Characterization of a *cbbR* allele that suppresses phenotypes associated with a *tktA/B* double mutant strain in *Sinorhizobium meliloti*

By

Ishani N. Patuwatha Withange

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

MSc.

Department of Microbiology
University of Manitoba
Winnipeg, Manitoba, Canada

Abstract

The genome of *Sinorhizobium meliloti* contains three genes that are annotated as transketolases. Transketolase activity is necessary for the pentose phosphate pathway, and under normal conditions the protein responsible for transketolase activity is encoded for by the gene *tktA.* A strain carrying a mutation in this gene is auxotrophic for aromatic amino acids and is also severely debilitated in its ability to establish an effective symbiosis. It has been shown that the phenotypes associated with a mutation in *tktA* can be partially suppressed by point mutations in a negative regulator that results in the expression of *tktB.* In the present study, mutants were isolated that suppressed carbon phenotypes associated with a *tktA*/*tktB* double mutant. The objective of the study was to identify and characterize this suppressor mutation.

The mutation was in *cbbR* which is annotated as a Lys-R type transcriptional regulator which is recognized as the key regulator of the **C**alvin **B**enson **B**assham (*cbb)* operon in CO_2 fixing organisms. In *S. meliloti*, the *cbb* operon encodes a third putative transketolase-*cbbT*. The results of qRT-PCR show that the *cbbR** mutation is correlated with the upregulation of the 10 genes in the *cbb* gene cluster which includes *cbbR*. Introduction of *cbbT* on a plasmid into a strain carrying mutations in *tktA*/*tktB* results in suppression of *tktA*/*tktB* associated phenotypes, suggesting that the enzyme encoded by *cbbT* plays a significant role in suppression of *tktA*/*tktB*. When inoculated onto alfalfa plants, strains carrying the *cbbR** mutation were able to partially suppress the fix^- phenotypes associated with *tkt* mutations. This suppression resulted in ~50% nitrogen fixing efficiency compared to wild-type. These strains were also highly uncompetitive for nodule occupancy, yet they exhibited nodulation kinetics that were like that of the wildtype. Overall, these observations suggest that the aberrant regulation of *cbb* genes can affect symbiotic development and nitrogen fixation in *S. meliloti*.

Acknowledgments

I would like to extend my gratitude to my supervisor Dr. Ivan Oresnik for all his support given to me through out past years. His mentorship and encouragement helped immensely for me to become more enthusiastic about research. I am highly grateful for his understanding and flexibility during difficult times I faced during my time in the lab. I would also like to thank my committee members Dr. Matthew Bakker and Dr. Karen Brassinga for all the advice and input provided for my work.

A huge thanks for all the past and present members of Oresnik lab for creating a friendly environment to work in from my very first day. Their support and friendship made me feel comfortable and help me do my work with peace of mind. A special thanks to Dr. Justin Hawkins for all his guidance and support given for my project and being there to answer any question I had even after leaving the lab. The department of microbiology also deserves an appreciation for making a friendly and fun environment for graduate students. All the graduate students and the staff have made this journey a memorable one for me and I am highly grateful for that.

I would like to thank my family, especially my parents and two brothers who have been the greatest strength and inspiration of mine. I could not have done this without their support. My loving husband Gayan deserves a huge appreciation for always having my back through thick and thin and helping me to come this far. Finally, I would like to thank my friends from back home for always checking in on me and for not letting me feel alone. Also, friends and community here in Winnipeg, I thank you all for your help from the moment I moved to Canada.

Dedication

This thesis is dedicated to my parents, Ajith Patuwatha Withanage and Chandrika Abeyrathne and to my husband Gayan Sanjana…….

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANOVA	Analysis of variance
BNF	Biological nitrogen fixation
CBB	Calvin Benson Bassham
CCRH	Curled colonized root hair
cDNA	Complementary deoxyribose nucleic acid
CFU	Colony forming unit
Cm	Chloramphenicol
DHAP	Dihydroxy-acetone phosphate
ED	Entner Doudoroff
EMP	Embden Meyerhof Parnas
EPS	Exopolysaccharide
F6P	Fructose-6-phosphate
G3P	Glyceraldehyde-3-phosphate
G6P	Glucose-6-phosphate
Gm	Gentamicin
IT	Infection thread
KDPG	2-Keto-3-deoxy-6-phosphogluconate
Km	Kanamycin
LB	Luria Bertani
LTTR	Lys-R Type transcriptional regulator
Nm	Neomycin
ORF	Open reading frame
PP	Pentose phosphate

qRT-PCR	Quantitative real time polymerase chain reaction
RuBP	Ribulose bis-phosphate
Sm	Streptomycin
SNP	Single nucleotide polymorphism
Tc	Tetracycline
TCA	Tri-carboxylic acid
TPP	Thiamine pyrophosphate
VMM	Vincent’s minimal medium
X5P	Xylulose-5-phosphate

Chapter 1: Literature Review

1.1 Biological nitrogen fixation

Nitrogen is an essential component for the biosynthesis of many of the key compounds of life. This includes, but is not restricted to, amino acids, nucleic acids, chlorophylls, and vitamins. Even though N_2 is the most abundant gas in the Earth's atmosphere, its inert nature makes it inaccessible for plants and animals to assimilate. To be made bioavailable, nitrogen needs to be reduced. The reduction of nitrogen gas occurs through three main processes: biological, industrial, and atmospheric nitrogen fixation.

The atmospheric nitrogen fixation by lightning and industrial nitrogen fixation by Haber-Bosch process contribute for about 5 Tg N/year and 121 Tg N/year respectively (Galloway *et al*., 2008). Biological nitrogen fixation (BNF) plays an important role in global nitrogen cycle by contributing approximately 230 Tg N/year fixed nitrogen. It is carried out by organisms with the enzyme nitrogenase which is encoded by the genes *nifHDK* (Galibert *et al*., 2001). This reaction converts atmospheric di nitrogen to ammonia (Galloway *et al*., 2008).

$$N_2 + 8H^+ + 8\ e^- + 16\text{MgATP} \rightarrow 2NH_3 + H_2 + 16\text{MgADP} + 16\ \text{Pi}$$

In addition to the structural genes for nitrogenase, several other genes that encode accessory proteins necessary for the biosynthesis of the metallo-centres for the functioning of nitrogenase also need to be present. The genes which encode nitrogenase and as well as these accessory components have only been found in the bacteria and archaea (Koirala & Brözel, 2021). Collectively these organisms are often referred to as diazotrophs.

A major group of nitrogen fixing organisms are the rhizobia. In this group the reduction of nitrogen occurs mainly through their symbiotic association with legumes. Studies have shown that the use of rhizobia as a commercial inoculum for legume crop cultivation increases BNF and hence the yield (Herridge *et al*., 2008).

1.1.1 *Rhizobium*-legume symbiosis

The *Rhizobium*-legume symbiosis is a mutualistic interaction where both organisms benefit. Bacteria receive photosynthetic products (carbon) as well as a protective environment from the host legume (Lodwig *et al*., 2003). In return the host legume receives fixed nitrogen in the form of NH_4^+ that is used to synthesize required proteins (Peoples & Craswell, 1992).

Compatibility between the host legume and Rhizobial strain is required to produce a successful interaction. The host specificity of Rhizobia can vary broadly. It is possible that a single species of legume can be nodulated by several Rhizobial species, or alternately, a single species of *Rhizobium* can nodulate a wide range of legumes (Bala & Giller, 2001). A few examples for different species of rhizobia and their host legumes include *Rhizobium leguminosarum* biovar *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) (Geddes & Oresnik, 2016). These symbiotic associations result in the formation of root nodules in which rhizobia differentiate into bacteroids and fix nitrogen (Jones *et al*., 2007).

1.1.2 Root nodule formation

The first step of the infection process is the signal exchange between the legume and free-living rhizobia in the soil rhizosphere. The root exudates of legumes contain signaling molecules such as flavonoids which are recognized by the NodD transcriptional activator proteins of Rhizobia (Gage, 2004). This event induces transcription of downstream *nod* genes that lead to production of a lipo-chito-oligosaccharide signaling molecule called Nod factor. Nod factors are released to the surrounding environment and are in return recognized by the receptors

on legume root hairs (Gust *et al*., 2012). The recognition of Nod factors triggers calcium spiking in the nucleus of plant root cells, which leads to a signaling cascade causing morphological changes in root hairs, such as root hair curling and the division in the inner cortical cells which gives rise to the nodule primordium (Capoen *et al*., 2011; Mitra *et al*., 2004).

Following the attachment of the bacteria to a root hair, the production of Nod factor induces root hair curling. The curling of the root hairs traps the bacteria on their surface forming an apoplastic compartment called curled colonized root hair (CCRH) (Geurts *et al*., 2005). The rhizobia trapped within the CCRH continue to divide and subsequently invade the root cells through the formation of a plant derived intracellular tube-like structure called the infection thread (Oldroyd *et al*., 2011). This structure extends through the root cells until it reaches the inner cortical cells. Rhizobia colonize the tip of the infection thread and enter the inner cortical cells by endocytosis. During the endocytosis, bacteria get surrounded by a plant-derived membrane called symbiosome membrane and the resulting structure is called the symbiosome. The Rhizobia continue to replicate inside the symbiosome and ultimately differentiate into nitrogen fixing bacteroids (Geddes & Oresnik, 2016). Hormone signaling in the root cells mediate the nodule organogenesis (Oldroyd *et al*., 2011) and give rise to the root nodules which consists of a tissue infected with rhizobia in the center surrounded by an uninfected tissue that connects to the plant vascular system (Udvardi & Poole, 2013).

1.1.3 *Sinorhizobium meliloti*

S. meliloti is a gram negative alpha proteobacterium belonging to the Rhizobiaceae family and forms a symbiotic relationship with the forage legume *Medicago sativa* (Alfalfa). This association is being studied as a model system for studying *Rhizobium*-Legume symbiosis

(Geddes *et al*., 2020; Geddes & Oresnik, 2014; Jones *et al*., 2007). In addition to alfalfa, *S. meliloti* can also form viable symbioses with other species in the genus *Medicago* including *M. trancatula*, *Melilotus* (sweet clover) as well as *Trigonella*. The most widely studied strain of *S. meliloti* is Rm1021 which is a streptomycin resistant derivative of the original wild-type strain SU47 (Meade *et al*., 1982).

The complete genome of *S. meliloti* Rm1021 was sequenced in 2001. It has a tripartite genome consists of the chromosome (3.6 Mbp) and two megaplasmids, pSymA (1.35 Mbp) and pSymB (1.68 Mbp) (Barnett *et al*., 2001; Capela *et al*., 2001; Finan *et al*., 2001; Galibert *et al*., 2001). The chromosome contains the major portion of the essential genes for growth such as the genes required for transcription, translation and DNA repair (Capela *et al*., 2001). The megaplasmid pSymA was identified as the symbiotic plasmid as the majority of the genes necessary for nitrogen fixation and nodule formation are localized on this replicon. It has been shown that the pSymA is not essential for the growth of *S. meliloti* as the pSymA cured strain could grow well indicating that it does not contain essential genes for growth (Oresnik *et al*., 2000). The megaplasmid pSymB contains some essential genes such as $tRNA^{arg}$, *engA* and genes for asparagine and thiamine biosynthesis (DiCenzo *et al*., 2013). It also contains genes that are involved in the synthesis and export of exopolysaccharides as well as genes that encode determinants necessary for small molecule transport and catabolism. It has been shown both experimentally as well as computationally using metabolic modelling that pSymB plays an important role in growth and survival of *S. meliloti* in diverse soil environments (DiCenzo *et al*., 2016; DiCenzo *et al*., 2014; Finan *et al*., 2001). It has been recently suggested that the replicon pSymB should be termed a chromid (DiCenzo & Finan, 2017).

1.1.4 Competition for nodule occupancy

Rhizobia exist in the soil in free-living saprophytic form in addition to the symbiotic association they form with legumes. In agricultural systems, the inability of the inoculant rhizobia to successfully outcompete the native rhizobia in soil to form an effective nitrogen fixing association with plant host has been identified as a major problem (Triplett & Sadowsky, 1992). The soil rhizosphere is rich in root exudates which consists of a wide array of carbohydrates, proteins, signal molecules and other nutrients (Oldroyd *et al*., 2011). Rhizobia must be able to utilize these nutrients and adapt to survive in the soil for successful nodule occupancy. There are several factors that affect competition in soil for nodule occupancy including genetic traits of rhizobia (Triplett & Sadowsky, 1992), soil conditions such as, water level (Krasova-Wade *et al*., 2006), pH levels (Geddes *et al*., 2014) and temperature (Rice & Olsen, 1988) and the ability of bacteria to utilize different carbon sources (Geddes & Oresnik, 2012b; Oresnik *et al*., 1998; Rivers *et al*., 2022).

Medium acidification and increased exopolysaccharide production when grown in the presence of specific carbon sources has been shown to correlate with increased competitiveness (Geddes *et al*., 2014). A strain possessing a mutation in the transketolase encoding gene (*tktA*), which is unable to acidify its growth medium (Hawkins *et al*., 2017), was impaired in exopolysaccharide production and shown to be ineffective root nodule formation in Alfalfa plants (Hawkins *et al*., 2018). Several studies have been conducted to determine the effect of differential ability to utilize carbon sources on competition for nodule occupancy. The mutants which are unable to catabolize rhamnose have also been shown to be less competitive than the wild-type *S. meliloti* (Rivers *et al*., 2022). In contrast, strains carrying mutations in genes

involved in De Ley-Doudoroff pathway for galactose catabolism have increased competitiveness for nodule occupancy (Geddes & Oresnik, 2012b).

While most of these studies have focused on the effects of single carbon compounds, there are recent approaches that have focused on system wide analysis to determine the role of carbon utilization for competition and nodule occupancy (Burghardt & diCenzo, 2023; Pini *et al*., 2017; Poole *et al*., 2018; Wheatley *et al*., 2020). The genes required to metabolize different carbon sources have been identified among the essential genes in rhizobia at different stages of root nodule formation including rhizosphere progression, root colonization and bacteroid differentiation (Pini *et al*., 2017; Wheatley *et al*., 2020). This suggests that the carbon metabolism in rhizobia is one of the key factors for competition for nodule occupancy.

1.2 Central Carbon metabolism

The ability of Rhizobia to metabolize different carbon sources available in the soil rhizosphere has been shown to play a role in their ability to compete for nodule occupancy (Geddes & Oresnik, 2012b; Hawkins *et al*., 2018). Historically it was well documented that Rhizobia could utilize diverse carbon compounds (Stowers, 1985). This was shown to be true with the completed genome sequence of *S. meliloti* (Finan *et al*., 2001; Galibert *et al*., 2001), which revealed a large number of genes dedicated to the utilization of different carbon sources including genes encoding 430 ABC transporters (Finan *et al*., 2001). Remarkably 235 of these are found on pSymB (Finan *et al*., 2001; Galibert *et al*., 2001). In addition to having a large number of genes dedicated to the transport and catabolism of various compounds, the genome of *S. meliloti* also contains all of the central carbon pathways including the Entner-Doudorof (ED) pathway, the Embden-Meyerhof-Parnas (EMP) pathway, the pentose phosphate (PP), the

tricarboxylic acid (TCA), as well as the Calvin-Benson-Bassam (CBB) pathway (Galibert *et al*., 2001).

1.2.1 Entner-Doudoroff and Embden Meyerhof-Parnas pathways

S. meliloti contains both the Entner-Doudoroff pathway (ED) as well as the EMP pathway (Arias *et al*., 1979; Geddes & Oresnik, 2014; Irigoyen *et al*., 1990). The EMP pathway is probably the most widely recognized glycolytic pathway. The pathway is usually thought of as a glycolytic pathway leading to the breakdown of glucose to pyruvate. The pathway can be thought of occurring in two stages. An upper EMP pathway which yields glyceraldehyde-3-phosphate (G3P), and the lower EMP pathway which converts G3P to pyruvate in a series of biochemical reactions (Figure 1).

The EMP pathway starts with the phosphorylation of glucose to glucose-6-phosphate (G6P). This is subsequently isomerized to fructose-6-phosphate (F6P) and phosphorylated to become fructose 1,6, bisphosphate. Fructose 1, 6 bisphosphate is acted on by fructose 1, 6 bisphospho-aldolase yielding dihydroxy-acetone phosphate (DHAP) and glyceraldehyde-3-phosphate (G3P). The dihydroxy-acetone-phosphate is isomerized by triose-phosphate-isomerase to become G3P (Figure 1). Glyceraldehyde-3-phosphate then is metabolized using the enzymes associated with the lower half of the EMP pathway, eventually yielding pyruvate.

The ED pathway involves a series of enzymatic reactions which converts G6P ultimately into G3P and pyruvate. The first dedicated step oft his pathway is carried out by G6P dehydrogenase which converts G6P into 6-phosphogluconate. This intermediate is subsequently acted on by 6-phosphono-

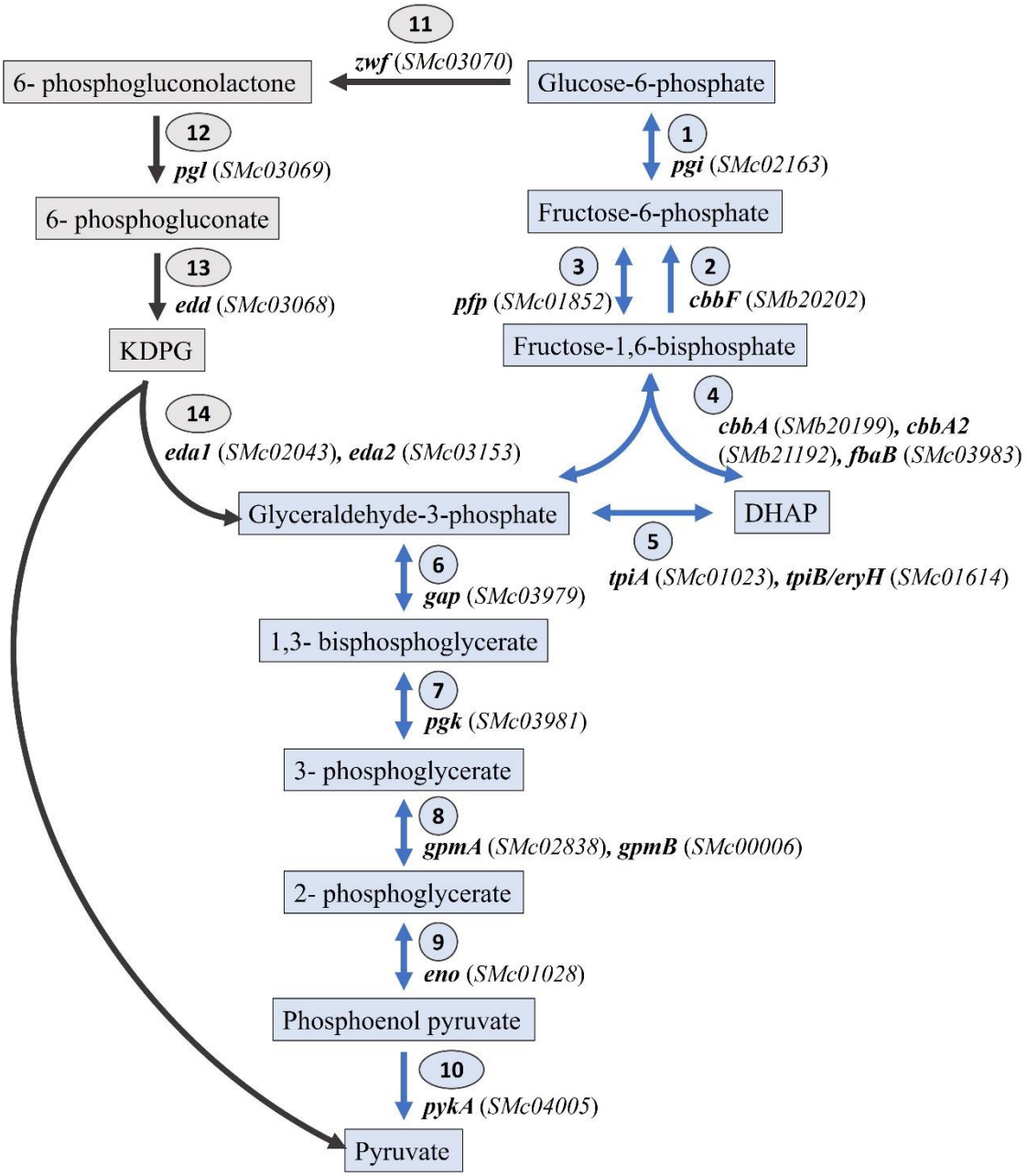

Figure 1. Enzymatic reactions for ED (shown in black) and EMP (shown in blue) pathways. The numbers from 1-10 represent the enzymes for EMP pathway, 11-14 represent the enzymes for ED pathway. The genes that are predicted or experimentally proven to encode the enzymes are written in bold. The locus tags for the genes are indicated inside the brackets. Enzymes; 1, Phosphoglucose isomerase; 2, Fructose bisphosphotase; 3, Pyrophosphate dependent phosphofructokinase; 4, Fructose bisphosphate aldolase; 5, Triose phosphate isomerase; 6, glyceraldehyde-3-phosphate dehydrogenase, 7, Phosphoglycerate kinase; 8, Phosphoglycerate mutase; 9, Enolase; 10, Pyruvate kinase; 11, Glucose-6P dehydrogenase; 12, 6-Phosphogluconolactonase; 13, 6-Phosphogluconate dehydratase; 14, 2-Keto-3-deoxy-6-phosphogluconate aldolase (Geddes & Oresnik, 2014)

glucono-lactonase and 6-phosphogluconate dehydratase to yield 2-keto-3-deoxy-6-phosphogluconate. This compound undergoes an aldol-cleavage that is catalyzed by 2-keto-3-deoxy6-phosphogluconate aldolase to yield G3P and pyruvate. Whereas the pyruvate produced is utilized in the TCA cycle or for biosynthetic reactions, ^{13}C labelling studies show that a large proportion of the G3P is used in a gluconeogenic manner (Fuhrer *et al*., 2005). Supporting these labelling studies, earlier work reported the absence of ATP-dependent phosphofructokinase activity in *S. meliloti* suggesting that this pathway is not being used for glycolysis (Arias *et al*., 1979; Irigoyen *et al*., 1990). However, there is a gene annotated as a probable pyrophosphate-fructose 6-phosphate 1-phosphotransferase in the annotated genome of *S. meliloti*. A recent study has shown that this gene encodes for an enzyme with phosphofructokinase activity dependent on inorganic phosphate instead of ATP which suggest the possibility of having a complete EMP pathway for glycolysis in *S. meliloti* (Kaur, 2021).

