Brigham Young University
BYU ScholarsArchive

All Theses and Dissertations

2013-03-20

Modulators of Symbiotic Outcome in Sinorhizobium meliloti

Matthew B. Crook
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

 Part of the Microbiology Commons

BYU ScholarsArchive Citation

Crook, Matthew B., "Modulators of Symbiotic Outcome in Sinorhizobium meliloti" (2013). *All Theses and Dissertations*. 3946.
https://scholarsarchive.byu.edu/etd/3946

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Modulators of Symbiotic Outcome in *Sinorhizobium meliloti*

Matthew Ben Crook, Jr.

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Joel S. Griffitts, Chair
Brent Nielsen
William R. McCleary
Jeff Maughan
David L. Erickson

Department of Microbiology and Molecular Biology

Brigham Young University

March 2013

Copyright © 2013 Matthew Ben Crook, Jr.

All Rights Reserved

ABSTRACT

Modulators of Symbiotic Outcome in *Sinorhizobium meliloti*

Matthew Ben Crook, Jr.
Department of Microbiology and Molecular Biology, BYU
Doctor of Philosophy

Microorganisms interact frequently with each other and with higher organisms. This contact and communication takes place at the molecular level. Microbial interactions with eukaryotes can be pathogenic or mutualistic. One of the best-studied symbioses is the complex interaction between nitrogen-fixing soil bacteria, termed rhizobia, and legumes. This symbiosis culminates in the elaboration of a new organ, the root nodule. Many of the molecular signals exchanged between the host plant and the invading rhizobia have been deduced, but there is still much that remains to be discovered. The molecular determinant of host range at the genus level of the plant host has been determined to be lipochitooligomers called Nod factors. The molecular determinants of host range at the species and cultivar level are less well-defined. Part of my work has been to identify and characterize accessory plasmids that disrupt the normal progression of symbiosis between legumes of the genus *Medicago* and their rhizobial symbiont, *Sinorhizobium meliloti*. A *cre–loxP*-based system capable of making large, defined deletions was developed for the analysis of these plasmids. This system is also being employed to cure the laboratory strain, *S. meliloti* Rm1021 of its two megaplasmids—a loss of nearly half of its genome. I have also done work to determine whether locally-collected sinorhizobia are native, invasive, or native with symbiosis genes acquired horizontally from invasive sinorhizobia. Finally, I have studied *Sinorhizobium meliloti* as a host by identifying an outer membrane porin that several bacteriophages use to adsorb to the *S. meliloti* cell surface.

Key words: *Sinorhizobium meliloti*, *Medicago*, symbiosis, host range, plasmid, pHRC017, bacteriophage receptor, RopA1

ACKNOWLEDGMENTS

First and foremost I would like to thank my wife, Leann, who over the last six years has tolerated my working late, my working weekends, my working holidays, and lately, the long hours spent preparing this dissertation. Her patience and loving support has been integral to my success. I would also like to thank my parents who bought me my first microscope and helped me with my first insect collection when I was in first grade. They always encouraged my curiosity and provided me with access to useful sources of knowledge. When at home I could often be found reading from a set of World Book Encyclopedias they kept on their shelves.

My mentor, Dr. Griffitts, has taken a personal interest in my development as a scientist and has helped me to take great strides in thinking analytically and scientifically. His tutelage and expertise have made a deep impact on my approach to scientific inquiry. My association with the other members of my graduate committee has also been exceptional. They always gave good advice during my committee meetings and provided me with new insights into my work. I have also received help, support, and camaraderie from other members of the MMBIO faculty. The office staff of the department, both past and present, has been exactly what they should be: helpful and knowledgeable. I always came away from my interactions with them feeling satisfied and encouraged. I would also like to thank Dr. Kim O'Neill and Dr. Paul Savage, who gave me research opportunities before I began my graduate program.

I would like to thank the many BYU undergraduates who have contributed to this work in some way: Joshua Bentley, Michael Bevins, Matthew Biggs, Divyesh Choudhri, Casey Crum, Alicia Draper, Krysta Felix, Jordan Guillory, Clarice Harrison, Daniel Lindsay, Gledi Peco, Shawna Rogers, and Elise Scoggin. I have also received invaluable advice and support from the other graduate students in the lab, Ryan VanYperen and Phil Bennellack, and from Skip Price, a

post-doctoral fellow in the lab who will be continuing the study of host range plasmids after I leave and who generated most of the data presented in **§2.1.6**.

I would also like to thank members of the BYU Computer Science Department who helped with the assembly and annotation of pHRC017 and its subsequent submission to GenBank: Jared C. Price, Spencer C. Clement, and Dr. Mark J. Clement. Camille Porter, a graduate student in the Biology Department, helped train me in phylogenetics and gave me access to the software I needed.

Finally, many other scientists have provided materials and expertise without which this project would not have been possible: P. Van Berkum of the USDA–ARS, M. Hynes of the University of Calgary, H. Krishnan of the University of Missouri, S. Long of Stanford University, I. Oresnik of the University of Manitoba, M. Standing in the microscopy lab here in the college, and Graham Walker of the Massachusetts Institute of Technology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A17	*Medicago truncatula* Gaertn. cv. A17
A20	*Medicago truncatula* Gaertn. cv. A20
ACC	1-aminocyclopropane-1-carboxylate
AM	arbuscular mycorrhizal
AR	*Medicago arabica* (L.) Huds.
CI	competition index
Cm^R	chloramphenicol resistance
CO	*Medicago constricta* Durieu
bp	base pairs
bv.	biovar
cv.	cultivar
DI	*Medicago disciformis* DC.
dpi	days post-inoculation
EPS I	exopolysaccharide I (i.e. succinoglycan)
EPS II	exopolysaccharide II (i.e. galactoglucan)
Gm^R	gentamicin resistance
GOC	gain-of-compatibility
HR	host range
IT	*Medicago italica* (Mill.) Fiori
Kb	kilobases
Km^R	kanamycin resistance
LB	lysogeny broth
LE	*Medicago lesinsii* E. Small
LU	*Medicago lupulina* L.
Mb	megabases
N	nitrogen
neo	neomycin resistance cassette
Nm^R	neomycin resistance
OD_{600}	optical density at $\lambda = 600$ nm
oriT	origin of conjugative transfer
oriV	origin of vegetative replication
PCR	polymerase chain reaction
Φ	phage
PR	*Medicago praecox* DC.
Rf^R	rifampicin resistance
ropA1	rhizobial outer membrane protein A1
ropA2	rhizobial outer membrane protein A2
RU	*Medicago rugosa* Desr.
SA	*Medicago sativa* L.
SD	standard deviation
SEM	standard error of the mean
Sm^R	streptomycin resistance
Sp^R	spectinomycin resistance
sv.	symbiovar (7)

T3E	type III effector
T3SS	type III secretion system
TA	toxin–antitoxin
Tc^R	tetracycline resistance
TE	*Medicago tenoreana* Ser.
tet	tetracycline resistance cassette
Tn*5*	transposon *5*
TY	tryptone–yeast extract broth
YEM	yeast extract–mannitol broth

1 INTRODUCTION

Legumes (Family Fabaceae or Leguminosae) are the third largest flowering plant family and exhibit astonishing diversity, with over 700 genera and more than 20,000 species (8). They evolved 50–60 million years ago along semi-arid regions of the Tethys seaway (2, 3), but now they are found throughout the world (**Figure 1.1-1**). The family Leguminosae is further subdivided into three subfamilies, the Caesalpinioideae, the Mimosoideae, and the Papilionoideae, the latter being the largest. Nodulation is observed less often in the Caesalpinioideae than in the other two legume subfamilies.

Legumes are ecologically important because they convert inert atmospheric dinitrogen to a biologically useful form, ammonia, through a symbiotic association with soil bacteria called rhizobia. This extremely beneficial process evolved only a few million years after the appearance of legumes (9). Most rhizobia belong to the class α-Proteobacteria, but there are some rhizobia

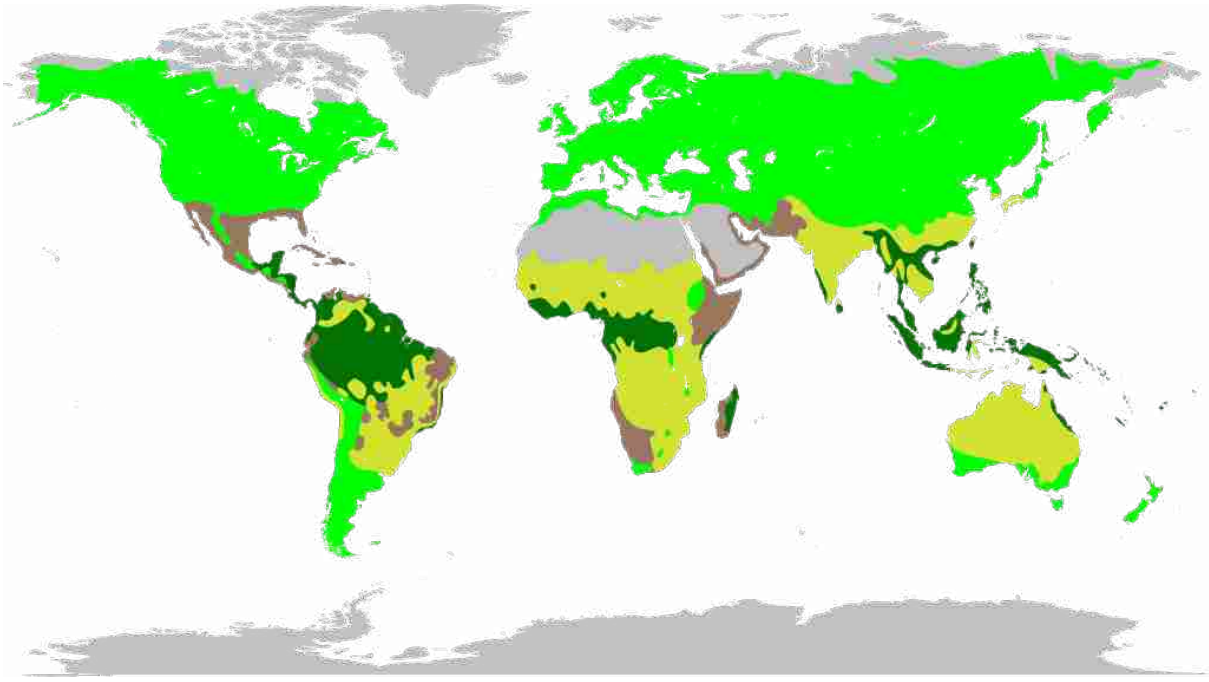

Figure 1.1-1. Legumes are found in four biomes throughout the world. The four major legume biomes are rainforest (dark green), temperate (green), grassland (yellow), and succulent (brown). Legumes are not found in areas of the map that are gray. Based on (2, 3).

among the β-Proteobacteria, including members of the genera *Burkholderia* and *Cupriavidus* (10). Rhizobia in association with legumes are responsible for fixing 1 to 3×10^{14} grams per year (11) and account for roughly half of the nitrogen introduced or returned to agricultural soil (12, 13). Economically important legumes include food crops, such as soybean (*Glycine max*), common bean (*Phaseolus vulgaris*), and pea (*Pisum sativum*) as well as important forage crops and green manures, such as clover (*Trifolium* spp.), alfalfa (*Medicago sativa*), and barrel medic (*M. truncatula*). My research has focused on the interaction between *Medicago* species, primarily *M. truncatula*, and their rhizobial symbionts, *Sinorhizobium meliloti* and *S. medicae*. The *M. truncatula–S. meliloti* symbiosis is a model system for the indeterminate legume symbiosis generally (14). Several *S. meliloti* isolates have been sequenced (15–17) and the genome of *M. truncatula* cv. A17 was recently published (18).

Interaction between rhizobia and legumes culminates in the formation of a new plant organ, termed a nodule, on the roots of the plant. Root nodules come in a variety of forms but can be generally broken down into two broad types, determinate and indeterminate. Cellular division at the nodule primordium of determinate nodules begins at the outer cortex and halts quickly. Subsequent growth in the size of the nodule is due to an increase in cellular volumes rather than in cell numbers (6). This gives rise to root nodules which are generally spherical in shape. Determinate nodules are generally found on tropical and sub-tropical legumes, such as soybean (*Glycine max*), bean (*Phaseolus vulgaris*), and bird's-foot trefoil (*Lotus corniculatus*) (6). As determinate nodules age, a senescent zone can appear in the center of the nodule where nitrogen fixation ceases.

In contrast, cellular division at the nodule primordium of indeterminate nodules begins at the inner cortex and develops into an apical meristem. Subsequent growth in the size of the

nodule is due to both increases in cellular volume and in cell numbers (6). This gives rise to nodules which are elongated, though branching may occur. Indeterminate nodules are generally found on temperate legumes, such as alfalfa (*Medicago sativa*), pea (*Pisum sativum*), and vetch (*Vicia sativa*) (6).

The continual addition of cells to the nodule by the dividing meristem occasions the need for a concomitant infection of new cells by the invading rhizobia. Thus there is a stratified structure to the indeterminate nodule that is lacking in the determinate nodule (**Figure 1.1-2**). Just below the apical meristem (Zone I) is an infection zone (Zone II) where rhizobia are found in the process of exiting the infection thread and becoming encased in a plant-derived membrane called the peribacteroid membrane. At this point the plant fills the nodule with leghemoglobin giving the nodule a characteristic pink color. Between Zones II and III is a narrow intermediate zone (II–III) where bacteroids stop elongating and they fill the cytosol of host cells almost

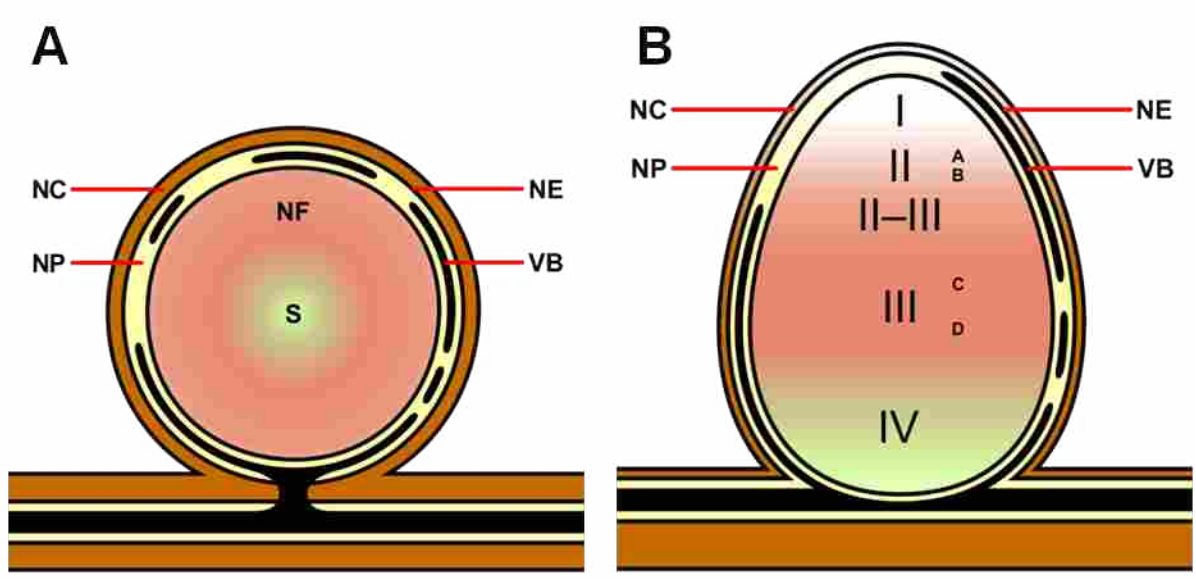

Figure 1.1-2. Root nodules can be classified as either determinate or indeterminate. The different zones of both types are shown. **A.** Determinate nodules have a nitrogen fixation zone (NF) and a senescent zone (S). **B.** Indeterminate nodules have a an apical meristem (I), an infection zone (II), an invasion zone (IIA), a pre-fixing zone (IIB), an intermediate zone (II–III), a nitrogen fixation zone (III), which consists of efficient nitrogen (IIIA) and inefficient (IIIB) zones, and a senescent zone (IV). Other features are labeled as follows: NC, nodule cortex; NE, nodule endodermis; NP, nodule parenchyma; and VB vascular bundle. Based on (6)

completely. In Zone III the bacteroids begin actively fixing nitrogen. Eventually the proximal layers of the root nodule senesce (Zone IV) (19). Nitrogen fixation ceases and the cells become green due to the breakdown of the heme moiety of hemoglobin.

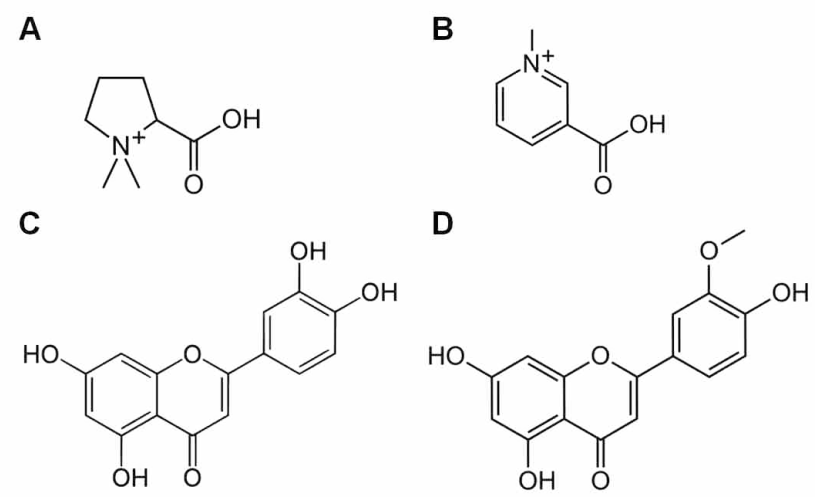

Figure 1.1-3. Key root exudates of *Medicago* species. The betaines **A.** stachydrine and **B.** trigonelline and the flavonoids **C.** luteolin and **D.** chrysoeriol.

The elaboration of a root nodule is a complex process with multiple signals exchanged between the rhizobial symbiont and the host plant. The first step is recruitment of rhizobia to the root hairs of the plant via the secretion of flavonoids, isoflavones, and betaines into the rhizosphere, which are recognized by the bacteria (13, 20). Many compounds have been identified in the root exudates of *M. sativa* and/or *M. truncatula*: 4,4′-dihydroxy-2′-methoxychalcone, chrysoeriol, cynaroside, daidzein, 7,4′-dihydroxyflavone, eriodictyol, liquiritigenin, luteolin, 6″-O-malonylononin, 3′,4′-dimethoxyluteolin, 5-methoxyluteolin, medicarpin, naringenin, stachydrine, and trigonelline (21–26). (For example structures of these molecules, see **Figure 1.1-3**.) Of these the most important for nodulation by *S. meliloti* is luteolin. Luteolin induces the *nodD*$_1$ gene of *S. meliloti* and interacts with its transcription factor product to stimulate transcription of the other *nod* genes (27, 28). The betaines stachydrine and trigonelline perform a similar function with respect to the *nodD*$_2$ gene of *S. meliloti* (25).

Induction of the *nod* genes culminates in bacterial production of lipochitooligosaccharides commonly referred to as Nod factors. The major active Nod factor of *S. meliloti* consists of four *N*-acetyl-D-glucosamine (GlcNAc) residues in β-1,4 linkage (chitin) *N*-acylated with C_{16} *bis*-unsaturated fatty acid at the nonreducing end and substituted with a sulphate ester at the reducing end (1) (**Figure 1.1-4**, left). Nod factors produced by other rhizobia can differ in the length and saturation of the fatty acid, the number of GlcNAc residues, and the type and placement of other chemical substituents (**Figure 1.1-4**, right). It is most common for *N*-methyl, *O*-acetyl, and *O*-carbamoyl groups to occur on the nonreducing residue and L-fucosyl, 2-*O*-methylfucosyl, 4-*O*-acetylfucosyl, acetyl, and sulfate ester groups to occur at the reducing residue (4). The specific structure of Nod factor is required to initiate nodule formation by the plant and is considered the major determinant of host range (29, 30). However, nod-independent rhizobial symbioses have been described (31–33). Lipochitooligosaccharides are also used by arbuscular mycorrhizal (AM) fungi as a signal to initiate colonization of the roots of host plants (34). Since the mycorrhizal symbiosis is more ancient than the rhizobial symbiosis, it is believed

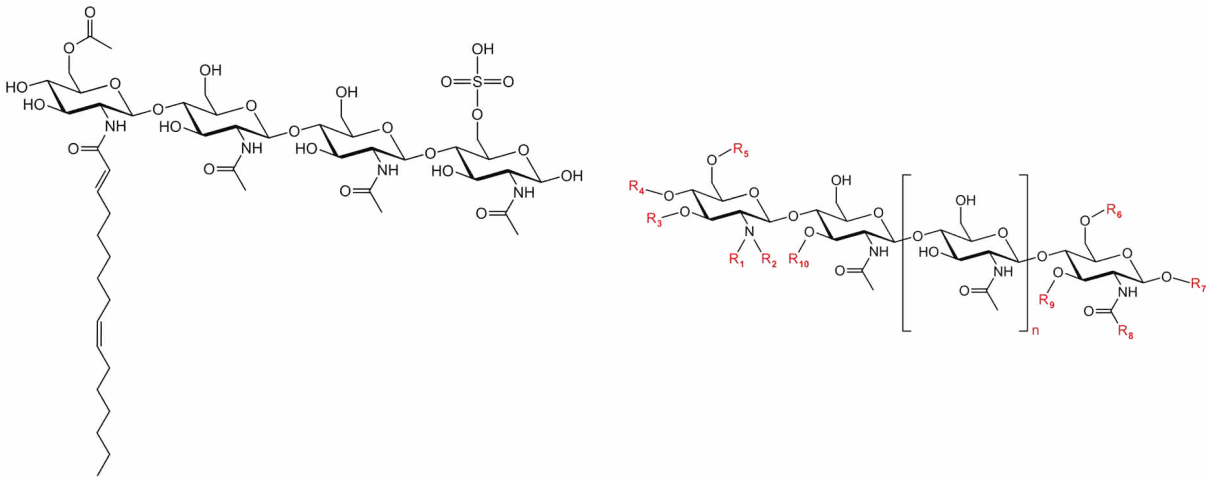

Figure 1.1-4. Specific and general structures of Nod factor. The structure of the symbiotically active Nod factor of *S. meliloti*, NodSm-IV (Ac,C16:2,S), is shown on the left (1). A generic Nod factor structure is shown on the right (4). R1 through R10 can be modified by various chemical substituents (see text). R1 is always a fatty acid chain and *n* indicates the number of repeats, from 0 to 2.

that rhizobia co-opted the signalling pathways of the AM fungi (35).

Recognition of the correct Nod factor by the host plant stimulates several physiological changes. The root hair colonized by the rhizobia deforms, entrapping the attached bacteria. Subsequently, an invagination, termed the infection thread, proceeds down through the root hair (a process which requires substantial remodeling of the cytoskeleton) and into underlying layers of the root cortex. The rhizobia colonize this infection thread by cellular division. Concomitantly, binding of Nod factor to the plant Nod factor receptor stimulates calcium spiking in and around the nucleus of the root hair cell, which may prepare the cell for the on-coming infection (36). Nod factor perception also results in the stimulation of root cortical cell division leading to the development of a nodule primordium and eventually to the apical meristem of the indeterminate nodule (37). The *nod*, *nol*, and *noe* genes encode enzymes responsible for the production, regulation, and export of Nod factor.

During rhizobial growth in the infection thread, several more molecular signals are elaborated by the bacterium. In the *Medicago–Sinorhizobium* symbiosis these signals are low-molecular-weight succinoglycan (EPS I), low-molecular-weight galactoglucan (EPS II), capsular polysaccharide (KPS), cyclic glucans, and lipopolysaccharide (LPS) (**Figure 1.1-5**). Each will be explained in more detail below.

The symbiotically active form of succinoglycan is made up of a backbone consisting of one galactose and seven glucose residues in various β linkages (38). Substituted onto this backbone are a pyruvate, an acetyl group, and a succinate (**Figure 1.1-5E**). Succinoglycan species with various succinate and malate substitutions are produced by *S. meliloti* (39), however at least one succinate is required for infection of *M. sativa* (40). Additional adducts to the succinoglycan may contribute to host specificity (39). Succinoglycan is produced in both high-

molecular weight (HMW) and low-molecular-weight (LMW) fractions, but only the LMW fraction is symbiotically active (41). Failure to produce the correct succinoglycan can result in defects in attachment to the root hair, root hair curling, and initiation and elongation of the infection thread (42). The *exo* genes encode enzymes for production, export, and regulation of succinoglycan.

Galactoglucan is not produced by *S. meliloti* Rm1021 due to a transposon insertion in the regulatory gene, *expR* (43). However, in revertants it is able to suppress the effects of succinoglycan mutants for nodulation of alfalfa (44). Galactoglucan is made up of glucose-β-

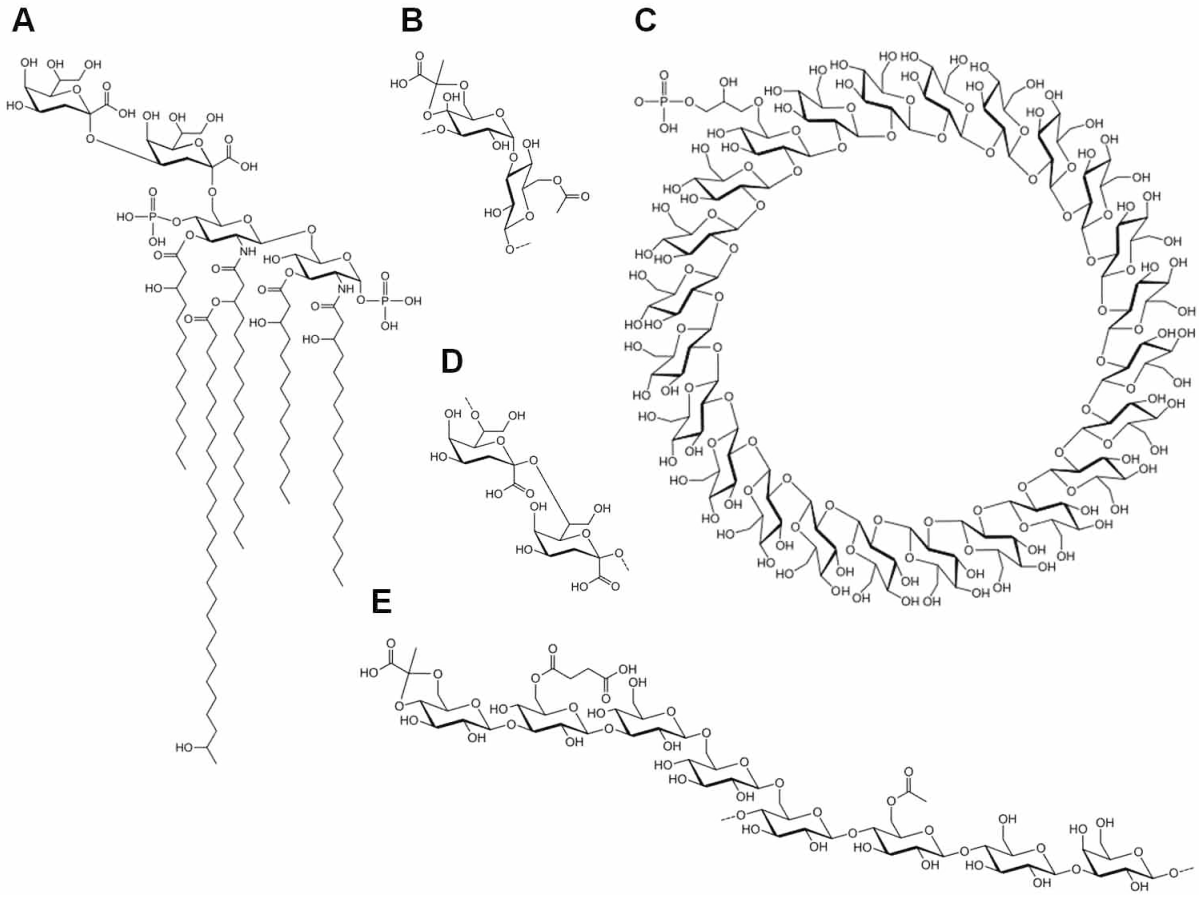

Figure 1.1-5. Secreted and surface polysaccharides of *S. meliloti*. A. A partial structure for the LPS of *S. meliloti*; **B.** the repeating subunit of galactoglucan (EPS II); **C.** a cyclic β-(1→2)-glucan 24-mer (the degree of substitution with *sn*-1-phosphoglycerol is variable); **D.** the repeating subunit of capsular polysaccharide (KPS), **E.** the repeating subunit of succinoglycan (EPS I).

(1→3)-galactose repeating units modified with acetyl and pyruvyl groups (45) (**Figure 1.1-5B**). As with succinoglycan, the LMW form of galactoglucan is symbiotically active (46). The *wg*- (formerly *exp*) genes are responsible for production of galactoglucan biosynthetic enzymes.

S. meliloti Rm41 produces a capsular polysaccharide (KPS) that can substitute for succinoglycan in nodulation of alfalfa (47). Like other rhizobial polysaccharides, its symbiotic function is dependent on chain length (48). *S. meliloti* Rm1021 lacks the necessary *rkpZ* gene required to produce symbiotically active KPS (49) but it does produce a symbiotically inactive capsular polysaccharide which consists of a lipid-linked homopolymer of 3-deoxy-D-manno-oct-2-ulosonic acid (50) (**Figure 1.1-5D**). The KPS of *S. meliloti* Rm41 is a highly polymerized disaccharide of glucuronic and pseudaminic acids decorated with butyryl and acetyl residues (51, 52). The *rkp* genes are responsible for production of capsular polysaccharide biosynthetic enzymes.

Cyclic glucans, also known as cyclosophoraoses, consist of seventeen to forty β-(1→2)-linked glucose residues, decorated with a variable number of phosphoglycerol moieties (53–55) (**Figure 1.1-5C**). They are produced by the cell in response to osmotic conditions and are maintained in the periplasm (56). Mutants deficient in the production of cyclic glucans are unable to initiate infection thread formation (57). A role for cyclic glucans in enhancing Nod factor delivery and/or suppressing the host immune response has also been postulated (58). The *ndv* and *cgm* genes are responsible for production and export of cyclic glucans.

When the cells colonizing the infection thread reach the root cortex they enter the plant cells enveloped in host-derived membranes. The bacteria terminally differentiate into a form known as bacteroids. The bacteroids exhibit altered cell morphology, becoming elongated and sometimes even branched, and they endoreduplicate their genome approximately 24 times (59).

This is thought to be due to the production of nodule-specific cationic peptides in hosts that produce indeterminate nodules (59–61). The plant cells that actively take up the rhizobia also exhibit polyploidy of their genome and increased cellular volume (37). At this stage the rhizobia begin fixing nitrogen and providing it to the plant in exchange for fixed carbon and other nutrients (62). Bacteroids in determinate nodules are able to de-differentiate upon release from the nodule, but bacteroids in indeterminate nodules are terminally differentiated.

The LPS of *S. meliloti* inhibits the immune response of its *M. truncatula* host (63). A variety of defects in the carbohydrate composition of LPS, particularly in the LPS core, can lead to the inability to persist as bacteroids within the host cells, decreased nitrogen fixation, and increased sensitivity to cationic peptides (64, 65). A defect in distribution of LPS fatty acids (*bacA*$^-$) also affects long-term survival of the bacteroid (66). This last finding is only true for survival in indeterminate nodules, suggesting that loss of LPS renders the bacteroid susceptible to nodule-specific cationic peptides produced by hosts that produce indeterminate nodules (67). Alternatively, it has been observed that *bacA*$^-$ mutants do not differentiate, which may incur sanctioning by the host plant (59). The *lps* genes are responsible for production of LPS biosynthetic enzymes.

Nitrogen fixation is an energy-intensive process because it involves the breaking of all three bonds of a triple-bonded dinitrogen:

$$N_2 + 16ATP + 8H^+ + 8e^- \rightarrow 2NH_3 + 16ADP + 16P_i + H_2$$

The reaction is catalyzed by a bacterial enzyme called nitrogenase. Nitrogenase houses several iron-sulfur clusters for electron transfer and an iron-molybdenum cofactor (FeMo-co) which lies in the active site (68). Nitrogen is then delivered to the plant as ammonia and as amino acids (69). The *nif* genes are responsible for production and regulation of nitrogenase.

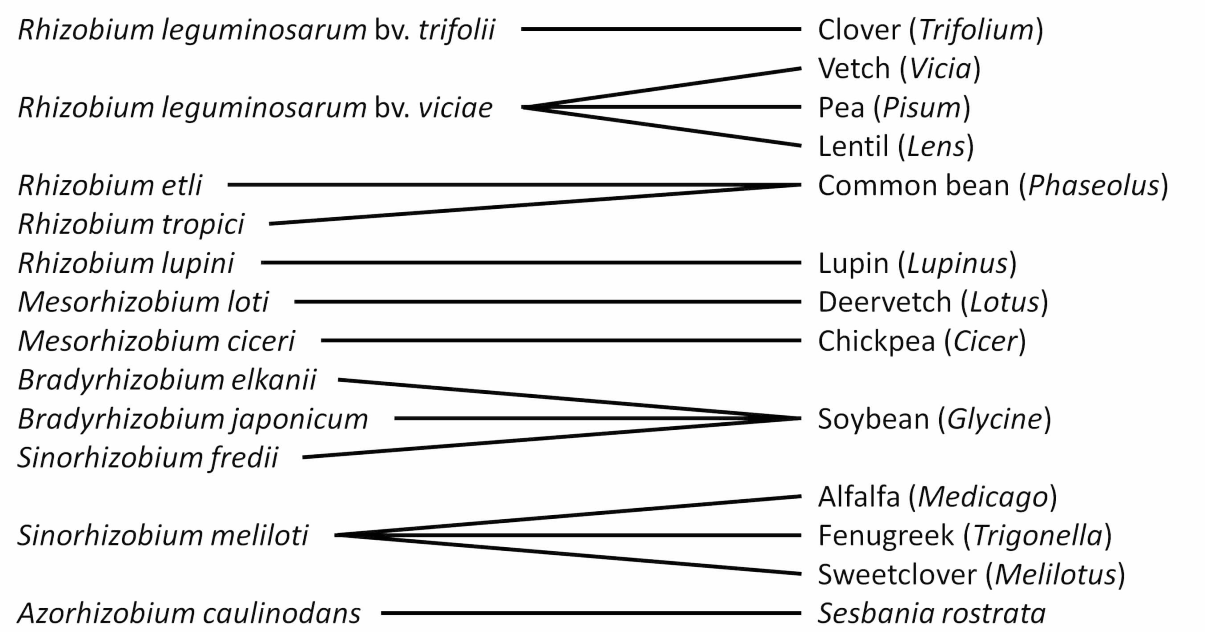

Figure 1.1-6. Cross-inoculation groups of legumes and their inoculant rhizobia.

In addition to its energy requirements for reduction of nitrogen gas, nitrogenase also has to be protected from oxygen, which can react to its metal centers, poisoning them (70). To create a microoxic environment for the rhizobia, the plant fills the nodule with leghemoglobin which binds up the oxygen. However, rhizobia are aerobic bacteria and use oxygen as their terminal electron acceptor (71). So to overcome the challenges of respiration in the microaerobic nodule, especially in the context of the high demand for electrons to contribute to nitrogen fixation, rhizobia have evolved an alternative electron transport chain which can donate electrons directly to nitrogenase (70). The *fix* genes encode production of the alternative electron transport chain, as well as other functions related to symbiotic nitrogen fixation (70).

Given the many molecular requirements to establish and maintain the legume–rhizobum symbiosis, alteration of any of the chemical signals can alter or abolish the host range of the rhizobium. I have described here the symbiotic program leading to nodulation of *Medicago truncatula* and/or *Medicago sativa* by *Sinorhizobium meliloti*. However, many other host–strain

combinations exist and they often have alternative molecular requirements (72). Due to the generally restrictive nature of partner recognition, host–strain pairs have historically been sorted into cross-inoculation groups (**Figure 1.1-6**). However, the relationships between rhizobia and their hosts are proving to be more entangled than once thought. For example, it is generally accepted that *Sinorhizobium meliloti* nodulates only members of the subtribe Trigonellinae (*Medicago*, *Melilotus*, and *Trigonella* (73)) and that *Rhizobium leguminosarum* bv. *viciae* is the symbiont of the common bean (*Phaseolus vulgaris*) (74). However, strains of *Sinorhizobium meliloti* which infect *Phaseolus vulgaris* (75, 76) and a strain of *Rhizobium* which infects *Medicago sativa* (77) have been recently reported.

Plasmids play a prominent role in the *Medicago–Sinorhizobium* symbiosis. The core *S. meliloti* genome consists of a ~3.6 Mb chromosome and two megaplasmids, pSymB (~1.7 Mb) and pSymA (~1.4 Mb). Many of the genes required for symbiosis are found on pSymA (*nod*, *nif*, and *fix* genes) and pSymB (*exo*, *eps*, and *rkp*). Additionally, the acquisition of accessory plasmids can alter the symbiotic program. Accessory plasmids can encode bacteriocins, which give them an advantage in the rhizosphere (78), or rhizobactin, a nitrogenous compound synthesized by bacteroids but catabolized by their free-living clonemates and which confers a competitive advantage for nodule occupancy (79). The 1-aminocyclopropane-1-carboxylate deaminase gene, *acdS*, which is often found on accessory plasmids, also confers competitive advantage for nodulation (80).

Several reports have implicated accessory plasmids in enhancing root nodule symbiosis for their rhizobial host. Plasmids other than the Sym plasmids of *R. leguminosarum* (81) and *R. etli* (82) have been reported to enhance nodulation. Two accessory plasmids of *R. tropici*, pRt899a from strain CIAT 899 (83) and an unidentified accessory plasmid of strain CFN299

enhance nodulation and competition for occupation of nodules (84). Loss of an accessory plasmid in *Rhizobium galegae* NBIMTC2250 was sometimes associated with a variety of symbiotic alterations, both positive and negative, but likely was not the cause (85). Loss of pTA2 from *S. meliloti* IZ450 (86) and a deletion in an accessory plasmid of *S. meliloti* P108 (87) rendered their host strains less effective and less competitive for nodule occupancy. An accessory plasmid in *S. fredii* USDA 206 enhances competition for nodulation (88), possibly due to altered levels of exopolysaccharide production (89). Plasmid pVS2 of *S. meliloti* 1076 carries *nod* and *nif* genes that may contribute to enhanced nodulation and/or nitrogen fixation (90). Additionally, transformation of *S. meliloti* CIAM 1759 with the pRi plasmid from *Agrobacterium rhizogenes* 15834 increased its competitiveness for nodule occupation (91). Plasmid pSmeSM11a, from *S. meliloti* SM11, a dominant strain recovered from a long-term field release experiment, contains copies of *nod* genes and an *acdS* gene (92). Whether or not pSmeSM11a is responsible for strain SM11 being dominant has not been determined. Plasmid pRmeGR4b of *S. meliloti* GR4 was shown to contain *nfe* genes which enhance nodulation (93–95). Plasmids related to pRmeGR4b also enhanced the ability to occupy nodules (96).

Less well-studied are instances of an accessory plasmid negatively affecting rhizobial symbiosis. A plasmid-cured derivative of *Mesorhizobium loti* NZP2037 was more effective at fixing nitrogen and was more competitive for nodulation than wildtype (97). Loss of the plasmid pRleF41b of *R. leguminosarum* bv. *viciae* F41 improved symbiotic effectiveness on *Pisum sativum* and reduced it when introduced to strain P13 (98). Loss of an accessory plasmid, pSmeSAF22c, from *S. meliloti* SAF22 improves symbiotic effectiveness on *M. sativa* (99). There are no reports of further characterization of the negative effects of these accessory plasmids, thus little is known about the role they play in altering symbiosis.

Pressure from bacteriophage infection may also affect symbiotic outcome. Rhizobial mutants leading to phage resistance have implicated Nod factor (100), succinoglycan (EPS I) (101), galactoglucan (EPS II) (102), lipopolysaccharide (LPS) (64, 103), capsular polysaccharide (KPS) (47, 104), and cyclic β-glucans (105). Due to their importance for symbiosis, altering any of these molecules to achieve phage resistance could have the additional cost of limiting host range (106). Thus susceptibility to rhizobiophage may tilt nodule occupancy in favor of less-efficient indigenous strains (107, 108). In some cases it could be an outer membrane protein involved in biosynthesis or transport of a signal molecule rather than the molecule itself that functions as a phage receptor, but this could still affect symbiosis (109). Subsequent analysis of these phage receptors has generally been limited to their effect on symbiosis. How they alter phage infectivity, in most cases, is not known.

