surface, or in an association with the host plant. The ability to catabolize organic compounds is important at all stages of its lifecycle including free-living in the rhizosphere, during the infection process of its host and in the intracellular nitrogen-fixing bacteroid form. This section will summarize the symbiotic phenotypes that have been identified in different carbon catabolism mutants of *S. meliloti*, and integrate them with the current understanding of *S. meliloti* carbon metabolism and symbiosis as discussed in sections 1.2-1.4.

### 1.5.1 Free-living catabolism

Prior to invasion of the plant, *S. meliloti* must grow and thrive saprophytically in the rhizosphere. The effectiveness of rhizobial inocula in agriculture is limited by the inability of commercial inocula to compete with indigenous rhizobia for nodule occupancy of legume crops (Triplett & Sadowsky, 1992). The ability to thrive in the rhizosphere and on the root surface could be influenced by the ability to extract energy from carbon compounds that are available there.

Most rhizobia, including *S. meliloti*, contain large multi-partite genomes. The large genomic size and complexity is reflective of a large metabolic capacity that could be advantageous towards adaptation to the diverse environment of the soil. Consistent with this idea, the metabolic potential of Rhizobium strains has been correlated with increased competitiveness (Wiebo *et al.*, 2007). The *S. meliloti* genome contains a large

number of ABC transporters that are inducible by organic compounds (Galibert *et al.*, 2001; Mauchline *et al.*, 2006). Of the two megaplasmsids contained by *S. meliloti*, pSymA has been described as a symbiotic plasmid, encoding the genes required for nod factor biosynthesis and nitrogen fixation (Barnett *et al.*, 2001). The second megaplasmid, pSymB encodes genes for exopolysaccharide biosynthesis as well as a large number of solute uptake systems and carbon metabolism genes and it has been suggested that pSymB may play a primary role in the ability to thrive in the diverse environment of the soil during saprophytic growth (Finan *et al.*, 2001).

Evidence for plasmids providing a competitive advantage in this way has been demonstrated in rhizobia. In *R. leguminosarum* bv. *viciae* and *Rhizobium etli* CFN42, plasmid curing experiments showed that strains lacking certain plasmids were less competitive for nodule occupancy than wild-type strains (Brom *et al.*, 1992; Hynes & O'Connell, 1990; Hynes & McGregor, 1990). This idea has recently been reinforced by an experiment analyzing the transcriptional response of *R. leguminosarum* bv. *viciae* grown in the rhizosphere of its host legume (pea), a non-host legume (alfalfa) and a non-legume (sugar beet) (Ramachandran *et al.*, 2011). Here, one of the six plasmids contained by *R. leguminosarum*, pRL8, was shown to be a pea rhizosphere specific plasmid. A large proportion of genes upregulated specifically in the pea rhizosphere are encoded on this plasmid, and mutants of many of these genes showed reduced competitiveness for pea rhizosphere colonization (Ramachandran *et al.*, 2011). Overall, carbon metabolism in the rhizosphere in *R. leguminosarum* was shown to be dominated by organic acids with a bias towards those that consist of one or two carbon atoms (Ramachandran *et al.*, 2011).

There has been no directed study of the carbon compounds that make up the rhizosphere of the legume hosts of *S. meliloti*. The *R. leguminosarum* transcriptional responses that were specific to the alfalfa rhizosphere suggest that the C<sub>4</sub> dicarboxylic acid malonate and the polysaccharide arabinogalactan may be prevalent (Ramachandran *et al.*, 2011). A *gfp* fusion to the *mela* promoter of *S. meliloti*, which is induced by galactose and galactosides, was used as a biosensor to show that galactosides are released from alfalfa seeds during germination and from roots of alfalfa seedlings in sterilized and unsterilized soil (Bringham *et al.*, 2001). The arabinose locus has also been shown to be induced by alfalfa seed exudate (Poysti *et al.*, 2007). One of the two putative ribulose-5-P epimerase genes, *ppe*, and the sugar binding protein *frcB* of the fructose transporter are induced by alfalfa root exudate (Zhang and Cheng, 2006). Transcription of *lacZ* reporter gene fusions to the trehalose catabolic gene *thuB* and the PQQ-dependent glucose dehydrogenase (*gcd*) was observed during growth on the alfalfa root surface in addition to other stages of infection (Bernardelli *et al.*, 2008; Jensen *et al.*, 2005).

An alternate approach to identifying carbon sources that may be important for the ability to thrive in the rhizosphere has been the isolation of genetic determinants of competition (Triplett & Sadowsky, 1992). Several different carbon utilization mutants have been isolated that have been shown to be impaired in competitiveness for nodule occupancy in *S. meliloti* and *R. leguminosarum*. The ability to utilize rhamnose is plasmid encoded in *R. leguminosarum* and mutants that are unable to utilize rhamnose show a severe competitive defect (Oresnik *et al.*, 1998). Mutants that were unable to catabolize inositol showed reduced competitiveness for nodule occupancy in both *S. meliloti*

Rm1021 and *R. leguminosarum* bv. *trifolii*, as well as *S. fredii* USDA191 (Fry *et al.*, 2001; Jiang *et al.*, 2001; Kohler *et al.*, 2010).

Some strains of *S. meliloti* and *R. leguminosarum* are capable of both the synthesis and catabolism of inositol derivatives called rhizopines (L-3-*O*-methyl-*scyllo*-inosamine and *scyllo*-inosamine). Rhizopine is synthesized in the nodules formed by these species, and exuded into the rhizosphere where it can be catabolized by other saprophytically growing rhizobia (few other bacteria possess the ability to catabolize rhizopine) (Murphy *et al.*, 1995). This has led to the “rhizopine concept” which suggests an altruistic synthesis of rhizopine in the nodules by rhizobia for catabolism by their siblings in the rhizosphere, providing them with a competitive advantage (Murphy *et al.*, 1995). Consistent with this concept, mutants in rhizopine catabolism genes had a reduced ability to compete for nodule occupancy (Gordon *et al.*, 1996).

Mutants that were unable to catabolize erythritol, glycerol and homoserine all have reduced competitiveness for nodule occupancy in *R. leguminosarum* bv. *viciae* (Ding *et al.*, 2012; Vanderlinde *et al.*, 2013; Yost *et al.*, 2006). The ability to catabolize the amino acid proline was shown to contribute to competitiveness in *S. meliloti* under drought conditions (Jiménez-Zurdo *et al.*, 1995; van Dillewijn *et al.*, 2001). Mutants of the PQQ-dependent glucose dehydrogenase showed a reduced competitiveness for nodule occupancy in *S. meliloti* (Bernardelli *et al.*, 2008). However, given that glucose catabolism would still be intact in this mutant it is possible that the competitive defect may be a result of other effects. For example, gluconate production by a periplasmic glucose dehydrogenase is involved in inorganic phosphate solubilization as well as the regulation of biocontrol traits in *Pseudomonas* spp. (de Werra *et al.*, 2009).

In most cases there is not significant evidence showing that competition is occurring specifically in the rhizosphere. An alternative possibility is that the carbon compounds may be utilized during portions of the infection process further downstream. It has been hypothesized that this may be the case for both rhamnose and inositol in *R. leguminosarum* based on the lack of significant differences in competitiveness observed at the level of root colonization and rhizosphere growth, respectively (Fry *et al.*, 2001; Oresnik *et al.*, 1998).

### ***1.5.2 Carbon catabolism and invasion***

Successful penetration of the IT by *S. meliloti* requires both continued biosynthesis of NF as well as biosynthesis of symbiotic exopolysaccharide (Jones *et al.*, 2007). The requirements for active biosynthesis of such macromolecules as well as for active proliferation in both the CCRH and the IT suggests the rhizobia must be in metabolically good shape to allow successful invasion. Despite the importance of this portion of the life cycle of *S. meliloti*, the carbon sources that are utilized to fuel invasion have not been elucidated.

It has been postulated that growth in the IT may be fuelled by the storage polymer polyhydroxybutyrate (PHB) in *S. meliloti* (Charles *et al.*, 1997). PHB accumulates as granules during normal growth in *S. meliloti* and other indeterminate nodule forming rhizobia that disappear during bacteroid differentiation (Lodwig *et al.*, 2005). It has therefore also been hypothesized that PHB may be an energy store to drive bacteroid differentiation (Prell & Poole, 2006). Consistent with the hypothesis that PHB may help fuel growth in the IT, mutants that were unable to synthesize or catabolize PHB were less



competitive than the wild type for nodule occupancy (Aneja *et al.*, 2005; Willis & Walker, 1998). However, mutants that were unable to synthesize PHB were at a more severe competitive disadvantage than mutants that were unable to catabolize PHB (Aneja *et al.*, 2005). Therefore, it was suggested that PHB synthesis may play a role in removing inhibitory metabolic intermediates during invasion (Aneja *et al.*, 2005). The study of the role of PHB in *S. meliloti* is complicated because rhizobia are also capable of accumulating the storage polymer glycogen in addition to PHB (Lodwig *et al.*, 2005). To resolve how these storage polymers function in the interactions of *S. meliloti* with *Medicago* spp. a study was conducted wherein mutants in PHB synthesis (*phbC*) were constructed in concert with mutants in the two putative glycogen synthases in *S. meliloti* (*glgA1*, *glgA2*) to resolve how these storage polymers function in the interactions of *S. meliloti* with *Medicago* spp (Wang *et al.*, 2007). Glycogen was not detectable in *glgA1* mutants, demonstrating that *glgA1* encodes the glycogen synthase in *S. meliloti* (Wang *et al.*, 2007). Glycogen levels were elevated in a *phbC* mutant, whereas PHB accumulation was reduced in the *glgA1* mutant (Wang *et al.*, 2007). Mutants in glycogen biosynthesis appeared to be unaffected in invasion of *Medicago* spp., but did show reduced nitrogen fixation phenotypes. Consistent with a reduced ability to compete for nodule occupancy, nodule formation delays were observed in *phbC* mutants, however EPS synthesis was significantly reduced in *phbC* mutants. Therefore, it is possible that PHB plays a role in fuelling EPS biosynthesis during invasion (Wang *et al.*, 2007).

Although the mechanism is unclear, the synthesis of exopolysaccharide is critical for invasion by *S. meliloti* (Jones *et al.*, 2007). EPS synthesis has also been shown to affect competition for nodule occupancy in other rhizobia (Triplett & Sadowsky, 1992).

The common precursor for the biosynthesis of the sugar subunits for EPSI, EPSII and cyclic  $\beta$ -glucan is glucose-6-P. Isotopic labeling experiments have demonstrated that during growth on hexoses, much of the carbon used for the biosynthesis of these polysaccharides is degraded through the ED and PP pathways and reformed before being used for biosynthesis (Gosselin *et al.*, 2001; Portais *et al.*, 1999). These results demonstrate a link between central metabolism and biosynthesis of these important symbiotic molecules. Consistent with the role of central metabolism in EPS biosynthesis, strains carrying phosphoglucose isomerase mutations were 80 % reduced in EPS production relative to wild type (Arias *et al.*, 1979).

Some studies have suggested a link between the regulation of carbon metabolism and EPS biosynthesis. The  $\alpha$ -galactoside locus was isolated on the basis of identifying genes downregulated by SyrA which is also known to effect exopolysaccharide regulation (Gage & Long, 1998). Mutations in components of the PTS-like SMCR system (*hpr* and *eIIA*) also showed dramatic effects on exopolysaccharide production (Pinedo *et al.*, 2008). Notably, these components are encoded in a locus with *exoS* and *chvI*; ExoS and ChvI comprise a two-component system that regulates exopolysaccharide biosynthesis in *S. meliloti* (Cheng & Walker, 1998a). In addition to dramatic effects on exopolysaccharide production, null mutants of *exoS* and *chvI* showed an inability to grow using over 21 different carbon sources (Bélanger *et al.*, 2009). During a global study of ChvI binding sites, a ChvI binding site was identified upstream of the putative galactose dehydrogenase *gal/SMc00588* (Bélanger & Charles, 2013). Despite these links it is unclear how carbon metabolism, or its regulation might influence exopolysaccharide biosynthesis during symbiosis.

Osmoprotection of *S. meliloti* can be accomplished by the accumulation of trehalose in the cytosol during growth in hyperosmotic conditions and the accumulation of the cyclic  $\beta$ -glucans as in the periplasm during growth in hypoosmotic conditions (Breedveld *et al.*, 1993; Dylan *et al.*, 1990a). Symbiotic phenotypes have been associated with the inability to synthesize either cyclic  $\beta$ -glucan, or trehalose. Mutants in cyclic  $\beta$ -glucan biosynthesis were impaired in their ability to nodulate the host plant; however, suppressor mutants that regained the ability to effectively nodulate alfalfa did not regain the ability to synthesize cyclic  $\beta$ -glucan and remained sensitive to hypoosmotic growth conditions (Dylan *et al.*, 1990b).

Mutants of three trehalose biosynthetic loci showed reduced osmotic tolerance and a triple mutant of all three loci showed a significantly reduced ability to compete for nodule occupancy (Domínguez-Ferreras *et al.*, 2009). Therefore, tolerance to hyperosmotic conditions may be more important than tolerance to hypoosmotic conditions during invasion. An intriguing relationship between trehalose biosynthesis and catabolism during invasion has been demonstrated in *S. meliloti*, where mutants of the trehalose catabolism genes *thuA* and *thuB* showed an increased ability to compete for nodule occupancy despite a reduced ability to colonize alfalfa roots (Jensen *et al.*, 2005). This increased competitive advantage occurs at the level of root hair infection (Jensen *et al.*, 2005). It was suggested that the increased ability to compete for nodule occupancy was the result of enhanced tolerance to osmotic stress in the infection thread as a result of an incidental accumulation of trehalose in the catabolic mutants (Jensen *et al.*, 2005). Trehalose biosynthesis is also tied to central metabolism as glucose-6-P dehydrogenase

was shown to be required for trehalose to be an efficient osmoprotectant (Barra *et al.*, 2003).

### **1.5.3 *Bacteroid* metabolism**

Evidence is overwhelmingly in favor of C<sub>4</sub> dicarboxylic acids as the carbon source in the bacteroid. The dicarboxylic acids are derived from sucrose photosynthate in plant cells, and transported through the SM to the bacteroid. There is evidence that the SM of legumes contains a dicarboxylic acid transporter, however a genetic locus encoding the transporter has not yet been identified (Udvardi & Poole, 2013). Bacterial genetics has defined the components for dicarboxylic acid transport and catabolism in *S. meliloti*. Mutants that are unable to transport or catabolize dicarboxylic acids display phenotypes consistent with normal infection and bacteroid development, but produce bacteroids that are devoid of nitrogen fixation (Fix<sup>-</sup>) (Lodwig & Poole, 2003). In contrast, mutants of hexose or polyol catabolism, as well as the key central catabolism enzymes G6P dehydrogenase and pyruvate carboxylase show no significant effect on nitrogen fixation in bacteroids (Cerveñansky & Arias, 1984; Dunn *et al.*, 2001; Stowers, 1985).

Dicarboxylic acid transporters were first identified in *R. leguminosarum* where they were shown to be necessary for growth on succinate, fumarate and malate as well as being necessary for nitrogen fixation in nodules (Finan *et al.*, 1983; Ronson *et al.*, 1981). In *S. meliloti*, *dct* mutants were subsequently isolated and also shown to be necessary for effective nitrogen fixation (Bolton *et al.*, 1986; Engelke *et al.*, 1987; Finan *et al.*, 1988; Jiang *et al.*, 1989; Watson *et al.*, 1988; Yarosh *et al.*, 1989). The C<sub>4</sub> dicarboxylic acid

transporter is DctA, of the major facilitator superfamily of transporters. DctA is regulated by a two-component system consisting of DctB and DctD, that induce *dctA* transcription in response to C<sub>4</sub> dicarboxylates (Yurgel & Kahn, 2004). DctB is a membrane-bound sensor kinase. Binding of dicarboxylates by DctB in the periplasm results in the phosphorylation of DctD by the periplasmic histidine kinase domain of DctB, and subsequent phosphorylation of the response regulator DctD. Phosphorylated DctD activates *dctA* transcription in an RpoN-dependent manner (Yurgel & Kahn, 2004, 2005; Zhou *et al.*, 2008).

C<sub>4</sub> dicarboxylic acids are catabolized through the TCA cycle. Consistent with this, the TCA cycle has been shown to be indispensable for nitrogen fixation in many rhizobia. Mutants of several TCA cycle enzymes in *S. meliloti* have been shown to display a Fix<sup>-</sup> phenotype. These include succinate dehydrogenase (Gardiol *et al.*, 1982) citrate synthase (Koziol *et al.*, 2009), isocitrate dehydrogenase (McDermott & Kahn, 1992), and malate dehydrogenase (Dymov *et al.*, 2004). Function of the TCA cycle during growth on dicarboxylates requires the active synthesis of acetyl-CoA from pyruvate. In *S. meliloti* bacteroids this role is predominantly performed by NAD<sup>+</sup> dependent malic enzyme (*dme*) rather than PEP carboxykinase (*pck*) or NADP<sup>+</sup> dependent malic enzyme (*tme*). Mutants of *dme* display a Fix<sup>-</sup> phenotype (Driscoll & Finan, 1993). Mutants of *tme* show no symbiotic phenotype, and despite encoding enzymes with the same biochemical activity that could functionally replace each other *in vitro*, overexpressed *tme* was unable to complement *dme* mutants for functional symbiosis *in planta* (Driscoll & Finan, 1996; Mitsch *et al.*, 2007). Strains carrying mutations in *pckA* did show a reduced nitrogen fixation phenotype (Finan *et al.*, 1991).

However, significant PEP carboxykinase activity has not been detected in bacteroids, whereas *dme* has been shown to be constitutively expressed (Driscoll & Finan, 1996; Finan *et al.*, 1991).

In addition to the synthesis of pyruvate for acetyl-CoA production, evidence is consistent with the requirement of gluconeogenesis to synthesize essential sugar intermediates for biosynthesis in the bacteroid. Several mutants of the lower EMP showed Fix – phenotypes, including enolase, G3P dehydrogenase and 3PG kinase (Finan *et al.*, 1988). Gluconeogenesis via malic enzyme requires pyruvate orthophosphate dikinase (*pod*) to synthesize PEP from pyruvate (Østerås *et al.*, 1997). Given that *pod* mutants show no significant symbiotic phenotype (Østerås *et al.*, 1997), it seems possible that the role of PEP carboxykinase may be to generate the PEP needed for gluconeogenesis through the lower EMP, while NAD<sup>+</sup> malic enzyme is responsible for generating the pyruvate for Acetyl-CoA synthesis.

## 1.6 Thesis Objectives

The overall goal of this thesis was to use bacterial genetics to investigate carbon metabolism in *S. meliloti*, and to assess the influence of carbon source utilization on the competitiveness for nodule occupancy of *S. meliloti* strains. The benefits to carrying out this research are two-fold. The genetics and physiology of carbon metabolism in *S. meliloti* are beginning to be well understood as described in this section. Augmenting this understanding with the continued characterization of carbon metabolism in *S. meliloti* will help to refine this organism as a model system for understanding carbon metabolism in the  $\alpha$ -proteobacteria. Secondly, carbon metabolism has been shown to influence the ability of rhizobial strains to compete for nodule occupancy of their legume hosts. Further investigation of the contribution of carbon metabolism to nodule competitiveness could lead to the development of superior inoculum strains, as well as contribute to the *Sinorhizobium-Medicago* model of plant-microbe interactions.

The first objective of this thesis was to define the genetic locus for erythritol catabolism in *S. meliloti* that was identified during the investigation of *S. meliloti* triose-phosphate isomerase mutants (Poysti & Oresnik, 2007). The erythritol locus of *S. meliloti* was of interest because the ability to catabolize erythritol has been shown to influence virulence in *Brucella* spp., and the ability to compete for nodule occupancy in *R. leguminosarum* (Burkhardt *et al.*, 2005; Yost *et al.*, 2006). The locus is also dramatically different in its content and complexity from the characterized loci in *Brucella* spp. and *R. leguminosarum*. Experiments were carried out to characterize the genes in the erythritol locus of *S. meliloti* and to assess the influence of the ability to utilize erythritol on competitiveness for nodule occupancy.

The second objective of the thesis was to identify and characterize the genetic locus (or loci) for galactose catabolism and transport in *S. meliloti*. The biochemical pathway of galactose catabolism has been defined in *S. meliloti* (Arias & Cerveñansky, 1986), but a genetic locus encoding the pathway of galactose catabolism or galactose transport has not been identified. Moreover, despite the prevalence of galactose in pea root mucilage (Knee *et al.*, 2001), and on the surface of alfalfa seedlings (Bringham *et al.*, 2001), the influence of the ability to catabolize galactose on competitiveness for nodule occupancy has not been described in Rhizobia. It was hypothesised that the ability to catabolize galactose in the rhizosphere may be important for the ability to compete for nodule occupancy in *S. meliloti*.



## **Chapter 2**

### **A Locus Necessary for the Transport and Catabolism of Erythritol in**

#### ***Sinorhizobium meliloti***

Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology. This work was carried out by Barney Geddes in collaboration with Nathan Poysti and Brad Pickering. NP performed the erythritol transport assays and competition for nodule occupancy assays. BP constructed the mutation in *SMc01615*. Technical assistance was provided by Heather Collins, Erin Loewen, Harry Yudistira and Roy Hutchings.

## 2.1 Introduction

*Sinorhizobium meliloti* is a Gram-negative soil microbe that can exist either as a free-living organism in the soil or in a symbiotic relationship with legume plants. *S. meliloti* can interact with the legume alfalfa to form root nodules, and within these nodules it can reduce nitrogen gas and provide this fixed nitrogen to the plant (Spaink, 2000). Prior to the colonization of the root and the initiation of infection threads that lead to the release of the bacteria into the root nodule, the bacteria must be able to survive in the rhizosphere. It has been shown that the ability to use organic compounds can strongly influence competition in the rhizosphere (Triplett & Sadowsky, 1992). In *R. leguminosarum*, the inability to utilize erythritol has been shown to affect a strain's ability to effectively compete for nodule occupancy (Yost *et al.*, 2006).

The ability to utilize erythritol as a sole carbon source is not universal among the *Rhizobiaceae* (Jordan, 1984; Stowers, 1985). Evidence has been presented that suggests the erythritol catabolic locus may have been horizontally transferred from *Brucella* to *R. leguminosarum* (Yost *et al.*, 2006). In addition to the *eryABCD* locus that had been previously shown to be necessary for erythritol catabolism (Sangari *et al.*, 2000), it was shown that two sets of genes flanking *eryABCD* of *R. leguminosarum* were also necessary for erythritol catabolism. These are *eryEFG* encoding three ABC transporter genes, and an operon consisting of a DeoR type regulator, a triose-phosphate isomerase and a putative ribulose-phosphate isomerase (Yost *et al.*, 2006).

The inability to utilize erythritol by *Brucella abortus* strain S19 has been correlated with attenuated virulence (Meyer, 1966; Meyer, 1967). This association prompted the elucidation of biochemical pathway for its catabolism (Sperry & Robertson,

1975a, b). Subsequently, the genes for the first three steps of this biochemical pathway were characterized (*eryA*, *eryB* and *eryC*) and found to appear in an operon with a negative regulator *eryD* (Sangari *et al.*, 2000; Sangari *et al.*, 1998). More recent work, using proteomic as well as comparative genomic approaches, has reinforced the role erythritol transport and catabolism may play in virulence (Burkhardt *et al.*, 2005).

In *S. meliloti*, screening for mutants unable to use rhamnose as a sole carbon source identified two regions: the genes homologous to the *R. leguminosarum* genes necessary for rhamnose catabolism, *rhaDI* and *rhaRSTPQMK* (Richardson & Oresnik, 2007; Richardson *et al.*, 2004; Richardson *et al.*, 2008), and a gene encoding a triose-phosphate isomerase (*tpiA*) (Poysti & Oresnik, 2007). The *S. meliloti* genome contains another gene, *tpiB*, that also encodes triose-phosphate isomerase activity (Poysti & Oresnik, 2007). Although both genes encode functional triose-phosphate isomerases, they appear to play distinct biochemical roles in the organism. TpiA is necessary for glycerol catabolism and has been hypothesized to play a role in central metabolism, whereas TpiB is necessary for the catabolism of erythritol (Poysti & Oresnik, 2007). It is of note that the over-expression of the wild-type *tpiB* can complement *tpiA* associated mutant phenotypes whereas it was shown that a *tpiB* mutation was unable to be genetically complemented by *tpiA* (Poysti & Oresnik, 2007). Since only a *tpiA/B* double mutant, and not a *tpiA* or *tpiB* single mutant, is impaired for autotrophic and gluconeogenic growth, *tpiB* is expressed at low levels in the absence of erythritol (Pickering & Oresnik, 2008; Poysti & Oresnik, 2007).

Our interest in erythritol stems from our initial studies with *tpiB* (Poysti & Oresnik, 2007). Additionally, as noted by Yost *et al.* (2006), the genetic content and

arrangement of the putative erythritol locus adjacent to *tpiB* is drastically different from that previously described in *R. leguminosarum* and *B. abortus* (Galibert *et al.*, 2001; Halling *et al.*, 2005; Young *et al.*, 2006). Therefore the goal of this work was to genetically characterize the locus and to determine what components are necessary for the transport and catabolism of erythritol.

## **2.2 Materials and Methods**

### **2.2.1 Bacterial strains, plasmids and media**

Bacterial strains and plasmids used and generated in this work are listed in Table 2.1. *S. meliloti* strains were grown routinely at 30° C on complex Luria-Bertani medium (Sambrook *et al.*, 1989) or on defined Vincent's minimal medium (VMM) (Vincent, 1970). Carbon sources were filter sterilized and added to VMM to a final concentration of 15 mM. When required, *S. meliloti* and *Escherichia coli* strains were grown in the following concentrations of antibiotic (µg/mL): chloramphenicol (Cm) 20; gentamicin (Gm) 20 or 60; kanamycin (Km) 20; neomycin (Nm) 200; rifampicin (Rf) 50; streptomycin (Sm) 200; tetracycline (Tc) 5. All antibiotics were filter sterilized before use.

### **2.2.2 DNA manipulations and constructions**

Standard techniques were used for plasmid isolation, restriction enzyme digests, ligations, transformations and agarose gel electrophoresis (Sambrook *et al.*, 1989). Oligonucleotide primers used for PCR amplification are listed in Table 2.2.

**Table 2.1.** Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
<b>Strain</b>		
<i>S. meliloti</i>		
Rm1021	SU47 <i>str-21</i> , Sm <sup>R</sup>	(Meade <i>et al.</i> , 1982)
RmG212	Rm1021 <i>lac</i> , Sm <sup>R</sup>	Lab collection
SRmA355	Rm5000 <i>tpiB1</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA428	Rm1021 <i>SMc01627</i> ::Tn5, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA449	$\Delta tpiA$	(Poysti & Oresnik, 2007)
SRmA465	$\phi$ (SRmA428) $\rightarrow$ Rm1021, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA515	$\Delta tpiA$ , <i>tpiB1</i> $\phi$ (SRmA355) $\rightarrow$ SRmA449, Gm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA584	Rm1021 <i>tpiB2</i> , Nm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA723	Rm1021 <i>SMc01627</i> ::Tn5-233, Gm <sup>R</sup> Sp <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA778	Rm1021 <i>ery-1</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA779	Rm1021 <i>ery-2</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA780	Rm1021 <i>ery-3</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA781	Rm1021 <i>ery-5</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA782	Rm1021 <i>eryA6</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA783	Rm1021 <i>ery-7</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA784	Rm1021 <i>eryB9</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA785	Rm1021 <i>ery-10</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA787	Rm1021 <i>eryC12</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA885	Rm1021 <i>eryD13</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA966	Rm1021 <i>ery-14</i> :: <i>nptII</i> , Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
<i>E. coli</i>		
DH5 $\alpha$	$\lambda$ $\Phi$ 80dlacZ <sup>o</sup> M15 <sup>o</sup> (lacZYA-argF)U169 recA1endA1 hsdR17(r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) supE44 thi-1 gyrA relA1	(Hanahan, 1983)
DH5 $\alpha$ Rif	Rifampicin resistant DH5 $\alpha$	(House <i>et al.</i> , 2004)
DH5 $\alpha$ $\lambda$ pir	$\lambda$ pir lysogen of DH5 $\alpha$	(House <i>et al.</i> , 2004)
MM294A	pro-82-thi-1 hsdR17 supE44	(Finan <i>et al.</i> , 1986)
MT607	MM294A recA56	(Finan <i>et al.</i> , 1986)
MT616	MT607 (pRK600)	(Finan <i>et al.</i> , 1986)
S17-1	<i>recA</i> derivative of MM294A with integrated RP4-2 (Tc::Mu::Km::Tn7)	(Simon <i>et al.</i> , 1983)
EcA101	MT607 $\Omega$ Tn5-B20, Kan <sup>R</sup>	(Clark <i>et al.</i> , 2001)
<b>Plasmids</b>		
pBG1	pKNOCK-Gm + 400 bp internal <i>eryD</i> fragment ( <i>Bam</i> HI / <i>Cla</i> I), Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG2	pRK7813 + <i>eryD</i> ( <i>Pst</i> I, <i>Bam</i> HI), Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG3	pRK7813/ <i>eryA</i> , ( <i>Eco</i> RI, <i>Hind</i> III), Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG13	pCO37/ <i>eryA</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG14	pCO37/ <i>eryC</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)

pBG18	pCO37/ <i>eryD</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBP149	pBS/ <i>SMc01615</i> , Ap <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBP151	pBS/ <i>SMc01615::nptII</i> , Ap <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBP173	pJQ200/ <i>SMc01615::nptII</i> , Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBluescript II SK	Cloning vector, Ap <sup>R</sup>	Stratagene
pCO37	pRK7813 containing <i>attB</i> sites, Gateway® compatible destination vector	(Jacob <i>et al.</i> , 2008)
pEL6	CX1 erythritol complementing cosmid, Tc <sup>R</sup>	(Wang <i>et al.</i> , 2006)
pHY112	pRK7813 / <i>eryB</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC1	pEL6 <i>ery-1::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC2	pEL6 <i>ery-2::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC3	pEL6 <i>ery-3::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC5	pEL6 <i>ery-5::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC6	pEL6 <i>eryA6::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC7	pEL6 <i>ery-7::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC9	pEL6 <i>eryB9::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC10	pEL6 <i>ery-10::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pHC12	pEL6 <i>eryC12::Tn5-B20</i> , Nm <sup>R</sup> Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pKNOCK-Gm	Suicide vector, Gm <sup>R</sup>	(Alexeyev, 1999)
pNP163	<i>tpiB</i> -complementing cosmid	(Poysti & Oresnik, 2007)
pPH1JI	IncP plasmid, Gm <sup>R</sup>	(Beringer <i>et al.</i> , 1978)
pRK600	pRK2013 <i>nptII::Tn9</i> , Cm <sup>R</sup>	(Finan <i>et al.</i> , 1986)
pRK602	pRK600Ω <i>Tn5</i> , Cm <sup>r</sup> , Nm <sup>R</sup>	(Finan <i>et al.</i> , 1985)
pRK7813	Broad-host-range vector, Tc <sup>R</sup>	(Jones & Gutterson, 1987)
pJQ200SK	Gene replacement suicide vector, Gm <sup>R</sup>	(Quandt & Hynes, 1993)
pMM22	Kan <sup>R</sup> fragment cloned as <i>SmaI</i> in pBlueScript	(Pickering & Oresnik, 2008)

Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

**Table 2.2.** PCR primers

Primer number	Primer name	Nucleotide sequence (5' → 3')
1	pKNOCK eryD F	ATATGGATCCTCGAGCTGGAGCAGAAGC
2	pKNOCK eryD R	ATATATCGATGACATCGCGATCCTCGAC
3	eryD 7813 F	ATATCTGCAGGAGGCTAGCTAGATGGCGATCATCGGTGG
4	eryD 7813 R	ATATGGATCCTCAAGCCAGCGTCTGCG
5	eryA 7813 F	ATATAAGCTTGGAGGCTAGCTAGATGCGCGACATCCTCATC
6	eryA 7813 R	ATATGAATTCTCATTGCGGTCTTTCCTGTTG
7	eryB 7813 F	ATATCTGCAGGAGGCTTATTCATAGTGAGTTCGAAACCGG
8	eryB 7813 R	ATATGGATCCTCACGCCGCAGCCTGCTGGC
9	deoR Kan F	ATATGGGCCCCAAGCAGTCCTCGATGAACA
10	deoR Kan R	ATATGAGCTCGACACTTCGTTTCCGGTCAG
11	1628 RT F	CAAGGATGCGACCTGGTATT
12	1628 RT R	GGATAGCTCTTGTTACGCCG
13	eryD RT F	ATCAGGAAGGCTTGGTGAAA
14	eryD RT R	ACCACCTCGCAATAGTCCAG
15	eryA RT F	CTGGACGGACACAGTCAAAA
16	eryA RT R	CTCTGTCGATCATCCAGGTG
17	eryC RT F	CAAGCAGTCCTCGATGAACA
18	eryC RT R	ATTCGAAGCATTGAATCCG
19	deoR RT F	TGAAACCGGAAGACAGACG
20	deoR RT R	CTGCTCGAGATCGTCAAGGT
21	eryA 1622 J F	CCGTCTCTCTGGGCATATCTAT
22	eryA 1622 J R	GCTTGCCGAGATATTCCTTG
23	1622 fucA J F	CTGGTCCACGATCCCTATGT
24	1622 fucA J R	CGATCATGTAGTGACAGGCG
25	fucA eryB J F	TCGAGGATTTGCCACTTAC
26	fucA eryB J R	CCCAGGTGATCGTAGAGGAA
27	eryB 1619 J F	CCGGATTTCGACCAGTTCCT
28	eryB 1619 J R	AGGGGTTTCAGATCCCTAAG
29	IS50 R 1	GACGATGAAGAGCAGAAG
30	DGEN 1	GGCCACGCGTCGACTAGTCAGNNNNNNNNNACGCC
31	IS50 R 2	TAGGAGGTCACATGGAGTCAGAT
32	DGEN 2	GGCCACGCGTCGACTAGTCAG

Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

To create an *eryD* mutation, pBG1 was first constructed by PCR amplification of a 400-bp internal fragment of *eryD* using Rm1021 genomic DNA as a template and primers 1 and 2. The amplicon was restricted and ligated into the suicide vector pKNOCK-Gm (Alexeyev, 1999) using *Bam*HI and *Cla*I. pBG1 was transformed into competent *E.coli* S17-1 and selected for using Gm 20. Single crossover *eryD* mutants, SRmA855 and SRmA856, were isolated by conjugating pBG1 into Rm1021 and SRmA449 respectively and selecting for Sm<sup>R</sup> and Gm<sup>R</sup> transconjugants as previously described (Poysti & Oresnik, 2007; Richardson *et al.*, 2004). Colonies were screened for phenotype and mutations were verified by sequencing.

To construct a strain carrying a mutation in *SMc01615*, the gene upstream of *tpiB*, putatively encoding a DeoR-type regulator, a 1,053-bp product was PCR amplified from Rm1021 genomic DNA with primers 9 and 10, digested with *Apa*I and *Sst*I, and cloned into pBluescript SK to yield pBP149. pBP149 was digested with *Sma*I into which a *Sma*I-digested *nptII* cassette from pMM22 was ligated to generate pBP151. pBP151 was subsequently recloned into pJQ200SK to produce pBP173. pBP173 was mobilized into Rm1021 with MT616 and selected for double recombinants. Colonies were screened onto selective media and sequenced to verify the insertion of *nptII* in *SMc01615*. One such mutant was designated SRmA966.

Two different methods were used to construct constitutively expressed erythritol genes. Open reading frames of *eryA*, *eryB* and *eryD* were amplified using primers 5 and 6, 7 and 8, and 3 and 4 respectively and cloned into the broad host-range vector, pRK7813. The 5' primers 5, 7 and 3 were all designed to have a ribosome binding site approximately 8 bp upstream of the annotated ATG to allow constitutive expression.



pBG2 was constructed by using *Pst*I and *Bam*HI restriction sites that were added to the primers to digest and ligate the 960 bp *eryD* ORF into pRK7813 such that it would be expressed from the *lacZ* promoter. pBG3 and pHY112 were constructed similarly using the *eryA* ORF with *Hind*III and *Eco*RI restriction sites and the *eryB* ORF with *Pst*I and *Bam*HI restriction sites. Additionally, the open reading frames for *eryA*, *eryC* and *eryD* were recombined into pCO37 from the *S. meliloti* ORFeome using the methods previously described (House *et al.*, 2004; Jacob *et al.*, 2008; Schroeder *et al.*, 2005).

### 2.2.3 Genetic manipulations

Conjugations between *E. coli* strains and *S. meliloti* were performed as previously described using the mobilizing strain MT616 (Finan *et al.*, 1985). Transposon mutagenesis of Rm1021 using pRK602 was carried out as previously described (Finan *et al.*, 1985). Mutants were routinely single colony purified three times and subsequently transduced into Rm1021 with phage  $\Phi$ M12 to ensure that the mutation was 100% genetically linked to the transposon (Finan *et al.*, 1985). The erythritol complementing cosmid pEL6 was isolated by conjugating a cosmid bank (Wang *et al.*, 2006) of *S. meliloti* into SRmA723 and plating on defined medium using erythritol as a sole carbon source to select for complementation. The cosmid pEL6 was then mutagenised by mating the cosmid into the Tn5-B20 carrying EcA101 (Clark *et al.*, 2001) and selecting for co-transfer of the cosmid and a Tn5-B20 marker when mated out of EcA101 and into DH5 $\alpha$  yielding pHC1- pHC12 essentially as previously described (Poysti *et al.*, 2007). The position of Tn5-B20 insertions was verified using arbitrary PCR as previously described (Miller-Williams *et al.*, 2006; Poysti *et al.*, 2007). Insertions confirmed to be

in the erythritol locus were recombined into Rm1021 as previously described (Glazebrook & Walker, 1991). The inserts were confirmed using arbitrary PCR and were transduced into Rm1021 yielding SRmA778-SRmA787.

#### **2.2.4 *β*-galactosidase assays**

Initial expression studies were carried out using *lacZ* fusions that were isolated in pEL6 (Table 2.1). Plasmids containing *lacZ* fusions were conjugated into the *lac* – derivative of Rm1021, RmG212. The assays were essentially carried out as previously described (Clark *et al.*, 2001; Oresnik *et al.*, 1998).

#### **2.2.5 RNA isolation and cDNA synthesis**

Bacterial cultures were grown to an OD<sub>600</sub> of approximately 0.8 in VMM medium containing either 15 mM erythritol and 15 mM glycerol or 15 mM glycerol alone. Cells were harvested by centrifugation and resuspended in TE buffer with lysozyme (0.4 mg/ml). RNA was then isolated using the QIAgen RNA isolation technique previously described (Barnett *et al.*, 2004). All RNA samples were treated twice with Qiagen on-column RNase-free DNase kit during the DNase step to remove DNA contamination.

The RNA was analyzed spectrophotometrically and the absorbance ratio of the sample at 260 nm and 280 nm was compared to ensure the purity of the RNA sample. The concentration of RNA was determined using the extinction coefficient for RNA (0.027 (μg/ml) cm<sup>-1</sup>). First strand cDNA synthesis was performed using 1 μg of RNA as suggested by the supplier. The quality of the cDNA synthesis was checked by electrophoresis and quantitated spectrophotometrically.

### **2.2.6 Junction PCR**

To determine the extent of transcripts, PCR reactions across putative junctions in the erythritol locus were performed using the cDNA from either induced cells grown on erythritol/glycerol or non-induced cells grown on glycerol. An identical PCR reaction using the RNA samples that were used for the cDNA synthesis as a template was carried out to control for possible genomic DNA contamination. A similar reaction was also performed using genomic template as a positive primer control. PCR reactions across three gene junctions were performed using primers 21 and 22, 23 and 24, 25 and 26 (Table 2.2).

### **2.2.7 Quantitative RT-PCR**

Approximately 100 ng of cDNA sample was used as a template for quantitative RT-PCR. Reactions were performed using the SYBR green RT-PCR kit from Invitrogen as recommended by manufactures. Primers 11 and 12 were used to analyze *SMc01628* transcription; 13 and 14 were used to analyze *eryD* transcription; 15 and 16 were used to analyze *eryA* transcription; 17 and 18 to analyze *eryC* transcription and finally 19 and 20 to analyze *SMc01615* transcription. Primers were designed upstream of all polar insertion mutations in strains in this work. The RT-PCR reaction was performed using a Cepheid Smart Cycler with the following program: Stage 1, 95°C for 120s, 1 time; Stage 2, 95°C for 15s followed by 60°C for 30s, repeated 40 times; Stage 3, melting curve analysis of PCR products.

### **2.2.8 Transport assay**

Transport assays were carried out essentially as previously described (Poysti *et al.*, 2007). Labelled [ $^{14}\text{C}$ -1] erythritol (370 MBq mmol $^{-1}$ ) was purchased from American Radiolabeled Chemicals Ltd. (St. Louis, Missouri). Transport assays were initiated by the addition of [ $^{14}\text{C}$ ] erythritol to a final concentration of 2  $\mu\text{M}$  to the samples and aliquots of 0.25 or 0.5 ml were withdrawn at appropriate time points and rapidly filtered through a Millipore 0.45  $\mu\text{m}$  Hv filter on a Millipore sampling manifold. Samples were taken every 15 or 20 seconds and continued for up to 2 minutes. The amount of radioactivity retained by the cells was quantitated by a liquid scintillation counter (Beckman LS6500).

### **2.2.9 Competition assay**

Alfalfa plants were grown as previously described except that Leonard jar assemblies that were made from Magenta jars (Sigma) were utilized (Oresnik *et al.*, 1994). Competition for nodule occupancy assays were carried out as previously described (Oresnik *et al.*, 1999).

## **2.3 Results**

### **2.3.1 Isolation of erythritol transport and catabolic mutants**

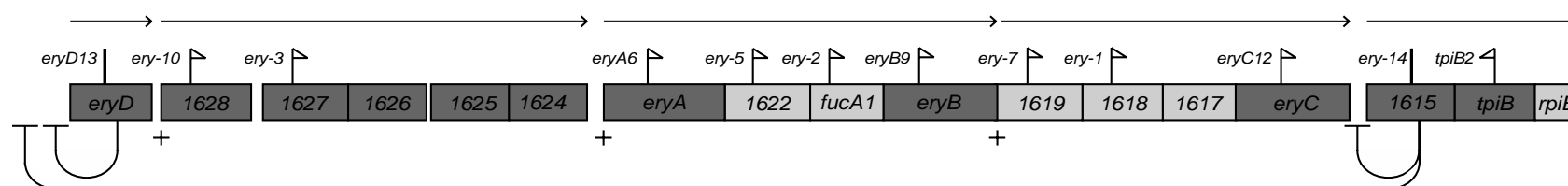
Comparison of the putative erythritol utilization locus of *S. meliloti* with that of *R. leguminosarum* shows that the organization of the catabolic genes is very different (Yost *et al.*, 2006). In addition, the ABC type transporter genes that are found downstream from *eryD* in *S. meliloti* (*SMc01628-SMc01624*) do not appear to be orthologues of

*eryEFG* in *R. leguminosarum*, suggesting that components involved in erythritol transport may not always be found in close proximity to the putative erythritol catabolic genes *eryA*, *eryB* and *eryC*. In an attempt to isolate genes necessary for the transport of erythritol a Tn5 mutagenesis of the wild type was carried out.

Approximately 3,000 transposon mutants were initially screened for their inability to grow on erythritol as a sole carbon source. The single mutant that was isolated was purified and designated SRmA428 (Table 2.1). The transposon was subsequently transduced into Rm1021 and the transductants were retested for phenotype. The transduced strain was designated SRmA465. The transposon was found to be within *SMc01627*, which is predicted to encode a permease that appears to be part of an ABC-type transporter located between the putative negative regulator *eryD* and the gene encoding the putative erythritol kinase, *eryA*.

Based on the close proximity of the putative start and stop codons found immediately up and downstream of each of the genes annotated as being necessary for erythritol catabolism, it was initially assumed that all the genes from *eryD* to *eryC* constituted a single operon (Figure 2.1). Combined with the ABC gene transporter gene locus, this suggests close proximity of all the erythritol utilization genes. Attempts to complement SRmA465 with the complementing cosmid for *tpiB*, pNP163 (Poysti & Oresnik, 2007), were unsuccessful, suggesting that pNP163 did not contain the entire locus necessary for erythritol catabolism. To obtain a complementing cosmid for the erythritol catabolic locus, a cosmid bank was mated *en masse* into SRmA465

**Figure 2.1.** Predicted ORFs and isolated mutant alleles in the erythritol locus of *S. meliloti*.



Vertical lines with flags represent the location and orientation of Tn5::B20 insertion alleles. Vertical lines without flags represent other insertion mutations. Arrows above the diagram represent transcriptional units as defined by this work. Lines below the diagram represent regulation as defined by this work while '+'s are positioned at the beginning of other inducible transcripts (see text for details). Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology

(Wang *et al.*, 2006) and the *S. meliloti* transconjugants that had regained their ability to use erythritol as a sole carbon source were isolated. One such isolate, that had the ability to complement the *SMc01627* mutation in SRmA465 as well as the *tpiB2* allele in SRmA584, was retained and designated pEL6 (Table 2.1).

To generate additional mutations in the erythritol locus a saturation mutagenesis of pEL6 was performed using Tn5::B20 to generate representative insertions across pEL6. Inserts were identified by generating a PCR fragment from the IS50 element on the Tn5-B20 using an arbitrary PCR protocol, and sequencing the products using a primer designed to the IS50 element (Poysti *et al.*, 2007). The plasmids carrying insertions within the erythritol locus were designated pHc1-12 (Table 2.1, Figure 2.1). Each of these insertions was marker exchanged into the chromosome of Rm1021 to yield the corresponding genomic mutation as described in Materials and Methods. These mutants were designated SRmA778-787 (Table 2.1). Each of these strains was unable to use erythritol as a sole carbon source. Also each strain containing an insertion in a gene downstream of *eryA* was unable to grow when erythritol was present with a secondary carbon source such as glycerol. Similar toxic effects have been previously noted for erythritol as well as other sugars in both *R. leguminosarum* and *E. coli* (Adhya & Shapiro, 1969; Power, 1967; Richardson *et al.*, 2004; Sperry & Robertson, 1975b; Yost *et al.*, 2006).

Although a good distribution of transposon insertions across the transport and catabolic genes of the locus were isolated, insertions in the two putative regulators were not (Figure 2.1). Therefore single crossover mutations in *eryD* and *SMc01615* were constructed giving rise to SRmA885 and SRmA966 respectively. Neither strain was able

to use erythritol as a sole carbon source. In addition it was found that SRmA966 was unable to grow on glycerol in the presence of erythritol. Since *eryD* and *SMc01615* are annotated as negative regulators, these phenotypes suggested polarity on downstream genes and were consistent with our hypothesis that the entire locus was transcribed as a single unit.

### ***2.3.2 Erythritol utilization genes are in different complementation groups***

To address the hypothesis that the erythritol locus in *S. meliloti* was a single operon, plasmids containing inserts within genes that were presumed to be necessary for erythritol catabolism were conjugated into each of SRmA778-SRmA787 to test for complementation. In addition, these plasmids were also conjugated into SRmA885 (*eryD*), SRmA966 (*SMc01615*), and SRmA584 (*tpiB*). The resulting transconjugants were screened for their ability to utilize erythritol as a sole carbon source. The results show that contrary to our initial assumptions, the erythritol locus in *S. meliloti* does not consist of a single transcript from *eryD* to *eryC* (Table 2.3). The locus appears to consist of at least 5 transcripts; a transcript containing *eryD*, a transcript encoding the components for one or more ABC transporters (*SMc01628-SMc01624*), a transcript that spans *eryA* through *eryB*, a transcript containing *eryC*, and a transcript containing the genes *SMc01615*, *tpiB*, and *rpiB*. Interestingly, since *eryD* appears to be transcribed independently, the inability of SRmA885 to use erythritol as a sole carbon source cannot be explained by polar effects on the ABC transporter. Similarly, SRmA428 (*SMc01627*) is not polar on the downstream erythritol catabolic genes.



**Table 2.3.** Erythritol catabolic genes are in separate complementation groups

Strain	Relevant genotype	Plasmid						
		---	pHC10	pHC3	pHC6	pHC9	pHC12	pEL6
Rm1021	Wild-type	+	+	+	+	+	+	+
SRmA855	<i>eryD</i>	-	+	+	+	+	+	+
SRmA785	<i>SMc01628</i>	-	-	-	+	+	+	+
SRmA780	<i>SMc01627</i>	-	-	-	+	+	+	+
SRmA782	<i>eryA</i>	-	+	+	-	+/-	+	+
SRmA784	<i>eryB</i>	-	+	+	+/-	-	+	+
SRmA787	<i>eryC</i>	-	+	+	+	+	-	+
SRmA966	<i>SMc01615</i>	-	+	+	+	+	+	+
SRmA584	<i>tpiB</i>	-	+	+	+	+	+	+

Growth is as follows: +, wild-type; +/-, marginal growth; -, no growth.

Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J.

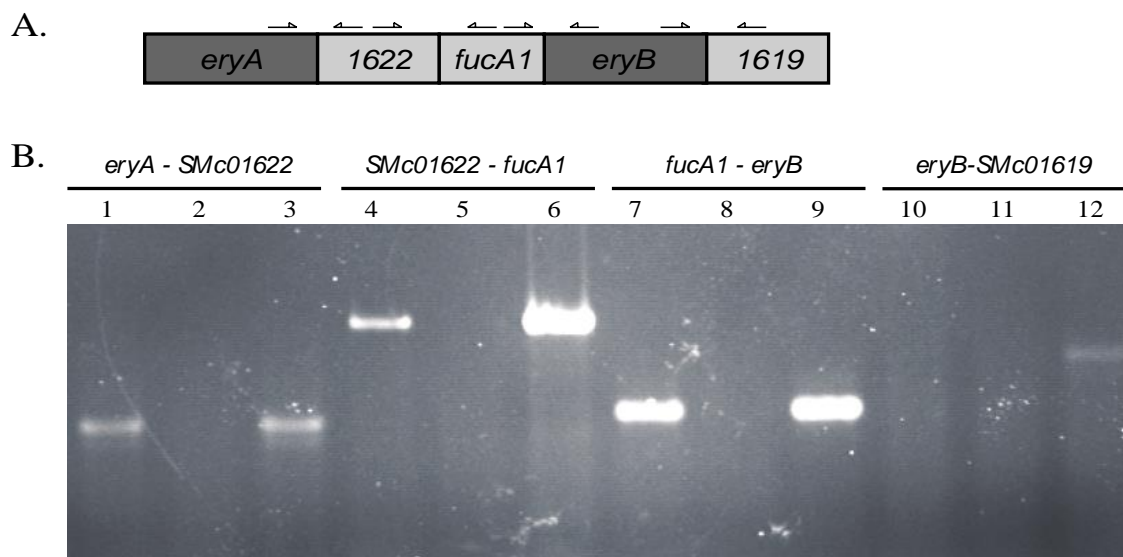
Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

### ***2.3.3 eryA and eryB are contained in a single transcript***

Although the complementation data are consistent with *eryA* and *eryB* being in a single transcript (Table 2.3), the data could be viewed as ambiguous with respect to *eryA* and *eryB* (Table 2.3). We note that SRmA782, containing *eryA*, as well as SRmA784, containing *eryB*, often grew poorly when carrying pHC6, pHC9, or pHC11. We reasoned that if *eryA*, *SMc01622*, *fucA1*, and *eryB* constituted a single transcript, they should all be present on a single mRNA. To test this, total RNA from erythritol-grown Rm1021 was isolated and used to create cDNA. Primers to the junctions of each of these genes were designed. In each case we were able to generate a PCR product from the cDNA suggesting that these genes indeed were within a single transcript (Figure 2.2). Primers designed to amplify the junction between *eryB* and *SMc01619* did not generate a PCR product from cDNA suggesting that *eryB* was the final gene in this transcript.

### ***2.3.4 Erythritol locus does not contain uncharacterized genes involved in erythritol catabolism***

The metabolic steps necessary for the breakdown of erythritol in *B. abortus* have been shown to consist of 5 steps (Sperry & Robertson, 1975a). Of the 5 reactions, the genes for only three of these, *eryA*, *eryB*, and *eryC* have been discovered (Sangari *et al.*, 2000). The two transcripts that contain *eryA*, *eryB* and *eryC* in *S. meliloti* include 5 other genes (Figure 2.1). These genes encode a putative oxidoreductase (SMc01622), a putative aldolase (FucA1), a putative class II aldolase/adducin protein (SMc01619), a putative sugar kinase (SMc01618), and a putative hydrolase (SMc01617). To determine if any of these genes encoded proteins that play a role in the catabolism of erythritol,

**Figure 2.2.** Junction PCR

**A.** Schematic diagram showing the arrangement of genes between *eryA* and *SMc01619*. Arrows represent approximate positions of primers that were used for junction PCR analysis. **B.** Junction PCR of erythritol induced cDNA. Lane 1-3, 4-6, 7-9, and 10-12 each represent primers that span the junction between two genes as labeled. Lanes 1, 4, 7 and 10 show the PCR product using cDNA synthesized from isolated RNA from erythritol induced wild-type cells as template. Lanes 2, 5, 8 and 11 show PCR product using isolated RNA from erythritol induced wild-type cells as template. Lanes 3, 6, 9 and 12, show PCR product using wild-type genomic DNA as template. Reproduced with modification from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

*eryA*, *eryB*, and *eryC* were introduced and over-expressed in each of the mutants that we had previously isolated (Table 2.4).

When *eryB* was overexpressed in SRmA781, SRmA779, and SRmA784, carrying mutations in *SMc01622*, *fucA1*, and *eryB* respectively, each strain was complemented for the inability to utilize erythritol as a sole carbon source (Table 2.4). SRmA781, carrying an *eryA* mutation, was not complemented by the introduction of *eryB*. Taken together these data strongly suggest that only *eryA* and *eryB* are necessary for erythritol breakdown and that the insertions in SRmA781 and SRmA779 are polar on *eryB*. Similarly, it was found that SRmA783 and SRmA778 were complemented by the introduction of *eryC* on a plasmid suggesting that proteins encoded by *SMc01619*, *SMc01618*, and *SMc01617* do not play a role in erythritol catabolism (Table 2.4). These data suggest that the transcript that contains *eryC*, also contains *SMc01617-SMc01619* (Table 2.3, Table 2.4).

Introduction of a plasmid containing *eryA* into SRmA782 did not complement the ability to utilize erythritol (Table 2.4). Although this is not surprising since all the data presented support the hypothesis that an *eryA* mutation has a polar affect on *eryB* (Table 2.3, Figure 2.2, Table 2.4), it is worth noting that when *eryA* was expressed in SRmA782, the strain was unable to grow on a medium that contained erythritol with a second carbon source that could support growth (Table 2.4). This correlates with previous suggestions that toxic effects are the result of the build-up of a phosphorylated intermediate (Adhya & Shapiro, 1969; Power, 1967; Richardson *et al.*, 2004; Sperry & Robertson, 1975b; Yost *et al.*, 2006).