Following the generation of pyruvate, its metabolism is carried out through the tricarboxylic acid (TCA) cycle. Studies have shown the expression and the activity of the enzymes involved in TCA cycle and is well characterized (Dunn, 1998; Stowers, 1985).

1.2.2 Pentose Phosphate pathway

S. meliloti has a complete pentose phosphate (PP) pathway which interconverts pentoses and hexoses. The intermediates of this pathway provide the precursors for essential aromatic amino acid biosynthesis. However, characterization of the PP pathway in rhizobia is limited (Geddes & Oresnik, 2014).

The PP pathway consists of two enzymatic branches namely, the oxidative and non-oxidative branch (Figure 2). The oxidative branch contains three irreversible enzymatic reactions

which convert G6P into ribulose-5-phosphate. The first two reactions which subsequently generate 6-phosphogluconate from G6P dehydrogenase and 6-phosphogluconolactonase are shared with the ED pathway. The conversion of 6-phosphogluconate into ribulose-5-phosphate is carried out by 6-phosphogluconate dehydrogenase. Ribulose-5-phosphate then enters the non-oxidative branch which is composed of a series of reversible reactions (Martínez-De Drets & Arias, 1972). This branch interchanges phosphorylated intermediates that act as precursors for aromatic amino acid, nucleic acid and exopolysaccharide biosynthesis. The two enzymes, transketolase (Tkt) and transaldolase (Tal) have been identified as the key enzymes in the non-oxidative branch of the PP pathway (Cervenansky & Arias, 1984). The activity of transketolase generates G3P and F6P which can be cycled back to the EMP pathway.

Transaldolase generates erythrose-4-phosphate which is the main precursor for aromatic amino acid biosynthesis. Moreover, the pentose sugars such as ribose and xylose and the 4-carbon polyol erythritol enter the central carbon metabolism through the non-oxidative branch (Geddes & Oresnik, 2014; Hawkins *et al*., 2018). Mutations in *tkt* and *tal* have been reported (Hawkins *et al*., 2018). Whereas a mutation in *tal* did not affect symbiosis or growth, a mutation in *tkt* greatly affected the physiology of the organism (Hawkins *et al*., 2018).

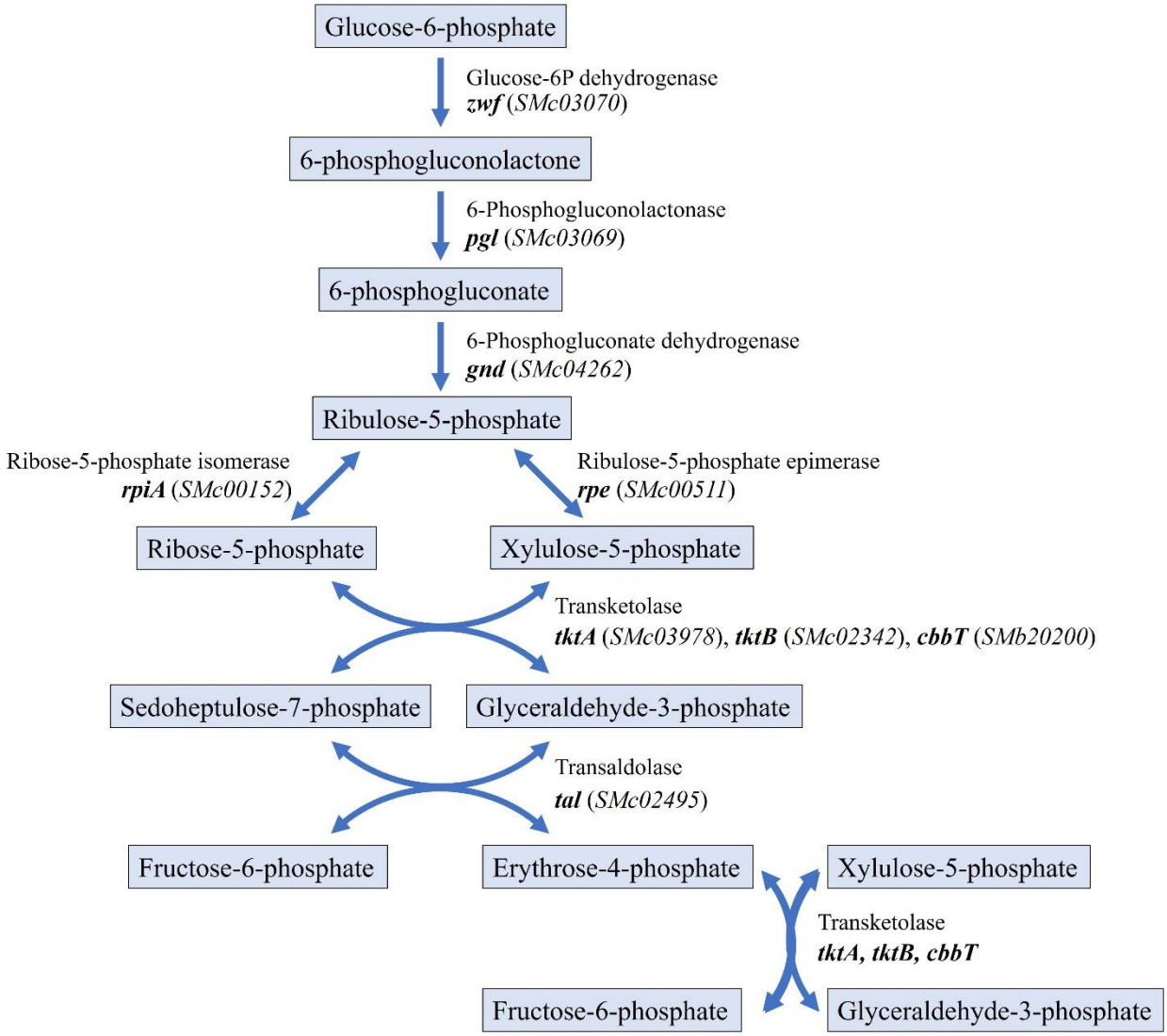

Figure 2. Pentose phosphate pathway of *S. meliloti*.

The arrows represent the enzymatic direction of enzymatic reactions and the annotated genes in *S. meliloti* genome that encode for the enzymes are in bold. The locus tags for the genes are indicated inside the brackets.

1.2.3 Transketolase

Transketolase is a thiamine diphosphate-dependent enzyme that catalyzes the rate limiting step of the non-oxidative branch of the pentose phosphate pathway. The enzymatic reactions catalyzed by transketolase utilize a ketose donor substrate and an aldose recipient substrate (Kochetov & Solovjeva, 2014). In PP pathway, it carries out the reversible conversion of ribose-5-phosphate and xylulose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate, and xylulose-5-phosphate and erythrose-4-phosphate to F6P and glyceraldehyde-3-phosphate (Schenk *et al*., 1998). The transketolase enzyme functions as a homodimer, formed through the interaction of PP (N-terminal) and Pyr (middle) domains of two identical subunits. Thiamine diphosphate and calcium are the two main co-factors for transketolase, interacting with the two active sites of the enzyme (Kochetov & Philippov, 1970).

In *S. meliloti* three genes are annotated as putative transketolases (*tktA*, *tktB* and *cbbT*). The gene *tktA* has been identified as the major tansketolase encoding gene and a mutation in this gene results in an auxotrophy and an inability to fix nitrogen in plants (Hawkins *et al*., 2018). Mutations in *tktB* did not produce such detrimental growth phenotypes, suggesting that *tktB* might encode for a minor transketolase similar to that in *E. coli* (Hawkins *et al*., 2018; Iida *et al*., 1993). Complementation of *tktA* with a plasmid carrying *tktB* has been shown to restore growth phenotypes that are due to *tktA* mutation, providing evidence that *tktB* encodes for an enzyme with partial activity. However, complementation with *cbbT* did not suppress any of the growth phenotypes associated with *tktA* (Hawkins *et al*., 2018). A recent study has identified a negative regulator that acts on *tktB* and upon mutation of the regulator, *tktB* is upregulated (Kaur, 2021). This upregulation of *tktB* in a *tktA* mutant background was able to partially restore the nitrogen fixation in plants (Kaur, 2021).

The gene *cbbT* is localized on pSymB together with a cluster of genes which are involved in Calvin Benson cycle for CO_2 fixation. However, there is no evidence available on the activity of *cbbT* up to date.

1.2.4 Phosphoketolase

Phosphoketolase is a thiamine pyrophosphate (TPP) dependent enzyme that is a key glycolytic enzyme in lactic acid bacteria and *Bifidobacteria* (Costelloe *et al*., 2008; Duggleby, 2006). About 128 putative eukaryotic and bacterial phosphoketolases have been identified to date that are recognized as a catalytic siblings of transketolase (Costelloe *et al*., 2008). Phosphoketolase catalyzes the cleavage of either fructose-6-phosphate (F6P) or xylulose-5-phosphate (X5P) into their respective aldose phosphate and acetyl phosphate (Sánchez *et al*., 2010). Based on this, two types of phosphoketolases have been defined. These are F6P phosphoketolase and X5P phosphoketolase. However, in some organisms, the enzyme has been found to have dual activity (Sánchez *et al*., 2010). Several bacterial metabolic engineering studies have been done utilizing phosphoketolases due to its ability to bypass some of the enzymatic reactions in the pentose phosphate and EMP pathways, resulting in less carbon loss (Krüsemann *et al*., 2018; Sonderegger *et al*., 2004). There are two genes in *S. meliloti* genome annotated as putative phosphoketolases (*SMc04146*, *SMa1084*). These genes have not been characterized and there is no evidence of phosphoketolases activity in *S. meliloti*.

1.3 Autotrophic growth of Rhizobia

It has been shown that some members of the Rhizobiaceae family are capable of growing autotrophically and form root nodules in the presence of another energy source (Hanus *et al*.,

1979).The enzymes required for the CO_2 fixation have been identified within the family. However, one of the key enzymes in CO_2 fixation, which is ribulose bisphosphate carboxylase activity was only found in *Bradyrhizobium japonicum* and *S. meliloti* (Manian & O'Gara, 1982).

Bradyrhizobium japonicum has been shown to be capable of chemoautotrophic growth using H_2 as the source of energy and CO_2 as the sole carbon source. Strains that have mutations in their H_2 uptake systems have been found to be defective in autotrophic growth (Hanus *et al*., 1979; Maier, 1981). In contrast, *S. meliloti* must be supplemented with genes required for H_2 uptake systems from *B. japonicum* to be able to grow chemoautotrophically (Lambert *et al*., 1985). *S. meliloti* has shown autotrophic growth in the presence of formate (Pickering & Oresnik, 2008). There are three gene clusters annotated as formate dehydrogenases (*fdoGHI* and *SMa0478* on the megaplasmid pSymA and *fdsABCDG* in the chromosome) in *S. meliloti* genome. In addition, pSymB contains a cluster of genes responsible for the Calvin Benson cycle for CO_2 fixation (Finan *et al*., 2001). The characterization of formate dehydrogenase genes in *S. meliloti* Rm1021 has revealed that the autotrophic growth of this strain in the presence of formate and bicarbonate is dependent only on the chromosomal *fds* locus. Moreover, the contribution of *cbb* operon was also confirmed by mutating the *cbb* operon gene *cbbF*, which resulted in an inability to grow on minimal medium with formate and bicarbonate (Pickering & Oresnik, 2008).

1.3.1 Calvin Benson Bassham cycle

Calvin Benson Bassham (CBB) cycle plays the most important role in CO_2 fixation, and this has been found in wide range of autotrophic organisms that belong to both eukaryotes and prokaryotes. The CBB cycle is a reductive pentose phosphate cycle which facilitates the assimilation of CO_2 by the reaction carried out by the key enzyme ribulose 1,5-bisphosphate

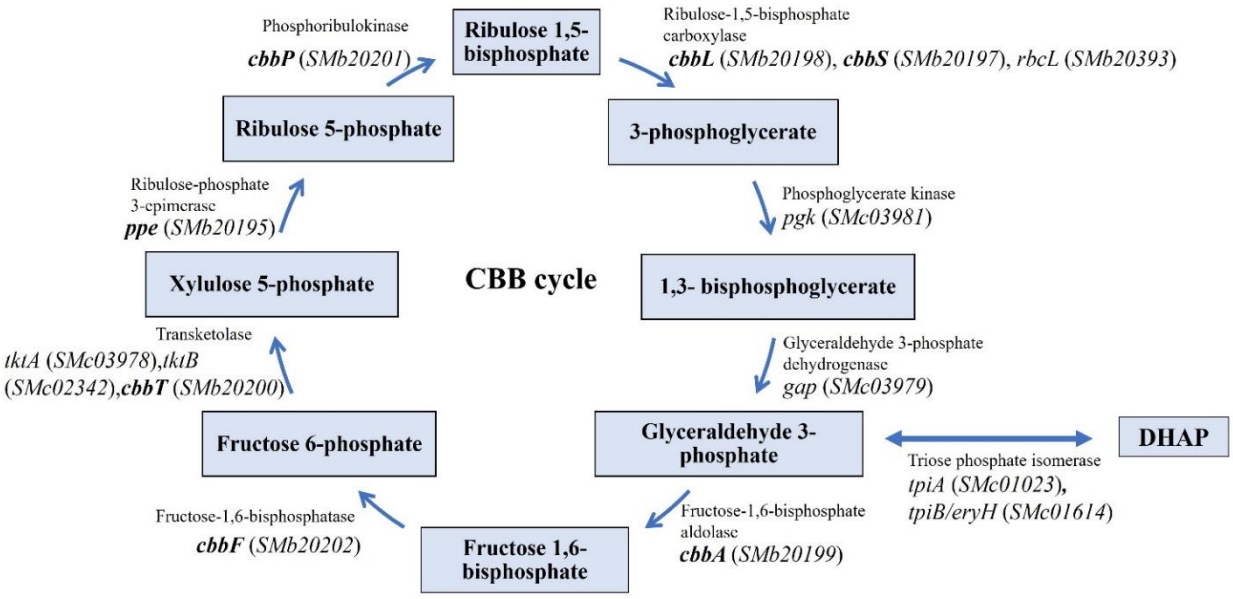

Figure 3 Calvin Benson-Bashham cycle in *S. meliloti*.
The arrows represent the direction of enzymatic reactions and the annotated genes in *S. meliloti* genome that encode for the enzymes are shown below. The genes that are localized in the *cbb* operon are in bold. The locus tags for the genes are indicated inside the brackets.

carboxylase (RuBisCO). RuBisCO catalyzes the reaction which uses atmospheric CO_2 and ribulose 1,5-bisphosphate (RuBP) to generate two 3C molecules which will be subsequently used to generate starch and regenerate RuBP (Kusian & Bowien, 1997) (Figure 3). *S. meliloti* contains the genes which could encode enzymes responsible for this entire cycle and the majority of those genes are clustered together in pSymB.

1.3.2 *cbb* genes and regulation

The *cbb* genes encode most of the enzymes required for the CBB cycle. Although the genes are often organized in an operon, the regulation and grouping of the genes can differ in different groups of organisms (Dangel & Tabita, 2015b). *S. meliloti* contains one of the most conserved gene arrangements (Kusian & Bowien, 1997). The 8 *cbb* genes (*cbbXSLATPF* and *ppe*) found in the *S. meliloti* genome are arranged in an apparent operon with a hypothetical protein (*SMb20194*) (Figure 4). The potential gene regulator *cbbR* is located adjacent to this cluster and is transcribed divergently. These genes are all localized on the chromid pSymB. The genes, *cbbS* and *cbbL* are annotated as subunits for the key enzyme RuBisCO. The genes *cbbX*, *cbbA*, *cbbT*, *cbbP* and *cbbF* are annotated as an ATPase, aldolase, putative transketolase, putative phosphoribulokinase and a putative bisphosphatase respectively. The gene *ppe* is annotated as a ribulose-phosphate 3-epimerase. A mutation in *cbbF* resulted a polar effect on the downstream genes terminating the autotrophic growth of *S. meliloti*, suggesting that these genes are required for CO_2 fixation and may be within an operon (Pickering & Oresnik, 2008).

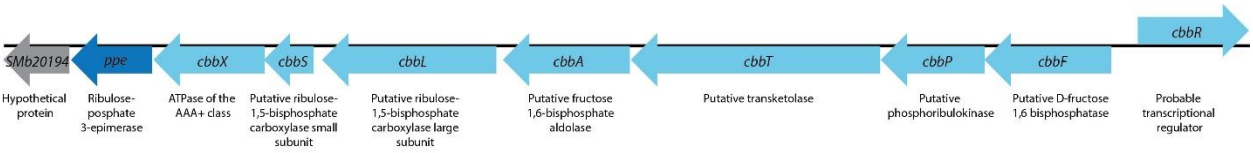

Figure 4. Genetic map of the *cbb* operon in *S.meliloti*.
Arrows indicate the open reading frames and the direction of transcription. The colors represent the putative functional class of each open reading frame according to the annotated genome of *S. meliloti.* (https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi).

1.3.3 *cbbR*; Lys-R type transcriptional regulator (LTTR)

The expression of genes in a *cbb* operon is highly regulated in the majority of CO_2 fixing organisms (Dangel *et al*., 2005). The key regulator is often a **L**ys-R **t**ype **t**ranscriptional **r**egulator (LTTR); *cbbR* (Dangel & Tabita, 2015b). The LTTRs are one of the most abundant class of regulators found in bacteria. These regulators can function as either positive or negative regulators (Lindquist *et al*., 1989; Schell, 1993). The characteristic features of these proteins are the N-terminus helix-turn-helix motif that acts as the DNA binding domain and the C-terminus recognition domain which consists of two distinct subdomains. LTTRs function as tetramers which binds to the promoter region of the operon (Schell, 1993).

cbbR is a LTTR which has been identified as the master regulator of the *cbb* cycle in autotrophic organisms. In most of the organisms, the *cbbR* gene is often located directly upstream of the *cbb* operon genes and the direction of transcription *cbbR* is often divergent to the *cbb* operon (Falcone & Tabita, 1993; Nargang *et al*., 1984). Gene regulation by CbbR is initiated with the binding of effector molecules to the effector domains of the protein. Like the other LTTRs, CbbR functions as a tetramer and the binding of the effector molecules to CbbR changes the angle of the bend occurs when CbbR binds to the promoter region (Schell, 1993). This alters the interaction with RNA polymerase to either initiate or increase the transcription. In addition to activation, the binding of some effector molecules can cause inhibition of transcription (Maddocks & Oyston, 2008). The effectors for CbbR are often metabolites from the pathways that it regulates. Effectors can differ from organism to organism (Maddocks & Oyston, 2008; Schell, 1993).