In this work I set out to identify determinants of host range that occur after Nod factor perception. Nod factor generally determines host specificity at the genus level of the legumes, but very little is known about what determines host specificity at the species or cultivar level. In the course of my investigations I identified several accessory plasmids in wild isolates of *S. meliloti* that were responsible for restricting host range. While pursuing this course, I also discovered an essential *S. meliloti* porin, RopA1, which serves as a receptor for several bacteriophages. While there is no evidence of a role for this porin in the legume symbiosis (except for its essentiality for cell viability), the study of it still fits within the context of host range since phage binding is dependent on the structure of the porin. However, in this scenario *S. meliloti* has taken the role of the host.

2 OPUS

2.1 Rhizobial Plasmids That Cause Impaired Symbiotic Nitrogen Fixation and Enhanced Host Invasion

2.1.1 Summary

The genetic rules that dictate legume–rhizobium compatibility have been investigated for decades, but the causes of incompatibility occurring at late stages of the nodulation process are not well understood. An evaluation of naturally diverse legume (genus *Medicago*) and rhizobium (genus *Sinorhizobium*) isolates has revealed numerous instances in which *Sinorhizobium* strains induce and occupy nodules that are only minimally beneficial to certain *Medicago* hosts. Using these ineffective strain-host pairs, we identified gain-of-compatibility (GOC) rhizobial variants. We show that GOC variants arise by loss of specific large accessory plasmids, which we call HR plasmids due to their effect on symbiotic host range. Transfer of HR plasmids to a symbiotically effective rhizobium strain can convert it to incompatibility, indicating that HR plasmids can act autonomously in diverse strain backgrounds. We provide evidence that HR plasmids may encode machinery for their horizontal transfer. On hosts in which HR plasmids impair N fixation, the plasmids also enhance competitiveness for nodule occupancy, showing that naturally occurring, transferrable accessory genes can convert beneficial rhizobia to a more exploitative lifestyle. This observation raises important questions about agricultural management, the ecological stability of mutualisms, and the genetic factors that distinguish beneficial symbionts from parasites. (110)

2.1.2 Introduction

In a specialized organ known as the root nodule, the legume–rhizobium symbiosis links the energy-harvesting process of photosynthesis with the energy-demanding process of biological nitrogen (N) fixation and is, thus, of great ecological and agricultural importance. The question of how plant-microbe compatibility is encoded in this symbiosis has been the subject of intense inquiry over the last three decades, with efforts focused on early recognition events that are required for the initiation of infection and nodule development (36, 111). In some instances, however, naturally occurring legume–rhizobium incompatibility (defined here as the inability of the symbiosis to support plant growth) becomes evident only after nodule induction, resulting in visible nodules that do not fix N (39, 112–115). This abortive nodulation phenomenon is particularly problematic for a host plant, as it invests in the assembly of nodules that will be of no use, while potentially foregoing opportunities to engage in symbiosis with more beneficial strains. Indeed, there is abundant documentation from agricultural settings of superior rhizobium inoculants being outcompeted for nodule occupancy by indigenous rhizobium strains that are less-effective N fixers (107, 116). Abortive nodulation may represent an important transition from a mutualistic lifestyle to that of a parasite or vice versa. While evolutionary and ecological models for such transitions have been widely addressed (117–121), the genetic mechanisms behind them have been difficult to resolve.

The N-fixing symbiosis between *Sinorhizobium meliloti* and members of the legume genus *Medicago* (alfalfa and its relatives) has become one of the best-understood symbioses at the level of molecular genetic mechanisms and signal exchange. Mutations in the standard *S. meliloti* laboratory strain (Rm1021) have proven useful for identifying genes that contribute to effective nodulation and N fixation on *Medicago* hosts with which it is naturally compatible.

More recently, the genetic and phenotypic diversity of wild *S. meliloti* isolates has been investigated (16, 112, 122). These studies have helped to develop a sufficient understanding of *S. meliloti* as a species to allow molecular-level analyses of naturally variable traits. Here, we used a large set of natural *S. meliloti* strains and a panel of diverse *Medicago* hosts to first identify symbiotically incompatible host-strain pairs and, then, to isolate *S. meliloti* derivatives that overcome the initial incompatible condition. We show that this phenotypic change is brought about by the loss of accessory plasmids that, in addition to limiting symbiotic host range, also confer a hypercompetitive phenotype with respect to nodule invasion.

2.1.3 Materials and Methods

2.1.3.1 Bacterial culture.

Escherichia coli and *S. meliloti* cultures were grown at 37 and 30°C, respectively, in lysogeny broth (LB) supplemented with appropriate antibiotics as follows (in micrograms per milliliter): chloramphenicol, 30; kanamycin, 30; neomycin, 100; rifampicin, 100; streptomycin, 200; and tetracycline, 5. Rhizobia were extracted from nodules by surface sterilization and maceration (details below).

2.1.3.2 Plasmid and strain construction.

Plasmids and strains used in this study are listed in **Table S5.2-1**. Plasmids were constructed using standard techniques with enzymes purchased from New England Biolabs (Ipswich, MA, U.S.A.) The high-fidelity polymerase *Pfx*50 (Invitrogen, Carlsbad, CA, U.S.A.) was used for insert amplification. All custom oligonucleotides were purchased from Invitrogen and are listed in **Table S5.2-2**. Mobilization of plasmids was accomplished by triparental mating with helper *E. coli* B001 (DH5α harboring plasmid pRK600). pRK600 expresses trans-acting proteins required for mobilization of plasmids harboring the RK2 transfer origin (*oriT*). For modification of HR plasmids to enable pRK600-mediated transfer, an *oriT*/*neo* cassette was introduced by single-crossover homologous recombination using the pUC-based plasmid pJG194 (123). pJG194 was targeted (nondisruptively) to the *acdS* locus in pHRC017, pHRB469, and pHRB800 and the *traA* locus in pHRC377. Fragments corresponding to the regions upstream of *acdS* and *traA* were amplified using primer pairs oJG1127 and oJG1128 and oMC089 and oMC090, respectively, followed by ligation into pJG194 to yield integration plasmids pJG461 (for C017), pJG463 (for B469), pJG476 (for B800), and pJG499 (for C377). These modified pJG194 constructs were introduced into *S. meliloti* strains by triparental mating. When

appropriate, *S. meliloti* was made tetracycline resistant by integration of pJG505 into the *rhaK–icpA* intergenic region of the chromosome (124). The *rhaK–icpA* fragment in pJG505 was amplified using primers oMC014 and oMC015. HR plasmid transfer experiments depicted in **Figure 2.1-4** were accomplished by triparental mating, using the helper strain B001 (described above). Briefly, a parent *S. meliloti* strain carrying an *oriT*/*neo*-modified HR plasmid was mixed with *A. tumefaciens* UBAPF2 (125) for approximately 12 h, followed by selection on neomycin and rifampicin (to counterselect the donor). The resulting strain was then mixed with the tetracycline-resistant *S. meliloti* recipient (e.g., B464*tet*) for approximately 12 h, followed by selection on neomycin and tetracycline (the latter to counterselect the *A. tumefaciens* donor). The plasmid transfer experiment was accomplished by biparental mating (i.e., by excluding helper strain B001) for approximately 12 h, followed by selection on neomycin and tetracycline (to counterselect the donor).

2.1.3.3 Plant growth and nodulation.

Medicago species and cultivars used in this study are listed in **Table S5.2-3**. Scarified and surface-sterilized seeds were allowed to germinate in Petri plates, and 2-day-old seedlings were planted in sterile Turface clay particles (Turface Athletics, Buffalo Grove, IL, U.S.A.) and were allowed to grow for 4 days before being inoculated with 1 ml of *S. meliloti* cells suspended in 2.6 mM KH_2PO_4 (adjusted to pH 7.0 with 2N KOH) to an optical density at 600 nm (OD_{600}) of approximately 0.1. Nodulation was allowed to proceed for two to six weeks, depending on the experiment. Plants were watered with standard nodulation medium (SNM). SNM contains (per liter) 0.35 g of KH_2PO_4, 0.25 g of $MgSO_4$, 0.15 g of $CaCl_2 \cdot 2H_2O$, and 2 ml of minor salts solution (500× minor salts [per liter] = 9.5 g of Na_2-EDTA$\cdot 2H_2O$, 7 g of $FeSO_4 \cdot 7H_2O$, 1.5 g of $ZnSO_4 \cdot 7H_2O$, 1.5 g of H_3BO_3, 1.5 g of $MnSO_4 \cdot H_2O$, 0.15 g of $Na_2MoO_4 \cdot 2H_2O$, 15 mg of $CuSO_4$,

and 15 mg of $CoCl_2$). The pH of this solution was adjusted to 7.0 with 2N KOH, and the medium was sterilized by autoclaving. Plants were maintained at approximately 27°C under fluorescent lamps (2.7 klux intensity, 16 h day length). To determine dry weight of plants, samples were placed in an oven at 50°C for 48 h and were then weighed on a fine balance.

2.1.3.4 Microscopy.

At 14 days after inoculation, whole nodules were excised, were fixed with 2% glutaraldehyde in 50 mM cacodylate (pH 7.2) for 2 h, were washed, were postfixed in 1% osmium tetroxide for 2 h, and were washed again with H_2O. Samples were then stained in 0.5% uranyl acetate overnight, followed by dehydration in an acetone series. Samples were then embedded in Spurr's resin and were cured for 48 h at 70°C. Sections were made using an ultramicrotome. For light microscopy, 600-nm sections were bound to a microscope slide (by heating at 70°C), were stained with toluidine blue, and were imaged with an Axio Imager.A1 microscope. For transmission electron microscopy, 80-nm sections were stained in Reynold's lead citrate for 10 min, followed by washing with H_2O. Sections were imaged with a Tecnai G2 T-12 transmission electron microscope.

2.1.3.5 Modified Eckhardt gel electrophoresis.

Modified Eckhardt gels were performed as previously described (81) with modifications. Briefly, bacteria were grown to an OD_{600} of 0.6 in LB. Culture (150 µl) was added to 500 µl of chilled 0.3% sarkosyl in 1× SBE (20× SBE = 500 ml of H_2O, 4 g of NaOH, 3.72 g of Na_2-EDTA·$2H_2O$, pH to 8.0 with boric acid). Each sample was pelleted and resuspended in 20 µl of lysis solution (1× SBE, 10 mg of sucrose per milliliter, 1 mg of lysozyme per milliliter, 40 µg of RNase A per milliliter), followed immediately by loading into a sodium dodecyl sulfate (SDS)–SBE minigel (1× SBE, 0.8% agarose, 0.5% SDS). Each sample remained in the well for 2 min,

followed by electrophoresis at 23 V for 10 min, followed by electrophoresis at 96 V for 90 min. The minigel was then stained for 1 h in 0.4 μg of ethidium bromide per milliliter and was destained for 10 min in H_2O prior to imaging.

2.1.3.6 Competition experiments.

Strains employed were C241 (C017 modified with a *neo* marker in pHRC017) and C017-GOC. Newly saturated cultures were diluted to an OD_{600} of 0.05 in 2.6 mM KH_2PO_4 (pH 7.0), were mixed at a 1:1 ratio, and were inoculated into LB broth (5 μl of cells per 4-ml culture) or onto plants (1 ml of suspension per plant). LB cultures were grown for 72 h, followed by dilution plating onto LB-streptomycin and LB-streptomycin-neomycin. Plants were harvested after 21 days. To evaluate rhizosphere colonization, whole-root systems were vortexed for 30 s in 5 ml of LB-10% glycerol and were dilution-plated onto LB with streptomycin (LB-Sm) and LB with streptomycin and neomycin (LB-Sm-Nm). To evaluate root nodule occupancy, the roots were then surface-sterilized in 75% ethanol for 30 s and in 1.5% hypochlorite for 30 s and were then washed in H_2O. Surface-sterilized whole roots were macerated in LB and 10% glycerol and were dilution-plated onto LB-Sm and LB-Sm-Nm. The competition index (CI) was calculated according to (126) as follows. When C017-GOC yielded more colony-forming units (CFU) (LB and rhizosphere), CI = (C017-GOC recovered/C241 recovered)/(C017-GOC inoculated/C241 inoculated); when C241 yielded more CFU (whole roots), CI = (C241 recovered/C017-GOC recovered)/(C241 inoculated/C017-GOC inoculated).

2.1.3.7 Sequencing and annotation of pHRC017, and comparison with other HR plasmids.

Original evidence of the existence of pHRC017 came from Illumina-sequencing of strains C017 and C017-GOC. DNA was purified in a manner similar to that described below. Sequence reads from both samples were mapped to the *S. meliloti* Rm1021 genome sequence

using GNUMAP (127). DNA segments from strain C017 that were absent in strain Rm1021 and C017-GOC were analyzed by BLASTn.

For pyrosequencing of pHRC017, the plasmid was marked with an *oriT*/*neo* cassette and was transferred to plasmid-free *A. tumefaciens* UBAPF2 to yield strain C382. C382 was grown overnight in LB, and a total of 6 ml of saturated culture was pelleted and resuspended in 2.4 ml of TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). Cells were then treated with 0.6% SDS and 150 µg of proteinase K per milliliter and were incubated at 42°C for 20 min. The resulting lysate was brought to 0.7 M NaCl and 0.6 % cetyltrimethylammonium bromide, followed by vortexing and incubation at 65°C for 10 min. The sample was then extracted with phenol and chloroform. The aqueous phase was isolated and treated with 80 µg of RNase A per milliliter, was incubated at 37°C for 10 min, and was re-extracted with chloroform. The aqueous phase was precipitated with isopropanol, was suspended in TE buffer, and was further purified over a Nucleobond Midi column, according to manufacturer's instructions. The 454 library preparation was performed according to the rapid library preparation protocol, followed by sequencing on the 454 Genome Analyzer FLX. Assembly into contigs was performed using Newbler (version 2.5.3) after

Table 2.1-1. Strain–host combinations evaluated in this study. Symbiotic phenotypes were scored approximately 30 days post inoculation (dpi) according to nitrogen fixation phenotype.[a]

	***Medicago* Host Plant**[b]				
Strain	**LU**	**IT**	**PR**	**A17**	**A20**
B464	Fix^+	Fix^+	Fix^+	Fix^+	Fix^+
C017	Fix^+	Fix^+	Fix^- *	Fix^- *	Fix^- *
B469	Fix^+	Fix^+	Fix^- *	Fix^- *	Fix^- *
B800	Fix^+	Fix^+	Fix^-	Fix^+	Fix^- *
C377	Fix^+	Fix^+	Fix^- *	Fix^-	Fix^- *

[a]Fix^+, effective N-fixing pairs; Fix^-, abortively nodulating pairs; Fix^- *, pairs exhibiting suppressible incompatibility (the GOC phenomenon).
[b]LU, *M. lupulina*; IT, *M. italica*; PR, *M. praecox*; A17, *M. truncatula* cv. A17; A20, *M. truncatula* cv. A20.

subtracting reads that corresponded to *A. tumefaciens*. Contig edge ambiguities were resolved by PCR. The previously introduced *oriT*/*neo* cassette was then removed from the assembly to reconstitute the native pHRC017 sequence. Reads were remapped to this assembly, using Newbler (version 2.5.3) to verify the sequence. The final assembly was annotated using Glimmer (version 3.02), and predicted open reading frames were assigned putative functions after BLASTx analysis. This sequence can be accessed in GenBank (accession number JQ665880).

Genetic regions on pHRC017 were compared with the other three HR plasmids (pHRB469, pHRB800, and pHRC377) by PCR. Primers oMC73, oMC74, oMC81, oMC82, and oMC209 through oMC292 were used for this analysis. For most markers, primers were designed to amplify approximately 500-bp regions (**Figure 2.1-6**).

2.1.4 Results

2.1.4.1 Isolation of rhizobia with increased host range.

Using numerous *S. meliloti* isolates from the United States Department of Agriculture collection (122), we have evaluated symbiotic properties on several *Medicago* species, as well as multiple cultivars of the model plant *Medicago truncatula* (**Table 2.1-1.**; a more detailed data set may be found in **Table S5.2-4**). Approximately half of all host-strain pairs resulted in an

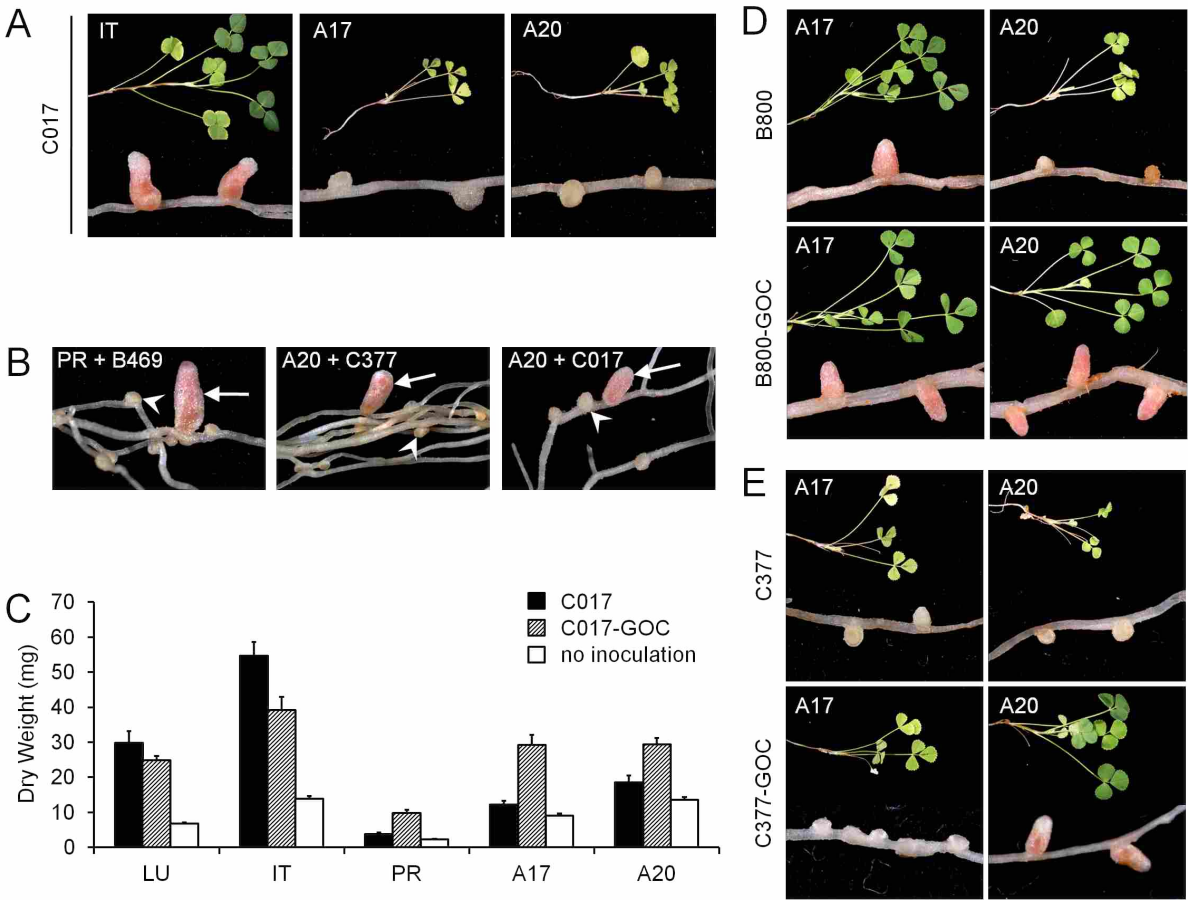

Figure 2.1-1. Host range restriction in *S. meliloti* is unstable. A. Representative shoot and nodule images from compatible and incompatible pairs, 30 dpi. **B.** Examples of GOC nodules (arrows) among incompatible nodules (arrowheads). **C.** Quantification of shoot dry mass 40 dpi with strain C017 (black; $n = 8$), C017-GOC (hatched; $n = 8$), or no inoculation (white; $n = 16$). Error bars represent SEM. **D.** Representative shoot and nodule images of B800 and B800-GOC inoculated onto *M. truncatula* cultivars A17 and A20, 30 dpi. Note that both strains are Fix^+ on A17, but only B800-GOC is Fix^+ on A20. (E) Representative shoot and nodule images of C377 and C377-GOC inoculated onto *M. truncatula* cultivars A17 and A20, 30 dpi. Note that both strains are Fix^- on A17 but only C377-GOC is Fix^+ on A20.

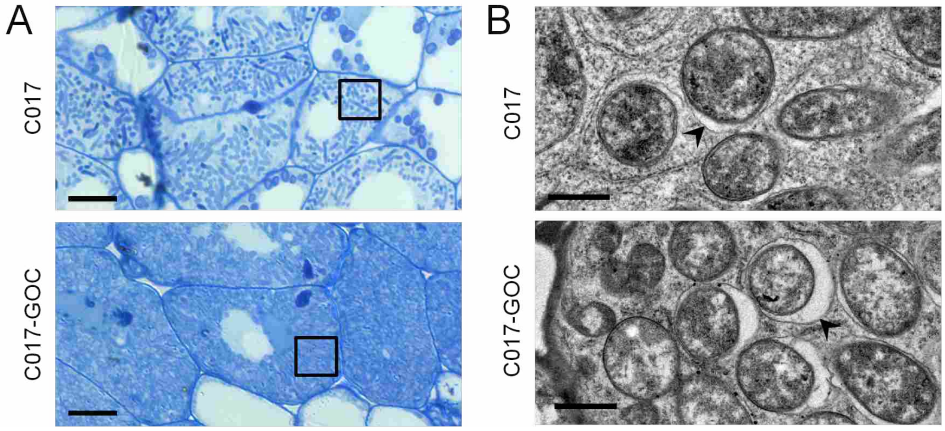

Figure 2.1-2. C017 and C017-GOC differ in their ability to occupy host plant cells. A. Light micrographs of A17 infected nodule cells, 14 dpi with strain C017 (top) or C017-GOC (bottom) from equivalent positions within the nodule. Inset boxes denote regions of nodule cell cytoplasm that are occupied by stained bacteroids. Scale bars: 10 µm. **B.** Transmission electron micrographs of nodules treated as in (**A**), showing bacteroid cross-sections. Arrowheads indicate plant-derived peribacteroid membranes. Scale bars: 1 µm.

ineffective symbiosis, evidenced by chlorotic shoots and small, white nodules (**Figure 2.1-1A**). In most cases, symbiotic incompatibility was host-conditioned, with strains exhibiting an effective N fixation phenotype (Fix^+) on some *Medicago* hosts and an ineffective phenotype (Fix^-) on others. This was apparent both for different species of *Medicago* (**Figure 2.1-1A**) and different cultivars of *M. truncatula* (**Figure 2.1-1D** and **E**). These strains, therefore, have the genetic capacity for N fixation, but that capacity is overridden by an unknown response by specific hosts that occurs after nodule initiation.

To better understand genetic mechanisms of host-conditioned incompatibility, several wild *S. meliloti* strains were screened on incompatible hosts in search of spontaneous mutants in which incompatibility was suppressed. These so-called gain-of-compatibility (GOC) mutants resulted in elongated, pink, N-fixing nodules among wild-type nodules that were small and white (**Figure 2.1-1B**). Rhizobia isolated from GOC nodules showed a stable Fix^+ phenotype after isolation and reinoculation (**Figure S5.1-1**). To confirm the host benefit brought about by GOC mutations, we measured shoot biomass from plants nodulated by *S. meliloti* C017 and its

derivative C017-GOC. C017-GOC provided a greater host benefit on plants with which the C017 parent was incompatible (**Figure 2.1-1C**, PR, A17, and A20). Conversely, on a host with which C017 was naturally compatible for N fixation (**Figure 2.1-1C**, LU and IT), the C017-GOC strain did not confer additional benefit and may even have slightly reduced biomass.

To assess how GOC derivatives affect nodule development on *M. truncatula* cv. A17, microscopic analysis was conducted on nodules that had been induced by either C017 (incompatible with A17) or C017-GOC (compatible with A17) (**Figure 2.1-2A** and **B**). Under both conditions, bacteria successfully invaded the developing nodule, followed by morphological differentiation into swelled, intracellular bacteroids. It was consistently observed that C017-GOC attained higher bacteroid densities in the nodule cell cytoplasm than C017. It was therefore concluded that C017 incompatibility on A17 occurs very late in the establishment of the symbiosis, after bacterial entry and after at least partial bacteroid differentiation within nodule cells.

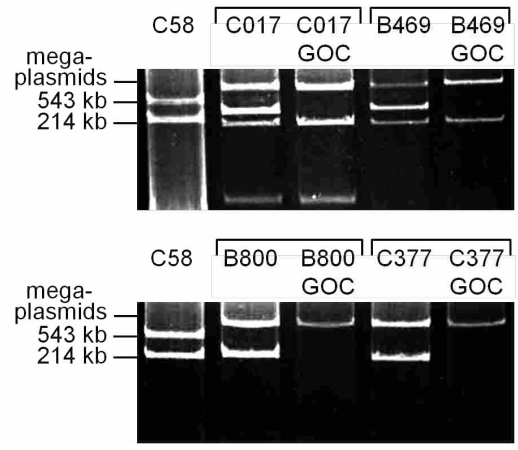

Figure 2.1-3. The GOC phenomenon is associated with the loss of an accessory plasmid. Eckhardt gels of parent strains and GOC derivatives. *A. tumefaciens* C58 plasmids (543 kb and 214 kb) are included as a molecular weight standards; *S. meliloti* symbiotic megaplasmids (>1 Mb) are indicated.

2.1.4.2 Large accessory plasmids can dictate symbiotic host range.

In hopes of finding a mutation responsible for the GOC phenotype, short-read, shotgun sequencing of strains C017 and C017-GOC was performed. Upon mapping reads to the reference Rm1021 genome using the GNUMAP algorithm (127), we discovered a significant amount of DNA in C017 that is absent in C017-GOC. Some of this DNA is

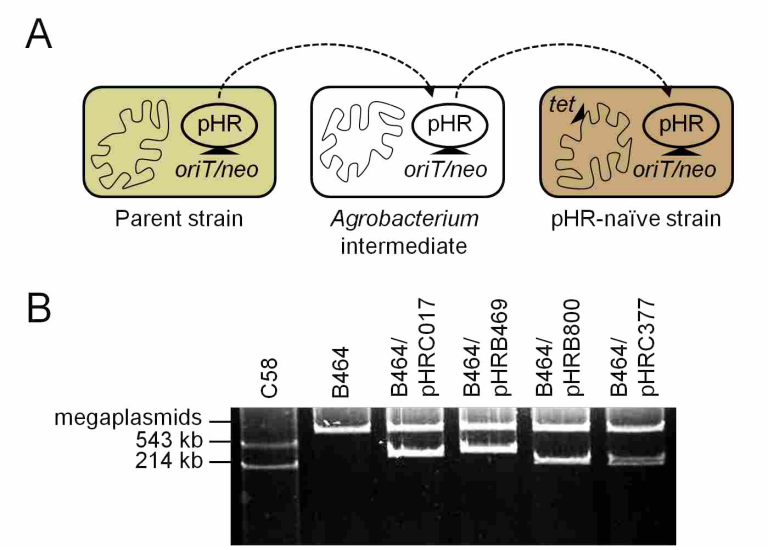

Figure 2.1-4. Host range plasmids can be transferred into different strain backgrounds. A. Diagram of the process used to introduce exogenous HR plasmids into plasmid-naïve strains. **B.** Eckhardt gels showing B464 harboring different exogenous HR plasmids. All HR plasmids are modified with the *oriT*/*neo* cassette to permit transfer and selection. All strains are modified with the *tet* cassette to permit counterselection.

similar to a known *Sinorhizobium* accessory plasmid, pSMeSM11a (92). Subsequent Eckhardt gel analysis of all of our GOC derivatives and corresponding parent strains revealed that the GOC phenomenon correlates with the curing of accessory plasmids in the size range of 200 to 300 kb (**Figure 2.1-3**). We generally refer to these as HR plasmids due to their effect on symbiotic host range (defined here as the range of plants benefitting from the symbiosis). Specific HR plasmid names assigned in this study (pHRC017, pHRB469, pHRB800, and pHRC377) are in reference to the parent strains in which they were initially discovered.

To confirm the role of HR plasmids in determining host compatibility, plasmid transfer experiments were performed. These experiments required modifying the plasmids with a genetic cassette containing a conjugative transfer origin (*oriT*) and a neomycin-resistance gene (*neo*). Recipient strains were modified with a tetracycline-resistance gene (*tet*) inserted at a neutral chromosomal location. This plasmid transfer procedure is diagrammed in **Figure 2.1-4A** and described below. In one experiment, pHRB469 was transferred from B469 into plasmid-free

Table 2.1-2. Symbiotic outcomes on multiple hosts after inoculation with B464 harboring different HR plasmids. Symbiotic phenotypes were scored approximately 30 days post inoculation (dpi) according to nitrogen fixation phenotype.[a]

	Medicago Host Plant				
Strain	**LU**	**IT**	**PR**	**A17**	**A20**
B464[b]	Fix^+	Fix^+	Fix^+	Fix^+	Fix^+
B464/pHRC017[c]	Fix^+	Fix^+	Fix^-	Fix^-	Fix^-
B464/pHRB469	Fix^+	Fix^+	Fix^-	Fix^-	Fix^-
B464/pHRB800	Fix^+	Fix^+	Fix^-	Fix^-	Fix^-
B464/pHRC377	Fix^+	Fix^+	Fix^+	Fix^-	Fix^-

[a] Fix^+, effective N-fixing pairs; Fix^-, abortively nodulating pairs; Fix^- *, pairs exhibiting suppressible incompatibility (the GOC phenomenon).
[b] All strains referred to are modified with the *tet* cassette to permit counterselection.
[c] All HR plasmids referred to are modified with the *oriT*/*neo* cassette to permit transfer and selection.

Agrobacterium tumefaciens UBAPF2 (125) and, subsequently, into strain B469-GOC*tet*. The resulting transconjugant showed restored incompatibility on host plants with which B469 is incompatible (**Figure S5.1-2**).

We next considered the question of whether HR plasmids could be accommodated in a *S. meliloti* strain (B464) that is naturally devoid of accessory plasmids and is an effective N fixer on all of the *Medicago* hosts in our panel (**Table 2.1-1**). After mobilization to the *A. tumefaciens* intermediate, all four HR plasmids were successfully transferred into B464*tet* (**Figure 2.1-4B**). The HR plasmids conferred upon B464*tet* nearly the same host range restriction as they did to their parental *S. meliloti* strains (**Figure 2.1-5A** and **B**, **Table 2.1-2**). One exception was that of pHRB800, which did not restrict N fixation on *M. truncatula* A17 in parental strain B800 (Fig. 1E) but did in strain B464*tet* (**Table 2.1-2**). This experiment shows that HR plasmids can restrict host range in a manner that is dominantly acting and largely (but not completely) independent of the genotype of the strain in which it resides. Further testing with all four HR plasmids in a variety of strain backgrounds revealed several other instances in which strain background can modulate the ultimate symbiotic consequence of harboring an HR plasmid (**Table S5.2-5**).

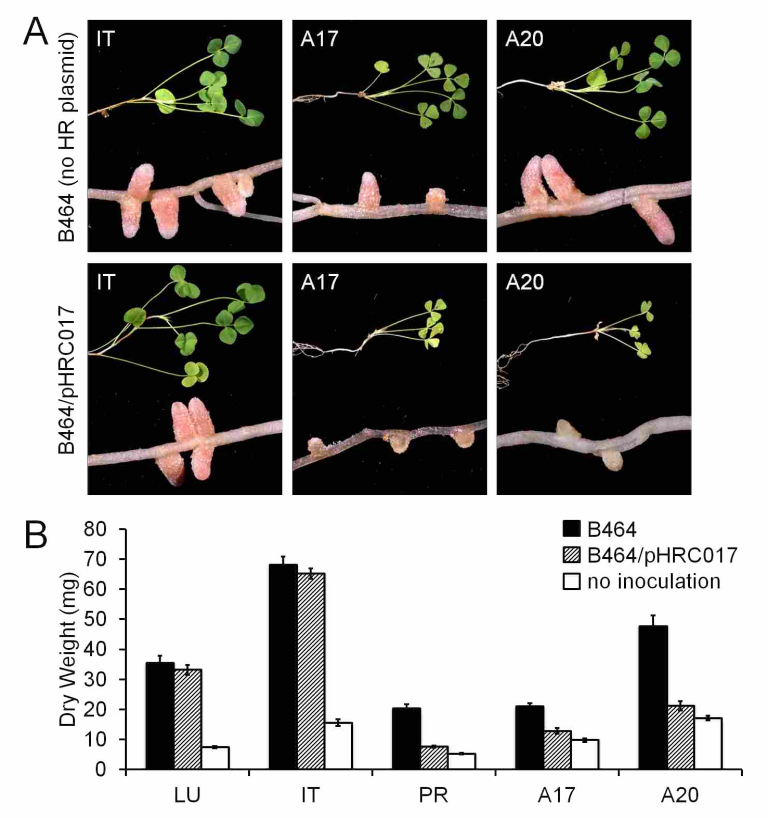

Figure 2.1-5. Abortive nodulation is governed by autonomously functioning accessory plasmids. A. Representative shoot and nodule phenotypes from plants inoculated with B464 with or without pHRC017 (30 dpi). **B.** Quantification of shoot dry mass 40 dpi with strain B464 (black; n = 8), B464/pHRC017 (hatched; n = 8), or no inoculation (white; n = 16). Error bars represent SEM. All HR plasmids referred to are modified with the *oriT*/*neo* cassette to permit transfer and selection. All strains are modified with the *tet* cassette to permit counterselection.

2.1.4.3 HR plasmids can confer a competitive advantage for nodulation.

Because HR plasmids cause an abortive nodulation phenotype on certain *Medicago* hosts, we sought to better understand how these plasmids might provide a fitness advantage for bacterial cells in which they reside, even if they interfere with nodule development and N fixation. Supposing that HR plasmids might provide some growth advantage to free-living cells in the soil or rhizosphere (128), competition experiments were performed using mixed populations of strains C017 and C017-GOC. First, the competition index was determined for a C017 and C017-GOC mixed population grown for approximately 15 generations in lysogeny

broth (LB). In this experiment, the GOC strain increased its representation in the population by approximately tenfold (**Figure 2.1-6A**). This can be accounted for by the fact that C017-GOC has a faster doubling time than C017 in LB (4.4 versus 5.5 h). Next, mixed populations were grown on plants for three weeks. These bacteria were recovered both from the root surface (rhizosphere) and from within the nodules of surface-sterilized roots. In order to sample as many nodules as possible, tests for nodule occupancy were performed on a whole-plant basis. While pHRC017 imposed a modest disadvantage in the rhizosphere, it reproducibly provided a competitive advantage for nodule occupancy (**Figure 2.1-6A**). Similar results were seen for competition between strains B469 and B469-GOC (data not shown).

We have considered how it is possible that an incompatible strain can achieve a larger population than its effective GOC counterpart, which induces larger and more densely colonized nodules. In one model, the plasmid-harboring incompatible strain may be more aggressive at stimulating nodule initiation, even though the resulting nodules are less beneficial to the host. In a second model, the plasmid may permit higher recoverability of rhizobia from nodules, even if the number of nodules induced by each strain in competition is equivalent. These models are not mutually exclusive. The colony counts reported in **Figure 2.1-6A** reflect tallies of white and pink nodules on plants from the competition experiments, suggesting that the plasmid-harboring C017 strain is more competitive at the level of nodule establishment. To investigate specific molecular mechanisms that could be responsible for HR plasmid control over host range determination and nodulation competitiveness, pHRC017 was sequenced. This was done by transferring pHRC017 into *A. tumefaciens* UBAPF2 and isolating total DNA, subjecting it to 454 sequencing. Sequence reads corresponding to *A. tumefaciens* were subsequently subtracted, and the remaining reads were assembled into a single circular molecule of 298,356 bp. An annotated record for pHRC017

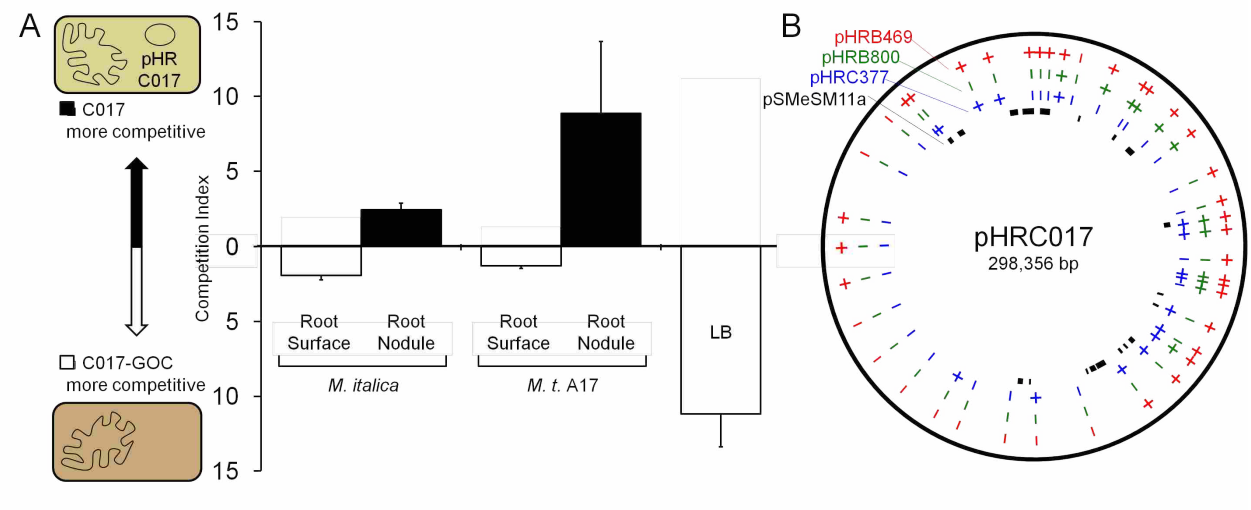

Figure 2.1-6. pHRC017 confers a competitive advantage for nodule occupancy and is genetically similar to other HR plasmids. A, The competitive indexes for C017 and C017-GOC grown in competition when inoculated onto a compatible host (*M. italica*; $n = 3$), onto an incompatible host (*M. t.* A17; $n = 6$), or into LB ($n = 6$). Error bars represent SEM. The HR plasmid in C017 was modified to contain the selectable *neo* gene, as described in Supporting Online Material. B, Diagram and annotation of pHRC017. Conservation of sequence in other HR plasmids (based on PCR tests) is indicated by + or –. Regions with BLASTn similarity to pSMeSM11a are indicated by black bars.

has been submitted to GenBank (accession number JQ665880). Among the approximately 430 potential open reading frames on this plasmid are several genes that may encode proteins of symbiotic consequence, such as a probable 1-aminocyclopropane-1-carboxylate (ACC) deaminase. ACC is the direct precursor to the phytohormone ethylene, and bacterial ACC deaminases are known to influence plant-microbe interactions (80, 129). Additionally, a type IV secretion system (T4SS) operon, which may be involved in horizontal plasmid transfer, is encoded on pHRC017. In tests to determine if HR plasmids are self-transmissible, plasmid-harboring strains were mixed with plasmid-naive strains. For pHRB469, clear evidence of transfer was observed (**Figure S5.1-3**) but transfer of pHRC017 was not. Transfer of the other HR plasmids was not tested.

Due to its similarity with pSMeSM11a, we performed BLAST analysis of pHRC017 in comparison with other sequenced *Sinorhizobium* accessory plasmids and observed only limited similarities (**Figure S5.1-4**). The plasmid pSINME01 from *S. meliloti* AK83 (16) exhibited the

highest similarity, matching approximately 25% of pHRC017, with regions of similarity being distributed throughout the pHRC017 sequence. We used polymerase chain reaction (PCR)-based markers to compare pHRC017 with the other HR plasmids, which have not yet been sequenced (**Figure 2.1-6B**). Of 41 markers, 30 markers amplify from pHRB469, 12 amplify from pHRB800, and 15 amplify from pHRC377. Six of the 41 pHRC017 markers amplify from all four HR plasmids (**Figure 2.1-6B**; additional details in **Table S5.2-6**). This analysis confirms that the four HR plasmids are genetically related but that a limited number of gene clusters are conserved in all four plasmids. How these genes contribute to host-conditioned compatibility and competition for nodule invasion awaits further characterization.

2.1.5 Discussion

We have presented the discovery and characterization of accessory plasmids in S. meliloti that block effective N fixation in a host-conditioned manner. Rhizobial symbiotic host range control is typically associated with Nod-factor signaling, which occurs at the earliest stages of nodule development. The HR plasmids described here appear to influence host range at a much later stage of symbiotic infection, and they act dominantly to limit host range. The idea that host range in the *Sinorhizobium–Medicago* symbiosis can be determined after Nod-factor perception is well documented (39, 112, 115), but genetic mechanisms for this phenomenon have been difficult to resolve. In various rhizobial genera, correlations between accessory plasmids and nodulation competitiveness (84, 93) as well as connections between accessory plasmids and reduced host benefit (97–99) have been reported. Our work highlights the autonomy with which HR plasmids can function in multiple strains, their effect on competitiveness for nodule invasion, and the host-conditioned nature of the phenomenon.