**Table 2.4.** Complementation of mutants with erythritol catabolic genes

Strain	Genotype	Plasmid											
		None			pBG3 ( <i>eryA</i> <sup>+</sup> )			pBG14 ( <i>eryB</i> <sup>+</sup> )			pHY112 ( <i>eryC</i> <sup>+</sup> )		
		Ery	Ery/glyc	Glyc	Ery	Ery/glyc	Glyc	Ery	Ery/glyc	Glyc	Ery	Ery/glyc	Glyc
Rm1021	Wild-type	+	+	+	+	+	+	+	+	+	+	+	+
SRmA782	<i>eryA</i>	-	+	+	-	-	+	-	+	+	-	+	+
SRmA781	<i>SMc01622</i>	-	-	+	-	-	+	+	+	+	-	-	+
SRmA779	<i>fucA1</i>	-	-	+	-	-	+	+	+	+	-	-	+
SRmA784	<i>eryB</i>	-	-	+	-	-	+	+	+	+	-	-	+
SRmA783	<i>SMc01619</i>	-	-	+	-	-	+	-	-	+	+	+	+
SRmA778	<i>SMc01618</i>	-	-	+	-	-	+	-	-	+	+	+	+
SRmA787	<i>eryC</i>	-	-	+	-	-	+	-	-	+	+	+	+

Growth is as follows: +, wild-type; +/-, marginal growth; -, no growth.

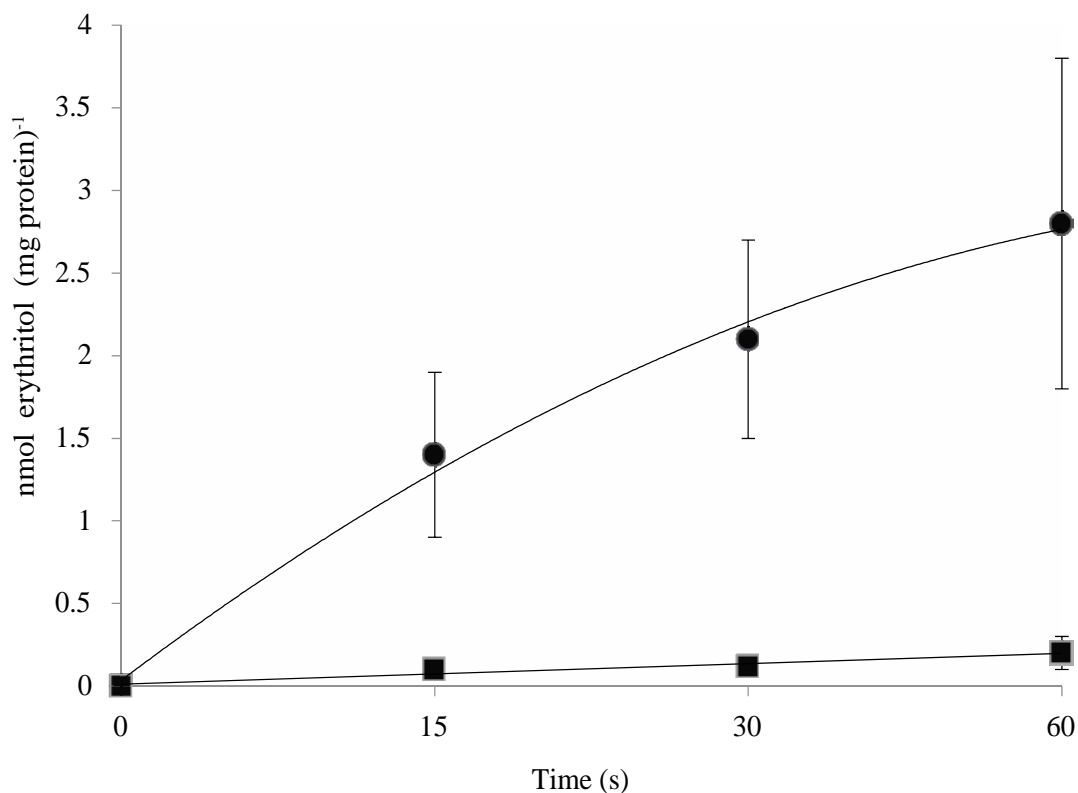
Ery/glyc media contained both 15 mM erythritol and 15 mM glycerol.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

***2.3.5 Components of the ABC-type transporter defined by SMc01628-SMc01624 are necessary for transport of erythritol***

BLAST analysis using Hyp-EryEFG from *R. leguminosarum* VF39SM against the *S. meliloti* Rm1021 database identified SMb20349-SMb20352 as having the highest similarity at the amino acid level (about 67% for EryEFG, about 47% for Hyp). Data from our transposon mutagenesis experiments show that inserts that affect the *SMc01628-SMc01624* lead to an inability to use erythritol as a sole carbon source (Table 2.3). Since duplication of function is not uncommon in *Rhizobium* (Oresnik *et al.*, 1998; Renalier *et al.*, 1987; Schlüter *et al.*, 1997; Schwedock & Long, 1992), we wanted to determine if erythritol transport was inducible in *S. meliloti* and if the transporter genes found between *eryA* and *eryD* were indeed necessary for erythritol transport.

Rm1021 grown in defined medium containing glucose or glycerol was not capable of taking up labelled erythritol at levels above those observed in dead cell controls. In contrast, Rm1021 grown in medium containing either erythritol or erythritol and glycerol displayed robust transport rates (Figure 2.3). To determine if the ABC transporter genes between *eryA* and *eryD* were necessary, SRmA465, carrying a Tn5 insertion in *SMc01627*, was grown in medium containing erythritol and glycerol. The data demonstrate that SRmA465 was unable to transport labelled erythritol into the cell suggesting that these gene(s) encode proteins used in an ABC transporter involved in the transport of erythritol (Figure 2.3).

**Figure 2.3.** Erythritol transport assays

*S. meliloti* wild-type (Rm1021, circles) and the erythritol transport mutant SRmA465 (SMc01627, squares), induced with defined medium containing erythritol and glycerol. Where not shown error bars are smaller than the symbol. Error bars represent standard deviation (n=3). Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

### ***2.3.6 EryD and SMc01615 negatively regulate their own transcripts and affect the expression of other ery transcripts***

We had previously shown that *tpiB* could be induced by low concentrations of erythritol thus allowing it to complement a *tpiA* mutation in Rm1021 (Poysti & Oresnik, 2007). Since *tpiB* is in an operon that includes a DeoR-type negative regulator, encoded by *SMc01615*, it was postulated that this regulator would regulate *tpiB* expression and that second site mutations that allowed phenotypic suppression of *tpiA* would be readily isolatable (Poysti & Oresnik, 2007). Although greater than  $10^{10}$  cells were plated, this class of second site mutations was never isolated, which led to the suggestion that *tpiB* may not be solely under the regulation of *SMc01615* (Poysti & Oresnik, 2007). To address the regulation of *tpiB* and the other components of the erythritol catabolic genes RT-PCR experiments were carried out.

Consistent with preliminary  $\beta$ -galactosidase assays using genetic fusions (Table 2.5), it was found that, in the presence of erythritol, components of the erythritol catabolic pathway were induced 5 to 10 fold when compared to transcript levels found in glycerol grown cells (Table 2.6). Similar analysis with the *eryD* mutant, SRmA885, showed greater than 10-fold expression of *eryD* in glycerol conditions compared to wild-type levels demonstrating that, consistent with the annotation, EryD is a negative regulator that regulates the expression of its own transcript (Table 2.6). Intriguingly, every other transcript in the erythritol locus was expressed at, or below, background levels when SRmA885 (*eryD*) was grown under inducing conditions; a phenotype more commonly associated with a positive regulator (Table 2.6). The inability of SRmA885 to induce erythritol transport and catabolic genes is consistent with its inability to grow



**Table 2.5.** Induction of erythritol locus  $\beta$ -galactosidase gene fusions by erythritol

Plasmid	Genotype	Glycerol	Erythritol
None	Wild-type	72(42)	154 (78)
pHC3	<i>SMc01627::Tn5-B20</i>	64(21)	1423 (90)
pHC6	<i>eryA::Tn5-B20</i>	115(100)	1090 (90)
pHC9	<i>eryB::Tn5-B20</i>	224(26)	932 (212)
pHC1	<i>SMc01618::Tn5-B20</i>	215 (20)	524 (195)

Values represent induction of gene fusions in the erythritol complementing cosmid pEL6, expressed in the *lac*- background RmG212.  $\beta$ -galactosidase activity is expressed in Miller units. Values are the average of 3 independent replicates. Values in parentheses represent standard error.

**Table 2.6.** Regulation of gene expression by *eryD* and *SMc01615* analyzed by qRT-PCR

Strain	Relevant genotype	Carbon source	<i>eryD</i>	<i>SMc01628</i>	<i>eryA</i>	<i>eryC</i>	<i>SMc01615</i>
Rm1021	Wild-type	Gly	1	1	1	1	1
Rm1021	Wild-type	Ery/glyc	11.0	19.4	16.9	11.9	6.5
SRmA885	<i>eryD</i>	Glyc	14.7	0.8	0.1	0.9	0.3
SRmA885	<i>eryD</i>	Ery/glyc	18.6	0.3	0.1	1.1	0.6
SRmA966	<i>SMc01615</i>	Glyc	5.5	0.7	0.6	1.4	14.6
SRmA782	<i>eryA</i>	Glyc	0.3	0.3	1.1	1.1	0.3
SRmA782	<i>eryA</i>	Ery/glyc	1.47	0.8	0.4	0.7	0.9

Data expressed as  $2^{\Delta\Delta Ct}$  and represents fold expression over uninduced Rm1021 grown in glycerol. The experiment also included *SMc00128* as an internal control (Krol & Becker, 2004). The table represents data from a single experiment. The experiment was repeated three times showing consistent results. Ery/glyc media contained both 15 mM erythritol for induction and 15 mM glycerol for growth.

Reproduced from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J.

Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

using erythritol as a sole carbon source (Table 2.3).

To address what role SMc01615 plays in the regulation of erythritol catabolic genes, a strain (SRmA966) containing a mutation in this gene was constructed. Since this strain is unable to grow in the presence of erythritol, only the effect of the mutation in non-inducing conditions could be analyzed. The results clearly show that lack of *SMc01615* resulted in a greater than 10-fold induction of its own transcript under non-inducing conditions. We note that in SRmA966 we observed an increased level of transcription of *eryD* suggesting that SMc01615 may also play a role in regulating *eryD* (Table 2.6).

### ***2.3.7 The inability to catabolise erythritol affects gene expression***

It is unclear if the regulators of *ery* catabolic genes respond to either erythritol and/or an erythritol metabolite as an effector molecule (Sangari *et al.*, 2000; Yost *et al.*, 2006). Since the organization of the operons containing the *ery* catabolic and transport genes is so different than either of *B. abortus* or *R. leguminosarum*, it was thought that investigating the effect of erythritol catabolic mutations on the regulation of these genes may provide some insight. Analysis of transcription in SRmA782 (*eryA*), revealed that only background levels of transcription occurred across the entire erythritol region when it was grown on media containing erythritol (Table 2.6). This suggests that in addition to its own induction, a functional EryA is essential for the induction of the entire erythritol locus including *tpiB* and *eryD*. In addition it suggests that the regulators of the erythritol locus likely respond to an intermediate of erythritol catabolism.

### ***2.3.8 The erythritol locus contains determinants for adonitol and arabitol utilization***

The transcripts that contain the genes necessary for erythritol catabolism (*eryA*, *eryB*, *eryC* and *tpiB*) also contain 6 other genes that are annotated as being involved in small molecule metabolism. Moreover, the genes that encode the components of the ABC transporter contain 2 ABC proteins and 2 permease components, formally making it possible that this may represent more than one transport system. To attempt to define what role the other genes in this cluster may play, representative insertion mutants from each of the transcripts (except *eryD*) were analyzed for carbon utilization phenotypes using Biolog PM1 and PM2 plates. The results from the PM1 and PM2 plates which suggested an inability to use a compound as a sole carbon source were verified by streaking the strains on defined media containing these as sole carbon sources. Intriguingly these results show that mutants with insertions in the ABC transporter transcript (SRmA783), the transcript containing *eryA* and *eryB* (SRmA782) as well as the transcript containing *eryC* (SRmA783) were unable to use adonitol or L-arabitol as a sole carbon source (Table 2.7). In contrast, the transcript that contains *tpiB* was only necessary for erythritol catabolism. In an effort to more precisely define which genes were necessary for the utilization of which sugar, other insertions in these transcripts were also tested. The majority of the strains carrying inserts in these transcripts were unable to use adonitol and L-arabitol in addition to erythritol (Table 2.7). This includes SRmA855 which contains an *eryD* mutation, suggesting that EryD may also regulate the locus during adonitol or L-arabitol catabolism. SRmA787, containing an insertion in *eryC*, was able to use adonitol and L-arabitol as sole carbon sources (Table 2.7). It is

**Table 2.7.** Polyol utilization phenotypes of mutations across the erythritol locus

Strain	Relevant genotype	Carbon source					
		Erythritol	Adonitol	L-Arabitol	D-Arabitol	Xylitol	Glycerol
Rm1021	Wild-type	+	+	+	+	-	+
SRmA885	<i>eryD</i>	-	-	-	+	ND	+
SRmA785	<i>SMc01628</i>	-	-	-	+	ND	+
SRmA782	<i>eryA</i>	-	-	-	+	ND	+
SRmA781	<i>SMc01622</i>	-	-	-	+	ND	+
SRmA779	<i>fucA1</i>	-	-	-	+	ND	+
SRmA784	<i>eryB</i>	-	+/-	-	+	ND	+
SRmA783	<i>SMc01619</i>	-	+/-	-	+	ND	+
SRmA778	<i>SMc01618</i>	-	-	-	+	ND	+
SRmA787	<i>eryC</i>	-	+	+	+	ND	+
SRmA966	<i>SMc01615</i>	-	+	+	+	ND	+
SRmA584	<i>tpiB</i>	-	+	+	+	ND	+

Growth is as follows: +, wild-type; +/-, marginal growth; -, no growth. ND indicates carbon utilization phenotype not determined. Reproduced with modification from Microbiology, volume 156, Geddes, B. A., B. S. Pickering, N. J. Poysti, H. Collins, H. Yudistira, and I. J. Oresnik, A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*, pp. 2970-2981, Copyright 2010, with permission from the Society for General Microbiology.

noteworthy to point out that the analysis also showed that although these strains were unable to grow on L-arabitol, they were still capable of utilizing D-arabitol (Table 2.7). It is also worth noting that Rm1021 could not grow using the closely related sugar alcohol xylitol as a sole carbon source (Table 2.7).

### ***2.3.9 Lack of erythritol, adonitol, and arabitol utilization does not affect competition for nodule occupancy in *S. meliloti****

It has recently been shown that erythritol utilization affects competition for nodule occupancy in *R. leguminosarum* (Yost *et al.*, 2006). In addition, erythritol utilization has been associated with avirulence in various *Brucella* strains (Köhler *et al.*, 2002). SRmA465 was used as representative strain since it lacked the ability to take up erythritol and was unable to grow on adonitol or arabitol (Figure 2.1, Table 2.7). Two different inoculation ratios (mutant:wild type) were tried and in every trial the ratio of strains isolated from the nodules was not significantly different from the inoculation ratio suggesting that the ability to utilize erythritol, adonitol, or arabitol does not play role in competition for nodule occupancy.

## **2.4 Discussion**

Our genetic and physical data show that the erythritol locus in *S. meliloti* consists of 5 transcripts. Of the catabolic genes present, only *eryA*, *eryB*, *eryC*, and *tpiB* appear to encode proteins that are necessary for erythritol catabolism (Table 2.3, Table 2.4). In addition, genes that encode the determinants for an ABC transporter are essential for the transport of erythritol into the cell (Figure 2.3). It is of note that genes encoding

SMb20349-51, which were predicted to be necessary for erythritol transport in *S. meliloti* on the basis of identity to the *R. leguminosarum* erythritol transporter (Yost *et al.*, 2006), do not have erythritol transport function based on our ability to show lack of erythritol transport in our mutants.

The original work outlining the catabolic pathway for erythritol catabolism provided evidence for 5 biochemical steps yielding dihydroxy-acetone phosphate (Sperry & Robertson, 1975a). Analyses done on *B. abortus*, *R. leguminosarum*, as well as the data presented here, only account for 3 of these genes as well as a dedicated triose-phosphate isomerase (Sangari *et al.*, 2000; Yost *et al.*, 2006). Complementing of each *eryA*, *eryB*, *eryC* and *tpiB* clearly show that the other genes at this locus do not encode enzymes that have the missing activities Table 2.4). Although our limited mutagenesis did provide evidence for the transporters necessary for erythritol, it did not identify either of these two missing activities. It should also be noted that although a mutation in the gene annotated as *rpiB* (*SMc01613*) did not affect the ability to use erythritol as a sole carbon source, the corresponding mutant in *R. leguminosarum* was unable to use erythritol (Poysti & Oresnik, 2007; Yost *et al.*, 2006). It is not known at this time how *rpiB* does, or does not, contribute to erythritol catabolism in either *R. leguminosarum* or *S. meliloti*. Conjugation of pEL6 (containing the entire *S. meliloti* *ery* locus) into *R. leguminosarum* strain Rlt100, a strain that does not have the capability of utilizing erythritol as a sole carbon source, allowed it to utilize erythritol. We suggest that perhaps the biochemical steps that were originally described in *B. abortus*, that are not accounted for by the genes present in the erythritol locus, may be due to enzymes that have enzymatic activity on a broad number of substrates and may not be specific for erythritol.

Alternatively, some of the erythritol associated proteins may have more than one enzymatic activity.

The regulation of erythritol catabolic genes in *R. leguminosarum* and *B. abortus* has been shown to be carried out by EryD which has been annotated as a negative regulator that falls into a DeoR category (Sangari *et al.*, 2000; Yost *et al.*, 2006). Consistent with this annotation, a *R. leguminosarum* strain carrying an *eryD* mutation was not diminished in its ability to grow on erythritol (Yost *et al.*, 2006). In contrast we found that *S. meliloti* carrying an *eryD* mutation was unable to use erythritol as a sole carbon source (Table 2.3). In addition if an *eryD* mutation was moved into a strain containing a *tpiA* mutation, it was unable to phenotypically suppress the inability to utilize glycerol associated with a *tpiA* mutation if grown on glycerol with 0.5 mM erythritol (Poysti & Oresnik, 2007). Taken together these data prompted a more thorough analysis of regulation that may be occurring at this locus.

Our data clearly show that EryD is a negative regulator (Table 2.6). In contrast, the transcription of the other operons at this locus appears to be at, or below, its basal level when *eryD* is absent (Table 2.6). These data are inconsistent with the hypothesis that *eryD* negatively regulates the genes encoding the ABC transporter, *eryA*, *eryB*, *eryC*, and *tpiB*. The data support the involvement of a positive regulator in erythritol catabolism in *S. meliloti*.

BLAST analysis clearly showed that EryD falls into the SorC subclass of DeoR regulators. We find it of interest that DalR, the regulator of arabinol utilization in *Klebsiella pneumonia* is also considered a member of this family (Heuel *et al.*, 1998). SorC has been shown to act as both a negative and a positive regulator of sorbitol



utilization in *K. pneumonia* (Wöhrl *et al.*, 1990). It has been shown that SorC is able to negatively regulate its own transcription. In addition Wohrl *et al.* (1990) were able to demonstrate lack of transcription of the sorbose operon in a non-polar background lacking *sorC*. Complementation of this mutation with a *sorC*<sup>+</sup> allele restored gene expression of the sorbitol operon. More recently the crystal structure of SorC has been solved, and based on the structural model, a model of the positive and negative transcriptional regulation has been proposed (de Sanctis *et al.*, 2009). The SorC data are analogous to what we observed with *eryD* (Table 2.6). Attempts at complementing the *eryD* mutation with pBG2 and pBG18 (*eryD*<sup>+</sup>) gave ambiguous results; in some cases transcription was observed, whereas in other cases it was not.

RT-PCR data also showed that SMc01615 negatively affects the regulation of its own transcript, as well as the transcription of *eryD* (Table 2.6). We note that SMc01615 contains a sugar-phosphate binding site (c10039 sugarP\_isomerase Super-family), suggesting that it does not respond directly to erythritol, but to a phosphorylated metabolite of the pathway. This is consistent with the data that show only basal levels of transcription of *SMc01615* in an *eryA* background. It is also supported by observation that *tpiB* is only necessary for erythritol catabolism in *S. meliloti* and that a strain carrying a *tpiA* mutation had its inability to grow on glycerol phenotypically suppressed when it was grown on glycerol in the presence of 0.5 mM erythritol, but not 0.5 mM adonitol or L-arabitol (Table 2.8). The constitutive expression of the *SMc01615* transcript in a *SMc01615* background suggests that although EryD is required for its induction, the effect is probably indirect and is a consequence of its effect on the catabolic genes

**Table 2.8.** Phenotypic suppression of a *tpiA* mutant by *tpiB* induction

Strain	Genotype	Glyc	Glyc/ery	Glyc/ado	Glyc/L-ara
Rm1021	Wild-type	+	+	+	+
SRmA455	$\Delta tpiA$	-	+	-	-
SRmA515	$\Delta tpiA, tpiB1$	-	-	-	-

Growth is as follows: +, wild-type; -, no growth. Glyc/ery. Glyc/ado, and Glyc/L-ara media contained 15 mM Glycerol with 0.5 mM erythritol, adonitol or L-arabitol, respectively. 0.5 mM of these carbon sources is not sufficient for growth of Rm1021.

(Table 2.6). The effect of SMc01615 on the transcription of *eryD* however, is likely direct.

In SRmA782 (*eryA*), transcription at the erythritol locus was abolished (Table 2.6). These data strongly support the hypothesis that transcription is dependent upon a metabolite from the erythritol pathway since even transcription of *eryD* and *SMc01615* were unable to be induced (Table 2.6). We note that EryD and SMc01615 contain different putative domains. EryD contains a domain that is classified as a part of the sugar-binding super family (cl04446) and SMc01615 contains a domain related to the sugar-phosphate isomerase super family (cl10039). Taken together this suggests that the regulators may not recognize the same metabolite.

The apparent role of this locus in erythritol catabolism was expected since *eryA*, *eryB* and *eryC* are all homologues of genes previously shown to be necessary for erythritol catabolism in *R. leguminosarum* and *B. abortus* (Sangari *et al.*, 2000; Yost *et al.*, 2006). The role of the locus in adonitol and L-arabitol metabolism is surprising as it was not suggested by the annotation of genes in the locus. The annotation suggests that the genes interspersed among the *ery* catabolic genes encode an oxidoreductase (*SMc01622*), an aldolase (*fucA1*), a putative class II aldolase/adducin protein (*SMc01619*), a sugar kinase (*SMc01618*), and a hydrolase (*SMc01617*). Based on the currently understood pathways of adonitol and arabitol catabolism as they were defined in *Aerobacter aerogenes* (Wood *et al.*, 1961), as well as what is known about polyol metabolism in *R. leguminosarum* (Primrose & Ronson, 1980; Ronson & Primrose, 1979; Stowers, 1985; Yost *et al.*, 2006), it is unclear which of these genes are necessary for the catabolism of either arabitol or adonitol. It may be pertinent to point out that orthologues

of *SMc01617-19* appear to form an independent operon in *Mesorhizobium loti* (*mlr3601*, *mlr3603*, and *mlr3604*), suggesting that these genes may in fact be a separate catabolic entity (Kaneko *et al.*, 2000).

An inability to catabolize erythritol has been shown to affect competition for nodule occupancy in *R. leguminosarum* bv. *viciae* (Fry *et al.*, 2001; Yost *et al.*, 2006), whereas the ability to catabolize adonitol in *R. leguminosarum* bv. *trifolii* did not affect competition for nodule occupancy (Oresnik *et al.*, 1998). Competition experiments indicated that a strain unable to transport erythritol, adonitol and L-arabitol, was as effective as the wild-type at competing for nodule occupancy on alfalfa. It would be of interest to determine whether the ability to catabolize erythritol by other rhizobia affects competition for nodule occupancy or whether there are host plant specific effects of erythritol catabolism on competition.

We are currently investigating the erythritol locus with respect to its makeup, gene arrangement and polyol catabolic ability throughout the rhizobia and other bacterial species. In addition, we are attempting to determine the precise function encoded by the genes that are necessary for adonitol and arabitol catabolism in *S. meliloti*. It is hoped that defining the roles of these genes will be of benefit both to the development of *S. meliloti* as a model organism for the study of plant/microbe interactions as well as providing insight into basic metabolic pathways.

### **Chapter 3**

#### **Genetic Characterization of a Complex Locus Necessary for the Transport and Catabolism of Erythritol, Adonitol, and L-Arabitol in *Sinorhizobium meliloti***

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology. This work was carried out by Barney Geddes with technical assistance from Travis McGill and Peiki Loay.

### 3.1 Introduction

Adonitol (ribitol) and L-arabitol, together with D-arabitol and xylitol comprise the 5-carbon sugar alcohols known as pentitols. These sugars range from being relatively abundant in nature and commonly found in many organisms (adonitol and D-arabitol), to being exceedingly rare and used, or synthesized, by relatively few organisms (L-arabitol) (Mortlock, 1984). Most of the work contributing to the current understanding of pentitol catabolism in bacteria has been carried out in *Klebsiella pneumoniae*, formerly known as *Aerobacter aerogenes* (Mortlock & Wood, 1964a, b; Mortlock *et al.*, 1965a; Mortlock *et al.*, 1965b; Wood *et al.*, 1961). The common route of their utilization involves oxidation to a 2-ketopentose (pentulose); D-ribulose, and L-xylulose for adonitol and L-arabitol respectively, and D-xylulose for D-arabitol and xylitol. This precedes a phosphorylation of the pentulose at the C-5 position. D-ribulose-5-phosphate and L-xylulose-5-phosphate are converted to D-xylulose-5-phosphate by epimerization, through which all pentitols enter central metabolism (Mortlock, 1984).

Rhizobia are robust in their metabolic capacity and have been studied for many years with respect to carbon metabolism since the use of carbon compounds is important in the rhizosphere prior to infection of the root tissue, as well as during symbiotic nitrogen fixation (Lodwig & Poole, 2003; Ramachandran *et al.*, 2011; White *et al.*, 2007). Enzymes that carry out pentitol oxidation reactions have also been identified in the Rhizobia. Two polyol dehydrogenases were identified from the extracts of mannitol-grown *S. meliloti* cells (Martinez de Drets & Arias, 1970). One of these acted specifically to oxidize D-sorbitol, while the other, which was identified as a D-arabitol dehydrogenase, was also capable of oxidizing D-mannitol. Subsequent work in *R.*

*leguminosarum* bv. *trifolii* identified five different polyol dehydrogenases including an inositol dehydrogenase, adonitol dehydrogenase, xylitol dehydrogenase, dulcitol dehydrogenase and D-arabitol dehydrogenase (Primrose & Ronson, 1980). While the inositol and adonitol dehydrogenases were specific, the xylitol dehydrogenase was capable of oxidizing xylitol and sorbitol, and the dulcitol dehydrogenase was capable of oxidizing dulcitol, xylitol and sorbitol (Primrose & Ronson, 1980). Many of these dehydrogenases were shown to be affected by glucose-mediated catabolite repression (Ronson & Primrose, 1979). More recently the genetic locus for inositol catabolism was identified in *R. leguminosarum* and *S. meliloti* (Fry *et al.*, 2001; Kohler *et al.*, 2010). In both of these organisms the ability to catabolize inositol was shown to be important for competition for nodulation (Fry *et al.*, 2001; Kohler *et al.*, 2010).

In *R. leguminosarum*, plasmid curing experiments identified erythritol catabolic and transport functions to be plasmid localized (Yost *et al.*, 2006). Characterization of the locus revealed that it contained genes necessary for the transport and catabolism of erythritol, and that the inability to catabolize erythritol resulted in a reduced ability to compete for nodule occupancy against the wild type (Yost *et al.*, 2006). In *S. meliloti*, the erythritol locus was initially identified during genetic characterization of triose-phosphate isomerase mutations (Poysti & Oresnik, 2007). The *S. meliloti* erythritol locus was subsequently characterized and it was found that although the locus contained the standard *eryA*, *eryB*, *eryC*, and *eryD* genes used for the catabolism of erythritol, the ability to catabolize erythritol did not affect competition for nodule occupancy in *S. meliloti*. The locus was also found to be necessary for the utilization of adonitol and L-arabitol as sole carbon sources (Geddes *et al.*, 2010; Poysti & Oresnik, 2007).

The genetic organization of the *S. meliloti* erythritol locus is different from other previously characterized erythritol loci (Sangari *et al.*, 2000; Yost *et al.*, 2006).

Characterization of this locus left 5 uncharacterized genes that encode proteins annotated as being involved in small molecule metabolism, but not necessary for erythritol catabolism. Since it had been shown that this locus was also necessary for L-arabitol and adonitol catabolism (Geddes *et al.*, 2010), we wished to genetically characterize this locus with respect to the catabolism and transport of these pentitols.

## 3.2 Materials and Methods

### 3.2.1 Bacterial strains, plasmids and media

Bacterial strains and plasmids used and generated in this work are listed in Table 3.1. *S. meliloti* strains were grown routinely at 30°C on either Luria-Bertani (LB) as a complex medium (Sambrook *et al.*, 1989) or Vincent's minimal medium (VMM) as a defined medium (Vincent, 1970). Carbon sources were filter sterilized and added to VMM to a final concentration of 15 mM. When required, *S. meliloti* and *Escherichia coli* strains were grown in the following concentrations of antibiotic (µg/mL): chloramphenicol (Cm) 20; gentamicin (Gm) 20 or 60; kanamycin (Km) 20; neomycin (Nm) 200; rifampicin (Rf) 50; streptomycin (Sm) 200; tetracycline (Tc) 5. All antibiotics were filter sterilized before use.

### 3.2.2 Genetic techniques

Conjugations and transductions were carried out essentially as previously described (Finan *et al.*, 1988; Finan *et al.*, 1984). The *S. meliloti* ORFeome platform was



**Table 3.1.** Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
<b>Strain</b>		
<i>S. meliloti</i>		
Rm1021	SU47 <i>str-21</i> , Sm <sup>R</sup>	(Meade <i>et al.</i> , 1982)
SRmA355	Rm5000 <i>tpiB1</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA428	Rm1021 <i>mptB</i> ::Tn5, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA449	Rm1021 $\Delta$ <i>tpiA</i>	(Poysti & Oresnik, 2007)
SRmA465	$\phi$ (SRmA428) $\rightarrow$ Rm1021 <sup>a</sup> , Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA515	$\Delta$ <i>tpiA</i> , <i>tpiB1</i> $\phi$ (SRmA355) $\rightarrow$ SRmA449 <sup>a</sup> , Gm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA584	Rm1021 <i>tpiB2</i> , Nm <sup>R</sup>	(Poysti & Oresnik, 2007)
SRmA723	Rm1021 <i>mptB</i> ::Tn5-233, Gm <sup>R</sup> Sp <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA766	Rm1021 <i>xykB</i> ::Tn5	(Geddes & Oresnik, 2012a)
SRmA780	Rm1021 <i>mptB3</i> , Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA782	Rm1021 <i>eryA6</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA784	Rm1021 <i>eryB9</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA785	Rm1021 <i>mptA10</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA787	Rm1021 <i>eryC12</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmA885	Rm1021 <i>eryD13</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmD208	Rm1021 $\Delta$ <i>lala</i>	(Geddes <i>et al.</i> , 2010)
SRmD209	Rm1021 $\Delta$ <i>rbtB</i>	(Geddes <i>et al.</i> , 2010)
SRmD210	$\Delta$ <i>tpiA</i> , <i>tpiB1</i> (SRmA515) $\Delta$ <i>pyc</i> , Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
SRmD211	Rm1021 $\Delta$ <i>pyc</i>	(Geddes & Oresnik, 2012a)
SRmD247	Rm1021 $\Delta$ <i>rbtA</i>	(Geddes & Oresnik, 2012a)
SRmD249	Rm1021 $\Delta$ <i>rbtC</i>	(Geddes & Oresnik, 2012a)
SRmD270	Rm1021 $\Delta$ <i>fucA1</i>	(Geddes & Oresnik, 2012a)
SRmD271	Rm1021 $\Delta$ <i>eryB</i>	(Geddes & Oresnik, 2012a)
<i>R. leguminosarum</i>		
Rlt100	<i>R. leguminosarum</i> bv. <i>trifolii</i> Sm <sup>R</sup> Wild-type	(Oresnik <i>et al.</i> , 1998)
Rlt103	Rlt100 <i>rbt-2</i> ::Tn5-B20	(Oresnik <i>et al.</i> , 1998)
<i>E. coli</i>		
DH5 $\alpha$	$\lambda$ $\Phi$ 80dlacZ <sup>o</sup> M15 <sup>o</sup> ( <i>lacZYA-argF</i> )U169 <i>recA1endA1 hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) <i>supE44 thi-1 gyrA relA1</i>	(Hanahan, 1983)
DH5 $\alpha$ Rif	Rifampicin resistant DH5 $\alpha$	(House <i>et al.</i> , 2004)
DH5 $\alpha$ $\lambda$ pir	$\lambda$ pir lysogen of DH5 $\alpha$	(House <i>et al.</i> , 2004)
MM294A	<i>pro-82-thi-1 hsdR17 supE44</i>	(Finan <i>et al.</i> , 1986)
MT607	MM294A <i>recA56</i>	(Finan <i>et al.</i> , 1986)
MT616	MT607 (pRK600)	(Finan <i>et al.</i> , 1986)
S17-1	<i>recA</i> derivative of MM294A with integrated RP4-2 (Tc::Mu::Km::Tn7)	(Simon <i>et al.</i> , 1983)
EcA101	MT607 $\Omega$ Tn5-B20, Kan <sup>R</sup>	(Clark <i>et al.</i> , 2001)

Plasmids		
pBG1	pKNOCK-Gm + 400 bp internal <i>eryD</i> fragment ( <i>Bam</i> HI / <i>Cla</i> I), Gm <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG2	pRK7813/ <i>eryD</i> ( <i>Pst</i> I, <i>Bam</i> HI), Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG3	pRK7813/ <i>eryA</i> , ( <i>Eco</i> RI, <i>Hind</i> III), Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG11	pJQ200SK/ <i>pyc</i> flanking regions	(Geddes & Oresnik, 2012a)
pBG13	pCO37/ <i>eryA</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG14	pCO37/ <i>eryC</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG18	pCO37/ <i>eryD</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pBG23	pCO37/ <i>rbtA</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012a)
pBG24	pCO37/ <i>fucA1</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012a)
pBG25	pCO37/ <i>lala</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012a)
pBG26	pCO37/ <i>rbtB</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012a)
pBG28	pJQ200SK/ <i>fucA1</i> flanking regions	(Geddes & Oresnik, 2012a)
pBG30	pJQ200SK/ <i>eryB</i> flanking regions	(Geddes & Oresnik, 2012a)
pBG31	pCO37/ <i>rbtC</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012a)
pCO37	pRK7813 containing <i>attB</i> sites, Gateway® compatible destination vector	(Jacob <i>et al.</i> , 2008)
pEL6	CX1 erythritol complementing cosmid, Tc <sup>R</sup>	(Wang <i>et al.</i> , 2006)
pHY112	pRK7813 / <i>eryB</i> , Tc <sup>R</sup>	(Geddes <i>et al.</i> , 2010)
pKNOCK-Gm	Suicide vector, Gm <sup>R</sup>	(Alexeyev, 1999)
pMK2014	FRT- <i>ccdB</i> -Cam <sup>R</sup> -FRT cassette, Pen <sup>R</sup>	(House <i>et al.</i> , 2004)
pMK2015	FRT- <i>ccdB</i> -Cam <sup>R</sup> -FRT cassette, Pen <sup>R</sup>	(House <i>et al.</i> , 2004)
pMK2016	<i>oriV oriT<sub>ColE1</sub></i> with FRT cassette from pMK2014, Str <sup>R</sup> Spc <sup>R</sup>	(House <i>et al.</i> , 2004)
pMK2017	<i>oriV<sub>R6K</sub> oriT<sub>RP4</sub></i> with FRT cassette from pMK2015, Tet <sup>R</sup>	(House <i>et al.</i> , 2004)
pRK600	pRK2013 <i>nptI::Tn9</i> , Cm <sup>R</sup>	(Finan <i>et al.</i> , 1986)
pRK602	pRK600ΩTn5, Cm <sup>R</sup> , Nm <sup>R</sup>	(Finan <i>et al.</i> , 1985)
pRK7813	Broad-host-range vector, Tc <sup>R</sup>	(Jones & Gutterson, 1987)
pJQ200SK	Gene replacement suicide vector, Gm <sup>R</sup>	(Quandt & Hynes, 1993)
pXINT129	λint and <i>xis</i> driven by P <sub>lac</sub> , Kan <sup>R</sup>	(Platt <i>et al.</i> , 2000)

<sup>a</sup> Designates that a strain was constructed by transduction. Donor lysate strain in brackets, transduced into the recipient, strain at arrowhead.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

used to construct deletions of *SMc01619*, *SMc01618*, *SMc01622*, and *SMc01617* to yield strains SRmD208, SRmD209, SRmD247 and SRmD249 respectively (House *et al.*, 2004; Schroeder *et al.*, 2005). Deletions were confirmed by sequencing a PCR product that contained the deletion junctions.

To generate deletions of *pyc*, *fucA1*, and *eryB* in strains SRmD211, SRmD270 and SRmD271 respectively, a splice overlap extension strategy was utilized as previously described (Richardson *et al.*, 2008). PCR primers were designed to facilitate cloning of the regions flanking these genes into pJQ200SK to create pBG11, pBG28 and pBG30. These were recombined into the chromosome of *S. meliloti* as single crossovers by conjugating them into Rm1021 and selecting for Sm and Gm resistance. Double crossovers were selected for by plating single crossovers onto LB containing 8% sucrose. Resulting colonies that were sucrose resistant and Gm sensitive were screened for deletions using carbon source phenotypes and PCR. In all cases deletions that were created were verified by amplifying the DNA containing the deletion junctions by PCR and sequencing the product.

The plasmid pCO37 was used to construct plasmids expressing genes found at the erythritol locus suspected of having roles in adonitol and L-arabitol utilization. Briefly, pCO37 (Jacob *et al.*, 2008) is a derivative of the broad host range plasmid pRK7813 that contains *attB* sites; thus making it a Gateway<sup>®</sup> compatible destination vector. Genes of interest were recombined from the *S. meliloti* ORFeome entry vectors as previously described (Geddes *et al.*, 2010). In this manner, pBG23-26 and pBG31 were constructed (Table 3.1). Identity of the resultant plasmids was confirmed by a combination of sequencing the insert and the demonstration of bioactivity.

Random Tn5 mutagenesis was carried out using pRK602 essentially as previously described (Finan *et al.*, 1985). SRmD766 was isolated by screening for an inability to use xylose as a sole carbon source. The location of the Tn5 was identified by arbitrary PCR and sequencing as previously described (Miller-Williams *et al.*, 2006). In addition, the phenotype was shown to be 100% co-transducible with the Nm<sup>R</sup> encoded by the Tn5.

### ***3.2.3 DNA manipulations***

Standard techniques were used for plasmid isolation, restriction enzyme digests, ligations, transformations, and agarose gel electrophoresis (Sambrook *et al.*, 1989). Nucleotide sequencing was carried by cycle-sequencing using a Big-Dye version 3.1 kit. Sequencing reactions were carried out as recommended by the manufacturer and resolved using an ABI 3130 sequencer.

### ***3.2.4 RNA isolation and cDNA synthesis***

Bacterial cultures were grown to an OD<sub>600</sub> of approximately 0.8 in VMM medium containing either 15 mM adonitol or L-arabitol and 15 mM glycerol or 15 mM glycerol alone. Cells were harvested by centrifugation and resuspended in TE (Tris 10 mM, EDTA 1 mM, pH= 8) buffer with lysozyme (0.4 mg/ml). RNA was then isolated using the Qiagen RNeasy isolation kits as previously described (Geddes *et al.*, 2010). All RNA samples were treated twice on-column with Qiagen RNase-free DNase to remove DNA contamination. RNA was routinely checked for DNA contamination by PCR and was analyzed spectrophotometrically. The concentration of RNA was determined using the extinction coefficient for RNA (0.027 (mg/ml) cm<sup>-1</sup>). First strand cDNA synthesis was

performed using 1 µg of RNA as suggested by the supplier. The quality of the cDNA synthesis was checked by electrophoresis and quantified spectrophotometrically.

### **3.2.5 Quantitative RT-PCR**

Approximately 200 ng of cDNA sample was used as a template for quantitative RT-PCR. Reactions were performed using the SYBR green RT-PCR kit from Invitrogen as recommended. Primers utilized have been previously described (Geddes *et al.*, 2010). The RT-PCR reaction was performed using a Cepheid Smart Cycler with the following program: Stage 1, 95°C for 120s, 1 time; Stage 2, 95°C for 15s followed by 60°C for 30s, repeated 40 times; Stage 3, melting curve analysis of PCR products.

### **3.2.6 Kinase assays**

Cultures of *R. leguminosarum* Rlt100 expressing pBG3 *in trans* were grown and lysed to create extracts as previously described (Oresnik & Layzell, 1994; Pickering & Oresnik, 2008). Erythritol kinase assays were carried out by adapting a fructose kinase assay previously described (Anderson & Sapico, 1975). Essentially, either erythritol, adonitol, L-arabitol, or D-arabitol were substituted for fructose. The assay was initiated by the addition of the polyol that was being measured for kinase activity. Rates were determined from linear portions of NADH oxidation data that were corrected for background oxidase activity and were proportional to the amount of extract.

### 3.2.7 Transport assays

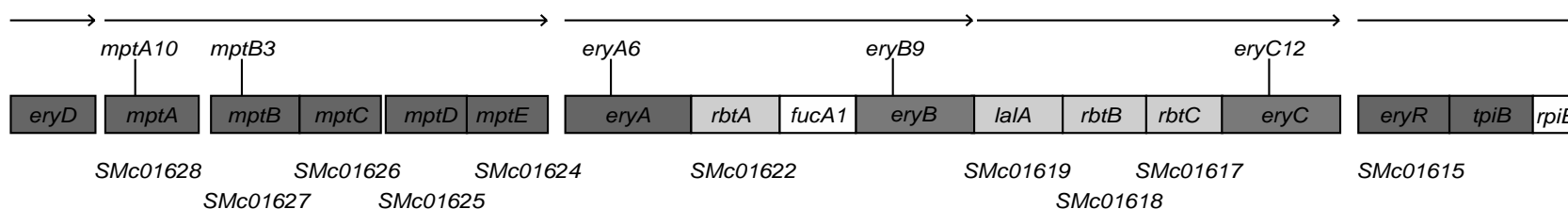
Transport assays were carried out essentially as previously described (Poysti *et al.*, 2007). Labelled [ $^{14}\text{C}$ -1] erythritol (370 MBq mmol $^{-1}$ ) was purchased from American Radiolabeled Chemicals Ltd. (St. Louis, Missouri). Transport assays were initiated by the addition of [ $^{14}\text{C}$ ] erythritol to a final concentration of 2  $\mu\text{M}$  to the samples; aliquots of 0.5 mL were withdrawn at appropriate time points and rapidly filtered through a Millipore 0.45  $\mu\text{m}$  Hv filter on a Millipore sampling manifold. The kinetics of uptake were linear for approximately 2 minutes (Geddes *et al.*, 2010), therefore samples were taken after 1 minute when competing solutes were added. Solute competition was carried out by adding 4  $\mu\text{M}$  or 10  $\mu\text{M}$  of unlabelled erythritol, adonitol or L-arabitol to samples along with 2  $\mu\text{M}$  labelled erythritol. The amount of radioactivity retained by the cells was quantified using a liquid scintillation counter (Beckman LS6500).

## 3.3 Results

### 3.3.1 Adonitol and L-arabitol compete with erythritol for transport

In *S. meliloti*, the genes necessary for the transport, catabolism, and regulation of erythritol are distributed over 5 transcripts (Geddes *et al.*, 2010) (Figure 3.1). In addition to containing the genes necessary for erythritol utilization, six other genes that are annotated as being involved in small molecule metabolism are found at this locus. One transcriptional unit contains five genes that encode components of an ABC transporter, including a sugar binding protein (*SMc01628*), 2 permeases (*SMc01626-27*) and 2 ABC proteins (*SMc01624-25*) (Figure 3.1).

**Figure 3.1.** Map of ORFs in the erythritol, adonitol and L-arabitol locus of *S. meliloti* Rm1021



Boxes represent ORFs. Vertical lines represent the location of Tn5-B20 insertions. Arrows above represent transcriptional units as defined in Geddes *et al.* 2010. Gene names are based on proposed annotation (See text). Systematic identifier numbers are included below the reannotated genes. Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

Strains carrying mutations in the genes encoding either the sugar binding protein (*SMc01628*) or the permease (*SMc01627*) were shown to be unable to utilize erythritol, adonitol or L-arabitol as sole carbon sources. Physical evidence was also presented demonstrating that this ABC transporter was required for intracellular accumulation of radiolabelled erythritol (Geddes *et al.*, 2010). Since these polyols can adopt identical stereochemistry over 3 carbons (Figure 3.2), it was reasoned that they may utilize a common transporter. To test this a transport competition assay was performed using unlabelled sugars to compete for transport with radiolabelled erythritol.

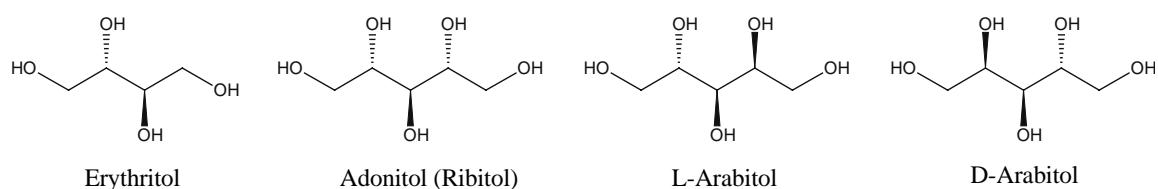
Unlabelled erythritol, adonitol and L-arabitol were competed against radiolabelled erythritol for transport at ratios of 2:1 and 5:1. The results show that the addition of unlabelled erythritol competing at a ratio of 2:1 reduced the accumulation of radiolabel in the cell after 1 minute. Increasing the ratio of unlabelled erythritol to 5:1 resulted in a further reduction of radiolabel accumulation (Figure 3.3). When the experiment was repeated using either adonitol or L-arabitol to compete with radiolabelled erythritol, a comparable reduction of radiolabelled erythritol in the cell was observed at both the 2:1 and 5:1 ratios. The ability of adonitol and L-arabitol to compete for transport with erythritol suggests that adonitol and L-arabitol share common determinants for transport with erythritol in *Sinorhizobium meliloti*. Since these genes are necessary for more than one polyol, we suggest that they should be reannotated as **multiple polyol transport**.

### ***3.3.2 EryB is not necessary for adonitol or L-arabitol catabolism***

The catabolism of erythritol to glyceraldehyde-3-phosphate is carried out by the actions of EryA, EryB, EryC, and TpiB in *S. meliloti* and *R. leguminosarum* bv. *viciae*

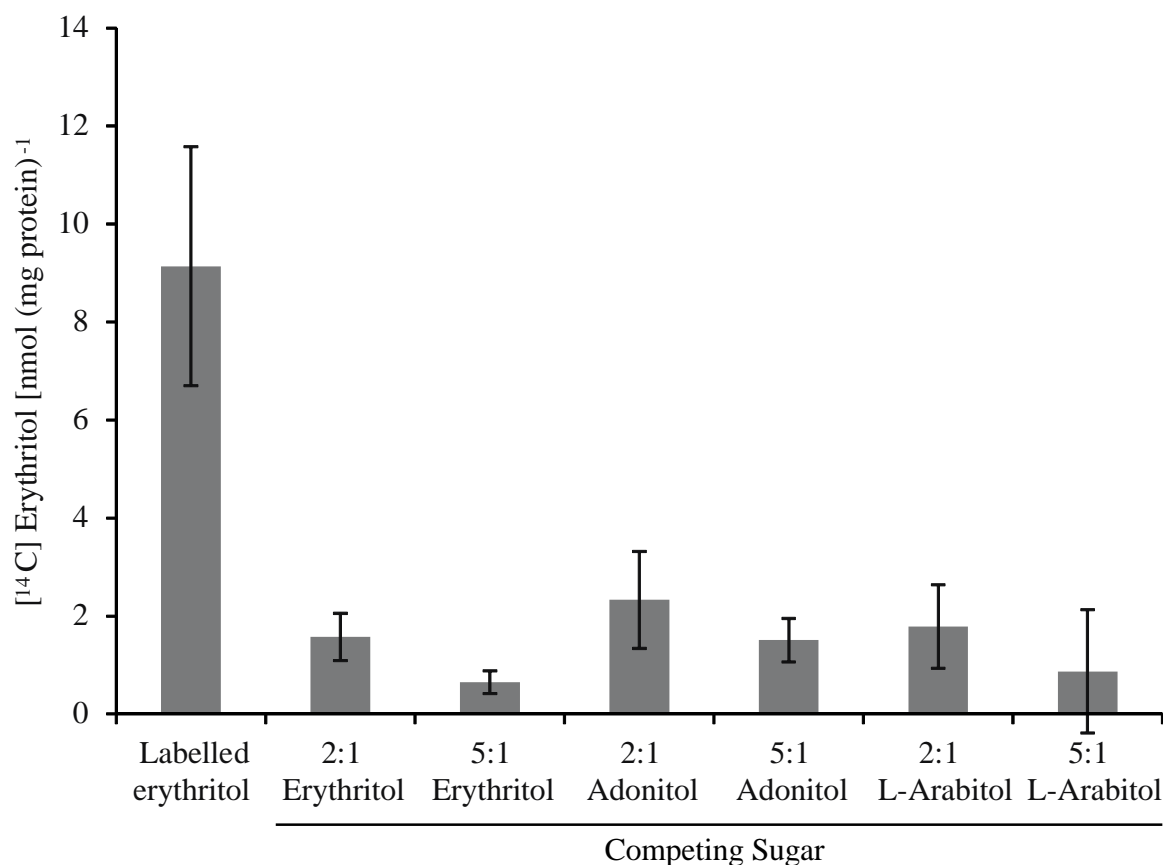


**Figure 3.2.** Chemical projections representing stereochemistry of erythritol, adonitol, L-arabitol and D-arabitol



Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

**Figure 3.3.** Competition of erythritol, adonitol and L-arabitol for transport with [ $^{14}\text{C}$ ] erythritol in *S. meliloti* Rm1021



2  $\mu\text{mol}$  of [ $^{14}\text{C}$ ] erythritol was competed against either 4  $\mu\text{mol}$  (2:1), or 10  $\mu\text{mol}$  (5:1) unlabelled erythritol, adonitol or L-arabitol. Accumulation of [ $^{14}\text{C}$ ] is expressed as nmol (mg protein) $^{-1}$  after one minute incubation. Data are expressed as the mean  $\pm$  standard deviation of three independent replicates. Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

(Geddes *et al.*, 2010; Poysti & Oresnik, 2007; Yost *et al.*, 2006). Since adonitol and L-arabitol share common determinants with erythritol transport (Figure 3.3), we were interested in determining whether any of the other genes necessary for erythritol catabolism were also necessary for the catabolism of these polyols.

A number of Tn5-B20 insertion mutants were previously generated at this locus and shown to be unable to utilize erythritol, adonitol or L-arabitol as sole carbon sources (Geddes *et al.*, 2010). Two of these mutants (SRmA784 and SRmA787), contained insertions at the end of transcriptional units, in *eryB* and *eryC* respectively (Figure 3.1). These strains were used to determine if the gene products of either *eryB* or *eryC* were necessary for adonitol and L-arabitol utilization.

It was found that SRmA787 (carries *eryC12::Tn5-B20*) was capable of utilizing adonitol and L-arabitol as sole carbon sources and unable to utilize erythritol (Table 3.2) (Geddes *et al.*, 2010), whereas SRmA784 (carries *eryB9::Tn5-B20*) was unable to utilize erythritol, adonitol, or L-arabitol (Geddes *et al.*, 2010). Introduction of pHY112 (*eryB*) into SRmA784 was able to complement the strain for the ability to utilize erythritol as a sole carbon source (Table 3.2) (Geddes *et al.*, 2010). However, when this construct was screened for complementation on adonitol and L-arabitol, it was found that it was unable to complement utilization of either pentitol as a sole carbon source (Table 3.2). It was hypothesized that the Tn5-B20 insert may be having polar effects on the transcriptional unit that contains *eryC*. To resolve this ambiguity we generated strain SRmD271, which contains an in-frame deletion that removed the DNA that encoded the first 485 of the 505 amino acids of EryB. SRmA271 was unable to grow using erythritol, but able to grow using adonitol or L-arabitol as sole carbon sources. It was also complemented for

**Table 3.2.** Growth of erythritol catabolic mutants on adonitol and L-arabitol

Strain	Genotype	Plasmid contains <sup>a</sup>	Carbon source						
			Glyc	Ery	Ery/glyc	Ado	Ado/glyc	L-Ara	L-Ara/glyc
Rm1021	Wild-type		+	+	+	+	+	+	+
SRmA782	<i>eryA6::Tn5-B20</i>	Vector	+	-	+	-	+	-	+
SRmA782 (pBG3)	<i>eryA6::Tn5-B20</i>	<i>eryA</i>	+	-	-	-	-	-	-
SRmA784	<i>eryB9::Tn5-B20</i>	Vector	+	-	-	-	-	-	-
SRmA784 (pHY112)	<i>eryB9::Tn5-B20</i>	<i>eryB</i>	+	+	+	-	-	-	-
SRmD271	$\Delta$ <i>eryB</i>	Vector	+	-	-	+	+	+	+
SRmD271 (pHY112)	$\Delta$ <i>eryB</i>	<i>eryB</i>	+	+	+	+	+	+	+
SRmA787	<i>eryC12::Tn5-B20</i>		+	-	-	+	+	+	+

Growth is indicated by +, for wild-type, or -, for no growth.

<sup>a</sup>Vector indicates strain carrying pRK7813. Genotypes correspond to genes expressed from pRK7813.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

erythritol utilization by pHY112 (Table 3.2). Therefore *eryB* is required for erythritol but not adonitol or L-arabitol utilization.

### 3.3.3 *EryA can use erythritol, adonitol, and L-arabitol as substrates.*

Introduction of the plasmid pBG3 (expresses *eryA*) into SRmA782 previously resulted in the inability to grow on medium containing both erythritol and glycerol (Geddes *et al.*, 2010). Based on other reported phenotypes (Adhya & Shapiro, 1969; Richardson *et al.*, 2004; Sperry & Robertson, 1975b; Yost *et al.*, 2006), it was reasoned that this may be the result of the build-up of a toxic phosphorylated intermediate because of the polar effects of the *eryA6* insertion on the downstream erythritol gene, *eryB* (Geddes *et al.*, 2010). In order to test whether EryA acts on adonitol and L-arabitol, we screened SRmA782 carrying pBG3 on defined medium containing either adonitol and glycerol, or L-arabitol and glycerol. Consistent with an ability to phosphorylate adonitol and L-arabitol we found that although SRmA782 was able to grow on defined medium when it contained only glycerol, it was unable to grow on media containing either adonitol and glycerol, or L-arabitol and glycerol when harbouring the plasmid pBG3 (Table 3.2).