1.4 Overall Thesis Goals and Hypothesis

Previous work in the lab had led to the characterization of the non-oxidative branch of the PPP (Hawkins *et al*., 2018) as well as suppressors which led to the constitutive expression of *tktB* (Kaur, 2021). Strains were constructed that had both *tktA* as well as *tktB* mutated. Even though the major pathway for pentose catabolism is PPP, studies have reported evidence for the possible existence of alternative ways for the utilization of pentose sugars (Duncan, 1981; Geddes & Oresnik, 2012a). Therefore, the original goal of this work was to investigate whether the PPP could be bypassed by using a well characterized phosphoketolase from *Bifidobacterium*. However, it was noticed that the *S. meliloti* genome contained two genes that were annotated phosphoketolases; one in the chromosome and the other in pSymA. The initial experiments undertaken were to attempt to mutate the annotated phosphoketolases in the chromosome and to express the codon optimized *Bifidobacterium* phosphoketolase in *S. meliloti* to determine what affect this might have on the symbiotic phenotype associated with transketolase mutants. During the work, we fortuitously isolated mutations that suppressed *tkt* in *S. meliloti*. Thus, the goal became to identify and characterize the suppressor mutation. Once the mutation was found to be in the CBB pathway, we hypothesized if this mutation led to an over-expression of the CBB pathway, then this strain might be more competitive in a carbon limited environment, and by extension more competitive for nodule occupancy. Experiments were designed to test this hypothesis.

Chapter 2: Materials and Methods

2.1 Bacterial strains, plasmids and media

Bacterial strains and plasmids used in this study are listed in Table 1. *S. meliloti* strains were grown at 28º C and *E. coli* strains were grown at 37 º C on Luria-Bertani (LB) medium as complex medium which consists of 10 g of Tryptone, 5 g of NaCl and 5 g of yest extract in 1 L of water. Vincent's minimal medium (VMM) supplemented with either ribose, xylose, glucose or succinate at 15 mM was used as the defined medium for *S. meliloti*. VMM was made by combining VMM A (containing 1 g K_2HPO_4, 1 g KH_2PO_4 and 0.6 g KNO_3 in 1 L of H_2O), VMM B (containing 0.05 g $FeCl_3$, 1.25 g $MgSO_4$ and 0.5 g $CaCl_2$ in 500 ml of H_2O) and VMM C (containing 0.01 g biotin, 0.01 g thiamine-HCl and 0.01 g Ca pantothenate in 100 ml H_2O, ×1000 stock) in 1000:100:1 ratio. Antibiotics were added to LB as necessary. The following concentrations were used: streptomycin, 200 µg/ml; neomycin, 200 µg/ml; gentamicin, 20 µg/ml (for *E. coli*) and 200 µg/ml (for *S. meliloti*); chloramphenicol, 20 µg/ml; tetracycline, 5 µg/ml. All carbon sources and antibiotics were filter sterilized before adding to the medium.

2.2 Genetic techniques, plasmid construction and mutant generation

Transformations and conjugations were conducted as described previously (Finan *et al*., 1988). Standard techniques were used for genomic and plasmid DNA extractions, restriction enzyme digestion, ligation and gel electrophoresis (Sambrook *et al*., 1989).

To construct insertion mutations in *SMc04146*, a 400 bp internal fragment from *SMc04146* was PCR amplified using the primers 01 and 02 (Table 2). The amplified fragment was then purified and digested with BamHI and XhoI restriction endonucleases. These sites were incorporated into the primers.

Table 1 List of bacterial strains and plasmids

Strain or plasmid	Genotype	Reference
Strains		
S. meliloti		
Rm1021	SU47 *str-21* Sm^r	(Meade *et al*., 1982)
SRmD397	Rm1021, *tktA*::Tn5 Sm^r Nm^r	(Hawkins *et al*., 2018)
SRmD410	Rm1021, *tktA*::Tn5 *stkR1* Sm^r Nm^r	(Kaur, 2021)
SRmD661	Rm1021, Δ*tktB* Sm^r	(Kaur, 2021)
SRmD663	ΦSRmD397 (*tktA*::Tn5) -> SRmD661 (Δ*tktB*) Sm^r Nm^r	(Kaur, 2021)
SRmD714	SRmD663, *cbbR** Sm^r Nm^r	This work
SRmD715	SRmD663, *cbbR** pIW01 Sm^r Nm^r Tc^r	This work
SRmD716	SRmD663, *cbbR** pIW01 Sm^r Nm^r Tc^r	This work
SRmD719	SRmD715, *cbbR** (ΔpIW01) Sm^r Nm^r	This work
SRmD720	SRmD716, *cbbR** (ΔpIW01) Sm^r Nm^r	This work
SRmD722	Rm1021 *SMc04146*::pKnock-Gm Gm^r	This work
SRmD725	SRmD661 *SMc04146*::pKnock-Gm Gm^r	This work
SRmD733	SRmD714, *cbbL*::pKnock-Gm Sm^r Nm^r Gm^r	This work
SRmD734	Rm1021, *dctA*::pKnock-Gm Sm^r Gm^r	This work
SRmD735	SRmD714, *dctA*::pKnock-Gm Sm^r Nm^r Gm^r	This work
SRmD740	Rm1021 *cbbR**	This work
SRmD741	SRmD714 *cbbT*::pKnock-Gm Sm^r Nm^r Gm^r	This work
SRmD742	Rm1021 *cbbT*::pKnock-Gm Gm^r	This work
SRmD743	Rm1021 *cbbL*:: pKnock-Gm Gm^r	This work
SRmD744	SRmD740 *cbbL*:: pKnock-Gm Gm^r	This work
SRmD762	SRmD740 *cbbT*:: pKnock-Gm Gm^r	This work
SRmD763	SRmD741 pIW12 Sm^r Nm^r Gm^r Tc^r	This work
SRmD764	SRmD663 pIW12 Sm^r Nm^r Tc^r	This work
E. coli		
MM294A	*pro-82 thi-1 hsdR17 supE44*	(Finan *et al*., 1985)
MT607	MM294A *recA56*	(Finan *et al*., 1985)
MT616	MT607(pRK600)	(Finan *et al*., 1985)
DH5α	λ^-Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(r_k^- m_k^-) *supE44 thi-1 gyrA relA1*	(Hanahan, 1983)
DH5αλpir	λ pir lysogen of DH5α	(House *et al*., 2004)
Plasmids		
pKnock-Gm	Suicide vector for insertional mutagenesis; R6K ori RK4 *oriT* Gm^r	(Alexeyev, 1999)
pJQ200SK	Gene replacement suicide vector Gm^r	(Quandt & Hynes, 1993)

pRK7813	Broad-host range cloning vector Tc^r	(Jones & Gutterson, 1987)
pCO37	pRK7813 containing *attB* sites; Gateway-compatible destination vector	(Jacob *et al*., 2008)
pIW01	pRK7813/codon optimized *pkt* from *Bifidobacterium adolescentis*	This work
pIW05	400bp internal fragment of *SMc04146* cloned into pKnock-Gm, Gm^r	This work
pIW07	400bp internal fragment of *cbbL* cloned into pKnock-Gm Gm^r	This work
pIW08	pJQ200SK/*cbbR* flanking regions Gm^r	This work
pIW09	400 bp internal fragment of *dctA* cloned into pKnock-Gm Gm^r	This work
pIW11	400 bp internal fragment of *cbbT* cloned into pKnock-Gm Gm^r	This work
pIW12	pRK7813/*cbbT* Tc^r	This work
pJH109	pCO37/*tktA* Tc^r	(Hawkins *et al*., 2018)
pJH110	pCO37/*cbbT* Tc^r	(Hawkins *et al*., 2018)
pJH116	pCO37/*tktB* Tc^r	(Hawkins *et al*., 2018)

Table 2 List of primers

Primer	Primer sequence 5'-3'	Primer name
	***SMc04146*::pKnock-Gm**	
01	GACGTA**GGATCC**CGGATATCAGGCGCATTCAG	phk_pKn (Fw)
02	CGATAG**CTCGAG**CTGAACCGCAACCGCATAATC	phk_pKn (Rv)
03	CGGCTGCGTCCTGCCTATC	phk_conf (Fw)
04	GGCGACAAGTGGTCGTCTTC	phk_conf (Rv)
	***phk*(pRK7813)**	
05	GATCAAGCTTGGAGGAATACATGACCTCGC	phkB (Fw)
06	AGTCGAATTCTCACTCGTTGTCGCCGGCC	phkB (Rv)
	***cbbT*::pKnock-Gm**	
07	CTAGAG**GGATCC**TCGACCGATCTCTCCACCTC	cbbT_pkn (Fw)
08	GACGTA**CTCGAG**GCTCTACGGTCTGCTCGTAA	cbbT_pkn (Rv)
09	GTCGGCATGGCGATCGCCGAA	cbbT_conf (Fw)
10	CGGCCCTTGAAATTACCCGG	cbbT_conf (Rv)
	***cbbL*::pKnock-Gm**	
11	GATGCC**GGATCC**GGCAAGCCGCTGCTGGGTG	cbbL_pKn (Fw)
12	CGAGTC**CTCGAG**GCCGCGTATAGGTGCCGTGA	cbbL_pKn (Rv)
13	GCAGTACTTCTGCTATGTCGC	cbbL_conf (Fw)
14	GAGGCGACGGGCATGACCT	cbbL_conf (Rv)
	cbbR*	
15	cgaattcctgcagcccggggACTGCGGGACCACGAGATCGAC	cbbR_L_(Fw)
16	cggtgaaccgCTGCCGGCGGATCGGCGT	cbbR_L_(Rv)
17	ccgccggcagCGGTTCACCGTGTCGCGC	cbbR_R_(Fw)
18	agggaacaaaagctggagctCCCGTGGGGTCGTTGTGC	cbbR_R_(Rv)
19	CGATCGAGGACAGCCTCATCG	cbbR_SNPconf (Fw)
20	GGATCGGTTCATCCCCGTCC	cbbR_SNPconf (Rv)
	qRT-PCR	
21	GGCGATCATCCTCTGGTCTT	cbbR_qRT (Fw)
22	TTTCGAGCGATATCCGGGTT	cbbR_qRT (Rv)
23	CACGATCTTCTCGGTGCTTC	cbbF_qRT (Fw)
24	ACTTGTAGGCCTCGACGAAA	cbbF_qRT (Rv)
25	GAAGATCCATCGCGACAAGG	cbbP_qRT (Fw)
26	GCGGGCAGATATAGTTCACG	cbbP_qRT (Rv)
27	GATGGCTCTGGAGGTCATCA	cbbT_qRT (FW)
28	GAAATTACCCGGCGAGATGG	cbbT_qRT (Rv)
29	GTTCGCAACTGCTTACCGAT	cbbA_qRT (Fw)
30	GCATCACGATATGGGTGTCC	cbbA_qRT (Rv)
31	GACCGCTACCTCTACTGCAT	cbbL_qRT (Fw)

32	ATAGCAGTCCAACCGACGAT	cbbL_qRT (Rv)
33	CCCTCATCCCAGAAACACCT	cbbS_qRT (Fw)
34	CTGGAATCGAAGGCGTTGAG	cbbS_qRT (Rv)
35	AGGCCTACTACCTCTACCGT	cbbX_qRT (Fw)
36	AAGTCTATGTGGTGGGCGAT	cbbX_qRT (Rv)
37	TCTGATGATCGAACCACCCA	ppe_qRT (Fw)
38	GCCGGTCCAGCACATATTC	ppe_qRT (Rv)
39	GACCTCAGGGCTCATGGAAT	SMb20194_qRT (Fw)
40	GGATTGGTCCTCGTCATGGA	SMb20194_qRT (Rv)
	***cbbT* (pRK7813)**	
41	GATCAAGCTTGGAGGAATACATGAATGTTTCGCAGCAG	cbbTorf (Fw)
42	AGTCGAATTCTCATGCCTCCTCGGATGATCTGACG	cbbTorf (Rv)
	***dctA*::pKnock**	
43	GACGTA**GGATCC**GGTGCTCTTCGGTATCTCGCT	dctA_pKn (Fw)
44	CGATAG**CTCGAG**GACGACCGAGCGCTTGCAGC	dctA_pKn (Rv)
45	GTCGGCCTCGTCGTCGCAA	dctA_conf (Fw)
46	GTTGCGGCAAGCGTGATGAAG	dctA_conf (Rv)

Highlighted in bold are restriction recognition sites.

Lowercase letters represent the overlapping region with the vector plasmid.

The plasmid vector pKnock-Gm was also digested with the same enzymes. The digested fragment was purified and cloned into the pKnock vector using T4 DNA ligase to create pIW05. The construct was then transformed into *E. coli* strain DH5αλpir, and subsequently conjugated into Rm1021 and SRmD661. A colony from each conjugation was single colony purified and the single crossover mutation of *SMc04146* gene was confirmed by PCR amplification using primers 03 and 04 (Table 2). The PCR product was nucleotide sequenced. The resultant strains were SRmD722 and SRmD725 respectively (Table 1).

The construction of SRmD741, SRmD742 and SRmD762 were carried out using a similar approach. Basically, a 400 bp internal fragment of *cbbT* was PCR amplified using primers 07 and 08. The fragment was then cloned into pKnock-Gm using BamHI and XhoI restriction enzymes to create pIW11. This construct was transformed into DH5αλpir and subsequently conjugated separately into SRmD714, Rm1021 and SRmD740 respectively and selected on $LB^{Sm,Gm}$. The knockout of the genes was confirmed by PCR amplification and sequencing using primers 09 and 10.

The construction of SRmD733, SRmD743 and SRmD744 were carried out by PCR amplifying an internal 400 bp fragment from *cbbL* using 11 and 12 primers and cloned into pKnock-Gm using BamHI and XhoI restriction enzymes to create pIW07. This construct was transformed into DH5αλpir and subsequently conjugated separately into SRmD714, Rm1021 and SRmD740 respectively and selected on $LB^{Sm,Gm}$. The final mutation was confirmed using primers 13 and 14.

The construction of SRmD734 and SRmD735 were carried out by PCR amplifying an internal 400 bp fragment from *dctA* using 43 and 44 primers and cloned into pKnock-Gm using BamHI and XhoI restriction enzymes to create pIW09. This construct was transformed into

DH5αλpir and subsequently conjugated separately into Rm1021 and SRmD714 respectively and selected on $LB^{Sm,Gm}$. The final mutation was confirmed using primers 45 and 46.

For the construction of expression constructs, the broad host range vector pRK7813 was used. To construct pIW01, the *phk* gene sequence obtained from *Bifidobacterium adolescentis* was first codon optimized to *S. meliloti* using the IDT codon optimization tool, then purchased. The sequence was then amplified using primers 05 and 06 and cloned into pRK7813 using HindIII and EcoRI. Then the construct was transformed into DH5α followed by conjugation into SRmD663 and selected on LB^{SmTc}. Isolated colonies were named SRmD714, SRmD715 and SRmD716. pIW12 was constructed similarly by PCR amplifying *cbbT* in *S. meliloti* using primers 41 and 42 and the construct was subsequently conjugated into SRmD741 and SRmD663.

To introduce the point mutation into Rm1021 to create SRmD740, the plasmid was constructed using Gibson assembly. Primers for the desired region of the *cbbR* gene were designed using NEBuilder Assembly tool (https://nebuilder.neb.com/#!/, New England Biolabs, Inc.). Two primer pairs (15,16 and 17,18) were designed to flank about 400 bp from both upstream and downstream of the position of the point mutation to be introduced. The single nucleotide change was incorporated at the desired position of both primer pairs. PCR was carried out using Q5 polymerase. Assembly of amplified fragments to the pJQ200SK vector was done using NEBuilder HiFi DNA Assembly Master mix from New England Biolabs, Inc. to construct pIW08. The vector was digested with BamHI and SacI and the vector and the insert were added in 1:2 molar ratio. The construct was then transformed into *E. coli* DH5α competent cells and selected for Gm resistance. Conjugation of pIW08 into Rm1021 was done utilizing MT616 as the mobility plasmid and selection for the single crossover was done on $LB^{Sm,Gm}$. Selected colonies were purified on $LB^{Sm,Gm}$ and 5 ml of LB was inoculated with purified colonies. The

double crossover was selected on LB containing 10% sucrose. The colonies that were resistant to sucrose and sensitive to Gm were selected and PCR amplified using primers 19 and 20. The introduction of the point mutation was verified by sequencing.

2.3 RNA extraction, cDNA synthesis and qRT-PCR

Total RNA was extracted using QIAgen RNeasy Kit, QIAshredder and DNase 1 kit. Strains Rm1021 and SRmD714 strains were grown in VMM supplemented with 15 mM ribose up to OD600: 0.4-0.6. 1 ml of culture was spun down and the supernatant was removed followed by resuspension of the pellet in 100 µl of TE+Lysozyme. RNA was isolated using QIAgen RNeasy Kit (Qiagen, Chatsworth, CA) under RNase free conditions and RNA samples were treated with QIAgen DNase kit to remove DNA contaminants during the purification process. Isolated RNA was quantified spectrophotometrically using Nano Drop (Thermofisher Scientific) and proceeded to cDNA synthesis.

cDNA synthesis was done using a VILO cDNA synthesis kit from Invitrogen. One µg of RNA was used to make a 14 µl reaction with 4 µl of 5x reaction mix, 2 µl of 10x reaction mix and x µl of H_2O. PCR was run using cDNA synthesis program with 25° C for 10 m, 42° C for 60 m and 85° C for 5 m. cDNA was diluted to an equivalent of 100 ng RNA added. Quantification and verification were done using a spectrophotometer and gel electrophoresis. qRT-PCR was conducted using SYBR Green RT-PCR kit from Invitrogen and approximately 200 ng of cDNA was used as the template for the reaction. A housekeeping gene for *S. meliloti SMc00128* was used as the internal control to compare the results (Krol & Becker, 2004). Three independent biological replicates were used and the results from qRT-PCR were analyzed using the $2^{\wedge-(\Delta\Delta Ct)}$ method (Livak & Schmittgen, 2001). Specific primers used for qRT-PCR are listed in Table 2.

2.4 DNA library preparation for whole genome sequencing

Bacteria were cultured in LB broth for two days, and genomic DNA was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen). Nanopore library preparation was carried out using the kit SQK-NBD11.24 (Oxford Nanopore). Reagents used were purchased from Oxford Nanopore unless otherwise indicated. DNA repair and end preparation was done using the NEBNext companion module from New England Biolabs. NEBNext FFPE DNA Repair Buffer (0.875 µl), Ultra II End-prep reaction buffer (0.875 µl), Ultra II End-prep enzyme mix (0.75 µl), NEBNext FFPE DNA Repair Mix (0.50 µl) were combined with 1000 ng of DNA in 12 µl nuclease-free water. The mixture was spun down and then incubated in a thermal cycler at 20°C for 5 minutes and 65°C for 5 minutes. 15 µl of AMPure XP Beads (AXP) was added to each end-prep reaction and samples were inverted by hand for 5 minutes at room temperature. The samples were briefly spun down, and the beads were pelleted using a magnet. The supernatant was pipetted off and the beads were gently washed with 70% ethanol. This was repeated twice, followed by a brief spin-down and removal of any residual ethanol. The AMPure bead pellet was then resuspended in 10µl of H_2O and incubated for 10 minutes at room temperature. Beads were then pelleted again on the magnet, and the eluate from each sample was collected into a clean 1.5 ml Eppendorf DNA LoBind tube and quantified using a Qubit 2.0 fluorometer. These samples were used for native barcoding.

For native barcode ligation, a unique barcode was selected for every sample to be run together on the same flow cell. The barcode ligation reaction mixture was prepared by mixing End-prepped DNA (7.5 µl), Native barcode (2.5 µl), Blunt/TA Ligase Master Mix (New England Biolabs) (10 µl) per sample followed by incubation at room temperature for 20 minutes. 2 µl of EDTA was added to each tube and mixed thoroughly to stop the reaction. Then the barcoded

samples were pooled in a clean 1.5 ml Eppendorf DNA LoBind tube. The pooled mixture was washed as described previously using AMPure XP Beads (AXP) (8 µl per sample). The pellet was resuspended in 35 µl of nuclease-free water and incubated for 10 minutes at 37°C followed by pelleting the beads using magnets until the elute was clear. 35 µl of eluate was collected into a clean 1.5 ml Eppendorf DNA LoBind tube to be used for adapter ligation and cleanup.