Our observations suggest that HR plasmids may control the synthesis of a signal that somehow disrupts the symbiotic dialogue. Whether this disruptive signal elicits a host defense response or merely prevents a required host response is unclear. This model is reminiscent of observations first made in the soybean–rhizobium symbiosis, in which type III effectors (T3E) elicit a negative response from certain soybean cultivars, yet potentiate symbiosis with others (130–132). More recently, the soybean response to T3E has been linked to specific leucine-rich repeat resistance gene alleles, demonstrating that rules for rhizobium–legume compatibility are governed by pathways that are typically attributed to pathogen immunity (133). Unlike T3E-mediated incompatibility, which is evident at a very early stage of symbiotic infection (133), HR

plasmid–mediated incompatibility blocks the symbiotic program at a very late stage, so that incompatible nodules can still harbor significant rhizobial populations.

As an alternative to the model described above, which involves the synthesis of a new compatibility-determining signal, HR plasmids may alter the normal synthesis of a signal encoded elsewhere in the rhizobial genome. In a previous study, it was suggested that differential succinylation of exopolysaccharide I (EPS I) could explain *S. meliloti* compatibility with *M. truncatula* cultivars A17 and A20 (39). These *M. truncatula* cultivars were also used in our study. With these observations in mind, one can envision a model in which HR plasmids cause the modification of some common signal such as EPS I or lipopolysaccharide. Such a modification could give rise to the restricted host ranges reported here. It is notable that pHRB800 controls host range at the cultivar level in *M. truncatula* (**Figure 2.1-1D**) in a manner similar to the differentially succinylating strains reported previously.

The evolutionary maintenance of HR plasmids is interesting to consider, since these plasmids appear to skew the exchange of benefits between host and symbiont, essentially converting a mutualistic relationship to a parasitic one. HR plasmids confer what may be thought of as a 'cheater' phenotype on *S. meliloti* cells that harbor them. If we simplistically assume that rhizobial populations possess some relatively stable balance of beneficial and nonbeneficial strains, then our observations relating to HR plasmids provide possible insights into mechanisms that maintain this balance. First, we observed that the negative effect of HR plasmids is host-conditioned, meaning that negative effects of cheating can be masked (**Table 2.1-1**). Second, we found that HR plasmids may erode host benefit in some strain backgrounds but not in others; that is, HR plasmids function in a 'strain-conditioned' manner (**Table 2.1-2** and **Table S5.2-5**). Strains are not dedicated to a specific lifestyle; they may either gain or lose exploitative

properties in single steps. Finally, we have shown that HR plasmids couple cheating functions with functions that increase competitiveness for nodule occupancy (**Figure 2.1-6A**). In this way, exploitative genes keep a foothold in spite of the fitness disadvantage brought about by the cost incurred to the host. These mechanisms would tend to support the maintenance of exploitative functions, even under conditions that would select for mutualism. However, we do not dismiss the possibility that selection for HR plasmid maintenance may be largely driven by factors unrelated to the symbiosis.

This study points to a potential scenario in which commercial rhizobium inoculants that have been optimized for nitrogen fixation might be outcompeted by indigenous rhizobia or even neutralized by horizontal transfer of HR plasmids from coexisting native rhizobia. There is some evidence that HR plasmids may be naturally transmissible, based on the presence of T4SS genes in at least two of them (pHRC017 and pHRB469) and the observed transfer of pHRB469 from one strain to another. Future work should be dedicated to better defining the ecological and agricultural significance of HR plasmids and further characterizing molecular mechanisms of late-stage incompatibility brought about by these plasmids. We are currently developing tools for creating large-scale deletions in HR plasmids, which will assist in the genetic dissection of plasmid-mediated host range determination.

2.1.6 Unpublished Data

To identify the plasmid-encoded gene or genes responsible for limiting host range, we constructed a mini-Tn*5* transposon carrying a *neo* cassette and an *oriT* (pJG310). We mutagenized pHRC017, pHRB469, or pHRB800 in an *A. tumefaciens* background (UBAPF2) and then mobilized them back into *S. meliloti*. This would ensure that only the plasmid was mutagenized, not the *S. meliloti* genome. However, despite repeated attempts we failed to recover Tn*5* insertions in any of the host range plasmids as evidenced by the fact that we never observed transfer of neomycin resistance to *S. meliloti*. When we mutagenized *S. meliloti* strains C017 and B469 with the same *oriT*/*neo*-bearing Tn*5*, and mated out to *A. tumefaciens* UBAPF2, we observed frequent insertion and subsequent transfer of the non-HR accessory plasmids (pSmeN6B7a, pSmeN6B7b, and pSme74B17a) but never of pHRC017 (=pSmeN6B7c) or pHRB469 (=pSme74B17b) (data not shown). This suggests that the host range plasmids are able to prevent Tn*5* insertions in *cis* but not in *trans*. The mechanism underlying this phenomenon is intriguing but has not yet been investigated.

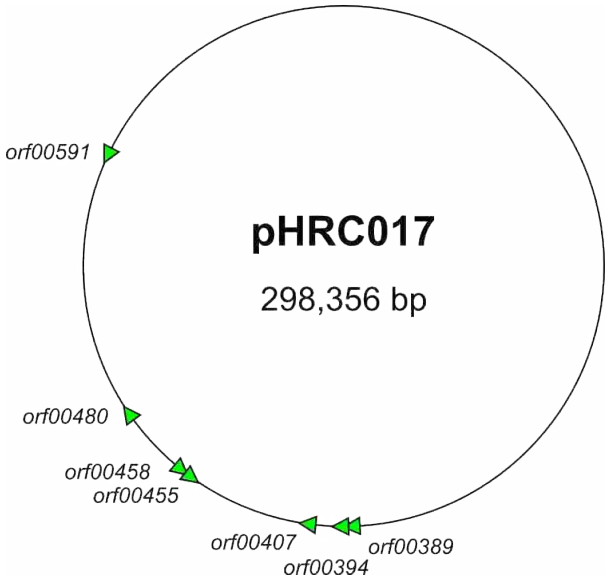

Figure 2.1-7. pHRC017 encodes seven different metalloaminopeptidases.

A post-doctoral fellow in the laboratory, Skip Price, then attempted the same procedure using an *oriT*-modified *himar1* transposon. Mutagenesis of pHRB800 and subsequent testing on *M. truncatula* cv. A20 yielded several insertions which had abolished the ability of pHRB800 to limit host range. These insertions mapped to a metallopeptidase, which is currently being further

characterized. With respect to the other host range plasmids, seven metallopeptidases are encoded on pHRC017 (*orf00389*, *orf00394*, *orf00407*, *orf00455*, *orf00458*, *orf00480*, and *orf00591*), but none of them share significant homology with the pHRB800 metallopeptidase (**Figure 2.1-7**).

Mutagenesis of the remaining host range plasmids (pHRC017, pHRB469, and pHRC377) yielded two interesting candidates. First, all three returned multiple insertions in an unusual *repAC* operon. Three types of replication systems have been described for rhizobial plasmids (134) (**Figure 2.1-8**). The most common is the *repABC* family (**Figure 2.1-8A**), which is found on many Sym plasmids. The *repC* gene is involved in replication and the *repA* and *repB* genes are responsible for plasmid partitioning. The *repC* gene is regulated by a counter-transcribed RNA, *incA* (also called *inc*α or *repE*). Less common is the *repAC* family (**Figure 2.1-8B**), which has a tightly-linked *repA* and *repC* with no counter-transcribed RNA. The third system is the ct-RNA–*repC* family (**Figure 2.1-8C**), which has a *repC* unlinked to any partitioning proteins but under the control of a counter-transcribed RNA.

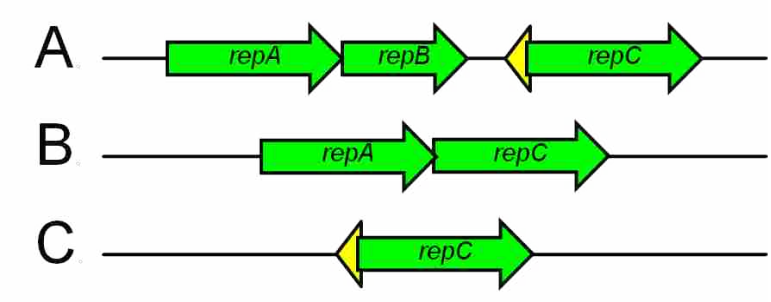

Figure 2.1-8. The three families of known rhizobial plasmid replication systems. A. The *repABC* family (which is regulated by a ct-RNA, yellow), **B.** the *repAC* family, and **C.** the ct-RNA–*repC* family. Based on (134).

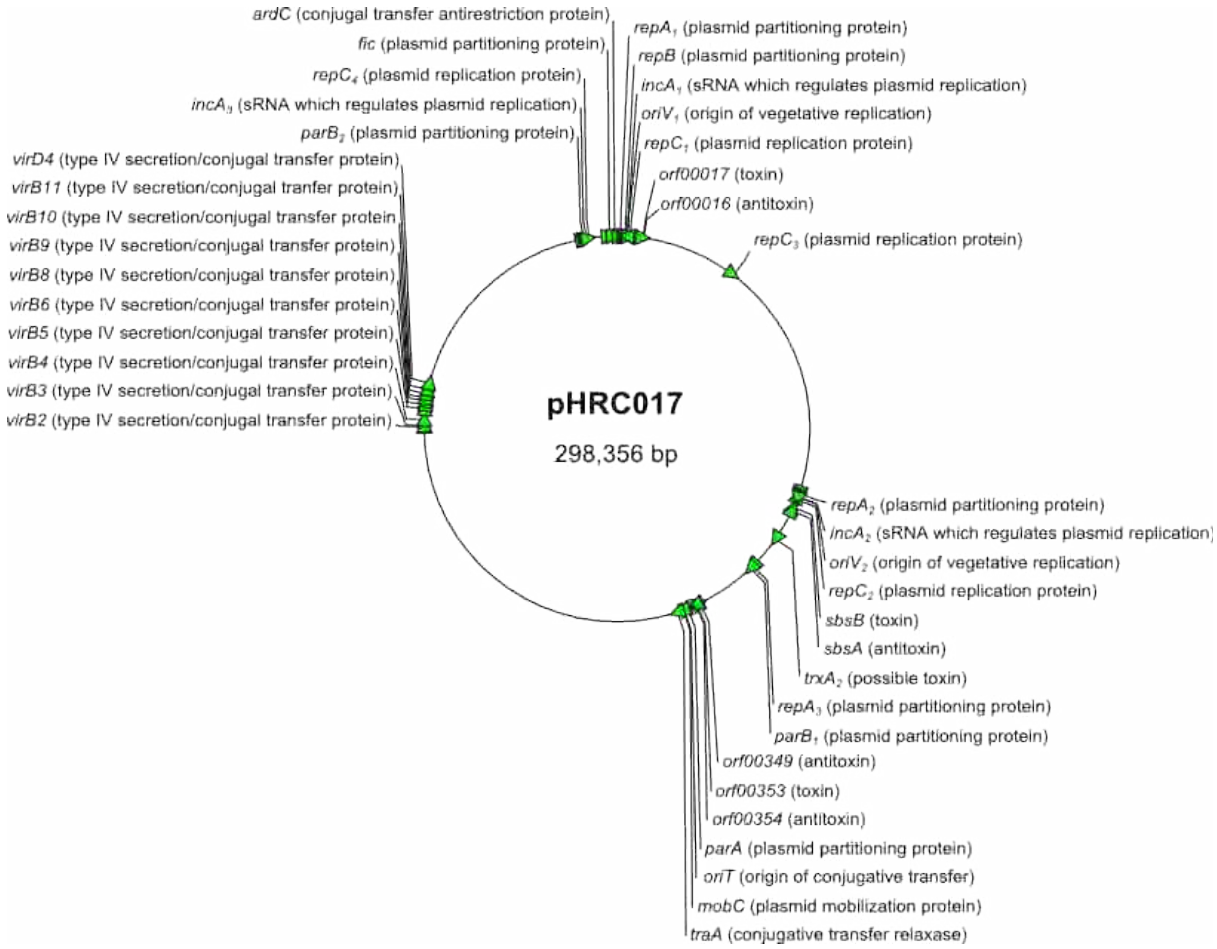

Figure 2.1-9. pHRC017 encodes several systems for replication, stability, and transfer. Indicated are four replication systems (*repA*$_1$*BC*$_1$, *repA*$_2$*C*$_2$, *repC*$_3$, and *parB*$_2$–*repC*$_4$ with a nearby *fic* and *ardC*), a set of partitioning genes unlinked to a replication system (*repA*$_3$–*parB*$_1$), four sets of toxin–antitoxin genes (*orf00016*–*orf00017*, *sbsAB*, *trxA*$_2$, and *orf000349*–*orf00353*–*orf00354*), and two transfer systems (an *oriT*-associated *mobC*–*traA* and a *virB2*–*virD4* type IV secretion system).

The sequenced plasmid pHRC017 encodes four replication loci (**Figure 2.1-9**). The first, *repA*$_1$*BC*$_1$, encodes a canonical *repABC* family replication system (**Figure 2.1-8A**). The second, *repA*$_2$*C*$_2$, is unlike previously described rhizobial plasmid replication systems in that it has a truncated, possibly pseudogenetic *repA* gene followed by an *incA*-regulated *repC* gene. The third, *repC*$_3$, only encodes a 189-bp fragment from the 3′ end of a *repC* gene and likely is nonfunctional. The last, *repC*$_4$, is an *incA*-regulated, GANTC-rich replication protein that is related to the replication system of the octopine-type pTi plasmids of *Agrobacterium tumefaciens*

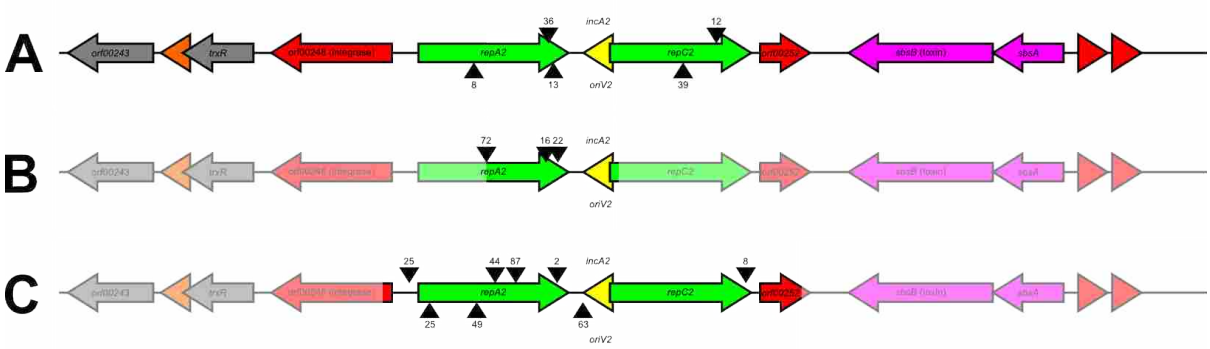

Figure 2.1-10. *himar1* mutagenesis for loss of host restriction returned multiple insertions in an unusal shared *rep* locus. A. *himar1* insertion sites in $repA_2C_2$ of pHRC017; **B.** *himar1* insertion sites in a similar locus of pHRB469; **C.** *himar1* insertion sites in a similar locus of pHRC377.

(135). It contains several nonsense and/or frameshift mutations and thus is a pseudogene. Just upstream is a truncated partitioning protein ($parB_2$) and ~4 kilobases downstream is a *fic* plasmid partitioning protein and an *ardC* conjugal transfer anti-restriction protein.

The majority of *himar1* insertions in HR-plasmids leading to expansion of host range mapped to the $repA_2C_2$ genes of pHRC017 and to similar loci in pHRB469 and pHRC377 (**Figure 2.1-10**). Insertions occurred in both genes but *repA* insertions may simply have a polar effect on *repC*. Besides being a unique replication system in a rhizobial plasmid, $repA_2C_2$ is also unusual in that the *repA* is phylogenetically related to *repA* genes from other accessory plasmids of *S. meliloti*, but the *repC* is more similar to the *repC* gene of the pSymA megaplasmid than it is to the *repC* genes found on other accessory plasmids of *S. meliloti* (**Figure 2.1-11**). Since ~24 copies of the *S. meliloti* genome are found in bacteroids of *M. truncatula* (59), we postulated a role for $repC_2$ in disregulating endoreduplication of pSymA. If too few copies of pSymA were produced, some necessary symbiotic signal might be insufficiently expressed, leading to abortion of the symbiosis; if too many copies of pSymA were produced, some offensive signal might be produced in excess, leading to an immune response from the plant. Alternatively, disruption of $repC_2$ may allow stable replication of the accessory plasmid in free-living bacteria but lead to an

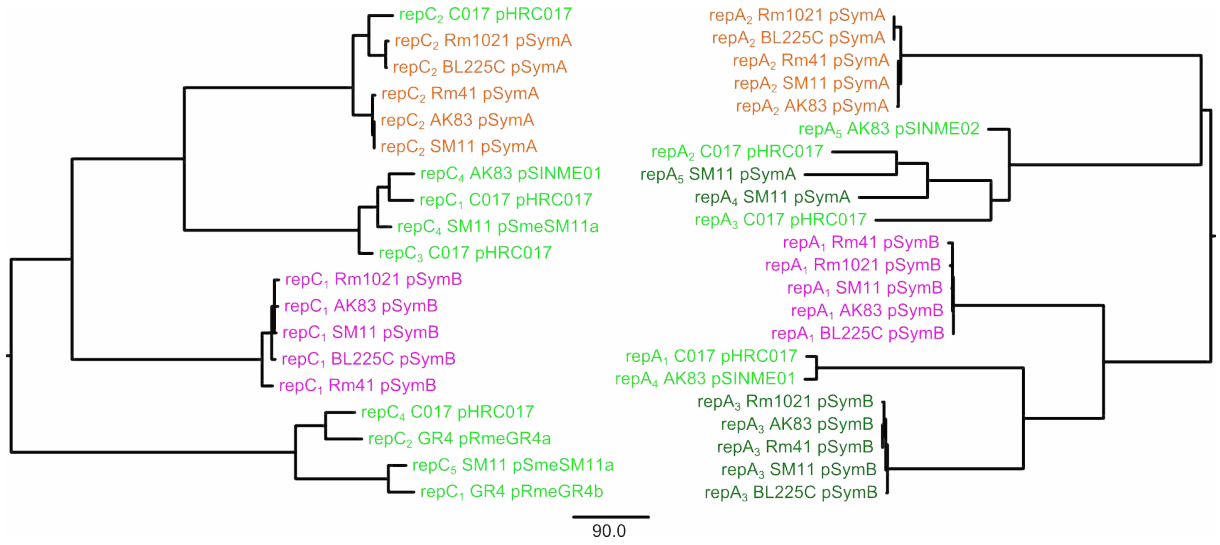

Figure 2.1-11. A comparison of phylogenetic trees of the *repC* genes and *repA* genes of *S. meliloti* megaplasmids and accessory plasmids. The *repC* tree is shown on the left and the *repA* tree is shown on the right. The primary *rep* genes of pSymB are shown in purple, the primary *rep* genes of pSymA are shown in orange, and the *rep* genes of accessory plasmids are shown in green.. Secondary *repA* genes found on pSymA and pSymB are shown in dark green.

increased rate of plasmid curing *in planta*. When these mutants were retested to confirm their phenotype, they produced both white and pink nodules suggesting that the *himar1* insertions in $repA_2C_2$ somehow accelerate plasmid curing *in planta* rather than providing a *bona fide* knockout of host range restriction.

The second gene of interest was identified in the *himar1* mutagenesis of pHRB469 and pHRC377: a LuxR-family transcriptional regulator similar to *orf00636* of pHRC017. It is predicted to have an N-terminal auto-inducer binding domain and a C-terminal helix-turn-helix DNA-binding domain. It is structurally related to the *E. coli* transcription factor, *sdiA*, which enhances cell division in response to N-acyl homoserine lactones (AHLs) (136); the *Salmonella enteric* transcription factor, *sdiA*, which detects auto-inducers of other species (137); the *traR* of *A. tumefaciens*, which increases replication and conjugal transfer of the pTi plasmid (138); and the *expR* of *Sinorhizobium meliloti*, which regulates EPS II biosynthesis (43, 140).

Further analysis of the pHRB800 metallopeptidase, the unusual *RepA*$_2$*C*$_2$, and the LuxR-family transcriptional regulator and confirmation of their role in restricting host range are ongoing.

2.1.7 Future Directions

In addition to the ongoing effort to identify host restriction genes on pHRC017, pHRB469, pHRB800, and pHRC377, there are several other outstanding questions. The majority of strains in our collection possess at least one accessory plasmid (data not shown) and most exhibit host range restriction to some degree (**Table S5.2-4**). But most strains, when tested, do not produce GOCs (data not shown). This could be because the host restriction is unrelated to the presence of accessory plasmids in that strain or because the plasmids are more stable. These strains could be intentionally cured of their accessory plasmids and then tested for altered host range. Alternatively, their accessory plasmids could be mobilized into an *S. meliloti* strain which is naturally devoid of accessory plasmids, such as B464, and then tested for their ability to restrict host range.

Accessory plasmids which affect host range have also been described for *Mesorhizobium loti* (97) and *R. leguminosarum* bv. *viciae* (98). But it is unknown how dependent the host restriction produced by these plasmids is on the strain–host pair. Our data suggest that for some strains of *S. meliloti* symbiosis is unaffected by the possession of certain HR plasmids (**Table S5.2-5**). However, it is unknown whether other species or genera of rhizobia would be affected by the HR plasmids of *S. meliloti*. To test this, HR plasmids can be mobilized into other rhizobia and inoculated onto a variety of hosts to see if they restrict host range in a heterologous system.

To determine bacterial genes which mask the deleterious effects of HR plasmids on symbiosis (**Table S5.2-5**), wildtype plasmid-free and GOC strains can be mutagenized with Tn*5*-267, a Tc^R-bearing transposon, prior to introduction (or re-introduction) of unmutagenized HR plasmids from *A. tumefaciens*. Because symbiosis genes in the *S. meliloti* strains could (and would) be disrupted, candidates with restricted host ranges would have to be cured of their HR

plasmids and retested to confirm that the phenotype is host range restriction and not loss of symbiosis independent of the presence of the HR plasmid.

2.2 Dissection of Rhizobial Plasmids Using *cre–loxP*

2.2.1 Summary

We report the development of a set of plasmid tools which utilize the DNA recombination properties of the *cre–LoxP* system to make targeted, large-scale deletions. These tools have been successfully used to make deletions up to 150 kilobases in accessory plasmids of *Sinorhizobium meliloti*, the microbial symbiont of alfalfa, but we expect that larger deletions are possible. These tools are also being employed to cure the laboratory strain, *Sinorhizobium meliloti* Rm1021 of its two megaplasmids (pSymB and pSymA), which will entail the loss of nearly half of its genetic information. To do so required additional genetic information concerning the location of any essential genes and toxin–antitoxin pairs. Steps have also been taken to prepare the Sym plasmids from two strains of *Sinorhizobium fredii*, both of which exhibit broad host ranges, into *S. meliloti* Rm1021 once it has been 'reset' by curing of its native Sym plasmids.

2.2.2 Introduction

Our lab has previously considered whether already-known symbiotic genes could be playing a role later in the symbiosis that would mediate late-stage incompatibility (141). To address this possibility, we developed a genetic tool, based on *cre–loxP* recombination, that would allow us to delete or upregulate a gene at a time of our choosing. Because of the difficulty presented by Tn*5* mutagenesis of HR plasmids, we adapted this *cre–loxP* system to generate new tools for making targeted, large-scale deletions. Large-scale deletions have previously been performed in *S. meliloti* using *sacB*-mediated homologous recombination between IS*50* elements of Tn*5* transposons (142) and the FRT/Flp system (143, 144). Once host range plasmids with larger deletions have been tested for an altered phenotype (e.g. loss of host range restriction), iteratively smaller deletions could be made within the original deletion frame until an exact gene could be pinpointed.

2.2.3 Materials and Methods

2.2.3.1 Bacterial culture.

Escherichia coli and *S. meliloti* cultures were grown at 37 and 30°C, respectively, in lysogeny broth (LB). *S. fredii* cultures were grown at 30°C in tryptone–yeast extract (TY) broth supplemented with appropriate antibiotics as follows (in micrograms per milliliter): chloramphenicol, 30; kanamycin, 30; neomycin, 100; rifampicin, 100; streptomycin, 200; and tetracycline, 5.

2.2.3.2 Plasmid and strain construction.

Plasmids and strains used in this study are listed in **Table S5.2-7**. Plasmids were constructed using standard techniques with enzymes purchased from New England Biolabs (Ipswich, MA, U.S.A.) The high-fidelity polymerase *Pfx*50 (Invitrogen, Carlsbad, CA, U.S.A.) was used for insert amplification. All custom oligonucleotides were purchased from Invitrogen and are listed in **Table S5.2-8**. Mobilization of plasmids was accomplished by triparental mating

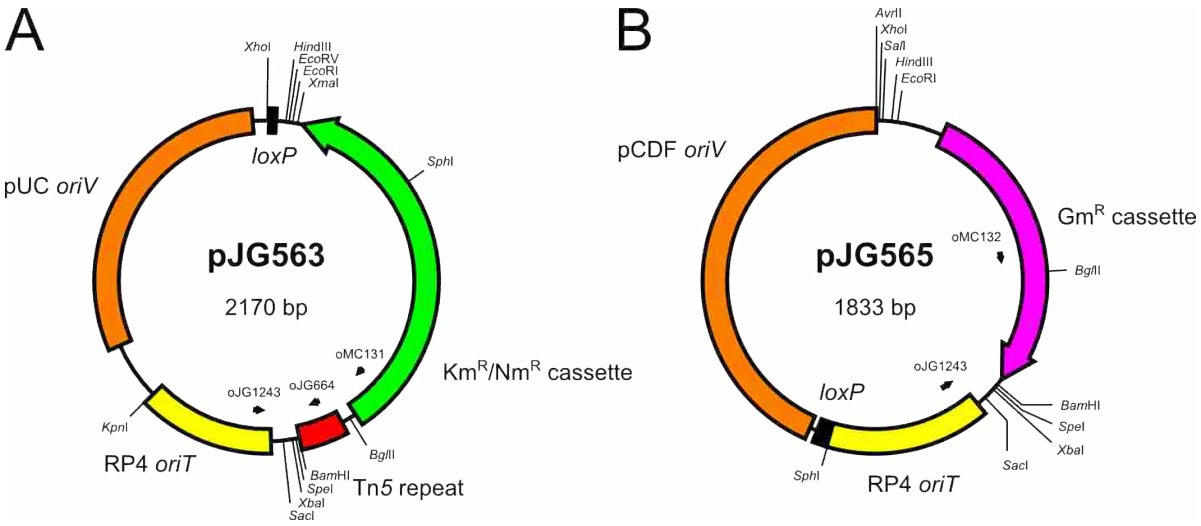

Figure 2.2-1. The upstream (pJG563) and downstream (pJG565) *loxP* integration plasmids. A. pJG563 consists of a pUC *oriV*, an RP4 *oriT*, a *kan*/*neo* cassette, a *loxP* site, one end of a Tn*5* transposon, and two polylinkers. The positions of useful primers (oJG664, oJG1243, and oMC131) are indicated. **B.** pJG565 consists of a pCDF *oriV*, an RP4 *oriT*, a gentamicin resistance cassette, a *loxP* site, and two polylinkers. The positions of useful primers (oJG1243 and oMC132) are indicated.

with helper *E. coli* B001 (DH5α harboring plasmid pRK600). pRK600 expresses trans-acting proteins required for mobilization of plasmids harboring the RK2 transfer origin (*oriT*). The upstream integration plasmid was constructed by annealing primers oMC119 and oMC120 together to produce the *loxP* sequence with XhoI- and HindIII-compatible overhangs. This product was then ligated into pJG194 which had been digested with XhoI and HindIII to produce plasmid pJG563 (**Figure 2.2-1A**). The *loxP* insert was sequence verified using primers oMC121 and oMC122. To construct the downstream integration plasmid, we amplified the gentamicin resistance cassette and RP4 *oriT* from pJG520 (an unpublished plasmid used in our lab) using primers oMC123 and oMC124 and the pCDF *oriV* from pAG101 (Novagen) using primers oMC125 (which included the *loxP* sequence) and oMC126. The two PCR products were then digested with AvrII and SphI and ligated together to yield pJG565 (**Figure 2.2-1B**). The ligation junctions were sequence verified using primers oMC127 through oMC130. Derivatives of pJG563 and pJG565 targeted to specific loci were constructed by amplifying approximately 300-bp fragments from the intended insertion site. The target fragment and either pJG563 or pJG565 were digested with BamHI and XbaI and then ligated together. The unstable *cre* plasmid was constructed by amplifying *cre* from pJG125 (141) using primers oJG1350 and oJG1320, which was then digested with BamHI, and ligated into pRK7813 to yield pJG577. The various plasmid constructs were introduced into *S. meliloti* strains by triparental mating. When appropriate, *S. meliloti* was made tetracycline resistant by integration of pJG505 (110) into the *rhaK–icpA* intergenic region of the chromosome (124). HR plasmid transfer was accomplished by triparental mating, using the helper strain B001 (described above). The *cre* plasmid pJG577 was similarly introduced by triparental mating. Following *cre*-mediated recombination between *loxP* sites pJG577 was cured by multiple passages in the absence of tetracycline selection.

For modification of *S. fredii* Sym plasmids to enable pRK600-mediated transfer, an *oriT*/*neo* cassette was introduced by single-crossover homologous recombination using the pUC-based plasmid pJG194 (123). pJG194 was targeted (nondisruptively) to the *y4xK–nopL* intergenic region. Fragments corresponding to the *y4xK–nopL* intergenic regions of *S. fredii* USDA 257 and *S. fredii* NGR234 were amplified using primer pairs oMC099 and oMC100 and oMC103 and oMC104, respectively, followed by ligation into pJG194 to yield integration plasmids pJG527 and pJG528. These modified pJG194 constructs were introduced into *S. fredii* strains by triparental mating, using the helper strain B001 (described above). Subsequent mobilization of the *oriT*/*neo*-modified Sym plasmid into *A. tumefaciens* UBAPF2 (125) was performed as described previously (110).

2.2.3.3 Modified Eckhardt gel electrophoresis.

Modified Eckhardt gels were performed as previously described (81) with modifications. Briefly, bacteria were grown to an OD_{600} of 0.6 in LB. Culture (150 µl) was added to 500 µl of chilled 0.3% sarkosyl in 1× SBE (20× SBE = 500 ml of H_2O, 4 g of NaOH, 3.72 g of Na_2-EDTA·$2H_2O$, pH to 8.0 with boric acid). Each sample was pelleted and resuspended in 20 µl of lysis solution (1× SBE, 10 mg of sucrose per milliliter, 1 mg of lysozyme per milliliter, 40 µg of RNase A per milliliter), followed immediately by loading into a sodium dodecyl sulfate (SDS)–SBE minigel (1× SBE, 0.8% agarose, 0.5% SDS). Each sample remained in the well for 5 min, followed by electrophoresis at 23 V for 10 min, followed by electrophoresis at 96 V for 90 min. The minigel was then stained for 1 h in 0.4 µg of ethidium bromide per milliliter and was destained for 10 min in H_2O prior to imaging.

2.2.3.4 Identification of putative toxin–antitoxin pairs on pSymB and pSymA.

Lists of putative toxin–antitoxin (TA) pairs were collected from two previous studies where TAs were identified in the *S. meliloti* Sym plasmids (145–149). The sequences for the pSymB and pSymA of *S. meliloti* Rm1021 (GenBank Accessions AL591985.1 and AE006469.1, respectively) were also individually submitted to the RASTA-Bacteria website (149) (http://bioinfo-mml.sjtu.edu.cn/TADB/) for automated TA prediction. Putative TA pairs which had not been identified previously were then compared to the GenBank database using BLASTp (150). Predicted TAs which did not share homology with other known or predicted TAs in the database were discarded.

2.2.3.5 Genomic alignment of *Sinorhizobium* Sym plasmids.

The following sequences (accession numbers in parentheses) were downloaded from the NCBI ftp website (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/): pSMED01 of *S. medicae* WSM419 (NC_009620.1), pSMED02 of *S. medicae* WSM419 (NC_009621.1), chromosome 2 of *S. meliloti* AK83 (NC_015596.1), chromosome 3 of *S. meliloti* AK83 (NC_015591.1), pSINMEB01 of *S. meliloti* BL225C (NC_017324.1), pSINMEB02 of *S. meliloti* BL225C (NC_017323.1), pSmeGR4c of *S. meliloti* GR4 (CP003936.1), pSmeGR4d of *S. meliloti* GR4 (CP003937.1), pSYMA of *S. meliloti* Rm41 (NC_018683.1), pSYMB of *S. meliloti* Rm41 (NC_018701.1), pSymA of *S. meliloti* Rm1021 (NC_003037.1), pSymB of *S. meliloti* Rm1021 (NC_003078.1), pSmeSM11c of *S. meliloti* SM11 (NC_017327.1), and pSmeSM11d of *S. meliloti* SM11 (NC_017326.1). Alignment was performed using progressiveMAUVE version 2.3.1 build 18 (151) (http://gel.ahabs.wisc.edu/mauve/).

2.2.4 Results

2.2.4.1 Deletion of the pSymB *exo* region.

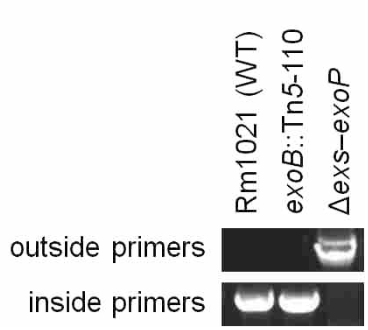

Figure 2.2-2. Confirmation of deletion of the *exo* genes of pSymB in *S. meliloti*. Primers that lay outside either end of the *exo* locus (oMC371 and oMC372) and primers that lay inside the *exo* locus (oMC293 and oMC294) were used to verify removal of the *exo* locus.

I first conducted a proof-of-principle experiment in a tractable strain, *S. meliloti* Rm1021. A 505-bp fragment of *SMb20931*, which lies at one end of the *exo* biosynthetic locus, was amplified from Rm1021 using primers oMC308 and oMC309 and ligated into pJG563 to produce plasmid pJG568. A 531-bp fragment of *thiD*, which is situated at the other end of the *exo* genes, was amplified from Rm1021 using primers oMC310 and oMC311 and ligated into pJG565 to yield plasmid pJG569. Both insertions were made independently and verified by PCR. I then transduced the *thiD*::pJG568 insertion (strain C582) into the *SMb20931*::pJG569 strain (C583). *cre*-mediated recombination between the *loxP* sites, upon introduction of pJG577, resulted in the removal of approximately 36 kilobases (Δ*exsH–exoP*) to yield strain C581. A subsequent PCR test confirmed the loss of the target region (**Figure 2.2-2**).

2.2.4.2 Large-scale deletions in HR plasmids.

I then applied this system to making large-scale deletions in two of the host range plasmids, pHRC017 and pHRB469. Based on the sequence conservation between pHRC017 and the other three host range plasmids (**Figure 2.1-6B**), I divided pHRC017 into quadrants, starting just downstream from the $repA_1BC_1$, and chose the first two quadrants for the first round of targeted deletions. Given the high degree of similarity between pHRC017 and pHRB469 in these quadrants, I used the same plasmids developed for pHRC017 to make similar deletions in pHRB469. For each host range plasmid, three deletions were attempted: removal of the first

quadrant, removal of the second quadrant, and removal of the first two quadrants (first half) simultaneously (as opposed to iteratively).

Working with pHRC017 and pHRB469 in their native context was complicated by the fact that both strains are to some degree resistant to both of our transducing phages (ΦM12 and ΦN3). So, for convenience, after one of the integration plasmids had been inserted, the host range plasmids were mobilized into a plasmid-free *Agrobacterium tumefaciens* background (strain UBAPF2) before attempting integration of the second plasmid. We are unaware of any transducing phages for *A. tumefaciens*, so introduction of the second integration plasmid had to take place in a strain which had already been modified with the first integration plasmid. This scenario is less ideal than utilizing transduction because the two integration plasmids have some sequence in common, creating the possibility of undesired insertion of the second integration plasmid into the first, rather than into the intended target. Besides having a 300-bp *oriT* in common, the pUC *oriV* and pCDF *oriV* (which are compatible) have ~81% sequence identity across ~500 bp. Thus there are four possible scenarios for off-target integration of the second plasmid (**Figure 2.2-3**). Each scenario produces a unique genetic structure which can be diagnosed using PCR: primers oMC122 and oMC129 will only amplify a product under the scenario shown in **Figure 2.2-3C**; primers oMC122 and oMC130 will only amplify a product under the scenario shown in **Figure 2.2-3D**; primers oMC121 and oMC130 will only amplify a product under the scenario shown in **Figure 2.2-3G**; and primers oMC121 and oMC129 will only amplify a product under the scenario shown in **Figure 2.2-3H**.

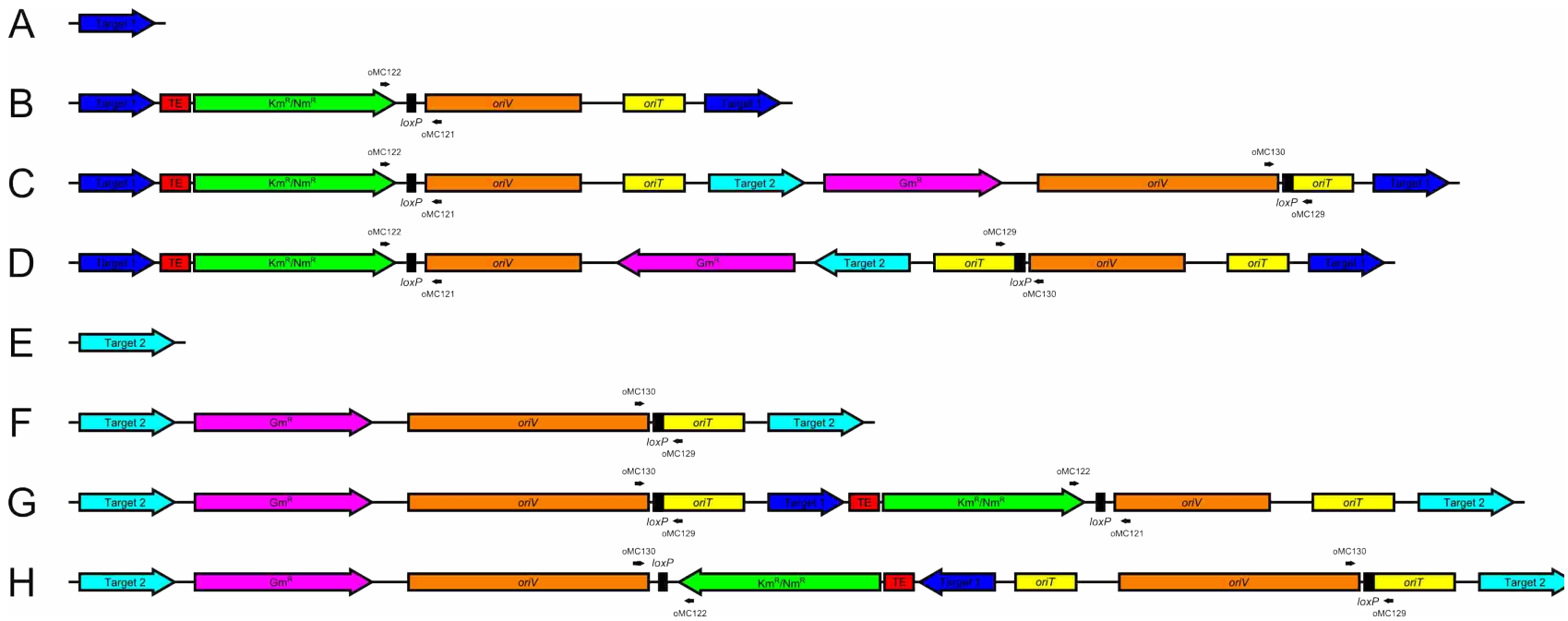

Figure 2.2-3. Scenarios for incorrect insertion of the second integrative plasmid. A. The first target gene. **B.** The first target gene with an integrated pJG563 derivative. **C.** Insertion of a pJG565 derivative into the *oriT* of the pJG563 derivative instead of the second target gene. **D.** Insertion of a pJG565 into the *oriV* of the pJG563 derivative instead of the second target gene. **E.** The second target gene. **F.** The second target gene with an integrated pJG565 derivative. **G.** Insertion of a pJG563 derivative into the *oriT* of the pJG565 derivative instead of the first target gene. **H.** Insertion of a pJG563 into the *oriV* of the pJG565 derivative instead of the first target gene. The locations of primers which can be used to diagnose these incorrect insertions are indicated.