To provide direct evidence for the involvement of EryA in the catabolism of L-arabitol and adonitol we performed erythritol, adonitol and arabitol kinase assays. To carry out these assays we chose *R. leguminosarum* bv. *trifolii* strain Rlt100 as a heterologous host to express *eryA* from pBG3 (Table 3.1). This *R. leguminosarum* host was chosen because it is unable to utilize either erythritol or L-arabitol as a sole carbon source. In addition a putative ribitol-2 dehydrogenase mutant in this background

(Rlt103), that is unable to use adonitol as a sole carbon source, was also available (Oresnik *et al.*, 1998).

Kinase activities were calculated from cell free extracts of Rlt100 carrying pBG3 or the empty vector pRK7813. Activities were calculated as the average of three independent replicates and corrected for background levels. Kinase activity of Rlt100 carrying pRK7813 did not deviate from background levels in the presence of erythritol, adonitol, L-arabitol or D-arabitol. When Rlt100 carrying pBG3 was assayed in the presence of erythritol, kinase activity was measured as  $17.0 \pm 1.4$   $\mu\text{mol}/\text{min}/\text{mg}$ . Assays using adonitol and L-arabitol as substrates resulted in similar activities of  $11.7 \pm 2.1$  and  $12.3 \pm 1.2$   $\mu\text{mol}/\text{min}/\text{mg}$  respectively, whereas using D-arabitol as a substrate resulted in a background level of activity of  $0.4 \pm 2.8$   $\mu\text{mol}/\text{min}/\text{mg}$ . Therefore, EryA of *S. meliloti* phosphorylates erythritol, adonitol and L-arabitol. The inability of EryA to phosphorylate D-arabitol suggests that the stereochemistry of the first three carbons is important for recognition by EryA (Figure 3.2).

#### ***3.3.4 Genes in the locus are uniquely necessary for adonitol and/or L-arabitol utilization***

The characterized pathways for adonitol and L-arabitol breakdown are initiated by oxidizing a hydroxyl on a C-2 carbon to yield a pentulose; D-ribulose and L-xylulose respectively. These ketopentuloses are subsequently phosphorylated, isomerized or epimerized to D-xylulose-5 phosphate and enter central metabolism (Mortlock *et al.*, 1965a). Since adonitol and L-arabitol can be phosphorylated by EryA directly, the

pathway for the degradation of these polyols in *S. meliloti* may be markedly different from what has been previously described.

In order to determine whether all the genes that are used during adonitol and L-arabitol catabolism in *S. meliloti* are contained within the erythritol locus, we conjugated the cosmid pEL6 into the *R. leguminosarum* strains Rlt100 and Rlt103 and screened for adonitol and L-arabitol utilization. The ability to utilize L-arabitol as a sole carbon source was conferred to Rlt100 and Rlt103 while maintaining pEL6 and the ability to utilize adonitol was restored to Rlt103 while maintaining the cosmid (Table 3.3).

Based on the ability of pEL6 to confer adonitol and L-arabitol utilization, we predict that the locus contains the complete set of adonitol and L-arabitol catabolic genes. To determine which genes are necessary for adonitol and L-arabitol catabolism, a series of in-frame deletions was constructed. Deletions were screened for carbon utilization and complemented by a plasmid-borne copy of each gene (Table 3.4).

It was found that strains SRmD247, SRmD209 and SRmD249, containing deletions of *SMc01622* (*rbtA*), *SMc01618* (*rbtB*) and *SMc01617* (*rbtC*) respectively, were unable to grow using L-arabitol and adonitol as a sole carbon source. This inability to utilize L-arabitol and adonitol was complemented *in trans* by the introduction of the corresponding gene on a plasmid (Table 3.4). SRmD208, containing a deletion of *SMc01619* (*lala*), was able to utilize adonitol as a sole carbon source but was unable to utilize L-arabitol. This phenotype was complemented by introducing *SMc01619* (*lala*) *in trans* (Table 3.4). To maintain consistency with other work describing adonitol catabolism, we propose that *SMc01622*, *SMc01618* and *SMc01617* be annotated as (**ribitol**) *rbtABC* respectively, and *SMc01619* as (**L-arabitol**) *lala*.

**Table 3.3.** Growth of *R. leguminosarum* with the *S. meliloti* erythritol cosmid on adonitol and L-arabitol

Strain	Genotype	Plasmid	Carbon source		
			Glycerol	Adonitol	L-Arabitol
Rlt100	Wild-type	pRK7813	+	+	-
Rlt100	Wild-type	pEL6	+	+	+
Rlt103	<i>rbt-2::Tn5-B20</i>	pRK7813	+	-	-
Rlt103	<i>rbt-2::Tn5-B20</i>	pEL6	+	+	+

Growth is as follows: +, wild-type, or -, no growth.



**Table 3.4.** Complementation analysis of strains containing deletions

Strain	Genotype <sup>a</sup>	Plasmid contains <sup>b</sup>	Carbon source			
			Erythritol	Adonitol	L-Arabitol	Glycerol
Rm1021	Wild-type		+	+	+	+
SRmD247	$\Delta$ <i>rbtA</i>	Vector	+	-	-	+
SRmD247 (pBG23)	$\Delta$ <i>rbtA</i>	<i>rbtA</i>	+	+	+	+
SRmD208	$\Delta$ <i>lalA</i>	Vector	+	-	+	+
SRmD208 (pBG25)	$\Delta$ <i>lalA</i>	<i>lalA</i>	+	+	+	+
SRmD209	$\Delta$ <i>rbtB</i>	Vector	+	-	-	+
SRmD209 (pBG26)	$\Delta$ <i>rbtB</i>	<i>rbtB</i>	+	+	+	+
SRmD249	$\Delta$ <i>rbtC</i>	Vector	+	-	-	+
SRmD249 (pBG31)	$\Delta$ <i>rbtC</i>	<i>rbtC</i>	+	+	+	+
SRmD270	$\Delta$ <i>fucA1</i>	Vector	-	+	+	+
SRmD270 (pBG24)	$\Delta$ <i>fucA1</i>	<i>fucA1</i>	-	+	+	+
SRmD270 (pHY112)	$\Delta$ <i>fucA1</i>	<i>eryB</i>	+	+	+	+

Growth is as follows: +, wild-type, or -, no growth.

<sup>a</sup>Correspondence of gene annotation and genomic systematic identifiers are as follows; *rbtA*, SMc01622; *rbtB*, SMc01618; *rbtC*, SMc01617; *lal*, SMc01619.

<sup>b</sup>Vector indicates strain carrying either pCO37 or pRK7813. Genotypes correspond to genes expressed from either pCO37 or pRK7813.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

A deletion of *fucAI* in SRmD270 was able to utilize adonitol and L-arabitol as sole carbon sources but unable to utilize erythritol. However, erythritol utilization was not complemented by the introduction of *fucAI* *in trans*. SRmD270 was complemented for erythritol utilization by *eryB* *in trans* suggesting that the deletion is affecting *eryB* transcription, which is required for erythritol catabolism (Table 3.4). This suggests that *fucAI* is not necessary for growth using erythritol, adonitol or L-arabitol.

### **3.3.5 xylB is necessary for D-arabitol but not L-arabitol or adonitol utilization**

The requirement of three and four unique catabolic genes for adonitol and L-arabitol utilization respectively is inconsistent with traditional pentitol catabolic pathways. We reasoned that determining whether *S. meliloti* uses either glycolysis or gluconeogenesis during catabolism of these substrates might provide insight into how they are catabolised. To determine this, we screened strains which were impaired in their glycolytic and/or gluconeogenic growth for the ability to utilize adonitol and L-arabitol as sole carbon sources.

SRmD211 is a pyruvate carboxylase (*pyc*) mutant and is incapable of growing on sugars which generate pyruvate during glycolytic growth in *S. meliloti*, such as glucose or glycerol (Table 3.5). Pyruvate carboxylase is an anapleurotic enzyme in *S. meliloti* and is required to resupply TCA cycle intermediates during glycolytic growth (Dunn *et al.*, 2001). SRmA515 is a *tpiA/tpiB* mutant. This strain is only capable of growth on sugars which produce both glyceraldehyde 3-phosphate and dihydroxyacetone phosphate (Poysti & Oresnik, 2007). Therefore, SRmA515 is incapable of growth on substrates such as succinate or arabinose that rely on gluconeogenesis to produce 6 carbon compounds.

**Table 3.5.** Carbon phenotypes of central catabolism mutants

Strain	Relevant genotype	Carbon source									
		Glycerol	Erythritol	Adonitol	L-Arabitol	D-Arabitol	Succinate	Arabinose	Glucose	Ribose	Xylose
Rm1021	Wild-type	+	+	+	+	+	+	+	+	+	+
SRmA766	<i>xylB</i>	+	+	+	+	-	+	+	+	+	-
SRmA515	$\Delta tpiA$ , <i>tpiB</i>	-	-	+	+	+	-	-	+	+	+
SRmA211	$\Delta pyc$	-	-	-	-	-	+	+	-	+	+
SRmA210	$\Delta pyc$ , $\Delta tpiA$ , <i>tpiB</i>	-	-	-	-	-	-	-	-	+	+

Growth is as follows: +, wild-type, or -, no growth.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

Similarly, this strain is unable to utilize substrates which enter central metabolism through DHAP such as glycerol or erythritol (Table 3.5).

Since D-xylulose is a common intermediate in traditional pentitol catabolism, we also utilized a strain containing a *xylB* mutation (SRmD766). XylB is necessary for conversion of D-xylulose to D-xylulose-5-phosphate. D-xylulose-5-phosphate is an entry point into the pentose phosphate pathway (Finan *et al.*, 1988).

These strains were screened on a number of pentoses and 3-, 4-, and 5- carbon polyols (Table 3.5). The results show that SRmA515 (*tpiA/tpiB*) was capable of growth on adonitol, L-arabitol, and D-arabitol suggesting that these compounds are not catabolised like erythritol which is converted to DHAP. SRmD211 (*pyc*) was incapable of growth on adonitol, L-arabitol, and D-arabitol (Table 3.5). Taken together these data suggest that *S. meliloti* grows glycolytically on all three pentitols. This is consistent with a traditional pentose phosphate pathway entry point for these pentitols. SRmA776 (*xylB*) was unable to grow using D-arabitol, but was able to grow using adonitol and L-arabitol as sole carbon sources (Table 3.5). This suggests that D-arabitol enters central metabolism in *S. meliloti* through a traditional route; oxidized to D-xylulose and subsequently phosphorylated by XylB. It also suggests that L-arabitol and adonitol are not converted to D-xylulose during their catabolism. Of note, SRmA210 (a *tpiA/tpiB* and *pyc* mutant) was able to utilize both xylose and ribose (Table 3.5).

### ***3.3.6 Bioinformatic characterization of genes necessary for adonitol and L-arabitol utilization***

Analyses of the protein coding sequences of the genes necessary for the adonitol and L-arabitol utilization were carried out using CDART (Geer *et al.*, 2002), InterPro Scan (Hunter *et al.*, 2009), as well as BLASTP (Altschul *et al.*, 1997) in an effort to determine putative biochemical function. The basic analysis suggests that, in addition to EryA, a dehydrogenase, a second kinase and a phosphatase are used during the catabolism of both adonitol and L-arabitol. A class II aldolase is also used during L-arabitol breakdown.

RbtA (SMc01622) belongs to IPR006140, and is predicted to be a 3-phosphoglycerate dehydrogenase (a 2 hydroxy-acid dehydrogenase). Since it is necessary for both adonitol and L-arabitol degradation, RbtA could follow EryA because the phosphorylation of adonitol or L-arabitol would produce a metabolite that is consistent with the needed substrate.

LalA (SMc01619) is predicted to be a member of the IPR001303 family of aldolases/adducins. IPR001303 is also related to IPR004661 which is represented by the enzyme L-ribulose-5-phosphate 4-epimerase. It has been shown that the epimerization reaction catalyzed by this domain is carried out via enolate intermediate (Luo *et al.*, 2001a). This suggests that LalA, which is necessary only for L-arabitol utilization, may be an epimerase.

RbtB (SMc01618) and RbtC (SMc01617), are predicted to be a carbohydrate kinase (IPR000577) of the FGGY family and a hydrolase/phosphatase belonging to the HAD II superfamily respectively. The categorization of RbtB in IPR000577 (FGGY

kinase family), and not in families such as IPR006082 (phosphoribulokinase) or IPR022463 (phosphofructokinase family) suggests that the substrate used by RbtB may be an unphosphorylated sugar.

### ***3.3.7 Adonitol and L-arabitol only induce transcripts necessary for their utilization***

Previous work has shown that each transcriptional unit in the locus was inducible by erythritol, and that this induction was dependant on the presence of both the regulator, EryD, and the erythritol kinase, EryA (Geddes *et al.*, 2010). SRmA885, a strain carrying an *eryD* mutation was unable to grow using erythritol, adonitol, and L-arabitol suggesting that *eryD* may be required for induction of the locus by adonitol and L-arabitol as well (Geddes *et al.*, 2010). Regulation of this locus is complex, and a second transcriptional regulator (SMc01615) is also induced by erythritol. SMc01615 affects the expression of its own transcript, which contains the erythritol catabolic gene *tpiB*, as well as the transcription of *eryD* (Geddes *et al.*, 2010). It was unclear whether adonitol or L-arabitol would affect induction of the locus in a similar or different manner from erythritol, so induction in response to the pentitols was measured by qRT-PCR.

Induction of each transcriptional unit that contains genes required for adonitol and L-arabitol transport and catabolism was observed (Table 3.6). Similar to what was previously observed for erythritol induction, induction of the locus in response to adonitol and L-arabitol was abolished in an *eryA6* background (Table 3.6). The transcriptional unit containing *SMc01615* (*eryR*, see below) was not induced by adonitol or L-arabitol (Table 3.6). SMc01615 was shown to regulate its own transcription (Geddes *et al.*, 2010), therefore the absence of induction by adonitol and L-arabitol suggests that it

**Table 3.6.** Induction by adonitol and L-arabitol

Strain	Relevant genotype	Carbon Source	<i>eryD</i>	<i>mptA</i>	<i>eryA</i>	<i>eryC</i>	<i>eryR</i>
Rm1021	Wild-type	Adonitol	4.7(0.7)	31.2(9.2)	11.8(1.5)	5.6(0.8)	1.0(0.2)
Rm1021	Wild-type	L-Arabitol	11.9(0.2)	67.4(14.3)	25.9(8.6)	18.1(2.9)	1.4(0.1)
SRmA782	<i>eryA6</i>	Adonitol	1.3(0.2)	4.2(3.3)	0.9 (1.0)	0.3(0.2)	0.3(0.1)
SRmA782	<i>eryA6</i>	L-Arabitol	2.2(0.8)	4.1(0.6)	0.4 (0.1)	0.4(0.1)	0.5(0.1)

Growth media contained either 15 mM adonitol or L-arabitol (for induction) and 15 mM glycerol (for growth). Data are expressed as  $2^{-\Delta\Delta C_t}$  and represent fold expression over uninduced Rm1021 grown in glycerol. The experiments include *SMc00128* as an internal control (Krol & Becker, 2004). The data presented are as the mean with standard deviation in brackets from three independent biological replicates. Correspondence of gene annotation and genomic systematic identifier numbers are as follows: *mptA*, *SMc01628*; *eryR*, *SMc01615*.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

does not respond to these pentitols. This is consistent with the hypothesis that SMc01615 is an erythritol specific regulator. We suggest that SMc01615 should be re-annotated as *eryR*.

### 3.4 Discussion

In this work we demonstrate that the polyols erythritol, adonitol, and L-arabitol are transported by an ABC transporter(s) that share at least one common determinant (Figure 3.1). In addition, we have shown that L-arabitol and adonitol are both phosphorylated by the erythritol kinase EryA. This may occur because these sugars have identical stereochemistry over 3 carbons (Figure 3.2). Quantitative RT-PCR experiments show that transcripts containing genes necessary for adonitol and L-arabitol utilization are induced by these sugars in an *eryA*-dependent manner.

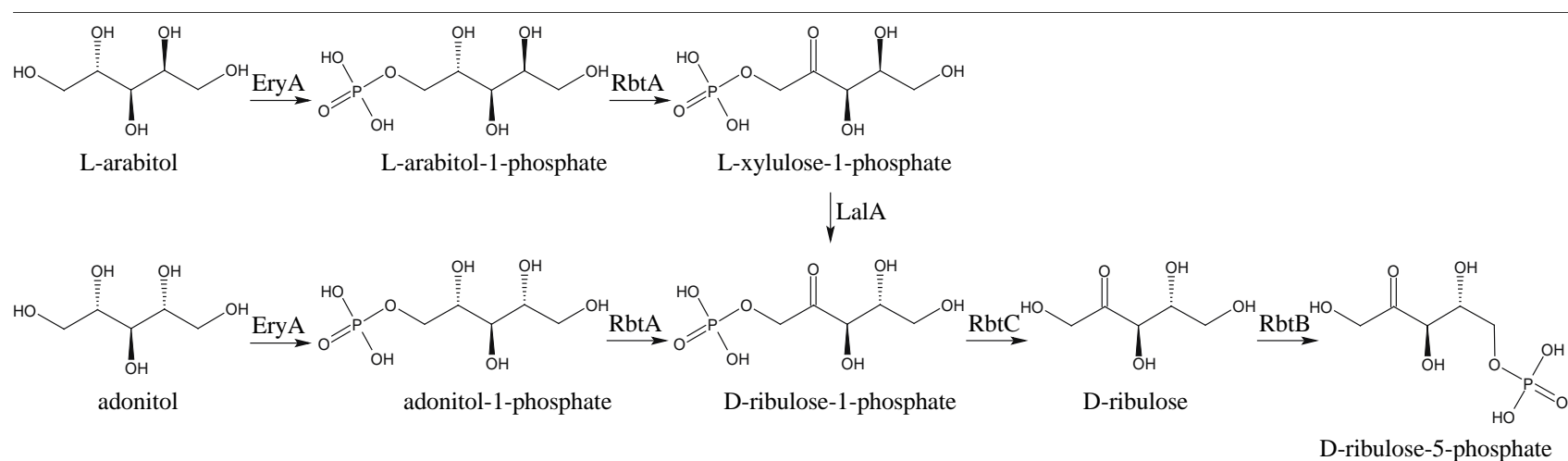
We had previously demonstrated that *eryC*, *tpiB* and *rpiB* are not required for adonitol and L-arabitol catabolism (Geddes *et al.*, 2010), and here we define that *fucAI*, and *eryB* are also not necessary for growth using adonitol and L-arabitol as sole carbon sources (Table 3.4). The ability of the cosmid containing the locus (pEL6) to confer the ability to grow using adonitol and L-arabitol as sole carbon sources suggests that all of the genes utilized during their breakdown are encoded by this locus. Our analysis has shown that three previously uncharacterised genes are necessary for growth using adonitol and L-arabitol as sole carbon sources (*rbtABC*), and one additional gene is required for L-arabitol catabolism (*lala*). This suggests that *eryA*, and *rbtABC* are the complete set of adonitol catabolic genes in *S. meliloti*, and that L-arabitol catabolism is facilitated by the same complement of genes, but also includes *lala*.



Pathways for the catabolism of polyols follow the pattern of a dehydrogenation to yield a 2 keto-sugar followed by a phosphorylation and, if necessary, an epimerization reaction to yield either a xylulose or ribulose-5-phosphate (Mortlock, 1984). Based on biochemical assays, phenotypic data (Table 3.4) and bioinformatic analysis, we hypothesise that adonitol and L-arabitol are not catabolised through such a pathway in *S. meliloti*. Despite differences we observed, phenotypes of mutations in central metabolism are consistent with adonitol and L-arabitol entering central metabolism through a 5 carbon pentose phosphate intermediate (Table 3.5), and not as a TCA cycle intermediate or as a product of an aldolase reaction as has been shown for other pentoses (Dahms & Anderson, 1969; Pedrosa & Zancan, 1974; Poysti & Oresnik, 2007; Poysti *et al.*, 2007; Richardson *et al.*, 2004).

Synthesis of the results from the bioinformatic analysis with the genetic data allows the construction of a putative pathway for the catabolism of adonitol and L-arabitol in *S. meliloti* (Figure 3.4). The logic for the pathway is as follows: the primary reaction is the phosphorylation of adonitol or L-arabitol by EryA. LalA is the only enzyme that is required uniquely for L-arabitol utilization. LalA is predicted to be a ribulose-1-phosphate 4-epimerase, therefore it is possible that an intermediate of L-arabitol utilization could be epimerized using LalA into an intermediate of the adonitol utilization pathway, thus explaining the use of all the remaining enzymes for the catabolism of both pentitols. This epimerization reaction is dependent upon a substrate that has a keto-group, therefore the L-arabitol derivative must be preceded by a dehydrogenase; RbtA is the only candidate gene at this locus that encodes a dehydrogenase. Since RbtA is necessary for the utilization of both pentitols, and its

**Figure 3.4.** Predicted pathway of adonitol and L-arabitol catabolism in *S. meliloti* Rm1021



The pathway is based on presented evidence (Table 3), as well as bioinformatic analysis of the genes required for adonitol and L-arabitol catabolism. See text for details.

Reproduced from Microbiology, volume 158, Geddes, B. A., and I. J. Oresnik, Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*, pp. 2180-2191, Copyright 2012, with permission from the Society for General Microbiology.

reaction would occur before the epimerization of an L-arabitol intermediate, we predict that this enzyme can function using both adonitol 1-phosphate and L-arabitol 1-phosphate as substrates; oxidation at C-2 would be most consistent with the predicted substrates of downstream genes. RbtB appears to be a pentulose kinase and not a phosphokinase. This suggests that its substrate is not a phospho-sugar. Since a phosphorylated intermediate has already been generated by EryA, RbtB must be preceded by a phosphatase. A candidate phosphatase is RbtC. The final kinase reaction would generate D-ribulose-5-phosphate, a logical entry point into central metabolism. Taken together the pathway as presented is in agreement with all the biological and bioinformatic evidence to this point.

The analysis of the locus has accounted for all the genes except *fucA1*. In *E. coli*, FucA is involved in the catabolism of the methyl-pentose fucose where it catalyzes the aldol cleavage of fucose 1-phosphate to lactaldehyde and dihydroxyacetone phosphate (Ghalambor & Heath, 1962). The breakdown of D-arabinose in *E. coli* has been shown to be dependent on *fucA*. In this context, FucA can recognize the intermediate D-ribulose 1-phosphate that is generated during the catabolism to generate glycolaldehyde and dihydroxyacetone phosphate (LeBlanc & Mortlock, 1971). The FucA1 found in this locus shares 48% identity with FucA of *E. coli*. The putative pathway predicts D-ribulose 1-phosphate as an intermediate in the catabolic pathway of adonitol and L-arabitol. However, the inability of mutations in *rbtB* and *rbtC* to grow using adonitol and L-arabitol as sole carbon sources suggests that this FucA1 activity is either not sufficient to sustain growth, or *S. meliloti* is unable to utilize the generated glycolaldehyde.

A screen to determine if any of the polyols were dependent on xylulose kinase for catabolism identified *xylB* as being necessary for D-arabitol catabolism (Table 3.5).

Biochemical analysis of *S. meliloti* had previously identified enzymatic activities of two different polyol dehydrogenases (Martinez de Drets & Arias, 1970). One of these had sorbitol dehydrogenase activity, while the other was able to oxidise mannitol and D-arabitol. Both of these dehydrogenases were shown to be induced by any of the three polyols (Martinez de Drets & Arias, 1970). More recently, a genome wide study that systematically mutated 78 genes that were in the short chain dehydrogenase family and identified *SMc01500* (*SmoC*) as involved in polyol metabolism (Jacob *et al.*, 2008). Subsequent testing of this strain showed that it is able to grow using erythritol, adonitol and L-arabitol as sole carbon sources, but unable to grow using D-arabitol, mannitol or sorbitol as sole carbon sources. The locus surrounding *smoS* contains many genes annotated as being involved in polyol catabolism, notably a putative mannitol dehydrogenase *mtlK* immediately downstream of *smoS*. Given the growth phenotypes of *SMc01500*, we hypothesise that the locus containing this mutation encodes the two dehydrogenases that were previously observed, and based on annotation, we predict that *smoS* encodes the sorbitol dehydrogenase, and *mtlK* (*SMc01501*) is necessary for oxidizing D-arabitol and mannitol. This suggests D-arabitol is broken down through a traditional catabolic pathway in *S. meliloti* using MtlK to form xylulose and XylB to phosphorylate it to xylulose-5-phosphate. This also corroborates the finding that the product of D-arabitol dehydrogenase was D-xylulose (Martinez de Drets & Arias, 1970).

*S. meliloti* has become a model organism for both the study of plant-microbe interactions as well as for transport and metabolism. Because of the large number of its genes that encode transporters and enzymes involved in metabolism, it provides an excellent platform for the discovery of novel biochemical processes. The delineation of

this complex locus has provided genetic and bioinformatic evidence that is consistent with the existence of a new pathway for the catabolism of adonitol and L-arabitol. We note that the genetic organization of these genes appears to be unique when compared to orthologues in other sequenced organisms. It is likely that a detailed study to elucidate the relationship of this locus to orthologues in other sequenced organisms, as well as work to provide direct biochemical evidence for each step of this pathway may provide basic knowledge that could be invaluable for understanding polyol metabolism in other diverse species that are sequenced.

## **Chapter 4**

### **Phylogenetic Analysis of Erythritol Catabolic Loci within the *Rhizobiales* and Proteobacteria**

Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria

## 4.1 Introduction

Operons are multigene arrangements transcribed as a single mRNA and are one of the defining features found in bacterial and archaeal genomes. This arrangement allows genes to be co-regulated, and members of operons are usually involved in the same functional pathway (Omata *et al.*, 1999; Osbourn & Field, 2009). Although operons are prominent features in the genomes of bacteria and archaea, the evolution and mechanisms that promote operon formation are still not resolved and a number mechanisms have been proposed (Fani *et al.*, 2005; Fondi *et al.*, 2009; Homma *et al.*, 2007; Muzzi *et al.*, 2008; Omelchenko *et al.*, 2003; Price *et al.*, 2006). These mechanisms involve dynamic genetic events that include gene transfer events, deletions, duplications, and recombinations (Muzzi *et al.*, 2008; Osbourn & Field, 2009; Price *et al.*, 2006). Since operons are prominent features in bacterial genomes, and often encode genes with metabolic potential, it may be assumed that their evolution is under some selection pressure, thus allowing prokaryotic cells to rapidly adapt, compete and grow under changing environmental conditions.

The metabolic capability of an organism can be a function of its genome size and gene complement and these greatly affect its ability to live in diverse environments. The alpha subdivision of the proteobacteria includes some organisms that are very similar phylogenetically but inhabit many diverse ecological niches, including a number of bacteria that can interact with eukaryotic hosts (Kuykendall *et al.*, 2012). The genome sizes of these organisms varies from about 1 MB for members of the genus *Rickettsia* to approximately 9 MB for members of the bradyrhizobia (Batut *et al.*, 2004). Comparative genomic studies of this group has led to the supposition that there has been two

independent reductions in genomic size, one which gave rise to the *Brucella* and *Bartonella*, the other which gave rise to the *Rickettsia* (Boussau *et al.*, 2004). In addition, it also suggests that there has been a major genomic expansion and that roughly correlates with the soil microbes within the order Rhizobiales (Boussau *et al.*, 2004). The genomes of Rhizobia are dynamic. Phylogenetic analysis of 26 different *Sinorhizobium* and *Bradyrhizobium* genomes recently showed that recombination has dominated the evolution of the core genome in these organisms, and that vertically transmitted genes were rare compared with genes with a history of recombination and lateral gene transfer (Tian *et al.*, 2012). In this manuscript we have utilized comparative genomics in a focused manner to investigate the evolution of genes and loci involved in the catabolism of the sugar alcohols erythritol, adonitol and L-arabitol, primarily within the alpha-proteobacteria.

The number of bacterial species that are capable of utilizing the common 4 carbon polyol, erythritol, as a carbon source is restricted (Wawskiewicz & Barker, 1968). Catabolism of erythritol has been shown to be important for competition for nodule occupancy in *Rhizobium leguminosarum* as well as for virulence in the animal pathogen *Brucella suis* (Burkhardt *et al.*, 2005). Genetic characterization of erythritol catabolic loci has only been performed in *R. leguminosarum*, *B. abortus* and *Sinorhizobium meliloti*. In these organisms erythritol is broken down to dihydroxyacetone-phosphate using the core erythritol catabolic genes *eryABC-tpiB* (Geddes & Oresnik, 2012a). During characterization of the erythritol locus of *S. meliloti*, it was observed that despite the close homologies of core erythritol genes, the genetic content and arrangement of the locus was drastically different from the previously characterized loci of *B. abortus* and *R.*



*leguminosarum* (Geddes *et al.*, 2010). In particular the locus encodes the catabolism of two 5-carbon pentitols (adonitol and L-arabitol) in addition to erythritol. It was shown that the ABC transporter encoded by *mptABCDE* and erythritol kinase encoded by *eryA* can also be used for adonitol and L-arabitol, and several genes in the locus are involved in adonitol and L-arabitol, but not erythritol catabolism including *lala-rbtABC* (Geddes & Oresnik, 2012a).

The differences between the erythritol loci in the sequenced *S. meliloti* strain Rm1021 (Galibert *et al.*, 2001), and *R. leguminosarum*, led us to question what the relationship of these erythritol catabolic loci may be to other putative erythritol catabolic loci in bacterial species. In this work we focus on this question by analyzing the content and synteny of loci containing homologs to the erythritol genes in other sequenced organisms. The results of the analysis lend support to several hypotheses regarding operon evolution, and in addition, the data predicts loci that may be involved in polyol transport and metabolism in other proteobacteria.

## 4.2 Materials and Methods

### 4.2.1 Identification of erythritol loci

The data set of erythritol loci utilized in this work was constructed in a two-step process. First BLASTN was used to identify sequenced genomes containing homologs to the core erythritol catabolic genes of *R. leguminosarum* and *S. meliloti* (Altschul *et al.*, 1997). The use of BLASTN rather than BLASTP at this stage allowed us to refine the search to bacteria with sequenced genomes. Furthermore, limiting the search to genes with highly similar sequences by using BLASTN allowed us to limit our search to only

genes that are likely involved in erythritol catabolism, since all of these genes encode proteins in highly ubiquitous families found throughout bacterial genomes. Initially BLASTN searches were performed using all the core erythritol genes shared between *R. leguminosarum* and *S. meliloti* (*eryA*, *eryB*, *eryC* and *eryD*). However, the search using *eryA* provided the most diverse data set that also showed a sharp drop in E-value and query coverage. Using either *eryA* from *R. leguminosarum*, or *eryA* from *S. meliloti* for the BLASTN search resulted in an identical data set. Genomes containing homologs to *eryA* were selected on the basis of E-values less than 1.00E-5. In cases where multiple strains of the same bacterial species were found to have highly homologous putative erythritol genes (>99% identity) only a single representative of the species was used to avoid redundancy. Additionally *B. melitensis* 16M and *B. suis* 1330 were chosen as representatives of the *Brucella* lineage despite a large number of *Brucella* species that were identified in our search due to the high degrees of similarity between their erythritol catabolic genes.

Second, the genetic region containing *eryA* in these organisms was identified and analyzed using the IMG Ortholog Neighborhood Viewer (<http://img.jgi.doe.gov>) (Markowitz *et al.*, 2010) in order to construct the gene maps (loci). The amino acid sequence of EryA from *S. meliloti* was used as a query for the IMG Ortholog Neighborhood Viewer search.

To analyze the genetic content of organisms in our data set, the amino acid sequence encoded by each gene involved in erythritol catabolism in *R. leguminosarum*, or in erythritol, adonitol or L-arabitol catabolism in *S. meliloti*, was individually used in a BLASTP search of the 19 genomes in the data set. The sugar binding proteins of the *S.*

*meliloti* and *R. leguminosarum* transporter were used as representatives of the entire ABC transporter. Identity cut-off values that were used to delineate potential homologs to erythritol proteins were unique to each query amino acid sequence. Cut-off values were as follows: MptA: 56%, EryD: 44%, EryA: 46%, RbtA: 50%, EryB: 65%, LalA: 49%, RbtB: 51%, RbtC: 40%, EryC: 68%, TpiB: 69%, EryR: 61%, EryG: 73%. These values were manually determined and generally correlated to a large drop in percentage identity within the BLASTP hits.

Homologs identified that were not within the primary *eryA* containing loci were used as a query within IMG-Ortholog neighborhood viewer to analyze the region surrounding them. Secondary loci containing homologs to some of these genes were identified in *Mesorhizobium* sp. and *Sinorhizobium fredii*. These loci are putative erythritol loci based on homology to known loci involved in erythritol catabolism in *Sinorhizobium meliloti* (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010), *R. leguminosarum* (Yost *et al.*, 2006) and *Brucella abortus* (Sangari *et al.*, 2000). Despite not having been experimentally verified we will refer to all loci in our data set as erythritol loci.

#### **4.2.2 Phylogenetic analysis**

Amino acid sequences of homologs to proteins previously shown to play a role in erythritol, adonitol or L-arabitol catabolism from each of the organisms in the data set were collected and used for phylogenetic analysis. The 16S *rDNA* and RpoD sequences were also extracted from the NCBI database for species examined in this study in order to obtain a potential species tree that could be compared with the various phylogenetic gene

trees obtained from the individual genes located within the polyol (i.e. erythritol, arabitol, and adonitol) utilization loci. Amino acid sequences were aligned using Clustal-X (Thompson *et al.*, 1997) and PRALINE (Simossis *et al.*, 2005) the resulting alignments were refined manually with the GeneDoc program v2.5.010 (Nicholas *et al.*, 1997).

Phylogenies were generated with maximum likelihood analysis (ML) as implemented in the Molecular Evolutionary Genetic Analysis package (MEGA5) (Tamura *et al.*, 2011) and with MrBayes (Ronquist & Huelsenbeck, 2003). MEGA5 was used to identify the most suitable substitution models for the aligned data sets. In order to evaluate support for the nodes observed in the ML phylogenetic trees bootstrap analysis (Felsenstein, 1985) was conducted by analysing 1000 pseudo replicates.

The MrBayes program (v3.1) was used for Bayesian analysis (Ronquist, 2004; Ronquist & Huelsenbeck, 2003) and the parameters set for amino acid alignments were mixed models and for the 16S rDNA gamma distribution with 4 rate categories. The models used (setting mixed model) for generating the final 50% majority rule trees were estimated by the program itself. The Bayesian inference of phylogenies was initiated from a random starting tree and four chains were run simultaneously for 1,000,000 generations; trees were sampled every 100 generations. The first 25 % of trees generated were discarded ("burn-in") and the remaining trees were used to compute the posterior probability values.

Phylogenetic trees were constructed for RpoD, 16S rDNA and all the key genes associated with the EryA genes. Phylogenetic trees were plotted with the TreeView program (Page, 1996) using MEGA5 and/or MrBayes tree outfiles. Final trees were annotated using Adobe Illustrator.

## 4.3 Results

### 4.3.1 Phylogenetic distribution of putative erythritol loci

Based on homology to *eryA* from *S. meliloti* and *R. leguminosarum* we have compiled a data set of 19 different putative erythritol loci from 19 different proteobacteria (Table 4.1). Previous studies suggested that erythritol loci may be restricted to the alpha-proteobacteria (Yost *et al.*, 2006). While a majority of the erythritol loci we identified followed this scheme, surprisingly we identified putative erythritol catabolic loci in *Verminephrobacter eiseniae* (a beta-proteobacterium) and *Escherichia fergusonii* (a gamma-proteobacterium). Erythritol loci are not widely distributed through the alpha-proteobacteria. A majority of the loci we identified were within the order Rhizobiales. Outside of the Rhizobiales we also identified erythritol loci in *Acidiphilium* species and *Roseobacter* species. Within the Rhizobiales, erythritol loci were notably absent from a large number of bacterial species such as *Rhizobium etli*, *Agrobacterium tumefaciens* and *Bradyrhizobium japonicum* that are closely related to other species that we have identified that contain erythritol loci. We also note that erythritol loci appear to plasmid localized only in *S. fredii* and *R. leguminosarum*. In all other cases the loci appear to be found on chromosomes.

### 4.3.2 Genetic content of loci

The genetic content of each of the organisms *ery* loci were analyzed by conducting a BLASTP search to the 19 genomes in our data set of the amino acid sequence of each gene associated with erythritol catabolism in *R. leguminosarum*, or

**Table 4.1.** Bacterial genomes used in this study containing erythritol loci

Genome	Accession number	Reference/ affiliation
<i>Sinorhizobium meliloti</i> 1021	AL591688.1	(Galibert <i>et al.</i> , 2001)
<i>Sinorhizobium medicae</i> WSM419	CP000738.1	(Reeve <i>et al.</i> , 2010b)
<i>Sinorhizobium fredii</i> NGR234	CP000874.1	(Schmeisser <i>et al.</i> , 2009)
<i>Mesorhizobium opportunism</i> WSM2075	CP002279.1	US DOE Joint Genome Institute
<i>Mesorhizobium loti</i> MAFF303099	BA000012.4	(Kaneko <i>et al.</i> , 2000)
<i>Mesorhizobium ciceri</i> bv. <i>biserrulae</i> WSM1271	CP002447.1	US DOE Joint Genome Institute
<i>Bradyrhizobium</i> sp. BTAi1	CP000494.1	(Giraud <i>et al.</i> , 2007)
<i>Bradyrhizobium</i> sp. ORS278	CU234118.1	(Giraud <i>et al.</i> , 2007)
<i>Agrobacterium radiobacter</i> K84	CP000629.1	(Slater <i>et al.</i> , 2009)
<i>Ochrobactrum anthropi</i> ATCC 49188	CP000759.1	(Chain <i>et al.</i> , 2011)
<i>Brucella suis</i> 1330	CP002998.1	(Tae <i>et al.</i> , 2011)
<i>Brucella melitensis</i> 16M	AE008918.1	(DelVecchio <i>et al.</i> , 2002)
<i>Acidiphilium multivorum</i> AIU301	AP012035.1	NITE Bioresource Information Center
<i>Acidiphilium cryptum</i> JF-5	CP000697.1	US DOE Joint Genome Institute
<i>Roseobacter denitrificans</i> Och114	CP000362.1	(Swingley <i>et al.</i> , 2007)
<i>Roseobacter litoralis</i> Och149	CP002623.1	(Kalhoefer <i>et al.</i> , 2011)
<i>Rhizobium leguminosarum</i> bv. <i>viciae</i> 3814	AM236086.1	(Young <i>et al.</i> , 2006)
<i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM1325	CP001623.1	(Reeve <i>et al.</i> , 2010a)
<i>Verminephrobacter eiseniae</i> EF01-2	CP000542.1	US DOE Joint Genome Institute
<i>Escherichia fergusonii</i> ATCC 35469	CU928158.2	Genoscope - Centre National de Sequencage

Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and *Proteobacteria*.

erythritol, adonitol or L-arabitol catabolism in *S. meliloti*. The results of the BLAST search are presented in Table 4.1, depicting the presence or absence of homologs to erythritol, adonitol or L-arabitol catabolic genes in each of the genomes that was investigated. Gene maps of erythritol loci were constructed based on the output of our IMG Ortholog Neighborhood Viewer searches and are depicted in Figure 4.1.

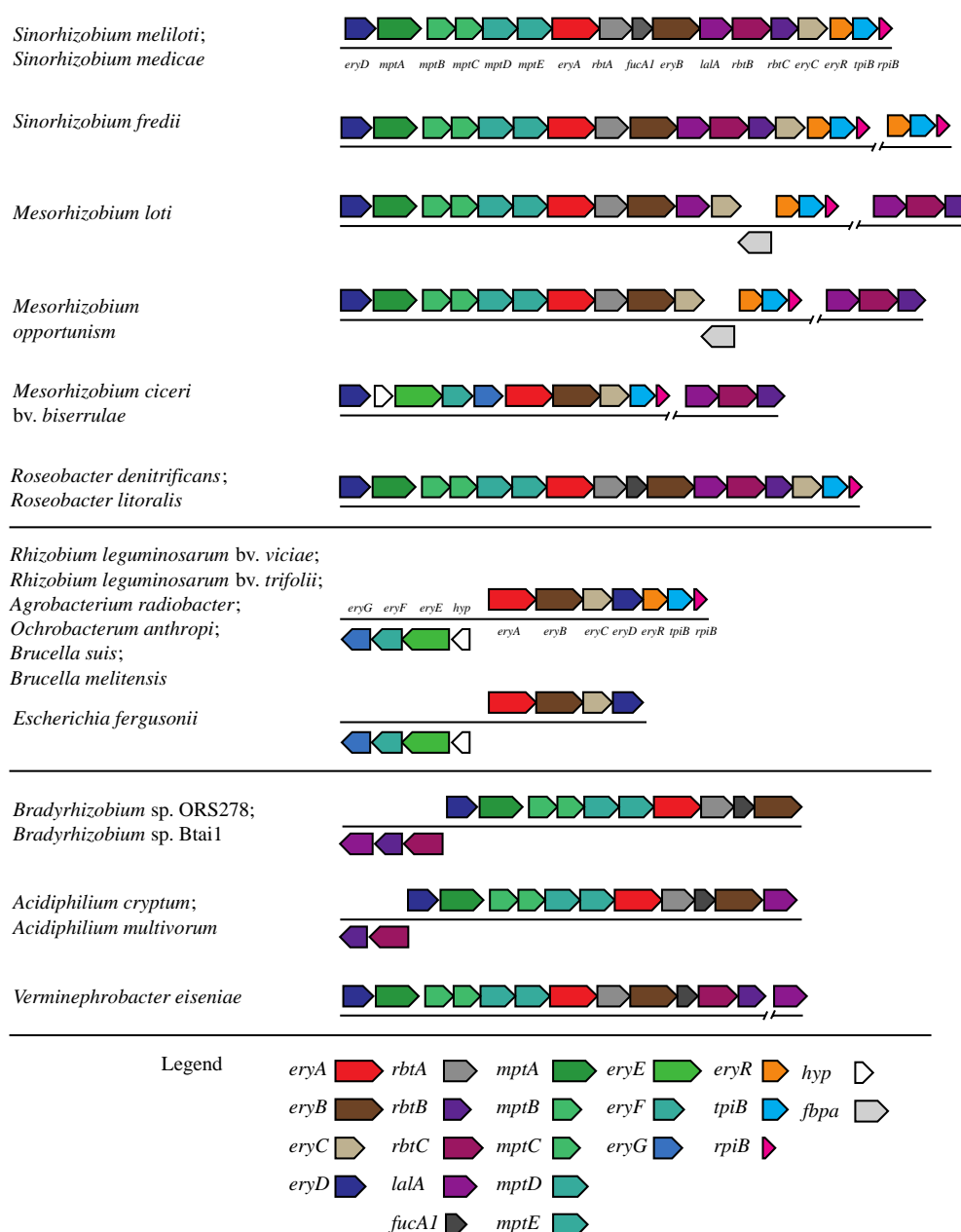
Genes encoding homologs to the core erythritol proteins EryA, EryB and EryD were ubiquitous throughout our data set (Table 4.1). With respect to the remaining genes, the genetic content of the species can be grouped into three broad categories. (1) Species that contain genes encoding homologs associated with erythritol, adonitol and L-arabitol catabolism. This includes *S. meliloti*, *S. medicae*, *S. fredii*, *M. loti*, *M. opportunism*, *M. ciceri*, *R. denitrificans* and *R. litoralis*. These genomes contained homologs to genes that encode enzymes specifically involved erythritol catabolism such as EryC, and TpiB as well as specifically involved in adonitol and L-arabitol catabolism including LalA, and RbtBC. They also contain genes encoding an ABC transporter homologous to the *S. meliloti* erythritol, adonitol and L-arabitol transporter (MptABCDE) and do not encode homologs to the *R. leguminosarum* erythritol transporter (EryEFG). One notable exception is *M. ciceri* which encodes EryEFG homologs rather than MptABCDE (Table 4.1). (2) Species that contain all the genes associated with erythritol catabolism, but lack the genes associated with adonitol or L-arabitol catabolism. These species include *R. leguminosarum* bvs. *viciae* and *trifolii*, *A. radiobacter*, *O. anthropi*, *B. suis*, *B. melitensis*, and *E. fergusonii*. These loci encode EryABC<sub>DR</sub>-TpiB as well as homologs to the *R. leguminosarum* ABC transporter EryEFG, but lack genes encoding

**Table 4.2.** Content of putative erythritol loci

Genome	Homologs involved in erythritol, adonitol and/or L-arabitol catabolism											
	EryA	EryB	EryD	EryC	EryG	EryR	TpiB	MptA	LalA	RbtA	RbtB	RbtC
<i>Sinorhizobium meliloti</i>	+	+	+	+	-	+	+	+	+	+	+	+
<i>Sinorhizobium medicae</i>	+	+	+	+	-	+	+	+	+	+	+	+
<i>Sinorhizobium fredii</i>	+	+	+	+	-	++	++	+	+	+	+	+
<i>Mesorhizobium opportunism</i>	+	+	+	+	-	+	+	+	+	+	+	+
<i>Mesorhizobium loti</i>	+	+	+	+	-	+	+	+	++	+	+	+
<i>Mesorhizobium ciceri</i> bv. <i>biserrulae</i>	+	+	+	+	+	-	+	-	+	-	+	+
<i>Roseobacter denitrificans</i>	+	+	+	+	-	-	+	+	+	+	+	+
<i>Roseobacter litoralis</i>	+	+	+	+	-	-	+	+	+	+	+	+
<i>Rhizobium leguminosarum</i> bv. <i>viciae</i>	+	+	+	+	+	+	+	-	-	-	-	-
<i>Rhizobium leguminosarum</i> bv. <i>trifolii</i>	+	+	+	+	+	+	+	-	-	-	-	-
<i>Agrobacterium radiobacter</i>	+	+	+	+	+	+	+	-	-	-	-	-
<i>Ochrobacterum anthropi</i>	+	+	+	+	+	+	+	-	-	-	-	-
<i>Brucella suis</i> 1330	+	+	+	+	+	+	+	-	-	-	-	-
<i>Brucella melitensis</i> 16M	+	+	+	+	+	+	+	-	-	-	-	-
<i>Escherichia fergusonii</i>	+	+	+	+	+	-	-	-	-	-	-	-
<i>Bradyrhizobium</i> sp. BTAi1	+	+	+	-	-	-	-	+	+	+	+	+
<i>Bradyrhizobium</i> sp. ORS278	+	+	+	-	-	-	-	+	+	+	+	+
<i>Acidiphilium multivorum</i>	+	+	+	-	-	-	-	+	+	+	+	+
<i>Acidiphilium cryptum</i>	+	+	+	-	-	-	-	+	+	+	+	+
<i>Verminephrobacter eiseniae</i>	+	+	+	-	-	-	-	+	+	+	+	+

+ indicates presence of homolog in the genome, - indicates absence of homolog in the genome, ++ indicates presence of 2 homologs in genome. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.



**Figure 4.1.** The genetic arrangement of putative erythritol loci in the proteobacteria

Genes are represented by colored boxes and identical colors identify genes that are believed to be homologous. Gene names are given below the boxes for *Sinorhizobium meliloti* and *Rhizobium leguminosarum*. Loci arrangements are depicted based on the output from the IMG Ortholog Neighborhood Viewer primarily using the amino acid sequence EryA from *Sinorhizobium meliloti*, and *Rhizobium leguminosarum*. Gene names in the legend generally correspond to the annotations in *R. leguminosarum* and *S. meliloti*. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

homologs to enzymes associated specifically with adonitol and L-arabitol catabolism or the *S. meliloti* transport protein MptABCDE. *E. fergusonii* contains the most minimal set of homologs to erythritol genes of all the genomes investigated, and did not encode EryR and TpiB. (3) Species that do not encode the specifically erythritol-associated EryC, EryR, and TpiB, but encode the adonitol/L-arabitol catabolic complement LalA-RbtABC and homologs to the *S. meliloti* polyol transporter MptABCDE including *Bradyrhizobium* spp. BTAi1 and ORS278, *A. multivorum*, *A. cryptum* and *V. eiseniae*.

#### 4.3.3 The genetic structure of erythritol loci

The genetic context of *eryA* in each of the genomes in our data set supported that each of these organisms contained an erythritol locus. A physical map of the loci in each of these organisms is depicted in Figure 4.1. Of note, a number of putative erythritol loci were identified in organisms with incomplete genome sequences at the time of analysis, and thus are not discussed here, including: *Octadecabacter antarcticus*, *Pelagibaca bermudensis*, *Enterobacter hormaechei*, *Fulvimarina pelagi*, *Aurantimonas* sp. SI85-9A1, *Roseibium* sp. TrichSKD4, *Burkholderia thailandensis* and *Stappia aggregata*.

The putative erythritol loci of bacteria in our data set ranged in genetic complexity with the loci from *S. meliloti* and *S. medicae* containing 17 different genes, to the simplest being the locus of *E. fergusonii*, which contained only two divergently transcribed operons that are homologous to the *eryEFG* and *eryABCD* loci of *R. leguminosarum*. A number of species contained loci that were identical in content and arrangement to the *R. leguminosarum* erythritol locus including members of the *Brucella*, *Ochrobacterum*, and *Agrobacterium*. The only species that contains a locus identical in

content and arrangement to *S. meliloti* is the closely related *Sinorhizobium medicae*. The locus of *Sinorhizobium fredii* NGR234, contains all but one of the genes (*fucA1*) found in the other *Sinorhizobium* loci (Table 4.2).

The loci of *Mesorhizobium* species were varied, however all three *Mesorhizobium* sp. contained an independent locus with homologs to *lala* and *rhtBC* elsewhere in the genome (Figure 4.1). Interestingly, while *M. loti* and *M. opportunism* both contain transporters homologous to *mptABCDE*, *Mesorhizobium ciceri* bv. *biserrulae* contains a transporter homologous to *eryEFG*. This operon also contains the same hypothetical gene that is found at the beginning of the *R. leguminosarum eryEFG* transcript. The transporters however, are arranged in a manner similar to that seen in *S. meliloti* and the gene encoding the regulator *eryD*, is found ahead of the transporter genes, whereas in *R. leguminosarum* and *Brucella*, *eryD* is found following *eryC* (Figure 4.1). We also note that whereas *M. loti* and *M. opportunism* both contain a putative fructose 1,6 bisphosphate aldolase gene between the *eryR-tpiB-rpiB* operon and *eryC*, a homolog to this is also gene is found adjacent to the *rpiB* in *Brucella*.

*Bradyrhizobium* sp. BTAi1, and ORS278, *A. cryptum* and *V. eiseniae* all have similar genetic arrangement to that of *S. meliloti*, except that they do not contain a homolog to *eryC*, or an associated *eryR-tpiB-rpiB* operon. These loci also differ primarily in their arrangement of *lala-rhtBC* (Figure 4.1).

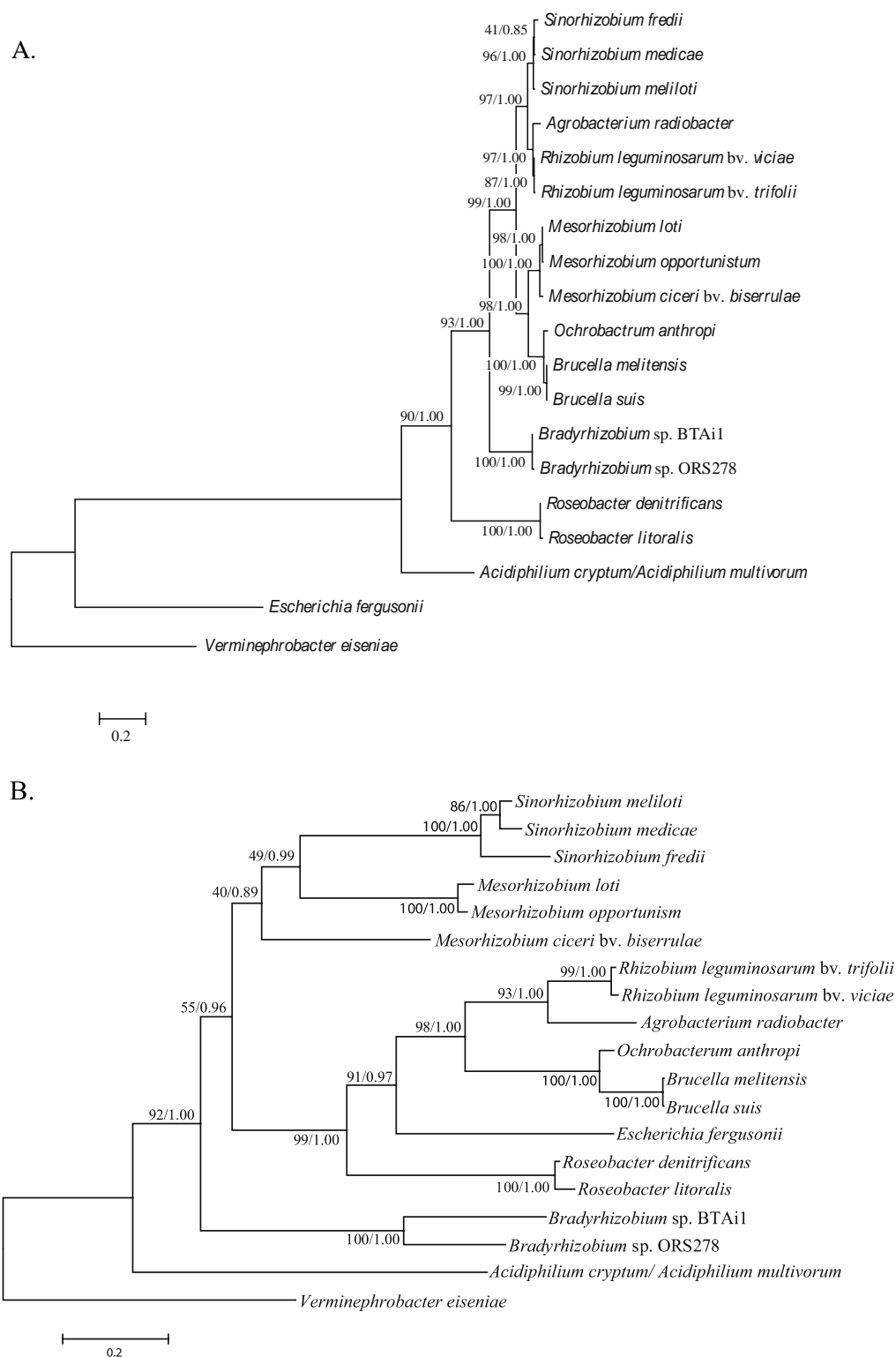
#### ***4.3.4 The phylogenies of erythritol proteins do not correlate with species phylogeny***

The amino acid sequences of RpoD were extracted from GenBank to analyze the phylogenetic relationships of the organisms examined in this study, using the most

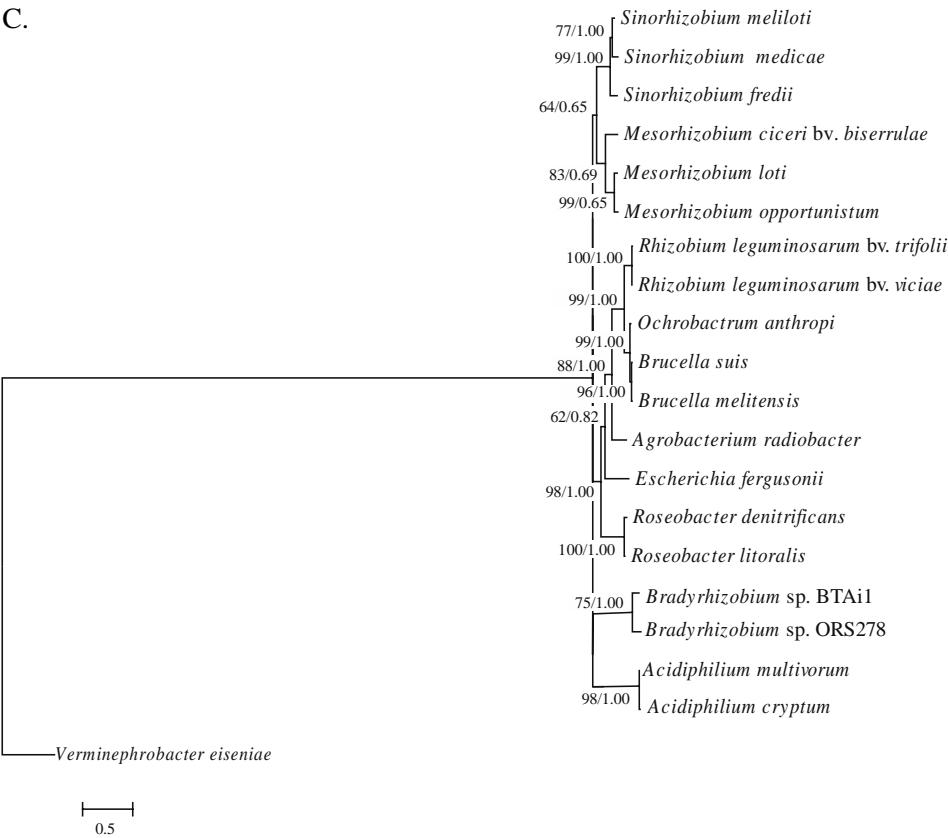
phylogenetically distant organism *V. eiseniae* as an out-group. The results of the RpoD sequence analyses was consistent with phylogenies that have been previously generated (Figure 4.2) (Crossman *et al.*, 2008). Initial comparison of the operon structures with the generated phylogenies suggested that the operon structure(s) did not correlate with the species phylogeny. Since the structure of some operons did not correspond well with the species phylogenies we wished to determine if operon structure did correlate with any of the erythritol genes found at the *S. meliloti* loci. Since homologs to EryA, EryB and EryD were ubiquitous through the data set, it was decided to construct phylogenies based on Maximum Likelihood and Bayesian analysis using the EryA, EryB and EryD data sets. The phylogenetic trees of EryB, EryC and EryD are presented in Figure 4.2 (B-D). *V. eiseniae* was also the most distant member with respect to the EryA phylogeny and again used as an outgroup. The phylogenetic trees of EryA, EryB and EryD were generally consistent (Figure 4.2). The species tree, based on RpoD, was included as a mirror tree with the EryA tree to demonstrate possible horizontal gene transfer events is presented in Figure 4.3.