For adapter ligation, pooled barcoded sample (30 µl), Adapter Mix II H (5 µl), NEBNext Quick Ligation Reaction Buffer (5X) (10 µl) and Quick T4 DNA Ligase (5 µl) from the NEBNext Quick Ligation Module (New England Biolabs) was mixed and incubated for 20 minutes at room temperature. 20 µl of AMPure XP Beads (AXP) were used to bind and pellet the DNA followed by washing twice with 125 µl of Long Fragment Buffer (LFB). Each time the pellet was fully resuspended in LFB before being pelleted on the magnet. The beads were pelleted again, dried, and then were resuspended in 15 µl Elution Buffer (EB). The suspension was incubated for 10 minutes at 37°C with periodic agitation. Then the beads were pelleted using magnet and the clear, colorless elute was collected and quantified with a Qubit fluorometer to obtain final library concentration. Approximately 10 fmol of library was used to load the flowcell for sequencing.

2.5 Whole genome sequencing using Nanopore MinION Mk1B

All genomes were sequenced using as previously described using the Nanopore MinION Mk1B (Hawkins *et al*., 2022). Sequencing was carried out using a Nanopore MinION Mk1B with kits SQK-LSK-109 and EXP-NBD-104, R10.3 flowcells, and base-calling was handled by Guppy-GPU (Wick *et al*., 2019). Default parameters were used for all software in the analysis.

Sequencing was stopped when on average there was enough data present to provide ~100x coverage on all genomes.

2.6 Genome assembly and variant calling

Reads were trimmed using BBduk (Bushnell *et al*., 2017), and then De-Novo genome assembly was carried out using Flye followed by 3 rounds of polishing using Minimap2 (Kolmogorov *et al*., 2019; Li, 2018). Genome completion was estimated using checkM (Parks *et al*., 2015) and found to be above 99% with an estimated 1% contamination in all assemblies. Once all genomes were assembled, Mauve was used to align all genomes to the assembled Rm1021 to determine if any major rearrangements had occurred.

To determine any mutations that may have arisen in the selected strains, reads were mapped to the lab De Novo assembled Rm1021 strain using Minimap2 with default parameters. The reads of Rm1021 were also mapped to the De Novo assembly using Minimap2 to determine any possible errors that may appear due to issues with assembly. After assembly a list of mutations for each strain compared to Rm1021 was obtained, limited to mutations that occurred in above 70% of the reads. Mutations in each strain were filtered to not include mutations that appeared when mapping Rm1021 reads to itself as these were likely errors in assembly.

2.7 Plant assays, competition for nodule occupancy and nodule kinetics

Plant nodulation assays were conducted as described previously (Poysti & Oresnik, 2007). Briefly, alfalfa seeds were surface sterilized using 1 % bleach and autoclaved H_2O and germinated on H_2O agar plates for 3days. The seedlings were then planted on sterilized Leonard jar assemblies containing 1:1 sand and vermiculite and the nitrogen free Jensen’s medium

comprises of 1 g $CaHPO_4$, 0.2 g K_2HPO_4, 0.2 g $MgSO_4.7H_2O$, 0.2 g NaCl, 0.1 g $FeCl_3$ and 1 ml of 1000x trace elements in 1 L of H_2O at pH 7.0 (Glazebrook & Walker, 1991). The 1000x trace elements solution comprises of 1.0 g H_3BO_3, 1.0 g $ZnSO_4.7H_2O$, 0.5 g $CuSO_4.5H_2O$, 0.5 g $MnCl_2.4H_2O$, 1.0 g $NaMoO_4.2H_2O$, 10 g EDTA, 2.0 g NaFeEDTA and 0.4 g Biotin in 1 L of H_2O. Then the seedlings were inoculated with approximately 10^7 bacteria suspended in 10 ml of sterile H_2O, after growing the seedlings for 2 days. Plants were grown for 28 days and the shoots were harvested and dried to measure the dry weights. There were 3 biological replicates for each treatment and 10 plants for each replicate.

For competition assays, the alfalfa plants were inoculated with the mixture of two strains Rm1201(WT) and SRmD714 at ratios 1:1, 1:5 and 1:10. The initial inoculum ratios were determined by diluting the inoculum mixture up to 10^{-5} and spread plating on LB^{Sm} followed by patching the single colonies on LB and LB^{Nm}. After growing for 28 days, the plants were harvested, and root nodules were collected. They were surface sterilized using 1% bleach followed by washing with sterile dH_2O. Then the nodules were placed in PCR tubes and crushed in 50µl of sterile dH_2O to create a suspension and spot plated on LB and LB^{Nm} to determine the final ratios. About 60 nodules per inoculum ratio were used for the determination of final percentages.

For nodule kinetics, the alfalfa seedlings were surface sterilized as previously described. Then they were germinated on water agar for 2 days and the seedlings were then transferred into Jensen's agar slants. The seedlings were allowed to grow for 2 days and inoculated with approximately 10^7 bacteria in 100µl of ddH_2O. The rate of nodule formation was observed over 21 days. Five independent biological replicates of wild-type Rm1021 and SRmD714 were used. Each replicate consisted of 10 seedlings.

2.8 Growth in soil

Soil preparation and the assessment of growth was done as described previously with few modifications (DiCenzo *et al*., 2014). Soil samples were obtained from agricultural fields in Carman and Kelburn (Manitoba, Canada). A third sample which consists of 1:1 sand and vermiculite was used as a reference. The soil samples were dried for about 9 days and the large material was removed. Then the dried soil was passed through a 2 mm sieve to remove the remaining large particles. The soil samples, and the sand/vermiculite mixture were autoclaved and let cool for 3 days. About 40 g from each preparation was added to 200 ml glass bottles. The sand/vermiculite was supplemented with Jensen's medium to provide the essential nutrients and autoclaved.

S. meliloti strains to be tested were grown in 5 ml of LB broth overnight at 28º C and washed with 0.85 % saline followed by washing with ddH_2O for three times. Then the cell pellets were resuspended in ddH_2O to an OD600 of 1 and a serial dilution was prepared up to 10^{-4} which contains approximately 2×10^5 CFU/ml. One ml from this dilution was used to inoculate each soil sample. Additional 1.38 ml of ddH_2O was added to achieve a soil mesocosm with 40 g dry weight of soil with 20 % moisture (w/v) and the samples were incubated at room temperature in dark. The moisture level was maintained by the addition of ddH_2O couple of times per week. To determine the competition between wildtype Rm1021 and desired mutant strains, the soil mesocosm was inoculated with both strains together.

In order to determine the cell density, 0.6 g of soil was removed aseptically from each sample and was resuspended in 1 ml of sterile 0.85 % saline. Samples were mixed vigorously by vortexing and the soil particles were pelleted by centrifugation at 60 g for 1 minute. The

supernatant was collected, and serial diluted for plating on LB^{Sm} and calculating the CFU/g to determine the growth in soil. CFU/g was calculated for a period of 34 days.

Chapter 3: Results

3.1 Isolation of a strain in *tktA/B* genetic background that grows on defined medium.

The pentose phosphate pathway (PPP) is key for pentose sugar metabolism. However, there is evidence that suggests the possibility of *S. meliloti* having alternative pathways or metabolic bypasses for pentose sugar metabolism (Duncan, 1981; Geddes & Oresnik, 2012; Geddes & Oresnik, 2014). A phosphoketolase reaction uses either fructose 6-phosphate or xylulose-5-phosphate to generate an aldose phosphate and acetyl phosphate, thus bypassing the PPP (Sánchez *et al*., 2010). Based on the annotated genome of *S. meliloti*, and the KEGG metabolic pathways map for Rm1021, it was found that *SMc04146* was annotated as a putative phosphoketolase.

A preliminary experiment was conducted to determine if *SMc04146* may play a role in bypassing the PPP by introducing pKnock-Gm insertion mutations in both the wild-type as well as SRmD661, which carries a deletion of *tktB.* This yielded SRmD722 and SRmD725 respectively. These mutants were tested for carbon metabolism phenotypes on defined medium supplemented with either ribose or xylose. The results showed that addition of this insertion did not have an effect on its ability to utilize ribose or xylose under the tested conditions (Table 3).

Phosphoketolase from *Bifidobacterium* sp. has been utilized in different metabolic engineering studies to bypass the core metabolic pathways for the purpose of carbon conservation (Krüsemann *et al*., 2018; Yang *et al*., 2016). As an alternative strategy to bypass the PPP it was reasoned that if this gene were expressed in an appropriate genetic background, it could bypass the PPP.

To carry this out the phophoketolase gene sequence obtained from *Bifidobacterium adolescentis* was codon optimized for expression in *S. meliloti* (Figure 5),

Table 3 Carbon metabolism phenotypes for *SMc04146* mutants

Strain	**Genotype**	**LB**	**LB^{Gm}**	**Ribose**	**Xylose**
Rm1021	Wild-type	++	-	++	++
SRmD397	*tktA::Tn5*	++	-	-	-
SRmD663	*tktA*::*Tn5, ΔtktB*	+	-	-	-
SRmD661	*ΔtktB*	++	-	++	++
SRmD722	*SMc04146*::pKnock	++	++	++	++
SRmD725	*ΔtktB,SMc04146*::pKnock	++	++	++	++

Growth; ++, like wild-type; +, slow growth; – , no growth

Abbreviations; LB, Luria Bertani; LB^{Gm}, Luria Bertani with gentamicin

Figure 5. Comparison of the phosphoketolase gene obtained from *Bifidobacterium* sp. and the codon optimized sequence for *S. meliloti*.
The upper row represents nucleotides. Bases highlighted in black are unchanged. Bases not highlighted are changed from the original sequence to that listed. The colored letters in the bottom row represents the encoded amino acids.

and the synthesized gene was subsequently cloned into pRK7813 such that it was constitutively expressed by the p_{tac} promoter that flanked the multiple cloning site yielding pIW01. This plasmid was subsequently conjugated into SRmD663 which carries mutations in both *tktA* and *tktB* and is unable to grow on ribose or xylose. When this was streaked onto defined medium containing either ribose or xylose the resulting strain showed similar growth phenotypes to its parent strain SRmD663, suggesting that the presence of the pIW01 was unable to restore the growth (Table 4).

It was observed that upon extended incubation (7 days), that several isolated colonies appeared on the plate streaked with SRmD663(pIW01). Three of these colonies were single colony purified and retested for their ability to utilize ribose as the sole carbon source. These strains were named SRmD714, SRmD715, and SRmD716.

The strains SRmD715 and SRmD716 both carried the plasmid pIW01 whereas SRmD714 did not. This suggests that the phenotypic reversion in SRmD714 was not due to the presence of the plasmid, but it could be due to a spontaneous mutation that arose in SRmD663. To determine if the reversion event in SRmD715 and SRmD716 was dependent upon the presence of pIW01, these strains were streaked onto LB agar that did not contain tetracycline. Approximately 10-20 of the resultant colonies were screened for loss of the plasmid. Two tetracycline sensitive colonies were isolated and designated SRmD719 (derived from SRmD715) and SRmD720 (derived from SRmD716). The results show that both SRmD719 and SRmD720 retained their ability to utilize ribose upon loosing pIW01, suggesting that the observed reversion phenotype was due to a genomic mutation and is independent of the phosphoketolase expressing construct (Table 5). Based on these preliminary experiments, we hypothesized that a second site suppressor mutation was present in these strains.

Table 4 *Bifidobacterium* phosphoketolase does not revert a transketolase mutant strain.

Strain	**Genotype**	**LB**	**LB^{Tc}**	**Ribose**	**Glucose**
Rm1021	Wild-type	++	-	++	++
SRmD663	*tktA*::*Tn5, ΔtktB*	+	-	-	-
SRmD663(pIW01)	*tktA*::*Tn5, ΔtktB* (*phk*)	+	+	-	-

Growth; ++, like wild-type; +, slow growth; –, no growth

Abbreviations; LB, Luria Bertani; LB^{Tc}, Luria Bertani with tetracycline.

Table 5 pIW01 does not contribute to phenotypic reversion of transketolase mutants.

Strain	**Genotype**	**LB**	**LB^{Tc}**	**Ribose**
Rm1021	Wild-type	++	-	++
SRmD663	*tktA*::*Tn5*, *ΔtktB*	+	-	-
SRmD714	*tktA*::*Tn5*, *ΔtktB*,	++	-	++
SRmD715 (pIW01)	*tktA*::*Tn5*, *ΔtktB*,	++	++	++
SRmD716 (pIW01)	*tktA*::*Tn5*, *ΔtktB*,	++	++	++
SRmD719	SRmD715 cured of pIW01	++	-	++
SRmD720	SRmD716 cured of pIW01	++	-	++

Growth; ++, like wild-type; +, slow growth; –, no growth

Abbreviations; LB, Luria Bertani; LB^{Tc}, Luria Bertani with tetracycline.

3.2 Identification of the point mutation in *cbbR*

To determine if SRmD714, SRmD715, and SRmD716 contained a mutation capable of reverting the phenotype of the parental strain SRmD663, complete genome sequencing was performed for all three strains using the Nanopore MinION Mk1B as previously described (Hawkins *et al*., 2022). Analysis of the sequenced genomes revealed that each of the strains contained the same mutation in a LysR-type regulator annotated as *cbbR* (Table 6). Since each of the three strains had the same mutation in *cbbR*, this suggests that they are siblings and not independent mutation events. The single nucleotide change was in the 811th nucleotide of *cbbR* and it resulted in a base change from T to C, which resulted in the codon TGG becoming CGG, thus changing the amino acid at the 271st position of CbbR from Tryptophan (W) to an Arginine (R) (Figure 6). This allele was termed *cbbR**. The carbon phenotypes of the *cbbR** strain SRmD714 was further confirmed growing on minimal medium (Table 7).

The structures of both wild-type and mutant CbbR protein were predicted using the Phyre.2 program to identify the position and possible effects that this mutation could have on the protein structure and function (Kelley *et al*., 2015). The mutation is predicted to be within the recognition domain or effector domain in the C terminus of the protein (Figure 7). This region is responsible for interacting with different effector molecules involved in the control of the transcription of the gene (Dangel & Tabita, 2015b).

SRmD663 had been previously characterized and suppressors had not been previously observed (Kaur, 2021). We wanted to determine the frequency of finding second site suppressors in this *tktA/B* background. To carry this out SRmD663 was incubated in VMM supplemented with ribose for 7 days and subsequently plated on the same medium. Putative suppressor mutants

Table 6 Nanopore sequencing parameters and SNP identification parameters

Sequencing parameters (bp)	**Location**	**SRmD714**	**SRmD715**	**SRmD716**
Average read length	Chromosome	14335	14951	14795
	pSymA	14983	15842	15586
	pSymB	14915	15638	15212
Number of reads	Chromosome	73028	33458	40063
	pSymA	32348	18362	20438
	pSymB	36781	21190	23082
N50	-	26093	25047	24227
Assembly parameters (bp)				
Coverage	-	202	156	178
Contigs	-	3	3	3
Total bases in contigs	-	6680560	6680565	6680568
Minimum contig length	-	1359895	1359904	1359903
Maximum contig length	-	3637329	3637311	3637331
SNP identification				
CDS position	Chromosome	516	516	516
	pSymB	811	811	811
Change	Chromosome	A -> C	A -> C	A -> C
	pSymB	T -> C	T -> C	T -> C
Codon change	Chromosome	GGA -> GGC	GGA -> GGC	GGA -> GGC
	pSymB	TGG -> CGG	TGG -> CGG	TGG -> CGG
AA change	Chromosome	None	None	None
	pSymB	W -> R	W -> R	W -> R
Polymorphism type	Chromosome	SNP (transversion)	SNP (transversion)	SNP (transversion)
	pSymB	SNP (transition)	SNP (transition)	SNP (transition)
Variant frequency	Chromosome	94.30%	93.60%	96.60%
	pSymB	96.40%	95.30%	98.50%
Protein effect	Chromosome	None	None	None
	pSymB	Substitution	Substitution	Substitution

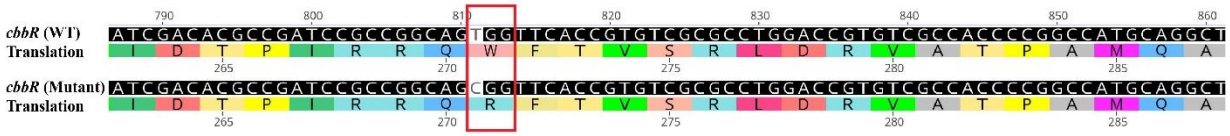

Figure 6. Sequence comparison of the wild type and mutant *cbbR*
The upper rows represent the base sequence and the bottom row represent the respective amino acids. The different colors represent different amino acids. The change identified is within the red block.

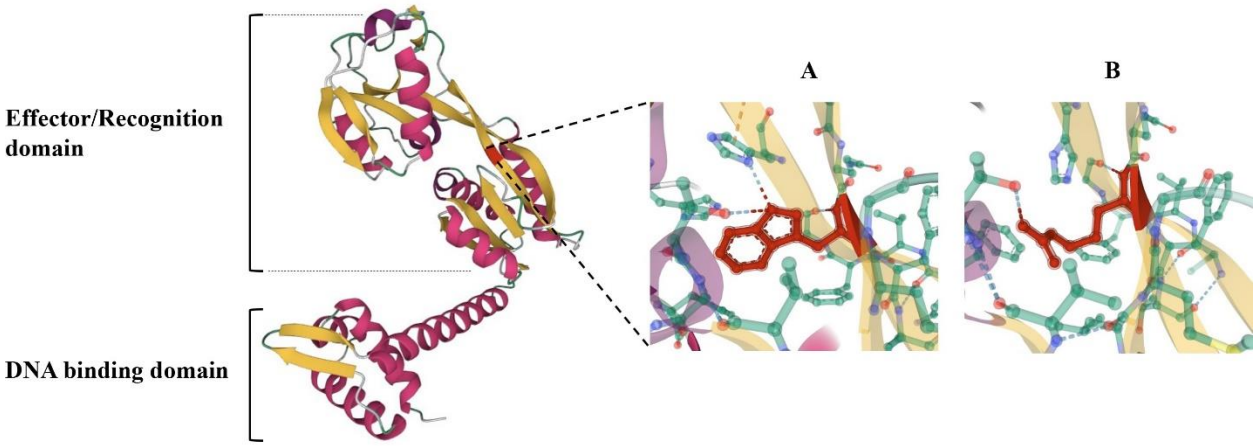

Figure 7. Predicted 3D structure of CbbR.
Panel A shows the wild-type residue (W), Panel B shows the mutant residue (R). The secondary structures of the protein are shown in different colors; alfa helix in magenta and beta strands in yellow. Structure prediction was done using Phyre.2 protein structure prediction program and visualization and modifications were done using Mol* Viewer.
(Sehnal *et al*., 2021)

Table 7 Carbon phenotypes of *cbbR** containing strains.

Strain	**Genotype**	**LB**	**LB^{Nm}**	**Ribose**	**Xylose**	**Glucose**	**Succinate**
Rm1021	Wild-type	++	-	++	++	++	++
SRmD397	*tktA*::*Tn5*	++	++	-	-	-	-
SRmD663	*tktA*::*Tn5*, *ΔtktB*	+	+	-	-	-	-
SRmD714	*tktA*::*Tn5*, *ΔtktB*, *cbbR**	++	++	++	++	++	++

Growth; ++, like wild-type; +, slow growth; –, no growth

Abbreviations; LB, Luria Bertani; LB^{Nm}, Luria Bertani with neomycin.

were observed at a frequency of $6.03\times10^{-8} \pm 1.2\times10^{-8}$. A few of these colonies were purified and tested for their ability to utilize either ribose, xylose, glucose or succinate as the sole carbon source and they were able to grow on defined medium without amino acid supplementation like the wild-type. These strains were not further characterized and were not retained.