Once incorrect insertions had been eliminated, pJG577 (*cre*) was introduced to mediate excision of the sequence between the introduced *loxP* sites. Deletions of quadrant one, quadrant two, and both quadrants together were successfully recovered for both pHRC017 and pHRB469 (**Figure 2.2-4**). PCR tests against targets expected to be lost in the pHRC017 deletions were conducted against all six deletion strains as well as against a spontaneous deletion in pHRB800 of indeterminate size which no longer restricted host range, and against *A. tumefaciens* UBAPF2 (125) as a negative control (**Table 2.2-1**). As expected, loci belonging to a particular quadrant of pHRC017 were lost as that quadrant was deleted. Consistent with our previous observations that pHRB469 had substantial sequence in common with the first two quadrants of pHRC017 (**Table S5.2-6**), deletions in pHRB469 corresponding to quadrants one and two of pHRC017 produced identical PCR results.

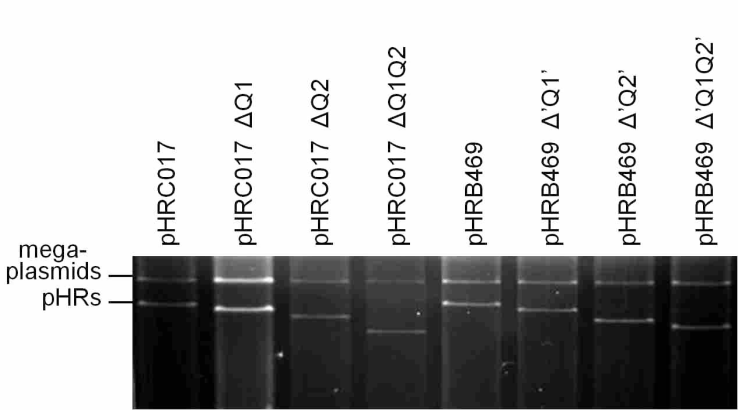

Figure 2.2-4. Large deletions can be made in host range plasmids using a *cre–loxP* system. An Eckhardt gel of intact host range plasmids (HR plasmids) and deletion derivatives. *S. meliloti* symbiotic megaplasmids (>1 Mb) are indicated.

Table 2.2-1. PCR confirmation of deletions in HR plasmids. + = target ORF was detected; – target ORF was not detected; NT = not tested.

	Target ORF	pHRC017	pHRC017 ΔQ1	pHRC017 ΔQ2	pHRC017 ΔQ1Q2	pHRB469	pHRB469 Δ'Q1'	pHRB469 Δ'Q2'	pHRB469 Δ'Q1Q2'	pHRB800	pHRB800 ΔX	UBAPF2
Q1	*orf00016*	+	–	+	–	+	–	+	+	+	+	–
	orf00054	+	–	+	–	+	–	+	+	+	+	–
	orf00066	+	–	+	–	+	–	+	+	+	+	–
	orf00070	+	–	+	–	+	–	+	+	+	+	–
	orf00087	+	–	+	–	+	–	+	+	+	+	–
	orf00104	+	–	+	–	+	–	+	+	+	+	–
	orf00163	+	–	+	–	+	–	+	+	+	+	–
	orf00172	+	–	+	–	+	–	+	+	+	+	–
Q2	*orf00223*	+	+	–	–	+	+	–	–	+	+	–
	orf00227	+	+	–	–	+	+	–	–	+	+	–
	orf00232	+	+	–	–	+	+	–	–	+	+	–
	orf00256	+	+	–	–	+	+	–	–	–	NT	–
	orf00269	+	+	–	–	+	+	–	–	–	NT	–
	orf00281	+	+	–	–	+	+	–	–	–	NT	–
	orf00318	+	+	–	–	+	+	–	–	–	NT	–
	orf00389	+	+	–	–	–	NT	NT	NT	–	NT	–
Q4	*orf00528*	+	NT	NT	NT	+	+	+	+	–	NT	–
	orf00567	+	NT	NT	NT	+	+	+	+	–	NT	–
	orf00636	+	NT	NT	NT	+	+	–	–	–	NT	–
	orf00689	+	NT	NT	NT	+	+	–	–	–	NT	–

Surprisingly, when a deletion in pHRB469 corresponding to the first half of pHRC017 was made, the PCR targets from quadrant one were retained. Since pHRB469 also shares some sequence with quadrant four of pHRC017 (**Table S5.2-6**), we also tested some loci from that region and discovered that some, but not all, had been lost both in the 'Q2' and the 'Q1Q2' deletions in pHRB469. This suggests some differences in the architecture of pHRB469 compared to pHRC017. The deletion in pHRB800, which abolished its ability to restrict host range, did not affect any of the loci shared with pHRC017 (**Table 2.2-1**).

The host range plasmids, modified with the various deletions, were then mobilized back into an *S. meliloti* background where they can be tested for altered phenotypes, such as increased host range, susceptibility to Tn*5* mutagenesis, or loss of competitiveness for nodule occupancy.

2.2.4.3 Identification of TA loci in *S. meliloti* Rm1021 Sym plasmids.

In order to cure pSymA and pSymB there are several obstacles. First, on pSymB there are two genes which have been shown to be essential, *engA* and an arginine tRNA (CCG) (143). Additionally, pSymB contains a cobalt uptake operon (*cbtJKL*) which is essential at trace concentrations of cobalt (152) and *SMb21254*, which can be deleted but with an impaired growth rate for the cell (143). Several other genes have been postulated to be essential: the potentially essential *minCDE* genes, one of the asparagine biosynthesis genes *asnB* or *asnO*, and/or the thiamine biosynthesis genes *thiD* and *thiCOSGE* (153), but a deletion analysis of pSymB did not report any difficulty in deleting these genes (143). The locations on pSymB of all of these genes are shown in **Figure 2.2-5**.

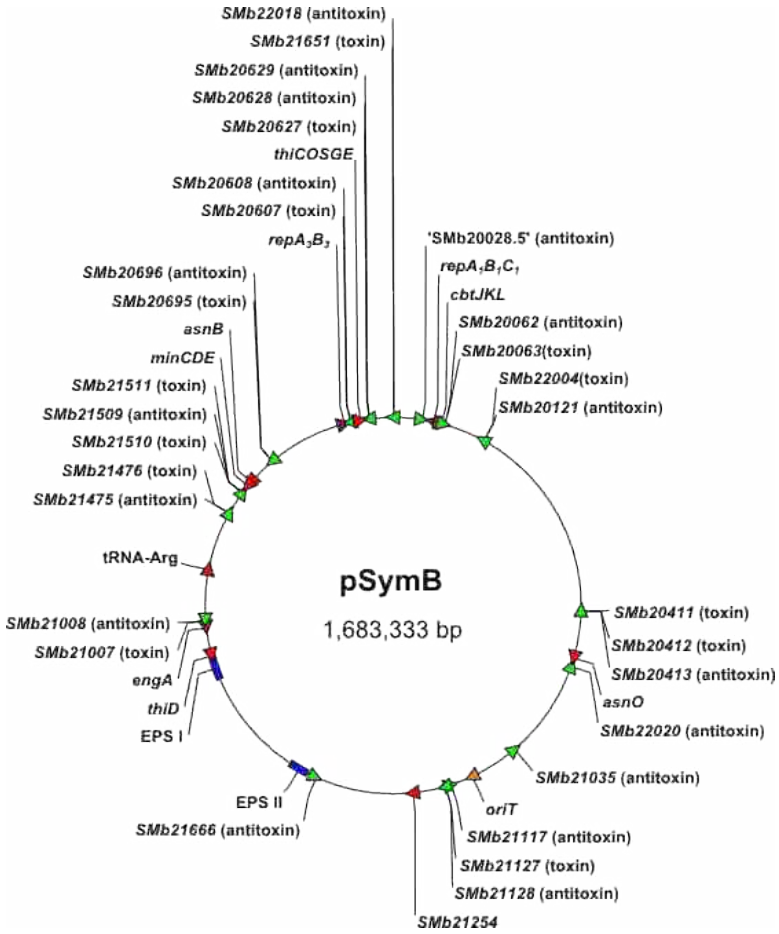

Figure 2.2-5. pSymB carries several toxins, antitoxins, and essential genes. Green, toxins and antitoxins; purple, replication genes; orange, *oriT*; blue, genes involved in symbiosis; red, genes known or suspected (at some point) to be essential (see text).

Second, the presence of toxin–antitoxin (TA) pairs on either replicon would complicate curing efforts. Several bioinformatic approaches have been applied to the *S. meliloti* genome to determine the location of potential TA pairs (145–149). In combination these analyses predicted 13 toxins and 17 antitoxins on pSymB (**Figure 2.2-5**; for more details see **Table S5.2-9**) and 24 toxins and 25 antitoxins on pSymA, as well as 3 toxin pseudogenes and 3 antitoxin pseudogenes (**Figure 2.2-6**; for more details see **Table S5.2-10**). Previous deletion analysis has determined that 3 TAs on pSymA (*SMa0471–SMa0473*; *SMa2105*; and *SMa2230–SMa2231*) and 0 TAs on pSymB were difficult to delete (143). While these TAs haven't proven to be problematic for small deletions across these replicons, they could be prohibitive when attempting to remove

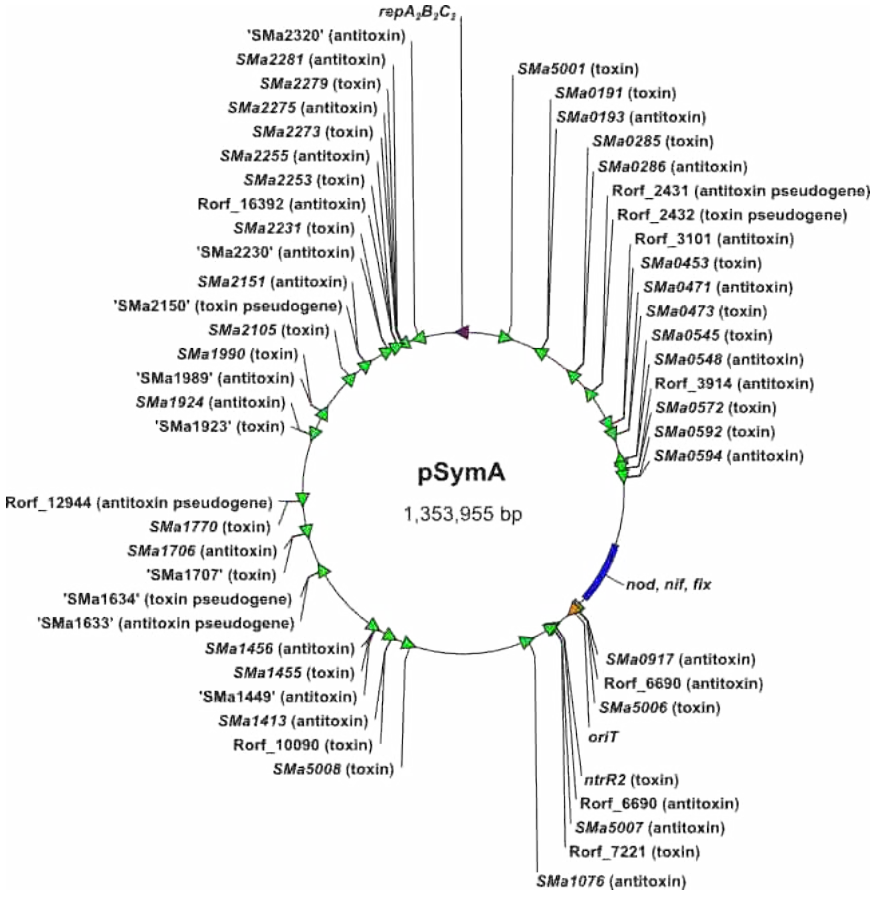

Figure 2.2-6. pSymA carries various toxins, antitoxins, and symbiotic genes. Green, toxins and antitoxins; purple, replication genes; orange, *oriT*; blue, genes involved in symbiosis.

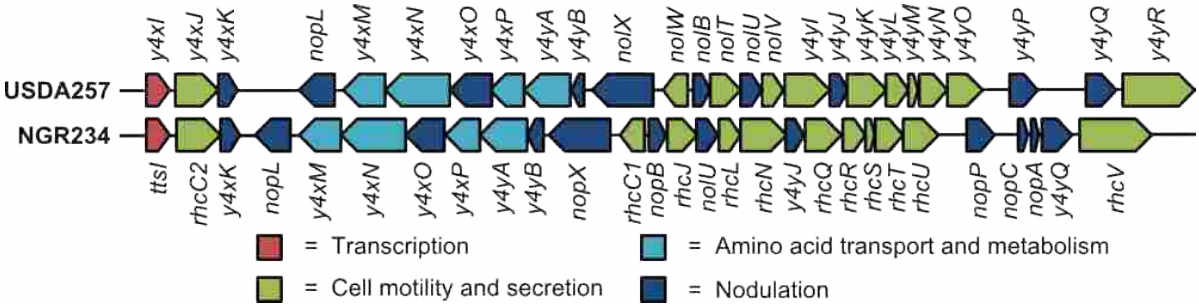

Figure 2.2-7. A pairwise alignment of the *nop–nol* regions of the Sym plasmids of *S. fredii* strains USDA257 and NGR234.

many or all of them at once.

Deletion of both pSymB and pSymA from *S. meliloti* Rm1021 would constitute a loss of 45.4% of its core genome. Besides the loss of known symbiotic loci (*fix*, *nif*, *nod*, *exo*, etc.), the loss of TA loci could have unpredictable consequences for the cell. For example, it has been determined that at least one TA pair in *S. meliloti*, *ntrPR*, plays a role in symbiosis (154, 155). Given the global ability of TAs to regulate transcriptional activity, drastic changes could take place in the *S. meliloti* transcriptome (156).

2.2.4.4 Mobilization of the Sym plasmids of *S. fredii*.

Loss of both pSymB and pSymA from *S. meliloti* Rm1021 would essentially erase its symbiotic specificity. Subsequent introduction of Sym plasmids from other rhizobia could effectively switch its host range. *Sinorhizobium fredii* NGR234 has the broadest host range of any known rhizobial strain, with the ability to nodulate legumes in 112 genera (157). *Sinorhizobium fredii* USDA 257 also has a broad host range, though not as large as strain NGR234, nodulating only 79 genera (157). The genome of *S. fredii* USDA 257 was recently published (158) but at the time I began work on this aspect of the project, only a fragment of the Sym plasmid sequence was available (131). Based on a pairwise alignment of the sequence of that fragment from the Sym plasmid of *S. fredii* USDA 257 (GenBank Accession AF229441.2)

and the Sym plasmid of *S. fredii* NGR234 (GenBank Accession U00090.2) (**Figure 2.2-7**), the *y4xK–nopL* intergenic region was chosen for nondisruptive integration of an *oriT*/*neo* cassette to facilitate transfer into plasmid-free *A. tumefaciens* UBAPF2 (125). Transfer was successfully performed and confirmed by Eckhardt gel and PCR (data not shown).

2.2.5 Discussion

The pJG563/pJG565 *loxP* system is functional for making large-scale deletions. Deletions of up to ~150 kilobases have been accomplished in this work, but even larger deletions are theoretically possible and attempts in pSymA and pSymB are currently underway. Its utility has been demonstrated here in *S. meliloti* but it should be generally useful in a variety of α-proteobacteria. The ideal scenario would involve use in a background for which there are general transducing phages available, such as *S. meliloti* Rm1021.

The use of this system for analyzing HR plasmids has currently been superseded by mutagenesis using *himar1* rather than Tn*5* (see **§2.1.6**). However, the deletions already made in pHRC017 and pHRB469 can be used to corroborate the data from *himar1* mutants. Furthermore, if the *repA*$_2$*C*$_2$ and LuxR-family regulator mutants are found to reduce plasmid stability rather than function directly in host range restriction, this system may still prove to be more useful than transposon mutagenesis for analysis of certain HR plasmids. The deletions in pHRC017 and pHRB469 that have already been made can also be tested for alteration of two other phenotypes associated with those plasmids: repulsion of Tn*5* insertions and competitiveness for nodule occupancy.

Other researchers in the lab are currently using derivatives of pJG563 and pJG565 to make large deletions in pSymA and pSymB with the goal to eventually eliminate these replicons from the cell. To overcome the lethality associated with deleting essential genes on pSymB, *engA* and the tRNA-Arg have been cloned into a mini-Tn*5* and relocated to the *S. meliloti* Rm1021 chromosome. Rather than clone the *cbtJKL* operon, additional cobalt will be added to the growth media used, thus obviating the need for these genes. There are currently no plans to

compensate for the slow-growth phenotype associated with *SMb21254*, but if deletion of pSymB proves to be troublesome, this may be required, as well.

Several previous studies have reported deletions in pSymB of various sizes, but never complete curing (142, 143, 159, 160). An alignment of pSymB replicons from five sequenced *S. meliloti* genomes (the recently released genome for *S. meliloti* GR4 {GenBank Accession CP003937.1} was not included) and one sequenced *S. medicae* genome (**Figure S5.1-5**) shows good conservation of synteny. Two regions show greater than average variability. The first, indicated by a box with a solid line, occurs in the vicinity of several *rkp* and *kps* genes and is due to high variability of those genes (possibly in response to phage pressure), the existence of a module of polysaccharide biosynthesis genes (*SMb21053–SMb21082*) in two of the strains (Rm1021 and BL225C), a ~50 kB rearrangement in WSM419 (indicated by a dashed line), and a number of transposons unique to each strain. The second, indicated by a box with a dotted line, occurs in the vicinity of the tRNA-arginine (CCG) and is due to the presence of putative prophage elements in three of the genomes (Rm1021, BL225C, and WSM419). The equivalent of this region in *S. fredii* is found on the chromosome, so it is postulated that this region, with its essential genes (*engA* and the tRNA-Arg, as well as the *bacA* gene which is important for symbiosis) translocated from the genome of an ancestor common to *S. meliloti* and *S. medicae* after they had diverged from *S. fredii* (161). The *S. meliloti* pSymB has also been reported to share a common ancestor with the linear plasmid of *Agrobacterium* species (162). It has been suggested that pSymB should be considered a second chromosome in *S. meliloti* rather than a megaplasmid (153, 163), despite its plasmid-like features (164). Alternatively, a new term for pSymB and other chromosome-like plasmids in other bacteria has been proposed: chromid (165). Thus successful deletion of pSymB would represent the first instance of a chromid being cured.

Deletions in pSymA have also been reported (143, 159) and in one case pSymA was successfully cured, but doing so required several rounds of deletions (166). This is most likely due to the presence of TA loci. A genomic alignment of pSymA shows much greater variability (**Figure S5.1-6**).

The Sym plasmid of *S. fredii* NGR234, pNGR234a, has previously been transferred to a strain of *A. tumefaciens* cured of its pTi plasmid. Only a small percentage of transconjugants were able to nodulate either of two hosts of *S. fredii* NGR234, *Psophocarpus tetragonolobus* or *Vigna unguiculata*, and nitrogen fixation did not occur (167). In this same study they transferred pNGR234a into a strain of *S. meliloti* with its *nod* and *nif* genes deleted. Nodulation of *P. tetragonolobus* and *V. unguiculata* occurred, but nitrogen fixation was variable (167). Since *S. fredii* NGR234 produces symbiotically active polysaccharides from genes encoded on its larger megaplasmid, pNGR234a, which is analogous to the *exo* genes of pSymB, transfer of pNGR234b might also be necessary before determining the host ranges of the transconjugant strains. The larger megaplasmid of *S. fredii* USDA 257 has cointegrated into the chromosome (158), so transfer may be difficult if not impossible.

2.2.6 Future Directions

pSymA has only ever been cured in the presence of an intact pSymB. And deletions in pSymB have always been performed in the presence of an intact pSymA. Thus there exists a possibility that both cannot be cured simultaneously due to synthetic lethality. If more genes required for viability in *S. meliloti* are revealed, these too would need to be cloned onto the chromosome prior to complete curing of the megaplasmids. Once both megaplasmids are absent from *S. meliloti* Rm1021, several analyses can be performed to investigate the adjustments made by the cell. Once both pSymA and pSymB have been cured it is possible that some chromosomal genes that weren't previously essential will become so due to the loss of a pSym gene with overlapping function. Saturation of the *S. meliloti* Rm1021 chromosome with Tn*5* and subsequent Illumina sequencing off the transposons might reveal additional essential loci not apparent in *S. meliloti* Rm1021 with its Sym plasmids intact. The alterations in metabolic phenotypes and transcription could be determined by Biolog array and RNA-seq, respectively.

In addition to testing for altered host range due to the introduction of plasmids from strains of *S. fredii*, plasmids from other rhizobia could also be introduced and the host ranges of the resulting transconjugants tested. The *loxP* deletion system could also be adapted to 'grab' sequences, mate them out into *E. coli*, and then transfer them elsewhere. For example, the *exo*-like locus from pNGR234b or from the linear *A. tumefaciens* chromosome can be mobilized into a pSymB-cured strain of *S. meliloti* to see if the succinoglycans of those strains would rescue infection thread invasion. Some rhizobia encode *nod* and *nif* functions on symbiosis islands rather than plasmids (e.g. *Bradyrhizobium* spp., *Mesorhizobium* spp., and *Azorhizobium caulinodans*). This technique could be used to transfer symbiosis islands into the pSym-cured strain of *S. meliloti* to evaluate their effect on host range.

Alternatively, this technique could be used to construct a core set of genes that would rescue symbiosis upon reintroduction to the strain cured of both its Sym plasmids. If all known pSym genes required for symbiosis did not rescue symbiosis in this experiment it might suggest that there are as-yet-unidentified pSym genes or that there are mechanistic constraints which cannot be duplicated by capturing and reassembling symbiosis genes.

2.3 Characterization of Sinorhizobia and Other Rhizobia Collected Primarily in the Lower Intermountain West

2.3.1 Summary

Several plant species which are known to form symbiotic root nodules in association with the rhizobia *Sinorhizobium meliloti* and *Sinorhizobium medicae* are found in Utah: *Melilotus alba* (white sweetclover), *Melilotus officinalis* (yellow sweetclover), *Medicago lupulina* (black medic), and *Medicago sativa* (alfalfa). All four of these species are native to the Mediterranean basin and/or the Middle East and are introduced in Utah, either as invasive weeds, as cultivated forage, or both. Despite being found in an exotic environment, they are still found to be nodulated. It is unknown whether the rhizobia in these root nodules are native, introduced with their foreign hosts, or native with symbiotic functions acquired horizontally from invading strains. In conjunction with high school biology classes we have initiated the collection of local isolates of *S. meliloti* and *S. medicae* antecedent to their characterization and phylogenetic analysis. Preliminary results are presented as well as characterization of non-*Sinorhizobium* rhizobia collected.

2.3.2 Introduction

The collection of, characterization of, and subsequent retesting of root nodule rhizobia is relatively easy and entails no health risks. Thus it is a useful and convenient system for introducing and/or reinforcing several concepts to students in secondary education, including sterile technique, streaking for pure culture, formulating and testing hypotheses, bacterial identification, PCR, gel electrophoresis, DNA sequencing and analysis, phage typing, statistical analysis, Koch's postulates, and phylogenetics. With this in mind we formed an outreach to local high school biology teachers and teachers of BYU undergraduate biology courses to provide them with materials and reagents and to help them integrate the study of the rhizobium–legume symbiosis into their classrooms. This outreach is called the Symbiosis Learning Consortium (SymLC).

The research focus in our lab is on the *Medicago–Sinorhizobium* symbiosis so we chose to retain this focus for the SymLC. The natural geographic range of *Medicago* species and species in the related genera, *Melilotus* and *Trigonella* (which are also nodulated by *Sinorhizobium*), is limited to the Mediterranean basin and southwest Asia (73). However, in modern times several species have achieved worldwide distribution. Available hosts in Utah are either cultivated (*Medicago sativa*) or are considered an invasive weed (*Melilotus alba*, *Melilotus officinalis*, *Medicago lupulina*, *Medicago polymorpha*, *Medicago sativa*, and *Trigonella corniculata*). In conjunction with this outreach to local high schools, we also collected local rhizobia and analyzed them, both to validate the research being done by the high school students and to generate rigorous data with an eye toward peer-reviewed publication, as has been done previously in other locales (128, 168–181).

2.3.3 Materials and Methods

2.3.3.1 Sample collection, bacterial isolation from root nodules, and culture conditions.

Root nodules were collected from any of the following invasive species of sub-tribe Trigonellinae: *Melilotus alba*, *Melilotus officinalis*, *Medicago lupulina*, or *Medicago sativa*. (*Medicago polymorpha* and *Trigonella corniculata* are rare in Utah and were never encountered by us.) A Google Map, exhibiting the collection locations for 2010 (green pins), 2012 (blue pins), as well as the countries of origin for several control strains (red pins), can be accessed at http://goo.gl/maps/OFZAl. Non-*Sinorhizobium* rhizobia collected in 2012 are indicated with green pins rather than blue. Excised nodules were surface-sterilized in 75% ethanol for 30 s and in 1.5% hypochlorite for 60 s. Following several rounds of washing, root nodules were then macerated in tryptone–yeast extract (TY) broth and 10% glycerol and streaked to single colonies on TY agar. Subsequently, *S. medicae* and *S. meliloti* cultures were grown at 30°C in TY or lysogeny broth (LB). Other rhizobia collected simultaneously are listed in **Table S5.2-18**. Recovery of bacteria from non-Trigonellinae nodules was performed identically except that yeast extract–mannitol (YEM) broth was used instead of TY.

2.3.3.2 Identification of unknown bacterial isolates and hosts.

For diagnosis between *S. meliloti* and *S. medicae,* three sets of primers were developed to amplify regions which exhibited length polymorphisms between the two species. Primers oMC293 and oMC294 amplify a fragment of the *exoP–thiD* intergenic region, which is 987 bp for *S. meliloti* and 692 bp for *S. medicae*. Primers oMC295 and oMC296 amplify a fragment from *SMc01522* to *ntrR1*. In *S. meliloti* there is an additional gene in this region, *ntrP*, which is lacking in *S. medicae*. Thus in *S. meliloti* the fragment is 1108 bp and in *S. medicae* it is only 309 bp. Primers oMC297 and oMC298 amplify a fragment of the *pyc–SMc03896* intergenic region,

which is 727 bp in *S. meliloti* and 486 bp in *S. medicae*. This last pair was chosen for use by the high school students. For more definite identification, 16S rDNA sequences were amplified and sequenced from unknown bacterial isolates using primers oJG1035 and oJG1036 or oJG1038, and then submitted to the SeqMatch program found at the Ribosomal Database Project (RDP) (182) (http://rdp.cme.msu.edu/index.jsp) set to compare only type strains of accepted species. When strain identity was ambiguous, alignment with the 16S rDNA sequences of the top hits, acquired from the List of Prokaryotic names with Standing in Nomenclature (183) (http://www.bacterio.cict.fr/), was performed using BLASTn (150). DNA from plant hosts was isolated from leaves macerated in liquid nitrogen using a Qiagen Blood & Cell Culture DNA Mini Kit according to manufacturer's instructions. Sequence was amplified from several commonly-used barcoding loci (184): the plastid *rbcL* gene using primers oMC091 and oMC092, the nuclear internal transcribed spacer region (ITS) using primers oMC093 and oMC094, and from the plastid *trnH–psbA* intergenic spacer using primers oMC095 and oMC096. The resulting sequences were BLASTed against the GenBank database and the known geographic ranges of the top hits from BONAP (http://www.bonap.org/) were consulted to see if they overlapped the collection site. Morphological characteristics were then compared to images available at the USDA Plants Database (http://plants.usda.gov/java/). Satisfaction of these three criteria (significant DNA similarity to vouchered specimens, occurrence in the known range, and correct morphological characteristics) were considered sufficient to confirm the identity of the host plant.

2.3.3.3 Plant growth and nodulation.

Plants used in this study are listed in **Table S5.2-11**. Surface-sterilized *Medicago sativa* seeds or scarified and surface-sterilized *Medicago praecox* seeds were allowed to germinate in

Petri plates, and 2-day-old seedlings were planted in sterile Turface clay particles (Turface Athletics, Buffalo Grove, IL, U.S.A.) and were allowed to grow for 4 days before being inoculated with 1 ml of *S. meliloti* or *S. medicae* cells suspended in 2.6 mM KH_2PO_4 (adjusted to pH 7.0 with 2N KOH) to an optical density at 600 nm (OD_{600}) of approximately 0.1. Nodulation was allowed to proceed for four weeks. Plants were watered with standard nodulation medium (SNM). SNM contains (per liter) 0.35 g of KH_2PO_4, 0.25 g of $MgSO_4$, 0.15 g of $CaCl_2 \cdot 2H_2O$, and 2 ml of minor salts solution (500× minor salts [per liter] = 9.5 g of Na_2-EDTA·$2H_2O$, 7 g of $FeSO_4 \cdot 7H_2O$, 1.5 g of $ZnSO_4 \cdot 7H_2O$, 1.5 g of H_3BO_3, 1.5 g of $MnSO_4 \cdot H_2O$, 0.15 g of $Na_2MoO_4 \cdot 2H_2O$, 15 mg of $CuSO_4$, and 15 mg of $CoCl_2$). The pH of this solution was adjusted to 7.0 with 2N KOH, and the medium was sterilized by autoclaving. Plants were maintained at approximately 27°C under fluorescent lamps (2.7 klux intensity, 16 h day length).

2.3.3.4 Modified Eckhardt gel electrophoresis.

Modified Eckhardt gels were performed as previously described (81) with modifications. Briefly, bacteria were grown to an OD_{600} of 0.6 in LB. Culture (150 µl) was added to 500 µl of chilled 0.3% sarkosyl in 1× SBE (20× SBE = 500 ml of H_2O, 4 g of NaOH, 3.72 g of Na_2-EDTA·$2H_2O$, pH to 8.0 with boric acid). Each sample was pelleted and resuspended in 20 µl of lysis solution (1× SBE, 10 mg of sucrose per milliliter, 1 mg of lysozyme per milliliter, 40 µg of RNase A per milliliter), followed immediately by loading into a sodium dodecyl sulfate (SDS)–SBE minigel (1× SBE, 0.8% agarose, 0.5% SDS). Each sample remained in the well for 2 min, followed by electrophoresis at 23 V for 10 min, followed by electrophoresis at 96 V for 90 min. The minigel was then stained for 1 h in 0.4 µg of ethidium bromide per milliliter and was destained for 10 min in H_2O prior to imaging.

2.3.3.5 Phylogenetic analysis.

All custom oligonucleotides were purchased from Invitrogen and are listed in **Table S5.2-12**. PCR amplification was conducted using *Taq* polymerase purchased from New England Biolabs (Ipswich, MA, U.S.A.). DNA sequences from sequenced genomes (accession numbers in parentheses) were downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/index.html): *Sinorhizobium medicae* WSM419 chromosome (CP000738.1), *Sinorhizobium meliloti* AK83 chromosome 1 (CP002781.1), *Sinorhizobium meliloti* BL225C chromosome (CP002740.1), *Sinorhizobium meliloti* Rm41 chromosome (HE995405.1), *Sinorhizobium meliloti* Rm1021 chromosome (AL591688.1), and *Sinorhizobium meliloti* SM11 chromosome (CP001830.1). *Sinorhizobium meliloti* GR4 chromosome (CP003933.1) was not available at the time this analysis was conducted. Sequences were aligned using the MUSCLE web server (185, 186) (http://www.ebi.ac.uk/Tools/msa/muscle/) with the default settings, and then manually adjusted using MacClade version 4.08 (187) (http://macclade.org/index.html). The phylogenetic reconstruction was conducted using maximum parsimony with 1000 replicates, implemented in PAUP* version 4.0 beta 10 (188) (http://paup.csit.fsu.edu/) using the default settings, and then visualized and exported using FigTree version 1.2.3 (http://tree.bio.ed.ac.uk/software/figtree/).

2.3.4 Results

2.3.4.1 Description of Strain Characteristics.

Root nodules were collected during the summers of 2010 and 2012. During 2010 eleven sites in Utah, Wyoming, and Montana were sampled, yielding 88 *Sinorhizobium* strains, of which 68 were *S. meliloti* and 20 were *S. medicae* (**Table S5.2-13**). All four host plants were represented with 20 strains isolated from *Melilotus alba*, 26 strains isolated from *Melilotus officinalis*, 32 strains isolated from *Medicago lupulina*, and 6 strains isolated from *Medicago sativa*. Interestingly, strains of *S. medicae* were only ever isolated from *M. lupulina*. This last observation could be due to specific activation of *S. medicae nolQS* genes by root exudates unique to *M. lupulina* (189).

During 2012 only three sites were sampled, all in Utah (**Table S5.2-14**), with a more carefully-defined sampling procedure: 1 m^2 plots were chosen which contained a large number of *Medicago* sp. Four plants were taken from each corner and four nodules were excised and processed from each plant, yielding a total of sixty-four root nodule isolates per collection site. *Medicago sativa* was sampled at two of the sites (Diamond Fork Canyon and Spring Hollow Road), while *Medicago lupulina* was sampled at the third site (Provo, Utah). These *Sinorhizobium* isolates haven't been further characterized, so the remainder of this discussion will focus on the strains collected in 2010.

We have previously observed that *S. medicae* strains are less likely to be able to grow on LB, possibly due to the high NaCl content (data not shown). When we tested the 88 strains from 2010 for the ability to grow on LB, six *S. medicae* strains were unable to grow on LB at all, nine grew poorly on LB, and five grew well on LB while one *S. meliloti* isolate grew poorly on LB and the remaining eighty-seven grew well (**Table S5.2-13**).

2.3.4.2 Accessory Plasmids in Utah Sinorhizobia.

To determine the accessory plasmid content of the strains collected in 2010, we performed Eckhardt gels (**Table S5.2-15**). The majority of collected strains had at least one accessory plasmid. Twenty-nine of the *S. meliloti* isolates had no accessory plasmids, twenty-seven had one accessory plasmid, seven had two accessory plasmids, and five had three accessory plasmids. Twelve of the *S. medicae* isolates had no accessory plasmids, six had one accessory plasmid, one had two accessory plasmids, and one had three accessory plasmids. No correlation between collection site and plasmid content or host species and plasmid content were apparent.

To test whether Utah accessory plasmids were similar to a known accessory plasmid in *S. meliloti*, pSmeSM11a (92), we designed primers to check for the presence of three genes from that plasmid: *repA*, *traA*, and *acdS*. The *repA* gene was chosen on the basis that it is involved in plasmid partitioning (134). The *traA* gene was chosen on the basis that it is involved in conjugal transfer of plasmids (190). And the *acdS* gene was chosen on the basis that it has been reported to improve nodulation (80, 191). These targets were also chosen on the basis that they are respectively found in two out of four, three out of four, and one out of four of the host range plasmids identified so far (see **Table S5.2-6**). Strains lacking accessory plasmids were excluded from this experiment. Four of sixteen strains tested possessed a *repA* gene similar to that of pSmeSM11a, twenty out of forty-six strains tested possessed a *traA* gene similar to that of pSmeSM11a, and thirty-two out of forty-seven strains tested possessed an *acdS* gene similar to that of pSmeSM11a (**Table S5.2-15**).

2.3.4.3 Host ranges of Utah Sinorhizobia.

A subset of the strains collected in the summer of 2012, including all strains which were shown to possess one or more accessory plasmids (**Table S5.2-15**), were inoculated onto a permissive host, *M. sativa*, and a more restrictive host, *M. praecox*, to evaluate their host ranges. Only one out of fifty-one strains tested (*S. medicae* U054) was unable to fix nitrogen on a *M. sativa* host. In addition to *S. medicae* U054, three additional strains (*S. meliloti* U016, U036, and U108, all isolated from *Melilotus officinalis*) did not fix nitrogen on *M. praecox*. Since all strains exhibiting host range restriction possessed multiple accessory plasmids, we inoculated several dozen *M. praecox* seedlings with each to see if we could recover a gain-of-compatibility (GOC) derivative for any of them. All four strains failed to produce GOCs (data not shown).

2.3.4.4 Phylogenetic analysis of Sinorhizobia.

As part of our analysis of Utah *S. meliloti* and *S. medicae* isolates, we wanted to compare them phylogenetically with *S. meliloti* and *S. medicae* isolates from the natural range of members of subtribe Trigonellinae. We wanted to determine whether local *S. meliloti* populations were native, introduced with their invasive hosts, or native with symbiotically-relevant genes horizontally acquired from foreign strains (172). Because of our interest in the rhizobium–legume symbiosis, we chose loci which were predominantly associated with symbiotic genes. Given that the megaplasmids carry a number of symbiotic genes, we chose three loci from pSymA (near the *fix*, *nif*, and *nod* genes), three loci from pSymB (near the *exo*, *thu*, and *wg*-(formerly *exp*) genes), and six loci from the chromosome (near the *cgm*, *feu* (*cyc*), *hem*, *ndv*, *ntr*, and *rkp* genes). A fourth pSymA locus was initially chosen on the basis that it was on the opposite side of the megaplasmid from the *nod* genes. Since it wasn't associated with any known symbiotic or housekeeping genes we simply labeled it the anti-*nod* locus. However, we were

unable to amplify this locus from three of the strains in our panel, so it was ultimately excluded. We chose intergenic regions rather than coding regions since they would be expected to be under less selective pressure and thus exhibit more sequence variation.

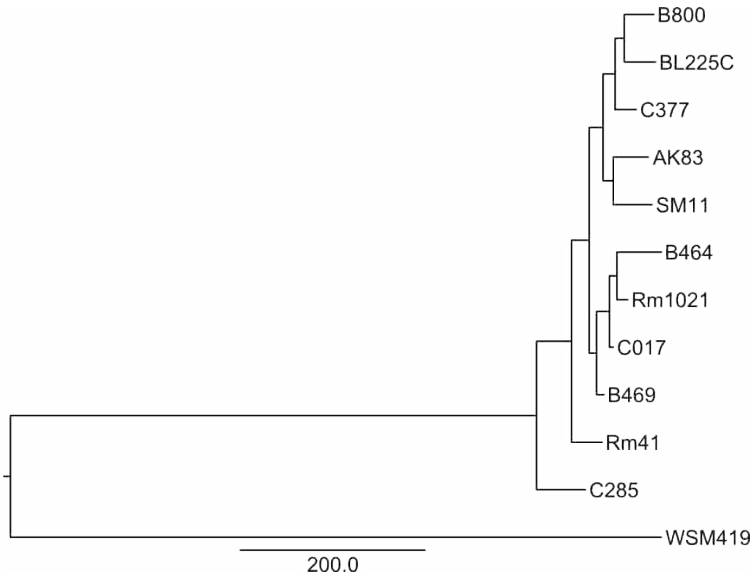

Figure 2.3-1. A maximum parsimony phylogenetic tree for all loci except anti-*nod*.

For a proof-of-principle experiment, we chose a panel of strains that included the four strains which exhibit the GOC phenomenon upon plasmid curing (B469, B800, C017, and C377) and two broad host range strains which do not possess any accessory plasmids (B464 and C285). We also retrieved corresponding sequences from GenBank for the five sequenced *S. meliloti* strains (AK83, BL225C, Rm41, Rm1021, and SM11) and the one sequenced *S. medicae* strain (WSM419) to use as an outgroup.

We performed an initial pairwise comparison of sequences (**Table S5.2-17**) to get a sense of how variable each locus was. At this point the genomic sequences from GenBank hadn't been included, with the exception of *S. meliloti* Rm1021. The chromosomal loci were relatively invariant. The pSymB loci were more variable, but the most sequence divergence was seen for the pSymA loci, which is consistent with previous reports (168, 180, 181, 192). After including the information from GenBank-curated genomes, we constructed phylogenetic trees for all sequenced loci (**Figure 2.3-1**), loci from each replicon (**Figure S5.1-7**), and each individual

locus (**Figure S5.1-8**, **Figure S5.1-9**, and **Figure S5.1-10**). As anticipated, based on the pairwise comparisons, the chromosomal loci yielded trees with very little branching—even when considered together. The pSymA and pSymB trees were much more resolved. Unfortunately, both pSymA and pSymB are transmissible (164), so they are presumably less reliable for reconstructing vertical descent. Thus for the purpose of conducting phylogenetic analysis in a classroom setting, we chose the *rkp* locus, which was the most variable of the chromosomal loci. Further analysis of the Utah sinorhizobia collected by our lab is ongoing.

To determine whether the inclusion or omission of the anti-*nod* sequence made a difference, we performed the analysis again with the anti-*nod* sequences included for those strains that could be PCR amplified in that locus. In the pSymA tree one branch lost resolution (**Figure S5.1-11A**, red asterisks) and the branch length was longer when anti-*nod* was included. In the overall tree, the only effect was an increase in branch length (**Figure S5.1-11B**).