The data show that there is a high degree of correlation between the loci configuration and the EryA phylogenetic tree (Figure 4.1, Figure 4.2, Figure 4.3). We note the similarity of the loci of *A. radiobacter* and *R. leguminosarum* to *Brucella* species and *O. anthropi* but not to the more closely related *Sinorhizobium* species. This suggests that a horizontal gene transfer may have occurred between these organisms. This is in agreement with what has been previously reported (Yost *et al.*, 2006). It also seems likely that a horizontal gene transfer event may have occurred between the *Brucella* and

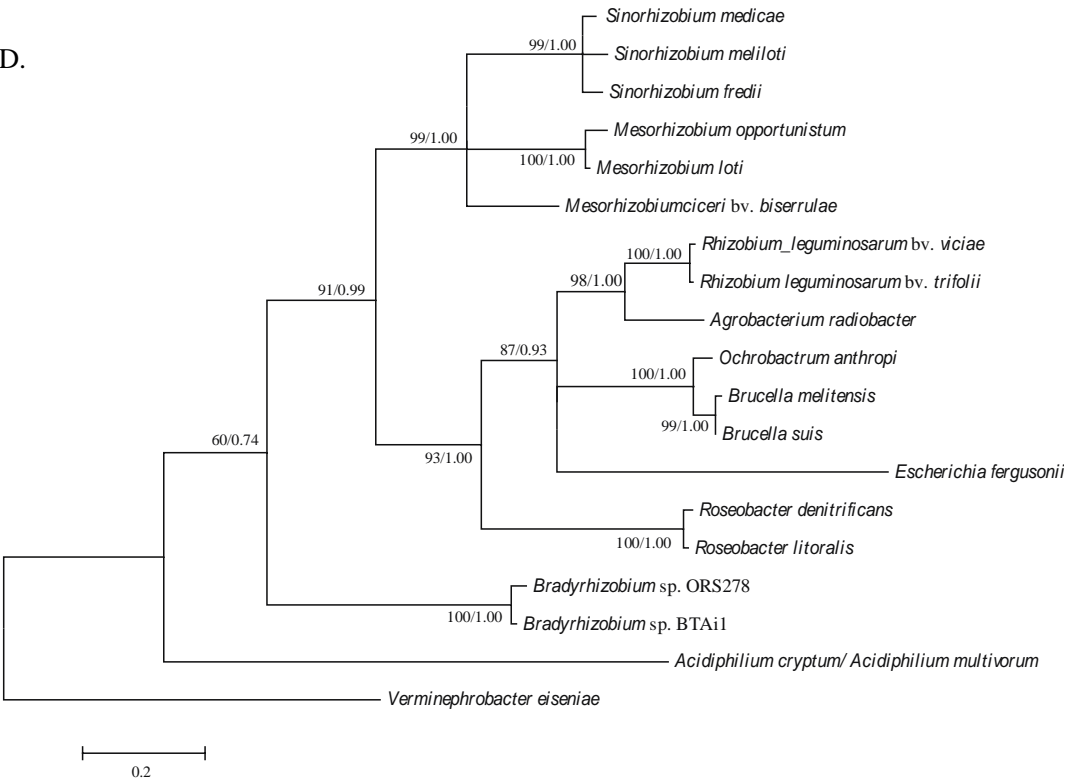
**Figure 4.2.** Phylogenetic trees of RpoD and the core erythritol genes EryA, EryB and EryD



C.

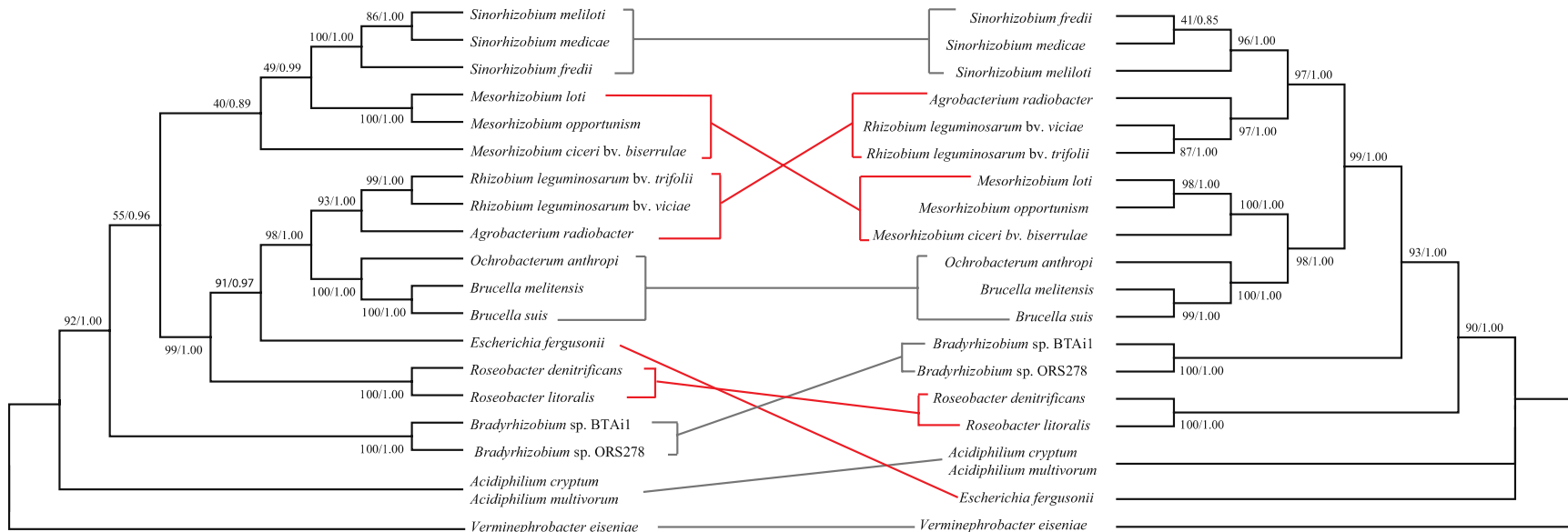


D.



RpoD species phylogenetic tree (**A.**), EryA phylogenetic tree (**B.**), EryB phylogenetic tree (**C.**) and EryD phylogenetic tree (**D.**) were constructed using ML and Bayesian analysis. Support for each clade is expressed as a percentage (Bayesian/ML, ie. posterior probability and bootstrap values respectively) adjacent to the nodes that supports the monophyly of various clades. The branch lengths are based on ML analysis and are proportional to the number of substitutions per site. Reproduced with modification from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

**Figure 4.3.** The phylogenetic tree of erythritol proteins does not correlate with species phylogeny; evidence for horizontal gene transfer



EryA phylogenetic tree (Left) and RpoD species tree (Right) were constructed using ML and Bayesian analysis. Support for each clade is expressed as a percentage (Bayesian/ML, ie. posterior probability and bootstrap values respectively) adjacent to the nodes that supports the monophyly of various clades. *V. eiseniae* was used as an outgroup for both trees since it was the most phylogenetically distant organism. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

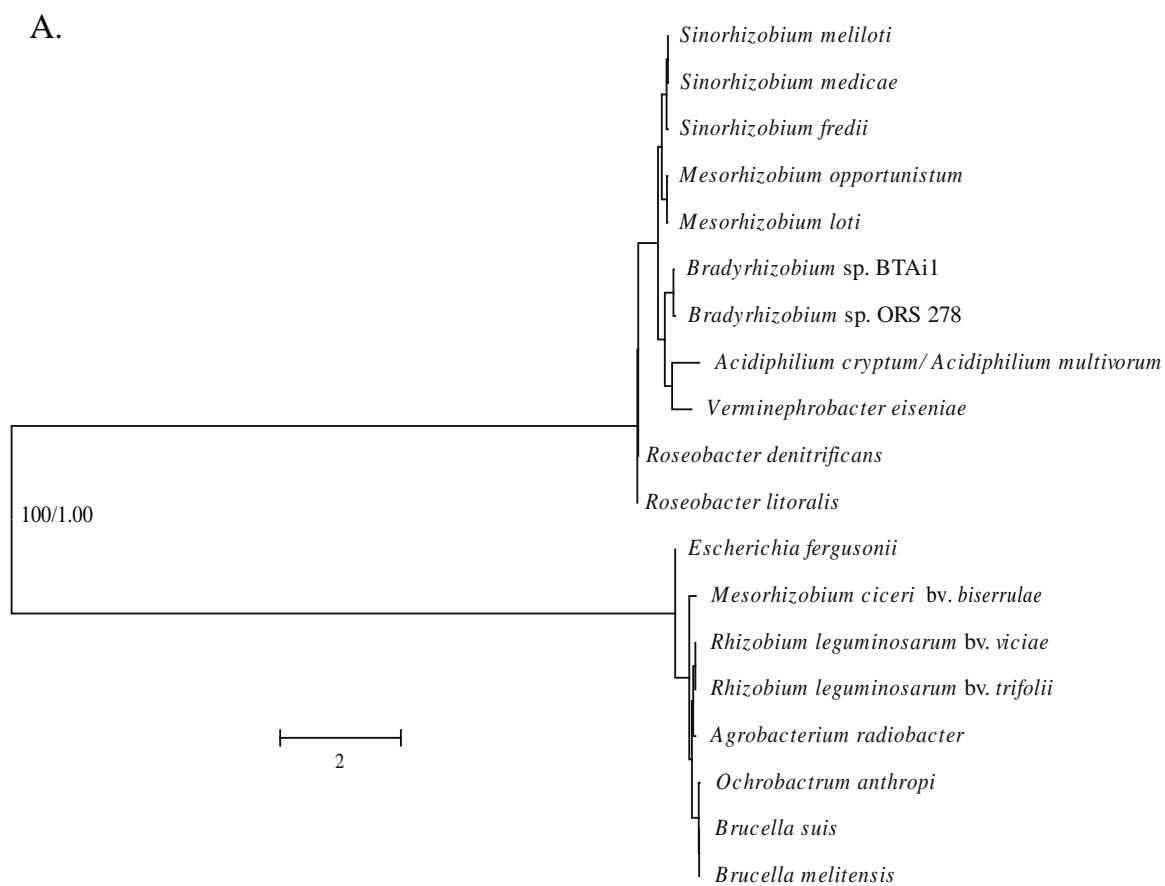


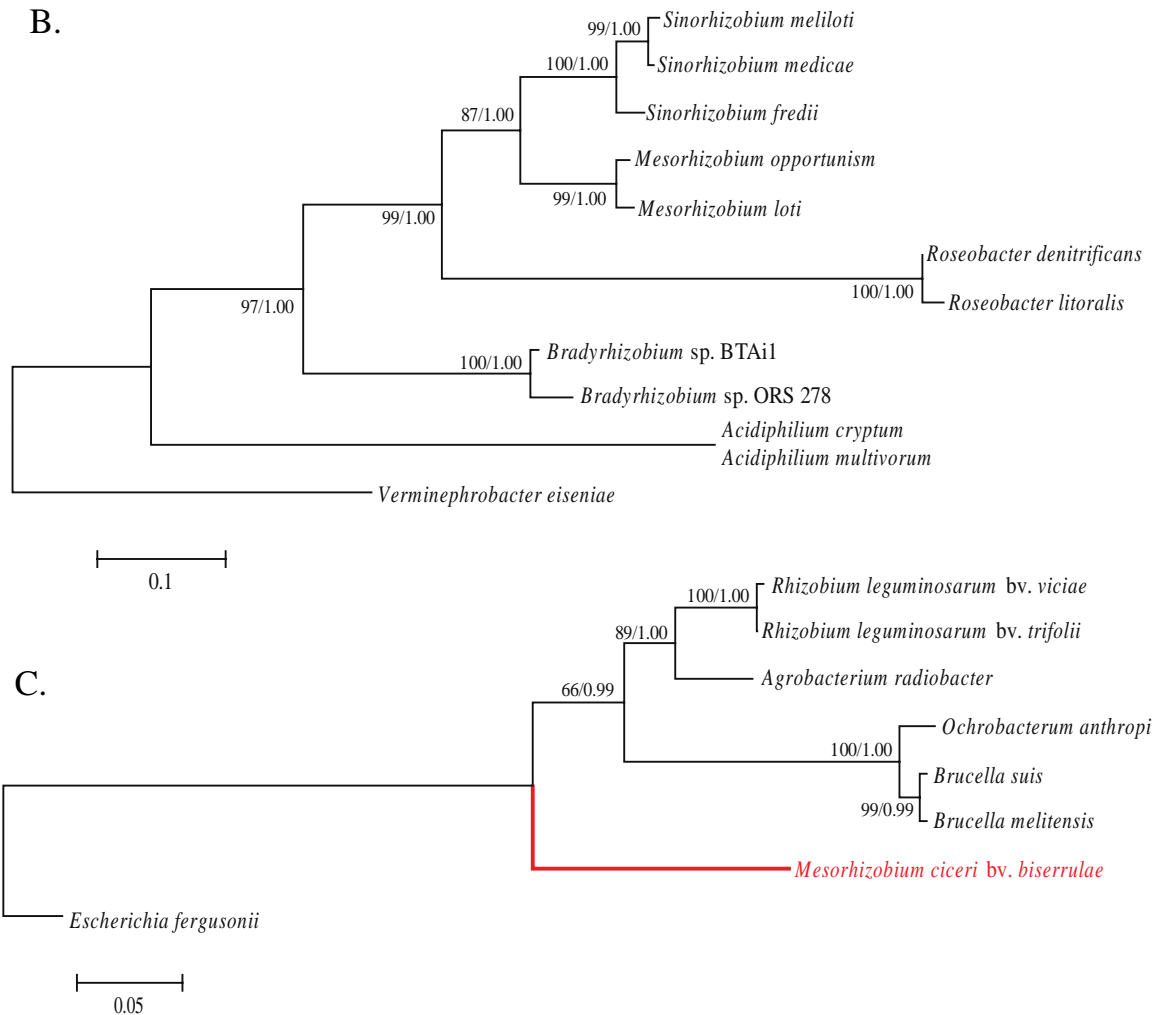
*E. fergusonii*. This may explain the unique occurrence of the loci's presence in a member of the gamma-proteobacteria. Finally, our mirror tree suggests that a horizontal gene transfer of the more complex erythritol locus may have occurred between *M. loti* and an ancestral species the *Sinorhizobium* species (Figure 4.3).

#### ***4.3.5 Modes of evolution for the polyol utilization loci***

Comparison of the phylogenetic trees of EryA, EryB and EryD to the arrangement and content of the loci led us to more thoroughly investigate the phylogenies of a number of proteins that stood out as unique within the data set. These phylogenies have led us to postulate modes of evolution that may have occurred in these loci.

BLASTP analysis showed a clear distinction between the type of transporter encoded by each of the loci and the remaining genetic content. In general, loci that contained adonitol/L-arabitol type genes contained a transporter homologous to the *S. meliloti* MptABCDE (Table 4.2, Figure 4.1). Loci that contained only erythritol genes contained a transporter homologous to the EryEFG of *R. leguminosarum*. One exception to this correlation was *M. ciceri* bv. *biserrulae* which contained a homologous transporter to EryEFG rather than MptABCDE. This is interesting because *M. ciceri* groups with the other *Mesorhizobia* in the EryABD trees. In order to analyze the evolution of these transporters more clearly, phylogenetic trees were constructed of homologs to EryG and homologs to MptA (Figure 4.4). In general the phylogenies are in agreement with the EryABD phylogenies, with the exception of *M. ciceri* which falls on a basal branch of the EryG phylogeny. The disparities between the EryG and EryABD phylogenies of *M. ciceri* strongly suggest that parts of its erythritol locus have a different

**Figure 4.4.** Phylogenetic trees of erythritol transporters



**A.** Unrooted phylogenetic tree including putative homologues to the sugar binding protein MptA of *Sinorhizobium meliloti* and EryG of *Rhizobium leguminosarum*. Support is provided for the node that clearly separates the putative homologues into two distinct and distant clades. Separate phylogenetic trees for erythritol transporters homologous to MptABCDE and EryEFG are also depicted using aligned amino acid sequences of the putative sugar binding proteins MptA (**B.**) and EryG (**C.**) as representatives of the transporters phylogenies. The branch that shows the anomalous placement of the *Mesorhizobium ciceri* bv. *biserrulae* within the tree of EryEFG homologs is highlighted in red. Trees were constructed using ML and Bayesian analysis. Support for each node is expressed as a percentage based on posterior probabilities (Bayesian analysis) and bootstrap values (ML). The branch lengths are based on ML analysis and are proportional to the number of substitutions per site. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

origin. This may have been the result of horizontal gene transfer of a second *R.*

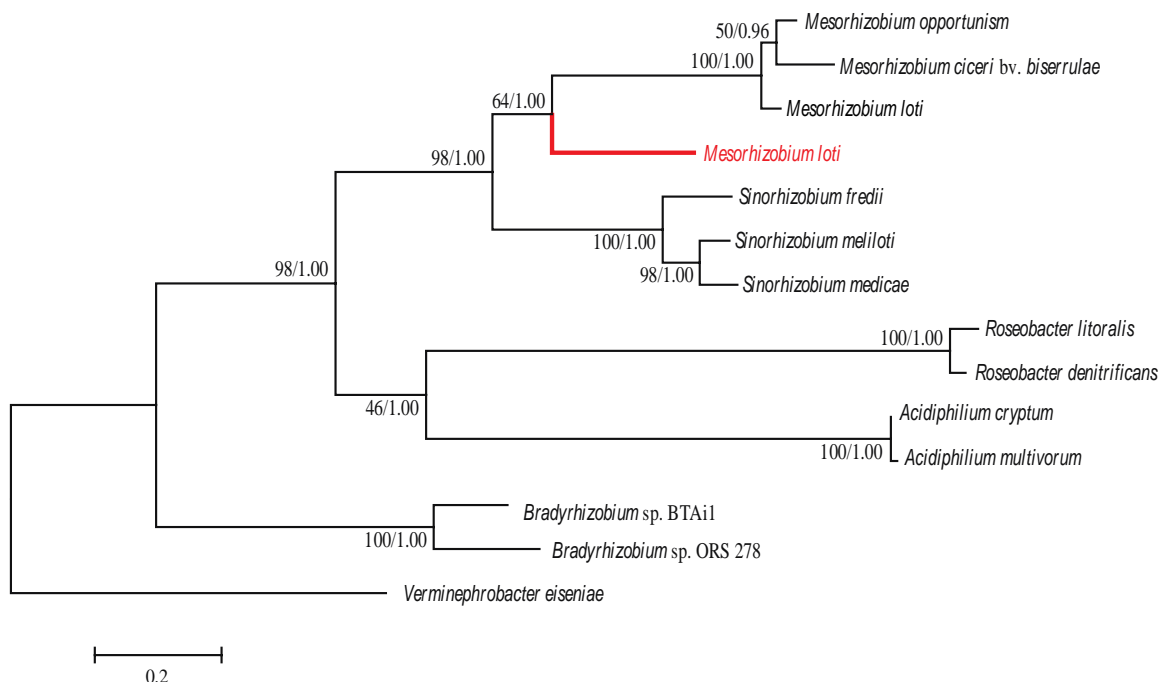
*leguminosarum*-type erythritol locus, followed by recombination between the two.

In two organisms, apparent duplications of genes were present. In *M. loti* one homolog of *lala* was present in the erythritol locus, while a second copy was present elsewhere in the genome adjacent to homologues of *rhtB* and *rhtC*, consistent with its location in the other two *Mesorhizobium* genomes. In *S. fredii* homologs to the apparent small operon that contains *eryR-tpiB-rpiB* were found both, as expected, in the erythritol locus, but also elsewhere on the chromosome in the same arrangement. To analyze the evolutionary history of these duplications phylogenetic trees were constructed for the *LalA* and *TpiB* homologs (Figure 4.5, Figure 4.6). The two copies of the *lala* gene in *M. loti* are most likely an example of paralogs, as they still group within the same clade among other *lala* homologs (Figure 4.5). The *tpiB* genes (Figure 4.6) in *S. fredii* are possible examples of xenologs (Koonin, 2005) as the phylogenetic tree shows that the two versions of the *tpiB* gene in *S. fredii* are only distantly related, with one homolog grouping within the expected clade that includes *S. medicae* and *S. meliloti* and the second homolog (not part of the main locus) showing monophyly with those found in a clade containing *R. leguminosarum* sp., *B. suis*, etc. (Figure 4.6).

#### 4.4 Discussion

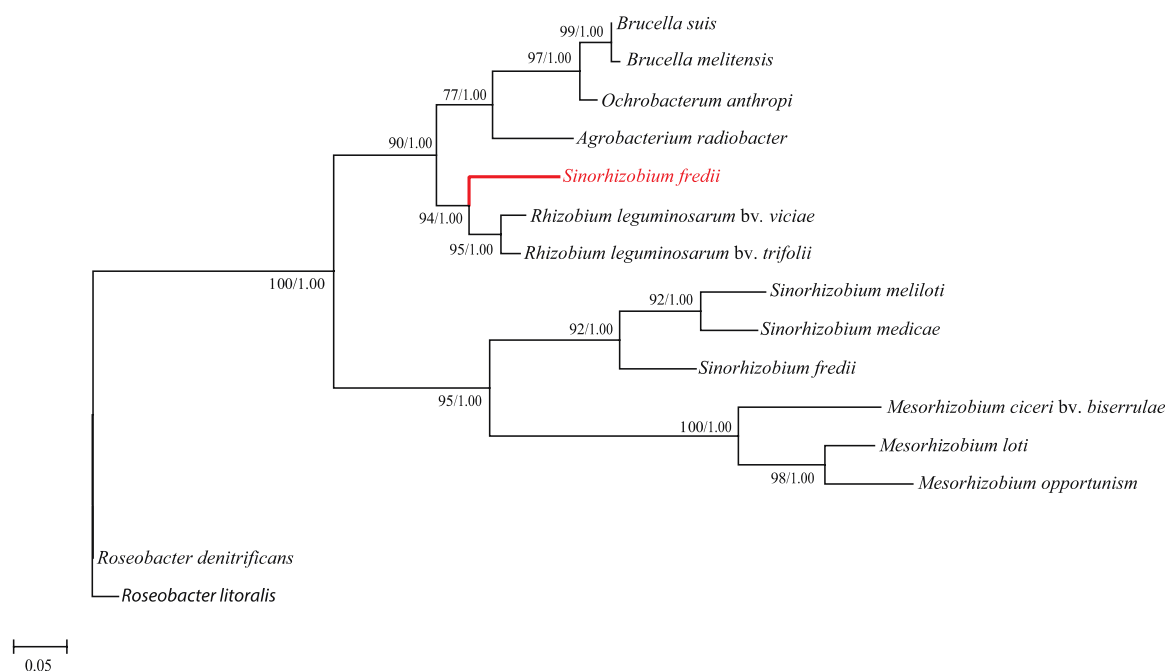
A number of models that are not mutually exclusive have been proposed to account for the formation and evolution of operons. Two broad aspects need to be considered, transfer of genes between organisms, as well as gathering and distributing genes within a genome. There is strong support for horizontal gene transfer as a driving

**Figure 4.5.** *Mesorhizobium loti* contains paralogs of LalA



The phylogeny of the L-arabitol catabolic gene LalA is depicted. *Mesorhizobium loti* contains a copy of *lalA* within an independent suboperon like the other *Mesorhizobium* species, as well as a second *lalA* homolog within the erythritol locus (Figure 4.1). The branch corresponding to the additional homolog within the erythritol locus is highlighted in red. The tree was constructed using ML and Bayesian analysis. Support for each node is expressed as a percentage based on posterior probabilities (Bayesian analysis) and bootstrap values (ML). The branch lengths are based on ML analysis and are proportional to the number of substitutions per site. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

**Figure 4.6.** *Sinorhizobium fredii* encodes TpiB xenologs



*Sinorhizobium fredii* contains a second suboperon that appears homologous to the *eryR-tpiB-rpiB* suboperon in the erythritol locus (Figure 4.1). The TpiB amino acid sequence was used as a representative of this suboperon to construct a phylogenetic tree. The branch corresponding to the TpiB encoded outside of the erythritol locus is highlighted in red. The tree was constructed using ML and Bayesian analysis. Support for each node is expressed as a percentage based on posterior probabilities (Bayesian analysis) and bootstrap values (ML). The branch lengths are based on ML analysis and are proportional to the number of substitutions per site. Reproduced from BMC Microbiology, volume 13, Geddes, B. A., Hausner, G., and I. J. Oresnik, Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria.

force for evolution of gene clusters (Lawrence & Roth, 1996). More recently, it has been shown that genes acquired by horizontal gene transfer events appear to evolve more quickly than genes that have arisen by gene duplication events (Treangen & Rocha, 2011). Within a genome the “piece-wise” model that suggests complex operons can evolve through the independent clustering of smaller “sub-operons” due to selection pressures for the optimization for equimolarity and co-regulation of gene products (Fondi *et al.*, 2009). Finally it has been suggested that the final stages of operon building can be the loss of “ORFan” genes (Fani *et al.*, 2005; Fondi *et al.*, 2009).

The data presented here provide examples supporting these models of operon evolution. The components of the polyol catabolic loci we have identified have been involved in at least 3 horizontal gene transfers within the proteobacteria (Figure 4.3). In addition, components such as the transporter *eryEFG* have been moved from the *R. leguminosarum* clade of loci into the *M. ciceri* bv. *biserrulae* polyol locus (Figure 4.4). The later species based on its phylogenetic position and category of polyol locus (*S. meliloti*) would have been expected to contain the *mtpA* gene. The presence of possible paralogs of *lala* (Figure 4.5) and the presence of *tpiB* xenologs (Figure 4.6) are also evidence for duplication and horizontal transfer events. Since *S. fredii* also contains a homolog to *tpiA* of *S. meliloti*, to our knowledge, this is the only example of an organism containing three triose-phosphate isomerases (Figure 4.6).

A striking example of a horizontal gene transfer and genetic rearrangement is exemplified by *M. ciceri* (Figure 4.1, Figure 4.3). It is likely that an exchange between *M. loti* and a common ancestor of *S. meliloti*, *S. medicae* and *S. fredii* NGR234 occurred. *M. loti* is located in the same clade as the *Brucella* and *O. anthropi* in the species tree

(Figure 4.3). Despite this, *M. loti* contains many of the genes corresponding to the adonitol and L-arabitol type loci of other species that cluster close to the base of the species tree such as *Bradyrhizobium* spp. (Figure 4.3). The presence of these factors in addition to the chimeric composition of the *M. loti* locus leads us to hypothesise that an ancestor of *M. loti* may have contained both an erythritol locus like that of the *Brucella* as well as a polyol type locus like that seen in the *Bradyrhizobia*, *A. cryptum* and *V. eiseniae*.

The *lala*, *rbtB*, *rbtC* suboperon appears to be the key component of the polyol locus in the *Bradyrhizobium* type loci (Table 4.2). Among the 19 loci identified, these three genes can be linked into a suboperon, embedded within the main locus (eg. *R. litoralis*) or split among two transcriptional units (see *A. cryptum* or *V. eiseniae*). As well, the gene module (or suboperon) *eryR*, *tpiB*-*rpiB* is presumable found in all erythritol utilizing bacteria. The acquisition of this module along with the *lala*, *rbtB* and *rbtC* suboperon may have allowed for the evolution of the more complex *S. meliloti* type locus (Figure 4.1).

The absence of *fucA* in *S. fredii* NGR234 and *M. loti* appears to be an example of the loss of an “ORFan” gene event having occurred. The gene is still present in *S. meliloti* however it has been shown that it is not necessary for the catabolism of erythritol, adonitol, or L-arabitol (Geddes & Oresnik, 2012a). It is likely that it was lost during the divergence of *M. loti* and *S. fredii* NGR234 from their common ancestors to *S. meliloti*. If this is true, it may be reasonable to assume that *fucA* may eventually also be lost from the *S. meliloti* erythritol locus.



In *S. meliloti*, erythritol uptake has been shown to be carried out by the proteins encoded by *mptABCDE* (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010), whereas in *R. leguminosarum* growth using erythritol is dependent upon the *eryEFG* (Yost *et al.*, 2006). Although both transporters appear to carry out the same function, the phylogenetic analysis clearly shows that they have distinct ancestors and may be best classified as analogues rather than orthologues (Figure 4.4). In addition, it has been shown that MptABCDE is also capable of transporting adonitol and L-arabitol (Geddes & Oresnik, 2012a). We note that these polyols appear to have stereo-chemical identity over three carbons and that EryA of *S. meliloti* can also use adonitol and L-arabitol as substrates (Geddes & Oresnik, 2012a). It is unknown whether EryA from *R. leguminosarum* has the ability to interact with these substrates.

The three distinct groups of loci we have identified probably correspond to the metabolic potential of these regions to utilize polyols. The locus of *S. meliloti* has been shown to contain the full complement of genes required to confer growth on using both erythritol and adonitol and L-arabitol as sole carbon sources (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010). Given that *S. fredii* NGR234 and *M. loti* each contain homologs to all of these genes, except for *fucA* which is not necessary for the catabolism of any of the sugars (Geddes & Oresnik, 2012a), it follows that these two loci may also be capable of catabolising all three polyols. It has also been established that the *B. abortus* and *R. leguminosarum* type loci are used for erythritol catabolism, and given the annotation and degree of relatedness (E value = 0) of proteins belonging to all species in the clade, it is not expected that these loci would be capable of breaking down additional polyols (Sangari *et al.*, 2000; Yost *et al.*, 2006). This is supported by the fact that the introduction

of the *R. leguminosarum* cosmid containing the erythritol locus into *S. meliloti* strains unable to utilize erythritol, adonitol, and L-arabitol were unable to be complemented for growth on adonitol and L-arabitol (Geddes & Oresnik, 2012a). It is however necessary to remember that some of identified loci are only correlated with polyol utilization based on our analysis and that basic biological function, such as the ability to utilize these polyols has not been previously described.

With the advent of newer generations of sequencing technologies a greater number of bacterial genomes will be sequenced. It is likely that more examples of rearrangements of catabolic loci through bacterial lineages will be observed. Since the acquisition of catabolic capacity of intermediary compounds does not directly affect an organism's ability to survive, we suggest that catabolic loci may make optimal models to observe operon evolution. Since the ability to catabolize erythritol is found in relatively few bacterial species, operons that encode erythritol and other associated polyols may be ideal models to observe operon evolution.

## **Chapter 5**

### **Inability to Catabolize Galactose Leads to an Increased Ability to Compete for Nodule Occupancy in *Sinorhizobium meliloti***

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology. This work was carried out by Barney Geddes with technical assistance from Peiki Loay, Amanda Appasamy and Roy Hutchings.

## 5.1 Introduction

*Sinorhizobium meliloti* is a Gram-negative soil microbe that is capable of fixing atmospheric nitrogen in a symbiotic relationship with legumes such as alfalfa (*Medicago sativa*). It also exists as a free-living saprophytic organism in the soil. Within the rhizosphere plants exude an array of organic compounds, including the signaling molecules that lead to the establishment of symbiosis between Rhizobia and their host plants (Oldroyd *et al.*, 2011). A problem in agriculture is the inability of commercial rhizobial inoculums to effectively compete with indigenous strains for nodule occupancy. This limits the ability of inoculums to increase legume crop yields (Triplett & Sadowsky, 1992). Several studies have shown that the ability to utilize organic compounds is important for competition for nodule occupancy in *S. meliloti* and *Rhizobium leguminosarum* (Ding *et al.*, 2012; Fry *et al.*, 2001; Jiménez-Zurdo *et al.*, 1995; Kohler *et al.*, 2010; Oresnik *et al.*, 1998; Wiebo *et al.*, 2007; Yost *et al.*, 2006).

One approach to isolating genetic determinants that affect competition for nodule occupancy is to identify catabolic genes involved in the utilization of sugars which are prevalent in plant exudate. Pea mucilage has been shown to support the growth of *R. leguminosarum*. Analysis of its sugar composition identified arabinose and galactose as accounting for over 60% of sugars in pea mucilage (Knee *et al.*, 2001). The genetic locus for arabinose catabolism has been identified in *S. meliloti*, however the inability to catabolise arabinose did not significantly affect competition for nodule occupancy (Poysti *et al.*, 2007). Alfalfa seed wash and root wash have been shown to contain galactose as well as several different galactosides. Additionally, a promoter for the  $\alpha$ -galactoside transport gene *mela* (*pmela*) is inducible by these compounds and was shown to be

induced in the rhizosphere of alfalfa seedlings in sterilized and unsterilized soil (Bringinghurst *et al.*, 2001). Galactose is present in plant cell walls as many different galacturonans and galactans (McNeil *et al.*, 1984). Plant derived polygalacturonase, an enzyme associated with cell wall degradation, has been shown to be induced during infection of *Medicago sativa* by *S. meliloti* (Muñoz *et al.*, 1998). Therefore, galactose utilization may be important during infection of the plant as well as colonization of the rhizosphere.

Galactose is catabolised through the De Ley-Doudoroff pathway in *S. meliloti* rather than the highly conserved Leloir pathway (Arias & Cerveñansky, 1986). The initial step of the De Ley-Doudoroff pathway is the reduction of D-galactose to D-galactono- $\gamma$ -lactone by a galactose dehydrogenase; D-galactono- $\gamma$ -lactone is then used in a lactonase reaction to generate D-galactonate; D-galactonate is dehydrated to 2-keto-3-deoxygalactonate which is phosphorylated to produce 2-keto-3-deoxy-6-phosphogalactonate (KDPgal). The pathway culminates with a KDPgal aldolase reaction which produces pyruvate and D-glyceraldehyde 3-phosphate (De Ley & Doudoroff, 1957). These enzymatic activities were identified in *S. meliloti* and shown to be inducible by galactose. A galactose mutant lacking KDP-gal aldolase activity was isolated and shown to be unable to utilize galactose as a sole carbon source (Arias & Cerveñansky, 1986).

In this work we identify a genetic locus involved in galactose catabolism on the chromosome of *S. meliloti*. We show that transport of galactose is redundant and partially facilitated by the arabinose transporter AraABC. Finally we demonstrate unusual

regulation in a mutant of galactose metabolism, and show that this mutant has an increased ability to compete for nodule occupancy compared to the wild-type strain.

## 5.2 Materials and Methods

### 5.2.1 Bacterial strains, plasmids and media

Bacterial strains and plasmids used and generated in this work are listed in Table 5.1. *S. meliloti* strains were grown routinely at 30°C on either Luria-Bertani (LB) as a complex medium (Sambrook *et al.*, 1989) or Vincent's minimal medium (VMM) as a defined medium (Vincent, 1970). Carbon sources were filter sterilized and added to VMM to a final concentration of 15 mM unless otherwise stated. Seed exudate was prepared as previously described (Mulligan & Long, 1985). When required, *S. meliloti* and *Escherichia coli* strains were grown in the following concentrations of antibiotic (µg/mL): chloramphenicol (Cm) 20; gentamicin (Gm) 20 or 60; kanamycin (Km) 20; neomycin (Nm) 200; rifampicin (Rf) 50; streptomycin (Sm) 200; tetracycline (Tc) 5. All antibiotics were filter sterilized before use. *A. tumefaciens* strains were grown at 30°C in MG/L (Young & Nester, 1988) (containing Tc 10 µg/mL when relevant) and induction broth (IB) containing 10 mM galactose and 5 µM acetosyringone as previously described (Charles *et al.*, 1994).

### 5.2.2 Genetic techniques

Conjugations and transductions were carried out essentially as previously described (Finan *et al.*, 1988; Finan *et al.*, 1984). Since selecting A348 derivatives with Tc can result in activation of a cryptic Tc<sup>R</sup> gene (Luo *et al.*, 2001b), Tc resistant

**Table 5.1.** Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
<b>Strain</b>		
<i>S. meliloti</i>		
Rm1021	SU47 <i>str-21</i> , Sm <sup>R</sup>	(Meade <i>et al.</i> , 1982)
SRmA240	RmG212 <i>araF</i> ::Tn5-B20, Nm <sup>R</sup>	(Poysti <i>et al.</i> , 2007)
SRmA503	Rm1021 <i>araA4</i> ::Tn5-B20, Nm <sup>R</sup>	(Poysti <i>et al.</i> , 2007)
SRmD144	Rm1021 <i>dgoK1</i> ::Tn5, Nm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD298	Rm1021 <i>galD</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD311	Rm1021 <i>dgoA</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD312	Rm1021 <i>SMc00883</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD313	Rm1021 <i>ilvD1</i> ::pKNOCK-Gm, Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD315	RmG212 <i>araF</i> ::Tn5-B20, <i>ilvD1</i> ::pKNOCK-Gm, Nm <sup>R</sup> , Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
STM.4.09.D04	Rm2011 <i>Sma0203</i> ::Tn5, Nm <sup>R</sup>	(Pobigaylo <i>et al.</i> , 2006)
STM.2.10.H05	Rm2011 <i>SMB21343</i> ::Tn5, Nm <sup>R</sup>	(Pobigaylo <i>et al.</i> , 2006)
STM.3.09.B05	Rm2011 <i>SMB20931</i> ::Tn5, Nm <sup>R</sup>	(Pobigaylo <i>et al.</i> , 2006)
<i>A. tumefaciens</i>		
A348		(Raineri <i>et al.</i> , 1993)
AT11043	A348 <i>virE358</i> ::Tn3-HoHo1	(Charles & Nester, 1993)
AT11055	A348 <i>chvE</i> ::Tn5, <i>virE358</i> ::Tn3-HoHo1	(Charles <i>et al.</i> , 1994)
<i>E. coli</i>		
DH5α	λ <sup>-</sup> Φ80dlacZ <sup>o</sup> M15 <sup>o</sup> ( <i>lacZYA-argF</i> )U169 <i>recA</i> <i>lndA1</i> <i>hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) <i>supE44</i> <i>thi-1</i> <i>gyrA</i> <i>relA1</i>	(Hanahan, 1983)
DH5α Rif	Rifampicin resistant DH5α	(House <i>et al.</i> , 2004)
DH5α λpir	λpir lysogen of DH5α	(House <i>et al.</i> , 2004)
EcA101	MT607ΩTn5-B20, Kan <sup>R</sup>	(Clark <i>et al.</i> , 2001)
MM294A	<i>pro-82-thi-1</i> <i>hsdR17</i> <i>supE44</i>	(Finan <i>et al.</i> , 1986)
MT607	MM294A <i>recA56</i>	(Finan <i>et al.</i> , 1986)
MT616	MT607 (pRK600)	(Finan <i>et al.</i> , 1986)
S17-1	<i>recA</i> derivative of MM294A with integrated RP4-2 (Tc::Mu::Km::Tn7)	(Simon <i>et al.</i> , 1983)
EcA101	MT607ΩTn5-B20, Kan <sup>R</sup>	(Clark <i>et al.</i> , 2001)
<b>Plasmids</b>		
pAA10	Galactose complementing cosmid, Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG56	pCO36 <i>dgoK1</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG57	pCO37 <i>araA</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG58	pKNOCK-Gm <i>dgoA</i> , Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG60	pKNOCK-Gm <i>SMc00883</i> , Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG62	pKNOCK-Gm <i>ilvD1</i> , Gm <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG65	pRK7813 <i>chvE</i> <i>gbpR</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)

---

pCO37	pRK7813 containing <i>attB</i> sites, Gateway® compatible destination vector	(Jacob <i>et al.</i> , 2008)
pPH1JI	IncP plasmid, Gm <sup>R</sup>	(Beringer <i>et al.</i> , 1978)
pKNOCK-Gm	Suicide vector, Gm <sup>R</sup>	(Alexeyev, 1999)
pMK2014	FRT- <i>ccdB</i> -Cam <sup>R</sup> -FRT cassette, Pen <sup>R</sup>	(House <i>et al.</i> , 2004)
pMK2015	FRT- <i>ccdB</i> -Cam <sup>R</sup> -FRT cassette, Pen <sup>R</sup>	(House <i>et al.</i> , 2004)
pRH1	pAA10 <i>fabG2</i> ::Tn5-B20, Tc <sup>R</sup> , Nm <sup>R</sup>	(Geddes & Oresnik, 2012b)
pRK600	pRK2013 <i>nptI</i> ::Tn9, Cm <sup>R</sup>	(Finan <i>et al.</i> , 1986)
pRK602	pRK600ΩTn5, Cm <sup>R</sup> , Nm <sup>R</sup>	(Finan <i>et al.</i> , 1985)
pRK7813	Broad-host-range vector, Tc <sup>R</sup>	(Jones & Gutterson, 1987)
pSL4	pUC19 <i>mob chvE gbpR</i>	(Doty <i>et al.</i> , 1993)
pXINT129	Kan <sup>R</sup> , λint and <i>xis</i> driven by P <sub>lac</sub>	(Platt <i>et al.</i> , 2000)
pZW1	CX1-derived arabinose complementing cosmid, Tc <sup>R</sup>	(Poysti <i>et al.</i> , 2007)
pZW3	pZW1, <i>araB8</i> ::Tn5-B20, Tc <sup>R</sup> , Nm <sup>R</sup>	(Poysti <i>et al.</i> , 2007)

---

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology



*A. tumefaciens* transconjugants were screened for the presence of plasmids using Eckhardt gel electrophoresis (Hynes & McGregor, 1990). Random Tn5 mutagenesis of Rm1021 was carried out using pRK602 as previously described to generate SRmD144 (Finan *et al.*, 1985). A cosmid bank was mated *en masse* into SRmD144 and pAA10 was isolated based on the ability to complement SRmD144 for galactose utilization (Friedman *et al.*, 1982). The ends of the cosmid were sequenced to ensure that it contained the entire galactose locus. Mutagenesis of the cosmid was carried out by mating pAA10 into EcA101 (Clark *et al.*, 2001) and selecting for co-transfer of the cosmid and a Tn5::B20 marker when mated out of EcA101 and into DH5 $\alpha$  Rif (Poysti *et al.*, 2007). pRH1 was identified based on the inability to complement SRmD144 for galactose utilization. The Tn5-B20 in pRH1 was then recombined into the chromosome of Rm1021 using pPH1JI as previously described to generate SRmD298 (Glazebrook & Walker, 1991). The positions of Tn5 and Tn5-B20 insertions were verified using arbitrary PCR as previously described (Miller-Williams *et al.*, 2006; Poysti *et al.*, 2007). In addition, phenotypes were shown to be 100% co-transducible with the Nm<sup>R</sup> encoded by the Tn5.

### 5.2.3 DNA manipulations

Standard techniques were used for plasmid isolation, restriction enzyme digests, ligations, transformations, and agarose gel electrophoresis (Sambrook *et al.*, 1989). Nucleotide sequencing was carried out by cycle-sequencing using a Big-Dye version 3.1 kit. Sequencing reactions were carried out as recommended by the manufacturer and resolved using an ABI 3130 sequencer.

To construct single crossover mutants of *dgoA*, *SMc00883*, and *ilvD1*, approximately 400-bp internal fragments of the genes were amplified using Rm1021 genomic DNA as template. The amplicon was restricted and ligated into the suicide vector pKNOCK-Gm to create pBG58, pBG60 and pBG62 in DH5 $\alpha$   $\lambda$ pir (Alexeyev, 1999). Plasmids were conjugated into Rm1021 and single crossovers were selected for using Sm and Gm as previously described (Richardson *et al.*, 2004). Single crossover mutations were confirmed by PCR and linkage in transduction to the Tn5 of SRmD144.

To construct plasmids expressing *araA* or *dgoK1*, the Gateway<sup>®</sup> compatible destination vector pCO37 was used. pCO37 is a derivative of the broad host range plasmid pRK7813 that contains *attB* sites (Jacob *et al.*, 2008). This facilitated the recombination of *araA* and *dgoK1* from an *S. meliloti* ORFeome entry vector into pCO37 as previously described (Geddes *et al.*, 2010). Identities of the resulting plasmids were confirmed by sequencing of the inserts.

To construct pBG65, the pUC19Mob plasmid pSL4 was digested with *EcoRI* to release a 3 kb *EcoRI* fragment containing *chvE* and *gbpR* of *A. tumefaciens*. This fragment was co-digested with pRK7813, ligated and transformed into *E. coli*. The resulting plasmid was purified and analyzed by restriction digest, followed by sequencing the ends of the 3 kb insertion to confirm the construction.

#### **5.2.4 Non-denaturing gel electrophoresis and dehydrogenase assays**

Cells were grown in defined medium, cell free lysates were prepared, and samples were separated using non-denaturing polyacrylamide gel electrophoresis (PAGE) as previously described (Pickering & Oresnik, 2008). Gels were developed using *p*-

nitroblue tetrazolium based dehydrogenase stain as previously described (Latner & Skillen, 1968). Samples were incubated either with or without  $\text{NAD}^+$  or  $\text{NADP}^+$ , in the presence or absence of substrate to determine the specificity of activity bands.

### **5.2.5 Transport competition assays**

Sugar uptake was measured using  $[1\text{-}^3\text{H}]$  arabinose ( $370\text{ GBq mmol}^{-1}$ ) and  $[1\text{-}^{14}\text{C}]$  arabinose ( $11.8\text{ GBq mmol}^{-1}$ ). Transport competition assays were carried out essentially as previously described with the following modifications (Glazebrook & Walker, 1991; Poysti *et al.*, 2007). The assay was initialized by the addition of  $1\text{ }\mu\text{M}$  labelled arabinose to cell culture which was grown in defined medium, washed twice and resuspended to an  $\text{OD}_{600}$  of 0.1-0.3. To measure competition,  $15\text{ }\mu\text{M}$  unlabelled sugar was added with the labelled arabinose to compete for uptake. Aliquots were withdrawn at 10-40s time-points for  $^3\text{H}$  arabinose, and 10s-5m for  $^{14}\text{C}$  arabinose. Aliquots were filtered through a Millipore  $0.45\text{ }\mu\text{m}$  Hv filter on a Millipore sampling manifold. The amount of radioactivity retained by the cells was quantitated by a liquid scintillation counter (Beckman LS6500).

### **5.2.6 RNA isolation, cDNA synthesis and quantitative RT-PCR**

Bacterial cultures were grown to an  $\text{OD}_{600}$  of approximately 0.4 in defined medium or undiluted seed exudate. Cells were harvested by centrifugation and resuspended in TE (Tris 10 mM, EDTA 1 mM, pH= 8) buffer with lysozyme (0.4 mg/ml). RNA was then isolated using the Qiagen RNAeasy kit as previously described (Geddes *et al.*, 2010). Approximately 200 ng of cDNA sample was used as a template for

quantitative RT-PCR. Reactions were performed using the SYBR green RT-PCR kit from Invitrogen as recommended by manufactures. The RT-PCR reaction was performed using a Cepheid Smart Cycler with the following program: Stage 1, 95°C for 120s, 1 time; Stage 2, 95°C for 15s followed by 60°C for 30s, repeated 40 times; Stage 3, melting curve analysis of PCR products. Primers unique to this work included galDF:

CGGATCGATCGTCAATTTCT, galDR: ATTCTTCCGTGCGCCACAAG, araAF:

GAGGGTCTGAAGCTCGACAG, araAR: ATCGACAGAACCGTCATAGGG, dgoAF:

AGCGGCCTCAAGTTCTTTC, and dgoAR: TAGTCGGCGAAATTGACCTC.

*SMc00128* was used as a reference gene because its expression level has been shown to be similar under a number of different conditions (Krol & Becker, 2004).

#### ***5.2.7 Competition for nodule occupancy assays***

Alfalfa seeds were surface sterilized and germinated on water-agar plates before planting in nitrogen-free medium. Competition for nodule occupancy was assessed by inoculating plants after 3 days of growth with a mixture of Rm1021 and SRmD144 at ratios of 1:1 and 5:1 respectively. Cultures of both strains were grown and diluted to the same OD<sub>600</sub>. A mixture of the strains was diluted 100-fold into 10 mL of sterile distilled water. Serial dilutions of the mixed inoculum spread-plated onto LB agar plates and at least 50 colonies were patched onto the appropriate antibiotics to differentiate the proportion of each strain in the inoculum. Plants were inoculated with the 10 mL inoculum mixture and grown for 28-35 days, after which time root nodules were manually removed. At least 50 root nodules from each experiment were sterilized and crushed into 100 µL sterile distilled water. Aliquots of each nodule extract were spotted

onto LB plates containing antibiotics for strain differentiation. Competition was assessed by comparing the proportion of the mutant strain found in the inoculum to the proportion that was isolated from nodules. Significance was evaluated using a paired T-test, assuming a P-value of less than 0.05 indicated a significant difference in competitiveness.

#### ***5.2.8 $\beta$ -galactosidase assays***

*A. tumefaciens* cells maintaining plasmids were subcultured from MG/L medium and grown overnight to an OD<sub>600</sub> of approximately 0.5 in induction broth containing 10 mM galactose and 5  $\mu$ M acetosyringone.  $\beta$ -galactosidase assays were performed essentially as described previously (Clark *et al.*, 2001). Three independent biological replicates were grown and used for the  $\beta$ -galactosidase assay.

### **5.3 Results**

#### ***5.3.1 Identification of the galactose locus***

In order to identify the locus for galactose catabolism in *S. meliloti*, a random Tn5 mutagenesis of the wild-type strain Rm1021 was performed. Approximately 5,000 mutants that contained Tn5 insertions were screened on defined medium that contained galactose as a sole carbon source. One mutant, SRmA144, was isolated that was unable to utilize galactose as a sole carbon source but able to utilize glycerol and glucose. Utilizing an arbitrary PCR protocol a fragment adjacent to the insert was generated and sequenced (Poysti & Oresnik, 2007). A BLASTN search of the *S. meliloti* database was performed using the nucleotide sequence of this amplicon, and the gene containing the

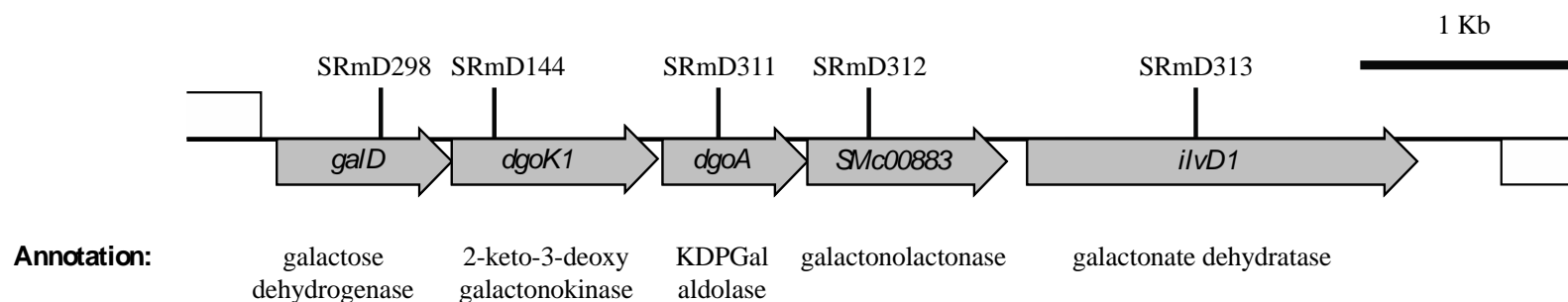
Tn5 was identified on the chromosome as *dgoK1* (Systematic identifier number SMc00881) (Figure 5.1).