3.3 Expression of *cbb* operon is highly induced in the *cbbR** strain.

In *S. meliloti*, the *cbb* operon consists of 10 genes which encode enzymes involved in the Calvin Benson cycle including a putative transketolase; *cbbT* (Figure 4). It has been previously reported that these genes are required for the autotrophic growth of *S. meliloti*, and a mutation in *cbbF* resulted in a polar effect terminating the autotrophic growth (Pickering & Oresnik, 2008). The *cbbR* gene is divergently transcribed and encodes a LysR-type transcriptional regulator which has been shown to control the expression of *cbb* genes in other organisms (Dangel & Tabita, 2015b).

The mutation in *cbbR* which restored the growth of SRmD663 is in the effector domain of CbbR (Figure 7). This domain functions as a controller for the transcription of *cbbR* upon binding of the effector molecules (Dangel & Tabita, 2015b). In other organisms it has been found that positive effector molecules such as RuBP induce the transcription of the operon, while negative effectors such as phosphoenol pyruvate inhibit or decrease the transcription (Dangel & Tabita, 2015b). Changes in the effector domain of CbbR can impact the affinity of effector molecules that control the levels of *cbb* transcription (Dangel & Tabita, 2015b). Therefore, it was hypothesized that the change of amino acid in the effector region from W to R might have caused a significant change that resulting in increased transcription of *cbb* operon leading to an

upregulation of *cbbT*, thus complementing the loss of the other two mutated transketolases. To test this hypothesis, qRT-PCR was performed targeting the 10 genes found at this locus. To carry this out the wild-type Rm1021 and SRmD714, which carries the *cbbR** allele, were grown overnight in VMM-ribose broth. The cells were pelleted, and the total RNA was extracted and assayed. The gene *SMc00128* was used as the internal control for the qRT-PCR analysis (Krol & Becker, 2004).

The results showed an increase in the expression of all genes compared to the wild-type (Figure 8,Figure 9). This is consistent with the hypothesis that *cbbFPTALSC*, *ppe* and *SMb20194* form an operon. Of note, the next gene to the left of *SMb20194* (as depicted) is transcribed on the other strand such that transcription is convergent. The data also shows that the *cbbR* is upregulated by 6-fold suggesting that *cbbR* autoregulates its own transcription (Figure 9). Consistent with our hypothesis, the expression of *cbbT* increased approximately 200-fold compared to the wild-type (Figure 9).

3.4 *cbbT* is responsible for the suppression of auxotrophic phenotypes of SRmD663

It was previously shown that the overexpression of *cbbT* was unable to restore the growth of a *tktA* mutant strain on minimal medium (Hawkins *et al*., 2018). However, it is evident that *cbbT* is highly upregulated due to the *cbbR** mutation and the carbon phenotypes of SRmD714 are similar to that of wild-type. To determine if the restoration of growth is due to *cbbT*, it was hypothesized that if an insertion mutation was constructed in *cbbT*, this should lead to the loss of phenotypic suppression. The mutation was constructed in SRmD714 background which carried a deletion of *tktB*, a Tn*5* insertion in *tktA*, as well as the *cbbR** allele. This mutant was designated

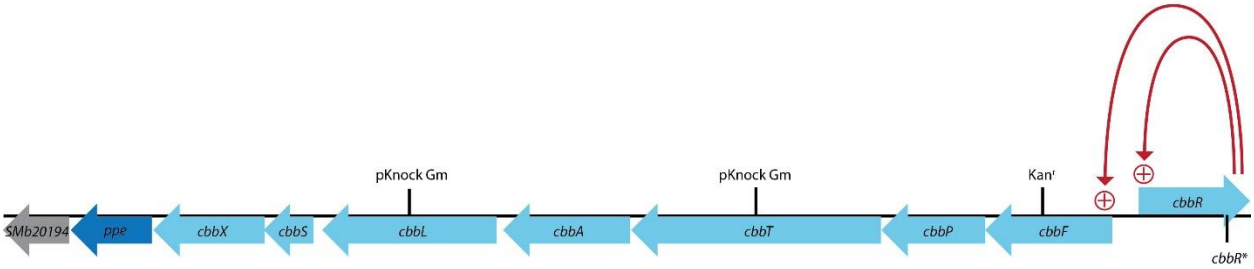

Figure 8. Genetic map of the *cbb* operon.
Arrows indicate the open reading frames and the direction of transcription. The vertical bars represent the approximate sites of mutations. The colors represent the putative functional class of each open reading frame according to the annotated genome of *S. meliloti* (https://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi). Red curved arrows show regions where CbbR may bind and affect transcription. Circled plus signs denote positive regulation.

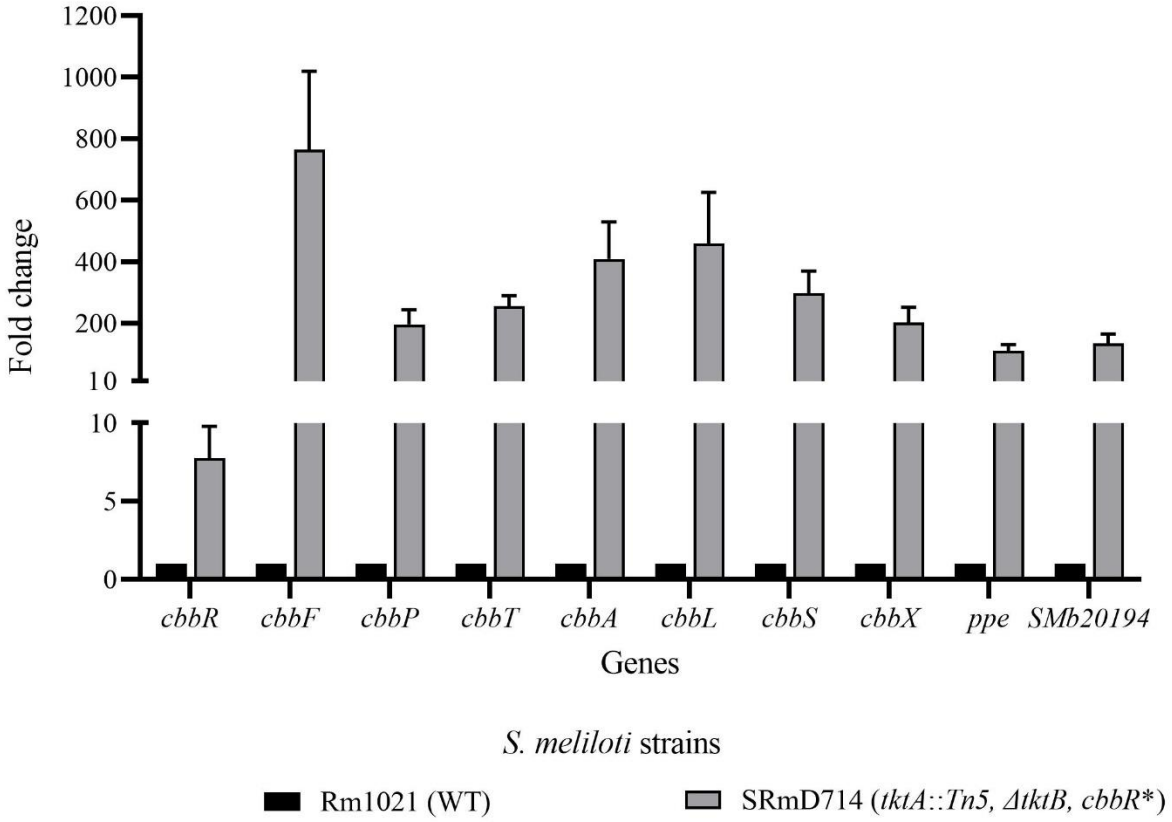

Figure 9. qRT-PCR analysis of *cbb* operon genes in *cbbR** mutant strain compared to the wild type Rm1021.
Graph shows the average fold change of two biological replicates.

as SRmD741. This mutation was also constructed in a Rm1021 background and was designated SRmD742. The results show that SRmD741 showed a similar phenotype on VMM supplemented with either ribose, glucose or succinate to that of either SRmD397 (containing *tktA*), or SRmD663 (containing *tktA*::Tn5 and *ΔtktB*) (Table 8). As anticipated, the *cbbT* insertion in SRmD742 did not have a phenotype discernable from the wild-type (Table 8).

Paradoxically a plasmid borne copy of *cbbT* was previously reported to not be able to complement a *tktA* mutation (Hawkins *et al*., 2018). Consistent with the results observed in the previous study, the introduction of pJH110 into SRmD663 did not restore the ability to grow on defined media, whereas the introduction of pJH109 and pJH116, containing *tktA* and *tktB* respectively, were able restore growth (Table 8).

Since the *cbbT* open reading frame in pJH110 had not been previously verified, it was reasoned that it may not be correct. To determine this, the complete *cbbT* gene sequence was PCR amplified from the *S. meliloti* genome and cloned into pRK7813 such that the transcription of *cbbT* was driven by the p_{tac} promoter that flanked the multiple cloning site. The insert was completely sequenced, and the plasmid was designated pIW12. Introduction of pIW12 into either SRmD741 or SRmD663 restored growth on defined media (Table 8). Collectively these results show that loss of *cbbT* is sufficient to reverse suppression and that the inability of pJH110 not being able to complement a *tktA* mutation is likely due to an error attributable to a construction artifact.

Table 8 Loss of *cbbT* in a *cbbR** background reverses suppression.

Strain	Genotype	LB	LB^{Tc}	LB^{Gm}	Rib	Glc	Suc
Rm1021	WT	++	-	-	++	++	++
SRmD397	*tktA*::*Tn5*	+	-	-	-	-	-
SRmD663	*tktA*::*Tn5*, Δ*tktB*	+	-	-	-	-	-
SRmD714	*tktA*::*Tn5*, Δ*tktB*, *cbbR**	++	-	-	++	++	++
SRmD741	*tktA*::*Tn5*, Δ*tktB*, *cbbR**, *cbbT*::pKnock-Gm	+	-	++	-	-	-
SRmD742	*cbbT*::pKnock-Gm	++	-	++	++	++	++
SRmD741 (pJH109)	*tktA*::*Tn5*, Δ*tktB*, *cbbR**, *cbbT*::pKnock-Gm (*tktA*$^{+}$)	++	++	++	++	++	++
SRmD741 (pJH116)	*tktA*::*Tn5*, Δ*tktB*, *cbbR** *cbbT*::pKnock-Gm (*tktB*$^{+}$)	++	++	++	++	++	++
SRmD741 (pJH110)	*tktA*::*Tn5*, Δ*tktB*, *cbbR** *cbbT*::pKnock-Gm (*cbbT*$^{+}$)	++	++	++	-	-	-
SRmD741 (pIW12)	*tktA*::*Tn5*, Δ*tktB*, *cbbR** *cbbT*::pKnock-Gm (*cbbT*$^{+}$)	++	++	-	++	++	++
SRmD663 (pIW12)	*tktA*::*Tn5*, Δ*tktB* (*cbbT*$^{+}$)	++	++	-	++	++	++

Growth; ++, like wild-type; +, slow growth; –, no growth

Abbreviations; LB, Luria Bertani; LB^{Tc}, Luria Bertani with tetracycline; LB^{Gm}, Luria Bertani with gentamicin; Rib, VMM ribose; Glc, VMM glucose; Suc, VMM succinate.

3.5 Growth of *cbbR** strain is not dependent on *cbbL*

RuBisCO is an essential enzyme for CO_2 fixation and consists of a large and small subunit that are presumably encoded by *cbbL* and *cbbS*. It has been shown that the RuBisCO activity is present during autotrophic growth of *S. meliloti* (Manian and O'Gara, 1982), and that the presence of the *cbb* operon is essential for this growth in *S. meliloti* Rm1021 (Pickering & Oresnik, 2008). In this work qRT-PCR results show that the expression of all the *cbb* genes including *cbbL* and *cbbS* are expressed (Figure 8,Figure 9). Since this may serve as a source of carbon for the organism we wished to determine if it could influence the growth phenotypes observed on VMM.

To determine if the carbon phenotypes observed in SRmD714 were dependent on RuBisCO, an insertion mutation in *cbbL* was constructed using pKnock-Gm yielding SRmD733. Since *cbbL* is directly upstream of *cbbS* it was reasoned that this should eliminate both genes that encode subunits necessary for RuBisCO activity. This insertion was also recombined into Rm1021. The resultant strain was designated SRmD743.

The results show that when these strains were tested on VMM supplemented with either ribose, glucose, or succinate that the phenotypes were unaffected (Table 9). These data are consistent with the hypothesis that the observed phenotypes are not dependent on *cbbL*. We note that although the results are consistent with the hypothesis, that lack of RuBisCO does not affect the tested phenotypes, these data are not definitive.

3.6 The *cbbR** allele does not suppress phenotypes associated with *dctA*.

In *S. meliloti dctA* encodes the dicarboxylic acid transporter which is an essential component for the transport of dicarboxylic acids such as succinate, malate and fumarate and

Table 9 A *cbbL* mutation does not alter carbon utilization phenotypes.

Strain	**Genotype**	**LB**	**LB^{Gm}**	**Rib**	**Glc**	**Suc**
Rm1021	Wild-type	++	-	++	++	++
SRmD663	*tktA*::*Tn5, ΔtktB*	+	-	-	-	-
SRmD714	*ΔtktB, tktA*::*Tn5, cbbR**	++	-	++	++	++
SRmD733	*ΔtktB, tktA*::*Tn5, cbbR*, cbbL*::*pKnock*	++	++	++	++	++
SRmD743	*cbbL*::*pKnock*	++	++	++	++	++

Growth; ++, like wild-type; +, slow growth; –, no growth.

Abbreviations; LB, Luria Bertani; LB^{Gm}, Luria Bertani with gentamicin; Rib, VMM ribose; Glc, VMM glucose; Suc, VMM succinate.

that mutations in *dctA* result in an inability to utilize succinate as the sole carbon source because of its inability to take up the compound (Finan *et al*., 1988; Labes & Finan, 1993; Ronson *et al*., 1981; Yarosh *et al*., 1989; Yurgel & Kahn, 2005). To grow using succinate entails the use of the EMP in a gluconeogenic manner. The CBB pathway fixes CO_2 and carbon leaves the cycle as fructose-6-phosphate. It was hypothesized that a strain carrying the *cbbR** allele may be able suppress a *dctA* mutation by providing metabolites capable of bypassing the *dctA* lesion. To test this hypothesis *dctA* was targeted for insertional mutagenesis using pKnock-Gm. An internal portion of the *dctA* gene was PCR amplified, cloned into pKnock-Gm, and introduced into the wild-type Rm1021 as well as SRmD714 which carries mutations in both *tktA* as well as *tktB* and contains the *cbbR** allele. The resultant strains were designated SRmD734 and SRmD735 respectively.

The results show that as expected SRmD734 was not capable of growing on defined medium containing succinate as a sole carbon source (Table 10). However, the results also show that the *cbbR** allele was unable to suppress the growth defects caused by the *dctA* mutation suggesting that a *cbbR** mutation cannot rescue the gluconeogenic phenotype associated with a *dctA* mutation (Table 10).

3.7 The *cbbR** allele partially suppress the *tkt*-associated symbiotic phenotype.

A *tktA* mutation was originally isolated because of its inability to acidify its growth medium (Hawkins *et al*., 2017). Subsequently it was found that a mutation in *tktA* also resulted in a strain that was unable to fix nitrogen when inoculated onto alfalfa plants (Hawkins *et al*., 2018). Suppressors that were isolated from rare effective nodules were demonstrated to have a mutation in a negative regulator that led to the expression of *tktB* (Kaur 2021). When strains carrying these

Table 10 Carbon phenotypes associated with *dctA* mutants.

Strain	**Genotype**	**LB**	**LB^{Gm}**	**Suc**	**Glc**	**Rib**
Rm1021	Wild-type	++	-	++	++	++
SRmD663	*tktA*::*Tn5*, *ΔtktB*	+	-	-	-	-
SRmD714	*ΔtktB*, *tktA*::*Tn5*, *cbbR**	++	-	++	++	++
SRmD734	*dctA*::pKnock-Gm	++	++	-	++	++
SRmD735	*ΔtktB*, *tktA*::*Tn5*, *cbbR**, *dctA*::pKnock-Gm	++	++	-	++	++

Growth; ++, like wild-type; +, slow growth; –, no growth.

Abbreviations; LB, Luria Bertani; LB^{Gm}, Luria Bertani with gentamicin; Rib, VMM ribose; Glc, VMM glucose; Suc, VMM succinate.

suppressors were inoculated onto alfalfa it was found that they could restore nitrogen fixation, however, the shoot dry weight was significantly lower compared to plants that were inoculated with the wild-type (Kaur, 2021). To determine whether the *cbbR** allele could restore nitrogen fixation in plants, alfalfa plants were inoculated with *cbbR** mutant strain SRmD714.

The results show that plants inoculated with SRmD714 were healthy and green indicating nitrogen fixation whereas the parental strain, SRmD663 was unable to fix nitrogen and resulted in small chlorotic plants. Plants inoculated with SRmD714 plants were shorter in stature than the wild-type and dry weights were significantly lower compared to the wild-type. The average dry weight of SRmD714 inoculated plants was about 50 mg per plant whereas the wild-type inoculated plants had about 100 mg per plant (Figure 10). This indicates that the *cbbR** allele can suppress the fix$^-$ phenotypes associated with the *tkt* mutations. Of note, the dry matter accumulation is ~50% that of the wild-type (Figure 10). This is similar to that what was previously observed previously with a SRmD410 that carries a mutation that results in the expression of *tktB* (Figure 10) (Kaur, 2021).

3.8 The *cbbR** allele reduces nitrogen fixation in wild-type background.

The *cbbR** allele was isolated in a genetic background that had both at *tktA* as well as a *tktB* mutation (Table 5). To better characterize this allele, we wished to introduce it into a wild-type background so as to eliminate the possibility of observations that could be affected by either of these mutations. Primers were designed such that a PCR product that contained the mutation found in SRmD663 could be cloned into pJQ200SK and recombined into Rm1021. Putative recombinants were screened by nucleotide sequencing. Several putative recombinants were

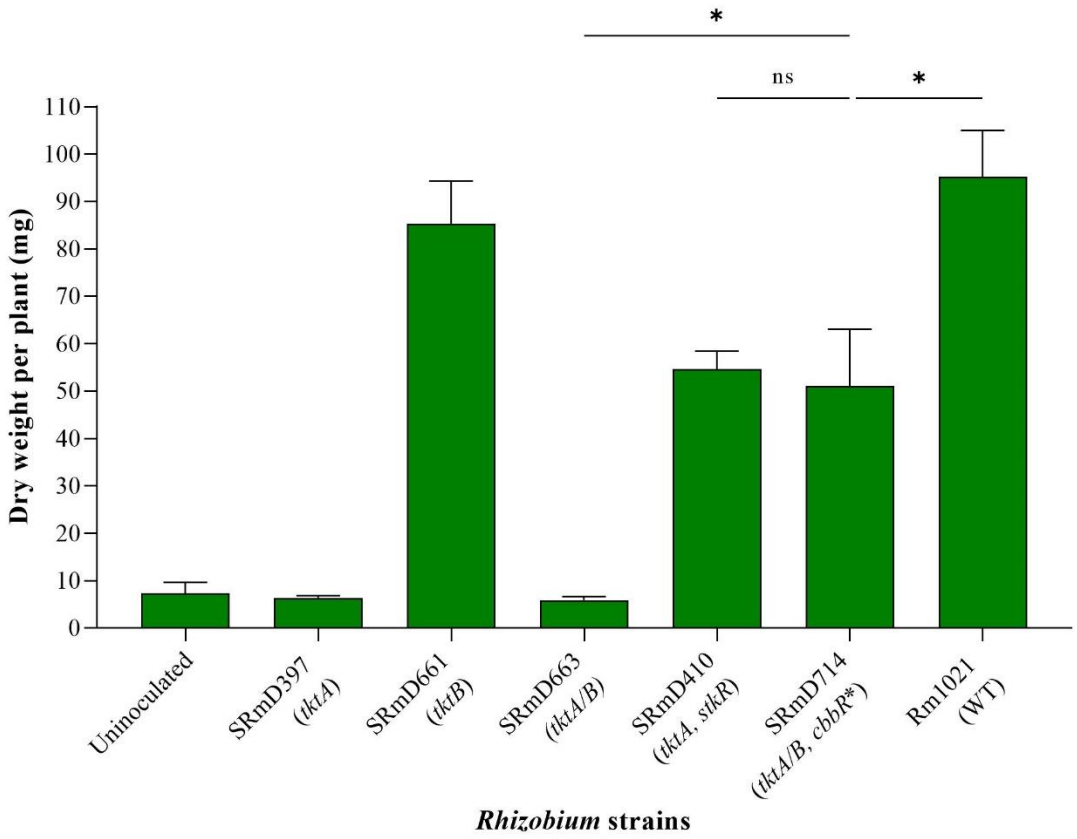

Figure 10. Shoot dry weight of alfalfa plants inoculated with *cbbR** mutant, different *tkt* mutants and wild-type strain Rm1021.