2.3.4.5 Other rhizobia collected.

Despite the large variety of legumes, very little is known about the ability of most legume species to nodulate or the identity of their root nodule symbionts (193). So, in addition to the sinorhizobia collected we also performed a limited sampling of other rhizobia occupying root nodules of various legumes (**Table S5.2-18**). They have been tentatively identified based on their 16S rDNA sequence. Five strains collected from white clover (*Trifolium repens*), three strains collected from red clover (*Trifolium pratense*), and two strains collected from alsike clover (*Trifolium hybridum*) were identified as *Rhizobium leguminosarum* and presumably belong to biovar *trifolii*, which is consistent with rhizobia known to nodulate members of the genus *Trifolium* (74). Two strains, identified as *Rhizobium pisi* (U075 and U122), were recovered from American vetch (*Vicia americana*) and a plant tentatively identified as bladder senna (*Colutea*

arborescens). If these relationships can be confirmed, they would represent novel strain–host pairs. Additionally, a *Bosea* sp. (U123) was collected from *V. americana* which couldn't be matched with confidence to any described member of the genus *Bosea* and thus may represent a novel species. Two strains, identified as *Mesorhizobium loti* (U095 and U098) were recovered from *Astragalus alpinus*. Nodulation of *A. alpinus* by a *Mesorhizobium* sp. has been reported previously (194). A *Mesorhizobium* species (strain U124) was recovered from the root nodules of purple crownvetch (*Securigeria varia*). Identification of *Mesorhizobium* in root nodules of *S. varia* (syn. *Coronilla varia*) has been reported previously (195, 196). More exact identification was not possible using the 16S rDNA sequence since *Mesorhizobium* sp. U124 matched the 16S rDNA from four type species equally: *M. gobiense*, *M. metallidurans*, *M. tarimense*, and *M. tianshanense.* One strain of *Mesorhizobium amorphae* (U125) was recovered from the root nodules of *Lotus corniculatus*, which is consistent with previous reports (197, 198). Finally, a strain of *Rhizobium huautlense* (U126) was recovered from the root nodules of a semi-aquatic legume in the genus *Sesbania*, either poison bean (*S. drummondii*) or bag pod (*S. vesicaria*), which is also consistent with a previous report (199).

The species designations of these new isolates can only be tentative. There is evidence of 16S rDNA horizontal transfer in rhizobia (200–206), which can confound identification and phylogenetics based on this locus. Thus further sequence analysis is needed, preferably using conserved housekeeping genes such as *recA*. The exact set of genes to use would be dependent on the genus that the particular strain is predicted to belong to.

2.3.5 Discussion and Future Directions

The use of the chromosomal *rkp* locus gives moderately poor resolution of phylogenetic relationships but is expected to be more indicative of vertical descent. Thus its use in a high school classroom setting is acceptable. However, none of the loci taken separately or together in replicons produced well-resolved trees. Only when all twelve loci (excluding anti-*nod*) were taken together could a single, well-supported tree be created (**Figure 2.3-1**). Thus for the greatest accuracy in analyzing the relationships between Utah sinorhizobia collected from different nodules, individual plants, different corners of the collection plot, and different collection sites, all twelve genetic loci should be used. Alternatively, combinations of fewer genetic loci could be tested for their ability to resolve phylogenetic relationships so long as loci from all three replicons were included. Even though trees for the different replicons weren't completely resolved, they will still prove useful in determining whether pSymA and pSymB exhibit any horizontal transmission.

With respect to the non-sinorhizobia, strains which appear to be significantly different from any described species (e.g. *Bosea* sp. U123) would require further characterization, including DNA–DNA hybridization with its closest relatives, determination of mol%G+C content of DNA, carbon sources it can utilize, antibiotic resistance, determination of the major respiratory quinone, determination of the major fatty acids, and determination of the polar lipid profile to confirm it as a new species (207). Ideally additional strains should be collected from the same host species to reinforce the description of the new rhizobial species. In addition to the taxonomic descriptors mentioned above, reinoculation onto the original host plant would be needed to confirm the ability of the strain to nodulate.

2.4 Identification of an Essential Porin That Serves as a Phage Receptor for Multiple *Sinorhizobium meliloti* Phages

2.4.1 Summary

The bacterial receptors for ΦM12 and ΦN3, the two transducing phages universally used in the laboratory strain *Sinorhizobium meliloti* Rm1021, have never been identified. We recovered strains which were resistant to either phage separately or to both simultaneously and mapped the phage resistance mutations. All mutations mapped so far fall within the open reading frame of *SMc02396*, which encodes an outer membrane porin. We have renamed it *ropA1* on the basis of homology with a similar gene in *Rhizobium leguminosarum* bv. *viciae* strain 248. Subsequent analysis determined that either RopA1, LPS, or both are required for infection by a larger collection of phages. Attempts to disrupt or delete *ropA1* have all failed, leading us to conclude that *ropA1* may be essential for viability in *S. meliloti*. Subsequent phylogenetic and genomic analysis revealed that *ropA1* homologs in other Rhizobiales are often duplicated but that the duplications are always more closely related to each other than to homologs in other species, suggesting multiple independent duplication events.

2.4.2 Introduction

Multiple infective phages are expected to exist for each bacterial species, but outside of *E. coli* and *Lactococcus lactis* very little is known about the cell surface receptors used by phage to gain entry to the cell (208–210). Adsorption of phage to the bacterial host is the key host range determinant (211). Phage adsorption takes place in two steps: first, reversible contact with the host cell surface followed by irreversible binding to the host receptor (212, 213). Essentially any molecule exposed on the bacterial cell surface is available as a phage receptor. Bacteriophage receptors in Gram-negative bacteria can be classified into four broad categories: outer membrane proteins, flagella, pili, and extracellular polysaccharides (213). Within this latter group, the lipopolysaccharide (LPS) layer of Gram-negative bacteria is a common phage target. Outer membrane protein receptors can be further divided into five subcategories: structural proteins, porins, enzymes, high-affinity substrate receptors, and secretion proteins (211). A variety of tactics are employed by bacteria to prevent phage infection, including alteration, down-regulation, or deletion of the receptor; obstructing access to the receptor (though production of exopolysaccharides, lipoproteins, or competitive inhibitors); blocking phage DNA entry (often a consequence of lysogeny); restriction of phage DNA; CRISPR-mediated immunity; and even programmed cell death (210). With respect to alteration of the receptor, deletion or down-regulation can be costly for the bacterium (214), so sequence alteration is the most effective short-term solution. Spontaneous phage resistance is often brought about by mutations that alter receptor structure.

Two transducing phages, ΦM12 and ΦN3, are extensively used for transduction in the *S. meliloti* laboratory strain Rm1021. ΦM12 was originally isolated from a commercial *S. meliloti* inoculant manufactured in the United States (215) and ΦN3 was originally isolated from soil

obtained from an alfalfa field in Coachella Valley, California (216). Despite the distance separating their respective collection sites, ΦM12 and ΦN3 are predicted to be similar based on their reactions to antisera (215). Several general transducing phages have been identified for various other strains of *S. meliloti* (299–302), but their use has been more limited than the Rm1021 phages. Despite the ubiquitous use of phages ΦM12 and ΦN3, the corresponding bacterial receptors have never been described. In this work we identify an essential outer membrane porin, RopA1, as a receptor for both ΦM12 and ΦN3. Furthermore, we show that RopA1 and LPS account for the entry pathways used by all *S. meliloti* phages tested from a larger panel of diverse phage isolates.

2.4.3 Materials and Methods

2.4.3.1 Growth conditions and phage susceptibility assays.

Escherichia coli and *S. meliloti* cultures were grown at 37°C and 30°C, respectively, in lysogeny broth (LB) supplemented as follows: 4 mM $CaCl_2 \cdot 2H_2O$ (Ca^{2+}), chloramphenicol (Cm, 30 μg/ml), (Km, 30 μg/ml), (Nm, 100 μg/ml), (Sm, 200 μg/ml), and tetracycline (Tc, 5 μg/ml). To evaluate phage resistance 2 μl of phage lysate (10^8 to 10^9 PFUs/ml) were spotted onto lawns of *S. meliloti* on LB-Sm-Ca^{2+} agar.

2.4.3.2 Isolation of phage-resistant mutants.

S. meliloti Rm1021 was grown overnight in LB-Sm-Ca^{2+} and then subcultured. When the subculture had reached an OD_{600} ~1.0 a 30 μl aliquot of concentrated phage lysate (10^8 to 10^9 PFUs/ml) of either ΦM12 or ΦN3 were added. After 0.5 h incubation phage-infected cultures were embedded in LB-top agar and incubated at 30°C for 2–3 days until resistant colonies began to appear. Resistant colonies were picked out using a sterile toothpick, spread on LB-Sm-Ca^{2+} agar, and spotted with 2 μl undiluted phage to confirm resistance.

2.4.3.3 Plasmid and strain construction.

Plasmids and strains used in this study are listed in **Table S5.2-20**. Plasmids were constructed using standard techniques with enzymes purchased from New England Biolabs (Ipswich, MA, U.S.A.) The high-fidelity polymerase *Pfx*50 (Invitrogen, Carlsbad, CA, U.S.A.) was used for insert amplification. All custom oligonucleotides were purchased from Invitrogen and are listed in **Table S5.2-21**. Mobilization of plasmids was accomplished by triparental mating with helper *E. coli* B001 (DH5α harboring plasmid pRK600). pRK600 expresses trans-acting proteins required for mobilization of plasmids harboring the RK2 transfer origin (*oriT*). Tn*5*-110 mini-transposon delivery and identification of transposon insertion sites by arbitrary

PCR were described previously (123). Phage-mediated transduction was also described previously (215, 216).

2.4.3.4 RopA1 structural prediction.

Most solved β barrel structures are found in *E. coli*, but none exhibit significant sequence homology with *ropA1* (217), so structural predictions had to rely entirely on *in silico* predictions. After the removal of a 22-aa signal sequence predicted by SignalP 4.0 (218), which ends at the consensus peptidase cleavage site (AQA) (219), the amino acid sequence of RopA1 was analyzed to confirm that it was a transmembrane β-barrel and to predict its secondary structure using PRED-TMBB (220) (http://biophysics.biol.uoa.gr/PRED-TMBB/).

2.4.3.5 Phage adsorption assays.

Cultures of *S. meliloti* strains were grown overnight in LB-Sm-Tc-Ca^{2+}, then subcultured and grown to an OD_{600} ~1.0 whereupon 30 µl of concentrated phage lysate (10^8 to 10^9 PFUs/ml) was added to 400 µl of bacterial culture (or 400 µl of LB as a control) and shaken at 225 rpm at 30°C for 1 h (the predetermined time point at which maximum phage adsorption was observed in wildtype *S. meliloti* Rm1021). Cultures were then centrifuged for 30 s at 13.2 krpm. The supernatant, which contained unadsorbed phage, was then added to a fresh 400 µl culture of wildtype *S. meliloti* Rm1021, shaken at 225 rpm at 30°C for 0.5 h, embedded in LB-top agar, and incubated at 30°C overnight.

2.4.3.6 Genetic knockouts.

Disruption-integration plasmids were introduced into *S. meliloti* Rm1021 via triparental mating performed on LB agar. Mating lawns were suspended in LB supplemented with 10% glycerol, serially diluted, and plated on selective medium (LB-Sm-Nm). PCR checks to verify

plasmid integration into intended targets was conducted using a vector-specific primer (oJG1243) and a primer upstream of the target site (see **Table S5.2-21**).

2.4.3.7 Genomic alignments.

The following sequences (accession numbers in parentheses) were downloaded from the NCBI ftp website (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/): *Agrobacterium tumefaciens* C58 circular chromosome (AE007869.2), *Azorhizobium caulinodans* ORS 571 chromosome (AP009384.1), *Bartonella bacilliformis* KC583 chromosome (CP000524.1), *Bradyrhizobium japonicum* USDA 110 chromosome (BA000040.2), *Brucella melitensis* bv. 1 str. 16M chromosome I (AE008917.1), *Mesorhizobium loti* MAFF303099 chromosome (BA000012.4), *Rhizobium leguminosarum* bv. *trifolii* WSM1325 chromosome (CP001622.1), *Sinorhizobium fredii* HH103 chromosome (HE616890.1), *Sinorhizobium fredii* NGR234 chromosome (CP001389.1), *Sinorhizobium fredii* USDA 257 chromosome (CP003563.1), *Sinorhizobium medicae* WSM419 chromosome (CP000738.1), *Sinorhizobium meliloti* AK83 chromosome 1 (CP002781.1), *Sinorhizobium meliloti* BL225C chromosome (CP002740.1), *Sinorhizobium meliloti* GR4 chromosome (CP003933.1), *Sinorhizobium meliloti* Rm41 chromosome (HE995405.1), *Sinorhizobium meliloti* Rm1021 chromosome (AL591688.1), and *Sinorhizobium meliloti* SM11 chromosome (CP001830.1). Initial alignments were performed using progressiveMAUVE version 2.3.1 build 18 (151) (http://gel.ahabs.wisc.edu/mauve/) and then manually adjusted.

2.4.3.8 Phylogenetic analysis.

The following protein sequences (accession numbers in parentheses) were downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/index.html): *Agrobacterium tumefaciens* C58 *Atu1020* (AAK86828.1), *Agrobacterium tumefaciens* C58 *Atu1021*

(AAK86830.1), *Agrobacterium tumefaciens* C58 *Atu4693* (AAK88757.1), *Azorhizobium caulinodans* ORS 571 *AZC_1213* (BAF87211.1), *Azorhizobium caulinodans* ORS 571 *AZC_3535* (BAF89533.1), *Bartonella bacilliformis* KC583 *BARBAKC583_0447* (ABM44571.1), *Bradyrhizobium japonicum* USDA 110 *bll4983* (BAC50248.1), *Bradyrhizobium japonicum* USDA 110 *bll5076* (BAC50341.1), *Bradyrhizobium japonicum* USDA 110 *bll6888* (BAC52153.1), *Brucella melitensis* bv. 1 str. 16M *BMEI1305* (AAL52486.1), *Brucella melitensis* bv. 1 str. 16M *BMEI1306* (AAL52487.1), *Mesorhizobium loti* MAFF303099 *mll4029* (BAB50784.1), *Mesorhizobium loti* MAFF303099 *mll6389* (BAB52694.1), *Mesorhizobium loti* MAFF303099 *mll7738* (BAB54137.1), *Mesorhizobium loti* MAFF303099 *mlr7740* (BAB54139.1), *Mesorhizobium loti* MAFF303099 *mlr7768* (BAB54159.1), *Rhizobium leguminosarum* bv. *trifolii* WSM1325 *Rleg_1139* (ACS55434.1), *Rhizobium leguminosarum* bv. *trifolii* WSM1325 *Rleg_2312* (ACS56587.1), *Rhizobium leguminosarum* bv. *trifolii* WSM1325 *Rleg_6754* (ACS59793.1), *Sinorhizobium meliloti* Rm1021 *SMc02396* (CAC45624.1), *Sinorhizobium meliloti* Rm1021 *SMc02400* (CAC45628.1). Sequences were aligned using the MUSCLE web server (185, 186) (http://www.ebi.ac.uk/Tools/msa/muscle/) with the default settings, and then manually adjusted using MacClade version 4.08 (187) (http://macclade.org/index.html). The phylogenetic reconstruction was conducted using maximum parsimony with 1000 replicates, implemented in PAUP* version 4.0 beta 10 (188) (http://paup.csit.fsu.edu/) using the default settings, and then visualized and exported using FigTree version 1.2.3 (http://tree.bio.ed.ac.uk/software/figtree/).

2.4.4 Results

2.4.4.1 *S. meliloti* mutants resistant to ΦM12 and ΦN3 map to the same gene.

Mapping mutations which confer resistance to a transducing phage presents a unique challenge because the ability to map the mutation by phage transduction is precluded by the very mutations being analyzed. To overcome this, we acquired *two* commonly-used transducing phages for *S. meliloti*, ΦM12 (215) and ΦN3 (216), and used each one to map the resistance mutations for the other. Fortuitously, several ΦM12-resistant mutants remained susceptible to ΦN3, and vice versa. Resistance mutations for both transducing phages mapped to the chromosomally encoded gene *SMc02396* (**Figure 2.4-1A**). *SMc02396* encodes a putative outer membrane porin that consists of a 16-pass transmembrane β-barrel and like many bacterial porins, it probably forms a trimer (221). Porin homotrimers are the most common form, but heterotrimers can occur (222). Due to the similarity of *SMc02396* to *ropA* (rhizobial outer membrane

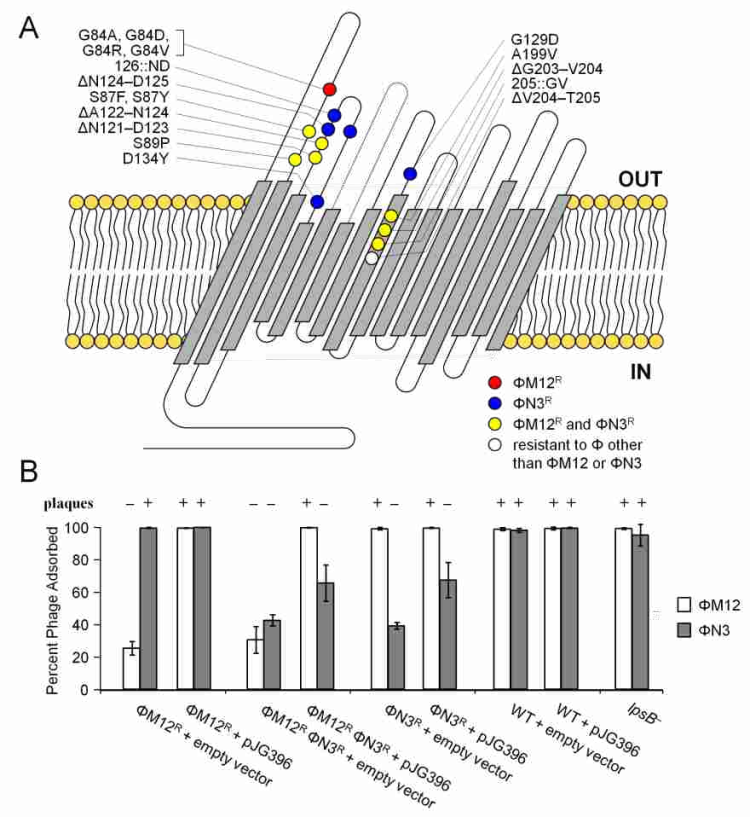

Figure 2.4-1. *ropA1* is the site of phage adsorption for ΦM12 and ΦN3. A. All mutations that conferred resistance to ΦM12 (red), ΦN3 (blue), both (yellow), or other *S. meliloti* phages in the collection (white) were found in *SMc02396* (*ropA1*), a 16-pass outer membrane β-barrel porin. The majority of mutations were found in predicted extracellular loops. The loop predicted to form the porin eyelet is indicated in gray. **B.** A ΦM12^{R} mutant (*ropA1*^{G84A}), a ΦM12^{R}ΦN3^{R} mutant (*ropA1*$^{\Delta G203–V204}$), and a ΦN3^{R} mutant (*ropA1*$^{\Delta N124–D125}$) were tested for phage adsorption (n = 3). Error bars represent SD. The ability of a given strain to form plaques is indicated below.

protein A) in *Rhizobium leguminosarum* bv. *viciae* 248 (223), we propose that *SMc02396* be renamed *ropA1*. Approximately 2 Kb downstream of *ropA1* is a duplicate gene, *SMc02400*, which also encodes an outer membrane porin. Based on its similarity to *ropA1* we propose *SMc02400* be renamed *ropA2*. Despite this similarity, no phage resistance alleles have mapped to *ropA2*. **Figure 2.4-1A** describes all resistance alleles of *ropA1* that have been sequenced to date. Some *ropA1* alleles confer resistance to ΦM12, some to ΦN3, and some confer simultaneous resistance to ΦM12 and ΦN3. It is interesting to note that all phage resistance mutations in *ropA1* are either point mutations or small insertions/deletions that do not alter the frame of the coding region. Large insertions/deletions, frameshift mutations, or nonsense mutations have not been observed.

2.4.4.2 RopA1 is the site of phage adsorption during infection.

To test whether ΦM12 and ΦN3 bind to RopA1, we measured adsorption of both phages to *ropA1* mutants that were resistant to ΦM12 only (*ropA1*^{G84A}), resistant to ΦN3 only (*ropA1*$^{\Delta N124–D125}$), or resistant to both (*ropA1*$^{\Delta G203–V204}$) in the presence of an empty vector (pRF771) or a complementing copy of *ropA1* (pJG396) (**Figure 2.4-1B**). In the case of ΦM12, expression of wildtype *ropA1* from pJG396 completely restored ΦM12 adsorption ($p<0.001$). However, we observed only slight restoration of ΦN3 adsorption upon reintroduction of *ropA1* ($p<0.1$). In the presence of an allele that simultaneously confers resistance to ΦM12 and ΦN3, pJG396 is more effective for restoring adsorption of ΦM12 than of ΦN3. In a plaquing assay, pJG396 restored ability to form plaques in *ropA1* mutant backgrounds resistant to ΦM12, but not for backgrounds resistant to ΦN3 (**Figure 2.4-1B**). Even when a given mutation conferred resistance to both phages, pJG396 restored plaquing by ΦM12 but not by ΦN3. Considering the possibility that resistance to ΦN3 acts dominantly, we then cloned *ropA1*$^{\Delta N124–D125}$ and

Table 2.4-1. *ropA1* and/or LPS are involved in phage infection for all phages tested. Various mutations in *ropA1* can lead to resistance to a variety of *S. meliloti* bacteriophages, including ΦM12 and ΦN3. An *lpsB* knockout, which alters the LPS core (see text), confers resistance to six of the ten phages. A knockout of a downstream paralog, *ropA2*, did not alter phage susceptibility.

Phage tested	WT	G84A	G84D	G84R	G84V	S87F	S87Y	S89P	ΔN121–D123	ΔA122–N124	ΔG203–V204	205::GV	ΔN124–D125	126::ND	G129D	D134Y	A199V	ΔV204–T205	*lpsB*⁻	*ropA2*⁻
ΦM12	S	R	R	R	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	S
ΦM7	S	R	R	R	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	S
ΦM19	S	R	R	R	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	S
ΦN3	S	S	S	S	S	R	R	R	R	R	R	R	R	R	R	R	R	S	S	S
ΦM1	S	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S
ΦM6	S	R	R	R	R	R	R	R	R	R	R	R	S	S	R	R	R	R	R	S
ΦM5	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S
ΦM9	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	R	S
ΦM10	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S
ΦM14	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S

ropA1$^{\Delta G203–V204}$ into pRF771 and introduced them into wild-type *S. meliloti* Rm1021. Ectopic expression of these resistant forms of RopA1 did not prevent ΦN3 from forming plaques on the transformed strains (data not shown), suggesting they are not dominant. To test whether *ropA1* expression level influences its ability to function as a receptor for ΦN3, we re-cloned *ropA1* and included several kilobases of upstream and downstream sequence. This fragment was cloned into pRK7813 (224) in both directions. Both forward- and reverse-orientation clones were mobilized into several ΦN3-resistant backgrounds and tested for susceptibility to ΦM12 and ΦN3. Consistent with previous tests, ΦM12 formed plaques on all of the resulting strains, but ΦN3 did not form plaques on any (data not shown).

2.4.4.3 RopA1 and/or LPS are involved in phage infection for all phages tested.

In addition to ΦM12 and ΦN3, we have acquired eight other *S. meliloti* phages from diverse sources (see **Table S5.2-20**). To test whether the requirement for *ropA1* was unique to

ΦM12 and ΦN3, or whether it was a general requirement for more phages in our collection, we tested all of our mutant strains against every phage (**Table 2.4-1**). Since LPS has previously been reported as a receptor for some of the phages in this collection (64), we also included an *lpsB*⁻ mutant. LpsB is a mannosyl/glucosyl transferase that may have a role both in incorporating GDP-mannose into Kdo_2-lipid IV_A as well as constructing the LPS core using ADP- or UDP-glucose (225, 226). Disruption results in drastic alteration of the LPS core in *S. meliloti* (64) but does not prevent attachment of the O antigen (65). Two out of ten phages required LPS only (ΦM10 and ΦM14), four out of ten required *ropA1* only (ΦM7, ΦM12, ΦM19, and ΦN3), and four out of ten required both (ΦM1, ΦM5, ΦM6, and ΦM9). The latter four probably use both LPS and RopA1 as co-receptors. We also considered the possibility that the downstream paralog, *ropA2*, might form a heterotrimer with *ropA1*, so we tested a knockout of *ropA2* for phage susceptibility. None of the phages tested required *ropA2* (**Table 2.4-1**).

2.4.4.4 *ropA1* appears to be essential for viability in *S. meliloti*.

Mutations in *ropA1* that conferred resistance to bacteriophages were always point mutations or insertions/deletions that were multiples of three base pairs (**Table 2.4-1**), suggesting that *ropA1* null alleles are not tolerated. Furthermore, a *ropA1* homologue in *Brucella melitensis*, *omp2b*, was reported to be essential (227), though no experimental evidence was provided. To test whether *ropA1* might be essential for viability, we first made several failed attempts to make an in-frame deletion of *ropA1* using the pJQ200sk *sacB* vector (228). Even with the partially complementing plasmid pJG396 (described above), deletion of the chromosomal copy or *ropA1* was not possible (data not shown). We then resorted to targeting the disruption of *ropA1* by internal fragment (single-crossover) disruption. This experiment was performed with thorough controls: three different *ropA1* internal fragments were tried, as well as seven arbitrarily chosen

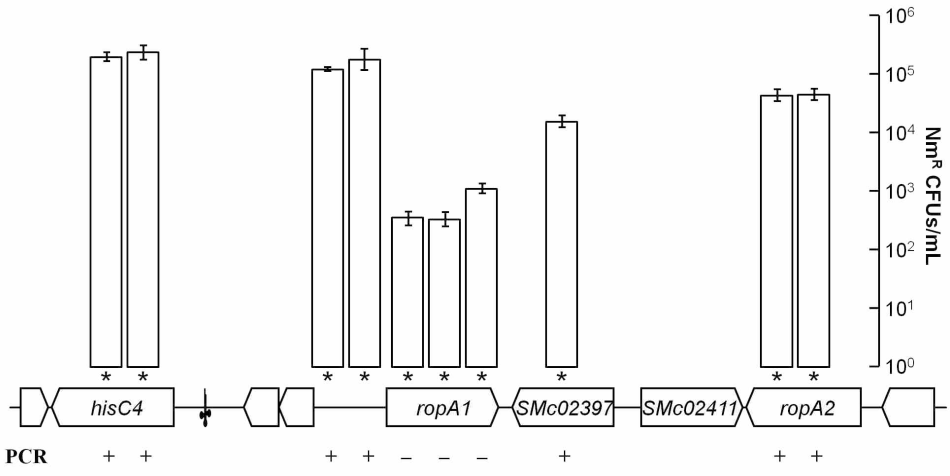

Figure 2.4-2. *ropA1* appears to be essential in *S. meliloti* but not *ropA2*. To determine whether *ropA1* or *ropA2* was essential, we designed disruption plasmids for ten loci (*) in the local genetic neighborhood. Integration, conferring resistance to neomycin, was observed for all loci but very rarely for the three targets in *ropA1* (n = 9). Error bars represent SEM. A subsequent PCR check (n = 4) determined that in the few colonies obtained for *ropA1* insertions that the integration plasmids were not inserted into *ropA1* but had non-specifically integrated elsewhere.

regions, upstream and downstream of the *ropA1* gene, that were not predicted to be essential (**Figure 2.4-2**). For these ten plasmid integration targets, PCR-based tests were designed to confirm that the intended integration events had occurred. All disruptions outside of *ropA1* successfully occurred, but no insertions in *ropA1* could be detected. This indicates that *ropA1* disruption leads to inviable cells.

2.4.4.5 *ropA1* orthologs in other Rhizobiales show evidence of frequent recent duplication.

The downstream ortholog of *ropA1* in *S. meliloti*, *ropA2*, shares 78.4% identity with *ropA1* at the amino acid level. This, combined with the fact that similar duplications have been reported for other Rhizobiales (229, 230), prompted us to investigate whether these duplication events were of ancient origin or whether they had occurred independently in multiple lineages. A phylogenetic comparison of various representative organisms in the Rhizobiales (**Figure 2.4-3A**) indicated that *ropA1* homologs were almost always most closely related to duplicates within the same genus rather than orthologs in other genera. This observation points to some selective

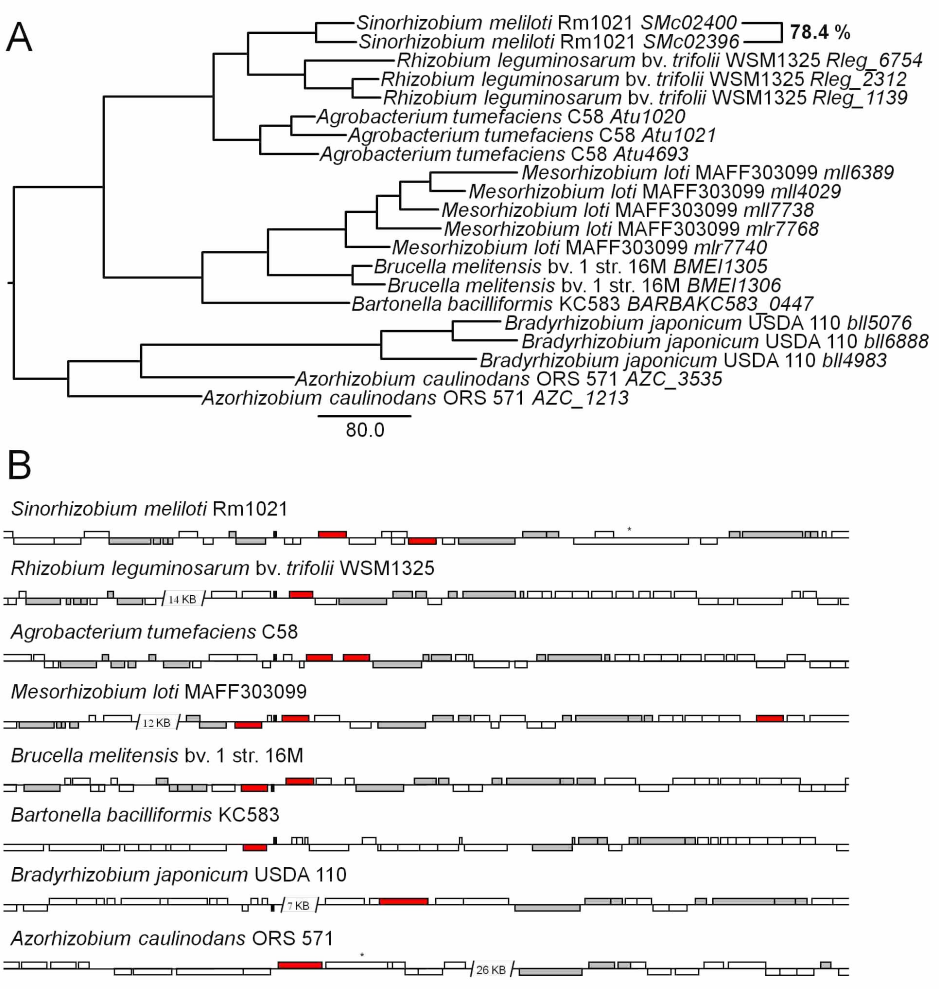

Figure 2.4-3. *ropA1* orthologues show evidence of multiple recent duplication events. A. Phylogenetic analysis of *ropA* homologues in various representative Rhizobiales species suggests multiple recent duplication events always within a species. **B.** Duplication of *ropA* homologues (red) frequently occurs in the vicinity of a tRNA-serine (black). Other syntenous genes are indicated in gray. Nearby unrelated porins are indicated with an asterisk (*).

pressure for *ropA* orthologs in many α-proteobacterial genera to independently duplicate. Considering that *S. meliloti ropA1* and *ropA2* are not functionally identical, these duplication events may give rise to functional diversification of *ropA* paralogs.

Given that *ropA1* and *ropA2* are so close together spatially, we performed a genomic alignment of *S. meliloti* Rm1021 with the same organisms used in the phylogenetic analysis (**Figure 2.4-3B**). The alignment confirmed that at least one copy of *ropA* lies in a conserved position in the genome of the various organisms as evidenced by the conservation of synteny

with certain genes both upstream (*amn* and *hisC*) and downstream (*slt*, *dapA*, *smpB*, *rpoZ*, and *relA*). Also of note is the presence in many strains of a tRNA-Ser nearby. In half of the strains examined a second copy of *ropA* was found nearby and in one case (*Mesorhizobium loti* MAFF303099) there was even a third copy within a few kilobases. An examination of other sequenced Rhizobiales genomes (e.g. *Bradyrhizobium* sp. BTAi1, *Nitrobacter hamburgensis* X14, *Ochrobactrum anthropi* ATCC 49188, *Parvibaculum lamentivorans* DS-1, *Pseudovibrio* sp. FO-BEG1, *Xanthobacter autotrophicus* Py2, etc.) gave further evidence for one or more duplications of *ropA* at this locus. It should also be noted that in contrast to most *Rhizobium* strains, *Rhizobium leguminosarum* bv. *viciae* 248 (which could not be included in the genomic alignment since it's genome has not yet been sequenced) has two copies of *ropA* in close proximity to each other (230).

2.4.5 Discussion

RopA1 is highly expressed in free-living *S. meliloti* (231) and likely forms a major portion of the *S. meliloti* outer membrane protein population. Thus it is a convenient target for phage binding. We have shown that certain alterations in the RopA1 amino acid sequence prevent infection by eight of the ten *S. meliloti* phages tested (**Table 2.4-1**). In the case of the two transducing phages, ΦM12 and ΦN3, every phage-resistant mutant tested was mutated in *ropA1*. Additionally, the adsorption of ΦM12 and ΦN3 to various *ropA1* mutant strains was reduced (**Figure 2.4-1B**). This confirms the role of RopA1 as a receptor for these phages. Previous work in *Rhizobium leguminosarum* correlates phage resistance with a loss of an antigen (232) later identified as RopA (223). But experiments to test RopA as a susceptibility factor or receptor were not performed.

This system is unique in that both ΦM12 binding and DNA injection (as evidenced by the formation of plaques) are completely restored by plasmid-based expression of *ropA1* but for ΦN3 binding is only partially restored and DNA injection (leading to plaque formation) is not observed (**Figure 2.4-1B**). The incomplete complementation phenomenon is not allele-specific, but phenotype-specific. Additionally, the apparent lethality brought on by a *ropA1* disruption seems not to be complemented by a plasmid since repeated attempts to delete or disrupt *ropA1* in the presence of a complementing plasmid failed (data not shown). This is why our evidence for the essentiality of *ropA1* has to depend on well-controlled negative data (**Figure 2.4-2**). We cannot currently explain the mechanistic basis for this incomplete complementation phenomenon.

Only two of the phages in our panel did not exhibit any detectable requirement for RopA1 for infection (**Table 2.4-1**). LPS is also a major component of the Gram-negative

bacterial cell surface and frequently occurs as a phage receptor (213). Our observation that the *lpsB*⁻ mutant was resistant to six of the ten phages is in agreement with a previous report (64). Three of the remaining phages in that study (ΦM7, ΦM12, and ΦM19) were reported to be unaffected by any of a variety of LPS mutants, suggesting it plays no role in infection by those phages. We show here that RopA1 serves as the receptor for all three, as well as for ΦN3. Even though ΦN3 was still able to infect an *lpsB*⁻ mutant, a role for LPS in the infectivity of ΦN3 cannot be ruled out since the O antigen is still attached to lipid A in *lpsB*⁻ mutants (65) and we did not test any O antigen mutants of *S. meliloti* Rm1021 for susceptibility to ΦN3.

The impossibility of disrupting *ropA1* under laboratory conditions leads us to conclude that *ropA1* is essential for viability in *S. meliloti*. Despite the general belief that porins play a critical role in outer membrane function and stability of Gram-negative bacteria (233), there are very few instances of a porin being shown to be essential. Members of the Omp85/BamA (β-barrel assembly machine protein A) family have been shown to be essential and are responsible for the assembly and insertion of proteins into the outer membrane (234), as well as lipids and LPS (235). Homologues are found throughout Gram-negative bacteria as well as in mitochondria (236). Two genes in *S. meliloti* Rm1021 belong to the same COG (COG0729): *SMc02094* and *SMc03097*. While we cannot rule out a role for RopA1 in outer membrane biogenesis, we can conclude that it does not belong to the Omp85/BamA family of porins based on lack of sequence homology (data not shown).

With the exception of Omp85/BamA homologs, no porins are reported to be essential in *Escherichia coli* (237, 238), *Pseudomonas aeruginosa* (239), *Haemophilus influenzae* (240), or *Salmonella enterica* (241). The *omp2b* gene of *Brucella melitensis* bv. 1 str. 16M has been reported to be essential but no experimental evidence is given (227). The *porB* gene of *Neisseria*

gonorrhea has also been reported to be essential, but, again, no experimental evidence is given (242, 243). The *porB* gene is not essential in the related species, *N. meningitides* (244). Additionally, Tsx, a nucleoside-specific porin, is essential for growth of *S. enterica* bv. Typhi on minimal medium (245) but not in serovar Typhimurium or other Gram-negative bacteria (246). Since both *ropA1* in *S. meliloti* (this report) and *omp2b* in *Brucella melitensis* (227) are believed to be essential, it may be that *ropA* homologs are essential in most Rhizobiales species which possess them. One possible exception would be the single *ropA* homolog in *Bartonella henselae*, *omp43*, which has been successfully disrupted (247). Homology-based searches of sequence databases do not suggest a specific function for RopA1 or why it is essential.

Bacteroids are differentiated, nitrogen-fixing forms of rhizobia that occupy host cells within the root nodule. Root nodules can be broadly classified as determinate or indeterminate based on whether the nodule has a persistent apical meristem. Bacteroids in determinate nodules can de-differentiate upon release from the nodule, but bacteroids in indeterminate nodules are terminally differentiated (6). Both *ropA1* and *ropA2* of *S. meliloti* are highly expressed in free-living conditions (231) but strongly down-regulated in bacteroids of *M. truncatula* (248), a legume which produces indeterminate nodules. Down-regulation of *ropA* and *ropA2* in *R. leguminosarum* has been observed for several hosts that form indeterminate nodules (pea, broadbean, vetch, clover), but in a host that forms determinate nodules (common bean) neither is down-regulated (249). The remarkable correlation between cells that are competent for proliferation and the expression of *ropA1* (**Figure 2.4-2**) suggests that its down-regulation in bacteroids could be involved in or be a side effect of the terminal nature of bacteroid differentiation in hosts that form indeterminate nodules.

The frequent occurrence of *ropA1* duplication at a particular genetic locus (**Figure 2.4-3B**) suggests a certain plasticity in this portion of Rhizobiales genomes. Acquisition, loss, or duplication of genes may be due to the insertion and incorrect excision of phage genomes (250). An examination of the genomes of sequenced *S. meliloti* strains AK83 and Rm41 revealed the presence of two independent prophages which have inserted into the tRNA-serine just upstream of *ropA1* (**Figure S5.1-12**). Further evidence includes the presence of many elements of potentially external origin, including transposable elements, prophages, integrases and recombinases, pilus assembly proteins, and toxin–antitoxin pairs (**Figure S5.1-12**). TA pairs are not elements of external origin *per se*, but in addition to being found in certain phage genomes (e.g. the *phd–doc* system of phage P1 (251)), they have also been implicated in phage immunity (252–254). Pilus proteins can be encoded on phage genomes (255), though they may be erroneously annotated capsid proteins (256, 257). Taken all together, the data suggest that *ropA1* resides in a genetic hotspot that may be a target for elements of external origin. Indeed, in a recent multi-genome analysis of *S. meliloti* and *S. medicae* the authors concluded that *ropA1* was the only chromosomal gene that showed evidence of horizontal transfer between the two species (258). This could be due to its proximity to the tRNA-serine since tRNAs are common integration sites for phages (259). Given that outer membrane porins are always going to be candidates for phage adsorption, if at least one copy of *ropA* is essential in each organism, duplication may provide a means to maintain the essential function of the porin while evolving resistance to phage. However, nearly all of the phages we tested required the essential porin, RopA1, and none required the non-essential duplicate, RopA2. The fact that these physically-proximate copies are always more similar to each other than to homologs in other species (**Figure 2.4-3A**) suggests that these are the result of a recent duplication. As an alternative to

recent duplication, juxtaposing two genes with such similar sequences could allow for gene conversion, which would drive sequence diversification (260), but retain within-species identity even in the event of an ancient duplication event.