The locus surrounding *dgoK1* contains 4 other annotated open reading frames encoded on the same strand as *dgoK1* (Figure 5.1). Analysis of the predicted protein coding sequences of *dgoK1* and surrounding genes was carried out using InterPro Scan (Hunter *et al.*, 2009), and BLASTP (Altschul *et al.*, 1997). As its annotation suggests, *dgoK1* is predicted to encode a 2-keto-3-deoxy galactonokinase (IPR007729). The gene immediately upstream of *dgoK1* was annotated as *fabG2* (suggested annotation *galD*), and is predicted to encode a NAD<sup>+</sup> or NADP<sup>+</sup> binding short-chain dehydrogenase (IPR002198). The gene immediately downstream of *dgoK1* is annotated as *dgoA*. It belongs to IPR000887 which includes 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase. The remaining two genes are downstream of *dgoA*. One is unannotated and represented by the systematic identifying number SMc00883 and the other is annotated *ilvD1*. SMc00883 has some similarity to the IPR013568 family which includes gluconolactonases. Finally *ilvD1* is predicted to encode a dehydratase belonging to IPR000581 which includes 6-phosphogluconate dehydratases. Based on our analysis of the protein coding sequences predicted to be encoded by genes at this locus we hypothesised that the locus encodes all five of the enzymes of the De Ley-Doudoroff pathway of galactose catabolism (Figure 5.1).

### 5.3.2 Genetic characterization of the galactose locus

To test the hypothesis that this locus contained all the genes necessary for the De Ley-Doudoroff pathway a complementing cosmid was first isolated. A cosmid bank was

**Figure 5.1.** Genetic map of the galactose catabolic locus of *S. meliloti*



Boxes indicate ORFs. Vertical lines indicate position of insertion mutations constructed or isolated, strain numbers are shown above each vertical line. Annotation below genes represents predicted function in the De Ley-Doudoroff pathway. Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

mated *en masse* into SRmD144. Transconjugants were selected for their ability to utilize galactose as a sole carbon source on defined medium and the resulting complementing cosmid, pAA10, was isolated. Subsequent Tn5-B20 mutagenesis of pAA10 yielded a single insert which was unable to complement SRmD144 for galactose utilization. The Tn5-B20 was found within the first gene in the locus, previously annotated as *fabG2*. This insert was subsequently recombined into the chromosome constructing SRmD298. SRmD298 was unable to utilize galactose as a sole carbon source. Based on bioinformatic analysis, gene context, and biochemical evidence (see later) this gene was renamed *galD* (galactose dehydrogenase).

Since the mutagenesis of pAA10 did not yield inserts in *dgoA*, *SMc00883*, or *ilvD1*, it was decided that mutations should be constructed in each of the remaining genes. Single cross-over mutations were constructed using internal fragments from *dgoK1*, *SMc00883*, and *ilvD1* that were cloned into pKNOCK-Gm and mobilized into Rm1021 to construct SRmD311, SRmD312, and SRmD313 respectively.

Mutations in the galactose locus were screened for their ability to grow on defined medium with galactose as a sole carbon source and tested for complementation by pAA10, pRH1 and pBG56 (Table 5.2). SRmD144 (*dgoK1*), SRmD298 (*galD*), and SRmD311 (*dgoA*) were all unable to utilize galactose as a sole carbon source, and complemented for galactose utilization by pAA10. SRmD312 (*SMc00883*) and SRmD313 (*ilvD1*) were able to grow using galactose as a sole carbon source. Therefore, despite the implication of a role in the De Ley-Doudoroff pathway based on their annotation and genetic context, *SMc00883* and *IlvD1* are either not used during galactose catabolism, or their biochemical activities are redundant with proteins encoded elsewhere



**Table 5.2.** Growth of strains on galactose

Strain	Genotype	Plasmid			
		pRK7813 empty vector	pAA10 <i>gal</i> locus	pRH1 <i>gal</i> locus <i>galD</i> ::Tn5-B20	pBG56 <i>dgoK1</i>
Rm1021	Wild type	+	+	+	+
SRmD298	<i>galD</i> ::Tn5-B20	-	+	-	-
SRmD144	<i>dgoK1</i> ::Tn5	-	+	-	+
SRmD311	<i>dgoA</i> ::pKNOCK-Gm	-	+	-	-
SRmD312	<i>SMc00883</i> ::pKNOCK-Gm	+	+	ND <sup>a</sup>	ND
SRmD313	<i>ilvD1</i> ::pKNOCK-Gm	+	+	ND	ND

Growth is as follows: +, wild type; -, no growth on defined media with galactose as a sole carbon source.

<sup>a</sup> Not Done

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

in the genome. The inability of pRH1 to complement SRmD144 or SRmD311 for galactose utilization suggests that *galD*, *dgoK1*, and *dgoA* are encoded by one transcriptional unit. The ability of pBG56 to complement SRmD144 shows that *dgoK1* is necessary for galactose catabolism.

Mutants were also screened for the ability to grow on defined medium that contained both 15 mM glycerol and 0.5 mM galactose as carbon sources. SRmD144 (*dgoK1*), SRmD298 (*galD*), SRmD312 (*SMc00883*) and SRmD313 (*ilvD1*) were able to grow in the presence of both sugars, whereas SRmD311 (*dgoA*) did not grow (Table 5.3). This type of toxic phenotype has previously been suggested to be the result of the build-up of a phosphorylated intermediate (Adhya & Shapiro, 1969; Power, 1967; Richardson *et al.*, 2004). The inability of a *dgoA* mutation to grow on this medium is consistent with the placement of the aldolase reaction in the De Ley-Doudoroff pathway following a phosphorylation; presumably carried out by DgoK1.

Mutants carrying plasmids were also tested for a toxic phenotype (Table 5.3). SRmD311 was complemented by pAA10 for the ability to grow on glycerol and galactose. Although SRmD298 (*galD*) was able to grow on glycerol and galactose, it was unable to grow carrying pBG56 (*dgoK1*). We hypothesised that SRmD298 was polar on *dgoK1* and *dgoA*, and the presence of a heterologous copy of *dgoK1* allowed galactose to be catabolised to the phosphorylated intermediate KDP-gal in the absence of *galD* and *dgoA*. Therefore *galD* may not be required for galactose catabolism. To provide evidence that SRmD298 was polar on *dgoK1* and *dgoA* we performed qRT-PCR analysis of *dgoA* expression in SRmD298 and SRmD144. Transcription of *dgoA* was detected in SRmD144 but not SRmD298.

**Table 5.3.** Toxic growth phenotypes of strains grown on galactose and glycerol

Strain	Genotype	Plasmid		
		pRK7813 empty vector	pAA10 <i>gal</i> locus	pBG56 <i>dgoK1</i>
Rm1021	Wild type	+	+	+
SRmD298	<i>galD</i> ::Tn5-B20	+	+	-
SRmD304	<i>dgoK1</i> ::Tn5	+	+	+
SRmD311	<i>dgoA</i> ::pKNOCK-Gm	-	+	-

Growth is as follows: +, wild type; -, no growth on defined media with 15 mM glycerol and 0.5 mM galactose as carbon sources.

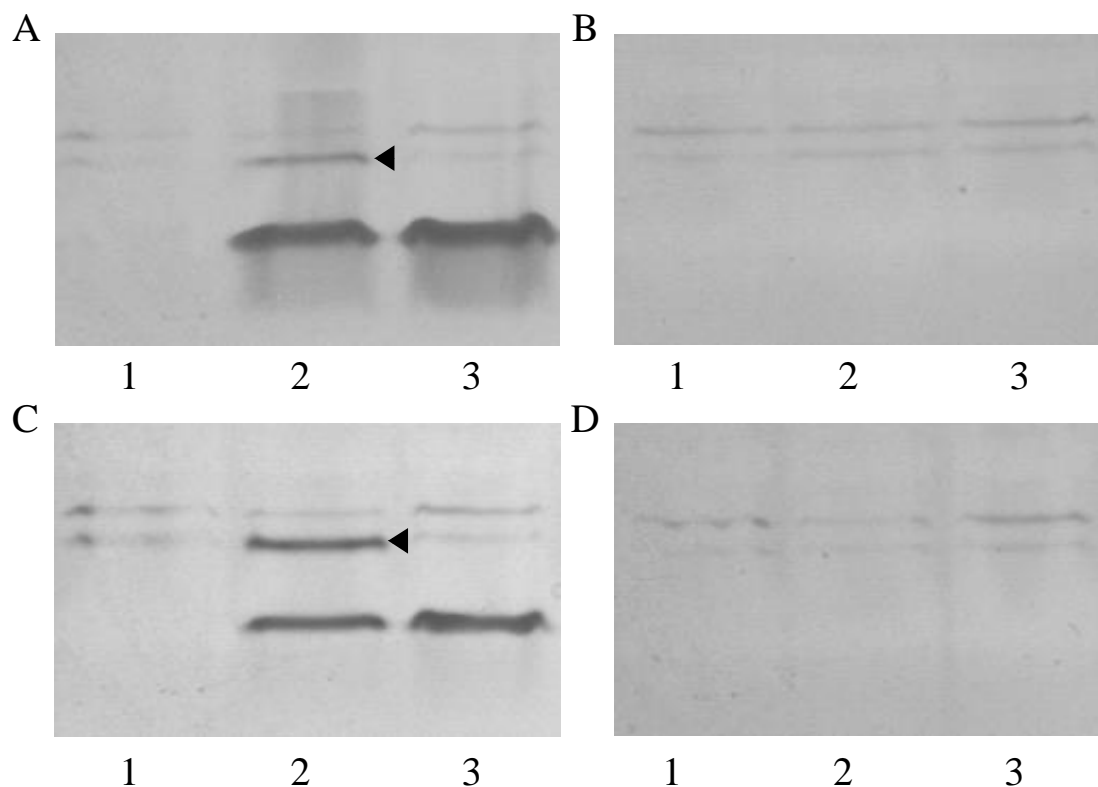
Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

### 5.3.3 *GalD* is a galactose dehydrogenase

Annotation suggested that *galD* encodes the galactose dehydrogenase used during the De Ley-Doudoroff pathway. However, genetic data suggested *galD* may not be necessary for galactose utilization. To provide evidence that *galD* encodes a galactose dehydrogenase we wished to demonstrate biochemical activity. Since both  $\text{NAD}^+$  and  $\text{NADP}^+$  dependent galactose dehydrogenase activities have previously been observed in *R. meliloti* L5-30 (Arias & Cerveñansky, 1986), it was reasoned that isozyme analysis might be more likely to provide unambiguous evidence. The *galD* mutant, SRmD298, and the wild type, Rm1021 were grown in defined medium with either glycerol, or glycerol and galactose as carbon sources. Cell free extracts were separated using non-denaturing PAGE and were stained for galactose dehydrogenase activity, both in the presence and absence of galactose, in assays that contained either  $\text{NAD}^+$  or  $\text{NADP}^+$ .

Faint bands of activities that were independent of growth condition, substrate, and either  $\text{NAD}^+$  or  $\text{NADP}^+$  were observed (Figure 5.2, panel B and panel D). Two prominent bands of activities appeared to be inducible by growth on a medium containing galactose (Figure 5.2). Whereas two distinct bands were present in the galactose induced wild-type lane, the upper band was absent from extracts of galactose-grown SRmD298, suggesting that this band corresponds to the activity encoded by *galD* (Figure 5.2). We note that this activity band was significantly more prominent in the presence of  $\text{NAD}^+$  than  $\text{NADP}^+$ ; suggesting that GalD may have a preference for  $\text{NAD}^+$ . The lower unidentified galactose-induced dehydrogenase activity band appeared to be more prominent in the presence of  $\text{NADP}^+$ .

**Figure 5.2.** Non denaturing PAGE of galactose-inducible enzyme activity



Extracts of Rm1021 grown on defined medium containing glycerol as a sole carbon source (Lane 1), Rm1021 grown on defined medium containing glycerol and galactose as sole carbon sources (Lane 2) and SRmD298 grown on defined medium containing glycerol and galactose as carbon sources (Lane 3) are shown in each panel. (A.) Stained in the presence of  $\text{NADP}^+$  and galactose. (B.) Stained in the presence of  $\text{NADP}^+$  and the absence of galactose. (C.) Stained in the presence of  $\text{NAD}^+$  and galactose. (D.) Stained in the presence of  $\text{NAD}^+$  in the absence of galactose. Arrow indicates the band predicted to correspond to GalD. Panels (B.) and (D.) also serve as loading controls for panels (A.) and (C.). Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

### 5.3.4 Identification of a galactose transporter

Candidate transporter genes were not identified by Tn5 mutagenesis. Therefore a more directed approach was taken towards identifying a galactose transporter. A number of ABC transporters had previously been shown to be induced by galactose or galactosides (Mauchline *et al.*, 2006); Specifically the ABC transporters that contain *SMa0203* (induced 19-fold by galactose), *SMb21343* (induced 6-fold by galactose), and *SMb20931* (induced by  $\alpha$  and  $\beta$ -galactosides, and 2-fold by galactose). Since mutations have been previously isolated in these regions (Pobigaylo *et al.*, 2006), these mutants were screened for the ability to use galactose as a sole carbon source on defined medium and found to grow as well as the wild-type Rm1021 (Table 5.4). The remaining genes that were identified to be induced by galactose and galactosides were  $\alpha$ -galactoside transport (*agpA*) and  $\beta$ -galactoside transport (*lacE*) genes (Mauchline *et al.*, 2006). Mutants of these genes have previously been shown to be able to utilize galactose as a sole carbon source so they were not tested further (Charles & Finan, 1991; Gage & Long, 1998).

A bioinformatic approach was also used to try to identify galactose transport genes. JGI-IMG Ortholog Neighborhood viewer was used to analyze the region surrounding homologues to *dgoK1* in related organisms (Markowitz *et al.*, 2010). We identified a set of three transport genes encoded downstream of a *dgoK1* homologue in *Brucella* sp.. These genes have been previously characterized in *Brucella suis* and shown to be necessary for galactose utilization (Alvarez-Martinez *et al.*, 2001). A BLASTP search of these transporter genes to the *S. meliloti* genome revealed that they shared 70-75% identity with *araABC* of *S. meliloti*.

**Table 5.4.** Growth phenotypes of mutants of galactose-induced transporters

Strain	Genotype	Carbon source	
		Glycerol	Galactose
Rm1021	Wild type	+	+
STM.4.09.D04	<i>Sma0203</i> ::Tn5, Nm <sup>R</sup>	+	+
STM.2.10.H05	<i>SMb21343</i> ::Tn5, Nm <sup>R</sup>	+	+
STM.3.09.B05	<i>SMb20931</i> ::Tn5, Nm <sup>R</sup>	+	+

Growth is as follows: +, wild type.

Previous characterization of *araABC* in *S. meliloti* demonstrated that they were necessary for arabinose but not galactose utilization (Poysti *et al.*, 2007). This was concluded on the basis of an *araA* mutant, SRmA503, being able to grow on solid defined medium with galactose. To further investigate whether *araABC* might play a role in galactose utilization the growth rate of an *araA* mutant in liquid defined medium was determined. We found that SRmA503 grew significantly more slowly than the wild type on defined medium containing galactose ( $p < 0.025$ ), whereas the growth rate on other carbon sources including glycerol, glucose, xylose, and fucose were not significantly different from Rm1021 (Table 5.5). It was hypothesised that AraABC plays a role in galactose transport in *S. meliloti*, however the transport of galactose into the cell is likely also facilitated by other transporters.

### ***5.3.5 Galactose and glucose compete with arabinose for transport***

It was previously shown that the transporter encoded by *araABC* was required for intracellular accumulation of  $^3\text{H}$  and  $^{14}\text{C}$  arabinose (Poysti *et al.*, 2007). To directly test the ability of AraABC to transport galactose, competition assays using unlabelled sugars to compete with radiolabelled arabinose for uptake were conducted. In the absence of a competing sugar we observed similar levels of  $^3\text{H}$  accumulation to those previously reported (Figure 5.3) (Poysti *et al.*, 2007). When the assay was repeated with a 15:1 ratio of unlabelled arabinose:  $^3\text{H}$  arabinose, radiolabel accumulation was reduced to background levels. An identical assay was carried out using unlabelled galactose, glucose, fucose, and xylose at a ratio of 15:1 to test whether any of these sugars could compete with arabinose for transport. We found that when galactose or glucose were



**Table 5.5.** Doubling time (h) of an *araA* mutant

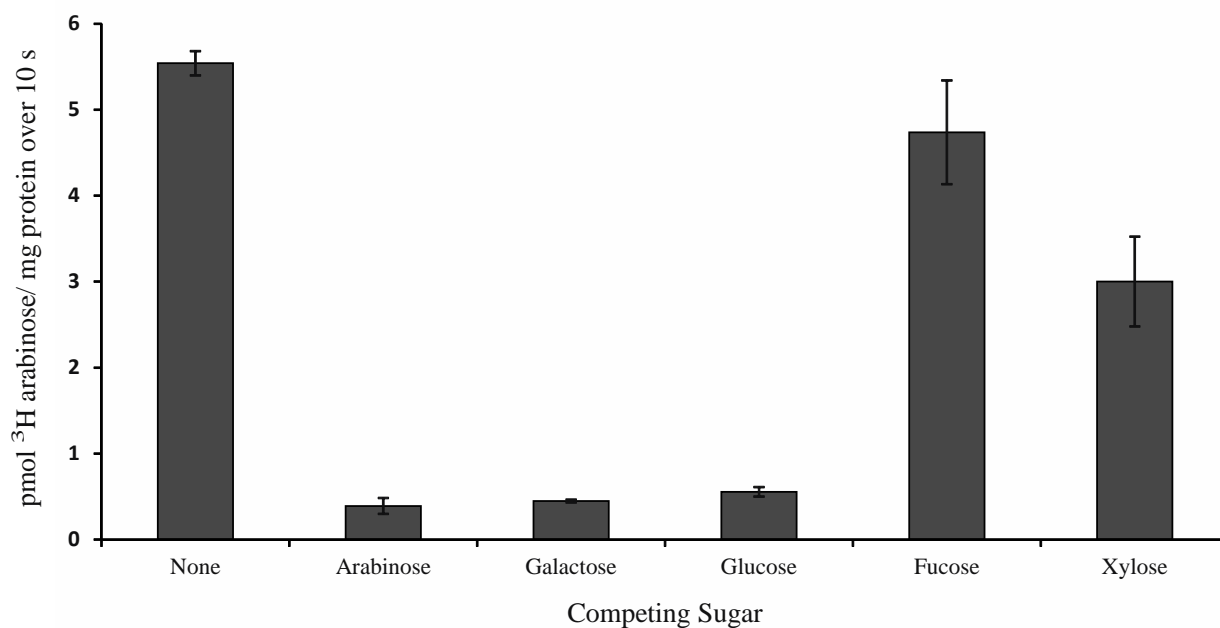
Carbon Source	Mean doubling time (h) $\pm$ SD		<i>P</i> value
	Rm1021 (wild type)	SRmA503 ( <i>araA</i> ::Tn5)	
Glycerol	7.6 $\pm$ 1.0	8.3 $\pm$ 0.3	<0.5
Galactose	6.8 $\pm$ 0.4	9.7 $\pm$ 1.1	<0.025
Glucose	8.0 $\pm$ 1.2	9.1 $\pm$ 0.4	<0.25
Arabinose	7.5 $\pm$ 0.4	28.9 $\pm$ 8.3	<0.05
Xylose	7.5 $\pm$ 0.6	8.7 $\pm$ 0.5	<0.1
Fucose	8.8 $\pm$ 1.4	9.3 $\pm$ 0.6	<0.5

Doubling times (h) were calculated from cultures growing in defined media over an 8 hour time interval. All cultures were in mid-log phase over the entire time interval. Numbers represent the mean  $\pm$  S. D. of 3 independent cultures.

The doubling time was calculated as  $\ln(2)/((\ln(N_2/N_1))/T)$ , where  $N_1$  is the initial OD<sub>600</sub>,  $N_2$  is the final OD<sub>600</sub>, and T is the time (in hours).

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

**Figure 5.3.** Competition for transport with [ $^3\text{H}$ ] arabinose in *S. meliloti* Rm1021



Accumulation of [ $^3\text{H}$ ] arabinose is expressed as pmol  $\text{mg}^{-1}$  over 10 seconds. Data are presented as the mean  $\pm$  standard deviation of three independent biological replicates. Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

added the accumulation of radiolabel was significantly reduced (Figure 5.3). Competition with fucose did not significantly reduce radiolabel accumulation, while competition with xylose moderately reduced radiolabel accumulation (Figure 5.3). Fructose, lyxose, ribose, mannose, and rhamnose were also tested for competition with  $^3\text{H}$  arabinose for uptake and showed no significant differences compared to uptake levels in the absence of competing sugar. The data is consistent with hypothesis that transport of glucose and galactose can be facilitated by AraABC.

### 5.3.6 Induction of the galactose locus

The locus that contains *araABC* had previously been shown to be inducible by arabinose, galactose, and seed exudate (Poysti *et al.*, 2007). Since galactose is in high concentrations in plant secretions (Knee *et al.*, 2001) we hypothesised that the galactose locus we identified would also be induced by these compounds. We used qRT-PCR to investigate the induction of the galactose locus, with primers to *galD* as a representative measure of gene expression of the locus Table 5.6. Induction was measured in the wild-type Rm1021 background and the galactose mutants SRmD144 and SRmD298. Effects on gene expression in SRmD298 and SRmD144 were consistent therefore only results for SRmD144 are shown.

Rm1021 grown in defined medium containing galactose showed induction of the galactose locus compared to cells grown in defined medium with glycerol. The induction of *galD* was modest (4-fold) compared to inductions that have been previously observed for other catabolic genes using qRT-PCR (Geddes *et al.*, 2010). Significant induction of *galD* by arabinose or seed exudate was not observed in Rm1021. The level of gene

**Table 5.6.** qRT-PCR analysis of galactose locus gene expression

Strain	Genotype	Carbon source	Fold <i>galD</i> expression <sup>b</sup>
Rm1021	Wild-type	Galactose <sup>b</sup>	3.9
Rm1021	Wild-type	Arabinose <sup>b</sup>	NC <sup>b</sup>
Rm1021	Wild-type	Seed exudate	NC
SRmD304	<i>dgoK1</i>	Glycerol	6.6
SRmD304	<i>dgoK1</i>	Galactose <sup>b</sup>	5.4
SRmD304	<i>dgoK1</i>	Arabinose <sup>b</sup>	4.9
SRmD304	<i>dgoK1</i>	Seed exudate	10.8

Data are expressed as  $2^{\Delta(\Delta C_t)}$  and represents fold expression over Rm1021 grown in glycerol. The experiment included *SMc00128* as an internal control (Krol & Becker, 2004). Results are from 3 independent biological replicates. Standard error between experiments was within 1 cycle threshold (2-fold change). NC indicates no significant change in expression level.

<sup>b</sup>Glycerol was included as a carbon source to support mutant growth.

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

expression in a *dgoK1* mutant grown on defined medium containing only glycerol was comparable to levels of *galD* gene expression found in Rm1021 grown on galactose (Table 5.6). This suggests that *galD* is constitutively expressed in a *dgoK1* background.

### ***5.3.7 Inability to catabolize galactose affects araABC expression***

Since the transport genes *araABC* have been shown to be induced by galactose (Poysti *et al.*, 2007), we were interested in determining if the transcription of *araABC* was also affected in the *dgoK* background. In the wild-type background, strong induction of *araA* was observed in cells grown on arabinose, and more moderate induction was observed in cells grown on galactose or seed exudate (Table 5.7). These results are consistent with induction of this locus that has been previously reported (Poysti *et al.*, 2007). In a *dgoK1* background we found that, while we did not observe constitutive expression of *araA*, we did observe an increased induction of the arabinose locus by galactose and seed exudate (Table 5.7).

### ***5.3.8 SRmD144 is more competitive than Rm1021 for nodule occupancy***

Our interest in galactose catabolism stemmed from the hypothesis that the catabolism of sugars that are present in high concentrations in the rhizosphere and in plant cell walls may be important for competition for nodule occupancy. To determine if galactose catabolism was important for competition for nodule occupancy SRmD144 and Rm1021 were competed and scored for nodule occupancy.

Initial competition experiments contained approximately equal proportions of the wild type and the mutant. The results of these experiments suggested that a greater

**Table 5.7.** qRT-PCR analysis of *araABC* gene expression

Strain	Genotype	Carbon source	Fold <i>araA</i> expression <sup>a</sup>
Rm1021	Wild type	Galactose <sup>c</sup>	5.4
Rm1021	Wild type	Arabinose <sup>c</sup>	12.8 <sup>b</sup>
Rm1021	Wild type	Seed exudate	7.6
SRmD304	<i>dgoK1</i>	Glycerol	NC
SRmD304	<i>dgoK1</i>	Galactose <sup>c</sup>	26.5
SRmD304	<i>dgoK1</i>	Arabinose <sup>c</sup>	15.0 <sup>b</sup>
SRmD304	<i>dgoK1</i>	Seed exudate	27.7 <sup>b</sup>

<sup>a</sup>Data are expressed as  $2^{\Delta(\Delta C_t)}$  and represents fold expression over Rm1021 grown in glycerol. The experiment included *SMc00128* as an internal control (Krol & Becker, 2004). Results are from 3 independent biological replicates. Standard error between experiments was within 1 cycle threshold (2-fold change). NC indicates no significant change in expression level.

<sup>b</sup>Standard error between 3 experiments was within 1.5 cycle threshold.

<sup>c</sup>Glycerol was included as a carbon source to support mutant growth

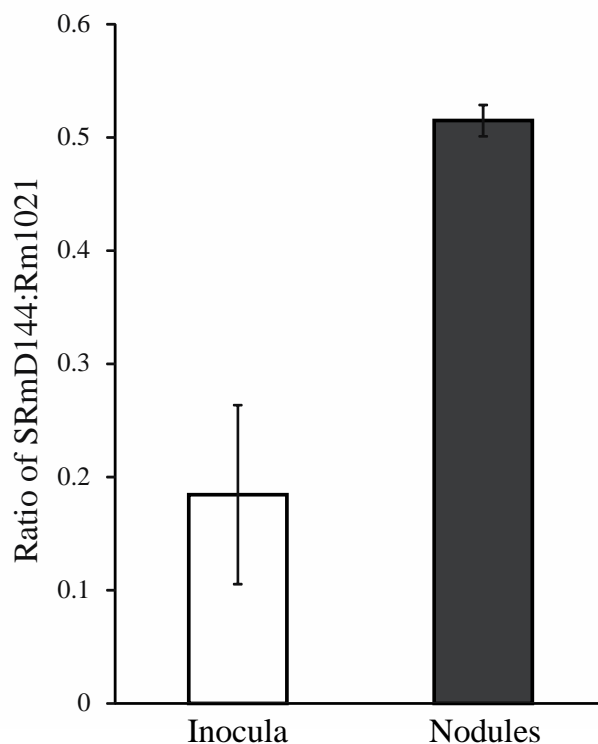
Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

proportion of nodules contained SRmD144 than Rm1021. To verify this, the experiment was repeated using wild-type cells in significantly higher concentrations in the inoculum (approximately 5:1). The results are presented in Figure 5.4, and clearly show that SRmD144 is significantly more competitive for nodule occupancy than the wild-type Rm1021.

### 5.3.9 *araA* does not complement an *Agrobacterium chvE* mutant for virulence

The *araA* homologue in *A. tumefaciens*, *chvE*, encodes a sugar binding protein that plays a role in the control of virulence gene expression in response to sugar binding during infection of the plant (He *et al.*, 2009). It was hypothesised that AraA may be playing a similar role in *S. meliloti*, and since the induction of *araA* by seed exudate was drastically increased in SRmD144, it may be involved in the increased competition phenotype. We reasoned that if *araA* was playing a similar role in *S. meliloti* it may be able to complement an *Agrobacterium chvE* mutant for virulence gene induction. To test this we used the *A. tumefaciens virE* reporter strains AT11043 (*virE358::Tn3-HoHo1*) and AT11055 (*virE358::Tn3-HoHo1*, *chvE::Tn5*). Induction of *virE* has previously been measured by  $\beta$ -galactosidase assay in these strains and shown to be dependent on the presence of *chvE* (Charles *et al.*, 1994). The plasmids pBG57, pBG65 and pZW3 were conjugated into AT11055, and compared for their ability to complement *virE* induction (Table 5.8). We observed  $\beta$ -galactosidase levels indicative of *virE* induction similar to those previously observed in AT11043 (Charles *et al.*, 1994). Induction of *virE* in the *chvE* mutant, AT11055, was not achieved under these conditions. When the transconjugant containing a copy of *chvE* on pBG65 was assayed, expression

**Figure 5.4.** Competition for nodule occupancy



SRmD144 was competed against the wild-type Rm1021. The data are presented as the mean  $\pm$  standard deviation of the proportion of SRmD144 found in the inoculum (light grey bar), and the proportion of SRmD144 recovered from the nodules (dark grey bar). Recovery ratios were significantly greater than inoculum ratios ( $p < 0.01$ ). Data represents three independent experiments each comprised of 10 plants. Approximately 100 nodules were assayed for each replicate.

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology



**Table 5.8.** Complementation of *A. tumefaciens chvE*

Strain	Genotype	Plasmid	Plasmid genotype	Mean $\beta$ -galactosidase activity (Miller units) $\pm$ SD
AT11043	<i>virE</i> :: Tn3-Hoho1	None	---	495 $\pm$ 56
AT11055	<i>virE</i> :: Tn3-Hoho1, <i>chvE</i> ::Tn5	None	---	30 $\pm$ 7
AT11055	<i>virE</i> :: Tn3-Hoho1, <i>chvE</i> ::Tn5	pZW3	SM <i>ara</i> cosmid; <i>araB</i> ::Tn5	31 $\pm$ 10
AT11055	<i>virE</i> :: Tn3-Hoho1, <i>chvE</i> ::Tn5	pBG57	SM <i>araA</i>	50 $\pm$ 9
AT11055	<i>virE</i> :: Tn3-Hoho1, <i>chvE</i> ::Tn5	pBG65	AT <i>chvE</i> 3KB <i>EcoRI</i> fragment	311 $\pm$ 7

<sup>a</sup>Data are presented in Miller units. Values are presented as the mean  $\pm$  S. D, n=3.

Reproduced from the Journal of Bacteriology, volume 194, Geddes, B. A., and I. J. Oresnik, Inability to catabolize galactose leads to an increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*, pp. 5044-5053, Copyright 2012, with permission from the American Society for Microbiology

of *virE* was restored. Neither construct containing the heterologous *S. meliloti* homologue *araA* (pBG57 or pZW3) was able to complement AT11055 for *virE* induction. Our results suggest that *araA* of *S. meliloti* cannot complement a *chvE* mutant of *A. tumefaciens* for virulence gene induction.

## 5.4 Discussion

In this work we have identified a locus required for galactose catabolism. To our knowledge this is the first report identifying the genes necessary for galactose catabolism in *S. meliloti*. It has been shown that galactose catabolism in *S. meliloti* proceeds via the De Ley-Doudoroff pathway rather than the more highly conserved Leloir pathway (Arias & Cerveñansky, 1986). The De Ley-Doudoroff pathway is carried out by a galactose dehydrogenase, lactonase, D-galactonate dehydratase, 2-keto-3-deoxy galactonokinase, and a 2-keto-3-deoxy-6-phosphogalactonate aldolase. Sequence analysis suggested that the locus we identified had the capacity to encode each of these enzymes using *galD*, *SMc00883*, *ilvD1*, *dgoK1*, and *dgoA* respectively (Figure 5.1).

Despite the close correlation between annotation of genes in the locus and the enzymes required for the De Ley-Doudoroff pathway, *SMc00883* and *ilvD1* were dispensable for galactose catabolism (Table 5.2). Previously enzymatic activities in the De Ley-Doudoroff pathway were shown to be inducible by L-arabinose as well as galactose (Arias & Cerveñansky, 1986). It is of note that the first three steps of galactose and arabinose catabolism in *S. meliloti* proceed through intermediates with similar stereochemistry (Arias & Cerveñansky, 1986; Duncan & Fraenkel, 1979). It was suggested that the dehydrogenase, lactonase, and dehydratase activities that were

observed could be redundant (Arias & Cerveñansky, 1986). We have shown that the galactose locus is not inducible by arabinose (Table 5.6). Taken together with the ability of mutants in *SMc00883* and *ilvD1* to grow using galactose as a sole carbon source our data are consistent with the suggestion that these activities are redundant in *S. meliloti*. Since the redundant activities were inducible by L-arabinose, and the previously characterized L-arabinose locus of *S. meliloti* contains a dehydratase (*araF*) (Poysti *et al.*, 2007) that shares 65 percent identity at the amino acid sequence level with *ilvD1*, we constructed the double mutant SRmD315 (*araF*::Tn5-B20, *ilvD1*::pKNOCK-Gm). SRmD315 was still able to utilize galactose as a sole carbon source. Characterization of the arabinose locus showed that genes encoding arabinose dehydrogenase and lactonase activity were found elsewhere in the genome (Poysti *et al.*, 2007). It is possible that these are also functionally redundant with genes in the galactose locus.

Although genetic evidence suggested that *galD* may not be necessary for galactose catabolism, biochemical evidence clearly shows that *galD* encodes a NAD<sup>+</sup>/NADP<sup>+</sup> galactose dehydrogenase (Figure 5.2). These findings are consistent with the previous work that showed galactose induced both NAD<sup>+</sup> and NADP<sup>+</sup> dependant galactose dehydrogenase activities (Arias & Cerveñansky, 1986). Additionally, a second non-specific dehydrogenase was also observed with activity on galactose. It is possible that *galD* encodes the biologically relevant galactose dehydrogenase, but we note that galactose dehydrogenase activity is also redundant with a non-specific dehydrogenase (Figure 5.2).

Despite extensive random mutagenesis, and directed mutations of ABC transporters that were shown to be inducible by galactose (Mauchline *et al.*, 2006), we

were unable to isolate a transporter mutant that was unable to utilize galactose as a sole carbon source. However, bioinformatic evidence led us to investigate the arabinose transporter AraABC for its ability to transport galactose. SRmA503 (*araA::Tn5-B20*) was shown to grow slower using galactose as a sole carbon source than the wild-type strain (Table 5.5). Furthermore, transport competition experiments show that galactose as well as glucose compete directly with arabinose for transport. The pentose xylose was also able to partially compete with arabinose for transport (Figure 5.3). The ability of AraABC to transport multiple sugars is consistent with phenotypes observed for homologous transporters in *Brucella suis* and *Agrobacterium tumefaciens*. In *B. suis* mutants of the *araABC* homologues *chvE-gguAB* were impaired for growth using glucose, galactose, arabinose, xylose, fucose and mannose as sole carbon sources (Alvarez-Martinez *et al.*, 2001). Homologues in *A. tumefaciens* (*chvE-mmsAB*) have recently been shown to play a role in glucose, galactose, arabinose, fucose, and xylose transport (Zhao & Binns, 2011). It is important to note that although AraABC may be able to facilitate galactose and glucose transport, it is likely that *S. meliloti* encodes other transporters for these sugars as it has been shown that uptake of radiolabelled glucose is not reduced in SRmA503 (Poysti *et al.*, 2007).

We have shown here that a mutant of galactose catabolism, SRmD144, is more competitive for nodule occupancy than the wild-type strain (Figure 5.4). This phenotype is in distinct contrast to previously investigated catabolic mutants with regards to symbiotic proficiency. The utilization of several carbon sources including glycerol, erythritol, rhamnose, inositol, and proline has been shown to be an important trait for competition of nodule occupancy in *R. leguminosarum* and *S. meliloti* (Ding *et al.*, 2012;

Fry *et al.*, 2001; Jiménez-Zurdo *et al.*, 1995; Kohler *et al.*, 2010; Oresnik *et al.*, 1998; Yost *et al.*, 2006). To our knowledge, this is the first catabolic mutant that has been shown to have an increased ability to compete for nodule occupancy. It is noteworthy that regulation in SRmD144 is unusual, and regulatory phenotypes including constitutive expression of the galactose locus and enhanced induction of the arabinose locus were observed (Table 5.6, Table 5.7). Therefore, it is possible that the symbiotic proficiency phenotype may be a result of regulatory effects on other loci that are involved in symbiotic competition rather than the inability to catabolise galactose. In addition, competition phenotypes have not been conserved between Rhizobia (Geddes *et al.*, 2010; Yost *et al.*, 2006), therefore it would be interesting to investigate whether mutants of galactose catabolism in other rhizobia would result in a similar increased competition phenotype.

## **Chapter 6**

### **Exopolysaccharide production in response to decreased pH is correlated with an increase in competition for nodule occupancy**

This chapter has been submitted for publication. Exopolysaccharide production in response to decreased pH is correlated with an increase in competition for nodule occupancy. Barney A. Geddes, Juan E. González and Ivan J. Oresnik. Technical assistance was provided by Rachel Wong.

## 6.1 Introduction

Collectively, the Rhizobium-legume symbiosis contributes over 40 million tons of nitrogen per year into agricultural systems (Herridge *et al.*, 2008). Enhancing the effectiveness of rhizobial inoculums is desirable as it helps alleviate the requirement for synthetic nitrogen fertilizers that are economically and environmentally costly. One factor that limits the effectiveness of rhizobial inoculums is the inability to compete with indigenous rhizobia in the soil for nodule occupancy of legume hosts (Triplett & Sadowsky, 1992).

*Sinorhizobium meliloti* engages in nitrogen fixing symbiosis with the agriculturally important forage *Medicago sativa* (alfalfa), and the model legume *Medicago truncatula*. Symbiosis between *S. meliloti* and its legume hosts begins with a signal exchange in the rhizosphere. *S. meliloti* produces Nod Factor (NF) in response to flavonoids secreted by the legume (Spaink, 2000). NF recognition by the legume host activates a complex signaling pathway that leads to root hair curling and nodule development. Root hair curling traps *S. meliloti* that have colonized the root hair within an apoplastic compartment called the curled colonized root hair (CCRH) (Oldroyd *et al.*, 2011). Intracellular invasion of the root hair, and subsequent cell layers, proceeds by the formation of a plant-derived infection thread (IT) (Oldroyd *et al.*, 2011). Rhizobia are eventually endocytosed by inner cortical cells and acquire a plant-derived symbiotic membrane before differentiating into nitrogen fixing bacteroids (Oldroyd *et al.*, 2011). Successful penetration of the IT requires both the continued biosynthesis of NF, and the synthesis of symbiotic exopolysaccharide (EPS) (Jones *et al.*, 2007).

*S. meliloti* produces two symbiotic EPSs: succinoglycan and EPSII.

Succinoglycan is a polymer of octosaccharide repeating units, composed of seven glucose and one galactose residue, and decorated with succinyl, pyruvyl and acetyl modifications (Reinhold *et al.*, 1994). EPSII is a polymer of a disaccharide repeating unit composed of an acetylated glucose and pyruvylated galactose residue (Her *et al.*, 1990). The ability to synthesize one of these EPSs is required for invasion of *M. sativa* by *S. meliloti* (Cheng & Walker, 1998b; Glazebrook & Walker, 1989; Leigh *et al.*, 1985). Succinoglycan and EPSII are both synthesized in high-molecular-weight (HMW) and low-molecular-weight (LMW) fractions. The LMW fractions of both succinoglycan and EPSII have been implicated as being particularly important for symbiosis (Battisti *et al.*, 1992; González *et al.*, 1996). Despite the extensive study of the genetics and regulation of symbiotic EPSs, their role in symbiosis remains unclear. Studies have suggested that they may play a role in dampening the legume immune response during invasion (Jones *et al.*, 2008), or protection from reactive oxygen species encountered during invasion (Lehman & Long, 2013).

The regulation of EPS synthesis is influenced by a number of environmental factors including nitrogen and phosphate limitation, metal ions, salt, osmolarity and pH (Breedveld *et al.*, 1990; Dilworth *et al.*, 1999; Hellweg *et al.*, 2009; Leigh *et al.*, 1985; Mendrygal & González, 2000; Miller-Williams *et al.*, 2006; Rossbach *et al.*, 2008). The regulation of succinoglycan and EPSII production is controlled by several regulatory systems in *S. meliloti* including ExoS/ChvI, SyrA/SyrM, and SinI/ExpR. (Cheng & Walker, 1998a; Glenn & Dilworth, 1981; Marketon *et al.*, 2003; Mulligan & Long, 1989). The ExoS/ChvI two-component system activates succinoglycan biosynthesis and



suppresses flagella biosynthesis in response to an unknown signal (Yao *et al.*, 2004). The system is also inhibited by physical interactions of a periplasmic protein, ExoR, with the sensor kinase ExoS (Chen *et al.*, 2008). SinI/ExpR is a quorum-sensing system that regulates EPSII biosynthesis (Marketon *et al.*, 2003). The wild type *S. meliloti* Rm1021 contains a natural insertion element disrupting *expR* and, as a result, is unable to synthesize EPSII under normal conditions (Pellock *et al.*, 2002).

In other rhizobia, EPSs are not essential for symbiosis, but rather have been shown to provide a competitive advantage for nodule occupancy (Triplett & Sadowsky, 1992). The ability to metabolize many different carbon sources has also been shown to influence competition for nodule occupancy in *S. meliloti* and *R. leguminosarum* (Ding *et al.*, 2012; Fry *et al.*, 2001; Jensen *et al.*, 2005; Jiménez-Zurdo *et al.*, 1995; Kohler *et al.*, 2010; Oresnik *et al.*, 1998; Vanderlinde *et al.*, 2013; Yost *et al.*, 2006). During the characterization of galactose catabolism in *S. meliloti*, we recently showed that a mutant that was unable to catabolize galactose had an increased ability to compete for nodule occupancy compared to the wild type (Geddes & Oresnik, 2012b). This phenotype was in contrast to the reduced ability to compete for nodule occupancy demonstrated in other carbon metabolism mutants (Ding *et al.*, 2012; Fry *et al.*, 2001; Jiménez-Zurdo *et al.*, 1995; Kohler *et al.*, 2010; Oresnik *et al.*, 1998; Vanderlinde *et al.*, 2013; Yost *et al.*, 2006).

In this work we describe further characterization of a galactose mutant of *S. meliloti* in pursuit of an explanation for increased competitiveness. Our data is consistent with the hypothesis that early succinoglycan production mediates increased competitiveness for nodule occupancy in galactose mutants of *S. meliloti*.

## 6.2 Materials and Methods

### 6.2.1 Bacterial strains, plasmid and media

Bacterial strains and plasmids used and generated in this work are listed in Table 6.1. *S. meliloti* strains were grown routinely at 30°C on either Luria-Bertani (LB) or Tryptone Yeast Extract (TY) as a complex medium (Sambrook *et al.*, 1989). Vincent's minimal medium (VMM) (Vincent, 1970), Rhizobium minimal medium (RMM) (Broughton *et al.*, 1986), Yeast extract mannitol (YEM) (Vincent, 1970) and ½ GYM (Glutamate yeast extract mannitol) (Dylan *et al.*, 1990a) were used as defined media. Carbon sources were filter sterilized and added to defined media to a final concentration of 15 mM when relevant. 0.02% Calcofluor was added to media to visualize succinoglycan production. Buffering of the media was accomplished using 50 mM HEPES when relevant. When required, *S. meliloti* and *Escherichia coli* strains were grown in the following concentrations of antibiotic (µg/mL): chloramphenicol (Cm) 20; gentamicin (Gm) 20 or 60; kanamycin (Km) 20; neomycin (Nm) 200; streptomycin (Sm) 200; tetracycline (Tc) 5. All antibiotics were filter sterilized before use.

### 6.2.2 DNA and genetic manipulations

Conjugations and transductions were carried out essentially as previously described (Finan *et al.*, 1988; Finan *et al.*, 1984). Standard techniques were used for plasmid isolation, restriction enzyme digests, ligations, transformations, and agarose gel electrophoresis (Sambrook *et al.*, 1989). To construct SRmD434, a 400-bp internal fragment of *SMc00588* was amplified, restricted and ligated into the suicide vector

**Table 6.1.** Bacterial strains and plasmids

Strain or plasmid	Genotype or phenotype	Reference or source
<b>Strain</b>		
<i>S. meliloti</i>		
Rm1021	SU47 <i>str-21</i> , Sm <sup>R</sup>	(Meade <i>et al.</i> , 1982)
Rm7055	Rm1021 <i>exoY</i> ::Tn5, Nm <sup>R</sup>	(Reuber & Walker, 1993)
SRmA363	Rm1021 <i>expR</i> <sup>+</sup>	(Miller-Williams <i>et al.</i> , 2006)
SRmD144	Rm1021 <i>dgoK1</i> ::Tn5, Nm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD288	Rm1021 <i>dgoK1</i> ::Tn5-233, Gm <sup>R</sup> , Sp <sup>R</sup>	This work
SRmD298	Rm1021 <i>galD</i> ::Tn5-B20, Nm <sup>R</sup>	(Geddes & Oresnik, 2012b)
SRmD304	<i>dgoK1</i> ::Tn5 Φ (SRmD304) → Rm1021, Nm <sup>R</sup>	This work
SRmD305	<i>dgoK1</i> ::Tn5 Φ (SRmD304) → SRmA363, Nm <sup>R</sup>	This work
SRmD394	<i>exoY</i> ::Tn5 Φ (Rm7055) → SRmA363	This work
SRmD413	<i>dgoK1</i> ::Tn5-233 Φ (SRmD288) → SRmD394, Nm <sup>R</sup> , Gm <sup>R</sup> , Sp <sup>R</sup>	This work
SRmD434	<i>SMc00588</i> ::pKNOCK-Gm, Gm <sup>R</sup>	This work
SRmD441	<i>galD</i> ::Tn5-B20 Φ (SRmD298) → Rm1021, Nm <sup>R</sup>	This work
SRmD436	<i>SMc00588</i> ::pKNOCK-Gm Φ (SRmD434) → SRmD441	This work
<i>E. coli</i>		
DH5α	λ Φ80 <i>dlacZ</i> <sup>°</sup> M15 <sup>°</sup> ( <i>lacZYA</i> - <i>argF</i> ) <i>U169 recA1 endA1 hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>-</sup> ) <i>supE44 thi-1 gyrA relA1</i>	(Hanahan, 1983)
DH5α λpir	λpir lysogen of DH5α	(House <i>et al.</i> , 2004)
MM294A	<i>pro-82-thi-1 hsdR17 supE44</i>	(Finan <i>et al.</i> , 1986)
MT607	MM294A <i>recA56</i>	(Finan <i>et al.</i> , 1986)
MT616	MT607 (pRK600)	(Finan <i>et al.</i> , 1986)
<b>Plasmid</b>		
pAA10	Galactose complementing cosmid, Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)
pBG56	pCO36 <i>dgoK1</i> , Tc <sup>R</sup>	(Geddes & Oresnik, 2012b)
pCO37	pRK7813 containing <i>attB</i> sites, Gateway® compatible destination vector	(Jacob <i>et al.</i> , 2008)
pHC60	pHC41, S65T <i>gfp</i>	(Cheng & Walker, 1998a)
pPH1JI	IncP plasmid, Gm <sup>R</sup>	(Beringer <i>et al.</i> , 1978)
pKNOCK-Gm	Suicide vector, Gm <sup>R</sup>	(Alexeyev, 1999)
pRK600	pRK2013 <i>nptI</i> ::Tn9, Cm <sup>R</sup>	(Finan <i>et al.</i> , 1986)
pRK607	pRK2013::Tn5-233	(De Vos <i>et al.</i> , 1986)
pRK7813	Broad-host-range vector, Tc <sup>R</sup>	(Jones & Gutterson, 1987)
pRW2	<i>gal</i> fragment in pKNOCK-Gm, Gm <sup>R</sup>	This work

pKNOCK-Gm, creating pRW2. The plasmid was conjugated into *S. meliloti* Rm1021 and single-crossover recombination events were selected for using Sm and Gm as previously described (Richardson *et al.*, 2004).

### ***6.2.3 Non-denaturing gel electrophoresis and dehydrogenase assays***

Cells were grown in VMM medium with galactose or galactose and glycerol as sole carbon sources. Lysates were prepared, and separated using non-denaturing polyacrylamide gel electrophoresis (PAGE) as previously described (Pickering & Oresnik, 2008). Gels were developed using *p*-nitroblue tetrazolium based dehydrogenase stain according to Latner & Skillen, 1968. Samples were incubated either with or without NAD<sup>+</sup>/NADP<sup>+</sup>, and galactose as previously described to determine the specificity of activity bands (Geddes & Oresnik, 2012b).

### ***6.2.4 Quantitative analysis of exopolysaccharide production***

EPS quantification was carried out as previously described with some modification (Marroquí *et al.*, 2001; Yurgel *et al.*, 2013). *S. meliloti* strains were grown in RMM with galactose at 30<sup>0</sup>C with shaking and EPS was quantified at various time-points. Culture density (OD<sub>600</sub>) and pH were measured as necessary. Bacterial cultures grown in RMM were inoculated with a 1/4,000 dilution of overnight cultures grown in TY to an OD<sub>600</sub> of 1.0. To quantify EPS, 50-mL aliquots of culture were harvested by centrifugation at 4°C at 13,000 x g for 15 minutes. The supernatants were mixed with 2.5 mM NaCl and 2.5 volumes of 100% ethanol and incubated at 4°C overnight. Precipitated polysaccharides were removed from the ethanol the next day by spooling with a glass

rod. Spooled EPS was transferred to a pre-weighed petri plate and dried overnight at 4°C. To control for possible pH effects on EPS precipitation, culture supernatants were adjusted to pH 7; no differences in EPS precipitation were observed.

#### **6.2.5 Exopolysaccharide composition analysis**

Characterization of the carbohydrate composition of EPSs was done as according to Marketon *et al.* (2003) with some modifications. Rm1021 and SRmD304 were grown in RMM with galactose as described above. EPS was isolated from SRmD304 after 44 hours of growth, and both SRmD304 and Rm1021 after 6 days of growth. Briefly, 1 L of cells was harvested by centrifugation, and EPS was precipitated and concentrated by lyophilization. EPS was further purified by dialysis using a 1,000-molecular-weight-cutoff membrane (Spectrum, [www.spectrumlabs.com](http://www.spectrumlabs.com)) against water for 3 days, with water changed twice a day. A fraction of the dialyzed exopolysaccharide solution was mixed with an equal volume of 4 M trifluoroacetic acid and hydrolyzed overnight at 100°C in a sealed glass vial. Analyses by high-pressure anion-exchange chromatography were performed on a Dionex DX 500 HPLC system with a CarboPac MA1 column using a pulsed amperometric detector with a gold working electrode. The samples were eluted isocratically using a 3-step protocol as follows: 1) 500 mM NaOH for 20 min. 2) A linear gradient from 500 mM NaOH to 1 M Na-Acetate and 500 mM NaOH over 20 min. 3) 1 M Na-Acetate, 500 mM NaOH for 20 min. This profile was optimized to elute monosaccharides such as glucose and galactose, as well as acidic sugars such as glucuronic acid. The carbohydrate composition of the exopolysaccharides was calculated as a glucose:galactose ratio based on the elution times of known glucose and galactose

standards.

#### **6.2.6 Exopolysaccharide molecular weight distribution**

Analysis of the molecular weight profile of EPS synthesized in Rm1021 and SRmD304 was performed as previously described with modifications (Glenn *et al.*, 2007). Rm1021 and SRmD304 were grown overnight in TY to an OD<sub>600</sub> of 1.0, and subcultured at a 1/200 dilution into 25 mL of RMM with galactose. Strains were then grown overnight at 30°C with constant shaking. The next day, the pH of cultures was periodically monitored, and cells were harvested at equivalent time-points, when they reached pHs of approximately 5.75 and 6.5 for SRmD304 and Rm1021, respectively. Cells were pelleted by centrifugation at 13,000 x g for 15 min. Cell pellets were washed with RMM salts adjusted to the pH of the culture medium. 1 mL cell suspensions were incubated with 6 µCi of D-[U- <sup>14</sup>C] glucose for 4 h with constant shaking at 30°C. Samples were then boiled at 100°C for 5 min and placed on ice for 15 min. The supernatant was recovered by centrifugation at 15,000 x g for 15 min., and dialyzed against 500 mL of water using a 1,000 molecular-weight-cutoff Pur-A-Lyzer dialysis kit (Sigma, <http://www.sigmaaldrich.com>). Dialysis was performed over 3 days, with water changed every 12 hours. A 300-µL volume of succinoglycan-containing supernatant was fractionated by column chromatography (1 cm by 100 cm) using Bio-Gel P-6 (Bio-Rad, <http://www.bio-rad.com>). Pre-equilibration and elution was performed using a pyridinium acetate buffer (0.1 M; pH 5.0) (González *et al.*, 1998). Fractions were collected in 1-mL aliquots, and radioactivity was detected by liquid scintillation with a Beckman LS6500 scintillation counter. Fractions were routinely corrected for

background D-[U-  $^{14}\text{C}$ ] glucose monomers. Determination of fractions containing HMW and LMW succinoglycan were based on previously reported distributions in *S. meliloti* using identical column chromatography conditions (Glenn *et al.*, 2007).

#### **6.2.7 RNA isolation, cDNA synthesis and quantitative RT-PCR**

For RNA isolation, bacterial cultures were grown overnight in TY to an OD<sub>600</sub> of 1. Rm1021 and SRmD304 were subcultured into 20 mL cultures of RMM with galactose at a dilution of 1/200. Cultures were incubated overnight with shaking at 30°C. The next day, the pH of cultures was monitored periodically, and cells were harvested at the same time-point, at pHs of 5.75 and 6.5 for SRmD304 and Rm1021, respectively. Cells were harvested by centrifugation and cell pellets were frozen immediately in liquid nitrogen. RNA was then isolated using the Qiagen RNeasy kit as previously described (Geddes *et al.*, 2010). Approximately 200 ng of cDNA sample was used as a template for quantitative RT-PCR. Reactions were performed using the SYBR green RT-PCR kit from Invitrogen as recommended by manufactures. The RT-PCR reaction was performed using a Cepheid Smart Cycler with the following program: Stage 1, 95°C for 120s, 1 time; Stage 2, 95°C for 15s followed by 60°C for 30s, repeated 40 times; Stage 3, melting curve analysis of PCR products. The RT-PCR primers used to measure exopolysaccharide and flagella gene expression have been previously described (Glenn *et al.*, 2007; Hoang *et al.*, 2008). Other RT-PCR primers used are as follows: *lpsB* foreword: CATACCGGATGGAGCAAGTT; *lpsB* reverse GAAGAGGTCGGTCCCTTTCT; *ndvB* foreword: ACAACACCTCCATCGCACAG;

*ndvB* reverse: ATCGATCTTGCTGACGCTTT; *nodA* foreword:

ACCACCAGGAGCTCTCAGAA; *nodA* reverse: GCGTATAAGCCCAGTTCAGC.

### **6.2.8 Fluorescent microscopy**

Plants were grown for fluorescent microscopy essentially as previously described with some modifications (Cheng & Walker, 1998b; Gage *et al.*, 1996). Briefly, alfalfa seeds were surface sterilized and germinated on 1.5% agar for 72 hours in the dark. Seedlings were placed on microscope slides covered in 2 mL of Jensen's agar and covered with dialysis membrane (14,000 molecular weight cutoff). Seedlings were inoculated with 100  $\mu$ L of *S. meliloti* strains grown in LB medium overnight to an OD<sub>600</sub> of approximately 0.5 and diluted 1/100 in sterile water. Slides with seedlings were incubated in 50 mL falcon tubes with 20 mL of liquid Jensen's medium, loosely covered with plastic cling wrap.

Alfalfa roots were examined by fluorescent microscopy using a Zeiss Axio Imager.Z1 equipped with an AxioCamMR digital camera. Rhizobium strains expressing GFP were visualized using Zeiss filter set 38 (Endow GFP). Lysotracker Red DND-99 fluorescence was observed using Zeiss filter set 20 (Rhodamine). Composite images were generated using Adobe Photoshop CS5.