Plants were grown in sterile Leonard jar assemblies containing sand and vermiculite with nitrogen-free Jensen's medium. Shoots were harvested 28 days after inoculation, dried, and shoot dry weights were determined. The data represents the average of three independent biological replicates with 10 plants per replicate. *, represent a significant difference compared between the dry weights of plants inoculated with different strains calculated using student's *t*-test ($p < 0.05$).

found that had appropriate T to C base change. One colony was purified and designated as SRmD740. When inoculated onto plants, SRmD740 was not significantly different from SRmD714 or SRmD410. Plants inoculated with each of these strains accumulated approximately 50% dry weight of plants that were inoculated with the wild-type (Figure 11). Based on these results, it indicates that even though *tktA* is expressed normally, the *cbbR** allele still reduced nitrogen fixation by ~50%.

3.9 *cbbR** growth in sand/vermiculite.

Since SRmD740 was constitutively expressing the *cbb* operon, it was hypothesized that this strain might have a competitive advantage in a carbon limited environment. It was reasoned that soil, or sand/vermiculite, in the absence of plants might provide such an environment. Preliminary experiments were carried out using soil samples from Carmen, Kelburn, as well as the sand/vermiculite watered with Jensen's medium that is typically used for nodulation experiments.

Soil samples were sterilized and tested to ensure sterility. The soil growth experiment was conducted as described previously with modifications (DiCenzo *et al*., 2014). Growth of the Rm1021 (Wild-type), SRmD714 (*cbbR** in a *tktA/B* background), SRmD740 (*cbbR** in a wild-type background), and SRmD663 (*tktA/B*) were observed over 34 days. What was found was that none of the strains appeared to grow in soils obtained from agricultural sites, however the pilot experiment using the sand/vermiculite did grow. Based on these findings, attempts to show growth in soil were not pursued and the experiments were carried out using the sand/vermiculite regime.

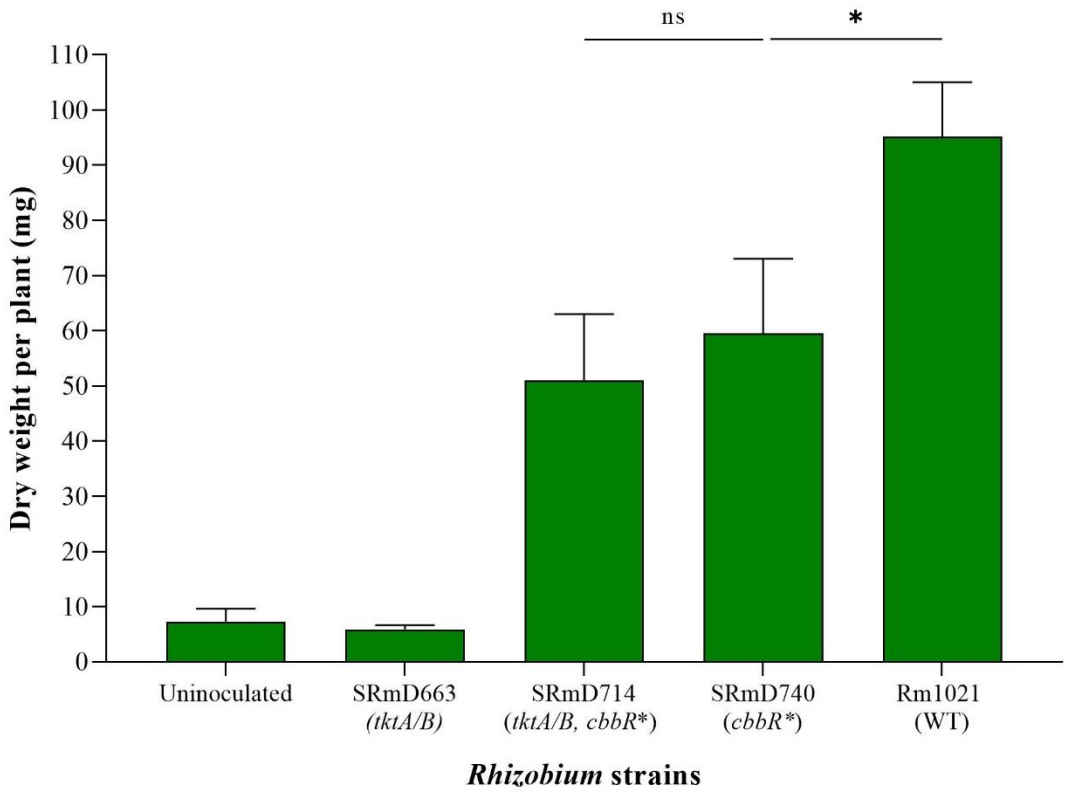

Figure 11. Shoot dry weight of alfalfa plants inoculated with SRmD740 (*cbbR**), different *tkt* mutants and wildtype strain Rm1021.

Plants were grown in sterile Leonard jar assemblies containing sand and vermiculite with nitrogen-free Jensen's medium. After growing for 28 days, the shoots were harvested and dried for 4-5 days and shoot dry weights were measured. The graph shows the average of three independent biological replicates with 10 plants per each replicate. *, represent a significant difference compared between the dry weights of plants inoculated with different strains calculated using student's *t*-test ($p<0.05$).

The growth in sand/vermiculite was carried out such that the Jensen's medium either did or did not contain nitrogen. All the tested strains grew well in the sand/vermiculite without added nitrogen except for the *tktA/B* mutant, SRmD663. The growth rates of wild-type, SRmD714 and SRmD740 were almost same, and they all reached an apparent carrying capacity approximately 12 days after inoculation. They appeared to form a stable population of about ~3×10^6 CFU/g. We were unable to detect the *tktA/B* mutant. This is consistent with the plant inoculation observations where that strain was unable to nodulate and fix nitrogen in alfalfa plants (Figure 12A). The addition of nitrogen to the medium did not have any effect on the growth of tested strains. The growth rates and the time at which they reached saturation were similar to that of the ones grown in nitrogen free medium (Figure 12B). The *tktA/B* mutant was not able to grow even with the nitrogen supplementation.

To determine whether there is any significant difference in the total bacteria present, the bacterial population at saturation was compared between all strains, as well as between the N supplemented and N free conditions. The results show that there was no significant difference in the population of wild-type, SRmD740 and SRmD714 (Figure 13). The only significant difference was found for SRmD663, which did not grow in either N-free or N-supplemented medium. Together, these data suggest that the SRmD714 ad SRmD740 both have a similar ability as the wild-type to survive, grow and establish stable populations in soil under ideal conditions.

3.10 A strain carrying *cbbR** is less competitive than the wild-type for nodule occupancy.

Strains carrying the *cbbR* allele fixed less nitrogen than the wild-type, yet they grew in sand/vermiculite at the same rate as the wild-type (Figure 10,Figure 12). We wished to determine

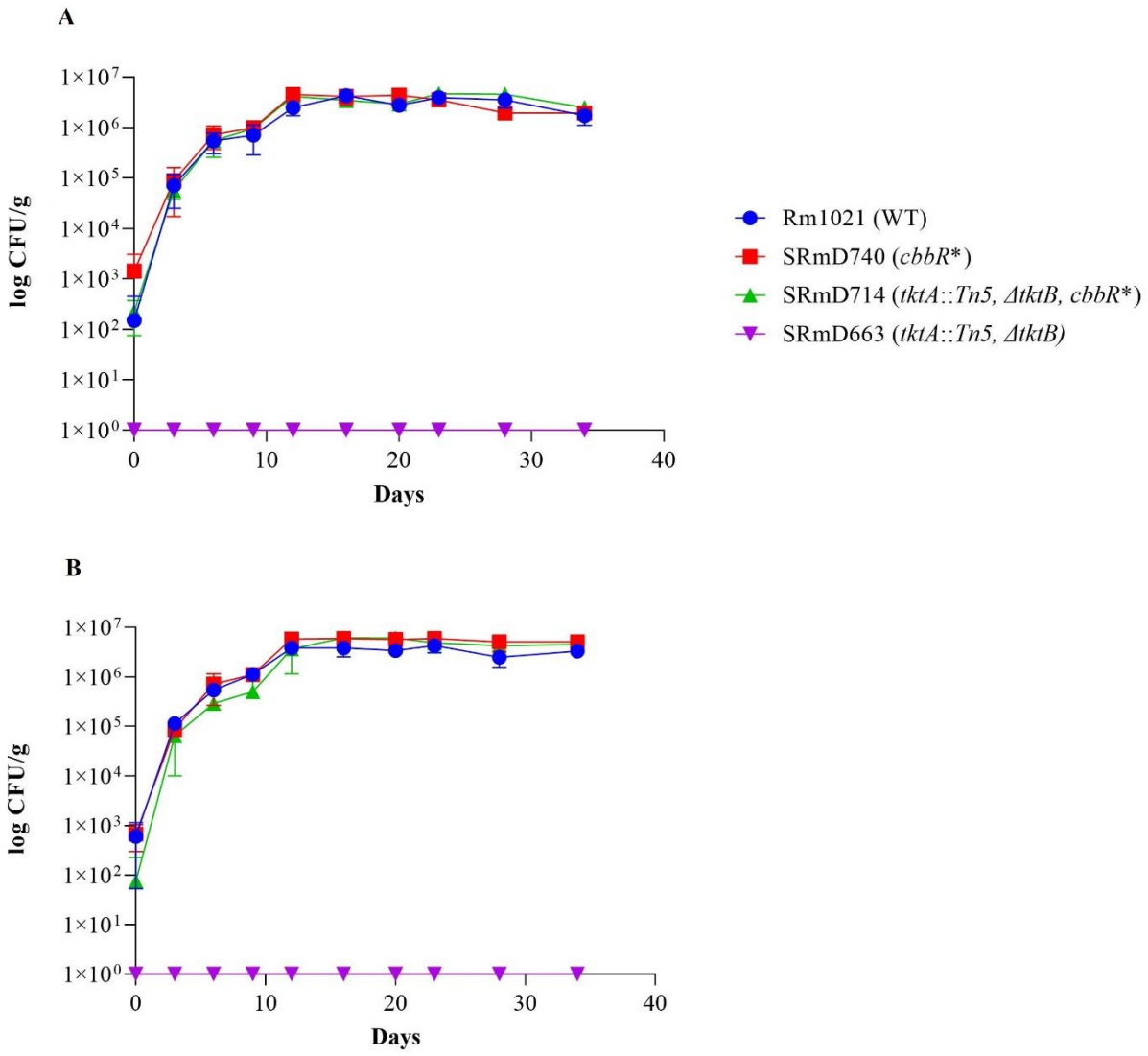

Figure 12. Growth of *S. meliloti* in sand/vermiculite.
Strains were grown overnight in LB, diluted, and inoculated into modified Magenta jars as described in Materials and Methods. Strains and relevant genotypes are as indicated in the legend: Rm1021, ; SRmD740, ; SRmD714, ; SRmD663, . **Panel A**, growth in sand/vermiculite soaked with nitrogen free Jesnson's medium. **Panel B**, growth in Jesnson's medium supplemented with nitrogen. Growth was measured over a period of 34 days and the data represent the average cfu/g of 4 biological replicates. Where not visible, error bars are smaller than the symbol.

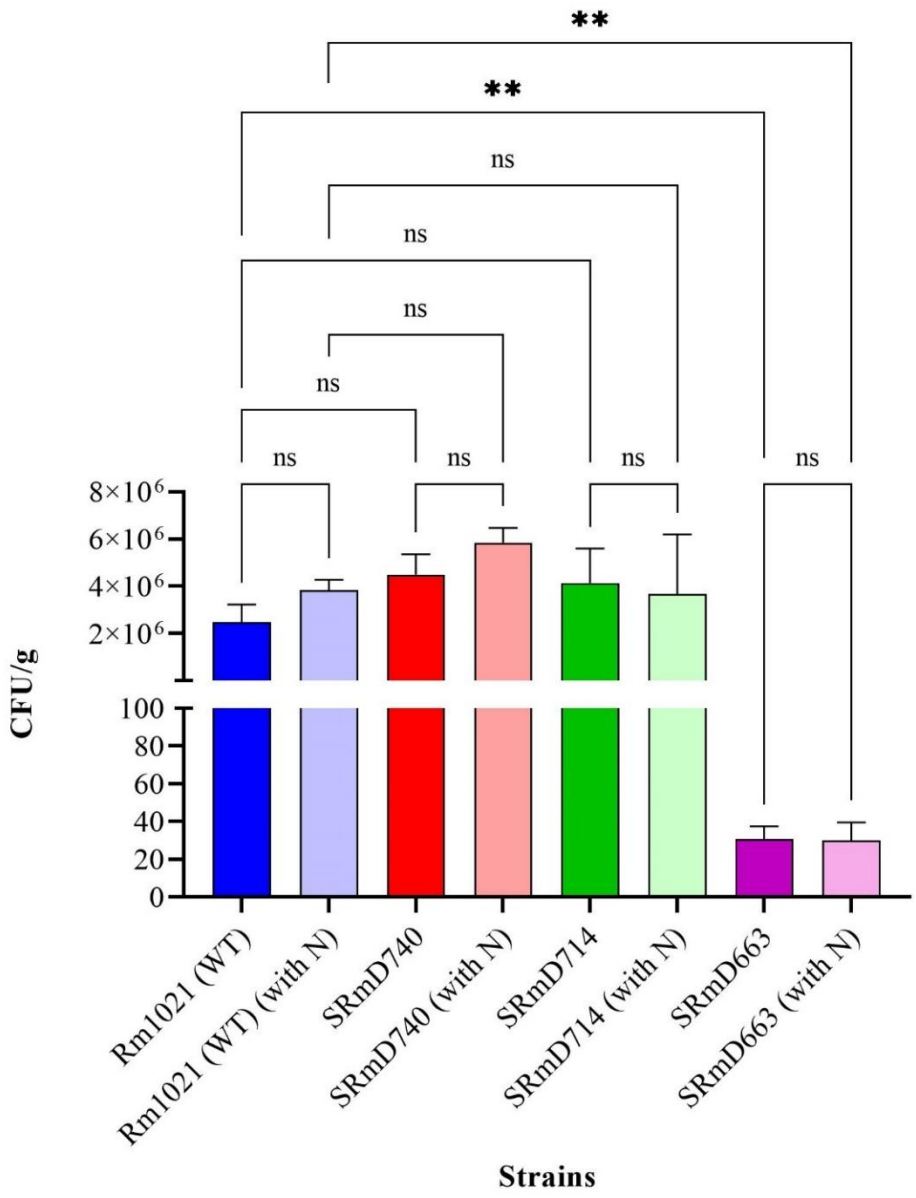

Figure 13. Comparison of total bacterial population in sand: vermiculite with or without nitrogen at 12 days post inoculation.

The data represent the average (cfu/g) ± standard error of 4 biological replicates. The values for SRmD663 were estimated based on the initial inoculum that was used. ** represent the significant difference between the tested strains calculated using an ANOVA ($p<0.05$)

if this allele would alter how quickly a strain could proceed through the infection process to produce nodules as well as if it affected a strain's ability to compete for nodule occupancy. To assay for nodule occupancy the presence of a marker that could be screened for is necessary. Since both SRmD714 and SRmD740 showed identical growth characteristics and symbiotic nitrogen fixation phenotypes (Figure 11), SRmD714 was chosen because it is neomycin resistant (Table 1). Nodulation kinetics were assessed by growing alfalfa seedlings inoculated with *S. meliloti* on Jensen's agar slants and nodule formation was observed and recorded every other day for a period of 21 days to determine the percentage plants nodulated and the average nodules per plant.

The results show that the nodulation kinetics of both strains were identical (Figure 14A) with it taking approximately 7 days to attain 50% nodulation and reaching 100% nodulation by day 19. Similarly, the average nodules per plant was not significantly different between the two strains with wild-type forming about 6.6 nodules per plant and those inoculated with SRmD714 forming about 5.7 nodules per plant at the end of 21 days (student t-test ($p > 0.05$)) (Figure 14B). These results suggest that rate of infection and nodule formation between the two strains is not different.

To assay competition for nodule occupancy SRmD714 was competed against the wild-type at three different ratios. The inocula were determined to contain 45.5%, 66.6% and 87.6% of SRmD714. At harvest, the plants were all healthy and green with no visible differences in growth between the different inoculum mixtures. The nodules were harvested after 28 days of growth to determine the final percentage of infection by the mutant strain. Despite the inoculum ratios, most of the nodules were infected with wild-type. The final percentages of nodules infected with *cbbR** mutant was significantly lower ($p<0.01$) compared to the initial inoculum

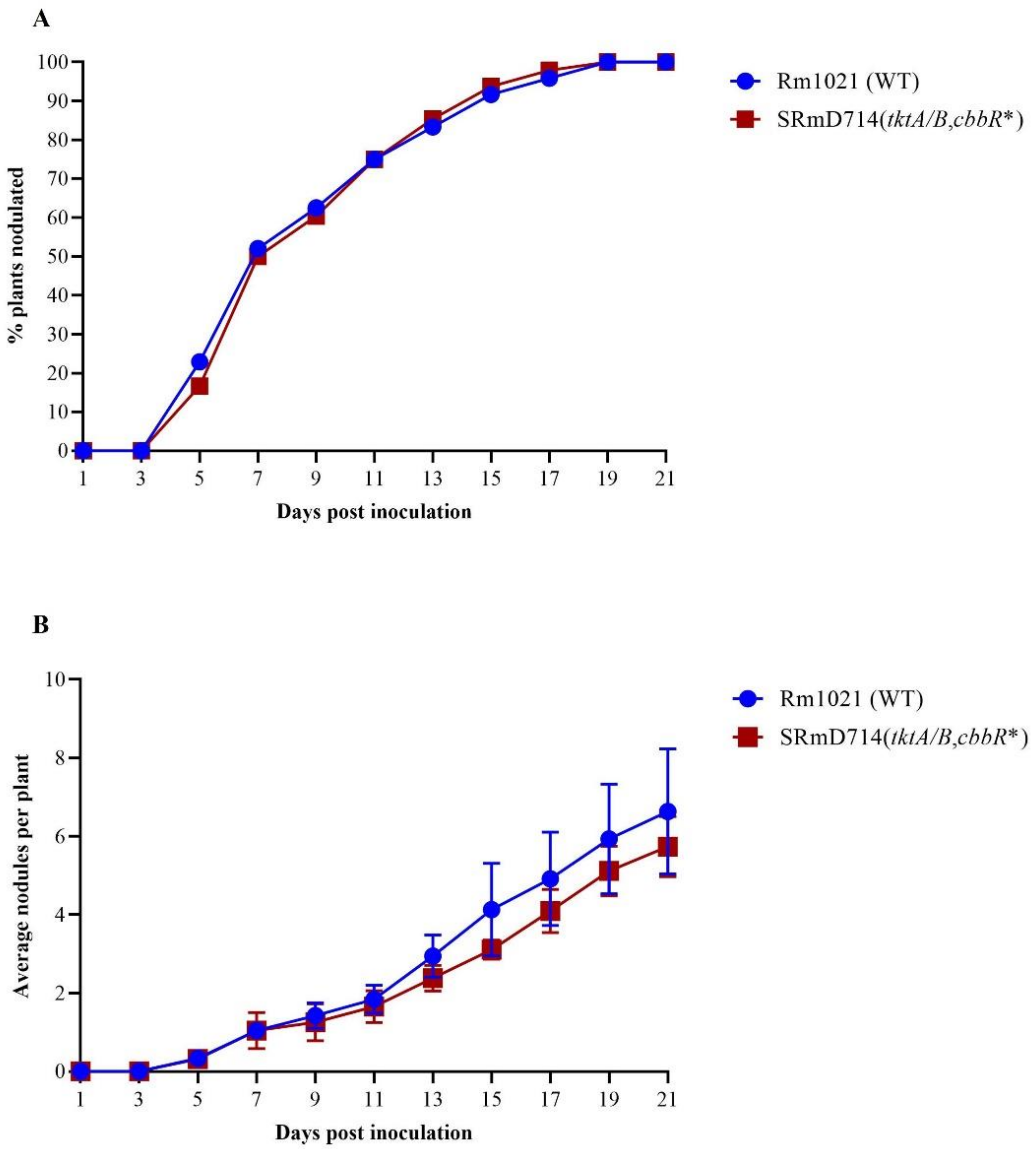

Figure 14. Nodulation kinetics of the Rm1021 (wild-type) and SRmD714 (*tktA/B, cbbR**). Nodulation kinetics assays were carried out as described in Materials and Methods. The data is presented as the percentage of plants nodulated (**A**), and average nodules per plant (**B**). Data represents the average ± standard error of five independent biological replicates which consisted of at least 10 plants per each replicate.