2.4.6 Future Directions

We were unable to disrupt *ropA1* in the presence of two different covering plasmids, thus our proof of its essentiality is based on well-controlled, negative data (**Figure 2.4-2**). But there remain several approaches to study the essentiality of *ropA1*. Rather than restore *ropA1* on a covering plasmid, a complementing copy can be delivered to the chromosome on a mini-Tn*5*. This provides a distinct advantage since it would be single-copy, thus eliminating deleterious effects on the membrane from overexpression of a porin. A biochemical approach would involve placing *ropA1* under the control of an IPTG-inducible promoter and then depleting IPTG from the culture medium (234).

We have identified RopA1 as the bacterial receptor for most of the phages in our panel. Yet when these phages were tested against a collection of wild *S. meliloti* isolates (122), they exhibited considerable diversity in their host ranges (**Table S5.2-22**). Some of this could be due to variation in the LPS species (64) or *ropA1* alleles of individual strains (**Table 2.4-1**). The LPS from the strains in our collection can be solubilized and analyzed by SDS-PAGE followed by silver staining. Differences in LPS species could be correlated with phage host ranges. Likewise, by determining the sequence of *ropA1* from the strains in our collection a correlation with host range might be possible.

Aside from an incompatible form of the cell surface receptor, there are several other mechanisms of phage resistance: restriction modification, lysogeny/prophage, CRISPR, obstruction of the receptor (including the use of extracellular matrix, lipoproteins, and competitive inhibitors), blocking uptake of phage DNA, or interference with phage reproduction/maturation/lysis (including programmed cell death), etc (210, 214). The latter is unlikely in *S. meliloti* since no CRISPR loci have been identified in any of the genomes

sequenced to date (261). The patterns in (**Table S5.2-22**) are likely dependent on the restriction systems of *S. meliloti* Rm1021, which is used to maintain our phages. Resistance due to DNA restriction can be tested by heat-inactivation of restriction enzymes (262) or by testing for altered host range after maintenance in a strain other than Rm1021. Resistance as a result of lysogeny can be tested for by attempting induction using UV irradiation or mitomycin C.

Most *ropA1* mutant alleles recovered so far were specifically selected for their resistance to ΦM12 or ΦN3 (**Table 2.4-1**). These alleles were then tested against the other phages in our collection. Given the variability in host ranges for these phage, mutations which confer resistance to other individual phages or combinations of phages, but not to ΦM12 or ΦN3, likely can be recovered. Additionally, spontaneous phage mutants which have recovered the ability to infect strains mutated in *ropA1* can be sequenced alongside their wildtype parents to determine their receptor-binding proteins.

Porins retain certain conserved features: the "latching loop", L2, serves to stabilize the center of the trimer and forms contacts with L1, L2, and L4 on the other subunits; the "eyelet", L3, acts as a gate (217). Most mutations in RopA1 mapped to L1 and L2 (**Figure 2.4-1**), suggesting a possible disruption of oligomerization in addition to mutating contact points for the phage tail proteins. 2D gel electrophoresis of RopA1 mutants would reveal whether or not multimeric forms of RopA1 can still be detected in these instances. The remainder of phage resistance mutations mapped to L4. It was observed previously that ΦT5 binding to FhuA in *E coli* converted it to an open conformation (263). Thus phage binding to L4 of RopA1 may alter porin structure and force open the gate in a similar manner. To test whether this is true, purified RopA1 could be embedded in a membrane bilayer and tested for changes in the electrophysiological characteristics of the channels in response to phage binding.

The majority of differences between RopA1 and RopA2 of *S. meliloti* Rm1021 lie in the extracellular loops (**Figure S5.1-13**). To test whether any given loop plays a role in the essential function of RopA1, loops of RopA2 could be substituted for loops of RopA1. However, analysis using the PROVEAN tool (264, 265) (http://provean.jcvi.org) suggests that some of the *ropA1* mutations already recovered are detrimental to its function (**Table S5.2-23**). Thus a less drastic approach would involve alanine scanning through the length of each loop.

Given the essentiality of *ropA1* for cell survival (**Figure 2.4-2**), knockout and deletion strains cannot be generated. Therefore it is unclear whether *ropA1* is required for symbiosis between *S. meliloti* and *M. truncatula* A17 aside from its essentiality for cell survival. Mutants that have been isolated so far are under two constraints: to alter phage binding and to retain the essential function of RopA1. Therefore our method for isolating *ropA1* mutants is unlikely to produce alleles that fulfill a third requirement of altering any putative symbiotic function. Nongenetic tests may be required to identify the symbiotic role, if any, of RopA1.

A central question is the extent to which *ropA* homologs (see **Table S5.2-19**) are essential in other Rhizobiales. Knockout experiments, such as the ones we conducted (**Figure 2.4-2**), could be performed in other members of the order as well as other strains of *S. meliloti*. Several defects have been reported for *S. meliloti* Rm1021 (266). Thus the possibility that *ropA1* essentiality in *S. meliloti* Rm1021 is a side-effect of its long domestication as a laboratory strain cannot be ruled out.

Despite the fact that *ropA1* is severely down-regulated in nodules, overexpression of *ropA1* did not affect symbiosis on *M. praecox* or *M. truncatula* cv. A17 (**Figure S5.1-14**). However, *ropA1* may still play some role in symbiosis in addition to its essentiality for cell viability. A *ropA* homolog in the related mammalian pathogen *Brucella abortus*, *omp2b*, has

been implicated in host cell invasion (230, 267–269) and in inhibiting apoptosis in host cells (227). Similarly *ropA1* may play a role in the uptake of rhizobia by plant cells or inhibition of an immune response during the initial stages of infection and then subsequently be down-regulated as the rhizobia differentiate into bacteroids.

A *B. japonicum* USDA 110 *ropA* homolog, *bll5076*, is expressed in bacteroids, but not until late in nodulation (270). Furthermore, a lupine *Bradyrhizobium* sp. *ropA* homolog is highly up-regulated in response to glyphosate-induced plant stress (271). This suggests that the appearance of *ropA* in bacteroids could be involved in or be a side effect of nodule senescence. This could be investigated by measuring the buildup of leghemolobin breakdown byproducts in nodules inoculated with wildtype and overexpressing strains. We have also tentatively proposed a role for the downregulation of *ropA1* in the terminal nature of bacteroid differentiation in indeterminate hosts. An initial test of this hypothesis would involve counting the number of PFUs from nodules inoculated with a *ropA1* overexpressing strain compared to a wildtype control.

Many porins are known to function as adhesins (233), including several RopA homologs: Omp43 in *Bartonella henselae* (272–274), Omp2b (275) and possibly Omp2a (276) of *Brucella abortus*. If RopA1 acts as an adhesin, one or both may play a role in adhesion to the root surface during symbiosis. However, *S. meliloti ropA1* lacks the RGD motif predicted to mediate binding for many porins (277). However, a role for RopA1 binding to host cell lectins or other host surface molecules cannot be ruled out at this point.

Two copies of *ropA* in the *M. loti* R7A symbiosis island are predicted to be regulated by the transcription factor NifA (278), which controls expression of the production of nitrogenase. *ropA1* in *S. meliloti* Rm1021 has a σ^{54} promoter (279) but it is unusual in that the NifA binding

site is downstream of the –24/–12 sequences (280) (**Figure S5.1-15**). Thus it could be that NifA negatively regulates *ropA1* in the nodule. *ropA2* does not have a σ^{54} promoter but does have an NtrC binding site. Both *ropA1* and *ropA2* possess a conserved upstream sequence (CTCGGGCCGTC), which occurs as an inverted repeat for *ropA2*, which may be a binding site for an unidentified regulator. Further analysis of the regulatory regions upstream of *ropA1* and *ropA2* could provide insights into their cellular roles.

We proposed a model where *ropA1* and *ropA2* could engage in gene conversion. Evidence for this model could be sought by sequencing the *ropA1* and *ropA2* genes for multiple strains to see if intragenic modules, particularly in the extracellular loops, are found to switch between the two paralogs.

3 CONCLUSIONS

While not essential for the rhizobium–legume symbiosis, accessory plasmids can modulate the exchange of signals between organisms and thus affect the outcome of said symbiosis. I have identified four unique accessory plasmids which restrict host range to varying degrees by at least two mechanisms. The exact nature of these mechanisms is still under investigation. In addition to restricting host range, two of these plasmids have been shown to improve the ability of the strain which harbors them to occupy root nodules. This combination may allow the rhizobia to cheat, accessing fixed carbon from the plant without providing fixed nitrogen in return.

The development of a *cre–loxP*-based system for making large deletions has made possible the dissection of these host range plasmids. Additionally, this *cre–loxP* system can be used to cure large, stable replicons such as the pSymB and pSymA megaplasmids of *S. meliloti*. Recent studies and bioinformatic analyses suggest that, with a few modifications, both megaplasmids can be cured. Not only will this give us unique insights into the biology of *S. meliloti*, but it may also convert the *S. meliloti* chromosome to a symbiotic 'blank slate'. Sym plasmids from other organisms could then be introduced to see if the host range of *S. meliloti* could be altered. Steps have already been taken to perform this experiment with the Sym plasmids from *S. fredii* strains NGR234 and USDA 257, which are symbionts of soybean (*G. max*).

The biosynthetic genes for the majority of symbiotically active molecules are encoded on these megaplasmids and both have been shown previously to be transmissible. The exact extent to which this transmission takes place in nature has not been determined, but it could impact the host range of recipient strains. Analysis of local populations of *Sinorhizobium* could give some

insight into this process. A pilot phylogenetic study identified several loci on the chromosome and on the megaplasmids which could be used to compare descent for all three replicons individually and together.

S. meliloti is itself a host to numerous bacteriophages. The highly-expressed outer membrane porin, RopA1, is the receptor or co-receptor for eight of the ten phages in our collection. The presence of the correct form of RopA1 is required for phage adsorption and subsequent infection, thus dictating, at least in part, the host range of the corresponding bacteriophages. RopA1 also appears to be essential for viability in *S. meliloti* Rm1021 under laboratory conditions.

4 CITATIONS

4.1 References

1. **Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Promé JC, Dénarié J.** 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature **344:**781–784.
2. **Schrire B, Lewis G, Lavin M.** 2005. Biogeography of the Leguminosae, p. 21–25. *In* Lewis G, Schrire B, Mackinder B, Lock M (ed.), Legumes of the world. Royal Botanic Gardens, Kew, Kew, England.
3. **Schrire BD, Lavin M, Lewis GP.** 2005, p 375–422. Plant diversity and complexity patterns—local, regional and global dimensions., The Royal Danish Academy of Sciences and Letters in Copenhagen, Denmark.
4. **D'Haeze W, Holsters M.** 2002. Nod factor structures, responses, and perception during initiation of nodule development. Glycobiology **12:**79R–105R.
5. **Struyve M, Moons M, Tommassen J.** 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J Mol Biol **218:**141–148.
6. **Hirsch AM.** 1992. Tansley Review No. 40. Developmental biology of legume nodulation. New Phytologist **122:**211–237.
7. **Rogel MA, Ormeño-Orrillo E, Martínez-Romero E.** 2011. Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst Appl Microbiol **34:**96–104.
8. **Doyle JJ, Luckow MA.** 2003. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. Plant Physiol **131:**900–910.
9. **Sprent JI.** 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. New Phytol **174:**11–25.

10. **Willems A.** 2006. The taxonomy of rhizobia: an overview. Plant and Soil **287:**3–14.

11. **Rees DC, Akif Tezcan F, Haynes CA, Walton MY, Andrade S, Einsle O, Howard JB.** 2005. Structural basis of biological nitrogen fixation. Philos Transact A Math Phys Eng Sci **363:**971–984.

12. **Zahran HH.** 1999. *Rhizobium*–legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. Microbiol Mol Biol Rev **63:**968–989.

13. **Gage DJ.** 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev **68:**280–300.

14. **Cook DR.** 1999. *Medicago truncatula*—a model in the making! Curr Opin Plant Biol **2:**301–304.

15. **Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, Weidner S, Wells DH, Wong K, Yeh KC, Batut J.** 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668–672.

16. **Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng JF, Goodwin L, Pitluck S, Land M, Hauser L, Woyke T, Mikhailova N, Ivanova N, Daligault H, Bruce D, Detter C, Tapia R, Han C, Teshima H, Mocali S, Bazzicalupo**

M, Biondi EG. 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. BMC Genomics **12:**1–15.

17. **Schneiker-Bekel S, Wibberg D, Bekel T, Blom J, Linke B, Neuweger H, Stiens M, Vorholter FJ, Weidner S, Goesmann A, Puhler A, Schluter A.** 2011. The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. J Biotechnol **155:**20–33.

18. **Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, Van de Peer Y, Proost S, Cook DR, Meyers BC, Spannagl M, Cheung F, De Mita S, Krishnakumar V, Gundlach H, Zhou S, Mudge J, Bharti AK, Murray JD, Naoumkina MA, Rosen B, Silverstein KA, Tang H, Rombauts S, Zhao PX, Zhou P, Barbe V, Bardou P, Bechner M, Bellec A, Berger A, Bergès H, Bidwell S, Bisseling T, Choisne N, Couloux A, Denny R, Deshpande S, Dai X, Doyle JJ, Dudez AM, Farmer AD, Fouteau S, Franken C, Gibelin C, Gish J, Goldstein S, González AJ, Green PJ, Hallab A, Hartog M, Hua A, Humphray SJ, Jeong DH, Jing Y, Jöcker A, Kenton SM, Kim DJ, Klee K, Lai H, Lang C, Lin S, Macmil SL, Magdelenat G, Matthews L, McCorrison J, Monaghan EL, Mun JH, Najar FZ, Nicholson C, Noirot C, O'Bleness M, Paule CR, Poulain J, Prion F, Qin B, Qu C, Retzel EF, Riddle C, Sallet E, Samain S, Samson N, Sanders I, Saurat O, Scarpelli C, Schiex T, Segurens B, Severin AJ, Sherrier DJ, Shi R, Sims S, Singer SR, Sinharoy S, Sterck L, Viollet A, Wang BB, Wang K, Wang M, Wang X, Warfsmann J, Weissenbach J, White DD, White JD, Wiley GB, Wincker P, Xing Y, Yang L, Yao Z, Ying F, Zhai J, Zhou L, Zuber A, Dénarié J, Dixon RA, May GD, Schwartz DC,**

Rogers J, Quétier F, Town CD, Roe BA. 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* **480:**520–524.

19. **Vasse J, de Billy F, Camut S, Truchet G.** 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol **172:**4295–4306.
20. **van Rhijn P, Vanderleyden J.** 1995. The *Rhizobium*–plant symbiosis. Microbiol Rev **59:**124–142.
21. **Maxwell CA, Hartwig UA, Joseph CM, Phillips DA.** 1989. A chalcone and two related flavonoids released from alfalfa roots induce *nod* genes of *Rhizobium meliloti*. Plant Physiol **91:**842–847.
22. **Hartwig UA, Maxwell CA, Joseph CM, Phillips DA.** 1990. Chrysoeriol and luteolin released from alfalfa seeds induce *nod* genes in *Rhizobium meliloti*. Plant Physiol **92:**116–122.
23. **Dakora FD, Joseph CM, Phillips DA.** 1993. Alfalfa (*Medicago sativa* L.) root exudates contain isoflavonoids in the presence of *Rhizobium meliloti*. Plant Physiol **101:**819–824.
24. **Peters NK, Frost JW, Long SR.** 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. Science **233:**977–980.
25. **Phillips DA, Joseph CM, Maxwell CA.** 1992. Trigonelline and stachydrine released from alfalfa seeds activate NodD2 protein in *Rhizobium meliloti*. Plant Physiol **99:**1526–1531.
26. **Peck MC, Fisher RF, Long SR.** 2006. Diverse flavonoids stimulate NodD1 binding to nod gene promoters in *Sinorhizobium meliloti*. J Bacteriol **188:**5417–5427.
27. **Mulligan JT, Long SR.** 1985. Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. Proc Natl Acad Sci U S A **82:**6609–6613.

28. **Honma MA, Asomaning M, Ausubel FM.** 1990. *Rhizobium meliloti nodD* genes mediate host-specific activation of *nodABC*. J Bacteriol **172:**901–911.

29. **Dénarié J, Debellé F, Promé JC.** 1996. *Rhizobium* lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annu Rev Biochem **65:**503–535.

30. **Long SR.** 1996. *Rhizobium* symbiosis: Nod factors in perspective. Plant Cell **8:**1885–1898.

31. **Bonaldi K, Gherbi H, Franche C, Bastien G, Fardoux J, Barker D, Giraud E, Cartieaux F.** 2010. The Nod factor-independent symbiotic signaling pathway: development of *Agrobacterium rhizogenes*-mediated transformation for the legume *Aeschynomene indica*. Mol Plant–Microbe Interact **23:**1537–1544.

32. **Okubo T, Fukushima S, Minamisawa K.** 2012. Evolution of *Bradyrhizobium–Aeschynomene* mutualism: living testimony of the ancient world or highly evolved state? Plant Cell Physiol **53:**2000–2007.

33. **Bonaldi K, Gargani D, Prin Y, Fardoux J, Gully D, Nouwen N, Goormachtig S, Giraud E.** 2011. Nodulation of *Aeschynomene afraspera* and *A. indica* by photosynthetic *Bradyrhizobium* sp. strain ORS285: the nod-dependent versus the nod-independent symbiotic interaction. Mol Plant–Microbe Interact **24:**1359–1371.

34. **Maillet F, Poinsot V, Andre O, Puech-Pagès V, Haouy A, Gueunier M, Cromer L, Giraudet D, Formey D, Niebel A, Martínez EA, Driguez H, Bécard G, Dénarié J.** 2011. Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. Nature **469:**58–63.

35. **Streng A, op den Camp R, Bisseling T, Geurts R.** 2011. Evolutionary origin of *Rhizobium* Nod factor signaling. Plant Signal Behav **6:**1510–1514.

36. **Oldroyd GE, Downie JA.** 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu Rev Plant Biol **59:**519–546.

37. **Foucher F, Kondorosi E.** 2000. Cell cycle regulation in the course of nodule organogenesis in *Medicago*. Plant Mol Biol **43:**773–786.

38. **Reinhold BB, Chan SY, Reuber TL, Marra A, Walker GC, Reinhold VN.** 1994. Detailed structural characterization of succinoglycan, the major exopolysaccharide of *Rhizobium meliloti* Rm1021. J Bacteriol **176:**1997–2002.

39. **Simsek S, Ojanen-Reuhs T, Stephens SB, Reuhs BL.** 2007. Strain–ecotype specificity in *Sinorhizobium meliloti–Medicago truncatula* symbiosis is correlated to succinoglycan oligosaccharide structure. *J Bacteriol* **189:**7733–7740.

40. **Leigh JA, Reed JW, Hanks JF, Hirsch AM, Walker GC.** 1987. *Rhizobium meliloti* mutants that fail to succinylate their calcofluor-binding exopolysaccharide are defective in nodule invasion. Cell **51:**579–587.

41. **Wang LX, Wang Y, Pellock B, Walker GC.** 1999. Structural characterization of the symbiotically important low-molecular-weight succinoglycan of *Sinorhizobium meliloti*. J Bacteriol **181:**6788–6796.

42. **Cheng HP, Walker GC.** 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. J Bacteriol **180:**5183–5191.

43. **Pellock BJ, Teplitski M, Boinay RP, Bauer WD, Walker GC.** 2002. A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. J Bacteriol **184:**5067–5076.

44. **Glazebrook J, Walker GC.** 1989. A novel exopolysaccharide can function in place of the calcofluor-binding exopolysaccharide in nodulation of alfalfa by *Rhizobium meliloti*. *Cell* **56:**661–672.

45. **Her GR, Glazebrook J, Walker GC, Reinhold VN.** 1990. Structural studies of a novel exopolysaccharide produced by a mutant of *Rhizobium meliloti* strain Rm1021. Carbohydr Res **198:**305–312.

46. **González JE, Reuhs BL, Walker GC.** 1996. Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. Proc Natl Acad Sci U S A **93:**8636–8641.

47. **Kereszt A, Kiss E, Reuhs BL, Carlson RW, Kondorosi Á, Putnoky P.** 1998. Novel *rkp* gene clusters of *Sinorhizobium meliloti* involved in capsular polysaccharide production and invasion of the symbiotic nodule: the *rkpK* gene encodes a UDP-glucose dehydrogenase. J Bacteriol **180:**5426–5431.

48. **Reuhs BL, Williams MN, Kim JS, Carlson RW, Côté F.** 1995. Suppression of the Fix^- phenotype of *Rhizobium meliloti exoB* mutants by *lpsZ* is correlated to a modified expression of the K polysaccharide. J Bacteriol **177:**4289–4296.

49. **Williams MN, Hollingsworth RI, Klein S, Signer ER.** 1990. The symbiotic defect of *Rhizobium meliloti* exopolysaccharide mutants is suppressed by $lpsZ^+$, a gene involved in lipopolysaccharide biosynthesis. J Bacteriol **172:**2622–2632.

50. **Fraysse N, Lindner B, Kaczynski Z, Sharypova L, Holst O, Niehaus K, Poinsot V.** 2005. *Sinorhizobium meliloti* strain 1021 produces a low-molecular-mass capsular polysaccharide that is a homopolymer of 3-deoxy-D-manno-oct-2-ulosonic acid harboring a phospholipid anchor. Glycobiology **15:**101–108.

51. **Reuhs BL, Carlson RW, Kim JS.** 1993. *Rhizobium fredii* and *Rhizobium meliloti* produce 3-deoxy-D-manno-2-octulosonic acid-containing polysaccharides that are structurally analogous to group II K antigens (capsular polysaccharides) found in *Escherichia coli*. J Bacteriol **175:**3570–3580.

52. **Reuhs BL, Geller DP, Kim JS, Fox JE, Kolli VS, Pueppke SG.** 1998. *Sinorhizobium fredii* and *Sinorhizobium meliloti* produce structurally conserved lipopolysaccharides and strain-specific K antigens. Appl Environ Microbiol **64:**4930–4938.

53. **Hisamatsu M, Amemura A, Koizumi K, Utamura T, Okada Y.** 1983. Structural Studies on Cyclic (1→2) -β-D-Glucans (Cyclosophoraoses) Produced by *Agrobacterium* and *Rhizobium*. Carbohydrate Research **121:**31–40.

54. **Miller KJ, Gore RS, Benesi AJ.** 1988. Phosphoglycerol substituents present on the cyclic β-(1→2)-glucans of Rhizobium meliloti 1021 are derived from phosphatidylglycerol. J Bacteriol **170:**4569–4575.

55. **Breedveld MW, Miller KJ.** 1994. Cyclic β-glucans of members of the family Rhizobiaceae. Microbiol Rev **58:**145–161.

56. **Breedveld MW, Miller KJ.** 1995. Synthesis of glycerophosphorylated cyclic (1→2) -β-glucans in *Rhizobium meliloti* strain 1021 after osmotic shock. Microbiology **141:**583–588.

57. **Dylan T, Ielpi L, Stanfield S, Kashyap L, Douglas C, Yanofsky M, Nester E, Helinski DR, Ditta G.** 1986. *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A **83:**4403–4407.

58. **Spaink HP.** 2000. Root nodulation and infection factors produced by rhizobial bacteria. Annu Rev Microbiol **54:**257–288.

59. **Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset AE, Barloy-Hubler F, Galibert F, Kondorosi Á, Kondorosi E.** 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*–legume symbiosis. Proc Natl Acad Sci U S A **103:**5230–5235.

60. **Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi Á, Kondorosi E.** 2003. A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. Plant Physiol **132:**161–173.

61. **Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaître B, Alunni B, Bourge M, Kucho K, Abe M, Kereszt A, Maroti G, Uchiumi T, Kondorosi E, Mergaert P.** 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science **327:**1122–1126.

62. **Udvardi MK, Day DA.** 1997. Metabolite transport across symbiotic membranes of legume nodules. Annu Rev Plant Physiol Plant Mol Biol **48:**493–523.

63. **Tellstrom V, Usadel B, Thimm O, Stitt M, Kuster H, Niehaus K.** 2007. The lipopolysaccharide of *Sinorhizobium meliloti* suppresses defense-associated gene expression in cell cultures of the host plant *Medicago truncatula*. Plant Physiol **143:**825–837.

64. **Campbell GR, Sharypova LA, Scheidle H, Jones KM, Niehaus K, Becker A, Walker GC.** 2003. Striking complexity of lipopolysaccharide defects in a collection of *Sinorhizobium meliloti* mutants. J Bacteriol **185:**3853–3862.

65. **Campbell GR, Reuhs BL, Walker GC.** 2002. Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. Proc Natl Acad Sci U S A **99:**3938–3943.

66. **Ferguson GP, Roop RM, 2nd, Walker GC.** 2002. Deficiency of a *Sinorhizobium meliloti* BacA mutant in alfalfa symbiosis correlates with alteration of the cell envelope. J Bacteriol **184:**5625–5632.

67. **Karunakaran R, Haag AF, East AK, Ramachandran VK, Prell J, James EK, Scocchi M, Ferguson GP, Poole PS.** 2010. BacA is essential for bacteroid development in nodules of galegoid, but not phaseoloid, legumes. J Bacteriol **192:**2920–2928.

68. **Halbleib CM, Ludden PW.** 2000. Regulation of biological nitrogen fixation. J Nutr **130:**1081–1084.

69. **Day DA, Poole PS, Tyerman SD, Rosendahl L.** 2001. Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. Cell Mol Life Sci **58:**61–71.

70. **Earl CD, Ronson CW, Ausubel FM.** 1987. Genetic and structural analysis of the *Rhizobium meliloti fixA*, *fixB*, *fixC*, and *fixX* genes. J Bacteriol **169:**1127–1136.

71. **Batut J, Boistard P.** 1994. Oxygen control in *Rhizobium*. Antonie Van Leeuwenhoek **66:**129–150.

72. **Masson-Boivin C, Giraud E, Perret X, Batut J.** 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol **17:**458–466.

73. **Small E.** 1989. Polythetic generic separation in tribe Trifolieae subtribe Trigonellinae (Leguminosae). Can. J. Bot. **67:**1480–1492.

74. **Jordan DC.** 1984. Family III. *Rhizobiaceae*., p. 234–242. *In* Krieg NR, Holt JG (ed.), *Bergey's Manual of Systematic Bacteriology*, vol. I. Williams and Wilkins, Co., Baltimore.

75. **Zurdo-Piñeiro JL, García-Fraile P, Rivas R, Peix A, León-Barrios M, Willems A, Mateos PF, Martínez-Molina E, Velázquez E, van Berkum P.** 2009. Rhizobia from Lanzarote, the Canary Islands, that nodulate *Phaseolus vulgaris* have characteristics in common with *Sinorhizobium meliloti* isolates from mainland Spain. Appl Environ Microbiol **75:**2354–2359.

76. **Mnasri B, Badri Y, Saidi S, de Lajudie P, Mhamdi R.** 2009. Symbiotic diversity of *Ensifer meliloti* strains recovered from various legume species in Tunisia. Syst Appl Microbiol **32:**583–592.

77. **Torres Tejerizo G, Del Papa MF, Soria-Diaz ME, Draghi W, Lozano M, Giusti Mde L, Manyani H, Megías M, Gil Serrano A, Puhler A, Niehaus K, Lagares A, Pistorio M.** 2011. The nodulation of alfalfa by the acid-tolerant *Rhizobium* sp. strain LPU83 does not require sulfated forms of lipochitooligosaccharide nodulation signals. J Bacteriol **193:**30–39.

78. **Mercado-Blanco J, Toro N.** 1996. Plasmids in rhizobia: The role of nonsymbiotic plasmids. Molecular Plant–Microbe Interactions **9:**535–545.

79. **Murphy PJ, Wexler W, Grzemski W, Rao JP, Gordon D.** 1995. Rhizopines—Their role in symbiosis and competition. Soil Biology & Biochemistry **27:**525–529.

80. **Ma W, Charles TC, Glick BR.** 2004. Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Sinorhizobium meliloti* increases its ability to nodulate alfalfa. Appl Environ Microbiol **70:**5891–5897.

81. **Hynes MF, McGregor NF.** 1990. Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol Microbiol **4:**567–574.

82. **Brom S, Garcia de los Santos A, Stepkowsky T, Flores M, Davila G, Romero D, Palacios R.** 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. J Bacteriol **174:**5183–5189.

83. **Barreto EF, Straliotto R, Baldani JI.** 2012. Curing of a non-symbiotic plasmid of the *Rhizobium tropici* strain CIAT 899 affected nodule occupancy and competitiveness of the bacteria in symbiosis with common beans. European Journal of Soil Biology **50:**91–96.

84. **Martínez-Romero E, Rosenblueth M.** 1990. Increased bean (*Phaseolus vulgaris* L.) nodulation competitiveness of genetically modified *Rhizobium* strains. Appl Environ Microbiol **56:**2384–2388.

85. **Gigova L, Petrova N, Vassileva V, Ignatov G.** 1997. Free-living and symbiotic characteristics of plasmid-cured derivatives of *Rhizobium galegae*. Plant Science **125:**87–96.

86. **Bromfield ES, Lewis DM, Barran LR.** 1985. Cryptic plasmid and rifampin resistance in *Rhizobium meliloti* influencing nodulation competitiveness. J Bacteriol **164:**410–413.

87. **Roumiantseva ML, Andronov EE, Sagulenko VV, Onishuk OP, Provorov NA, Simarov BV.** 2004. [Instability of a cryptic plasmid in *Sinorhizobium meliloti* P108 during symbiosis with alfalfa (*Medicago sativa*)]. Genetika **40:**454–461.

88. **Barbour WM, Elkan GH.** 1990. Physiological characteristics and competitive ability of plasmid-cured derivatives of *Rhizobium fredii* USDA 206. Archives of Microbiology **154:**1–4.

89. **Barbour WM, Elkan GH.** 1989. Relationship of the presence and copy number of plasmids to exopolysaccharide production and symbiotic effectiveness in *Rhizobium fredii* USDA 206. Appl Environ Microbiol **55:**813–818.

90. **Rastogi VK, Bromfield ESP, Whitwill ST, Barran LR.** 1992. A cryptic plasmid of indigenous *rhizobium meliloti* possesses reiterated *nodC* and *nifE* genes and undergoes DNA rearrangement. Canadian Journal of Microbiology **38:**563–568.

91. **Novikova NI, Pavlova EA.** 1993. Enhanced competitiveness for nodulation of *Medicago sativa* by *Rhizobium meliloti* transconjugants harboring the root-inducing plasmids of *Agrobacterium rhizogenes* strain 15834. Fems Microbiology Ecology **12:**61–68.

92. **Stiens M, Schneiker S, Keller M, Kuhn S, Pühler A, Schlüter A.** 2006. Sequence analysis of the 144-kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. Appl Environ Microbiol **72:**3662–3672.

93. **Sanjuan J, Olivares J.** 1991. A NifA–NtrA regulatory system activates transcription of *nfe*, a gene locus involved in nodulation competitiveness of *Rhizobium meliloti*. Arch Microbiol **155:**543–548.

94. **Soto MJ, Zorzano A, Mercado-Blanco J, Lepek V, Olivares J, Toro N.** 1993. Nucleotide sequence and characterization of *Rhizobium meliloti* nodulation competitiveness genes *nfe*. J Mol Biol **229:**570–576.

95. **Soto MJ, Zorzano A, Garcia-Rodriguez FM, Mercado-Blanco J, Lopez-Lara IM, Olivares J, Toro N.** 1994. Identification of a novel *Rhizobium meliloti* nodulation efficiency *nfe* gene homolog of *Agrobacterium* ornithine cyclodeaminase. Mol Plant–Microbe Interact **7:**703–707.

96. **Villadas PJ, Velazquez E, Martinez-Molina E, Toro N.** 1995. Identification of nodule-dominant *Rhizobium meliloti* strains carrying pRmeGR4b-type plasmid within indigenous soil populations by PCR using primers derived from specific DNA sequences. Fems Microbiology Ecology **17:**161–168.

97. **Pankhurst CE, Macdonald PE, Reeves JM.** 1986. Enhanced nitrogen fixation and competitiveness for nodulation of *Lotus pedunculatus* by a plasmid-cured derivative of *Rhizobium loti*. J Gen Microbiol **132:**2321–2328.

98. **Selbitschka W, Lotz W.** 1991. Instability of cryptic plasmids affects the symbiotic effectivity of *Rhizobium leguminosarum* bv. *viceae* strains. Molecular Plant–Microbe Interactions **4:**608–618.

99. **Velázquez E, Mateos PF, Pedrero P, Dazzo FB, Martínez-Molina E.** 1995. Attenuation of symbiotic effectiveness by *Rhizobium meliloti* SAF22 related to the presence of a cryptic plasmid. Appl Environ Microbiol **61:**2033–2036.

100. **Stacey G, Paau AS, Noel KD, Maier RJ, Silver LE, Brill WJ.** 1982. Mutants of *Rhizobium japonicum* defective in nodulation. Archives of Microbiology **132:**219–224.

101. **Leigh JA, Signer ER, Walker GC.** 1985. Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. Proc Natl Acad Sci U S A **82:**6231–6235.

102. **Ugalde RA, Handelsman J, Brill WJ.** 1986. Role of galactosyltransferase activity in phage sensitivity and nodulation competitiveness of *Rhizobium meliloti*. J Bacteriol **166:**148–154.

103. **Defives C, Werquin M, Mary P, Hornez JP.** 1996. Roles of exopolysaccharides and lipopolysaccharides in the adsorption of the Siphovirus phage NM8 to *Rhizobium meliloti* M11S Cells. Curr Microbiol **33:**371–376.

104. **Campbell GR, Reuhs BL, Walker GC.** 1998. Different phenotypic classes of *Sinorhizobium meliloti* mutants defective in synthesis of K antigen. J Bacteriol **180:**5432–5436.

105. **Dylan T, Nagpal P, Helinski DR, Ditta GS.** 1990. Symbiotic pseudorevertants of *Rhizobium meliloti ndv* mutants. J Bacteriol **172:**1409–1417.

106. **Hashem FM, Angle JS.** 1990. Rhizobiophage effects on nodulation, nitrogen fixation, and yield of field-grown soybeans (*Glycine max* L. Merr.). Biology and Fertility of Soils **9:**330–334.

107. **Dowling DN, Broughton WJ.** 1986. Competition for nodulation of legumes. Annu Rev Microbiol **40:**131–157.

108. **Vlassak KM, Vanderleyden J, Graham PH.** 1997. Factors influencing nodule occupancy by inoculant rhizobia. Critical Reviews in Plant Sciences **16:**163–229.

109. **Putnoky P, Deák V, Békási K, Pálvölgyi A, Maász A, Palágyi Z, Hoffmann G, Kerepesi I.** 2004. H protein of bacteriophage 16-3 and RkpM protein of *Sinorhizobium meliloti* 41 are involved in phage adsorption. J Bacteriol **186:**1591–1597.

110. **Crook MB, Lindsay DP, Biggs MB, Bentley JS, Price JC, Clement SC, Clement MJ, Long SR, Griffitts JS.** 2012. Rhizobial plasmids that cause impaired symbiotic nitrogen fixation and enhanced host invasion. Mol Plant–Microbe Interact **25:**1026–1033.

111. **Gibson KE, Kobayashi H, Walker GC.** 2008. Molecular determinants of a symbiotic chronic infection. Annu Rev Genet **42:**413–441.

112. **Béna G, Lyet A, Huguet T, Olivieri I.** 2005. *Medicago–Sinorhizobium* symbiotic specificity evolution and the geographic expansion of *Medicago*. J Evol Biol **18:**1547–1558.

113. **Bladergroen MR, Badelt K, Spaink HP.** 2003. Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. Mol Plant–Microbe Interact **16:**53–64.

114. **Ruvkun GB, Sundaresan V, Ausubel FM.** 1982. Directed transposon Tn*5* mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. *Cell* **29:**551–559.

115. **Tirichine L, de Billy F, Huguet T.** 2000. *Mtsym6*, a gene conditioning *Sinorhizobium* strain-specific nitrogen fixation in *Medicago truncatula*. Plant Physiol **123:**845–851.

116. **Triplett EW, Sadowsky MJ.** 1992. Genetics of competition for nodulation of legumes. Annu Rev Microbiol **46:**399–428.

117. **Foster KR, Wenseleers T.** 2006. A general model for the evolution of mutualisms. J Evol Biol **19:**1283–1293.

118. **Jones EI, Ferrière R, Bronstein JL.** 2009. Eco-evolutionary dynamics of mutualists and exploiters. Am Nat **174:**780–794.

119. **Nowak MA.** 2006. Five rules for the evolution of cooperation. Science **314:**1560–1563.

120. **Oono R, Anderson CG, Denison RF.** 2011. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. Proc Biol Sci **278:**2698–2703.

121. **Sachs JL, Mueller UG, Wilcox TP, Bull JJ.** 2004. The evolution of cooperation. Q Rev Biol **79:**135–160.

122. **van Berkum P, Elia P, Eardly BD.** 2006. Multilocus sequence typing as an approach for population analysis of *Medicago*-nodulating rhizobia. J Bacteriol **188:**5570–5577.

123. **Griffitts JS, Carlyon RE, Erickson JH, Moulton JL, Barnett MJ, Toman CJ, Long SR.** 2008. A *Sinorhizobium meliloti* osmosensory two-component system required for cyclic glucan export and symbiosis. Mol Microbiol **69:**479–490.

124. **Arango Pinedo C, Gage DJ.** 2009. Plasmids that insert into the rhamnose utilization locus, *rha*: a versatile tool for genetic studies in *Sinorhizobium meliloti*. J Mol Microbiol Biotechnol **17:**201–210.

125. **Hynes MF, Simon R, Pühler A.** 1985. The development of plasmid-free strains of *Agrobacterium tumefaciens* by using incompatibility with a *Rhizobium meliloti* plasmid to eliminate pAtC58. Plasmid **13:**99–105.

126. **Ellermeier CD, Hobbs EC, González-Pastor JE, Losick R.** 2006. A three-protein signaling pathway governing immunity to a bacterial cannibalism toxin. Cell **124:**549–559.

127. **Clement NL, Snell Q, Clement MJ, Hollenhorst PC, Purwar J, Graves BJ, Cairns BR, Johnson WE.** 2010. The GNUMAP algorithm: unbiased probabilistic mapping of oligonucleotides from next-generation sequencing. Bioinformatics **26:**38–45.

128. **Hartmann A, Giraud JJ, Catroux G.** 1998. Genotypic diversity of *Sinorhizobium* (formerly *Rhizobium*) meliloti strains isolated directly from a soil and from nodules of alfalfa (*Medicago sativa*) grown in the same soil. Fems Microbiology Ecology **25:**107–116.

129. **Nonaka S, Sugawara M, Minamisawa K, Yuhashi K, Ezura H.** 2008. 1-aminocyclopropane-1-carboxylate deaminase enhances *Agrobacterium tumefaciens*-mediated gene transfer into plant cells. Appl Environ Microbiol **74:**2526–2528.

130. **Deakin WJ, Broughton WJ.** 2009. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. Nat Rev Microbiol **7:**312–320.

131. **Krishnan HB, Lorio J, Kim WS, Jiang G, Kim KY, DeBoer M, Pueppke SG.** 2003. Extracellular proteins involved in soybean cultivar-specific nodulation are associated with pilus-like surface appendages and exported by a type III protein secretion system in *Sinorhizobium fredii* USDA257. Mol Plant–Microbe Interact **16:**617–625.

132. **Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X.** 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol Microbiol **28:**1381–1389.

133. **Yang S, Tang F, Gao M, Krishnan HB, Zhu H.** 2010. R gene-controlled host specificity in the legume–rhizobia symbiosis. Proc Natl Acad Sci U S A **107:**18735–18740.

134. **Mazur A, Koper P.** 2012. Rhizobial plasmids—replication, structure and biological role. Central European Journal of Biology **7:**571–586.

135. **Pinto UM, Flores-Mireles AL, Costa ED, Winans SC.** 2011. RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *repC* and functions only in *cis*. Mol Microbiol **81:**1593–1606.

136. **Wang XD, de Boer PA, Rothfield LI.** 1991. A factor that positively regulates cell division by activating transcription of the major cluster of essential cell division genes of *Escherichia coli*. EMBO J **10:**3363–3372.

137. **Michael B, Smith JN, Swift S, Heffron F, Ahmer BM.** 2001. SdiA of *Salmonella enterica* is a LuxR homolog that detects mixed microbial communities. J Bacteriol **183:**5733–5742.

138. **Pappas KM, Winans SC.** 2003. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. Mol Microbiol **48:**1059–1073.

139. **Geddes BA, Oresnik IJ.** 2012. Inability to catabolize galactose leads to increased ability to compete for nodule occupancy in *Sinorhizobium meliloti*. J Bacteriol **194:**5044–5053.

140. **Gao M, Chen H, Eberhard A, Gronquist MR, Robinson JB, Rolfe BG, Bauer WD.** 2005. *sinI*- and *expR*-dependent quorum sensing in *Sinorhizobium meliloti*. J Bacteriol **187:**7931–7944.