Visualization of curled colonized root hairs and infection threads was performed 7 days post-inoculation. Slides with alfalfa plants were incubated overnight with 8  $\mu$ M Lysotracker Red DND-99 in liquid Jensen's medium. Destaining was performed by incubating in liquid Jensen's medium for 30 minutes.



### **6.2.9 Competition for nodule occupancy assays**

Competition for nodule occupancy was assessed as previously described (Geddes *et al.*, 2010). Briefly, alfalfa seeds were surface sterilized and germinated for 48 hours in the dark on 1.5% agar. Seedlings were then planted in nitrogen-free media consisting of Leonard jars assemblies containing 10 alfalfa plants. Seedlings were inoculated after 2 days growth. The inocula consisted of 10 mL/ jar of an approximately 5:1 ratio of SRmD304:Rm1021, and 5:1 and 1:1 ratios of SRmD394:SRmD413, grown overnight in LB to equivalent optical densities, and diluted 100-fold in sterile water. Serial dilutions of the mixed inoculum spread onto LB agar plates and at least 50 colonies were patched onto the appropriate antibiotics to differentiate the proportion of each strain in the inoculum. Nodule occupancy was assessed after 30 days. At least 50 root nodules from each experiment were sterilized and crushed into 100  $\mu$ L sterile distilled water. Aliquots of each nodule extract were spotted onto LB plates containing antibiotics for strain differentiation. Competition was assessed by comparing the proportion of the mutant strain found in the inoculum to the proportion that was isolated from nodules. Significance was evaluated using a Students paired T-test, assuming a  $P < 0.05$  indicated a significant difference in competitiveness.

### **6.2.10 Nodule kinetics assays**

Alfalfa seed sterilization and germination were performed as described above. Alfalfa seedlings were planted in test tube slants containing 10 mL of Jensen's agar. After 2 days of growth, slants were inoculated with 100  $\mu$ L of rhizobium strains grown as described above and diluted in sterile water. Nodule kinetics were assessed by counting

the number of nodules formed on each plant daily over a 14-day period. Collectively over 100 plants over 5 independent replicates were monitored for nodule kinetics.

## 6.3 Results

### 6.3.1 Altered exopolysaccharide production in a galactose mutant of *S. meliloti*

We had previously shown that a mutant of the gene encoding 2-keto-3-deoxy-galactonokinase, *dgoK1*, (SRmD144) had an increased ability to compete for nodule occupancy in *S. meliloti* (Geddes & Oresnik, 2012b). This phenotype was of particular interest, as the inability to catabolize many other carbon sources resulted in a reduced ability to compete for nodule occupancy rather than an increased ability (Ding *et al.*, 2012; Fry *et al.*, 2001; Jiménez-Zurdo *et al.*, 1995; Kohler *et al.*, 2010; Oresnik *et al.*, 1998; Vanderlinde *et al.*, 2013; Yost *et al.*, 2006). In pursuit of a rationalization for the increased competitiveness of SRmD144, tests were conducted in search of additional phenotypes displayed by this strain, beyond an inability to utilize galactose as a sole carbon source. As part of this effort, we screened for effects on exopolysaccharide production since galactose is a subunit of both exopolysaccharides produced by *S. meliloti* (succinoglycan and EPSII) (Leigh & Walker, 1994).

The wild-type strain Rm1021 is unable to synthesize EPSII under normal conditions because of a mutation in the response regulator *expR* (Glazebrook & Walker, 1989; Pellock *et al.*, 2002). In order to assess effects on EPSII production, as well as succinoglycan production, the *dgoK1::Tn5* allele was transduced into both Rm1021 and an *expR*<sup>+</sup> variant of Rm1021, SRmA363, resulting in the construction of SRmD304 and SRmD305, respectively.

The presence of succinoglycan can be visualized under UV-light using the fluorescent dye Calcofluor (Finan *et al.*, 1985; Leigh *et al.*, 1985). EPSII production in *S. meliloti* can be visualized as a mucoid colony phenotype (Glazebrook & Walker, 1989; Miller-Williams *et al.*, 2006). Strains were screened on LB, VMM, RMM, YEM and ½ GYM media, in the presence or absence of galactose for effects on exopolysaccharide production. The Calcofluor UV-fluorescence phenotypes of Rm1021 and SRmD304 are presented in Table 6.2. Differences in UV-fluorescence between Rm1021 and SRmD304 were not observed on any of the medium tested when grown in the absence of galactose. In the presence of galactose, SRmD304 showed altered UV-fluorescence when grown on RMM, YEM and ½ GYM media. On all three media SRmD304 exhibited a dull-blue fluorescence under UV-light. Dull-blue fluorescence is in contrast to a bright blue-green fluorescence observed in the wild-type on RMM and YEM, in both strains in the absence of galactose, and as previously described (Finan *et al.*, 1985; Leigh *et al.*, 1985). On ½ GYM, it was in contrast to the absence of fluorescence after 3 days of growth. SRmA363 and SRmD305 showed the same UV-fluorescence phenotypes as Rm1021 and SRmD304 respectively.

SRmA363 and SRmD305 were also used to test for effects on EPSII production under the same conditions. The colony mucoidy phenotypes for these strains are presented in Table 6.3. Similar to UV-fluorescence, no significant differences were observed when grown on any of the media in the presence of glycerol. In the presence of

**Table 6.2.** Calcofluor UV-fluorescence phenotypes of a galactose mutant

Strain	Genotype	Carbon source	Medium				
			LB	VMM <sup>a</sup>	RMM	YEM	½ GYM <sup>a</sup>
Rm1021	Wild-type	Glycerol	+ <sup>b</sup>	+	+	+	- <sup>d</sup>
Rm1021	Wild-type <sup>+</sup>	Galactose	+	+	+	+	-
SRmD304	<i>dgoK1::Tn5</i>	Glycerol	+	+	+	+	-
SRmD304	<i>dgoK1::Tn5</i>	Galactose	+	+	+ (dull blue) <sup>c</sup>	+ (dull blue)	+ (dull blue)

Phenotypes were recorded after 3 days of growth

<sup>a</sup> Galactose containing medium also contained 15 mM glycerol to support growth.

Phenotypes were recorded after 3 days of growth

<sup>b</sup> + indicates typical bright blue-green Calcofluor fluorescence previously reported in the literature (Finan *et al.*, 1985; Leigh *et al.*, 1985).

<sup>c</sup> + (dull blue) indicates dull blue Calcofluor fluorescence, rather than bright blue-green Calcofluor fluorescence previously reported (Finan *et al.*, 1985; Leigh *et al.*, 1985).

<sup>d</sup> – indicates the absence of fluorescence under UV-light.

**Table 6.3.** Colony mucoidy phenotypes of a galactose mutant in an *expR*<sup>+</sup> background

Strain	Genotype	Carbon source	Medium				
			LB	VMM <sup>a</sup>	RMM	YEM	½ GYM <sup>a</sup>
SRmA363	<i>expR</i> <sup>+</sup>	Glycerol	+	+	+	+	+
SRmA363	<i>expR</i> <sup>+</sup>	Galactose	+	+	+	+	+
SRmD305	<i>expR</i> <sup>+</sup> , <i>dgoK1</i> ::Tn5	Glycerol	+	+	+	+	+
SRmD305	<i>expR</i> <sup>+</sup> , <i>dgoK1</i> ::Tn5	Galactose	+	+	-	-	-

Mucoid phenotypes are as follows: +, mucoid colony; -, non-mucoid colony

Phenotypes were recorded after 3 days of growth

<sup>a</sup> Galactose containing medium also contained 15 mM glycerol to support growth.

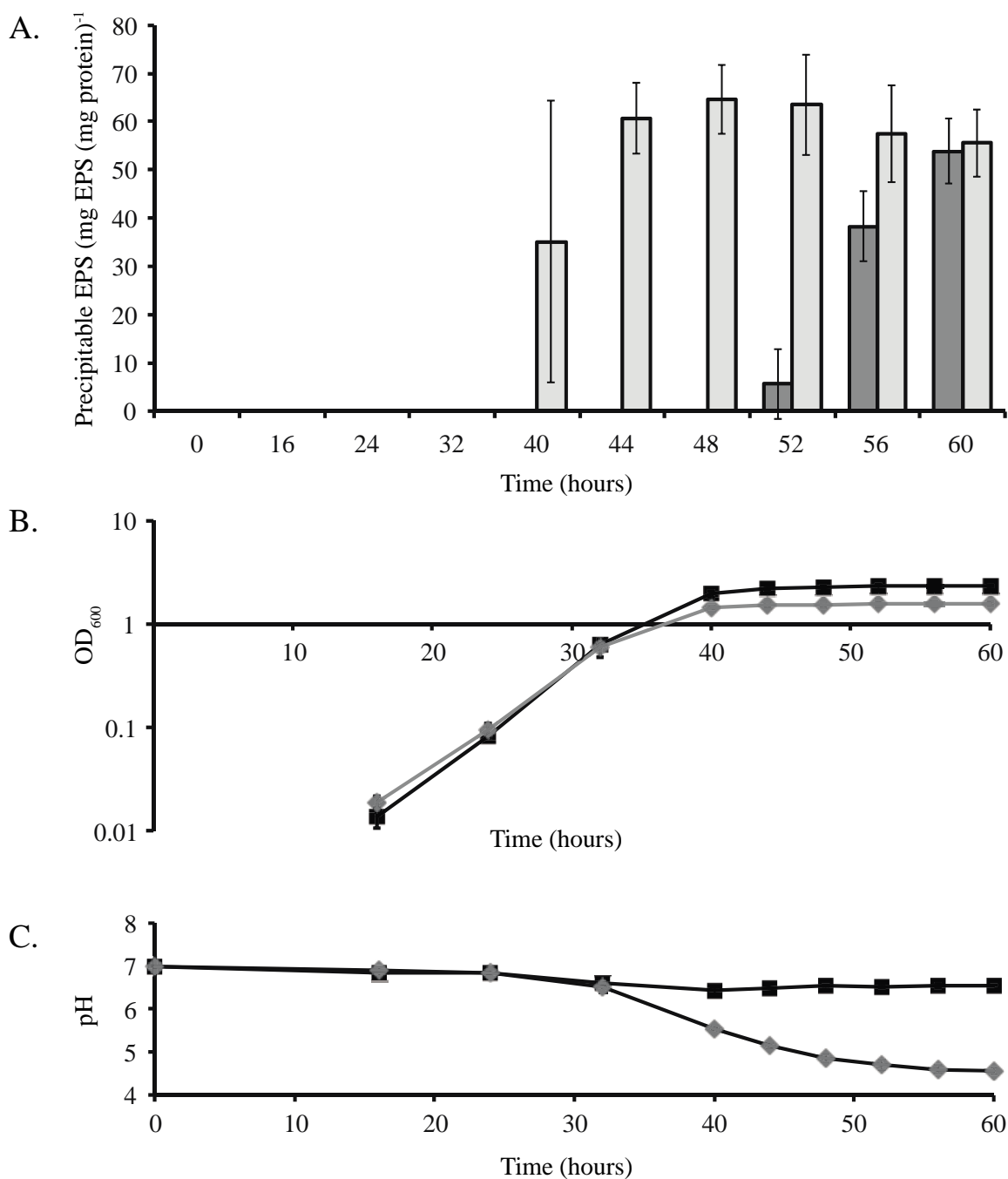
galactose, significantly reduced colony mucoidy of SRmD305 was observed compared to SRmA363 when grown on RMM, YEM and ½ GYM.

The reduction in colony mucoidy was consistent with the hypothesis that EPSII production is reduced in SRmD305 when grown on several defined media in the presence of galactose. The dull-blue UV-fluorescence phenotype we observed in galactose mutants was not consistent with previously observed phenotypes associated with succinoglycan production. We were interested in further characterizing exopolysaccharide production in SRmD304 in an effort to understand the atypical UV-fluorescence phenotype we observed.

### ***6.3.2 Early exopolysaccharide production and acidification of culture media by SRmD304***

In order to characterize altered EPS production in SRmD304 we wished to quantitatively assay EPS production in liquid medium over time. Traditionally, the anthrone-sulfuric acid method has been used to quantify EPS accumulation in *S. meliloti* culture supernatant (Glazebrook & Walker, 1989; Mendrygal & González, 2000). However, the EPS phenotypes we had observed were dependent on the presence of galactose in the culture media, which would interfere with this method. As an alternative, we used an EPS-precipitation based method that had previously been used to quantify HMW EPS production in rhizobia (Marroquí *et al.*, 2001; Yurgel *et al.*, 2013).

The results show that precipitable EPS accumulated in the SRmD304 culture at significantly earlier time-points than in Rm1021 (Figure 6.1 A). Precipitable EPS was detected 12 hours earlier in the SRmD304 culture than the Rm1021 culture

**Figure 6.1.** Exopolysaccharide production by Rm1021 and SRmD304

Dark grey bars, and black squares represent Rm1021. Light grey bars and gray diamonds represent SRmD304. Strains were grown in RMM with 15 mM galactose. EPS was precipitated from 50 mL of culture supernatant, dried and weighed at regular intervals over 60 hours of growth (A.). Optical density (B.) and pH (C.) of cultures was monitored at each of the time-point. Standard deviation (SD) of three independent replicates is shown by error bars. Where error bars are absent, SD was smaller than the marker.

(at 40 hours and 52 hours respectively). Growth-rates of both cultures were similar during the logarithmic growth phase (Figure 6.1 B). Coincident with early EPS production, a sharp drop in pH was observed in the SRmD304 cultures (Figure 6.1 C).

Acidic pH has previously been shown to induce the expression of genes for succinoglycan biosynthesis *S. meliloti* (Hellweg *et al.*, 2009). We hypothesized that early EPS production was a result of the acidification of the SRmD304 culture medium. Therefore, precipitable EPS was quantified in cultures of SRmD304 grown in presence or absence of 50 mM HEPES as a buffering agent. After 48 hours of growth, the pH of the unbuffered culture was  $4.95 \pm 0.01$  and  $45.1 \pm 2.5$  mg EPS/ mg protein was precipitated, whereas the pH of the buffered culture was  $6.36 \pm 0.03$  and no precipitable EPS was detected. These data are consistent with the hypothesis that acidification of SRmD304 culture media results in early EPS production in the galactose mutant.

We postulated that acidification of growth medium may explain the unusual Calcofluor phenotypes we observed. To test this, various culture media were buffered to pH 6.5 and 6.0 with HEPES, and the Calcofluor fluorescence of Rm1021 was observed after several days of growth. While the biomass visible on the plate was similar at pHs of 6.5 and 6.0, the typical bright blue-green fluorescence was observed at pH 6.5, whereas it was not apparent at pH 6.0.

### ***6.3.3 Early exopolysaccharide produced in SRmD304 is succinoglycan***

Since the Calcofluor UV-fluorescence phenotypes were ambiguous, we wished to directly determine the molecular identity of the EPS that was precipitated at earlier time-points in SRmD304. Therefore, we purified precipitable EPS and used anion exchange



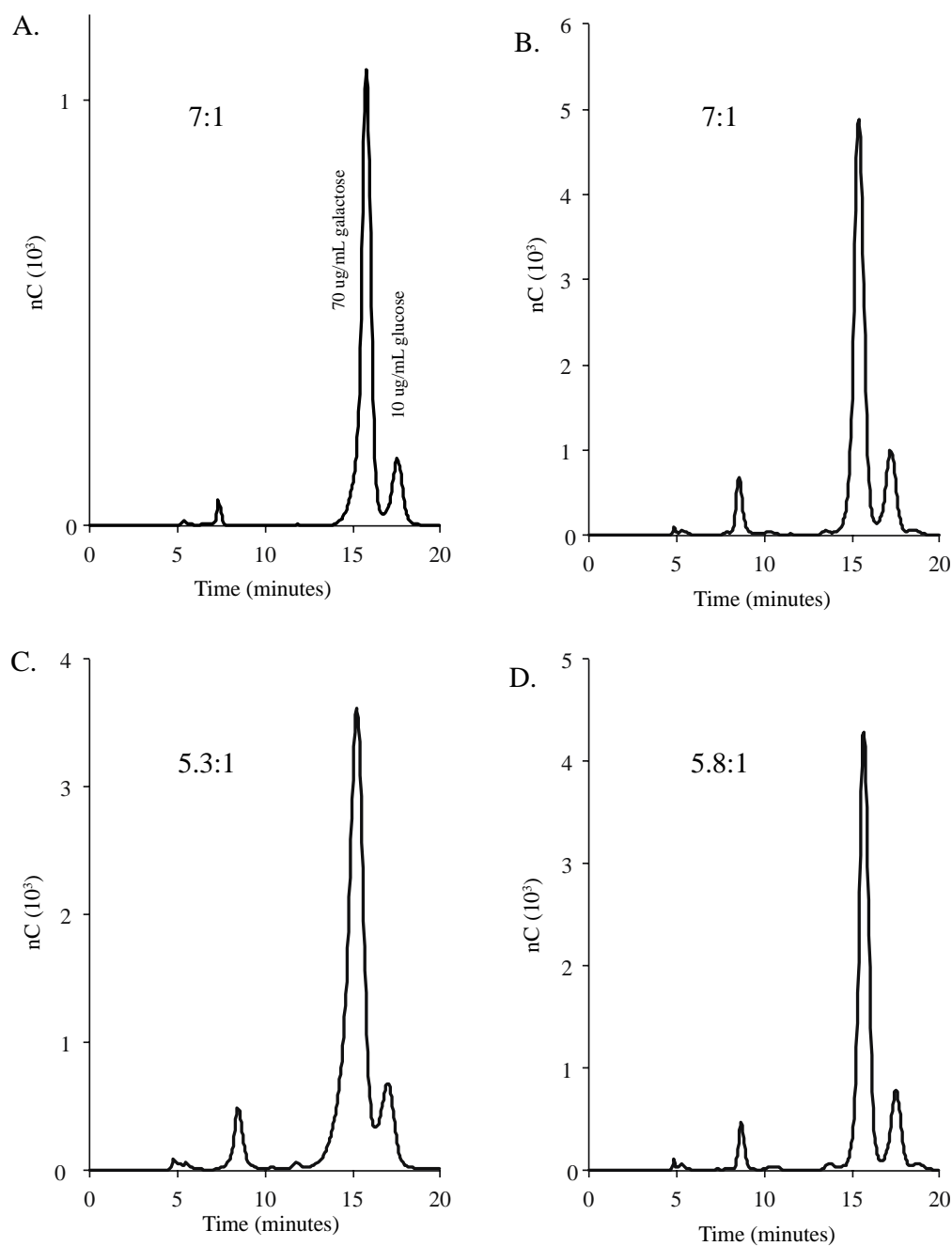
chromatography to determine the composition of its sugar subunits. EPS was precipitated and purified from an SRmD304 culture grown for 44 hours in RMM medium with galactose. As a control, precipitation was also performed on an Rm1021 culture grown under equivalent conditions; significant amounts of precipitable EPS from the wild-type culture were not observed.

The EPS isolated from the SRmD304 culture showed significant peaks that ran coincident with the monosaccharides glucose and galactose (Figure 6.2 A, Figure 6.2 B). No other significant peaks were observed. Quantification and normalization of the electrochemical response of the glucose and galactose peaks showed that glucose and galactose were present at a 7:1 ratio. This ratio of glucose and galactose is consistent with the production of exclusively succinoglycan at the early time-point in SRmD304. EPS was also precipitated and purified from Rm1021 and SRmD304 grown for 6 days in RMM medium with galactose (Figure 6.2 C, Figure 6.2 D). EPS purified from these cultures showed glucose:galactose ratios of (5.3:1) and (5.8:1), respectively, suggesting that the precipitated EPS was succinoglycan.

### **Succinoglycan is synthesized in HMW and LMW fractions at early time-points**

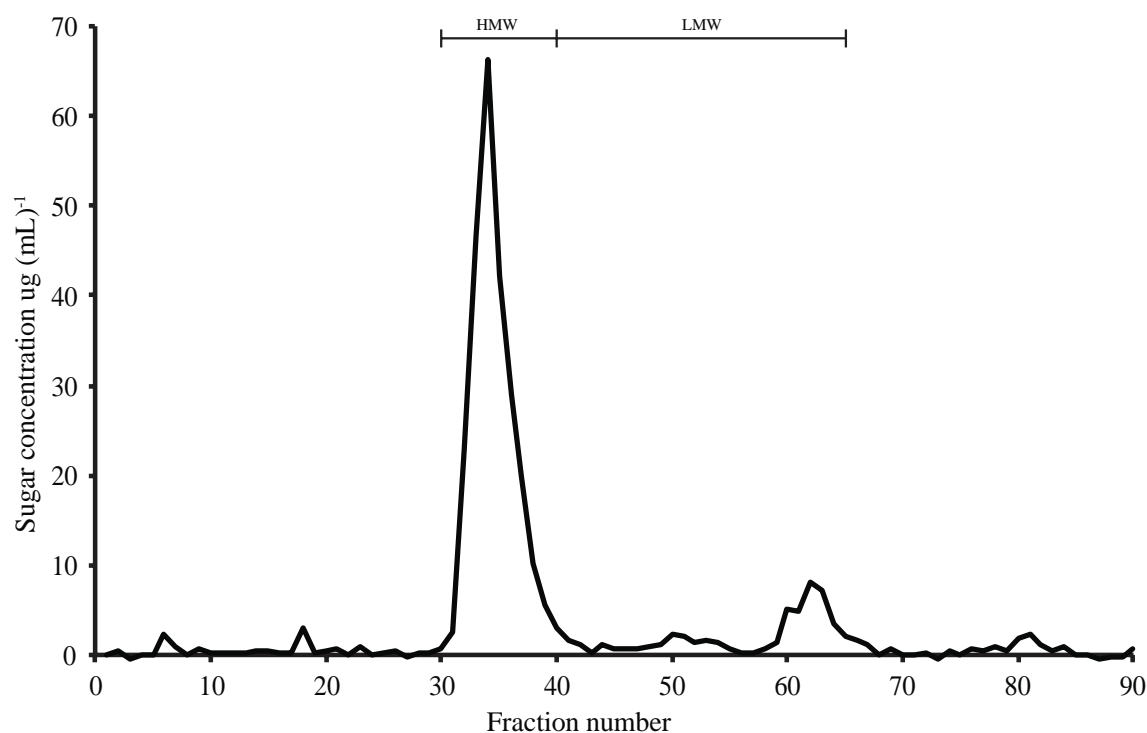
To analyze the molecular weight profile (HMW vs. LMW) of precipitated EPS, purified EPS was separated by column chromatography and quantified using the anthrone-sulfuric acid method (Morris, 1948). However, significant quantities of hexoses were only detected in fractions consistent with HMW succinoglycan (Figure 6.2).

**Figure 6.2.** Sugar composition of precipitated EPS.



Electrochemical response of glucose and galactose (A), precipitated EPS from SRmD304 after 44 hours growth (B), precipitated EPS from Rm1021 after 6 days growth (C), and precipitated EPS from SRmD304 after 6 days growth (D) using anion exchange chromatography. Ratio of glucose:galactose calculated by normalizing the electrochemical responses to A. are presented in the top left corners.

**Figure 6.3.** Molecular weight profile of precipitated EPS

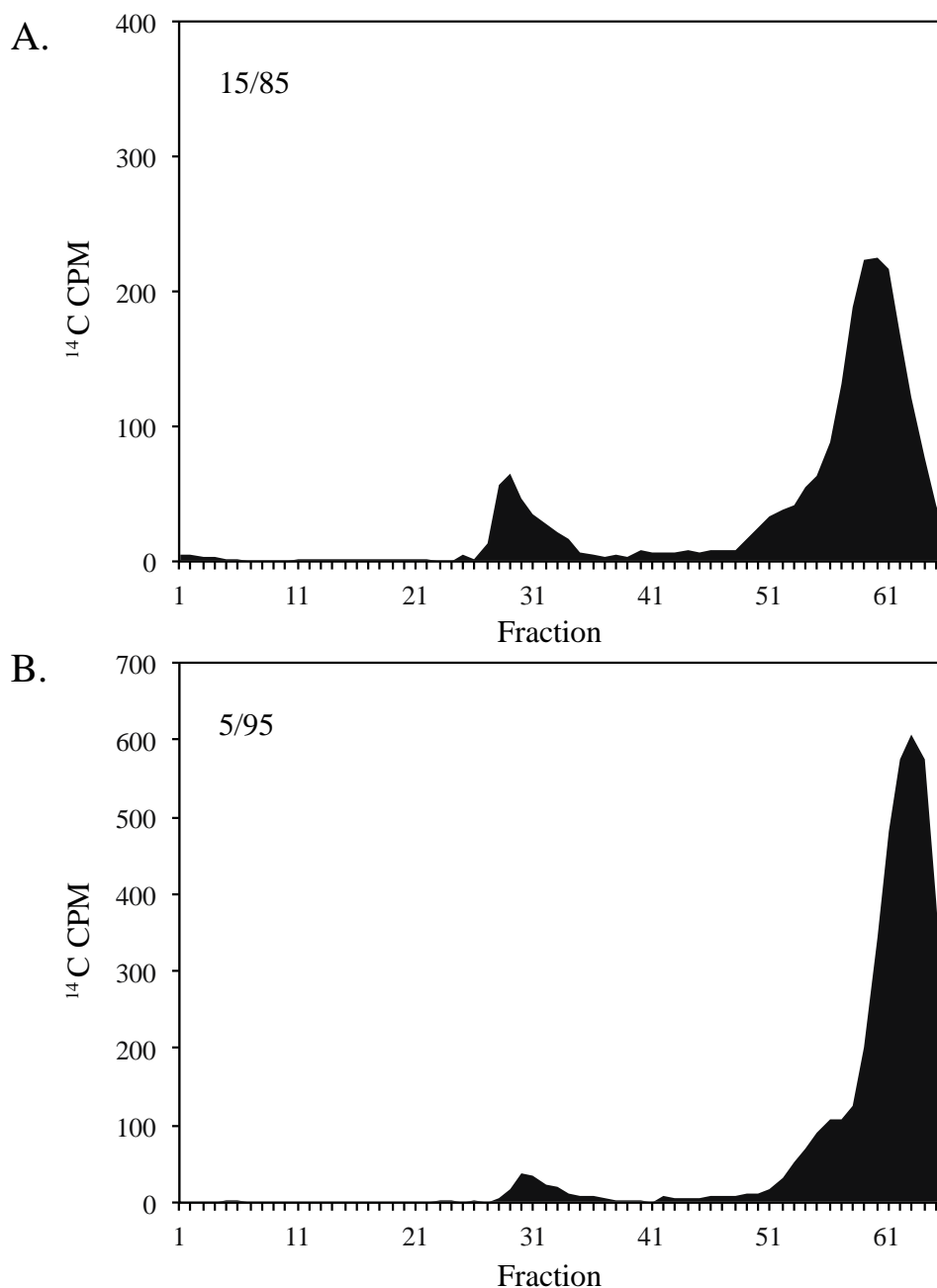


Representative example of precipitated EPS was separated by column chromatography. Hexose content of fractions was quantified using the anthrone-sulfuric acid method. All precipitated EPS samples showed similar profiles. Fractions containing HMW succinoglycan and LMW succinoglycan were determined based on previous experiments using the same column and matrix to separate radiolabelled succinoglycan (Glenn *et al.*, 2007).

To determine if LMW fractions of succinoglycan were also produced at early time-points, cells were labeled with D-[U-<sup>14</sup>C] glucose to allow direct observation of the molecular weight profile of succinoglycan in culture supernatant as previously described (Glenn *et al.*, 2007). The results showed substantial accumulation of LMW fractions of succinoglycan under these conditions. A distribution of approximately 15% HMW succinoglycan and 85% LMW succinoglycan was observed (Figure 6.4 A). These values are similar to those previously reported for wild type *S. meliloti* (Glenn *et al.*, 2007). We also assayed Rm1021 under these conditions and observed a distribution of approximately 5% HMW and 95% LMW succinoglycan (Figure 6.4 B).

#### ***6.3.4 Gene expression correlates with early succinoglycan production***

Characterization of the composition and molecular-weight distribution of early exopolysaccharide production in SRmD304 suggested that we observed early succinoglycan biosynthesis in the galactose mutant compared to wild-type. We hypothesized that we should observe increased expression of genes involved in succinoglycan biosynthesis under these culture conditions. Therefore, RNA was isolated from Rm1021 and SRmD304 grown under the conditions of early EPS synthesis. Following cDNA synthesis, gene expression was analyzed by qRT-PCR. The data are shown in Table 6.4. Consistent with the early synthesis of succinoglycan in SRmD304, we observed increased expression of the succinoglycan biosynthesis genes *exoY* and *exoH*, and the endoglycanases responsible for generation of LMW fractions of succinoglycan *exoK* and *exsH* compared to Rm1021.

**Figure 6.4.** Molecular-weight profile of EPS synthesized by Rm1021 and SRmD304

$^{14}\text{C}$  succinoglycan from the supernatants of SRmD304 (A) and Rm1021 (B) cultures was separated by column chromatography. The ratio of HMW to LMW succinoglycan is indicated in the top left corner as a percentage. Fractions containing HMW succinoglycan and LMW succinoglycan were determined based on previous experiments using the same column and matrix to separate radiolabelled succinoglycan (Glenn *et al.*, 2007). The experiment was repeated twice showing consistent results.

**Table 6.4.** Changes in gene expression in SRmD304

Gene	Gene product function	Fold change in gene expression <sup>a</sup>
<i>exoY</i>	Succinoglycan galactosyltransferase	6.7
<i>exoH</i>	Succinoglycan succinyltransferase	7.8
<i>exoK</i>	Succinoglycan endoglycanase	6.0
<i>exsH</i>	Succinoglycan endoglycanase	7.8
<i>flgG</i>	Flagellar basal-body rod protein	0.15 <sup>b</sup>
<i>lpsB</i>	LPS core biosynthesis	NC
<i>ndvB</i>	β-(1,2) glucan transmembrane protein	NC
<i>expE2</i>	EPSII glycosyltransferase	NC
<i>nodA</i>	Nod Factor N-acyltransferase	NC

<sup>a</sup>Data expressed as  $2^{\Delta\Delta Ct}$  and represents fold expression in SRmD304 over Rm1021 grown in RMM with 15 mM galactose. The experiment also included *SMc00128* as an internal control (Krol & Becker, 2004). The table represents data from a single experiment. The experiment was repeated three times showing consistent results. The experiment was also repeated in triplicate using ½ GYM medium and showed consistent results.

<sup>b</sup>0.15 is equivalent to a 6.7 fold reduction in gene expression.

NC = no change

To investigate possible effects on other biologically relevant processes, representative genes of flagella (*flgB*), cyclic  $\beta$ -(1,2) glucan (*ndvB*), LPS (*lpsB*), EPSII (*expE2*) and Nod Factor biosynthesis (*nodA*) were tested for changes in gene expression. Decreased expression of the flagella biosynthesis gene *flgB* was observed. Significant changes in the expression of the other genes were not observed (Table 6.4).

### ***6.3.5 SMc00588 encodes a galactose dehydrogenase that is required for galactose-dependent culture medium acidification***

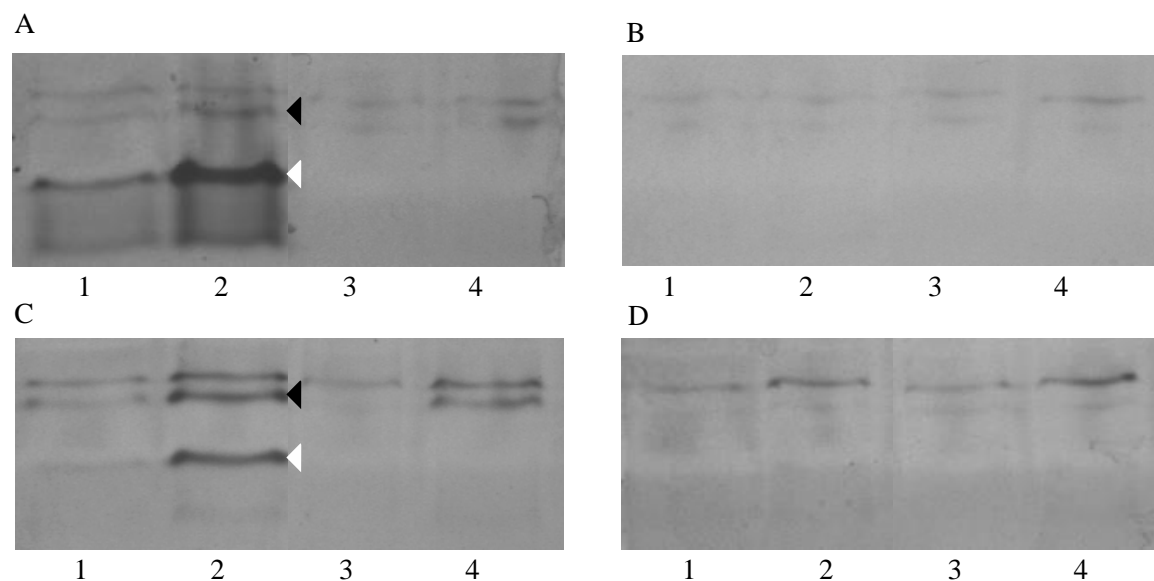
The strain SRmD304 carries a mutation in the galactonokinase *dgoK1*. Thus, galactose catabolism is blocked at the step of the phosphorylation of 2-keto-3-deoxy-galactonate in this mutant (Arias & Cerveñansky, 1986; Geddes & Oresnik, 2012b). We wanted to determine if medium acidification was the result of a metabolic defect resulting from the build-up of intermediates of galactose catabolism. Galactose catabolism in *S. meliloti* begins with the oxidation of galactose by galactose dehydrogenase (Arias & Cerveñansky, 1986). Therefore, we measured medium acidification and EPS production in the galactose dehydrogenase mutant SRmD441 (*galD::Tn5-B20*, polar on *dgoK1*). In this background similar levels of medium acidification (pH  $5.32 \pm 0.36$ ), and early EPS production ( $29.96 \pm 14.01$  mg EPS/ mg protein) were observed after growth for 44 hours in defined medium containing galactose, under the same conditions described for Figure 6.1. However, it was previously shown that *S. meliloti* contains a second unidentified galactose-inducible dehydrogenase that was capable of the oxidation of galactose and arabinose (Geddes & Oresnik, 2012b).

Based on annotation and similarity to L-arabinose/ galactose dehydrogenases, identified in *Agrobacterium tumefaciens* and *Azospirillum brasiliense* (Watanabe *et al.*, 2006; Zhao & Binns, 2011), we hypothesized that the putative galactose dehydrogenase *SMc00588* encoded for the redundant galactose dehydrogenase activity we had observed. Therefore, we constructed a single cross-over disruption in this gene using the suicide vector pKNOCK-Gm. The resulting mutant SRmD434 (*SMc00588::pKNOCK-Gm*) was unable to grow using L-arabinose as a sole carbon source and grew slowly using galactose as a sole carbon source. Assays for galactose and L-arabinose dehydrogenase activities using non-denaturing polyacrylamide gel electrophoresis confirmed that *SMc00588* is a galactose, L-arabinose dehydrogenase (Figure 6.5). During the characterization of L-arabinose catabolism in *S. meliloti*, an L-arabinose dehydrogenase was not identified (Poysti *et al.*, 2007). This evidence shows that *SMc00588* encodes an L-arabinose dehydrogenase in *S. meliloti*.

To block galactose catabolism at the first step, the oxidation of galactose, we constructed a double mutant of *galD* and *SMc00588*, SRmD436. To test if galactose oxidation was required for medium acidification, this strain (also missing *dgoK1* due to polar effects of the *galD* mutation) was grown as described for Figure 6.1, for 44 hours in RMM containing galactose. After 44 hours of growth SRmD436 (*galD::Tn5-B20*, *SMc00588::pKNOCK-Gm*) showed wild-type pH levels ( $6.57 \pm 0.02$ ) and the absence of early EPS production, whereas SRmD441 (*galD::Tn5-B20*) showed medium acidification and early EPS production at similar levels to SRmD304 as described above. These results



**Figure 6.5.** Non denaturing PAGE of galactose-inducible enzyme activity



Extracts of Rm1021 grown on defined medium containing glycerol as a sole carbon source (Lane 1), Rm1021 grown on defined medium containing glycerol and galactose as sole carbon sources (Lane 2), SRmD436 grown on defined medium containing glycerol and galactose as carbon sources (Lane 3) and SRmD434 grown on defined medium containing glycerol and galactose as carbon sources (Lane 4) are shown in each panel. **(A)** Assayed in the presence of  $\text{NADP}^+$  and galactose. **(B)** Assayed in the presence of  $\text{NADP}^+$  and the absence of galactose. **(C)** Assayed in the presence of  $\text{NAD}^+$  and galactose. **(D)** Assayed in the presence of  $\text{NAD}^+$  in the absence of galactose. Black arrow indicates the band predicted to correspond to GalD. White arrow indicates the band predicted to correspond to SMc00588. Panels **(B)** and **(D)** also serve as loading controls for panels **(A)** and **(C)**.

are consistent with the hypothesis that the observed media acidification resulted from a metabolic defect dependent on the intracellular oxidation of galactose.

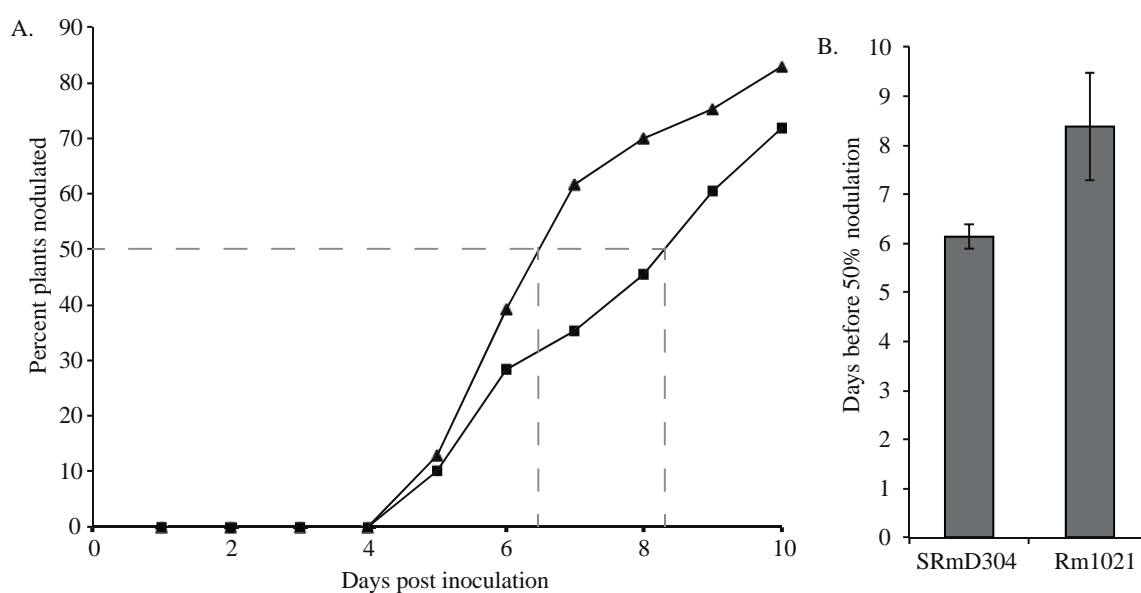
### ***6.3.6 Early nodulation in an *S. meliloti* galactose mutant***

To continue to characterize the increased competitiveness that was previously observed in the *dgoK1* mutant SRmD144, we investigated the nodulation kinetics in plants inoculated with SRmD304 compared to Rm1021. The kinetics of nodule formation was monitored for over 100 alfalfa plants over 5 independent replicates in Rm1021 and SRmD304. The effects observed over the first 10 days post-inoculation are shown in Figure 6.6 A. Nodulation of the first plants occurred at the same time in both Rm1021 and SRmD304 at 5 days post-inoculation. Following the nodulation of the first plants on day 5, a portion of plants became nodulated at an earlier time-point in SRmD304 compared to Rm1021. This phenotype was the most exacerbated as 50% of plants reached nodulation: 50% of plants inoculated with SRmD304 became nodulated by approximately 6 days post-inoculation, whereas 50% of plants inoculated with Rm1021 became nodulated by approximately 8 days post-inoculation (Figure 6.6 B).

### ***6.3.7 EPS phenotypes in galactose mutants correlate with competition for nodule occupancy***

Increased production of succinoglycan has recently been correlated with enhanced symbiosis of *S. meliloti* 1021 with *M. truncatula* (Jones, 2012), and galactose has been shown to be prevalent near the root surface of *M. sativa* seedlings (Bringham *et al.*, 2001). We hypothesized that enhanced succinoglycan production during invasion may

**Figure 6.6.** Effects on nodule kinetics in a galactose mutant of *S. meliloti*



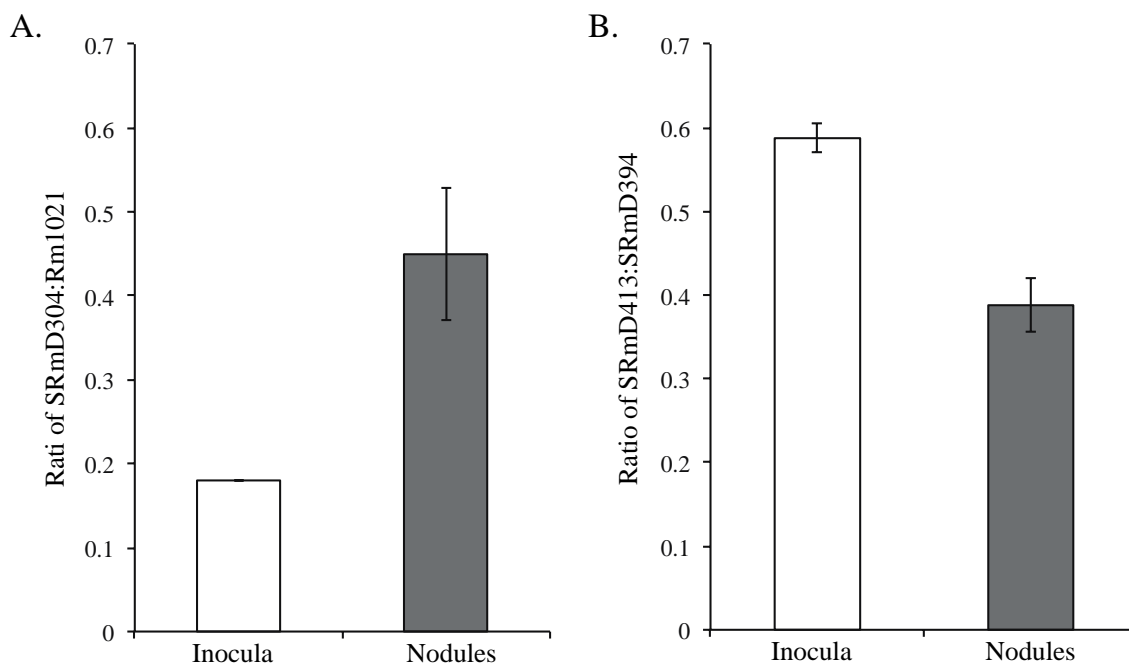
**A.** Percent of plants nodulated over time after inoculation with Rm1021 or SRmD304. Data is the sum of 100 plants over 5 independent replicates. **B.** Average number days post-inoculation before 50 % of plants became nodulated. The average and SD was calculated over 3 independent replicates that contained over 20 plants each.

have mediated increased competitiveness for nodule occupancy in SRmD144. If the effects we had observed on exopolysaccharide production in SRmD144 were occurring during the invasion of alfalfa, we predicted that *S. meliloti* using EPSII to invade *M. sativa*, rather than EPSI should have a reduced ability to compete for nodule occupancy, as we observed reduced synthesis of EPSII in *expR*<sup>+</sup> galactose mutants of *S. meliloti*. Conversely, if the increased competitiveness was due to other phenomena, the competitive advantage may be maintained.

To carry out this experiment we constructed a mutant that was unable to synthesize succinoglycan in the *expR*<sup>+</sup> background that is capable of EPSII synthesis, SRmD394 (*exoY*, *expR*<sup>+</sup>). SRmD394 was used in a competition for nodule occupancy experiment with a galactose mutant with the same genetic background, SRmD413 (*dgoK*, *exoY*, *expR*<sup>+</sup>). Colony mucoidy phenotypes in these backgrounds were consistent with those described in Table 3. Significantly reduced production of EPSII was observed in SRmD413 compared to SRmD394 when grown on defined media containing galactose.

The results from the competition experiments are presented in Figure 6.7. As previously reported, the *dgoK* mutant in an Rm1021 background was significantly more competitive for nodule occupancy than Rm1021 (Figure 6.7 A). When the competition experiment was performed in strains that used EPSII rather than succinoglycan to invade, we observed a significantly reduced ability to compete for nodule occupancy in mutants that contained the *dgoK*::Tn5 allele (Figure 6.7 B). These data are consistent with the hypothesis that *in vitro* effects on exopolysaccharide production we have described in this work occur *in planta* during the invasion of *M. sativa* by *S. meliloti*, and influence the ability of *S. meliloti* strains to compete for nodule occupancy.

**Figure 6.7.** Competition for nodule occupancy



**A.** SRmD304 competed against Rm1021 for nodule occupancy of *M. sativa*. The data are presented as the mean  $\pm$  standard deviation of the proportion of SRmD304 present in the inoculum (white bar) and isolated from nodules (dark grey bar) data are the average of 3 independent replicates, each comprised of 30 plants. **B.** SRmD413 competed against SRmD394 for nodule occupancy (using EPSII for invasion) of *M. sativa*. The data are presented as the mean  $\pm$  standard deviation of the proportion of SRmD413 present in the inoculum (white bar) and isolated from nodules (dark grey bar) data are the average of 2 independent replicates, each comprised of 30 plants. Reduced competitiveness of SRmD413 was also observed over 3 independent replicates using a 20% proportion of SRmD413 in the inoculum.

### **6.3.8 Acidification of the curled colonized root hair during invasion by *S. meliloti***

A signal that activates succinoglycan biosynthesis in *S. meliloti* during the invasion of host plants has not been described. Based on the previous data, we hypothesized that exposure to acidic environments during the invasion of *M. sativa* by *S. meliloti* may contribute to the activation of succinoglycan biosynthesis. We utilized fluorescent microscopy with an acidotropic dye, LysoTracker Red DND-99, to test for the possible acidic nature of curled colonized root hairs, and infection threads in *M. sativa* root hair cells during invasion by *S. meliloti*. LysoTracker dyes are composed of a fluorophore conjugated to a weakly basic amine that is permeable to plant tissues. In acidic compartments the amine is believed to become protonated and prevent free diffusion of the stain, resulting in the trapping of the fluorophore in acidic compartments. As a result, LysoTracker Red DND-99 provides qualitative evidence of an acidic pH in a compartment where its fluorescence co-localizes. We used rhizobia expressing GFP to visualize and define the compartments of the curled colonized root hair and infection thread during invasion. Recently, LysoTracker dyes have been used to demonstrate the acidic nature of the peribacteroid space in nitrogen fixing bacteroids *in planta* (Pierre *et al.*, 2013).

Invasion events on plant roots by *S. meliloti* were visualized by fluorescent microscopy using an endow green fluorescent protein filter set. Curled colonized root hairs and infection threads occupied by *S. meliloti* were analyzed for co-localization of red fluorescence from LysoTracker Red DND-99. Several curled colonized root hairs and infection threads were visible on most roots. Co-localization of red fluorescence with green fluorescence was observed in approximately 80% of curled colonized root hairs.

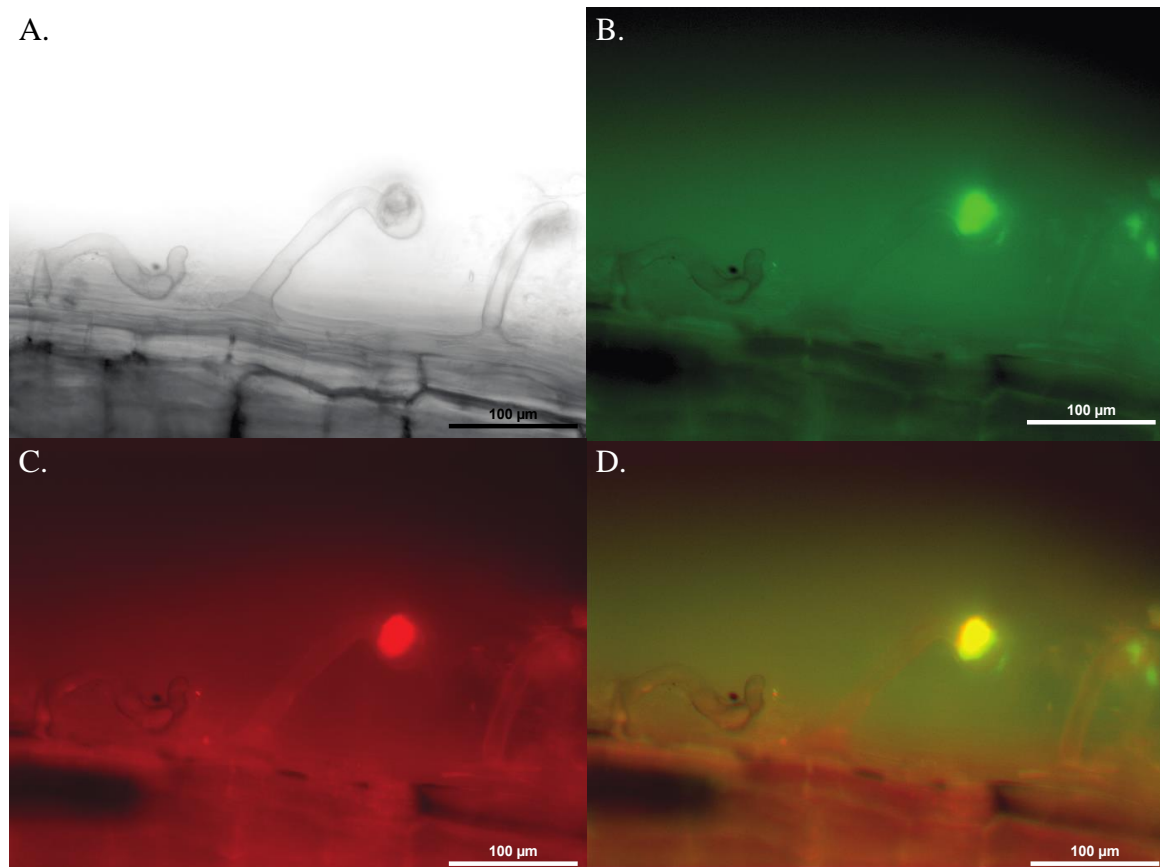
Co-localization was visualized in both Rm1021 and SRmD304. A representative image of the co-localization of fluorescence in a curled colonized root hair is presented in Figure 6.8. Co-localization of red fluorescence with green fluorescence was not observed in any of the infection threads visualized. Co-localization of red fluorescence with the green fluorescence of *S. meliloti* strains in curled colonized root hairs of *M. sativa* is consistent with an acidic nature of this compartment during invasion by *S. meliloti*.

## 6.4 Discussion

In this work we have shown that the galactose mutant SRmD304 produces the symbiotic exopolysaccharide at earlier time-points than the wild type in defined media containing galactose. We have also demonstrated the suppression of EPSII synthesis under similar conditions in backgrounds that contained an intact copy of the *expR* allele. We showed that the observed EPS phenotypes correlated with competitiveness for nodule occupancy in galactose mutants. We have also provided evidence that the CCRH may be an acidic compartment.

The observed early synthesis of succinoglycan was dependent on the ability of SRmD304 to acidify its culture medium. The induction of genes associated with succinoglycan biosynthesis following a transition to growth in acidic pHs has previously been reported (Hellweg *et al.*, 2009). Consistent with this study we observed the induction of genes for succinoglycan biosynthesis and the repression of genes associated with flagella biosynthesis in the acidic SRmD304 cultures. These data are consistent with the hypothesis that the acidification of its environment by SRmD304 resulted in the

**Figure 6.8.** The curled colonized root hair is an acidic compartment



**A.-D.** Confocal images of a curled colonized root hair from a root 7 days after inoculation with *S. meliloti* expressing *gfp*, stained with Lysotracker Red DND-99. **A.** Differential interference of contrast (DIC) image of curled colonized root hair (center). **B.** Green fluorescence of *S. meliloti* invading the curled colonized root hair expressing *gfp* viewed through endow GFP filter set. **C.** Red fluorescence of Lysotracker Red DND-99 viewed through rhodamine filter set. **D.** Merged image showing the co-localization of green and red fluorescence of GFP and Lysotracker Red DND-99.



induction of genes for succinoglycan biosynthesis and the early production of succinoglycan we observed.

We set out to identify possible explanations for the increased competitiveness for nodule occupancy we had previously observed in an *S. meliloti* galactose mutant (Geddes & Oresnik, 2012b). Three hypotheses can be generated regarding the phenotypes we observed *in vitro* and the increased competitiveness for nodule occupancy in SRmD144/SRmD304. 1) Early EPS production and medium acidification *in vitro* are unrelated to increased competitiveness. 2) Increased competitiveness results from environment acidification that occurs during invasion but is independent of effects on EPS production. 3) Increased competitiveness is mediated by effects on EPS production, resulting from acidification of its environment by galactose mutants during invasion. We believe that several lines of evidence, taken together, favor the last hypothesis. Galactose containing compounds shown to be prevalent on the surface of alfalfa roots in sterilized and unsterilized soils (Bringhurst *et al.*, 2001). Enhanced production of succinoglycan has recently been shown to enhance symbiosis of *S. meliloti* with *M. truncatula* (Jones, 2012). The competitiveness for nodule occupancy of galactose mutants in this study correlated with effects on EPS production we observed *in vitro*. If increased competitiveness for nodule occupancy was due to other factors in SRmD304, we would have expected a galactose mutant to maintain a competitive advantage over a strain that was unaffected in its ability to utilize galactose when using EPSII to invade *M. sativa*, rather than succinoglycan.

The correlation between competitiveness and the ability to acidify the environment of *S. meliloti* in response to hexoses is present in another example in the

literature. During growth using glucose as a carbon source, *S. meliloti* is able to acidify its environment by the periplasmic oxidation of glucose using a membrane bound pyrroloquinoline-quinone dependent glucose dehydrogenase (Bernardelli *et al.*, 2001; Gosselin *et al.*, 2001). Mutants of the gene encoding the PQQ-linked glucose dehydrogenase, *gcd*, had a reduced ability to compete for nodule occupancy, despite being unimpaired in growth rate when using glucose as a sole carbon source (Bernardelli *et al.*, 2001; Gosselin *et al.*, 2001). Interestingly, *gcd* mutants also showed a delay in nodulation kinetics (Bernardelli *et al.*, 2008), in contrast to the early nodulation we observed in a proportion of the plants inoculated with SRmD304.