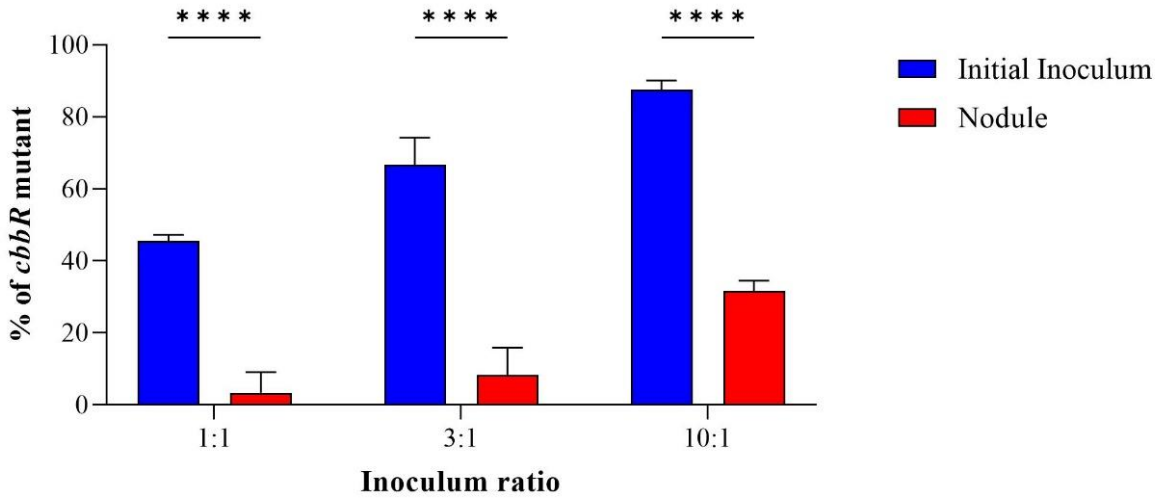

Figure 15. Competition for nodule occupancy of Rm1021 (wild-type) and SRmD714 (*tktA/B, cbbR**).

Rm1021 and SRmD714 were inoculated onto plants at three initial inoculum ratios as indicated. Blue bars represent the percentage of the mutant in the initial inoculum. Red bars represent the percentage of nodules occupied by the mutant strain as determined by isolating the bacteria from the nodules. The data represent the average ± standard error of 3 independent biological replicates at each ratio.

**** Represents significant difference compared between initial inoculum and nodule isolate calculated using student *t*-test ($p < 0.0001$).

percentage (Figure 15). With respect to competition for nodule occupancy these data clearly show that at every ratio of inocula, SRmD714 was significantly less competitive for nodule occupancy (Figure 15).

Chapter 4: Discussion

In this study a mutation that suppressed the phenotypes of a strain carrying *tktA* and *tktB* mutations was isolated. The point mutation was in the gene *cbbR* which is annotated as a Lys-R type transcriptional regulator. The gene *cbbR* is found on pSymB adjacent to an operon containing genes that has been annotated as necessary for the CBB cycle in *S. meliloti* and have been shown to be involved in autotrophic growth (Pickering & Oresnik, 2008). *cbbR* is recognized as the key regulator of the *cbb* operon in the majority of CO_2 fixing organisms (Dangel *et al*., 2005; Dangel & Tabita, 2015a; Van Keulen *et al*., 2003). A strain containing the suppressor was shown to have the genes *cbbFPTALSX* as well as *ppe* and *SMb20194* all up regulated and appear to form an operon (Figure 8). With the assumption that a strain carrying the *cbbR** allele was fixing CO_2, it was shown that constitutive CO_2 fixation did not confer an advantage in a carbon limited environment (Table 9), and that it compromised nitrogen fixation (Figure 11), as well as its ability to compete for nodule occupancy (Figure 15).

The genes that encode for CBB cycle enzymes can be either localized in a single cluster or multiple clusters within the genomes of CO_2 fixing organisms (Dangel & Tabita, 2015a; Esparza *et al*., 2010; Gibson & Tabita, 1993). In *S. meliloti*, the *cbb* genes are clustered together on pSymB which is considered as one of the most conserved *cbb* gene arrangements (Finan *et al*., 2001; Kusian & Bowien, 1997). The CBB pathway consists of eleven reactions (Figure 3). Eight of the genes in this operon are recognized to encode enzymes required to carry out the reactions in CBB cycle (Figure 4). The terminal gene in the operon is the hypothetical gene *SMb20194*. Nothing is known about its role within the cell. Based on rudimentary searches, the encoded protein appears to contain a thioredoxin-like fold, and similarities to an alkyl hydroperoxide reductase.

The determinants encoding the three missing activities in CBB cycle are *pgk*, *gap* and *tpi*. These are all found on the chromosome of Rm1021. Of note, *pgk* and *gap* are found in an operon and have been shown to be necessary for gluconeogenesis (Finan *et al*., 1988), whereas *tpi* is found as a single gene on the chromosome and is not required for gluconeogenesis but has been shown to be necessary for glycerol utilization as well as for autotrophic growth (Pickering & Oresnik, 2008; Poysti & Oresnik, 2007). It has been shown that the CbbR has the ability to bind and regulate the expression of *cbb* operons despite their physical location (Esparza *et al*., 2010; Gibson & Tabita, 1993). It is of note that the *cbbR* encoded transcriptional regulator in the chemoautotroph *Xanthobacter flavus* has been shown the ability to regulate not only the *cbb* operon, but also *gap-pgk* operon (Van Keulen *et al*., 2003).

Mutants of *Rhodobacter* sp. with amino acid substitutions in the effector domain of CbbR have shown to result in an increased expression (Dangel *et al*., 2005). These mutations change the conformation of the effector binding site and thereby alter the interaction with RNA polymerase (Dangel & Tabita, 2015a; Van Keulen *et al*., 2003). However, these amino acid substitutions seem to be specific for the different CbbRs (Dangel & Tabita, 2015b). The point mutation identified in the present study is also localized in the effector domain of the CbbR protein based on the predicted 3D structure (Figure 7). Presumably this must alter the effector binding or the interaction of CbbR* with RNA polymerase resulting in expression of downstream genes in the *cbb* operon.

Gene regulation by CbbR normally depends on an effector molecule binding to CbbR (Dangel & Tabita, 2015b). Positive effector binding can either initiate or increase the transcription by altering the interaction with RNA polymerase (Schell, 1993). The effector molecules for CbbR are often the intermediates of the CBB pathway, and are usually organism

specific (Dangel & Tabita, 2015b). RuBP is recognized as one of the main positive effectors in many organisms especially in *Rhodobacter* sp. (Dangel *et al*., 2005; Dubbs *et al*., 2004). PEP has been reported as a main negative effector in *Ralstonia eutropha* where the interaction of PEP with CbbR effector domain enhance the binding of the protein to the *cbb* promoter, but severely inhibit the transcription (Grzeszik *et al*., 2000). The possible effectors for *S. meliloti* CbbR have not been reported yet.

Transketolase is a key enzyme in the PPP as well as the CBB pathway. The *S. meliloti* genome contains five genes with the potential of encoding transketolase activity (Capela *et al*., 2001). These include *tktA* (*SMc03978*), *tktB* (*SMc02342*), *cbbT* (*SMb20200*), as well as a putative transketolase containing an alpha and beta chain encoded by *SMc00270* and *SMc00269* respectively. Mutations have been isolated or constructed in *tktA* as well as *tktB* (Hawkins *et al*., 2018; Kaur, 2021). It has been shown that *tktA* encodes the enzyme that is normally used in Rm1021 (Hawkins *et al*., 2018). The gene *tktB* is negatively regulated by *SMc02340* and can suppress a *tktA* mutant phenotype if *SMc02340* is inactivated (Kaur, 2021), and CbbT has been shown to be functional during autotrophic growth (Pickering and Oresnik, 2008). The biochemical identity of the enzyme encoded by *SMc00270* and *SMc00269* is less clear. These two genes do encode an enzyme that contains transketolase domains but have far lower overall identities to CbbT. When compared to CbbT, TktA (S=1890, E=0) and TktB (S=1440, E= 1.75^{-167}) are strongly supported as being transketolases, whereas SMc00270 and SMc00269 scores are not as strongly supported; S=276, E=2.15^{-26} and S=162, E=1.36^{-12} respectively (*Toulouse - Sinorhizobium meliloti 1021*, n.d.). These two genes also encode putative dehydrogenase domains suggesting that these proteins may carry out a transketolase-type reaction, but their

primary activity is not likely to use xylulose-5-phosphate as a keto-donor and either ribose 5-or erythrose 4-phosphate as keto-acceptor substrates.

The role of the *cbbT* encoded enzyme in the PP pathway has been inferred but not directly shown (Pickering & Oresnik, 2008). Despite the sequence similarity between CbbT and TktA, it has been reported that *cbbT* carried on a multicopy plasmid could not complement a *tktA* mutation (Hawkins *et al*., 2018). Although this was reported, it was not pursued (Hawkins *et al*., 2018). In the present study, *cbbT* was independently PCR amplified and expressed using pRK7813. The results clearly show that it can restore the growth of *tktA/B* mutant strain in minimal medium, indicating that *cbbT* encodes for an active transketolase. A possible reason for this discrepancy may be that Hawkins *et al*., (2018a) had constructed the plasmid carrying *cbbT* using the *S. meliloti* ORFeome (Hawkins *et al*., 2018). When the ORFeome was constructed each of the open reading frames in the library were not completely sequenced (Schroeder *et al*., 2005). The construct from present work is completely sequenced. The most likely explanation is the previous plasmid carrying *cbbT* contains sequence errors.

When the *cbbR** allele was introduced into the wild-type (SRmD740) and used to inoculate alfalfa plants it resulted in reduced plant biomass indicating a reduced nitrogen fixing efficiency. This suggests that overexpression of the genes encoding the CBB pathway affects nitrogen fixation. To fix 6 CO_2 to yield 1 F6P using the CBB pathway takes 12 NADPH and 18 ATP. This continual demand for energy and reductant may directly affect the ATP and reductant necessary for the nitrogenase-thus leading to a lower nitrogen fixation rates. Alternately it may be that an increase of the enzymes necessary for the CBB pathway may be acting like a sink for carbon skeletons being provided by the plant to the bacteroid away from the TCA cycle, thus

reducing the potential to generate the reductant and ATP necessary for nitrogenase. Both of these scenarios are not mutually exclusive and can be used to rationalize lower nitrogen fixation rates.

A precedent for CO_2 fixation affecting symbiotic nitrogen fixation exists in the literature (Giraud & Fleischman, 2004; Gourion *et al*., 2011). *Bradyrhizobium* sp. strain ORS278 contains a full CBB pathway with an operon that is syntenic with *S. meliloti* (Gourion *et al*., 2011). When a strain that carried a mutation in *cbbL1,* which encodes for a large subunit of type 1 RuBisCO, was tested symbiotically it resulted in reduced nitrogenase activity and plants with foliage chlorosis or stem thinning when compared to the wild-type (Gourion *et al*., 2011). This suggested that CBB plays a role in the symbiosis and a model was proposed where the CBB pathway essentially was used as an electron sink for the developing nodules prior to full expression of nitrogenase (Gourion *et al*., 2011). *S. meliloti* CbbL shows a 76% amino acid similarity to the *Bradyrhizobium* sp. OR278 CbbL1 (Gourion *et al*., 2011). Although the affect observed is contrary to what has been observed in this work (Figure 10), it does provide a precedent to demonstrate that altering the supply of reductants can influence nitrogen fixation. We do note that a similar reduction in nitrogen fixation was also observed with just the overexpression of *tktB* (Kaur, 2021), suggesting that mechanism of reduction may be due to several factors.

Strains that carry the *cbbR** allele are uncompetitive for nodule occupancy (Figure 15). Although they were uncompetitive, they did not show any growth defects in sand/vermiculite pots supplied with plant growth solution (Figure 12), and they had nodulation kinetics identical to that of the wild-type (Figure 14). Most competition for nodule occupancy phenotypes are due to a loss of function mutations (Ding *et al*., 2012; Hawkins *et al*., 2018). The “nodule sanctioning hypothesis”, suggests that the host plant favours rhizobia that are efficient at nitrogen fixation

and that explains competition for nodule occupancy (Denison, 2000; Kiers *et al*., 2003). It is clear that in the field legumes can be infected with multiple *Rhizobium* strains (Bala & Giller, 2001). However, the efficiency of nitrogen fixation by the different rhizobia can differ. In order to gain the maximum fixed nitrogen output, plants have been shown to favour the highly efficient nitrogen fixers by allocating resources such as carbon toward more efficient nodules (Westhoek *et al*., 2021). Models have been developed to predict the ability of legumes to sanction the rhizobia based on their efficiency in nitrogen fixation. Essentially, when rhizobia that are highly effective at nitrogen fixation are present, they had an advantage over rhizobia that had intermediate nitrogen fixation rates and that those rhizobia that were unable to fix nitrogen were eliminated (Kiers *et al*., 2006; Westhoek *et al*., 2021).

This hypothesis can partially explain the observations that have been made (Figure 15). A problem that arises is that nodules that are sanctioned often display as being younger and less developed. Differences in nodule sizes were not observed or quantitated in this work. If sanctioning was occurring prior to the full development of the nodule, it would imply that it was occurring either early stages of infection such as at colonization of the curled root hair or during infection thread development.

Chapter 5: Future Directions

The work in this thesis describes the isolation and the characterization of a novel allele that results in the overexpression of the CBB pathway. The work can proceed in a number of directions and can be used as a platform to address developmental as well as biochemical questions.

Does CbbR affect the expression of genes other than the cbb operon? The work presented used the assumption that only the *cbb* operon was affected by the *cbbR** allele. There is clear evidence in the literature that CbbR can affect a number of different operons in other organisms (Van Keulen *et al*., 2003). We note that three of the genes encoding enzymes necessary for growing autotrophically are not present in the *cbb* operon. RNAseq experiments would be insightful to determine if these genes, or others are either directly or indirectly affected by constitutive CbbR activity. Additionally, gel shift mobility assays and DNase I footprint analysis can be carried out to better define how CbbR interacts with its cognate promoters as well as what metabolites CbbR may be responding to.

Does the cbbR lead to fixation of CO_2?* A basic assumption in this thesis was that expression of the *cbb* operon would result in being able to fix CO_2. From the literature we know that *S. meliloti* can grow autotrophically, yet this has not been demonstrated for strains carrying the *cbbR** allele. This can be addressed directly by carrying out $^{14}CO_2$ assays.

What effect does the cbbR allele have on metabolism?* Carrying out a metabolomic experiment to quantify central carbon metabolite pools could provide evidence to how carbon might be portioned in strains carrying this allele. This can possibly provide indirect evidence on the role it may be playing symbiotically.

At what stage of infection do strains carrying cbbR alleles become uncompetitive?* Strain carrying the *cbbR** allele is highly uncompetitive for nodule occupancy even though it

does not exhibit delays during early nodulation stages. This could be a result of possible nodule sanctioning by alfalfa plants. A chromogenic approach such as one used by (Mendoza-Suárez *et al*., 2020) can be used to determine the distribution of the different strains within the root nodules of plants co-inoculated with wild-type and strain carrying the *cbbR**allele. Briefly, this would involve staining the entire nodulated root and being able to determine the occupancy of every nodule based on chromogenic activity of the inoculated strain. This could be used to determine directly if nodule sanctioning is occurring. Based on the ability of these mutant strains to survive in soil and nodulate the plants at a similar rate to the wild-type, it is possible that the competition is not happening at the earlier stages of infection. The use of fluorescence markers to directly assay rates of infection by mixed inocula could provide this type of information.

Literature Cited

Alexeyev, M. F. (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**(5), 5–7.

Arias, A., Cervenansky, C., Gardiol, A., & Martinez-Drets, G. (1979). Phosphoglucose isomerase mutant of *Rhizobium meliloti*. J. Bacteriol., **137**(1), 409–414. DOI:10.1128/jb.137.1.409-414.1979

Bala, A., & Giller, K. E. (2001). Symbiotic specificity of tropical tree rhizobia for host legumes. New Phytol., **149**(3), 495–507. DOI:10.1046/j.1469-8137.2001.00059.x

Barnett, M. J., Fisher, R. F., Jones, T., Komp, C., Abola, A. P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Huizar, L., Hyman, R. W., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Palm, C., … Long, S. R. (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. U.S.A., **98**(17), 9883–9888. DOI:10.1073/pnas.161294798

Burghardt, L. T., & diCenzo, G. (2023). The evolutionary ecology of rhizobia: Multiple facets of competition before, during, and after symbiosis with legumes. Curr. Opin. Microbiol., **72**, 102281. DOI:10.1016/j.mib.2023.102281

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. PLOS ONE, **12**(10), e0185056. DOI:10.1371/journal.pone.0185056

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dréano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., … Galibert, F. (2001). Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. U.S.A., **98**(17). DOI:10.1073/pnas.161294398

Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J. A., Morris, R. J., Ané, J. M., & Oldroyd, G. E. D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. Proc. Natl. Acad. Sci. U.S.A., **108**(34), 14348–14353. DOI:10.1073/pnas.1107912108

Cervenansky, C., & Arias, A. (1984). Glucose-6-phosphate dehydrogenase deficiency in pleiotropic carbohydrate-negative mutant strains of *Rhizobium meliloti*. J. Bacteriol., **160**(3), 1027–1030. DOI:10.1128/jb.160.3.1027-1030.1984

Costelloe, S. J., Ward, J. M., & Dalby, P. A. (2008). Evolutionary analysis of the TPP-dependent enzyme family. J. Mol. Evol., **66**(1), 36–49. DOI:10.1007/s00239-007-9056-2

Dangel, A. W., Gibson, J. L., Janssen, A. P., & Tabita, F. R. (2005). Residues that influence *in vivo* and *in vitro* CbbR function in *Rhodobacter sphaeroides* and identification of a specific region critical for co-inducer recognition: CbbR in *R. sphaeroides*. Mol. Microbiol., **57**(5), 1397–1414. DOI:10.1111/j.1365-2958.2005.04783.x

Dangel, A. W., & Tabita, F. R. (2015a). Amino acid substitutions in the transcriptional regulator CbbR lead to constitutively active CbbR proteins that elevate expression of the *cbb* CO_2 fixation operons in *Ralstonia eutropha* (Cupriavidus necator) and identify regions of CbbR necessary for gene activation. Microbiology, **161**(9), 1816–1829. DOI:10.1099/mic.0.000131

Dangel, A. W., & Tabita, F. R. (2015b). CbbR, the master regulator for microbial carbon dioxide fixation. J. Bacteriol., **197**(22), 3488–3498. DOI:10.1128/JB.00442-15

Denison, R. F. (2000). Legume sanctions and the evolution of symbiotic cooperation by rhizobia. Am. Nat., **156**(6), 567–576.

DiCenzo, G. C., Checcucci, A., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., Finan, T. M., Galardini, M., & Fondi, M. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. Nat. Commun., **7**. DOI:10.1038/ncomms12219

DiCenzo, G. C., & Finan, T. M. (2017). The divided bacterial genome: Structure, function, and evolution. Microbiol. Mol. Biol. Rev., **81**(3), e00019-17. DOI:10.1128/MMBR.00019-17

DiCenzo, G. C., MacLean, A. M., Milunovic, B., Brian Golding, G., & Finan, T. M. (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLoS Genetics, **10**(10), e1004742.

DiCenzo, G. C., Milunovic, B., Cheng, J., & Finan, T. M. (2013). The tRNAarg gene 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**(2), 202–212. DOI:10.1128/JB.01758-12

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**(5), 1369–1378. DOI:10.1099/mic.0.057281-0.