141. **Harrison CL, Crook MB, Peco G, Long SR, Griffitts JS.** 2011. Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. *Appl Environ Microbiol* **77:**3916–3922.

142. **Charles TC, Finan TM.** 1991. Analysis of a 1600-kilobase *Rhizobium meliloti* megaplasmid using defined deletions generated *in vivo*. *Genetics* **127:**5–20.

143. **Milunovic B.** 2011. Deletion analysis of the *Sinorhizobium meliloti* genome. McMaster University.

144. **House BL, Mortimer MW, Kahn ML.** 2004. New recombination methods for *Sinorhizobium meliloti* genetics. Appl Environ Microbiol **70:**2806–2815.

145. **Pandey DP, Gerdes K.** 2005. Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* **33:**966–976.

146. **Makarova KS, Wolf YI, Koonin EV.** 2009. Comprehensive comparative-genomic analysis of Type 2 toxin–antitoxin systems and related mobile stress response systems in prokaryotes. *Biology Direct* **4**.

147. **Arcus VL, McKenzie JL, Robson J, Cook GM.** 2010. The PIN-domain ribonucleases and the prokaryotic VapBC toxin–antitoxin array. *Protein Engineering Design and Selection* **24:**33–40.

148. **Sevin EW, Barloy-Hubler F.** 2007. RASTA-Bacteria: a web-based tool for identifying toxin–antitoxin loci in prokaryotes. *Genome Biol* **8:**R155.

149. **Shao Y, Harrison EM, Bi D, Tai C, He X, Ou HY, Rajakumar K, Deng Z.** 2011. TADB: a web-based resource for Type 2 toxin–antitoxin loci in bacteria and archaea. *Nucleic Acids Res* **39:**D606–611.

150. **Mount DW.** 2007. Using the Basic Local Alignment Search Tool (BLAST). CSH Protoc **2007:**pdb top17.

151. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147.

152. **Cheng JJ, Poduska B, Morton RA, Finan TM.** 2011. An ABC-type cobalt transport system is essential for growth of *Sinorhizobium meliloti* at trace metal concentrations. *J Bacteriol* **193:**4405–4416.

153. **Finan TM, Weidner S, Wong K, Buhrmester J, Chain P, Vorhölter FJ, Hernández-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:**9889–9894.

154. **Olah B, Kiss E, Gyorgypal Z, Borzi J, Cinege G, Csanadi G, Batut J, Kondorosi A, Dusha I.** 2001. Mutation in the *ntrR* gene, a member of the *vap* gene family, increases the symbiotic efficiency of *Sinorhizobium meliloti*. Mol Plant–Microbe Interact **14:**887–894.

155. **Bodogai M, Ferenczi S, Bashtovyy D, Miclea P, Papp P, Dusha I.** 2006. The *ntrPR* operon of *Sinorhizobium meliloti* is organized and functions as a toxin–antitoxin module. Mol Plant–Microbe Interact **19:**811–822.

156. **Puskas LG, Nagy ZB, Kelemen JZ, Ruberg S, Bodogai M, Becker A, Dusha I.** 2004. Wide-range transcriptional modulating effect of *ntrR* under microaerobiosis in *Sinorhizobium meliloti*. *Mol Genet Genomics* **272:**275–289.

157. **Pueppke SG, Broughton WJ.** 1999. *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. Mol Plant–Microbe Interact **12:**293–318.

158. **Schuldes J, Rodriguez Orbegoso M, Schmeisser C, Krishnan HB, Daniel R, Streit WR.** 2012. Complete genome sequence of the broad-host-range strain *Sinorhizobium fredii* USDA257. J Bacteriol **194:**4483.

159. **Hynes MF, Quandt J, O'Connell MP, Pühler A.** 1989. Direct selection for curing and deletion of *Rhizobium* plasmids using transposons carrying the *Bacillus subtilis sacB* gene. *Gene* **78:**111–120.

160. **Charles TC, Newcomb W, Finan TM.** 1991. *ndvF*, a novel locus located on megaplasmid pRmeSU47b (pEXO) of *Rhizobium meliloti*, is required for normal nodule development. *J Bacteriol* **173:**3981–3992.

161. **Dicenzo G, Milunovic B, Cheng J, Finan TM.** 2012. $tRNA^{arg}$ 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.

162. **Wong K, Golding GB.** 2003. A phylogenetic analysis of the pSymB replicon from the *Sinorhizobium meliloti* genome reveals a complex evolutionary history. Can J Microbiol **49:**269–280.

163. **Wong K, Finan TM, Golding GB.** 2002. Dinucleotide compositional analysis of *Sinorhizobium meliloti* using the genome signature: distinguishing chromosomes and plasmids. Funct Integr Genomics **2:**274–281.

164. **Blanca-Ordóñez H, Oliva-García JJ, Pérez-Mendoza D, Soto MJ, Olivares J, Sanjuán J, Nogales J.** 2010. pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. J Bacteriol **192:**6309–6312.

165. **Harrison PW, Lower RP, Kim NK, Young JP.** 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. Trends Microbiol **18:**141–148.

166. **Oresnik IJ, Liu SL, Yost CK, Hynes MF.** 2000. Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *J Bacteriol* **182:**3582–3586.

167. **Broughton WJ, Wong CH, Lewin A, Samrey U, Myint H, Meyer H, Dowling DN, Simon R.** 1986. Identification of *Rhizobium* plasmid sequences involved in recognition of *Psophocarpus*, *Vigna*, and other legumes. J Cell Biol **102:**1173–1182.

168. **Bailly X, Giuntini E, Sexton MC, Lower RP, Harrison PW, Kumar N, Young JP.** 2011. Population genomics of *Sinorhizobium medicae* based on low-coverage sequencing of sympatric isolates. ISME J **5:**1722–1734.

169. **Bailly X, Olivieri I, De Mita S, Cleyet-Marel JC, Béna G.** 2006. Recombination and selection shape the molecular diversity pattern of nitrogen-fixing *Sinorhizobium* sp. associated to *Medicago*. Mol Ecol **15:**2719–2734.

170. **Bailly X, Olivieri I, Brunel B, Cleyet-Marel JC, Béna G.** 2007. Horizontal gene transfer and homologous recombination drive the evolution of the nitrogen-fixing symbionts of *Medicago* species. J Bacteriol **189:**5223–5236.

171. **Rangin C, Brunel B, Cleyet-Marel JC, Perrineau MM, Béna G.** 2008. Effects of *Medicago truncatula* genetic diversity, rhizobial competition, and strain effectiveness on the diversity of a natural sinorhizobium species community. Appl Environ Microbiol **74:**5653–5661.

172. **Silva C, Kan FL, Martínez-Romero E.** 2007. Population genetic structure of *Sinorhizobium meliloti* and *S. medicae* isolated from nodules of *Medicago* spp. in Mexico. FEMS Microbiol Ecol **60:**477–489.

173. **Carelli M, Gnocchi S, Fancelli S, Mengoni A, Paffetti D, Scotti C, Bazzicalupo M.** 2000. Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa cultivars in Italian soils. Appl Environ Microbiol **66:**4785–4789.

174. **Elboutahiri N, Thami-Alami I, Udupa SM.** 2010. Phenotypic and genetic diversity in *Sinorhizobium meliloti* and *S. medicae* from drought and salt affected regions of Morocco. BMC Microbiol **10:**15.

175. **Talebi MB, Bahar M, Saeidi G, Mengoni A, Bazzicalupo M.** 2008. Diversity of *Sinorhizobium* strains nodulating *Medicago sativa* from different Iranian regions. FEMS Microbiol Lett **288:**40–46.

176. **Roumiantseva ML, Onishchuk OP, Belova VS, Kurchak ON, Simarov BV.** 2011. [Polymorphism of *Sinorhizobium meliloti* strains isolated from diversity centers of alfalfa in various soil and climatic conditions.] Russian Journal of Genetics: Applied Research **1:**97–102.

177. **Sun S, Guo H, Xu J.** 2006. Multiple gene genealogical analyses reveal both common and distinct population genetic patterns among replicons in the nitrogen-fixing bacterium *Sinorhizobium meliloti*. Microbiology **152:**3245–3259.

178. **Jebara M, Mhamdi R, Aouani ME, Ghrir R, Mars M.** 2001. Genetic diversity of *Sinorhizobium* populations recovered from different *Medicago* varieties cultivated in Tunisian soils. Can J Microbiol **47:**139–147.

179. **Paffetti D, Scotti C, Gnocchi S, Fancelli S, Bazzicalupo M.** 1996. Genetic diversity of an Italian *Rhizobium meliloti* population from different *Medicago sativa* varieties. Appl Environ Microbiol **62:**2279–2285.

180. **Roumiantseva ML, Andronov EE, Sharypova LA, Dammann-Kalinowski T, Keller M, Young JP, Simarov BV.** 2002. Diversity of *Sinorhizobium meliloti* from the Central Asian Alfalfa Gene Center. Appl Environ Microbiol **68:**4694–4697.

181. **Bromfield ESP, Behara AMP, Singh RS, Barran LR.** 1998. Genetic variation in local populations of *Sinorhizobium meliloti*. Soil Biology & Biochemistry **30:**1707–1716.

182. **Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM.** 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res **37:**D141–145.

183. **Euzéby JP.** 1997. List of bacterial names with standing in nomenclature: a folder available on the internet. Int J Syst Bacteriol **47:**590–592.

184. **Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH.** 2005. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci U S A **102:**8369–8374.

185. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797.

186. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5:**113.

187. **Maddison WP, Maddison DR.** 1989. Interactive analysis of phylogeny and character evolution using the computer program MacClade. Folia Primatol (Basel) **53:**190–202.

188. **Swofford D.** 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

189. **Plazanet C, Réfrégier G, Demont N, Truchet G, Rosenberg C.** 1995. The *Rhizobium meliloti* region located downstream of the *nod* box *n6* is involved in the specific nodulation of *Medicago lupulina*. FEMS Microbiol Lett **133:**285–291.

190. **Turner SL, Knight KA, Young JP.** 2002. Identification and analysis of rhizobial plasmid origins of transfer. FEMS Microbiol Ecol **42:**227–234.

191. **Tittabutr P, Awaya JD, Li QX, Borthakur D.** 2008. The cloned 1-aminocyclopropane-1-carboxylate (ACC) deaminase gene from *Sinorhizobium* sp. strain BL3 in Rhizobium sp. strain TAL1145 promotes nodulation and growth of *Leucaena leucocephala*. Syst Appl Microbiol **31:**141–150.

192. **Giuntini E, Mengoni A, De Filippo C, Cavalieri D, Aubin-Horth N, Landry CR, Becker A, Bazzicalupo M.** 2005. Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of *Sinorhizobium meliloti* natural strains. *BMC Genomics* **6:**158.

193. **Allen ON, Allen EK.** 1981. The Leguminosae: A source book of characteristics, uses, and nodulation. The University of Wisconsin Press, Madison, Wisconsin.

194. **Poinsot V, Bélanger E, Laberge S, Yang GP, Antoun H, Cloutier J, Treilhou M, Dénarié J, Promé JC, Debellé F.** 2001. Unusual methyl-branched α,β-unsaturated acyl chain substitutions in the Nod factors of an arctic rhizobium, *Mesorhizobium* sp. strain N33 (*Oxytropis arctobia*). J Bacteriol **183:**3721–3728.

195. **Tan ZY, Wang ET, Peng GX, Zhu ME, Martínez-Romero E, Chen WX.** 1999. Characterization of bacteria isolated from wild legumes in the north-western regions of China. Int J Syst Bacteriol **49 Pt 4:**1457–1469.

196. **Yang W, Kong Z, Chen W, Wei G.** 2012. Genetic diversity and symbiotic evolution of rhizobia from root nodules of *Coronilla varia*. Syst Appl Microbiol.

197. **Sotelo M, Irisarri P, Lorite MJ, Casaretto E, Rebuffo M, Sanjuán J, Monza J.** 2011. Diversity of rhizobia nodulating *Lotus corniculatus* grown in northern and southern regions of Uruguay. Applied Soil Ecology **49:**197–207.

198. **Estrella MJ, Muñoz S, Soto MJ, Ruiz O, Sanjuán J.** 2009. Genetic diversity and host range of rhizobia nodulating *Lotus tenuis* in typical soils of the Salado River Basin (Argentina). Appl Environ Microbiol **75:**1088–1098.

199. **Vinuesa P, Silva C, Lorite MJ, Izaguirre-Mayoral ML, Bedmar EJ, Martínez-Romero E.** 2005. Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. Syst Appl Microbiol **28:**702–716.

200. **van Berkum P, Terefework Z, Paulin L, Suomalainen S, Lindstrom K, Eardly BD.** 2003. Discordant phylogenies within the *rrn* loci of rhizobia. *J Bacteriol* **185:**2988–2998.

201. **Gaunt MW, Turner SL, Rigottier-Gois L, Lloyd-Macgilp SA, Young JP.** 2001. Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int J Syst Evol Microbiol* **51:**2037–2048.

202. **Eardly BD, Wang FS, van Berkum P.** 1996. Corresponding 16S rRNA gene segments in Rhizobiaceae and *Aeromonas* yield discordant phylogenies. *Plant and Soil* **186:**69–74.

203. **Young JPW, Haukka KE.** 1996. Diversity and phylogeny of rhizobia. *New Phytologist* **133:**87–94.

204. **Eardly BD, Nour SM, van Berkum P, Selander RK.** 2005. Rhizobial 16S rRNA and *dnaK* genes: mosaicism and the uncertain phylogenetic placement of *Rhizobium galegae*. *Appl Environ Microbiol* **71:**1328–1335.

205. **Sullivan JT, Eardly BD, van Berkum P, Ronson CW.** 1996. Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus*. *Appl Environ Microbiol* **62:**2818–2825.

206. **Eardly BD, Wang FS, Whittam TS, Selander RK.** 1995. Species limits in *Rhizobium* populations that nodulate the common bean (*Phaseolus vulgaris*). Appl Environ Microbiol **61:**507–512.

207. **Stackebrandt E, Frederiksen W, Garrity GM, Grimont PA, Kämpfer P, Maiden MC, Nesme X, Rosselló-Mora R, Swings J, Trüper HG, Vauterin L, Ward AC, Whitman WB.** 2002. Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol **52:**1043–1047.

208. **Forde A, Fitzgerald GF.** 1999. Bacteriophage defence systems in lactic acid bacteria. Antonie Van Leeuwenhoek **76:**89–113.

209. **Klaenhammer TR.** 1984. Interactions of bacteriophages with lactic streptococci, p. 1–29. *In* Allen IL (ed.), Advances in Applied Microbiology, vol. Volume 30. Academic Press.

210. **Labrie SJ, Samson JE, Moineau S.** 2010. Bacteriophage resistance mechanisms. Nat Rev Microbiol **8:**317–327.

211. **Rakhuba DV, Kolomiets EI, Dey ES, Novik GI.** 2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Pol J Microbiol **59:**145–155.

212. **Letellier L, Boulanger P, Plançon L, Jacquot P, Santamaria M.** 2004. Main features on tailed phage, host recognition and DNA uptake. Front Biosci **9:**1228–1339.

213. **Lindberg AA.** 1973. Bacteriophage receptors. Annu Rev Microbiol **27:**205–241.

214. **Hyman P, Abedon ST.** 2010. Bacteriophage host range and bacterial resistance. Adv Appl Microbiol **70:**217–248.

215. **Finan TM, Hartweig E, LeMieux K, Bergman K, Walker GC, Signer ER.** 1984. General transduction in *Rhizobium meliloti*. J Bacteriol **159:**120–124.

216. **Martin MO, Long SR.** 1984. Generalized transduction in *Rhizobium meliloti*. J Bacteriol **159:**125–129.

217. **Tamm LK, Hong H, Liang B.** 2004. Folding and assembly of β-barrel membrane proteins. Biochim Biophys Acta **1666:**250–263.

218. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785–786.

219. **Izard JW, Kendall DA.** 1994. Signal peptides: exquisitely designed transport promoters. Mol Microbiol **13:**765–773.

220. **Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ.** 2004. PRED-TMBB: a web server for predicting the topology of β-barrel outer membrane proteins. Nucleic Acids Res **32:**W400–404.

221. **Djordjevic MA, Chen HC, Natera S, Van Noorden G, Menzel C, Taylor S, Renard C, Geiger O, Weiller GF.** 2003. A global analysis of protein expression profiles in

Sinorhizobium meliloti: discovery of new genes for nodule occupancy and stress adaptation. Mol Plant–Microbe Interact **16:**508–524.

222. **Gehring KB, Nikaido H.** 1989. Existence and purification of porin heterotrimers of *Escherichia coli* K12 OmpC, OmpF, and PhoE proteins. J Biol Chem **264:**2810–2815.
223. **de Maagd RA, Mulders IH, Canter Cremers HC, Lugtenberg BJ.** 1992. Cloning, nucleotide sequencing, and expression in *Escherichia coli* of a *Rhizobium leguminosarum* gene encoding a symbiotically repressed outer membrane protein. J Bacteriol **174:**214–221.
224. **Jones JD, 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:**299–306.
225. **Kanipes MI, Kalb SR, Cotter RJ, Hozbor DF, Lagares A, Raetz CR.** 2003. Relaxed sugar donor selectivity of a *Sinorhizobium meliloti* ortholog of the *Rhizobium leguminosarum* mannosyl transferase LpcC. Role of the lipopolysaccharide core in symbiosis of Rhizobiaceae with plants. J Biol Chem **278:**16365–16371.
226. **Lagares A, Hozbor DF, Niehaus K, Otero AJ, Lorenzen J, Arnold W, Pühler A.** 2001. Genetic characterization of a *Sinorhizobium meliloti* chromosomal region in lipopolysaccharide biosynthesis. J Bacteriol **183:**1248–1258.
227. **Laloux G, Deghelt M, de Barsy M, Letesson JJ, De Bolle X.** 2010. Identification of the essential *Brucella melitensis* porin Omp2b as a suppressor of Bax-induced cell death in yeast in a genome-wide screening. PLoS One **5:**e13274.
228. **Quandt J, Hynes MF.** 1993. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. Gene **127:**15–21.

229. **Ficht TA, Bearden SW, Sowa BA, Adams LG.** 1989. DNA sequence and expression of the 36-kilodalton outer membrane protein gene of *Brucella abortus*. Infect Immun **57:**3281–3291.

230. **Roest HP, Bloemendaal CJ, Wijffelman CA, Lugtenberg BJ.** 1995. Isolation and characterization of *ropA* homologous genes from *Rhizobium leguminosarum* biovars *viciae* and *trifolii*. J Bacteriol **177:**4985–4991.

231. **Ampe F, Kiss E, Sabourdy F, Batut J.** 2003. Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. Genome Biology **4**.

232. **de Maagd RA, Wientjes FB, Lugtenberg BJ.** 1989. Evidence for divalent cation (Ca^{2+})-stabilized oligomeric proteins and covalently bound protein-peptidoglycan complexes in the outer membrane of *Rhizobium leguminosarum*. J Bacteriol **171:**3989–3995.

233. **Fairman JW, Noinaj N, Buchanan SK.** 2011. The structural biology of β-barrel membrane proteins: a summary of recent reports. Curr Opin Struct Biol **21:**523–531.

234. **Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J.** 2003. Role of a highly conserved bacterial protein in outer membrane protein assembly. Science **299:**262–265.

235. **Genevrois S, Steeghs L, Roholl P, Letesson JJ, van der Ley P.** 2003. The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. EMBO J **22:**1780–1789.

236. **Gentle I, Gabriel K, Beech P, Waller R, Lithgow T.** 2004. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. J Cell Biol **164:**19–24.

237. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H.** 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol **2:**2006.0008.

238. **Yamamoto N, Nakahigashi K, Nakamichi T, Yoshino M, Takai Y, Touda Y, Furubayashi A, Kinjyo S, Dose H, Hasegawa M, Datsenko KA, Nakayashiki T, Tomita M, Wanner BL, Mori H.** 2009. Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. Mol Syst Biol **5:**335.

239. **Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson MV, Manoil C.** 2003. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **100:**14339–14344.

240. **Akerley BJ, Rubin EJ, Novick VL, Amaya K, Judson N, Mekalanos JJ.** 2002. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. Proc Natl Acad Sci U S A **99:**966–971.

241. **Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK.** 2009. Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. Genome Res **19:**2308–2316.

242. **Bauer FJ, Rudel T, Stein M, Meyer TF.** 1999. Mutagenesis of the *Neisseria gonorrhoeae* porin reduces invasion in epithelial cells and enhances phagocyte responsiveness. Mol Microbiol **31:**903–913.

243. **Fudyk TC, Maclean IW, Simonsen JN, Njagi EN, Kimani J, Brunham RC, Plummer FA.** 1999. Genetic diversity and mosaicism at the por locus of *Neisseria gonorrhoeae*. J Bacteriol **181:**5591–5599.

244. **Tommassen J, Vermeij P, Struyve M, Benz R, Poolman JT.** 1990. Isolation of *Neisseria meningitidis* mutants deficient in class 1 (*porA*) and class 3 (*porB*) outer membrane proteins. Infect Immun **58:**1355–1359.

245. **Bucarey SA, Villagra NA, Martinic MP, Trombert AN, Santiviago CA, Maulen NP, Youderian P, Mora GC.** 2005. The *Salmonella enterica* serovar Typhi *tsx* gene, encoding a nucleoside-specific porin, is essential for prototrophic growth in the absence of nucleosides. Infect Immun **73:**6210–6219.

246. **Nieweg A, Bremer E.** 1997. The nucleoside-specific Tsx channel from the outer membrane of *Salmonella typhimurium*, *Klebsiella pneumoniae* and *Enterobacter aerogenes*: functional characterization and DNA sequence analysis of the *tsx* genes. Microbiology **143 (Pt 2):**603–615.

247. **Vayssier-Taussat M, Le Rhun D, Deng HK, Biville F, Cescau S, Danchin A, Marignac G, Lenaour E, Boulouis HJ, Mavris M, Arnaud L, Yang H, Wang J, Quebatte M, Engel P, Saenz H, Dehio C.** 2010. The Trw type IV secretion system of *Bartonella* mediates host-specific adhesion to erythrocytes. PLoS Pathog **6:**e1000946.

248. **Barnett MJ, Toman CJ, Fisher RF, Long SR.** 2004. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote–host interaction. Proc Natl Acad Sci U S A **101:**16636–16641.

249. **Roest HP, Goosenderoo L, Wijffelman CA, Demaagd RA, Lugtenberg BJJ.** 1995. Outer membrane protein changes during bacteroid development are independent of

nitrogen fixation and differ between indeterminate and determinate nodulating host plants of *Rhizobium leguminosarum*. Molecular Plant–Microbe Interactions **8:**14–22.

250. **Wibberg D, -**

ter A.

2011. Complete genome sequencing of *Agrobacterium* sp. H13-3, the former *Rhizobium lupini* H13-3, reveals a tripartite genome consisting of a circular and a linear chromosome and an accessory plasmid but lacking a tumor-inducing Ti-plasmid. J Biotechnol **155:**50–62.

251. **Lehnherr H, Maguin E, Jafri S, Yarmolinsky MB.** 1993. Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when prophage is retained. J Mol Biol **233:**414–428.

252. **Engelberg-Kulka H, Hazan R, Amitai S.** 2005. *mazEF*: a chromosomal toxin–antitoxin module that triggers programmed cell death in bacteria. J Cell Sci **118:**4327–4332.

253. **Hazan R, Engelberg-Kulka H.** 2004. *Escherichia coli mazEF*-mediated cell death as a defense mechanism that inhibits the spread of phage P1. Mol Genet Genomics **272:**227–234.

254. **Lewis K.** 2000. Programmed death in bacteria. Microbiology and Molecular Biology Reviews **64:**503–514.

255. **Ceyssens P-J.** 2009. Isolation and characterization of lytic bacteriophages infecting *Pseudomonas aeruginosa*. Katholieke Universiteit Leuven, Flanders, Belgium.

256. **de Mello Varani A, Souza RC, Nakaya HI, de Lima WC, Paula de Almeida LG, Kitajima EW, Chen J, Civerolo E, Vasconcelos AT, Van Sluys MA.** 2008. Origins of

the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PLoS One **3:**e4059.

257. **Karaolis DK, Somara S, Maneval DR, Jr., Johnson JA, Kaper JB.** 1999. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature **399:**375–379.

258. **Epstein B, Branca A, Mudge J, Bharti AK, Briskine R, Farmer AD, Sugawara M, Young ND, Sadowsky MJ, Tiffin P.** 2012. Population genomics of the facultatively mutualistic bacteria *Sinorhizobium meliloti* and *S. medicae*. PLoS Genet **8:**e1002868.

259. **Reiter WD, Palm P, Yeats S.** 1989. Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res **17:**1907–1914.

260. **Paquet JY, Diaz MA, Genevrois S, Grayon M, Verger JM, de Bolle X, Lakey JH, Letesson JJ, Cloeckaert A.** 2001. Molecular, antigenic, and functional analyses of Omp2b porin size variants of *Brucella* spp. J Bacteriol **183:**4839–4847.

261. **Grissa I, Vergnaud G, Pourcel C.** 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics **8:**172.

262. **Edwards RA, Helm RA, Maloy SR.** 1999. Increasing DNA transfer efficiency by temporary inactivation of host restriction. Biotechniques **26:**892–894.

263. **Bonhivers M, Ghazi A, Boulanger P, Letellier L.** 1996. FhuA, a transporter of the *Escherichia coli* outer membrane, is converted into a channel upon binding of bacteriophage T5. EMBO J **15:**1850–1856.

264. **Choi Y.** 2012. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein, p. 414–417, Proceedings of the ACM

Conference on Bioinformatics, Computational Biology and Biomedicine. ACM, Orlando, Florida.

265. **Choi Y, Sims GE, Murphy S, Miller JR, Chan AP.** 2012. Predicting the functional effect of amino acid substitutions and indels. PLoS One **7:**e46688.

266. **Terpolilli JJ, O'Hara GW, Tiwari RP, Dilworth MJ, Howieson JG.** 2008. The model legume *Medicago truncatula* A17 is poorly matched for N_2 fixation with the sequenced microsymbiont *Sinorhizobium meliloti* 1021. New Phytol **179:**62–66.

267. **Lamontagne J, Forest A, Marazzo E, Denis F, Butler H, Michaud JF, Boucher L, Pedro I, Villeneuve A, Sitnikov D, Trudel K, Nassif N, Boudjelti D, Tomaki F, Chaves-Olarte E, Guzman-Verri C, Brunet S, Cote-Martin A, Hunter J, Moreno E, Paramithiotis E.** 2009. Intracellular adaptation of *Brucella abortus*. J Proteome Res **8:**1594–1609.

268. **Cha SB, Rayamajhi N, Lee WJ, Shin MK, Jung MH, Shin SW, Kim JW, Yoo HS.** 2012. Generation and envelope protein analysis of internalization-defective *Brucella abortus* mutants in professional phagocytes, RAW 264.7. FEMS Immunol Med Microbiol **64:**244–254.

269. **Marchesini MI, Connolly J, Delpino MV, Baldi PC, Mujer CV, DelVecchio VG, Comerci DJ.** 2011. *Brucella abortus* choloylglycine hydrolase affects cell envelope composition and host cell internalization. PLoS One **6:**e28480.

270. **Lim CW, Park JY, Lee SH, Hwang CH.** 2010. Comparative proteomic analysis of soybean nodulation using a supernodulation mutant, SS2-2. Biosci Biotechnol Biochem **74:**2396–2404.

271. **de María N, Guevara Á, Serra MT, García-Luque I, González-Sama A, García de Lacoba M, de Felipe MR, Fernández-Pascual M.** 2007. Putative porin of *Bradyrhizobium* sp. (*Lupinus*) bacteroids induced by glyphosate. Appl Environ Microbiol **73:**5075–5082.

272. **Burgess AW, Anderson BE.** 1998. Outer membrane proteins of *Bartonella henselae* and their interaction with human endothelial cells. Microb Pathog **25:**157–164.

273. **Burgess AW, Paquet JY, Letesson JJ, Anderson BE.** 2000. Isolation, sequencing and expression of *Bartonella henselae omp43* and predicted membrane topology of the deduced protein. Microb Pathog **29:**73–80.

274. **Dabo SM, Confer AW, Saliki JT, Anderson BE.** 2006. Binding of *Bartonella henselae* to extracellular molecules: identification of potential adhesins. Microb Pathog **41:**10–20.

275. **Campbell GA, Adams LG, Sowa BA.** 1994. Mechanisms of binding of *Brucella abortus* to mononuclear phagocytes from cows naturally resistant or susceptible to brucellosis. Vet Immunol Immunopathol **41:**295–306.

276. **He Y, Xiang Z.** 2010. Bioinformatics analysis of *Brucella* vaccines and vaccine targets using VIOLIN. Immunome Res **6 Suppl 1:**S5.

277. **Ruoslahti E.** 1996. RGD and other recognition sequences for integrins. Annu Rev Cell Dev Biol **12:**697–715.

278. **Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot RM, Fleetwood DJ, McCallum NG, Rossbach U, Stuart GS, Weaver JE, Webby RJ, De Bruijn FJ, Ronson CW.** 2002. Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti strain* R7A. J Bacteriol **184:**3086–3095.

279. **Barrios H, Valderrama B, Morett E.** 1999. Compilation and analysis of σ^{54}-dependent promoter sequences. Nucleic Acids Res **27:**4305–4313.

280. **Gussin GN, Ronson CW, Ausubel FM.** 1986. Regulation of nitrogen fixation genes. Annual Review of Genetics **20:**567–591.

281. **Douglas SE.** 1995. DNA Strider—An inexpensive sequence-analysis package for the Macintosh. Molecular Biotechnology **3:**37–45.

282. **González-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fernández-Riverola F, Posada D.** 2010. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res **38:**W14–18.

283. **Mao C, Qiu J, Wang C, Charles TC, Sobral BW.** 2005. NodMutDB: a database for genes and mutants involved in symbiosis. Bioinformatics **21:**2927–2929.

284. **Kibbe WA.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res **35:**W43–46.

285. **Eardly BD, Materon LA, Smith NH, Johnson DA, Rumbaugh MD, Selander RK.** 1990. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. Appl Environ Microbiol **56:**187–194.

286. **Griffitts JS, Long SR.** 2008. A symbiotic mutant of *Sinorhizobium meliloti* reveals a novel genetic pathway involving succinoglycan biosynthetic functions. Mol Microbiol **67:**1292–1306.

287. **Hamilton RH, Fall MZ.** 1971. The loss of tumor-initiating ability in *Agrobacterium tumefaciens* by incubation at high temperature. Experientia **27:**229–230.

288. **Meade HM, Long SR, Ruvkun GB, Brown SE, Ausubel FM.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. J Bacteriol **149:**114–122.

289. **Olsen P, Wright S, Collins M, Rice W.** 1994. Patterns of reactivity between a panel of monoclonal antibodies and forage *Rhizobium* strains. Appl Environ Microbiol **60:**654–661.

290. **Finan TM, Kunkel B, De Vos GF, Signer ER.** 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J Bacteriol **167:**66–72.

291. **Penmetsa RV, Cook DR.** 2000. Production and characterization of diverse developmental mutants of *Medicago truncatula*. Plant Physiol **123:**1387–1398.

292. **Trinick MJ.** 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania grandiflora* and their affinities with other rhizobial groups. Journal of Applied Bacteriology **49:**39–53.

293. **Chen H, Batley M, Redmond J, Rolfe BG.** 1985. Alteration of the effective nodulation properties of a fastgrowing broad host range *Rhizobium* due to changes in exopolysaccharide synthesis. Journal of Plant Physiology **120:**331–349.

294. **Keyser HH, Bohlool BB, Hu TS, Weber DF.** 1982. Fast-growing rhizobia isolated from root nodules of soybean. Science **215:**1631–1632.

295. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640–6645.

296. **Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, Garrity GM, Tiedje JM.** 2005. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res **33:**D294–296.

297. **Grant SG, Jessee J, Bloom FR, Hanahan D.** 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc Natl Acad Sci U S A **87:**4645–4649.

298. **Wells DH, Long SR.** 2002. The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. Mol Microbiol **43:**1115–1127.

299. **Kowalski M.** 1970. Transducing phages of *Rhizobium meliloti*. Acta Microbiol Pol A **2:**109–113.

300. **Małek W.** 1990. Properties of the transducing phage Ml of *Rhizobium meliloti*. Journal of Basic Microbiology **30:**43–50.

301. **Casadesús J, Olivares J.** 1979. General transduction in *Rhizobium meliloti* by a thermosensitive mutant of bacteriophage DF2. J Bacteriol **139:**316–317.

302. **Sik T, Horváth J, Chatterjee S.** 1980. Generalized transduction in *Rhizobium meliloti*. Mol Gen Genet **178:**511–516.

303. **Eardly BD, Materon LA, Smith NH, Johnson DA, Rumbaugh MD, Selander RK.** 1990. Genetic structure of natural populations of the nitrogen-fixing bacterium *Rhizobium meliloti*. Appl Environ Microbiol. **56:**187–94.

4.2 Computational Resources

4.2.1 Software

1. A Plasmid Editor (APE) v.2.0.37 2003–2009, Wayne Davis, Department of Biology, University of Utah. http://biologylabs.utah.edu/jorgensen/wayned/ape/
2. FigTree Tree Figure Drawing Tool v.1.2.3 2006–2009, Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh. http://tree.bio.ed.ac.uk/
3. MacClade v.4.08 2005, David R. Maddison and Wayne P. Maddison, Sinauer Associates, Inc. http://macclade.org/macclade.html
4. PAUP* v.3.1 1993, David L. Swofford, Department of Scientific Computing, Florida State University. http://paup.csit.fsu.edu/.
5. DNA Strider, v.1.4f6 (281)
6. 4Peaks v.1.7.2 2005–2006, Alexander Griekspoor, Tom Groothuis, and John Timmer, Neefix Laboratories. http://nucleobytes.com/index.php/4peaks

4.2.2 Web Resources

1. ALTER http://sing.ei.uvigo.es/ALTER/ (282)
2. Biota of North America Program (BONAP) http://www.bonap.org/
3. Google Maps https://maps.google.com/
4. List of Prokaryotic names with Standing in Nomenclature (LPSN) http://www.bacterio.cict.fr/ (183)
5. MUSCLE, v.3.8 http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=muscle (185, 186)
6. NodMutDB http://nodmutdb.vbi.vt.edu/ (283)
7. Nucleic Acid Sequence Massager http://www.attotron.com/cybertory/analysis/seqMassager.htm
8. Oligo Calc http://www.basic.northwestern.edu/biotools/oligocalc.html (284)
9. PROVEAN http://provean.jcvi.org/index.php (264, 265)
10. Ribosomal Database Project (RDP) http://rdp.cme.msu.edu/index.jsp (182)

4.3 Image Attributions

The map of the legume biomes is from Wikimedia Commons:

http://commons.wikimedia.org/wiki/File:Legume_Biogeography.svg.

The diagrams of different betaine and flavonoid structures are from Wikimedia Commons:

http://commons.wikimedia.org/wiki/File:Chrysoeriol.svg,

http://commons.wikimedia.org/wiki/File:Luteolin.svg,

http://commons.wikimedia.org/wiki/File:Stachydrine.svg, and

http://commons.wikimedia.org/wiki/File:Trigonelline.svg.

The drawings of Nod factor structures are from Wikimedia Commons:

http://commons.wikimedia.org/wiki/File:NodSm-IV_(Ac,C16-2,S).svg and

http://commons.wikimedia.org/wiki/File:Generic_Nod_Factor structure.svg.

The diagrams of different extracellular polysaccharides are from Wikimedia Commons:

http://commons.wikimedia.org/wiki/File:Sinorhizobium_meliloti_strain_Rm1021_capsular_polysaccharide.svg,

http://commons.wikimedia.org/wiki/File:Cyclosophorotetracosaose.svg,

http://commons.wikimedia.org/wiki/File:Sinorhizobium_meliloti_galactoglucan_(EPS_II).svg,

http://commons.wikimedia.org/wiki/File:LPS_of_Sinorhizobium_species.svg, and

http://commons.wikimedia.org/wiki/File:Sinorhizobium_meliloti_monosuccinylated_succinoglycan_(EPS_I).svg.

The diagrams of different nodule structures are from Wikimedia Commons:

http://commons.wikimedia.org/wiki/File:Determinate_Nodule_Zones_Diagram.svg, and

http://commons.wikimedia.org/wiki/File:Indeterminate_Nodule_Zones_Diagram.svg.

5 APPENDICES

5.1 Supplemental Figures

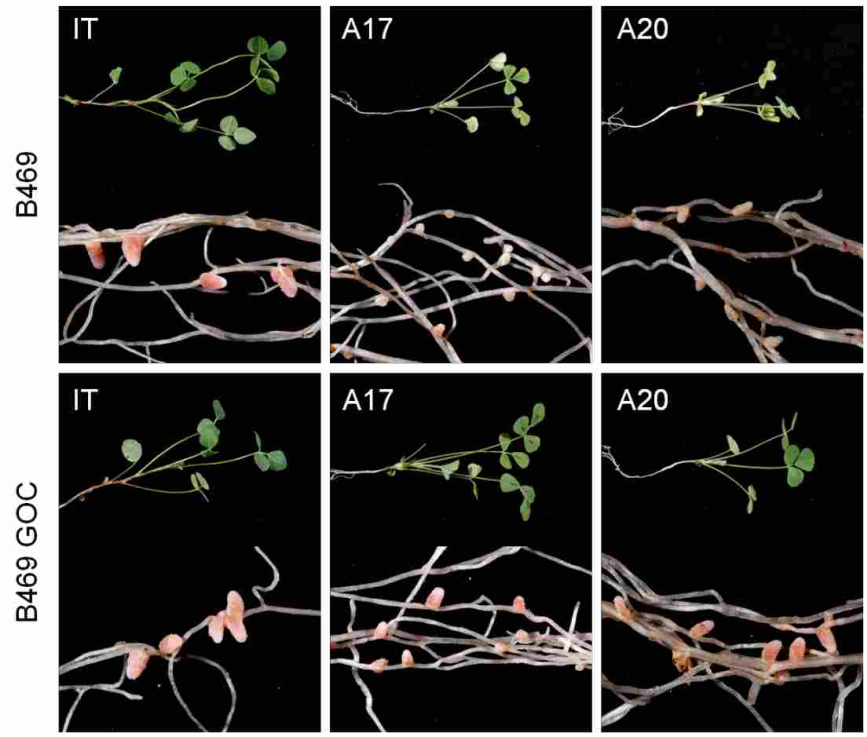

Figure S5.1-1. Rhizobia extracted from GOC nodules exhibit a stable Fix$^+$ phenotype. Representative root and shoot phenotypes from *M. italica* (IT), *M. truncatula* cv. A17, and *M. truncatula* cv. A20 inoculated with strains B469 and B469-GOC, 30 dpi.

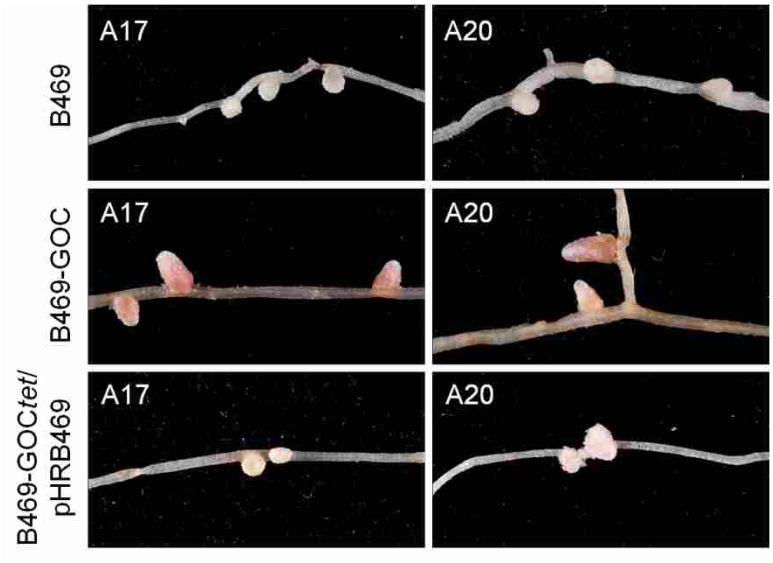

Figure S5.1-2. Re-introduction of an HR plasmid restores wild-type incompatibility. Plasmid pHRB469 (modified with the *oriT*/*neo* cassette to permit transfer and selection) was returned to B469-GOC*tet*, resulting in strain B469-GOC*tet*/pHRB469. Rhizobia were inoculated onto *M. truncatula* plants and nodules were imaged 30 dpi.