A role for acidic pH in the interactions of microbes with eukaryotic hosts is not unprecedented. The intracellular survival of *Brucella* spp. in human monocytes was dependent on an acidic intraphagosomal pH (Rittig *et al.*, 2001). Virulence genes in *A. tumefaciens* are activated by the two-component system VirA/VirG in cooperation with the sugar binding protein ChvE in response to phenolic compounds, acidic pH and hexoses (Charles & Nester, 1993; He *et al.*, 2009). Acid-inducible genes in *A. tumefaciens* are also regulated by a two-component system that is homologous to the ExoS-ChvI component system of *S. meliloti*, ChvG-ChvI (Li *et al.*, 2002). A mechanism contributing to the induction of the ChvG-ChvI system by acidic pHs has recently been demonstrated. Essentially the *A. tumefaciens* homolog of ExoR that inhibits ChvG in the periplasm is degraded under acidic conditions, allowing activation of the system, and the induction of acid-inducible genes by ChvI (Wu *et al.*, 2012). In *S. meliloti*, the ExoS-ChvI system has also been shown to be activated by the transient degradation of ExoR

(Lu *et al.*, 2012). Interestingly, a ChvI binding site has been identified upstream of the galactose dehydrogenase *SMc00588* (Bélanger & Charles, 2013).

In this study we have provided evidence that the CCRH may be an acidic compartment. It is possible that acidic pHs play a role in the activation of succinoglycan biosynthesis in the CCRH in *S. meliloti* through the activation of the ExoS-ChvI system, as in *A. tumefaciens*. It is unclear whether the acidic environment of the CCRH is plant-derived or the result of the action of *S. meliloti* on its environment. It is also unclear whether the acidic nature of the CCRH is transient or persistent throughout invasion. Our working hypothesis is that the galactose mutant is more competitive for nodule occupancy because of an increased acidification of its environment in response to galactose in the CCRH, resulting in enhanced succinoglycan production.

## **Chapter 7**

### **Conclusions**

## 7.1 Thesis Conclusions and Observations

The goals of this thesis were to characterize the genetic loci for erythritol and galactose catabolism and assess their influence on the ability of *S. meliloti* to compete for nodule occupancy. As described in the previous sections, these goals were carried out and led to many interesting conclusions about *S. meliloti* carbon metabolism and biology. These conclusions, along with outstanding questions and general observations about *S. meliloti* carbon metabolism and biology are discussed in this section.

It was hypothesized that the genetic locus that was identified adjacent to *tpiB* was responsible for erythritol catabolism in *S. meliloti*. This hypothesis was confirmed during the genetic characterization of the erythritol locus that was later shown to also be involved in the catabolism of adonitol and L-arabitol. Erythritol was shown to be catabolized by EryA, EryB, EryC and TpiB, as has been demonstrated in *R. leguminosarum* and *Brucella* spp. (Geddes *et al.*, 2010; Sangari *et al.*, 2000; Yost *et al.*, 2006). The catabolism of erythritol to DHAP begins with the phosphorylation of erythritol by EryA, followed by oxidations by EryB and EryC generating 3-keto-erythrose-4-phosphate (Figure 1.8) (Sperry & Robertson, 1975a, b). The further catabolism of erythritol requires one additional oxidation reaction, followed by decarboxylation to yield DHAP. The biochemical activities of these enzymes have been identified in *B. abortus*. The oxidation of 3-keto-erythrose-4-phosphate was performed by a membrane-bound dehydrogenase coupled to the electron transport chain, and the decarboxylation by a soluble decarboxylase (Sperry & Robertson, 1975a, b). Genes that encode these enzymes have not been identified.

We hypothesized that the erythritol locus of *S. meliloti* may encode these enzymes based on the presence of many additional carbon catabolism genes, but it was shown that the other catabolic genes in the locus were not required for erythritol catabolism (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010). Extensive mutagenesis failed at identifying other loci involved in erythritol catabolism in *S. meliloti*, and our phylogenetic analysis did not identify any candidate genes in the putative erythritol loci of other  $\alpha$ -proteobacteria (Geddes *et al.*, 2013; Geddes *et al.*, 2010). Therefore, genes that encode the final two steps of erythritol catabolism in bacteria remain unelucidated. One approach to identifying the full complement of genes required for the catabolism of a carbon compound is through global gene expression analysis following growth on the compound of interest as a sole carbon source. This approach has recently been used during the characterization of a novel pathway of hydroxyproline catabolism in *S. meliloti* (White *et al.*, 2012). Microarray analysis during growth using erythritol as a sole carbon source has been recently performed in *B. abortus* and *B. melitensis*, but candidate genes for these activities were not identified (Petersen *et al.*, 2013; Rodríguez *et al.*, 2012).

Another unresolved aspect of erythritol catabolism is the unique requirement for TpiB. It was previously shown that both *tpiA* and *tpiB* encode functional triose-phosphate isomerases. However, while *tpiB* was able to complement a *tpiA* mutant for growth using rhamnose or glycerol as sole carbon sources, *tpiA* was unable to complement a *tpiB* mutant for growth using erythritol as a sole carbon source (Poysti & Oresnik, 2007). A future direction for the study of erythritol catabolism in *S. meliloti* could be to pursue an explanation for this phenomenon. One hypothesis is that TpiB

interacts with another protein involved in erythritol catabolism, and that this interaction is required for successful erythritol catabolism. Such an interaction is not without precedent, for example the rhamnose sugar kinase, RhaK, is required for rhamnose ABC transport activity in *R. leguminosarum*, and the kinase activity can be genetically uncoupled from transport (Richardson & Oresnik, 2007; Rivers & Oresnik, 2013). Interestingly, decarboxylase activity was unable to be separated from triose-phosphate isomerase activity during the characterization of erythritol catabolism in *B. abortus* (Sperry & Robertson, 1975a).

Regulation of the genetic loci for erythritol catabolism in *R. leguminosarum* and *B. abortus* is carried out by EryD (Sangari *et al.*, 2000; Yost *et al.*, 2006). In these organisms the transcription of the erythritol operon *eryABCD* was shown to be repressed by EryD and induced by erythritol. Strains carrying mutations in *eryD* showed constitutive expression of *eryABCD* and were unaffected in the ability to utilize erythritol as a sole carbon source (Sangari *et al.*, 2000; Yost *et al.*, 2006). We demonstrated that the regulation of the erythritol locus of *S. meliloti* proceeds differently than in these closely related organisms. The genes for erythritol (and adonitol and L-arabitol) catabolism are distributed over 5 transcriptional units that are all regulated by EryD (Geddes *et al.*, 2010). EryD acts as both a negative and positive regulator of these transcripts, and a strain carrying a mutation in *eryD* was unable to grow using erythritol, adonitol or L-arabitol as sole carbon sources (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010). Induction by EryD was dependent on intermediates of erythritol, adonitol, or L-arabitol catabolism (Geddes & Oresnik, 2012a; Geddes *et al.*, 2010).

We also assigned function to a putative negative regulator encoded upstream of *tpiB*, *SMc01615* (renamed *eryR*). We showed that EryR is a negative regulator of its own transcriptional unit (contains *tpiB*), and also influences the transcription of *eryD*. EryR repression was only lifted during growth with erythritol as a sole carbon source (Geddes *et al.*, 2010). Homologs to *eryR* are present in the erythritol loci of *R. leguminosarum* and *Brucella* spp. (Geddes *et al.*, 2013).

It has recently been shown that growth using erythritol as a sole carbon source resulted in the induction of virulence genes in *B. melitensis* (Petersen *et al.*, 2013). Therefore, given that erythritol catabolism effects the ability to compete for nodule occupancy in *R. leguminosarum* it would be interesting to perform a similar global expression study in this organism to determine if increased competitiveness may be a result of regulatory effects on other loci involved in the invasion of legume hosts.

The ABC transporter encoded by the erythritol locus, *SMc01628-24* (renamed *mptABCD*) is not homologous to the erythritol ABC transporter of *R. leguminosarum* (*eryEFG*) that was shown to be divergently transcribed from *eryABCD* (Yost *et al.*, 2006). Phylogenetic analysis showed that putative erythritol loci in the  $\alpha$ -proteobacteria contained homologs to one of these two ABC transporters. The presence of a homologous transporter to that of *S. meliloti* correlated with the predicted capacity of the loci to encode adonitol and L-arabitol catabolism (Geddes *et al.*, 2013). Consistent with this, *mptABCD* was shown to be required for growth using erythritol, adonitol or L-arabitol as sole carbon sources (Geddes *et al.*, 2010). Transport competition experiments showed that adonitol and L-arabitol competed with radiolabeled erythritol for transport (Geddes & Oresnik, 2012a). This evidence is consistent with the hypothesis that



components of an ABC transporter are shared for the transport of all three polyols. It is still unclear whether all components encoded by *mptABCD* comprise an ABC transporter that is used for all three polyols, or if only some of the components are shared. Attempts to purify the sugar binding protein MptA in order to determine if it was able to bind all three polyols were unsuccessful, but future work could be directed at determining how MptABCD contribute to the transport of erythritol, adonitol and L-arabitol.

Introducing a cosmid containing the erythritol locus of *S. meliloti* conferred the ability to utilize adonitol and L-arabitol in a heterologous host (Geddes & Oresnik, 2012a). This finding suggested that the locus contained all of the genes necessary for the catabolism of adonitol and L-arabitol. We used defined deletions to demonstrate the complement of genes used for adonitol and L-arabitol catabolism in *S. meliloti*. This resulted in ascribing function to the remaining genes that were contained in the erythritol locus. We showed that *SMc01620*, *SMc01618* and *SMc01617* (renamed *rbtABC*) were required for adonitol catabolism, and that these three genes as well as *SMc01619* (renamed *lalA*) were required for L-arabitol catabolism. We also demonstrated that the erythritol kinase EryA directly phosphorylates adonitol and L-arabitol.

The common route of adonitol and L-arabitol catabolism involves oxidation to a 2-keto-pentulose, followed by phosphorylation yielding D-ribulose-5-phosphate and L-xylulose-5-phosphate respectively. D-ribulose-5-phosphate and L-xylulose-5-phosphate enter central catabolism via conversion to D-xylulose-5-phosphate (Mortlock, 1984). Our data were inconsistent with such a pathway for adonitol and L-arabitol catabolism in *S. meliloti*. Based on the genetic and bioinformatics data a putative novel pathway for adonitol and L-arabitol catabolism was proposed (Figure 3.4). Future work should be

directed towards biochemically defining the pathways of adonitol and L-arabitol catabolism in *S. meliloti*. Based on phylogenetic evidence, this may be a novel pathway that is present in a number of  $\alpha$ -proteobacteria (Geddes *et al.*, 2013).

During our characterization of the catabolism of some polyols, synthesis of the literature on polyol metabolism in *S. meliloti* allowed us to predict a genetic locus that may be involved in the catabolism of sorbitol, mannitol and D-arabitol (see sections 1.4.5.2., and 3.4). It would be of interest to genetically define this locus and verify these predictions to further define polyol catabolism in *S. meliloti*. Mannitol is well-known for its ability to enrich for rhizobia during environmental isolation, and is considered a favorable carbon source for rhizobia (Stowers, 1985; Vincent, 1970). Genetic characterization of this locus could also lead to evaluating the influence of the ability to utilize mannitol, as well as sorbitol and D-arabitol on competitiveness for nodule occupancy.

A genetic tool was developed and used during the characterization of adonitol and L-arabitol catabolism to predict their entry-points into central metabolism. Essentially, this involved using strains that were unable to grow using glycolysis (pyruvate carboxylase mutant), or gluconeogenesis (double triose-phosphate isomerase mutant) to determine if carbon sources entered central metabolism through the TCA cycle, such as arabinose, or were catabolized glycolytically, such as glucose. This allowed us to predict that adonitol and L-arabitol entered central metabolism through the PP pathway, rather than the TCA cycle like arabinose.

One particularly interesting finding was that a strain carrying mutations in pyruvate carboxylase, and both triose-phosphate isomerases was still able to grow

utilizing ribose and xylose as sole carbon sources (Geddes & Oresnik, 2012a). The ability to grow around such blocks in metabolism suggests that the PP pathway is not the exclusive route of catabolism for these pentoses. This result supported the previous suggestion of a ribose and xylose “metabolic bypass” in *S. meliloti* based on the ability of ribose kinase and xylose isomerase mutants to continuously incorporate  $^{14}\text{C}$  ribose and  $^{14}\text{C}$  xylose at rates consistent with the wild type (Duncan, 1981). The presence of an alternate pathway that could convert ribose or xylose into a TCA intermediate or acetate would be consistent with these genetic and physiological data. There is precedence for a number of such pathways for xylose catabolism (Dahms, 1974; Heath *et al.*, 1958; Stephens *et al.*, 2007). We attempted to isolate mutants in genes that may encode this metabolic bypass by performing a Tn5 mutagenesis in a mutant background that lacked pyruvate carboxylase and both triose-phosphate isomerases, but despite screening over 10,000 colonies, no candidate mutants were isolated. It would be of interest to continue to attempt to define the metabolic bypass for ribose and xylose catabolism in *S. meliloti*. Given the massive endoreduplication that occurs during bacteroid development, the demand for the biosynthesis of ribose from TCA cycle intermediates must be immense. It is possible that the metabolic bypass may contribute to the generation of ribose for nucleotide biosynthesis in the bacteroid during growth on TCA cycle intermediates.

Another objective of this thesis was to identify and characterize a genetic locus for galactose catabolism. Using a Tn5 mutagenesis screen for the inability to grow using galactose as a sole carbon source we isolated a strain carrying a mutation in *dgoK1* (Geddes & Oresnik, 2012b). The annotations of the other genes in the locus that contained *dgoK1* were consistent with the capacity of this locus to encode the De Ley-

Doudoroff pathway of galactose catabolism in its entirety. Genetic characterization of the locus revealed that only genes predicted to encode the two final steps of the De Ley-Doudoroff pathway were indispensable for galactose catabolism: *dgoK1* and *dgoA* (Geddes & Oresnik, 2012b). Genes predicted to encode galactonolactonase (*SMc00883*) and galactonate dehydratase (*ilvD1*) were not required for growth using galactose as a sole carbon source. A galactose dehydrogenase was shown to be encoded by *fabG2* (renamed *galD*), however the galactose-inducible activity of a second galactose/L-arabinose dehydrogenase was detected (Geddes & Oresnik, 2012b).

It was hypothesized that the redundant galactose/L-arabinose dehydrogenase may be encoded by *SMc00588* based on similarity to L-arabinose/galactose dehydrogenases of *A. brasiliense* and *A. tumefaciens* (Watanabe *et al.*, 2006; Zhao & Binns, 2011). A targeted mutation was constructed in *SMc00588*, and it was shown to encode the galactose/L-arabinose dehydrogenase we had previously detected. A strain carrying a mutation in *SMc00588* was unable to grow using L-arabinose as a sole carbon source. Previous work characterizing a genetic locus of L-arabinose catabolism was unable to identify an L-arabinose dehydrogenase (Poysti *et al.*, 2007). Our data is consistent with the hypothesis that L-arabinose dehydrogenase is encoded by *SMc00588* in *S. meliloti*. It would be of interest to characterize this strain for symbiotic phenotypes for two reasons. First, an L-arabinose mutant has been previously suggested to give rise to nodules that did not fix nitrogen on alfalfa (Duncan, 1981); mutants of the previously characterized arabinose locus showed no symbiotic phenotype (Poysti *et al.*, 2007). Second, a binding site for the response regulator ChvI, of the key symbiotic two-component system ExoS/ChvI has been identified upstream of *SMc00588* (Bélanger & Charles, 2013).

Future work should be directed at determining if this gene is important for nitrogen fixation or competitiveness for nodule occupancy in *S. meliloti*.

It is unclear whether the putative galactonolactonase SMc00883, and dehydratase IlvD1 do or do not contribute to galactose catabolism in *S. meliloti*. One hypothesis is that they are used for the De Ley-Doudoroff pathway but their activities are redundant with other activities in the genome, as is the case for GalD. An approach that could be taken to address this hypothesis is performing a Tn5 mutagenesis screen for the inability to utilize galactose as a carbon source in backgrounds that contain mutations of SMc00883 and *ilvD1*. Early results using this approach isolated a Tn5 mutant in SMa0203 that was unable to utilize galactose, only in backgrounds that also contained a mutation in SMc00883. SMa0203 appears to be the first gene in an operon that encodes an ABC transporter and a putative gluconolactonase (SMa0196). The possibility that SMa0196 is a galactonolactonase that is redundant with SMc00883 for galactose catabolism should be addressed in the future.

A dedicated galactose transporter was not identified during these studies. Rather we showed that galactose was able to compete for transport through the arabinose ABC transporter AraABC, as is glucose (Geddes & Oresnik, 2012b). A reduced growth rate of strains carrying *araA* mutations using galactose as a sole carbon source suggests that some transport of galactose through this transporter is biologically relevant, but other transport systems must also be present that allow the transport of galactose and glucose.

A common theme that has emerged throughout this thesis is the ability of multiple sugars to compete for transport through ABC transporters. Erythritol, adonitol and L-arabitol were shown to compete for transport through MptABCD (Geddes & Oresnik,

2012a), and galactose, glucose and L-arabinose for AraABC (Geddes & Oresnik, 2012b). Similar phenomena have been reported for the fructose ABC transporter with ribose and mannose, and  $\alpha$ -glucoside transporters in *S. meliloti* (Glenn & Dilworth, 1981; Lambert *et al.*, 2001). More research is required to determine if competition is representative of transport by these ABC transporters. For example, the maltose binding protein of *E. coli* forms hydrogen bonds with malto-oligosaccharides at the reducing end, whereas the transmembrane segment of the maltose ABC transporter binds three glycosyl units from the non-reducing end of the sugar. Therefore some modified malto-oligosaccharides are unable to be transported by the system despite having a high binding affinity for maltose binding protein (Oldham *et al.*, 2013).

The regulation of galactose catabolism remains unclear. We observed typical induction of the galactose locus and *araABC* transporter by galactose, but a strain carrying a mutation in *dgoKI* showed unusual regulatory phenomena. Constitutive expression of the galactose locus, and enhanced induction of the *araABC* locus by galactose and seed exudate was observed in this strain (Geddes & Oresnik, 2012b). One hypothesis for the constitutive expression of the galactose locus in this mutant is that there is a small intracellular pool of galactose that is continually being turned over by the De Ley-Doudoroff pathway during free-living growth. A mutation in the De Ley-Doudoroff pathway may result in an accumulation of galactose and subsequent induction of the galactose locus. Enhanced induction of *araABC* by galactose is easier to conceive, as galactose would be transported into the cell and not completely metabolized under these conditions. Galactose or another inducing metabolite may then accumulate and result in enhanced induction of *araABC*. More work is needed to define the

transcriptional regulators and inducing molecules that regulate galactose catabolism in *S. meliloti*.

One of the primary objectives during throughout this thesis was to consider the contribution of carbon metabolism to competitiveness for nodule occupancy in *S. meliloti*. We hypothesized that the inability to catabolize erythritol would lead to a reduced ability to compete for nodule occupancy based on its effects on the nodulation competitiveness of *R. leguminosarum* and virulence of *Brucella* spp.. We determined that the inability to catabolize erythritol did not lead to a reduced competitiveness for nodule occupancy in *S. meliloti* (Geddes *et al.*, 2010). This phenotype is of interest because other reduced competitiveness for nodule occupancy phenotypes have been conserved between organisms: rhamnose and inositol have been shown to be important for competition for nodule occupancy in both *S. meliloti* and *R. leguminosarum* (Fry *et al.*, 2001; Kohler *et al.*, 2010; Oresnik *et al.*, 1998) (Rivers and Oresnik, unpublished). It is unclear whether the differences in the influence of erythritol catabolism on competitiveness for nodule occupancy are the result of bacteria-derived or host-derived effects. The rhizosphere sugar profiles from *M. sativa* and *P. sativum* are significantly different based on the difference in induction profiles of genes involved in sugar metabolism in *R. leguminosarum* grown in the rhizospheres of these legumes (Ramachandran *et al.*, 2011). Although, increased expression of erythritol genes were not observed in either case (Ramachandran *et al.*, 2011).

Moving forward, completion for nodule occupancy experiments should be performed on a broader range of legume hosts. For example, competition for nodule occupancy experiments could be performed using erythritol mutants on the hosts *M.*

*truncatula*, *Melilotus* and *Trigonella*. This may help define whether competitiveness for nodule occupancy is a legume-derived trait or differential based on rhizobial species. Additionally, studying the influence of carbon catabolism on competition for nodule occupancy into *S. fredii* would provide a unique opportunity to study the diversity of host-dependent effects on competitiveness for nodule occupancy. *S. fredii* NGR234 has a sequenced genome, and is able to nodulate legumes of 112 different genera, and 1 non-legume, *Parasponia andersonii* (Pueppke & Broughton, 1999; Schmeisser *et al.*, 2009).

A notable observation that came from a synthesis of the literature for nodulation competitiveness and carbon metabolism in rhizobia is that many carbon sources that have been shown to effect competition for nodulation in *R. leguminosarum* and *S. meliloti* enter central metabolism through DHAP. These include rhamnose, inositol, glycerol and erythritol, all of which have been shown to affect competitiveness for nodule occupancy in *R. leguminosarum* (Ding *et al.*, 2012; Fry *et al.*, 2001; Oresnik *et al.*, 1998; Yost *et al.*, 2006). Based on a DHAP entry-point of central metabolism for all these sugars, it could be expected that a triose-phosphate isomerase mutant would be dramatically impaired in competitiveness for nodule occupancy. Strikingly, a strain carrying a *tpiA* mutation in *S. meliloti* was not impaired in competitiveness for nodule occupancy (Poysti & Oresnik, 2007). It would be interesting to determine if this phenotype is conserved in *R. leguminosarum*.

It was hypothesized that the ability to catabolize galactose may be important for competitiveness for nodule occupancy in rhizobia, based on the evidence for the prevalence of galactose in the rhizosphere. In distinct contrast to our hypothesis, the inability to catabolize galactose proved to be advantageous towards competitiveness for



nodule occupancy in *S. meliloti* (Geddes & Oresnik, 2012b). We also showed that a strain that was unable to catabolize galactose was able to nodulate a proportion of an *M. sativa* population earlier than the wild type.

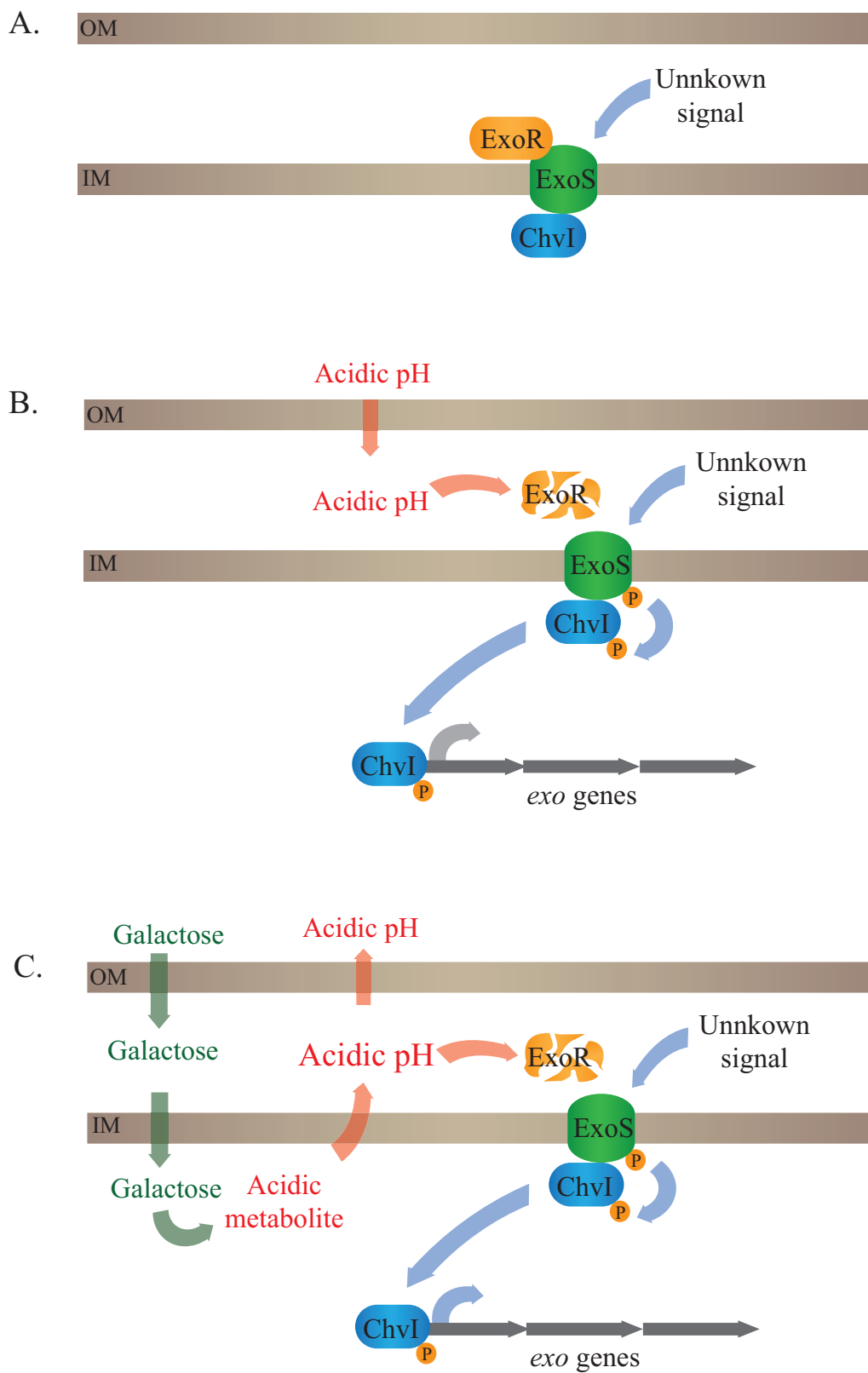
These phenotypes were of interest because of their contrast to other reported nodulation competitiveness phenotypes, wherein competition for nodule occupancy is typically reduced or unaffected. Therefore, we pursued an explanation for the increased competitiveness that resulted from the inability to catabolize galactose. We showed that the inability to catabolize galactose because of a block in *dgoK1* resulted in acidification of the culture medium of galactose mutants; medium acidification effected exopolysaccharide production in these mutants. These effects were manifested in the early production of succinoglycan in a wild-type background, as well as reduced production of EPSII in an *expR*<sup>+</sup> background. We showed that effects on competition for nodule occupancy in *S. meliloti* strains correlated with the effects on exopolysaccharide production we observed. This led us to hypothesize that increased competitiveness for nodule occupancy by a strain carrying a *dgoK1* mutation resulted from enhanced succinoglycan production during invasion.

These observations, coupled with current literature in *S. meliloti* and *A. tumefaciens* led us to hypothesized that the CCRH generated during *S. meliloti* invasion of *Medicago* spp. may be an acidic environment. We were able to provide evidence to support this hypothesis based on the co-localization of the acidotropic dye Lysotracker Red DND-99 with GFP expression *S. meliloti* in the CCRH. Lysotracker co-localization is a qualitative indicator of an acidic compartment. These results correlate well with the current understanding of the regulation of succinoglycan biosynthesis by the ExoS-ChvI

two-component system in *S. meliloti* as well as the homologous ChvG-ChvI two-component system in *A. tumefaciens* (Cheng & Walker, 1998a, b; Li *et al.*, 2002; Lu *et al.*, 2012; Wu *et al.*, 2012; Yao *et al.*, 2004). A model based on the synthesis of this literature and our results for a mechanism for the induction of succinoglycan biosynthesis in the CCRH is presented in Figure 7.1. Briefly, ExoS responds to an unknown signal as ExoR is transiently degraded by acidic pH; degradation of ExoR allows the phosphorylation and activation of ChvI by ExoS. ChvI modulates the transcription of regulatory targets, including the activation of succinoglycan biosynthesis. The production of succinoglycan facilitates successful invasion by *S. meliloti* and penetration of the plant-derived infection thread. The acidic pH of the CCRH may contribute to the coordination of this induction. In a galactose mutant this process may be exacerbated, resulting in either increased, or earlier succinoglycan production that may facilitate a greater likelihood of a successful invasion event by *S. meliloti*.

Several questions remain unanswered about the nature of the acidic pH we observed in the CCRH. It is unclear whether acidic pH in the CCRH is persistent or transient. The nature of the pH inside the IT also remains unclear, as we did not observe co-localization of LysoTracker with the IT. The molecular mechanisms behind the acidification of the CCRH, and whether these mechanisms are derived from the plant, the bacteria, or both remain to be elucidated. It is tempting to speculate that acidification of the environment by *S. meliloti* in response to hexoses, perhaps mediated by periplasmic PQQ-linked glucose dehydrogenase, contributes to acidification in the CCRH.

**Figure 7.1.** Model for the activation of succinoglycan synthesis by acidic pH.



**A.** Induction of the ExoS-ChvI two-component system (by an unknown signal) is inhibited by physical interactions of the periplasmic protein ExoR with the sensor kinase ExoS (Chen *et al.*, 2008). **B.** ExoR proteolysis allows the induction of the ExoS-ChvI system. An unknown signal triggers the autophosphorylation of ExoS, leading to the phosphorylation of ChvI. ChvI-P activates the transcription of the genes for succinoglycan biosynthesis, and regulates other targets (not shown) (Chen *et al.*, 2008; Lu *et al.*, 2012; Yao *et al.*, 2004). The signal that leads to ExoR proteolysis may be acidic pH as has been shown for the ChvG-ChvI system in *A. tumefaciens* (Wu *et al.*, 2012). **C.** Model for the activation of succinoglycan gene expression in a galactose mutant. The inability to catabolize galactose leads to acidification by an unknown metabolite. Acidic pH may activate the ExoS-ChvI system by resulting in the degradation of ExoR, as in B. Early or enhanced activation of the system in response to galactose on the root surface may result in an increased competitiveness for nodule occupancy.

Future investigations of the influence of carbon metabolism on competition for nodule occupancy in rhizobia should be directed towards providing context to increased competitiveness. The following key question should be addressed: At what stage of the symbiotic process is altered competitiveness manifested? Competitiveness in the rhizosphere and on the root surface could be assessed by quantification of the growth of rhizobial strains in these environments over time. Competitiveness at the level of successful IT formation can be assessed using fluorescent microscopy by quantifying CCRH and IT formation by differentially labeled rhizobial strains. Gage *et al.* (1996) were also able to quantify the growth rate of *S. meliloti* in the IT. It would be interesting to perform similar analysis on strains that have been hypothesized to be affected in their ability to grow in the IT such as rhamnose mutants (Oresnik *et al.*, 1998). The presence of sugars at the various stages of infection could also be monitored using promoters of loci for sugar metabolism fused to reporter genes.

It is also important to consider the need for continued directed functional characterization arising from the growing trend to assign function based on homology. As described in section 1, many paradigm differences exist between *S. meliloti* and other model organisms. Therefore establishing models for catabolism in different bacterial lineages will help refine the accuracy of functional association from bacteria in similar lineages. The large metabolic potential of *S. meliloti* also has the advantage of allowing the characterization of pathways of peripheral metabolism that may not be present in model bacteria with smaller metabolic capacity.

From a biological standpoint several key questions need to be addressed in the future with respect to carbon metabolism. Transcriptome studies such as those performed

by Ramachandran *et al.* (2011) should be performed on *S. meliloti* during interaction with its hosts to help push our understanding of the carbon experienced in the rhizosphere by *S. meliloti* beyond a number of disparate phenotypes. As studies like this move further it is clear that it is also critical to consider the rhizosphere in a spatial and temporal manner. The profile of compounds present in root exudates of *Arabidopsis* plants has recently been shown to change dramatically during plant development, based on both direct quantitation of compounds and global expression profiles of bacteria in the rhizosphere (Chaparro *et al.*, 2013). Exudate profile is also variable spatially along the axis of the root (Herron *et al.*, 2013). This seems an important consideration, particular since rhizobia are biased towards invasion of specific zones of the plant root (Bauer, 1981). Biosensors for specific carbon sources are a new promising technology that may help with studies such as this (Bourdès *et al.*, 2012; Herron *et al.*, 2013). Moving further into the future, the influence of the microbial community in the soil environment on exudate profiles and the response of rhizobia to these changes will prove insightful.

Overall the work presented in this thesis has contributed to our understanding of carbon metabolism and competition for nodule occupancy in *S. meliloti*. Specifically, this was accomplished through the characterization of genetic loci for erythritol, adonitol and L-arabitol catabolism, as well as for galactose catabolism. The increased understanding of carbon metabolism in *S. meliloti* will provide an important framework of functional characterization for future studies to build upon. Moreover, the characterization of a carbon catabolism mutant led us to make inferences about *S. meliloti* symbiosis that help further the understanding of plant-microbe interactions in the *S. meliloti*-*M. truncatula* model system. This finding is a proof of principle for the ability

of functional characterization to contribute to a more broad understanding of the biology of an organism. Future interest lies in the continued characterization of carbon metabolism in rhizobia, and its contribution to the invasion of legume hosts.

### **Literature Cited**



Adhya, S. L. & Shapiro, J. A. (1969). The galactose operon of *E. coli* K-12. I. Structural and pleiotropic mutations of the operon. *Genetics* 62, 231-247.

Alexeyev, M. (1999). The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of Gram-negative bacteria. *Biotechniques* 26, 824-828.

Allan Downie, J. (1994). Signalling strategies for nodulation of legumes by rhizobia. *Trends Microbiol* 2, 318-324.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J. H., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389-3402.

Alvarez-Martinez, M., Machold, J., Weise, C., Schmidt-Eisenlohr, H., Baron, C. & Rouot, B. (2001). The *Brucella suis* homologue of the *Agrobacterium tumefaciens* chromosomal virulence operon *chvE* is essential for sugar utilization but not for survival in macrophages. *J Bacteriol* 183, 5343-5351.

Amor, B. B., Shaw, S. L., Oldroyd, G. E., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Dénarié, J. & Gough, C. (2003). The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J* 34, 495-506.

Ampomah, O. Y., A., A., Hansen, E., Svenson, J., Huser, T., Jensen, J. B. & Bhuvaneswari, T. V. (2013). The *thuEFGKAB* operon of *Rhizobia* and *Agrobacterium tumefaciens* codes for transport of trehalose, maltitol, and isomers of sucrose and their assimilation through the formation of their 3-keto derivatives. *J Bacteriol* 195, 3797-3807.

Anderson, R. L. & Sapico, V. L. (1975). D-fructose (D-mannose) kinase. *Methods Enzymol* 42, 39-43.

Ané, J.-M., Kiss, G. B., Riely, B. K., Penmetsa, R. V., Oldroyd, G. E., Ayax, C., Lévy, J., Debellé, F., Baek, J.-M. & other authors (2004). *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. *Science* 303, 1364-1367.

Aneja, P. & Charles, T. C. (1999). Poly-3-hydroxybutyrate degradation in *Rhizobium* (*Sinorhizobium*) *meliloti*: isolation and characterization of a gene encoding 3-hydroxybutyrate dehydrogenase. *J Bacteriol* 181, 849-857.

Aneja, P., Zachertowska, A. & Charles, T. (2005). Comparison of the symbiotic and competition phenotypes of *Sinorhizobium meliloti* PHB synthesis and degradation pathway mutants. *Canadian journal of microbiology* 51, 599-604.

Ardisson, S., Kobayashi, H., Kambara, K., Rummel, C., Noel, K. D., Walker, G. C., Broughton, W. J. & Deakin, W. J. (2011). Role of BacA in lipopolysaccharide synthesis,

peptide transport, and nodulation by *Rhizobium* sp. strain NGR234. *J Bacteriol* 193, 2218-2228.

Ardourel, M., Lortet, G., Maillet, F., Roche, P., Truchet, G., Promé, J. C. & Rosenberg, C. (1995). In *Rhizobium meliloti*, the operon associated with the *nod* box n5 comprises *nodL*, *noeA* and *noeB*, three host- range genes specifically required for the nodulation of particular *Medicago* species. *Mol Microbiol* 17, 687-699.

Arias, A. & Martinez de Drets, G. (1976). Glycerol metabolism in *Rhizobium*. *Can J Microbiol* 22, 150-153.

Arias, A. & Cerveñansky, C. (1982). Transport and catabolism of D-mannose in *Rhizobium meliloti*. *J Bacteriol* 151, 1069–1072.

Arias, A. & Cerveñansky, C. (1986). Galactose Metabolism in *Rhizobium meliloti* L5-30. *J Bacteriol* 167, 1092-1094.

Arias, A., Cerveñansky, C., Gardiol, A. & Martinez de Drets, G. (1979). Phosphoglucose isomerase mutant of *Rhizobium meliloti*. *J Bacteriol* 137, 409-414.

Atkinson, E. M., Palcic, M. M., Hindsgaul, O. & Long, S. R. (1994). Biosynthesis of *Rhizobium meliloti* lipooligosaccharide Nod factors: NodA is required for an N-acyltransferase activity. *Proc Natl Acad Sci USA* 91, 8418-8422.

Avetisyan, A., Jensen, J. B. & Huser, T. (2013). Monitoring trehalose uptake and conversion by single bacteria using laser tweezers Raman spectroscopy. *Anal Chem*, 7624-7270.

Barnett, M. J., Toman, C. J., Fisher, R. F. & Long, S. R. (2004). A dual-genome Symbiosis-Chip for coordinate study of signal exchange and development in a prokaryote-host interaction *Proc Natl Acad Sci USA* 101, 16636-16641.

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F. & other authors (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc Natl Acad Sci USA* 98, 9883-9888.

Barra, L., Bowser, L., Pica, N., Gouffi, K., Walker, G. C., Blanco, C. & Trautwetter, A. (2003). Glucose 6-phosphate dehydrogenase is required for sucrose and trehalose to be efficient osmoprotectants in *Sinorhizobium meliloti*. *FEMS Microbiol Lett* 229, 183-188.

Barsch, A., Carvalho, H. G., Cullimore, J. V. & Niehaus, K. (2006). GC–MS based metabolite profiling implies three interdependent ways of ammonium assimilation in *Medicago truncatula* root nodules. *J Biotech* 127, 79-83.

Battisti, L., Lara, J. C. & Leigh, J. A. (1992). Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. *Proc Natl Acad Sci USA* 89, 5625-5629.

- Batut, J., Andersson, S. G. E. & O'Callaghan, D. (2004). The evolution of chronic infections strategies in the  $\alpha$ -proteobacteria. *Nature Rev Microbiol* 2, 933-945.
- Bauer, W. D. (1981). Infection of legumes by rhizobia. *Ann Rev Plant Physiol* 32, 407-449.
- Bélangier, L. & Charles, T. C. (2013). Members of the *Sinorhizobium meliloti* ChvI regulon identified by a DNA binding screen. *BMC Microbiology* 13, 132.
- Bélangier, L., Dimmick, K. A., Fleming, J. S. & Charles, T. C. (2009). Null mutations in *Sinorhizobium meliloti* *exoS* and *chvI* demonstrate the importance of this two-component regulatory system for symbiosis. *Mol Micro* 74, 1223-1237.
- Benson, D. R. & Silvester, W. (1993). Biology of *Frankia* strains, actinomycete symbionts of actinorhizal plants. *Microbiol Rev* 57, 293-319.
- Beringer, J. E., Beynon, J. L., Buchanan-Wollason, A. V. & Johnston, A. W. B. (1978). Transfer of the drug resistance transposon Tn5 to *Rhizobium*. *Nature* 276, 633-634.
- Bernardelli, C. E., Luna, M. F., Galar, M. L. & Boiardi, J. L. (2001). Periplasmic PQQ-dependent glucose oxidation in free-living and symbiotic rhizobia. *Curr Microbiol* 42, 310-315.
- Bernardelli, C. E., Luna, M. F., Galar, M. L. & Boiardi, J. L. (2008). Symbiotic phenotype of a membrane-bound glucose dehydrogenase mutant of *Sinorhizobium meliloti*. *Plant and Soil* 313, 217-225.
- Bolton, E., Higgisson, B., Harrington, A. & O'Gara, F. (1986). Dicarboxylic acid transport in *Rhizobium meliloti*: isolation of mutants and cloning of dicarboxylic acid transport genes. *Arch Microbiol* 144, 142-146.
- Bourdès, A., Rudder, S., East, A. K. & Poole, P. S. (2012). Mining the *Sinorhizobium meliloti* transportome to develop FRET biosensors for sugars, dicarboxylates and cyclic polyols. *PloS one* 7, e43578.
- Boussau, B., Karlberg, E. O., Frank, A. C., Legault, B. & Andersson, S. G. E. (2004). Computational inference of scenarios for  $\alpha$ -proteobacterial genome evolution. *Proc Natl Acad Sci USA* 101, 9722-9727.
- Breedveld, M., Zevenhuizen, L. & Zehnder, A. (1990). Osmotically induced oligo- and polysaccharide synthesis by *Rhizobium meliloti* SU-47. *J Gen Appl Microbiol* 136, 2511-2519.
- Breedveld, M. W., Dijkema, C., Zevenhuizen, L. P. & Zehnder, A. J. (1993). Response of intracellular carbohydrates to a NaCl shock in *Rhizobium leguminosarum* biovar *trifolii* TA-1 and *Rhizobium meliloti* SU-47. *J Gen Microbiol* 139, 3157-3163.

- Bringhurst, R. M. & Gage, D. G. (2000). An AraC-like transcriptional activator is required for induction of genes needed for  $\alpha$ -galactoside utilization in *Sinorhizobium meliloti*. *FEMS Microbiol Lett* 188, 23-27.
- Bringhurst, R. M. & Gage, D. J. (2002). Control of inducer accumulation plays a key role in succinate-mediated catabolite repression in *Sinorhizobium meliloti*. *J Bacteriol* 184, 5385-5392.
- Bringhurst, R. M., Cardon, Z. G. & Gage, D. J. (2001). Galactosides in the rhizosphere: Utilization by *Sinorhizobium meliloti* and development as a biosensor. *Proc Natl Acad Sci USA* 98, 4540-4545.
- Brom, S., de los Santos, A. G., Stepkowsky, T., Flores, M., Davila, G., Romero, D. & Palacios, R. (1992). Different plasmides of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. *J Bacteriol* 174, 5183-5189.
- Broughton, W., Wong, C., Lewin, A., Samrey, U., Myint, H., Meyer, H., Dowling, D. & Simon, R. (1986). Identification of *Rhizobium* plasmid sequences involved in recognition of *Psophocarpus*, *Vigna*, and other legumes. *J Cell Biol* 102, 1173-1182.
- Burkhardt, S., Jiménez de Bagüés, M. P., Liautard, J. P. & Kohler, S. (2005). Analysis of the behaviour of *eryC* mutants of *Brucella suis* attenuated in macrophages. *Infect Immun* 73, 6782-6790.
- Cabanes, D., Boistard, P. & Batut, J. (2000). Symbiotic induction of pyruvate dehydrogenase genes from *Sinorhizobium meliloti*. *Mol Plant Microbe Interact* 13, 483-493.
- Campbell, G. R., Reuhs, B. L. & Walker, G. C. (2002). Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. *Proc Natl Acad Sci USA* 99, 3938-3943.
- Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M. & other authors (2001). Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc Natl Acad Sci USA* 98, 9877-9882.
- Catoira, R., Galera, C., de Billy, F., Penmetsa, R. V., Journet, E.-P., Maillet, F., Rosenberg, C., Cook, D., Gough, C. & other authors (2000). Four genes of *Medicago truncatula* controlling components of a Nod factor transduction pathway. *Plant Cell* 12, 1647-1665.
- Cerveñansky, C. & Arias, A. (1984). Glucose-6-phosphate dehydrogenase deficiency in pleiotropic carbohydrate-negative mutant strains of *Rhizobium meliloti*. *J Bacteriol* 160, 1027-1030.
- Chain, P. S., Lang, D. M., Comerchi, D. J., Malfatti, S. A., Vergez, L. M., Shin, M., Ugalde, R. A., Garcia, E. & Tolmasky, M. E. (2011). Genome of *Ochrobactrum anthropi*

ATCC 49188 T, a versatile opportunistic pathogen and symbiont of several eukaryotic hosts. *J Bacteriol* 193, 4274-4275.

Chaparro, J. M., Badri, D. V., Bakker, M. G., Sugiyama, A., Manter, D. K. & Vivanco, J. M. (2013). Root exudation of phytochemicals in *Arabidopsis* follows specific patterns that are developmentally programmed and correlate with soil microbial functions. *PLoS one* 8, e55731.

Charles, T. C. & Finan, T. M. (1991). Analysis of a 1600-kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated *in vivo*. *Genetics* 127, 5-20.

Charles, T. C. & Nester, E. W. (1993). A chromosomally encoded two-component sensory transduction system is required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol* 175, 6614-6625.

Charles, T. C., Singh, R. S. & Finan, T. M. (1990). Lactose utilization and enzymes encoded in *Rhizobium meliloti*: implications for population studies. *J Gen Microbiol* 136, 2497-2502.

Charles, T. C., Doty, S. L. & Nester, E. W. (1994). Construction of *Agrobacterium* strains by electroporation of genomic DNA and its utility in analysis of chromosomal virulence mutants. *Appl Environ Microbiol* 60, 4192-4194.

Charles, T. C., Cai, G.-q. & Aneja, P. (1997). Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* 146, 1211-1220.

Chen, E. J., Sabio, E. A. & Long, S. R. (2008). The periplasmic regulator ExoR inhibits ExoS/ChvI two- component signalling in *Sinorhizobium meliloti*. *Mol Micro* 69, 1290-1303.

Cheng, H.-P. & Walker, G. C. (1998a). Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *J Bacteriol* 180, 20-26.

Cheng, H.-P. & Walker, G. C. (1998b). Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* 180, 5183-5191.

Clark, S. R. D., Oresnik, I. J. & Hynes, M. F. (2001). RpoN of *Rhizobium leguminosarum* bv. *viciae* strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation. *Mol Gen Genet* 264, 623-633.

Conway, T. (1992). The Entner-Doudoroff pathway: history, physiology and molecular biology. *FEMS Microbiol Rev* 9, 1-27.

Cordoba, E., Shishkova, S., Vance, C. P. & Hernández, G. (2003). Antisense inhibition of NADH glutamate synthase impairs carbon/nitrogen assimilation in nodules of alfalfa (*Medicago sativa* L.). *Plant J* 33, 1037-1049.

Crossman, L. C., Castillo-Ramírez, S., McAnnula, C., Lozano, L., Vernikos, G. S., Acosta, J. L., Ghazoui, Z. F., Hernández-Lucas, I., Meakin, G. & other authors (2008). A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. *PLoS ONE* 3, e2567.

Dahms, A. S. (1974). 3-deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. *Biochem Biophys Res Commun* 60, 1433-1439.

Dahms, A. S. & Anderson, R. L. (1969). 2-keto-3-deoxy-L-arabonate aldolase and its role in a new pathway of L-arabinose degradation. *Biochem Biophys Res Commun* 36, 809-814.

Dahms, A. S. & Anderson, R. L. (1972a). D-Fucose metabolism in a Pseudomonad: II. Oxidation of D-fucose to D-fucono- $\gamma$  lactone by an L-aranino-aldolase dehydrogenase and hydrolysis of the lactone by a lactonase. *J Biol Chem* 247, 2228-2232.

Dahms, A. S. & Anderson, R. L. (1972b). D-Fucose metabolism in a Pseudomonad: III. Conversion of D-fuconate to 2-keto-3-deoxy-D-fuconate by a dehydratase. *J Biol Chem* 247, 2233-2237.

Dahms, A. S. & Anderson, R. L. (1972c). D-Fucose metabolism in a Pseudomonad: I. Oxidation of D-fucose to  $\gamma$ -fuconolactone by a D-aldohexose dehydrogenase. *J Biol Chem* 247, 2222-2227.

Dahms, A. S. & Anderson, R. L. (1972d). D-Fucose metabolism in a Pseudomonad: IV. Cleavage of 2-keto-3-deoxy-D-fuconate to pyruvate and d-lactate by 2-keto-3-deoxy-arabonate aldolase. *J Biol Chem* 247, 2238-2241.

Day, D. A., Poole, P., Tyerman, S. D. & Rosendahl, L. (2001). Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. *Cell Mol Life Sci* 58, 61-71.

De Ley, J. & Doudoroff, M. (1957). The metabolism of D-galactose in *Pseudomonas saccharophila*. *J Biol Chem* 227, 745-757.

de Philip, P., Batut, J. & Boistard, P. (1990). *Rhizobium meliloti* FixL is an oxygen sensor and regulates *R. meliloti* *nifA* and *fixK* genes differently in *Escherichia coli*. *J Bacteriol* 172, 4255-4262.

de Sanctis, D., McVey, C. E., Enguita, F. J. & Carrondo, M. A. (2009). Crystal structure of the full-length sorbitol operon regulator SorC from *Klebsiella pneumoniae*: structural evidence for a novel transcriptional regulation mechanism. *J Mol Biol* 387, 759-770.

- De Vos, G. F., Walker, G. C. & Signer, E. R. (1986). Genetic manipulations in *Rhizobium meliloti* utilizing two new transposon Tn5 derivatives. *Mol Gen Genet* 204, 485-491.
- de Vries, G. E., van Brussel, A. & Quispel, A. (1982). Mechanism of regulation of glucose transport in *Rhizobium leguminosarum*. *J Bacteriol* 149, 872-879.
- de Werra, P., Péchy-Tarr, M., Keel, C. & Maurhofer, M. (2009). Role of gluconic acid production in the regulation of biocontrol traits of *Pseudomonas fluorescens* CHA0. *Appl Environ Microbiol* 75, 4162-4174.
- DeVecchio, V. G., Kapatral, V., Redkar, R. J., Patra, G., Mujer, C., Los, T., Ivanova, N., Anderson, I., Bhattacharyya, A. & other authors (2002). The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. *Proc Natl Acad Sci USA* 99, 443-448.
- Demont, N., Debellé, F., Aurelle, H., Dénarié, J. & Promé, J. (1993). Role of the *Rhizobium meliloti* *nodF* and *nodE* genes in the biosynthesis of lipo-oligosaccharidic nodulation factors. *J Biol Chem* 268, 20134-20142.
- diCenzo, G., Milunovic, B., Cheng, J. & Finan, T. M. (2012). tRNA<sup>arg</sup> and *engA* are essential genes on the 1.7-Mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. *J Bacteriol* 195, 202-212.
- Dickstein, R., Bisseling, T., Reinhold, V. N. & Ausubel, F. M. (1988). Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. *Genes Dev* 2, 677-687.
- Dilworth, M. J., Rynne, F. G., Castelli, J. M., Vivas-Marfisi, A. I. & Glenn, A. R. (1999). Survival and exopolysaccharide production in *Sinorhizobium meliloti* WSM419 are affected by calcium and low pH. *Microbiology* 145, 1585-1593.
- Ding, H., Yip, C. B., Geddes, B. A., Oresnik, I. J. & Hynes, M. F. (2012). Glycerol utilization by *Rhizobium leguminosarum* requires an ABC transporter and affects competition for nodulation. *Microbiology* 158, 1369-1378.
- Djordjevic, M. A. (2004). *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. *Proteomics* 4, 1859-1872.
- Domínguez-Ferreras, A., Soto, M. J., Pérez-Arnedo, R., Olivares, J. & Sanjuán, J. (2009). Importance of trehalose biosynthesis for *Sinorhizobium meliloti* osmotolerance and nodulation of alfalfa roots. *J Bacteriol* 191, 7490-7499.
- Doty, S. L., Chang, M. & Nester, E. W. (1993). The chromosomal virulence gene, *chvE*, of *Agrobacterium tumefaciens* is regulated by a LysR family member. *J Bacteriol* 175, 7880-7886.

- Driscoll, B. T. & Finan, T. M. (1993). NAD<sup>+</sup>-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. *Mol Microbiol* 7, 865-873.
- Driscoll, B. T. & Finan, T. M. (1996). NADP<sup>+</sup>-dependent malic enzyme of *Rhizobium meliloti*. *J Bacteriol* 178, 2224-2231.
- Driscoll, B. T. & Finan, T. M. (1997). Properties of NAD<sup>+</sup>- and NADP<sup>+</sup>-dependent malic enzymes of *Rhizobium (Sinorhizobium) meliloti* and differential expression of their genes in nitrogen-fixing bacteroids. *Microbiology* 143, 489-498.
- Duncan, M. J. (1981). Properties of Tn5-induced carbohydrate mutants in *Rhizobium meliloti*. *J Gen Microbiol* 122, 61-67.
- Duncan, M. J. & Fraenkel, D. G. (1979).  $\alpha$ -Ketoglutarate dehydrogenase mutant of *Rhizobium meliloti*. *J Bacteriol* 137, 415-419.
- Dunn, M. F. (1998). Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *FEMS Microbiol Rev* 22, 105-123.
- Dunn, M. F., Araíza, G. & Finan, T. M. (2001). Cloning and characterization of the pyruvate carboxylase from *Sinorhizobium meliloti* Rm1021. *Arch Microbiol*, 355-363.
- Dylan, T., Helinski, D. & Ditta, G. S. (1990a). Hypoosmotic adaptation in *Rhizobium meliloti* requires  $\beta$ -(1-2)-glucan. *J Bacteriol* 172, 1400-1408.
- Dylan, T., Nagpal, P., Helinski, D. & Ditta, G. S. (1990b). Symbiotic pseudorevertants of *Rhizobium meliloti* *ndv* mutants. *J Bacteriol* 172, 1409-1417.
- Dymov, S. I., Meek, D. J., Steven, B. & Driscoll, B. T. (2004). Insetion of transposon Tn5tac1 in the *Sinorhizobium meliloti* malate dehydrogenase (*mdh*) gene results in conditional polar effects on downstream TCA cycle genes. *Mol Plant Microbe Interact* 17, 1318-1327.
- Egelhoff, T., Fisher, R., Jacobs, T., Mulligan, J. & Long, S. (1985). Nucleotide sequence of *Rhizobium meliloti* 1021 nodulation genes: *nodD* is read divergently from *nodABC*. *DNA* 4, 241-248.
- Ehrhardt, D. W., Wais, R. & Long, S. R. (1996). Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. *Cell* 85, 673-681.
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P. & Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* 417, 962-966.
- Engelke, T., Jagadish, M. N. & Puhler, A. (1987). Biochemcial and genetical analysis of *Rhizobium meliloti* mutants defective in C<sub>4</sub>-dicarboxylate transport. *J Gen Microbiol* 133, 3019-3029.
- Ertl, G. & Jennings, J. (1991). Catalytic ammonia synthesis. *Plenum, New York*, 109-132.