Dubbs, P., Dubbs, J. M., & Tabita, F. R. (2004). Effector-mediated interaction of CbbR $_{I}$ and CbbR $_{II}$ regulators with target sequences in *Rhodobacter capsulatus*. J. Bacteriol., **186**(23), 8026–8035. DOI:10.1128/JB.186.23.8026-8035.2004.

Duggleby, R. G. (2006). Domain relationships in thiamine diphosphate-dependent enzymes. Acc. Chem. Res., **39**(8), 550–557. DOI:10.1021/ar068022z.

Duncan, M. J. (1981). Properties of Tn5-induced carbohydrate mutants in *Rhizobium meliloti*. J. Gen. Microbiol., **122**(1), 61–67. DOI:10.1099/00221287-122-1-61.

Dunn, M. F. (1998). Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. FEMS Microbiol. Rev., **22**(2), 105–123. DOI:10.1111/j.1574-6976.1998.tb00363.x.

Esparza, M., Cárdenas, J. P., Bowien, B., Jedlicki, E., & Holmes, D. S. (2010). Genes and pathways for CO_2 fixation in the obligate, chemolithoautotrophic acidophile, *Acidithiobacillus ferrooxidans*, Carbon fixation in *A. ferrooxidans*. BMC Microbiol., **10**(1), 229. DOI:10.1186/1471-2180-10-229.

Falcone, D. L., & Tabita, F. R. (1993). Complementation analysis and regulation of CO_2 fixation gene expression in a ribulose 1,5-bisphosphate carboxylase-oxygenase deletion strain of *Rhodospirillum rubrum*. J. Bacteriol., **175**(16), 5066–5077. DOI:10.1128/jb.175.16.5066-5077.1993.

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**(4), 869–877. DOI:10.1016/0092-8674(85)90346-0.

Finan, T. M., Oresnik, I. J., & Bottacin, A. (1988). Mutants of *Rhizobium meliloti* defective in succinate metabolism. J. Bacteriol., **170**(8), 3396–3403. DOI:0021-9193/88/083396-08$02.00/0.

Finan, T. M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorhölter, F. J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., & Pühler, A. (2001). The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2-fixing endosymbiont *Sinorhizobium meliloti*. Proc. Natl. Acad. Sci. U.S.A., **98**(17), 9889–9894. DOI:10.1073/pnas.161294698.

Fuhrer, T., Fischer, E., & Sauer, U. (2005). Experimental identification and quantification of glucose metabolism in seven bacterial species. J. Bacteriol., **187**(5), 1581–1590. DOI:10.1128/JB.187.5.1581-1590.2005.

Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol. Mol. Biol. Rev., **68**(2), 280–300. DOI:10.1128/mmbr.68.2.280-300.2004.

Galibert, F., Finan, T. M., Long, S. R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnet, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., … Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science, **293**(5530), 668–672. DOI:10.1126/science.1060966.

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**(5878), 889–892. DOI:10.1126/science.1136674.

Geddes, B. A., González, J. E., & Oresnik, I. J. (2014). Exopolysaccharide production in response to medium acidification is correlated with an increase in competition for nodule occupancy. Mol. Plant Microbe Interact., **27**(12), 1307–1317. DOI:10.1094/MPMI-06-14-0168-R.

Geddes, B. A., Kearsley, J., Morton, R., diCenzo, G. C., & Finan, T. M. (2020). The genomes of rhizobia. In *Advances in Botanical Research* (Vol. 94, pp. 213–249). Elsevier. DOI:10.1016/bs.abr.2019.09.014.

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**(8), 2180–2191. DOI:10.1099/mic.0.057877-0.

Geddes, B. A., & Oresnik, I. J. (2012b). Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. J. Bacteriol., **194**(18), 5044–5053. DOI:10.1128/JB.00982-12.

Geddes, B. A., & Oresnik, I. J. (2014). Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. Can. J. Microbiol., **60**(8), 491–507. DOI:10.1139/cjm-2014-0306.

Geddes, B. A., & Oresnik, I. J. (2016). The mechanism of symbiotic nitrogen fixation. In *The mechanistic benefits of microbial symbionts* (Vol. 2, pp. 69–97). Advances in Environmental Microbiology.

Geurts, R., Fedorova, E., & Bisseling, T. (2005). Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. Curr. Opin. Plant Biol., **8**(4), 346–352. DOI:10.1016/j.pbi.2005.05.013.

Gibson, J. L., & Tabita, F. R. (1993). Nucleotide sequence and functional analysis of *cbbR*, a positive regulator of the calvin cycle operons of *Rhodobacter sphaeroides*. J. Bacteriol., **175**(18), 5778–5784. DOI:10.1128/jb.175.18.5778-5784.1993.

Giraud, E., & Fleischman, D. (2004). Nitrogen-fixing symbiosis between photosynthetic bacteria and legumes. Photosynth. Res., **82**(2), 115–130. DOI:10.1007/s11120-004-1768-1.

Glazebrook, J., & Walker, G. C. (1991). Genetic techniques in *Rhizobium meliloti*. Meth. Enzymol., **204**(C), 398–418. DOI:10.1016/0076-6879(91)04021-F.

Gourion, B., Delmotte, N., Bonaldi, K., Nouwen, N., Vorholt, J. A., & Giraud, E. (2011). Bacterial RuBisCO is required for efficient *Bradyrhizobium/Aeschynomene* symbiosis. PLoS ONE, **6**(7), e21900. DOI:10.1371/journal.pone.0021900.

Grzeszik, C., Jeffke, T., Schäferjohann, J., Kusian, B., & Bowien, B. (2000). Phosphoenolpyruvate is a signal metabolite in transcriptional control of the *cbb* CO_2 fixation operons in *Ralstoniaeutropha*. J. Mol. Microbiol. Biotechnol., **2**(3), 311–320.

Gust, A. A., Willmann, R., Desaki, Y., Grabherr, H. M., & Nürnberger, T. (2012). Plant LysM proteins: Modules mediating symbiosis and immunity. Trends Plant Sci., **17**(8), 495–502. DOI:10.1016/j.tplants.2012.04.003.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol., **166**(4), 557–580. DOI:10.1016/S0022-2836(83)80284-8.

Hanus, F. J., Maier, R. J., & Evans, H. J. (1979). Autotrophic growth of H_2-uptake-positive strains of *Rhizobium japonicum* in an atmosphere supplied with hydrogen gas. Proc. Natl. Acad. Sci. U.S.A., **76**(4), 1788–1792. DOI:10.1073/pnas.76.4.1788.

Hawkins, J. P., Geddes, B. A., & Oresnik, I. J. (2017). Common dyes used to determine bacterial polysaccharides on agar are affected by medium acidification. Can. J. Microbiol., **63**(6), 559–562. DOI:10.1139/cjm-2016-0743.

Hawkins, J. P., Ordonez, P. A., & Oresnik, I. J. (2018). Characterization of mutations that affect the nonoxidative pentose phosphate pathway in *Sinorhizobium meliloti*. J. Bacteriol., **200**(2), 1–18. DOI:10.1128/JB.00436-17.

Hawkins, J. P., Ordonez, P. A., & Oresnik, I. J. (2022). Complete genome sequences of *Rhizobium gallicum* M101 and two potential new *Rhizobium* species isolated from soils in central Canada. Microbiol. Resour. Announc., **2**(8), 9–12. DOI:10.1128/mra.00216-22.

Herridge, D. F., Peoples, M. B., & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil, **311**(1–2), 1–18. DOI:10.1007/s11104-008-9668-3.

House, B. L., Mortimer, M. W., & Kahn, M. L. (2004). New recombination methods for *Sinorhizobium meliloti* genetics. Appl. Environ. Microbiol., **70**(5), 2806–2815. DOI:10.1128/AEM.70.5.2806-2815.2004.

Iida, A., Teshiba,' And, S., & Mizobuchi2t, K. (1993). Identification and characterization of the *tktB* gene encoding a second transketolase in *Escherichia coli* K-12. J. Bacteriol., **175**(17), 5375–5383.

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., Adham, 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**(7), 979–987. DOI:10.1094/MPMI-21-7-0979.

Jones, J. D. G., & Gutterson, N. (1987). An efficient mobilizable cosmid vector, pRK7813, and its use in a rapid method for marker exchange in *Pseudomonas fluorescens* strain HV37a. Gene, **61**(3), 299–306. DOI:10.1016/0378-1119(87)90193-4.

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.

Kaur, S. (2021). Identification and characterization of Pentose Phosphate and Embden Meyerhof Parnas pathway mutants in *Sinorhizobium meliloti*. University of Manitoba.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc., **10**(6), 845–858. DOI:10.1038/nprot.2015.053.

Kiers, E. T., Rousseau, R. A., & Denison, R. F. (2006). Measured sanctions: Legume hosts detect quantitative variation in *Rhizobium* cooperation and punish accordingly. Evol. Ecol. Res., **8**, 1077–1086.

Kiers, E. T., Rousseau, R. A., West, S. A., & Denison, R. F. (2003). Host sanctions and the legume–rhizobium mutualism. Nature, **425**(6953), 78–81. DOI:10.1038/nature01931.

Kochetov, G. A., & Philippov, P. P. (1970). Calcium: Co-factor of transketolase from baker's yeast. Biochem. Biophys. Res. Commun. **38**(5): 930–933.

Kochetov, G. A., & Solovjeva, O. N. (2014). Structure and functioning mechanism of transketolase. BBA, **1844**(9), 1608–1618. DOI:10.1016/j.bbapap.2014.06.003.

Koirala, A., & Brözel, V. S. (2021). Phylogeny of nitrogenase structural and assembly components reveals new insights into the origin and distribution of nitrogen fixation across Bacteria and Archaea. Microorganisms, **9**(8), 1662. DOI:10.3390/microorganisms9081662.

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, Pavel. A. (2019). Assembly of long error-prone reads using repeat graphs. Nat. Biotechnol., **37**, 540–546.

Krasova-Wade, T., Diouf, O., Ndoye, I., Sall, C. E., Braconnier, S., & Neyra, M. (2006). Water-condition effects on rhizobia competition for cowpea nodule occupancy. Afr. J. Biotechnol., **5**(16), 1457–1463.

Krol, E., & Becker, A. (2004). Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. Mol. Genet. Genomics, **272**(1), 1–17. DOI:10.1007/s00438-004-1030-8.

Krüsemann, J. L., Lindner, S. N., Dempfle, M., Widmer, J., Arrivault, S., Debacker, M., He, H., Kubis, A., Chayot, R., Anissimova, M., Marlière, P., Cotton, C. A. R., & Bar-Even, A. (2018). Artificial pathway emergence in central metabolism from three recursive phosphoketolase reactions. FEBS Journal, **285**(23), 4367–4377. DOI:10.1111/febs.14682.

Kusian, B., & Bowien, B. (1997). Organization and regulation of *cbb* CO_2 assimilation genes in autotrophic bacteria. FEMS Microbiol. Rev., **21**(2), 135–155. DOI:10.1111/j.1574-6976.1997.tb00348.x.

Labes, M., & Finan, T. M. (1993). Negative regulation of σ54-dependent *dctA* expression by the transcriptional activator DctD. J. Bacteriol., **175**(9), 2674–2681. DOI:10.1128/jb.175.9.2674-2681.1993.

Lambert, G. R., Cantrell, M. A., Hanus, F. J., Russell, S. A., Haddad, K. R., & Evans, H. J. (1985). Intra-and interspecies transfer and expression of *Rhizobium japonicum* hydrogen uptake genes and autotrophic growth capability. Proc. Natl. Acad. Sci. U.S.A., **82**, 3232–3236.

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics, **34**(18), 3094–3100. DOI:10.1093/bioinformatics/bty191.

Lindquist, S., Lindberg, F., & Normark, S. (1989). Binding of the *Citrobacter freundii* AmpR regulator to a single DNA site provides both autoregulation and activation of the inducible ampC r-Lactamase Gene. J. Bacteriol., **171**.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods, **25**(4), 402–408. DOI:10.1006/meth.2001.1262.

Lodwig, E. M., Hosie, A. H. F., Bourdes, A., Findlay, K., Allaway, D., Karunakaran, R., Downie, J. A., & Poole, P. S. (2003). Amino-acid cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. Letters to Nature, **422**(17), 722–726.

Maddocks, S. E., & Oyston, P. C. F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology, **154**(12), 3609–3623. DOI:10.1099/mic.0.2008/022772-0.

Maier, R. J. (1981). *Rhizobium japonicum* mutant strains unable to grow chemoautotrophically with H_2. J. Bacteriol., **145**(1), 533–540.

Manian, S. S., & O'Gara, F. (1982). Derepression of ribulose bisphosphate carboxylase activity in *Rhizobium meliloti*. FEMS Microbiol. Lett., **14**(2), 95–99. DOI:10.1111/j.1574-6968.1982.tb08642.x.

Martínez-De Drets, G., & Arias, A. (1972). Enzymatic basis for differentiation of *Rhizobium* into fast- and slow-growing groups. J. Bacteriol., **109**(1), 467–470. DOI:10.1128/jb.109.1.467-470.1972.

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., & Ausubel, F. M. (1982). Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. J. Bacteriol., **149**(1), 114–122. DOI:10.1128/jb.149.1.114-122.1982.

Mendoza-Suárez, M. A., Geddes, B. A., Sánchez-Cañizares, C., Ramírez-González, R. H., Kirchhelle, C., Jorrin, B., & Poole, P. S. (2020). Optimizing *Rhizobium*- legume symbioses by simultaneous measurement of rhizobial competitiveness and N_2 fixation in nodules. Proc. Natl. Acad. Sci. U.S.A., **117**(18), 9822–9831. DOI:10.1073/pnas.1921225117.

Mitra, R. M., Gleason, C. A., Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E. D., & Long, S. R. (2004). A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. Proc. Natl. Acad. Sci. U.S.A., **101**(13), 4701–4705. DOI:10.1073/pnas.0400595101.

Nargang, F., McIntosh, L., & Somerville, C. (1984). Nucleotide sequence of the ribulosebisphosphate carboxylase gene from *Rhodospirillum rubrum*. Mol. Gen. Genet., **193**, 220–224. DOI:10.1007/BF00330671.

Oldroyd, G. E. D., 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. DOI:10.1146/annurev-genet-110410-132549.

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**(12), 3582–3586. doi:10.1128/JB.182.12.3582-3586.2000.

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. t*rifolii*: Evidence for a plant-inducible rhamnose locus involved in competition for nodulation. Mol. Plant Microbe Interact.., **11**(12), 1175–1185. DOI:10.1094/MPMI.1998.11.12.1175.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research, **25**(7), 1043–1055. DOI:10.1101/gr.186072.114.

Peoples, M. B., & Craswell, E. T. (1992). Biological nitrogen fixation: Investments, expectations and actual contributions to agriculture. Plant and Soil, **141**(1–2), 13–39. DOI:10.1007/BF00011308.

Pickering, B. S., & Oresnik, I. J. (2008). Formate-dependent autotrophic growth in *Sinorhizobium meliloti*. J. Bacteriol., **190**(19), 6409–6418. DOI:10.1128/JB.00757-08.

Pini, F., East, A. K., Appia-Ayme, C., Tomek, J., Karunakaran, R., Mendoza-Suárez, M., Edwards, A., Terpolilli, J. J., Roworth, J., Downie, J. A., & Poole, P. S. (2017). Bacterial biosensors for in vivo spatiotemporal mapping of root secretion. Plant Physiol., **174**(3), 1289–1306. DOI:10.1104/pp.16.01302.

Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: From saprophytes to endosymbionts. Nat. Rev. Microbiol., **16**(5), 291–303. DOI:10.1038/nrmicro.2017.171.

Poysti, N. J., & Oresnik, I. J. (2007). Characterization of *Sinorhizobium meliloti* triose phosphate isomerase genes. J. Bacteriol., **189**(9), 3445–3451. DOI:10.1128/JB.01707-06.

Quandt, J., & Hynes, M. F. (1993). Versatile suicide vectors which allow direct selection for gene replacement in gram-negative bacteria. Gene, **127**(1), 15–21. DOI:10.1016/0378-1119(93)90611-6.

Rice, W. A., & Olsen, P. E. (1988). Root-temperature effects on competition for nodule occupancy between two *Rhizobium meliloti* strains. Biol. Fertil. Soils, **6**(2), 137–140. DOI:10.1007/BF00257663.

Rivers, D. M. R., Kim, D. D., & Oresnik, I. J. (2022). Inability to catabolize rhamnose by *Sinorhizobium meliloti* Rm1021 affects competition for nodule occupancy. Microorganisms, **10**(4). DOI:10.3390/microorganisms10040732.

Ronson, C. W., Lyttleton, P., & Robertson, J. G. (1981). C4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. Proc. Natl. Acad. Sci. U.S.A., **78**(7), 4284–4288. DOI:10.1073/pnas.78.7.4284.

Sambrook, J., Fritsch, E., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Sánchez, B., Zúñiga, M., González-Candelas, F., De Los Reyes-Gavilán, C. G., & Margolles, A. (2010). Bacterial and eukaryotic phosphoketolases: Phylogeny, distribution and evolution. J. Mol. Microbiol. Biotechnol., **18**(1), 37–51. DOI:10.1159/000274310.

Schell, M. A. (1993). Molecular biology of the Lys-R family transcriptional regulators. Annu. Rev. Microbiol, **47**, 597–626.

Schenk, G., Duggleby, R. G., & Nixon, P. F. (1998). Properties and functions of the thiamin diphosphate dependent enzyme transketolase. Int. J. Biochem. Cell B., **30**, 1297–1318.

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**(10), 5858–5864. DOI:10.1128/AEM.71.10.5858-5864.2005.

Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., Velankar, S., Burley, S. K., Koča, J., & Rose, A. S. (2021). Mol* Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res., **49**(W1), W431–W437. DOI:10.1093/nar/gkab314.

Sonderegger, M., Schümperli, M., & Sauer, U. (2004). Metabolic engineering of a phosphoketolase pathway for pentose catabolism in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol., **70**(5), 2892–2897. DOI:10.1128/AEM.70.5.2892-2897.2004.

Stowers, M. D. (1985). Carbon metabolism of *Rhizobium* species. Annu. Rev. Microbiol., **27**, 89–108.

Triplett, E. W., & Sadowsky, M. (1992). Genetics of competition for nodulation of legumes. Annu. Rev. Microbiol., 399–428.

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. Annu. Rev. Plant Biol., **64**, 781–805. DOI:10.1146/annurev-arplant-050312-120235.

Van Keulen, G., Ridder, A. N. J. A., Dijkhuizen, L., & Meijer, W. G. (2003). Analysis of DNA binding and transcriptional activation by the LysR-type transcriptional regulator CbbR of *Xanthobacter flavus*. J. Bacteriol., **185**(4), 1245–1252. DOI:10.1128/JB.185.4.1245-1252.2003.

Westhoek, A., Clark, L. J., Culbert, M., Dalchau, N., Griffiths, M., Jorrin, B., Karunakaran, R., Ledermann, R., Tkacz, A., Webb, I., James, E. K., Poole, P. S., & Turnbull, L. A. (2021). Conditional sanctioning in a legume– *Rhizobium* mutualism. Proc. Natl. Acad. Sci. U.S.A., **118**(19), e2025760118. DOI:10.1073/pnas.2025760118.

Wheatley, R. M., Ford, B. L., Li, L., Aroney, S. T. N., Knights, H. E., Ledermann, R., East, A. K., Ramachandran, V. K., & Poole, P. S. (2020). Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. Proc. Natl. Acad. Sci. U.S.A., **117**(38), 23823–23834. DOI:10.1073/pnas.2009094117.

Wick, R. R., Judd, L. M., & Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol., **20**(1), 129. DOI:10.1186/s13059-019-1727-y.

Yang, X., Yuan, Q., Zheng, Y., Ma, H., Chen, T., & Zhao, X. (2016). An engineered non-oxidative glycolysis pathway for acetone production in *Escherichia coli*. Biotechnol. Lett., **38**(8), 1359–1365. DOI:10.1007/s10529-016-2115-2.

Yarosh, O. K., Charles, T. C., & Finan, T. M. (1989). Analysis of C $_4$ -dicarboxylate transport genes in *Rhizobium meliloti*. Mol. Microbiol., **3**(6), 813–823. DOI:10.1111/j.1365-2958.1989.tb00230.x.

Yurgel, S. N., & Kahn, M. L. (2005). *Sinorhizobium meliloti dctA* mutants with partial ability to transport dicarboxylic acids. J. Bacteriol., **187**(3), 1161–1172. DOI:10.1128/JB.187.3.1161-1172.2005.