- Fani, R., Brilli, M. & Lio, P. (2005). The origin and evolution of operons: the piecewise building of the proteobacterial histidine operon. *J Mol Evol* 60, 370-390.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783-789.
- Fennington, G. J. & Hughes, T. A. (1996). The fructokinase from *Rhizobium leguminosarum* biovar *trifolii* belongs to group I fructokinase enzymes and is encoded separately from other carbohydrate metabolism enzymes. *Microbiology* 142, 321-330.
- Ferguson, G. P., Datta, A., Baumgartner, J., Roop, R. M., Carlson, R. W. & Walker, G. C. (2004). Similarity to peroxisomal-membrane protein family reveals that *Sinorhizobium* and *Brucella* BacA affect lipid-A fatty acids. *Proc Natl Acad Sci USA* 101, 5012-5017.
- Finan, T. M., Wood, J. M. & Jordan, D. C. (1983). Symbiotic properties of C<sub>4</sub>-dicarboxylic acid transport mutants of *Rhizobium leguminosarum*. *J Bacteriol* 154, 1403-1413.
- Finan, T. M., Oresnik, I. & Bottacin, A. (1988). Mutants of *Rhizobium meliloti* defective in succinate metabolism. *J Bacteriol* 170, 3396-3403.
- Finan, T. M., Kunkel, B., de Vos, G. F. & Signer, E. R. (1986). Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J Bacteriol* 167, 66-72.
- Finan, T. M., McWhinnie, E., Driscoll, B. & Watson, R. J. (1991). Complex symbiotic phenotypes result from gluconeogenic mutations in *Rhizobium meliloti*. *Mol Plant Microbe Interact* 4, 386-392.
- Finan, T. M., Hartweg, E., Lemieux, K., Bergman, K., Walker, G. C. & Signer, E. R. (1984). General transduction in *Rhizobium meliloti*. *J Bacteriol* 159, 120-124.
- Finan, T. M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldau, G. A., Deegan, S., Walker, G. C. & Signer, E. R. (1985). Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. *Cell* 40, 869-877.
- Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F. J., Hernández-Lucas, I., Becker, A., Cowie, A. & other authors (2001). The complete sequence of the 1,683-kb pSymB megaplasmid from the N<sub>2</sub>-fixing endosymbiont *S. meliloti*. *Proc Natl Acad Sci USA* 98, 9889-9894.
- Fisher, R. F. & Long, S. R. (1993). Interactions of NodD at the *nod* box: NodD binds to two distinct sites on the same face of the helix and induces a bend in the DNA. *J Mol Biol* 233, 336-348.
- Fondi, M., Emiliani, G. & Fani, R. (2009). Origin and evolution of operons and metabolic pathways. *Res Microbiol* 69, 512-526.

- Foussard, M., Garnerone, A. M., Ni, F., Soupène, E., Boistard, P. & Batut, J. (1997). Negative autoregulation of the *Rhizobium meliloti* *fixK* gene is indirect and requires a newly identified regulator, *FixT*. *Mol Microbiol* 25, 27-37.
- Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J. & Ausubel, F. M. (1982). Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* 18, 289-296.
- Fry, J., Wood, M. & Poole, P. S. (2001). Investigation of myo-inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. *Mol Plant Microbe Interact* 14, 1016-1025.
- Fuhrer, T., Fisher, E. & Sauer, U. (2005). Experimental identification and quantification of glucose metabolism in seven bacterial species. *J Bacteriol*, 1581-1590.
- Gage, D. J. (2002). Analysis of infection thread development using Gfp-and DsRed-expressing *Sinorhizobium meliloti*. *J Bacteriol* 184, 7042-7046.
- Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev* 68, 280-300.
- Gage, D. J. & Long, S. R. (1998).  $\alpha$ -galactoside uptake in *Rhizobium meliloti*: isolation and characterization of *agpA*, a gene encoding a periplasmic binding protein required for melibiose and raffinose utilization. *J Bacteriol* 180, 5739-5748.
- Gage, D. J., Bobo, T. & Long, S. R. (1996). Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). *J Bacteriol* 178, 7159-7166.
- Galbraith, M. P., Feng, S. F., Borneman, J., Triplett, E. W., de Bruijn, F. J. & Rossbachl, S. (1998). A functional myo-inositol catabolism pathway is essential for rhizopine utilization by *Sinorhizobium meliloti*. *Microbiology* 144, 2915-2924.
- Galibert, F., Finan, T. M., Long, S. R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A. & other authors (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293, 668-672.
- Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P. & Sutton, M. A. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* 320, 889-892.
- Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C., Green, P. & other authors (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry* 70, 153-226.
- Gardiol, A., Arias, A., Cerveñansky, C. & Martinez de Drets, G. (1982). Succinate dehydrogenase mutant of *Rhizobium meliloti*. *J Bacteriol* 151, 1621-1623.

- Gardiol, A., Arias, A., Cerveñanský, C., Gaggero, C. & Martínez-Drets, G. (1980). Biochemical characterization of a fructokinase mutant of *Rhizobium meliloti*. *J Bacteriol* 151, 1621–1623.
- Geddes, B. A. & Oresnik, I. J. (2012a). Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol, and L-arabitol in *Sinorhizobium meliloti*. *Microbiology* 158, 2180–2191.
- Geddes, B. A., Hausner, G. & Oresnik, I. J. (2013). Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and Proteobacteria. *BMC Microbiology* 13, 46.
- Geddes, B. A., Pickering, B. S., Poysti, N. J., Yudistira, H., Collins, H. & Oresnik, I. J. (2010). A locus necessary for the transport and catabolism of erythritol in *Sinorhizobium meliloti*. *Microbiology* 156, 2970–2981.
- Geddes, G. A. & Oresnik, I. J. (2012b). Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. *J Bacteriol* 194, 5044–5505.
- Geer, L. Y., Domrachev, M., Lipman, D. J. & Bryant, S. H. (2002). CDART: protein homology by domain architecture. *Genome Res* 12, 1619–1623.
- Geremia, R. A., Mergaert, P., Geelen, D., Van Montagu, M. & Holsters, M. (1994). The NodC protein of *Azorhizobium caulinodans* is an N-acetylglucosaminyltransferase. *Proc Natl Acad Sci USA* 91, 2669–2673.
- Geremia, R. A., Cavaignac, S., Zorreguieta, A., Toro, N., Olivares, J. & Ugalde, R. (1987). A *Rhizobium meliloti* mutant that forms ineffective pseudonodules in alfalfa produces exopolysaccharide but fails to form  $\beta$ -(1-2) glucan. *J Bacteriol* 169, 880–884.
- Ghalambor, M. A. & Heath, E. C. (1962). The metabolism of L-fucose. II. The enzymatic cleavage of L-fuculose-1-phosphate. *J Biol Chem* 237, 2423–2426.
- Gilles-Gonzalez, M. A., Ditta, G. S. & Helinski, D. R. (1991). A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. *Nature* 350, 170–172.
- Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J. C., Jaubert, M., Simon, D., Cartieaux, F. & other authors (2007). Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* 316, 1307–1312.
- Glazebrook, J. & Walker, G. C. (1989). A novel exopolysaccharide can function in place of the calcofluor-binding exopolysaccharide in nodulation of alfalfa by *Rhizobium meliloti*. *Cell* 56, 661–672.
- Glazebrook, J. & Walker, G. C. (1991). Genetic techniques in *Rhizobium meliloti*. *Methods Enzymol* 204, 398–418.

- Glazebrook, J., Ichige, A. & Walker, G. (1993). A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes Dev* 7, 1485-1497.
- Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B. & Oldroyd, G. E. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* 441, 1149-1152.
- Glenn, A. R. & Dilworth, M. J. (1981). The uptake and hydrolysis of disaccharides by fast- and slow-growing species of *Rhizobium*. *Arch Microbiol* 129, 233-239.
- Glenn, S. A., Gurich, N., Feeney, M. A. & González, J. E. (2007). The ExpR/Sin Quorum-Sensing System Controls Succinoglycan Production in *Sinorhizobium meliloti*. *J Bacteriol* 189, 7077-7088.
- González, J. E., Reuhs, B. L. & Walker, G. C. (1996). Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. *Proc Natl Acad Sci USA* 93, 8636-8641.
- González, J. E., Semino, C. E., Wang, L.-X., Castellano-Torres, L. E. & Walker, G. C. (1998). Biosynthetic control of molecular weight in the polymerization of the octasaccharide subunits of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Proceedings of the National Academy of Sciences* 95, 13477-13482.
- Gordon, D. M., Ryder, M. H., Heinrich, K. & Murphy, P. J. (1996). An experimental test of the rhizopine concept in *Rhizobium meliloti*. *Appl Environ Microbiol* 62, 3991-3996.
- Gosselin, I., Wattraint, O., Riboul, D., Barbotin, J. & Portais, J. (2001). A deeper investigation on carbohydrate cycling in *Sinorhizobium meliloti*. *FEBS Lett*, 45-49.
- Halling, S. M., Peterson-Burch, B. D., Bricker, B. J., Zuerner, R. L., Quing, Z., Li, L., Kapur, V., Alt, D. P. & Olsen, S. C. (2005). Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. *J Bacteriol* 187, 2715-2726.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166, 557-570.
- Hanus, F. J., Maier, R. J. & Evans, H. J. (1979). Autotrophic growth of H<sub>2</sub>-uptake-positive strains of *Rhizobium japonicum* in an atmosphere supplied with hydrogen gas. *Proc Natl Acad Sci USA* 76, 1788-1792.
- He, F., Nair, G. R., Soto, C. S., Chang, Y., Hsu, L., Ronzone, E., DeGrado, W. F. & Binns, A. N. (2009). Molecular basis of ChvE function in sugar binding, sugar utilization, and virulence in *Agrobacterium tumefaciens*. *J Bacteriol* 191, 5802-5813.

- Heath, E., Hurwitz, J., Horecker, B. & Ginsburg, A. (1958). Pentose fermentation by *Lactobacillus plantarum* I. The cleavage of xylulose 5-phosphate by phosphoketolase. *J Biol Chem* 231, 1009-1029.
- Hellweg, C., Pühler, A. & Weidner, S. (2009). The time course of the transcriptomic response of *Sinorhizobium meliloti* 1021 following a shift to acidic pH. *BMC microbiology* 9, 37.
- Her, G. R., Glazebrook, J., Walker, G. C. & Reinhold, V. N. (1990). Structural studies of a novel exopolysaccharide produced by a mutant of *Rhizobium meliloti* strain Rm 1021. *Carbohydr Res* 198, 305-312.
- Herridge, D. F., Peoples, M. B. & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* 311, 1-18.
- Herron, P. M., Gage, D. J., Pinedo, C. A., Haider, Z. K. & Cardon, Z. G. (2013). Better to light a candle than curse the darkness: illuminating spatial localization and temporal dynamics of rapid microbial growth in the rhizosphere. *Front Plant Sci* 4.
- Heuel, H., Shakeri-Garakani, A., Turgut, S. & Lengeler, J. (1998). Genes for D-arabitol and ribitol catabolism from *Klebsiella pneumonia*. *Microbiology* 144, 1631-1639.
- Hirsch, S., Kim, J., Muñoz, A., Heckmann, A. B., Downie, J. A. & Oldroyd, G. E. (2009). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *Plant Cell* 21, 545-557.
- Hoang, H. H., Gurich, N. & González, J. E. (2008). Regulation of motility by the ExpR/Sin quorum-sensing system in *Sinorhizobium meliloti*. *J Bacteriol* 190, 861-871.
- Homma, K., Fukuchi, S., Gojobori, T. & Nishikawa, K. (2007). Gene cluster analysis method identifies horizontally transferred genes with high reliability and indicates that they provide the main mechanism of operon gain in 8 species of  $\gamma$ -proteobacteria. *Mol Biol Evol* 24, 805-813.
- Honma, M., Asomaning, M. & Ausubel, F. (1990). *Rhizobium meliloti* nodD genes mediate host-specific activation of nodABC. *J Bacteriol* 172, 901-911.
- Horváth, B., Yeun, L. H., Domonkos, Á., Halász, G., Gobbato, E., Ayaydin, F., Miró, K., Hirsch, S., Sun, J. & other authors (2011). *Medicago truncatula* IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. *Mol Plant Microbe Interact* 24, 1345-1358.
- House, B. L., Mortimer, M. W. & Kahn, M. L. (2004). New recombination methods for *Sinorhizobium meliloti* genetics. *Appl Environ Microbiol* 70, 2806-2815.
- Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L. & other authors (2009). InterPro: the integrative protein signature database. *Nucleic Acids Res* 37, D11-D215.

- Hynes, M. F. & O'Connell, M. P. (1990). Host plant effect on competition among strains of *Rhizobium leguminosarum*. *Can J Microbiol* 36, 864-890.
- Hynes, M. F. & McGregor, N. F. (1990). Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. *Mol Microbiol* 4, 567-574.
- Irigoyen, J. J., Sanchez-Diaz, M. & Emerich, D. W. (1990). Carbon metabolism enzymes of *Rhizobium meliloti* cultures and bacteroids and their distribution within alfalfa nodules. *Appl Environ Microbiol* 56, 2587-2589.
- Jacob, A. I., Adhamn, S. A. I., Capstick, D. S., Clark, S. R. D., Spence, T. & Charles, T. C. (2008). Mutational analysis of the *Sinorhizobium meliloti* short-chain dehydrogenase/reductase family reveals substantial contribution to symbiosis and catabolic diversity. *Mol Plant Microbe Interact* 21, 979-989.
- Jelesko, J. G. & Leigh, J. A. (1994). Genetic characterization of a *Rhizobium meliloti* lactose utilization locus. *Mol Microbiol* 11, 165-173.
- Jensen, J. B., Peters, N. K. & Bhuvaneswari, T. V. (2002). Redundancy in periplasmic binding protein-dependent transport systems for trehalose, sucrose, and maltose in *Sinorhizobium meliloti*. *J Bacteriol* 184, 2978-2986.
- Jensen, J. B., Ampomah, O. Y., Darrah, R., Peters, N. K. & Bhuvaneswari, T. V. (2005). Role of trehalose transport and utilization in *Sinorhizobium meliloti*-alfalfa interactions. *Mol Plant Microbe Interact* 18, 694-702.
- Jiang, G., Krishnan, A. H., Kim, Y.-W., Wacek, T. J. & Krishnan, H. B. (2001). A functional *myo*-inositol dehydrogenase gene is required for efficient nitrogen fixation and competitiveness of *Sinorhizobium fredii* USDA191 to nodulate soybean (*Glycine max* [L.] Merr.). *J Bacteriol* 183, 2595-2604.
- Jiang, J., Gu, B. H., Albright, L. M. & Nixon, T. B. (1989). Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum* *dct* genes. *J Bacteriol* 171, 5244-5253.
- Jiménez-Zurdo, J. I., van Dillewijn, P., Soto, M. J., de Filipe, M. R., Olivares, J. & Toro, N. (1995). Characterization of a *Rhizobium meliloti* proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. *Mol Plant Microbe Interact* 8, 492-498.
- John, M., Röhrig, H., Schmidt, J., Wieneke, U. & Schell, J. (1993). *Rhizobium* NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *Proc Natl Acad Sci USA* 90, 625-629.
- Jones, J. D. G. & Gutterson, N. (1987). An efficient mobilizable cosmid vector and its use in rapid marker exchange in *Pseudomonas fluorescens* strain HV37a. *Gene* 61, 299-306.

- Jones, K. M. (2012). Increased production of the exopolysaccharide succinoglycan enhances *Sinorhizobium meliloti* 1021 symbiosis with the host plant *Medicago truncatula*. *J Bacteriol* 194, 4322-4331.
- Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E. & Walker, G. C. (2007). How rhizobial symbionts invade plants: the *Sinorhizobium*–*Medicago* model. *Nat Rev Microbiol* 5, 619-633.
- Jones, K. M., Sharopova, N., Lohar, D. P., Zhang, J. Q., VandenBosch, K. A. & Walker, G. C. (2008). Differential response of the plant *Medicago truncatula* to its symbiont *Sinorhizobium meliloti* or an exopolysaccharide-deficient mutant. *Proc Natl Acad Sci USA* 105, 704-709.
- Jordan, D. C. (1984). *Rhizobiaceae*. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 234-241. Edited by N. R. Kreig. Baltimore: Williams & Wilkins.
- Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., Ramuz, M. & Allardet-Servent, A. (1998). Unconventional genomic organization in the  $\alpha$  subgroup of the Proteobacteria. *J Bacteriol* 180, 2749-2755.
- Kalhoefer, D., Thole, S., Voget, S., Lehmann, R., Liesegang, H., Wollher, A., Daniel, R., Simon, M. & Brinkhoff, T. (2011). Comparative genome analysis and genome-guided physiological analysis of *Roseobacter litoralis*. *BMC Genomics* 12, 324.
- Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A. & other authors (2000). Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* 7, 331-338.
- Knee, E. M., Gong, F., Gao, M., Teplitski, M., Jones, A. R., Foxworthy, A., Mort, A. J. & Bauer, W. D. (2001). Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. *Mol Plant Microbe Interact* 14, 775-784.
- Kohler, P. R., Choong, E.-L. & Rossbach, S. (2011). The RpiR-like repressor IolR regulates inositol catabolism in *Sinorhizobium meliloti*. *J Bacteriol* 193, 5155-5163.
- Kohler, P. R. A., Zheng, J. Y., Schoffers, E. & Rossbach, S. (2010). Inositol catabolism, a key pathway in *Sinorhizobium meliloti* for competitive host nodulation. *Appl Environ Microbiol* 76, 7972-7980.
- Köhler, S., Foulongne, V., Ouahrani-Bettache, S., Bourg, G., Teyssier, J., Ramuz, M. & Liautard, J. P. (2002). The analysis of the intramacrophagic virulome of *Brucella suis* deciphers the environment encountered by the pathogen inside the macrophage host cell. *Proc Natl Acad Sci USA* 99, 15711-15716.
- Kondorosi, E., Buire, M., Cren, M., Iyer, N., Hoffmann, B. & Kondorosi, A. (1991). Involvement of the *syrM* and *nodD3* genes of *Rhizobium meliloti* in *nod* gene activation and in optimal nodulation of the plant host. *Mol Microbiol* 5, 3035-3048.

- Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet* 39, 309-338.
- Koziol, U., Hannibal, L., Rodríguez, M. C., Fabiano, E., Kahn, M. L. & Noyz, F. (2009). Deletion of citrate synthase restores growth of *Sinorhizobium meliloti* 1021 aconitase mutants. *J Bacteriol* 191, 7581-7586.
- Krol, E. & Becker, A. (2004). Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. *Mol Gen Genet* 272, 1-17.
- Kuykendall, L. D., Shao, J. Y. & Hartung, J. S. (2012). Conservation of gene order and content in the circular chromosomes of *Candidatus Liberibacter asiaticus* and other *Rhizobiales*. *PLoS ONE* 74, e34673.
- Lambert, A., Østerås, M., Mandon, K., Poggi, M. C. & Rudulier, D. L. (2001). Fructose uptake in *Sinorhizobium meliloti* is mediated by a high-affinity ATP-binding cassette transport system. *J Bacteriol* 183, 4709-4717.
- Latner, A. L. & Skillen, A. W. (1968). *Isozymes in biology and medicine*. New York, NY: Academic Press.
- Lawrence, J. G. & Roth, J. R. (1996). Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* 143, 1843-1860.
- LeBlanc, D. J. & Mortlock, R. P. (1971). Metabolism of D-arabinose: a new pathway in *Escherichia coli*. *J Bacteriol*, 90-96.
- Lehman, A. P. & Long, S. R. (2013). Exopolysaccharides from *Sinorhizobium meliloti* Can Protect against H<sub>2</sub>O<sub>2</sub>-Dependent Damage. *J Bacteriol* 195, 5362-5369.
- Leigh, J. A. & Walker, G. C. (1994). Exopolysaccharides of *Rhizobium*: synthesis, regulation and symbiotic function. *Trends Genet* 10, 63-67.
- Leigh, J. A., Signer, E. R. & Walker, G. C. (1985). Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. *Proc Natl Acad Sci USA* 82, 6231-6235.
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C. & Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* 344, 781-784.
- Lessie, T. G. & Phibbs, P. V. (1984). Alternative pathways of carbohydrate utilization in pseudomonads. *Annu Rev Microbiol* 38, 359-388.
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E. & other authors (2004). A putative Ca<sup>2+</sup> and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* 303, 1361-1364.



- Li, L., Jia, Y., Hou, Q., Charles, T. C., Nester, E. W. & Pan, S. Q. (2002). A global pH sensor: *Agrobacterium* sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. *Proc Natl Acad Sci USA* 99, 12369-12374.
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T. & Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* 302, 630-633.
- Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T. & Geurts, R. (2005). Formation of organelle-like N<sub>2</sub>-fixing symbiosomes in legume root nodules is controlled by DMI2. *Proc Natl Acad Sci USA* 102, 10375-10380.
- Lin, E. C. C. (1996). Dissimilatory pathways for sugars, polyols, and carboxylates. In *Escherichia coli and Salmonella Cellular and Molecular Biology*, vol. 1, pp. 307-341. Edited by F. C. Neidhardt. Washington, D. C: ASM Press.
- Lodwig, E. & Poole, P. (2003). Metabolism of *Rhizobium* bacteroids. *Crit Rev Plant Sci* 22, 37-78.
- Lodwig, E., Leonard, M., Marroqui, S., Wheeler, T., Findlay, K., Downie, J. & Poole, P. (2005). Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. *Mol Plant Microbe Interact* 18, 67-74.
- Lois, A. F., Weinstein, M., Ditta, G. S. & Helinski, D. (1993). Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. *J Biol Chem* 268, 4370-4375.
- Lu, H.-Y., Luo, L., Yang, M.-H. & Cheng, H.-P. (2012). *Sinorhizobium meliloti* ExoR is the target of periplasmic proteolysis. *J Bacteriol* 194, 4029-4040.
- Luo, Y., Samuel, J., Mosimann, S. C., Lee, J. E., Tanner, M. E. & Strynadka, N. (2001a). The structure of L-ribulose-5-phosphate 4 epimerase: an aldolase-like platform for epimerization. *Biochemistry* 40, 14763-14771.
- Luo, Z., Clemente, T. E. & Farrand, S. K. (2001b). Construction of a derivative of *Agrobacterium tumefaciens* C58 that does not mutate to tetracycline resistance. *Mol Plant Microbe Interact* 14, 98-103.
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S. & other authors (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* 425, 637-640.
- Manian, S. S. & O'Gara, F. (1982a). Induction and regulation of ribulose biphosphate carboxylase activity in *Rhizobium japonicum* during formate-dependent growth. *Arch Microbiol* 131, 51-54.
- Manian, S. S. & O'Gara, F. (1982b). Derepression of ribulose biphosphate carboxylase activity in *Rhizobium meliloti*. *FEMS Microbiol Lett* 14, 95-99.

- Marketon, M. M., Glenn, S. A., Eberhard, A. & González, J. E. (2003). Quorum sensing controls exopolysaccharide production in *Sinorhizobium meliloti*. *J Bacteriol* 185, 325-331.
- Markowitz, V. M., Chen, I. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Anderson, I., Lykidis, A. & other authors (2010). The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res* 38 (suppl 1), D382-D290.
- Marlow, V. L., Haag, A. F., Kobayashi, H., Fletcher, V., Scocchi, M., Walker, G. C. & Ferguson, G. P. (2009). Essential role for the BacA protein in the uptake of a truncated eukaryotic peptide in *Sinorhizobium meliloti*. *J Bacteriol* 191, 1519-1527.
- Marroquí, S., Zorreguieta, A., Santamaría, C., Temprano, F., Soberón, M., Megías, M. & Downie, J. A. (2001). Enhanced symbiotic performance by *Rhizobium tropici* glycogen synthase mutants. *Journal of bacteriology* 183, 854-864.
- Marsh, J. F., Rakocevic, A., Mitra, R. M., Brocard, L., Sun, J., Eschstruth, A., Long, S. R., Schultze, M., Ratet, P. & other authors (2007). *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol* 144, 324-335.
- Martinez de Drets, G. & Arias, A. (1970). Metabolism of some polyols by *Rhizobium meliloti*. *J Bacteriol* 103, 97-103.
- Martinez de Drets, G., Arias, A., Rovira, A. D. & Cutinella, M. (1974). Fast- and slow growing rhizobia: differences in sucrose utilization and invertase activity. *Can J Microbiol* 20, 605-609.
- Martínez de Drets, G. & Arias, A. (1972). Enzymatic basis for differentiation of *Rhizobium* into fast- and slow-growing groups. *J Bacteriol* 109, 467-470.
- Masson-Boivin, C., Giraud, E., Perret, X. & Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol* 17, 458-466.
- Mattevi, A., de Kok, A. & Perham, R. N. (1992). The pyruvate dehydrogenase multienzyme complex. *Curr Opin Struct Biol* 2, 877-887.
- Mauchline, T. H., Fowler, J. E., East, A. K., Sartor, A. L., Zaheer, R., Hosie, A. H. F., Poole, P. S. & Finan, T. M. (2006). Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc Natl Acad Sci USA* 103, 17933-17938.
- McDermott, T. R. & Kahn, M. L. (1992). Cloning and mutagenesis of the *Rhizobium meliloti* isocitrate dehydrogenase gene. *J Bacteriol* 174, 4790-4797.
- McNeil, M., Darvill, A. G., Fry, S. C. & Albersheim, P. (1984). Structure and function of the primary cell walls of plants. *Ann Rev Biochem* 53, 625-663.

- Meade, H. M., Long, S. R., Ruvkin, G. B., Brown, S. E. & Ausubel, F. M. R. (1982). Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J Bacteriol* 149, 114-122.
- Mendrygal, K. E. & González, J. E. (2000). Environmental regulation of exopolysaccharide production in *Sinorhizobium meliloti*. *J Bacteriol* 182, 599-606.
- Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A. & Kondorosi, E. (2003). A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol* 132, 161-173.
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F. & other authors (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. *Proc Natl Acad Sci USA* 103, 5230-5235.
- Meyer, M. E. (1966). Metabolic characterization of the genus *Brucella*. V. Relationship of strain oxidation rate of i-erythritol to strain virulence for guinea pigs. *J Bacteriol* 92, 584-588.
- Meyer, M. E. (1967). Metabolic characterization of the genus *Brucella*. VI. Growth stimulation by i-erythritol compared with strain virulence for guinea pigs. *J Bacteriol* 93, 996-1000.
- Middleton, P. H., Jakab, J., Penmetsa, R. V., Starker, C. G., Doll, J., Kaló, P., Prabhu, R., Marsh, J. F., Mitra, R. M. & other authors (2007). An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. *Plant Cell* 19, 1221-1234.
- Miller, K. J. & Wood, J. M. (1996). Osmoadaptation by rhizosphere bacteria. *Annu Rev Microbiol* 50, 101-136.
- Miller-Williams, M., Loewen, P. C. & Oresnik, I. J. (2006). Isolation of salt-sensitive mutants of *Sinorhizobium meliloti* strain Rm1021. *Microbiology* 152, 2049-2052.
- Milunovic, B., Morton, R. A. & Finan, T. M. (2014). Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J Bacteriol* 196, 811-824.
- Mitsch, M. J., Cowie, A. & Finan, T. M. (2007). Malic enzyme cofactor and domain requirements for symbiotic N<sub>2</sub> fixation by *Sinorhizobium meliloti*. *J Bacteriol* 189, 160-168.
- Morieri, G., Martinez, E. A., Jarynowski, A., Driguez, H., Morris, R., Oldroyd, G. E. & Downie, J. A. (2013). Host- specific Nod- factors associated with *Medicago truncatula* nodule infection differentially induce calcium influx and calcium spiking in root hairs. *New Phytol* 200, 656-662.

- Morris, D. L. (1948). Quantitative determination of carbohydrates with Dreyer's anthrone reagent. *Science* 105, 254-255.
- Mortimer, M. W., McDermott, T. R., York, G. M., Walker, G. C. & Kahn, M. L. (1999). Citrate synthase mutants of *Sinorhizobium meliloti* are ineffective and have altered cell surface polysaccharides. *J Bacteriol* 181, 7608-7613.
- Mortlock, R. P. (1984). The utilization of pentitols in the studies of the evolution of enzyme pathways. In *Microorganisms as Model Systems for Studying Evolution*, pp. 1-21. Edited by R. P. Mortlock. New York: Plenum Press.
- Mortlock, R. P. & Wood, W. A. (1964a). Metabolism of pentoses and pentitols by *Aerobacter aerogenes*. I. Demonstration of pentose isomerase, pentulokinase, and pentitol dehydrogenase enzyme families. *J Bacteriol* 88, 838-844.
- Mortlock, R. P. & Wood, W. A. (1964b). Metabolism of pentoses and pentitols by *Aerobacter aerogenes*. II. Mechanism of acquisition of kinase, isomerase, and dehydrogenase activity. *J Bacteriol* 88, 845-849.
- Mortlock, R. P., Fossitt, D. D. & Wood, W. A. (1965a). A basis for utilization of unnatural pentoses and pentitols by *Aerobacter aerogenes*. *Proc Natl Acad Sci USA* 54, 572-579.
- Mortlock, R. P., Fossitt, D. D., Petering, D. H. & Wood, W. A. (1965b). Metabolism of pentoses and pentitols by *Aerobacter aerogenes*. III. Physical and immunological properties of pentitol dehydrogenases and pentulokinases. *J Bacteriol* 89, 129-135.
- Mulligan, J. T. & Long, S. R. (1985). Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. *Proc Natl Acad Sci USA* 82, 6609-6613.
- Mulligan, J. T. & Long, S. R. (1989). A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. *Genetics* 122, 7-18.
- Muñoz, J. A., Coronado, C., Pérez-Hormaeche, J., Kondorosi, A., Ratet, P. & Palomares, A. J. (1998). MsPG3, a *Medicago sativa* polygalacturonase gene expressed during the alfalfa-*Rhizobium meliloti* interaction. *Proc Natl Acad Sci USA* 95, 9687-9692.
- Murphy, P., Wexler, W., Grzemski, W., Rao, J. & Gordon, D. (1995). Rhizopines—their role in symbiosis and competition. *Soil Biol Biochem* 27, 525-529.
- Muzzi, A., Moschioni, M., Covacci, A., Rappuoli, R. & Donati, C. (2008). *Streptococcus pneumoniae* is driven by positive selection and recombination. *PLoS ONE* 3, e3660.
- Nicholas, K. B., Nicholas Jr, H. B. & Deerfield, D. W. I. I. (1997). GeneDoc: analysis and visualization of genetic variation. *EMBNEW News* 4, 14.

- Niel, C., Guillaume, J. B. & Bechet, M. (1977). Mise en évidence de deux enzymes présentant une activité  $\beta$ -galactosidasique chez *Rhizobium meliloti*. *Can J Microbiol* 23, 1178-1181.
- Noxon, J. (1976). Atmospheric nitrogen fixation by lightning. *Geophys Res Lett* 3, 463-465.
- Oldham, M. L., Chen, S. & Chen, J. (2013). Structural basis for substrate specificity in the *Escherichia coli* maltose transport system. *Proc Natl Acad Sci USA* 110, 18132-18137.
- Oldroyd, G. E., Murray, J. D., Poole, P. S. & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* 45, 119-144.
- Omata, T., Price, G. D., Badger, M. R., Okamura, M., Gohta, S. & Ogawa, T. (1999). Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proc Natl Acad Sci USA* 96, 13571-13576.
- Omelchenko, M. V., Makarova, K. S., Wolf, Y. I., Rogozin, I. B. & Koonin, E. V. (2003). Evolution of mosaic operons by horizontal gene transfer and gene displacement *in situ*. *Genome Biol* 4, R55.
- Oresnik, I. J. & Layzell, D. B. (1994). Composition and distribution of adenylates in soybean (*Glycine max* L.) nodule tissue. *Plant Physiol* 104, 217-225.
- Oresnik, I. J., Charles, T. C. & Finan, T. M. (1994). Second site mutations specifically suppress the Fix- phenotype of *Rhizobium meliloti ndvF* mutations on alfalfa: identification of a conditional *ndvF*-dependent mucoid colony phenotype. *Genetics* 136, 1233-1243.
- Oresnik, I. J., Twelker, S. & Hynes, M. F. (1999). Cloning and characterization of a *Rhizobium leguminosarum* gene encoding a bacteriocin with similarities to RTX toxins. *Appl Environ Microbiol* 65, 2833-2840.
- Oresnik, I. J., Liu, S.-L., Yost, C. K. & Hynes, M. F. (2000). Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *J Bacteriol* 182, 3582-3586.
- Oresnik, I. J., Pacarynuk, L. A., O'Brien, S. A. P., Yost, C. K. & Hynes, M. F. (1998). Plasmid encoded catabolic genes in *Rhizobium leguminosarum* bv. *trifolii*: evidence for a plant-inducible rhamnose locus involved in competition for nodulation. *Mol Plant Microbe Interact* 11, 1175-1185.
- Osbourn, A. E. & Field, B. (2009). Operons. *Cell Mol Life Sci* 66, 3755-3775.
- Østerås, M., Driscoll, B. T. & Finan, T. M. (1995). Molecular and expression analysis of the *Rhizobium meliloti* phosphoenolpyruvate carboxykinase (*pckA*) gene. *J Bacteriol* 177, 1639-1648.

- Østerås, M., Driscoll, B. T. & Finan, T. M. (1997). Increased pyruvate orthophosphate dikinase activity results in an alternative gluconeogenic pathway in *Rhizobium (Sinorhizobium) meliloti*. *Microbiology* 143, 1639-1648.
- Ott, T., Sullivan, J., James, E. K., Flemetakis, E., Günther, C., Gibon, Y., Ronson, C. & Udvardi, M. (2009). Absence of symbiotic leghemoglobins alters bacteroid and plant cell differentiation during development of *Lotus japonicus* root nodules. *Mol Plant Microbe Interact* 22, 800-808.
- Page, R. D. M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12, 357-358.
- Pedrosa, F. O. & Zancan, G. T. (1974). L-Arabinose metabolism in *Rhizobium japonicum*. *J Bacteriol* 119, 336-338.
- Pellock, B. J., Teplitski, M., Boinay, R. P., Bauer, W. D. & Walker, G. C. (2002). A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. *J Bacteriol* 184, 5067-5076.
- Perret, X., Staehelin, C. & Broughton, W. J. (2000). Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev* 64, 180-201.
- Peters, N. K., Frost, J. W. & Long, S. R. (1986). A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* 233, 977-980.
- Petersen, E., Rajashekara, G., Sanakkayala, N., Eskra, L., Harms, J. & Splitter, G. (2013). Erythritol triggers expression of virulence traits in *Brucella melitensis*. *Microbes and Infection*.
- Pickering, B. S. & Oresnik, I. J. (2008). Formate-dependent autotrophic growth in *S. meliloti*. *J Bacteriol* 190, 6409-6418.
- Pierre, O., Engler, G., Hopkins, J., Brau, F., Boncompagni, E. & Hérouart, D. (2013). Peribacteroid space acidification: a marker of mature bacteroid functioning in *Medicago truncatula* nodules. *Plant Cell Environ* 36, 2059-2070.
- Pinedo, C. A. & Gage, D. J. (2009). HPrK Regulates succinate-mediated catabolite repression in the Gram-Negative symbiont *Sinorhizobium meliloti*. *J Bacteriol* 191, 298-309.
- Pinedo, C. A., Bringhurst, R. M. & Gage, D. J. (2008). *Sinorhizobium meliloti* mutants lacking PTS enzymes HPr or EIIA are altered in diverse processes including carbon metabolism, cobalt requirements and succinoglycan production. *J Bacteriol* 190, 2947-2956.
- Platt, R., Drescher, S. K., Park, S. K. & Phillips, G. J. (2000). Genetic system for reversible integration of DNA constructs and *lacZ* gene fusions into the *Escherichia coli* chromosome. *Plasmid* 43, 12-23.

- Pobigaylo, N., Wetter, D., Szymczak, S., U., S., Kurtz, S., Meyer, F., Nattkemper, T. W. & Becker, A. (2006). Construction of a large signature-tagged mini-Tn5 transposon library and Its application to mutagenesis of *Sinorhizobium meliloti*. *Appl Environ Microbiol* 72, 4329-4337.
- Poole, P. S., Blyth, A., Reid, C. J. & Walters, K. (1994). myo-Inositol catabolism and catabolite regulation in *Rhizobium leguminosarum* bv. *viciae*. *Microbiology* 140, 2787-2795.
- Portais, J., Tavernier, P., Gosselin, I. & Barbotin, J. (1999). Cyclic organization of the carbohydrate metabolism in *Sinorhizobium meliloti*. *Eur J Biochem* 265, 473-480.
- Power, J. (1967). The L-rhamnose genetic system of *E. coli* K-12. *Genetics* 55, 557-568.
- Poysti, N. J. & Oresnik, I. J. (2007). Characterization of *Sinorhizobium meliloti* triose phosphate isomerase genes. *J Bacteriol* 189, 3445-3451.
- Poysti, N. J., Loewen, E. D., Wang, Z. & Oresnik, I. J. (2007). *Sinorhizobium meliloti* pSymB carries genes necessary for arabinose transport and catabolism. *Microbiology* 153, 727-736.
- Prell, J. & Poole, P. (2006). Metabolic changes of rhizobia in legume nodules. *Trends Microbiol* 14, 161-168.
- Prell, J., Mulley, G., Haufe, F., White, J., Williams, A., Karunakaran, R., Downie, J. & Poole, P. (2012). The PTSNtr system globally regulates ATP- dependent transporters in *Rhizobium leguminosarum*. *Mol Micro* 84, 117-129.
- Price, M. N., Arkin, A. P. & Alm, E. J. (2006). The life-cycle of operons. *PLoS Genet* 2, e96.
- Primrose, S. B. & Ronson, C. W. (1980). Polyol metabolism by *Rhizobium trifolii*. *J Bacteriol* 141, 1109-1114.
- Pueppke, S. G. & Broughton, W. J. (1999). *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Mol Plant Microbe Interact* 12, 293-318.
- Quandt, J. & Hynes, M. F. (1993). Versatile suicide vectors which allow direct selection for gene replacement in Gram negative bacteria. *Gene* 127, 15-21.
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S. & other authors (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425, 585-592.
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Jurkiewicz, A., Fukai, E., Quistgaard, E. M., Albrektsen, A. S., James, E. K., Thirup, S. & other authors (2007). LysM domains

mediate lipochitin–oligosaccharide recognition and Nfr genes extend the symbiotic host range. *The EMBO journal* 26, 3923-3935.

Raineri, D. M., Boulton, M. I., Davies, J. W. & Nester, E. W. (1993). VirA, the plant-signal receptor, is responsible for the Ti plasmid-specific transfer of DNA to maize by *Agrobacterium*. *Proc Natl Acad Sci USA* 90, 3549-3553.

Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A. & Poole, P. (2011). Adaptation of *Rhizobium leguminosarum* to peas, alfalfa, and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biol* 12, R106.

Ramírez-Trujillo, J., Encarnación, S., Salazar, E., de los Santos, A. G., Dunn, M., Emerich, D., Calva, E. & Hernández-Lucas, I. (2007). Functional characterization of the *Sinorhizobium meliloti* acetate metabolism genes *aceA*, *SMc00767*, and *glcB*. *J Bacteriol* 189, 5875-5884.

Reeve, W., O'Hara, G., Chain, P., Ardley, J., Brau, L., Nandesena, K., Tiwari, R., Copeland, A., Nolan, M. & other authors (2010a). Complete genome sequence of *Rhizobium leguminosarum* bv. *trifolii* strain WSM1325, an effective microsymbiont of annual Mediterranean clovers. *Standards in Genomic Sciences* 2, 347-356.

Reeve, W., Chain, P., O'Hara, G., Ardley, J., Nandesena, K., Bräu, L., Tiwari, R., Malfatti, S., Kiss, H. & other authors (2010b). Complete genome sequence of the *Medicago* microsymbiont *Ensifer* (*Sinorhizobium*) *medicae* strain WSM419. *Standards in Genomic Sciences* 2, 77-86.

Reinhold, B. B., Chan, S. Y., Reuber, T. L., Marra, A., Walker, G. C. & Reinhold, V. N. (1994). Detailed structural characterization of succinoglycan, the major exopolysaccharide of *Rhizobium meliloti* Rm1021. *J Bacteriol* 176, 1997-2002.

Renalier, M., Batut, J., Ghai, J., Terzaghi, B., Gherardi, M., Garnerone, A., Vasse, J., Truchet, G., Huguët, T. & other authors (1987). A new symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional *fix* gene repeat and a *nod* locus. *J Bacteriol* 169, 2231-2238.

Reuber, T. L. & Walker, G. C. (1993). Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* 74, 269-280.

Richardson, J. S. & Oresnik, I. J. (2007). L-rhamnose transport in *Rhizobium leguminosarum* is dependent upon RhaK, a sugar kinase. *J Bacteriol* 189, 8437-8446.

Richardson, J. S., Hynes, M. F. & Oresnik, I. J. (2004). A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. *trifolii*. *J Bacteriol* 186, 8433-8442.

Richardson, J. S., Carpena, X., Switalta, J., Perez-Luque, R., Donald, L. J., Loewen, P. C. & Oresnik, I. J. (2008). RhaU of *Rhizobium leguminosarum* is a rhamnose mutarotase. *J Bacteriol* 190, 2903-2910.



- Rittig, M. G., Alvarez-Martinez, M.-T., Porte, F., Liautard, J.-P. & Rouot, B. (2001). Intracellular survival of *Brucella* spp. in human monocytes involves conventional uptake but special phagosomes. *Infect Immun* 69, 3995-4006.
- Rivers, D. & Oresnik, I. J. (2013). RhaK dependant ABC-transport of rhamnose in *R. leguminosarum*: genetic separation of kinase and transport activities *J Bacteriol* 195, 3424-3432.
- Robledo, M., Jiménez-Zurdo, J., Velázquez, E., Trujillo, M., Zurdo-Piñeiro, J., Ramírez-Bahena, M., Ramos, B., Díaz-Mínguez, J., Dazzo, F. & other authors (2008). *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. *Proc Natl Acad Sci USA* 105, 7064-7069.
- Roche, P., Debellé, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Dénarié, J. & Promé, J.-C. (1991). Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* 67, 1131-1143.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C. & other authors (2009). A safe operating space for humanity. *Nature* 461, 472-475.
- Rodríguez, M. C., Viadas, C., Seoane, A., Sangari, F. J., López-Goñi, I. & García-Lobo, J. M. (2012). Evaluation of the effects of erythritol on gene expression in *Brucella abortus*. *PloS one* 7, e50876.
- Ronquist, F. (2004). Bayesian inference of character evolution. *Trends Ecol Evol* 19, 475-481.
- Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574.
- Ronson, C. W. & Primrose, S. B. (1979). Effect of glucose on polyol metabolism by *Rhizobium trifolii*. *J Bacteriol* 139, 1075-1078.
- Ronson, C. W., Lyttleton, P. & Robertson, J. G. (1981). C<sub>4</sub>-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proc Natl Acad Sci USA* 78, 4284-4288.
- Rosbach, S., Mai, D. J., Carter, E. L., Sauviac, L., Capela, D., Bruand, C. & de Bruijn, F. J. (2008). Response of *Sinorhizobium meliloti* to elevated concentrations of cadmium and zinc. *Appl Environ Microbiol* 74, 4218-4221.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. A. (1989). *Molecular Cloning Manual 2nd ed.* Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press.

- Sangari, F. J., Agüero, J. & García-Lobo, J. M. (2000). The genes for erythritol catabolism are organized as an inducible operon in *Brucella abortus*. *Microbiology* 146, 487-495.
- Sangari, F. J., Grilló, M. J., Jiménez de Bagüés, M. P., González-Carreró, M. I., García-Lobo, J. M., Blasco, J. M. & Agüero, J. (1998). The defect in the metabolism of erythritol of *Brucella abortus* B19 vaccine strain is unrelated with its attenuated virulence in mice. *Vaccine* 16, 1640-1645.
- Santos, R., Hérouart, D., Sigaud, S., Touati, D. & Puppo, A. (2001). Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction. *Mol Plant Microbe Interact* 14, 86-89.
- Schlüter, A., Patschkowski, T., Quandt, J., Selinger, B., Weidner, S., Krämer, M., Zhou, L., Hynes, M. F. & Priefer, U. (1997). Functional and regulatory analysis of the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39. *Mol Plant Microbe Interact* 10, 605-616.
- Schmeisser, C., Liesegang, H., Krysciak, D., Bakkou, N., Le Quéré, A., Wollherr, A., Heinemeyer, I., Morgenstern, B., Pommerening-Röser, A. & other authors (2009). *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. *Appl Environ Microbiol* 75, 4035-4045.
- Schroeder, B. K., House, B. L., Mortimer, M. W., Yurgel, S. N., Maloney, S. C., Ward, K. L. & Kahn, M. L. (2005). Development of a functional genomics platform for *Sinorhizobium meliloti*: Construction of an ORFeome. *Appl Environ Microbiol* 71, 5858-5864.
- Schwedock, J. S. & Long, S. R. (1992). *Rhizobium meliloti* genes involved in sulfate activation: the two copies of *nodPQ* and a new locus, *saa*. *Genetics* 132, 899-909.
- Simon, R., Priefer, U. & Pühler, A. (1983). A broad host range mobilization system for *in vivo* engineering: Transposon mutagenesis in gram-negative bacteria. *Bio/Techniques* 1, 784-791.
- Simossis, V. A., Kleinjung, J. & Heringa, J. (2005). Homology-extended sequence alignment. *Nucleic Acids Res* 33, 816-824.
- Slater, S. C., Goldman, B. S., Goodner, B., Setubal, J. C., Farrand, S. K., Nester, E. W., Burr, T. J., Banta, L., Dickerman, A. W. & other authors (2009). Genome sequences of three agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* 191, 2501-2511.
- Smil, V. (2004). *Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production*. MIT press.

- Soto, M. J., Sanjuan, J. & Olivares, J. (2001). The disruption of a gene encoding a putative arylesterase impairs pyruvate dehydrogenase complex activity and nitrogen fixation in *Sinorhizobium meliloti*. *Mol Plant Microbe Interact* 14, 811-815.
- Spaink, H. P. (2000). Root nodulation and infection factors produced by rhizobial bacteria. *Ann Rev Microbiol* 54, 257-288.
- Spaink, H. P., Wijfjes, A. H., Drift, K. M., Haverkamp, J., Thomas- Oates, J. E. & Lugtenberg, B. J. (1994). Structural identification of metabolites produced by the NodB and NodC proteins of *Rhizobium leguminosarum*. *Mol Microbiol* 13, 821-831.
- Sperry, J. F. & Robertson, D. C. (1975a). Erythritol catabolism by *Brucella abortus*. *J Bacteriol* 121, 619-630.
- Sperry, J. F. & Robertson, D. C. (1975b). Inhibition of growth by erythritol catabolism in *Brucella abortus*. *J Bacteriol* 124, 391-397.
- Steele, T. T., Fowler, C. W. & Griffiths, J. S. (2009). Control of gluconate utilization in *Sinorhizobium meliloti*. *J Bacteriol* 191, 1355-1358.
- Stephens, C., Christen, B., Fuchs, T., Sundaram, V., Watanabe, K. & Jenal, U. (2007). Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. *J Bacteriol* 189, 2181-2185.
- Stowers, M. D. (1985). Carbon metabolism in *Rhizobium* species. *Annu Rev Microbiol* 39, 89-108.
- Swanson, J. A., Mulligan, J. & Long, S. (1993). Regulation of *syrM* and *nodD3* in *Rhizobium meliloti*. *Genetics* 134, 435-444.
- Swingley, W. D., Sadekar, S., Mastrian, S. D., Matthies, H. J., Hao, J., Ramos, H., Acharya, C. R., Conrad, A. L., Taylor, H. L. & other authors (2007). The complete genome sequence of *Roseobacter denitrificans* reveals a mixotrophic rather than photosynthetic metabolism. *J Bacteriol* 189, 683-690.
- Tae, H., Shallom, S., Settlage, R., Preston, D., Adams, L. G. & Garner, H. R. (2011). Revised Genome Sequence of *Brucella suis* 1330. *J Bacteriol* 193, 6410.
- Tamura, K., Peterson, D., Peterson, N. D., Stetcher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28, 2731-2739.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25, 4876-4882.
- Tian, C. F., Zhou, Y. J., Zhang, Y. M., Li, Q. Q., Zhang, Y. Z., Li, D. F., Wang, S., Wang, J., Gilbert, L. B. & other authors (2012). Comparative genomics of rhizobia

- nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *Proc Natl Acad Sci USA* 109, 8629-8634.
- Treangen, T. J. & Rocha, E. P. C. (2011). Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet* 7, e1001284.
- Triplett, E. W. & Sadowsky, M. (1992). Genetics of competition for nodulation of legumes. *Annu Rev Micorbiol* 46, 399-428.
- Ucker, D. S. & Signer, S. R. (1978). Catabolite-repression-like phenomenon *Rhizobium meliloti*. *J Bacteriol* 136, 1197-1200.
- Udvardi, M. & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol* 64, 781-805.
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A. & other authors (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122-1126.
- van Dillewijn, P., Soto, M. a. J., Villadas, P. J. & Toro, N. (2001). Construction and environmental release of a *Sinorhizobium meliloti* Strain genetically modified to be more competitive for alfalfa nodulation. *Appl Environ Microbiol* 67, 3860-3865.
- Vanderlinde, E. M., Hynes, M. F. & Yost, C. K. (2013). Homoserine catabolism by *Rhizobium leguminosarum* bv. *viciae* 3841 requires a plasmid- borne gene cluster that also affects competitiveness for nodulation. *Environmental Microbiology* 16, 205-217.
- Vincent, J. M. (1970). *A manual for the practical study of root-nodule bacteria*. Oxford, England: Blackwell Scientific Publications.
- Vitousek, P. M., Hättenschwiler, S., Olander, L. & Allison, S. (2002). Nitrogen and nature. *AMBIO* 31, 97-101.
- Wang, C., Meek, D. J., Panchal, P., Borluyka, N., Archibald, F. S., Driscoll, B. T. & Charles, T. C. (2006). Isolation of poly-hydroxybutyrate metabolism genes from complex microbial communities by phenotypic complementation of bacterial mutants. *Appl Environ Microbiol* 72, 384-391.
- Wang, C., Saldanha, M., Sheng, X., Shelswell, K. J., Walsh, K. T., Sobral, B. W. & Charles, T. C. (2007). Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with Medicago sp. *Microbiology* 153, 388-398.
- Wang, D., Griffiths, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T. & Long, S. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* 327, 1126-1129.

- Watanabe, S., Kodak, T. & Makino, K. (2006). Cloning, expression, and characterization of bacterial L-arabinose 1-dehydrogenase involved in an alternative pathway of L-arabinose metabolism. *J Biol Chem* 281, 2612-2623.
- Watson, R. J., Chan, Y. K., Wheatcroft, R., Yang, A.-F. & S., H. (1988). *Rhizobium meliloti* genes required for C<sub>4</sub>-dicarboxylate transport and symbiotic nitrogen fixation located on a megaplasmid. *J Bacteriol* 170, 927-934.
- Wawskiewicz, E. J. & Barker, H. A. (1968). Erythritol metabolism by *Propionibacterium pentosaceum*. *J Biol Chem* 243, 1948-1956.
- White, C., Gavina, J. M., Morton, R., Britz-McKibbin, P. & Finan, T. M. (2012). Control of hydroxyproline catabolism in *Sinorhizobium meliloti*. *Mol Microbiol* 85, 1133-1147.
- White, J., Prell, J., James, E. K. & Poole, P. (2007). Nutrient sharing between symbionts. *Plant Physiol* 144, 604-614.
- Wiebo, J., Marek-Kozaczuk, M., Kubik-Komar, A. & Skorupska, A. (2007). Increased metabolic potential of *Rhizobium spp.* is associated with bacterial competitiveness. *Can J Microbiol* 53, 957-967.
- Willis, L. B. & Walker, G. C. (1998). The *phbC* (poly- $\beta$ -hydroxybutyrate synthase) gene of *Rhizobium (Sinorhizobium) meliloti* and characterization of *phbC* mutants. *Can J Microbiol* 44, 554-564.
- Willis, L. B. & Walker, G. C. (1999). A novel *Sinorhizobium meliloti* operon encodes an  $\alpha$ -glucosidase and a periplasmic-binding-protein dependent transport system for  $\alpha$ -glucosides. *J Bacteriol* 181, 4176-4184.
- Wöhrl, B. M., Wehmeier, U. F. & Lengeler, J. W. (1990). Positive and negative regulation of expression of the L-sorbose (*sor*) operon by SorC in *Klebsiella pneumoniae*. *Mol Gen Genet* 224, 193-200.
- Wood, M. & Stanway, A. (2001). Myo-inositol catabolism by *Rhizobium* in soil: HPLC and enzymatic studies. *Soil Biol Biochem* 33, 375-379.
- Wood, W. W., McDonough, M. J. & Jacobs, L. B. (1961). Ribitol and D-arabitol utilization by *Aerobacter aerogenes*. *J Biol Chem* 236, 2190-2195.
- Wu, C.-F., Lin, J.-S., Shaw, G.-C. & Lai, E.-M. (2012). Acid-Induced Type VI Secretion System Is Regulated by ExoR-ChvG/ChvI Signaling Cascade in *Agrobacterium tumefaciens*. *PLoS pathogens* 8, e1002938.
- Yao, S.-Y., Luo, L., Har, K. J., Becker, A., Rüberg, S., Yu, G.-Q., Zhu, J.-B. & Cheng, H.-P. (2004). *Sinorhizobium meliloti* ExoR and ExoS proteins regulate both succinoglycan and flagellum production. *J Bacteriol* 186, 6042-6049.

- Yarosh, O. K., Charles, T. C. & Finan, T. M. (1989). Analysis of C<sub>4</sub>-dicarboxylate transport genes in *Rhizobium meliloti*. *Mol Microbiol* 3, 813-823.
- York, G. M. & Walker, G. C. (1998). The succinyl and acetyl modifications of succinoglycan influence susceptibility of succinoglycan to cleavage by the *Rhizobium meliloti* glycanases ExoK and ExsH. *J Bacteriol* 180, 4184-4191.
- Yoshida, K.-i., Yamaguchi, M., Morinaga, T., Kinehara, M., Ikeuchi, M., Ashida, H. & Fujita, Y. (2008). myo-Inositol catabolism in *Bacillus subtilis*. *J Biol Chem* 283, 10415-10424.
- Yost, C. K., Rath, A. M., Noel, T. C. & Hynes, M. F. (2006). Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. *viciae*. *Microbiology* 152, 2061-2074.
- Young, C. & Nester, E. W. (1988). Association of the VirD2 protein with the 5' end of T strands in *Agrobacterium tumefaciens*. *J Bacteriol* 170, 3367-3374.
- Young, J. P. W., Crossman, L. C., Johnston, A. W., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., Wexler, M., Curson, A. R. J., Todd, J. D. & other authors (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* 7, R34.
- Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F., Gouzy, J. & other authors (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520-524.
- Yurgel, S. N. & Kahn, M. L. (2004). Dicarboxylate transport by rhizobia. *FEMS Microbiol Rev* 28, 489-501.
- Yurgel, S. N. & Kahn, M. L. (2005). Sinorhizobium meliloti *dctA* mutants with partial ability to transport dicarboxylic acids. *J Bacteriol* 187, 1161-1172.
- Yurgel, S. N., Rice, J. & Kahn, M. L. (2013). Transcriptome Analysis of the Role of GlnD/GlnBK in Nitrogen Stress Adaptation by Sinorhizobium meliloti Rm1021. *PLoS one* 8, e58028.
- Zhao, J. & Binns, A. N. (2011). Characterization of the *mmsAB-araD1* (*gguABC*) Genes of *Agrobacterium tumefaciens*. *J Bacteriol* 193, 6586-6596.
- Zhou, Y.-F., Nan, B., Nan, J., Ma, Q., Panjikar, S., Liang, Y.-H., Wang, Y. & Su, X.-D. (2008). C<sub>4</sub> Dicarboxylates Sensing Mechanism Revealed by the Crystal Structures of DctB Sensor Domain. *J Bacteriol* 190, 49-61.