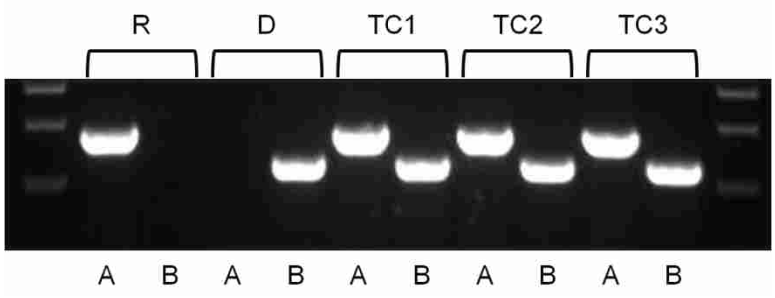

Figure S5.1-3. pHRB469 is self-transmissible. PCR confirmation of conjugation of pHRB469 into the B464 background. Donor (D; B469*oriT*/*neo*), recipient (R; B464*tet*), and transconjugant strains (TC1–3) were checked by PCR for the presence of the *tet* (A) and *neo* (B) cassettes.

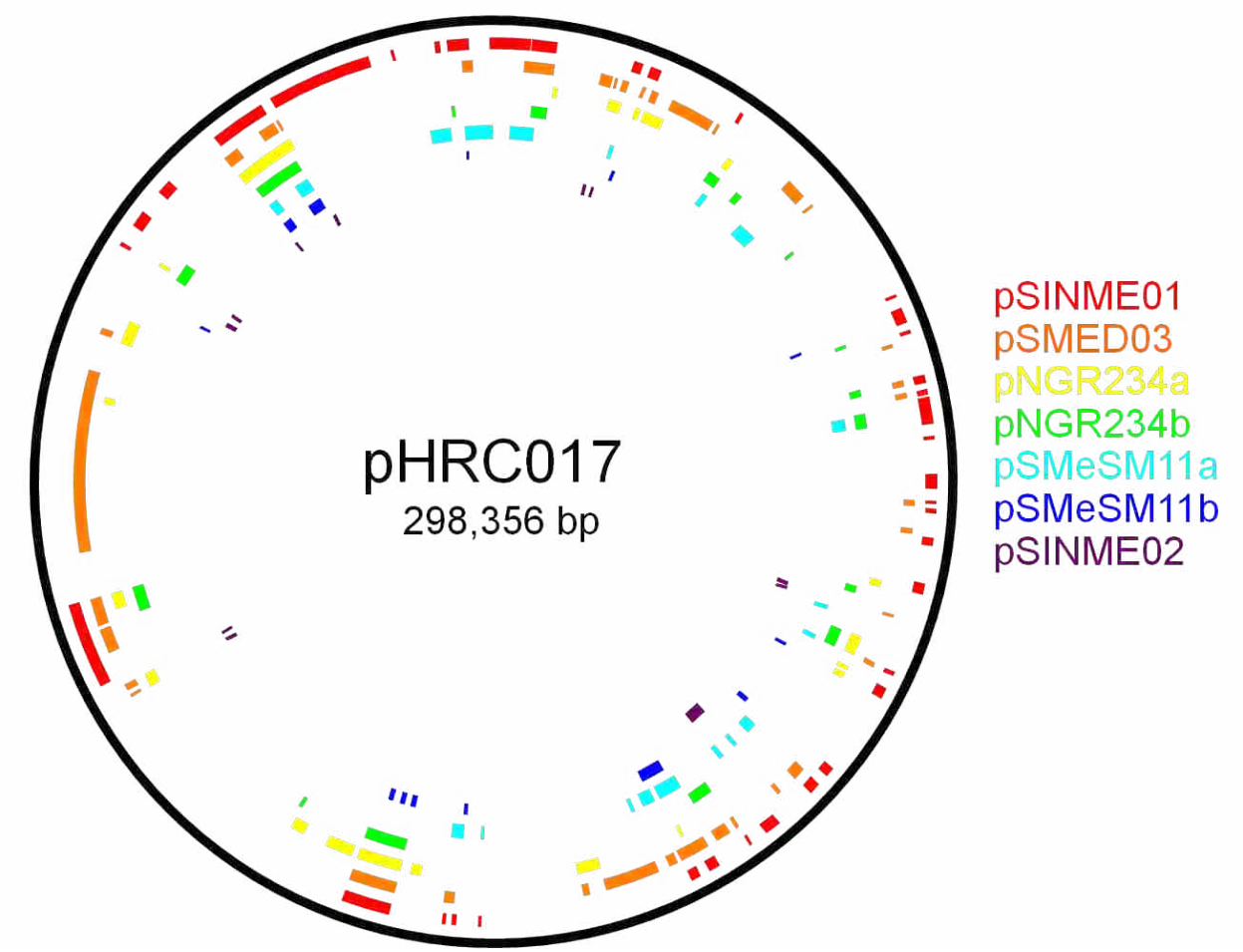

Figure S5.1-4. pHRC017 has limited similarity with other sequenced *Sinorhizobium* plasmids. Bars indicate BLASTn similarity with pSINME01 (red; GenBank Accession No. CP002784.1), pSMED03 (orange; GenBank Accession No. CP000741.1), pNGR234a (yellow; GenBank Accession No. NC000914.2), pNGR234b (green; GenBank Accession No. NC012586.1), pSMeSM11a (cyan; GenBank Accession No. DQ145546.1), pSMeSM11b (blue; GenBank Accession No. EF066650.1), pSINME02 (purple; GenBank Accession No. CP002785.1).

Figure S5.1-5. A genomic alignment of pSymB from several *S. meliloti* genomes and one *S. medicae* genome. A hyper-variable region which is located in the vicinity of a tRNA-arginine is boxed with a dotted line. A hyper-variable in the vicinity of *kps* and *rkp* genes is boxed with a solid line. A ~50 kB inversion in *S. medicae* is indicated with a dashed line. From top to bottom the strains are as follows: *S. meliloti* Rm1021, *S. meliloti* AK83, *S. meliloti* BL225C, *S. meliloti* Rm41, *S. meliloti* SM11, and *S. medicae* WSM419.

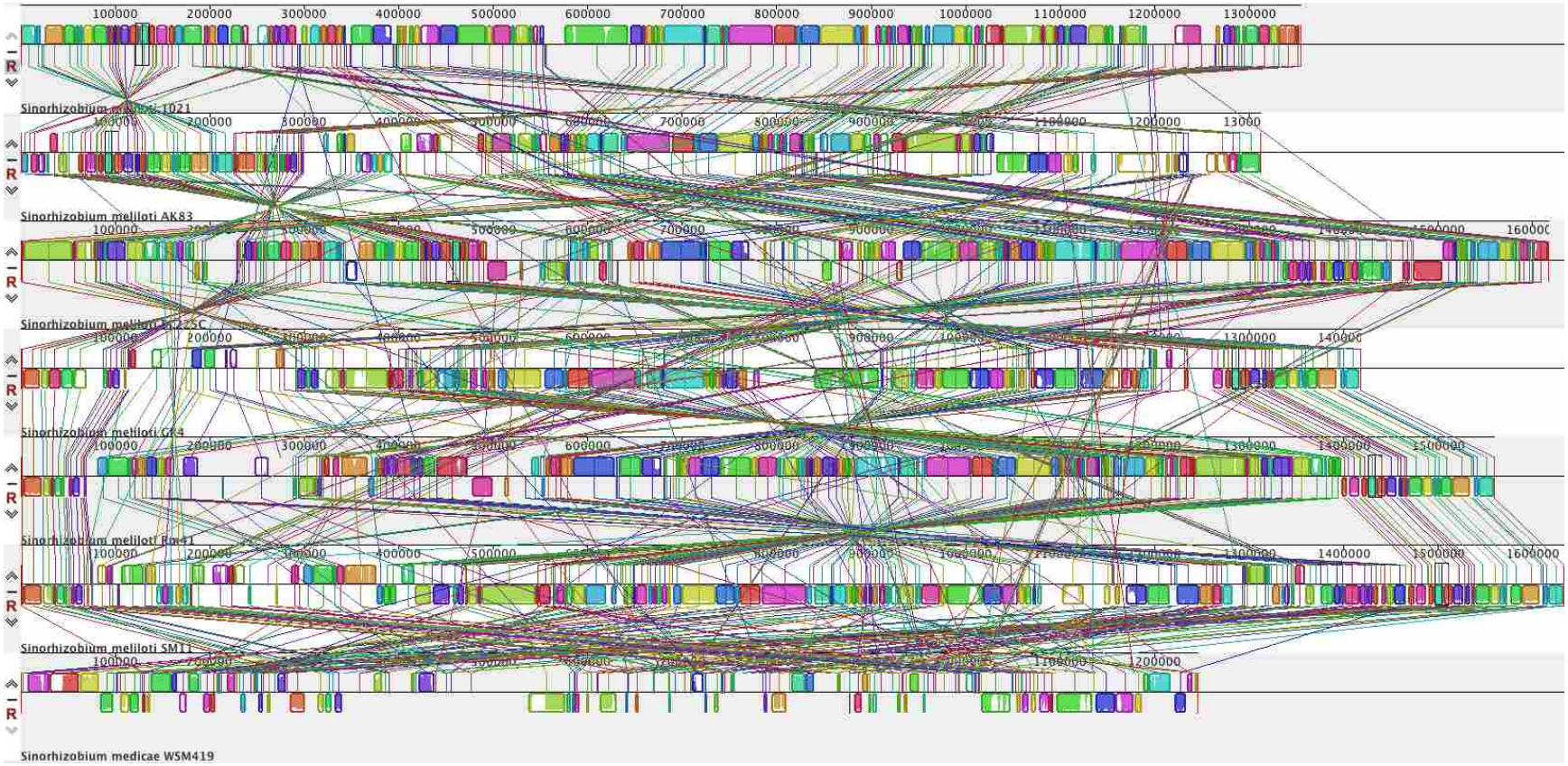

Figure S5.1-6. A genomic alignment of pSymA from several *S. meliloti* genomes and one *S. medicae* genome. Much more sequence variation can be noted, both intraspecific (between *S. meliloti* strains) and interspecific (between *S. meliloti* and *S. medicae*), compared to an alignment of pSymB. From top to bottom the strains are as follows: *S. meliloti* Rm1021, *S. meliloti* AK83, *S. meliloti* BL225C, *S. meliloti* Rm41, *S. meliloti* SM11, and *S. medicae* WSM419.

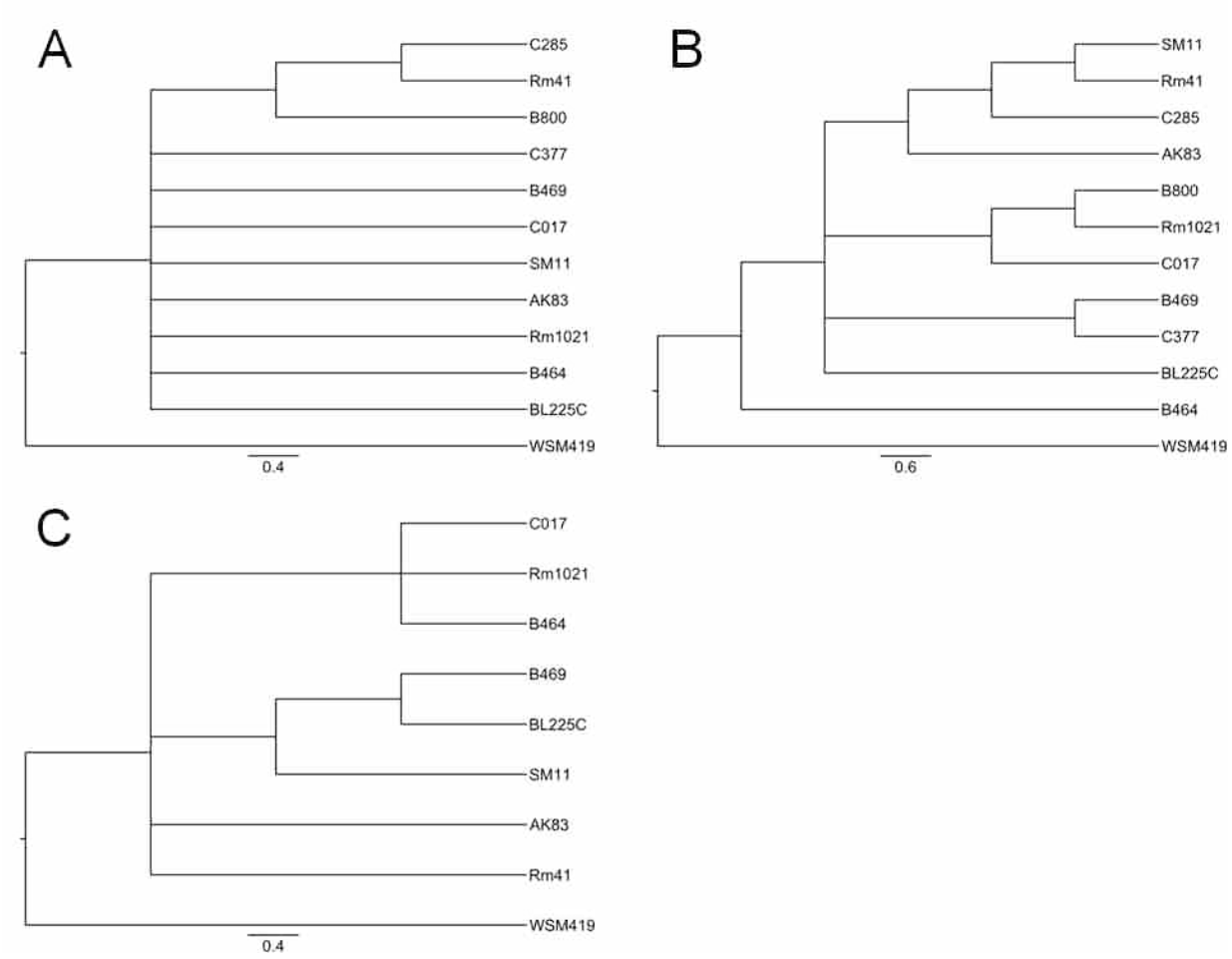

Figure S5.1-7. Comparison of consensus maximum parsimony trees for the different *S. meliloti* replicons. Sequences were concatenated from the *S. meliloti* **A.** chromosome (6 loci), **B.** pSymB (3 loci), or **C.** pSymA (3 loci). A fourth locus, anti-*nod*, was excluded on the basis that it could not be amplified from all strains.

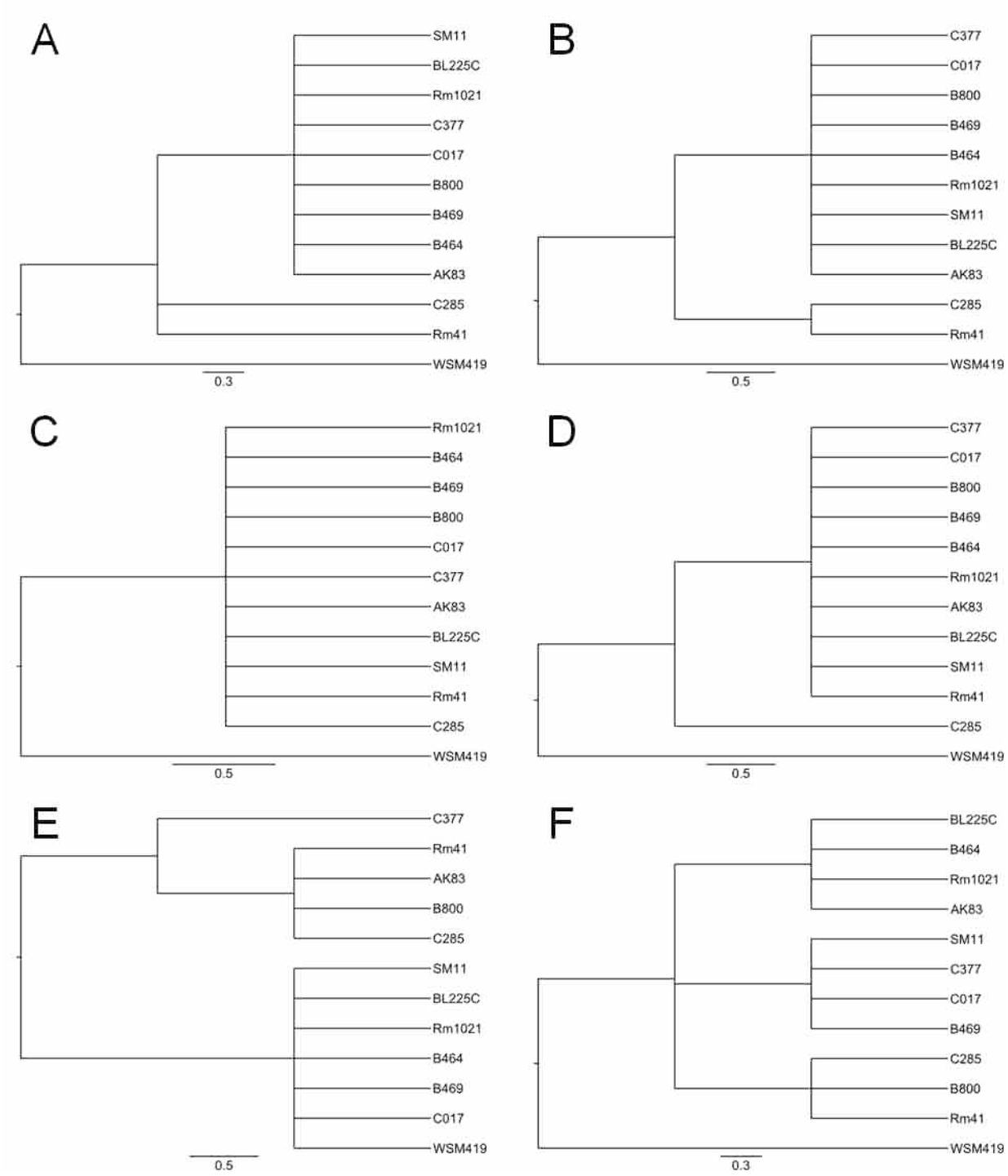

Figure S5.1-8. Comparison of consensus maximum parsimony trees for six different *S. meliloti* chromosomal loci. A. *cgm*, **B.** *cyc*, **C.** *hem*, **D.** *ndv*, **E.** *ntr*, **F.** *rkp*.

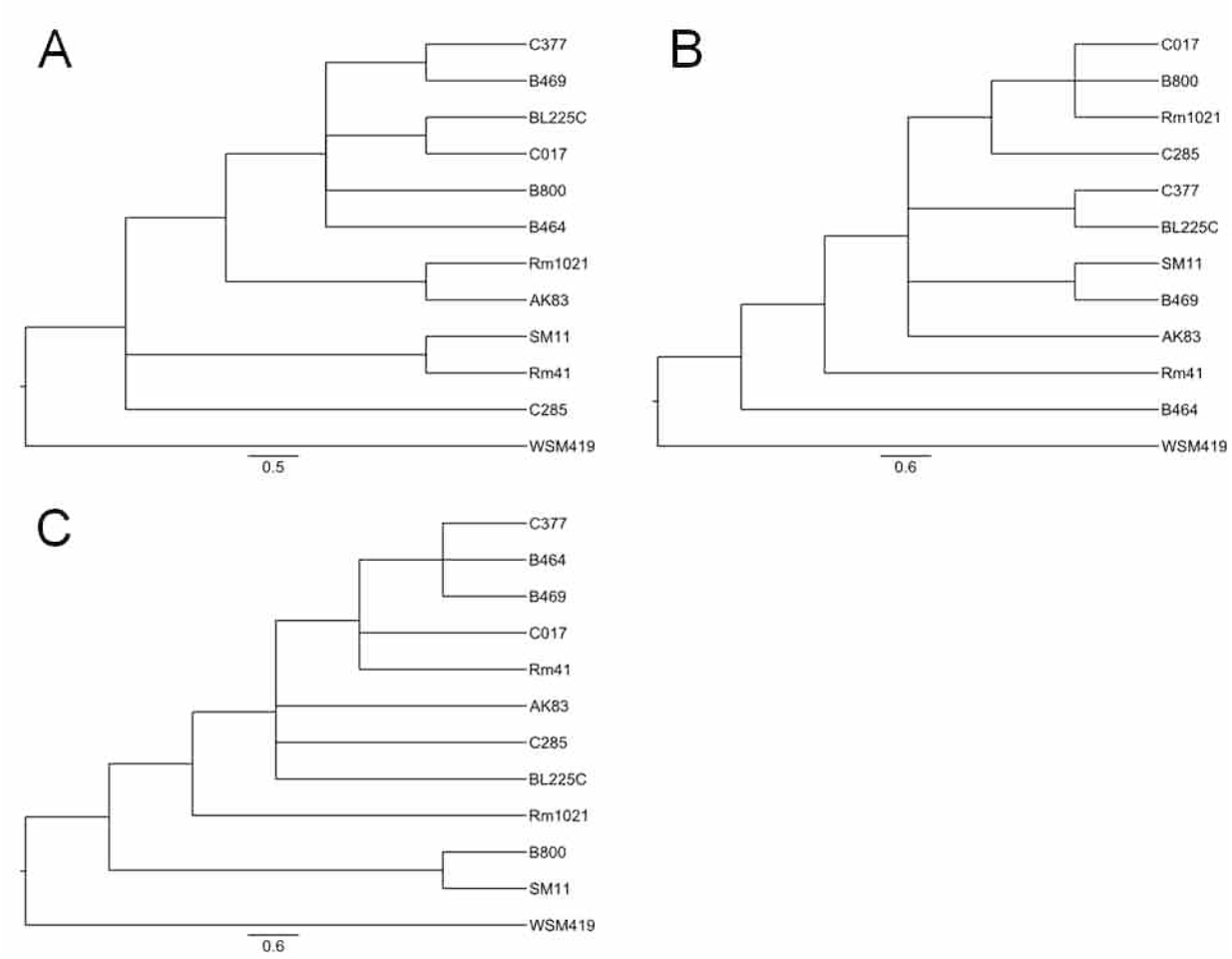

Figure S5.1-9. Comparison of consensus maximum parsimony trees for three different *S. meliloti* pSymB loci. A. *exo*, **B.** *thu*, **C.** *wg-*.

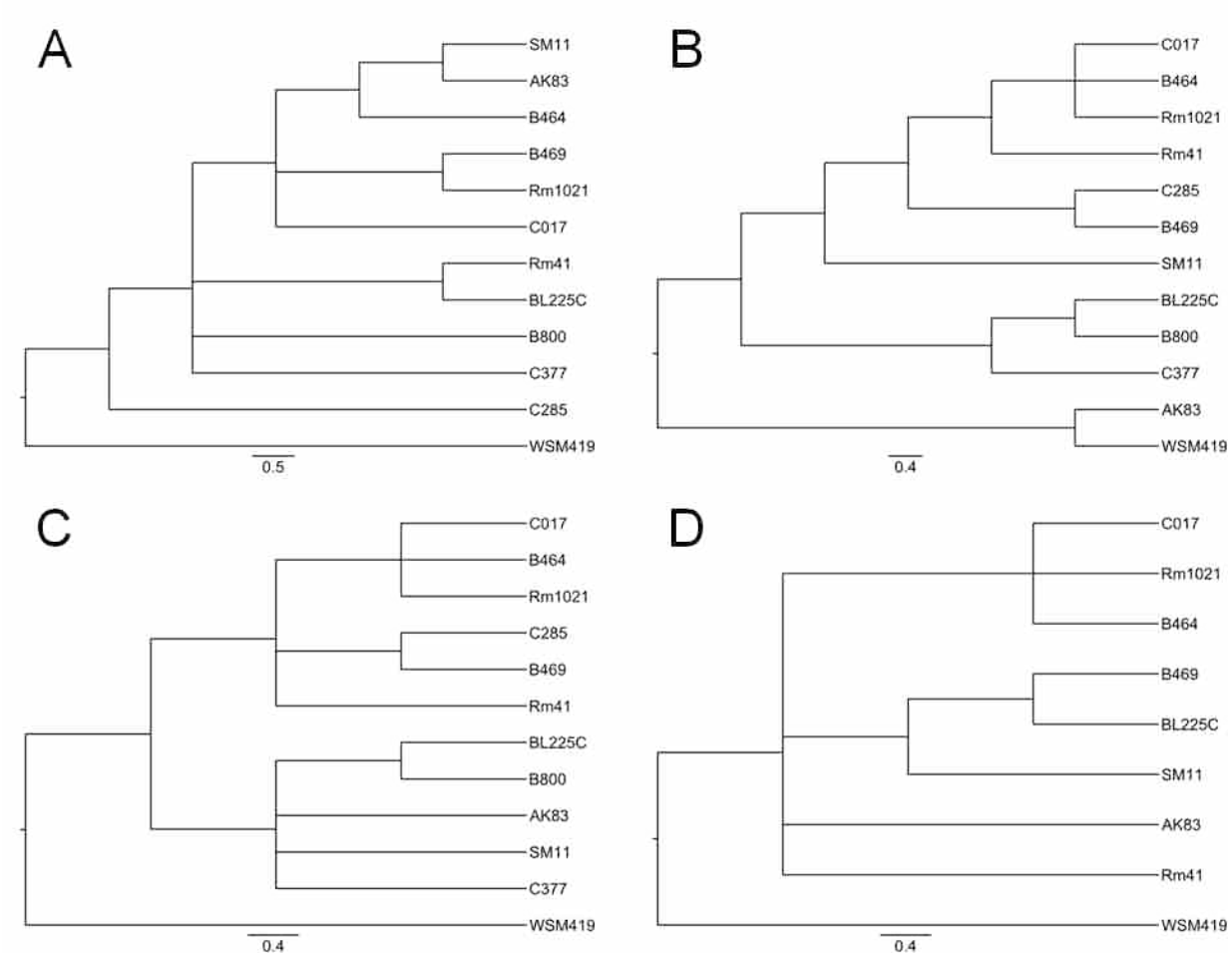

Figure S5.1-10. Comparison of consensus maximum parsimony trees for four different *S. meliloti* pSymA loci. A. *fix*, **B.** *nif*, **C.** *nod*, **D.** anti-*nod*.

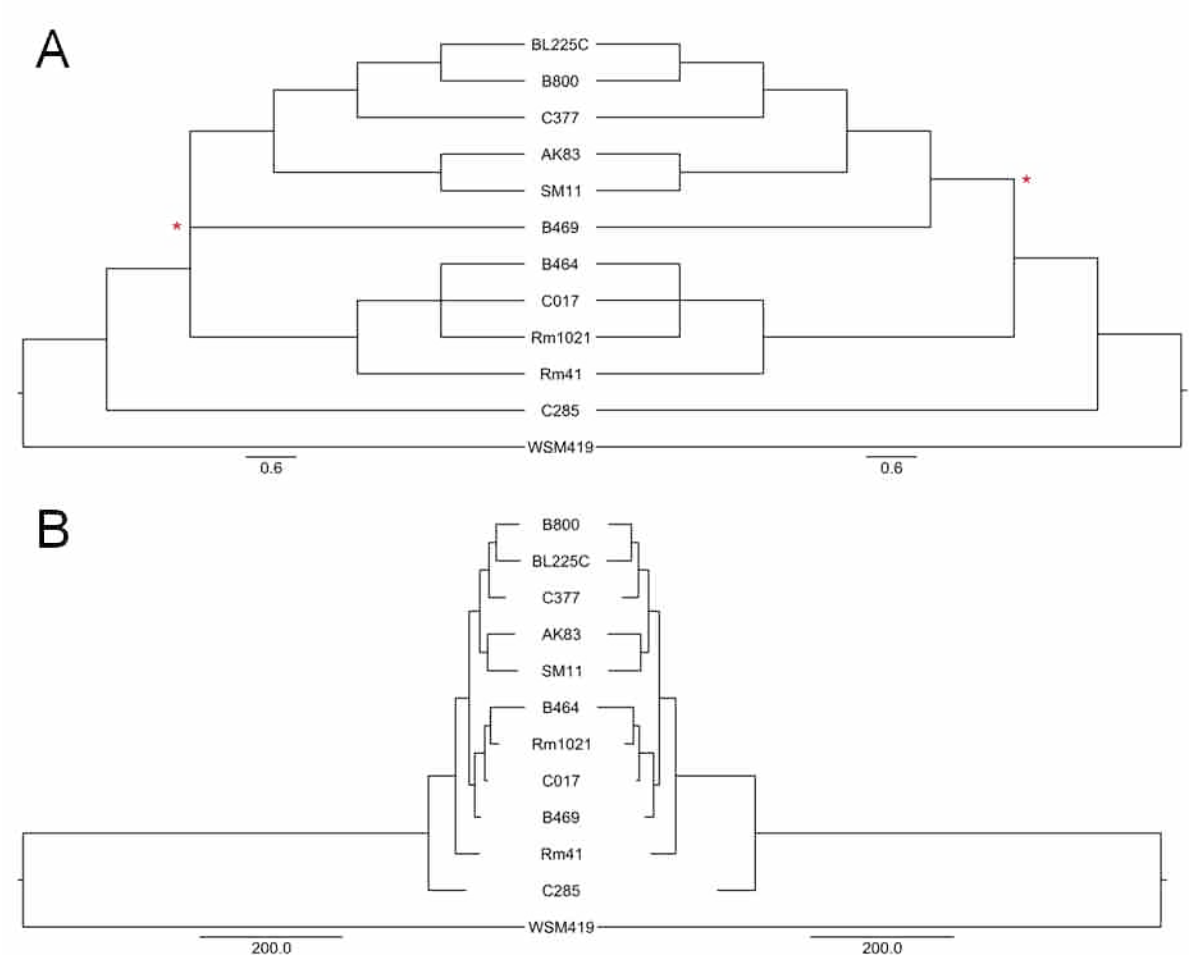

Figure S5.1-11. Including anti-*nod* has minimal effect on construction of *S. meliloti* phylogenetic trees. A. Consensus maximum parsimony trees for the pSymA replicon that exclude (left) or include (right) anti-*nod* sequences. The resolution of one branch (red asterisk) was affected and branches were lengthened. **B.** Maximum parsimony trees for *S. meliloti* genomes that exclude (left) or include (right) anti-*nod* sequences. Branches were lengthened but no other effect is apparent.

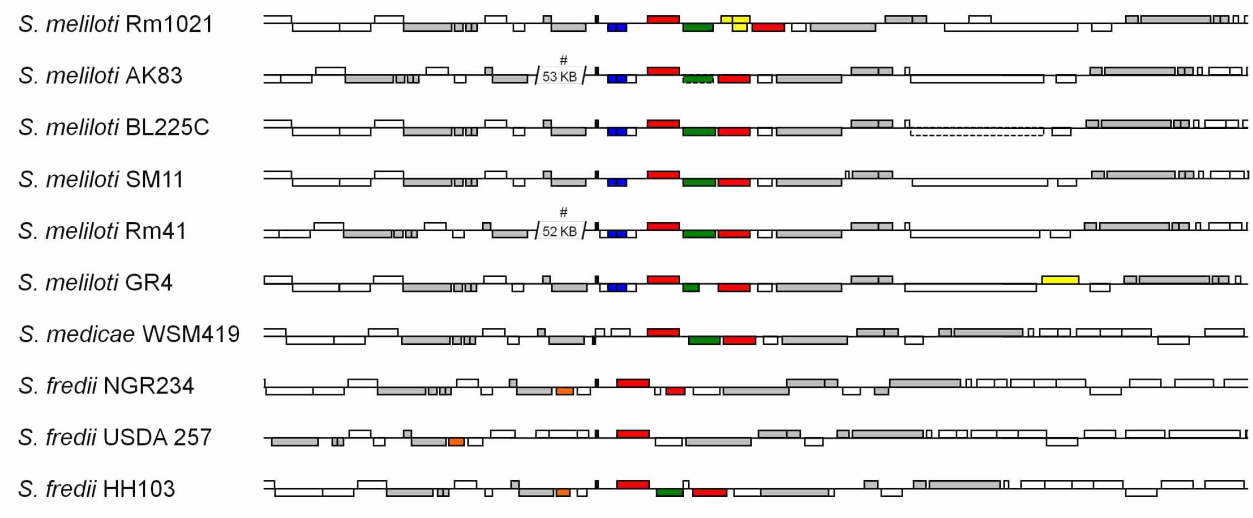

Figure S5.1-12. *ropA1* and *ropA2* reside in a hotspot for external element activity. An alignment of sequenced *Sinorhizobium* genomes reveals a nearby integrase or recombinase (green; 7/10), transposons (yellow; 2/10), pilus assembly proteins (orange; 3/10), two different prophages (#; 2/10), and toxin–antitoxin pairs (blue; 6/10). *ropA* homologues are indicated in red, a conserved tRNA-serine in black, and other nearby syntenous genes in gray. Dashed lines indicate genes which aren't annotated in that particular genome, possibly because they are pseudogenes.

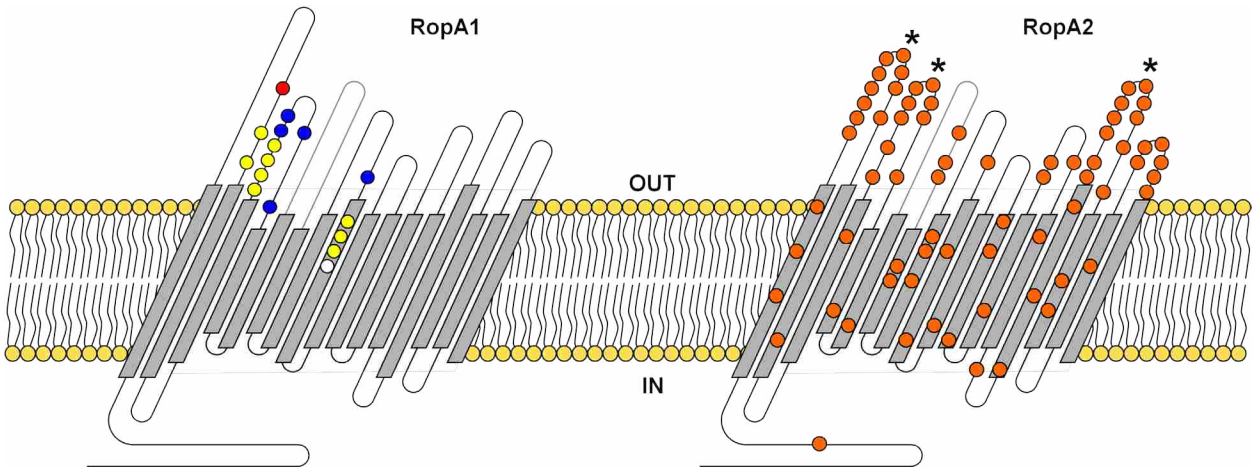

Figure S5.1-13. The majority of differences between RopA1 and RopA2 are predicted to be extracellular. RopA1 mutations conferring phage resistance are colored as in Figure 1. Amino acid differences between RopA1 and RopA2 are indicated by orange circles on RopA2. Most lie in predicted loops L1, L2, L7, and L8. Extracellular loops which are predicted to be longer or shorter in RopA2 are indicated with asterisks (*).

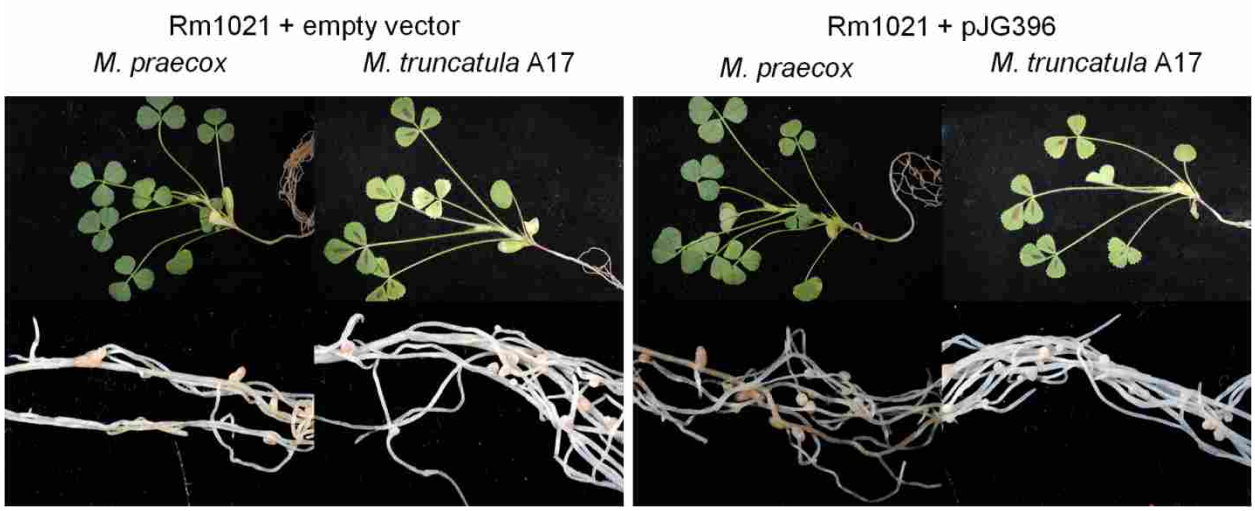

Figure S5.1-14. Overexpression of *ropA1* does not affect symbiosis. A strain overexpressing *ropA1* did not affect symbiosis with *M. truncatula* A17 or *M. praecox* as compared with an empty vector control.

Figure S5.1-15. *ropA1* and *ropA2* share little conservation of upstream sequence but predicted transmembrane strands are conserved. The –24/–12 σ^{54} promoter, IHF binding site (– indicates that it is on the antisense strand), NifA binding site, NtrC binding site, and ribosome binding site (RBS) are indicated in bold. Conserved *ropA* upstream palindrome sequences (CRUP) are indicated in bold/underline. The signal peptide is underlined. Transmembrane β strands are indicated in red, extracellular loops in blue, and intracellular turns in green. The conserved terminal phenylalanine (5) is indicated in bold red.

```
                                                  -24         -12
ropA1 CGCGGAGCGCTCTTGCCCGGTTTTATCCCCCTTTCTGGCCATTTTTTGCCCCTTTCCGGAACAGAT
ropA2 GAGCCAGCCGAGATATGCGGCTCCGCGGTATTCCGGCGGCGGCAACGCACATGTCGAATGGTGCTA

                                                                    -IHF
ropA1 GGCCGGGAGTGACCTCGCGCTAAACCCCCGTCGCTAAAGGATTTTCTTAACGCTTTGTTAATTCCC
ropA2 CTCGCGCTATCGCTCGTACCGGGCCTGGGACAATACGTTCCAGCCCTATAACGGCCCGAGGCGCCA
                                                                   CRUP

            CRUP                                NifA
ropA1 TCGGGCCGTCTGGGGCGATTTTGTGTCGTTTCAGCAACAGGGATCAGGGAAGGGCAGGGCAAAGGA
ropA2 GTGCTATTCGCCCTATAGCTGATGAGTTTCGCACCTAAGACGTGAGTCTGAAGGGAGGAGCCGGCG
                                                  NtrC

ropA1 GTCGATCTGGTAAGTTGGTCGTTGCGCGCCCCACCCGGTTTTGTATTTTCAACTCCGGAAGCAAGG
ropA2 CGCCTCCCTCGCCCGCATGCTCGGGCCGTCATTCTTGCGCGGCACCGACAGTTCATCACCGAATCC
                                 CRUP

ropA1 GCGGTTGATCGTCCGACTTCTTGAACGTTGGGGATGGGGCAGGGAACGCTGTCCTTCAGCAAAAAA
ropA2 CCTGCCATTTACCGCCTGATTCATCCGCGGCATGTAAGCCCGTGGCCTTGTATCTCCTATCCAGTC

RopA1                                  RBS      M  N  I  K  S  L  L  L  G
ropA1 AACCCAGACCCGTTTGAAACTTTTGACTGGAGGTCAGAAATGAACATCAAGAGCCTTCTTCTCGGC
           ||   | |||    || | ||||||||||||||||||||| |||||||||||||||| |||
ropA2 TCATAGGAATCATTTCTCACGTCTGACTGGAGGTCAGAAATGAATATCAAGAGCCTTCTTCTTGGC
RopA2                                  RBS      M  N  I  K  S  L  L  L  G

RopA1 S  A  A  A  L  A  A  V  S  G  A  Q  A  A  D  A  I  V  A  A  E  P
ropA1 TCCGCTGCTGCTCTCGCAGCAGTCTCCGGCGCCCAGGCTGCCGACGCGATCGTCGCTGCCGAGCCG
      ||  ||||||||| |  ||  |||||||||||||| ||  |||||||||||  |||||||||||  ||||||
ropA2 TCGGCTGCTGCTATGGCGGCAGTCTCCGGCGCGCATGCTGCCGACGCTATCGTCGCTGCAGAGCCG
RopA2 S  A  A  A  M  A  A  V  S  G  A  H  A  A  D  A  I  V  A  A  E  P

RopA1 E  P  M  E  Y  V  R  V  C  D  A  F  G  T  G  Y  F  Y  I  P  G  T
ropA1 GAGCCCATGGAATACGTTCGCGTCTGCGACGCTTTCGGCACGGGCTACTTCTACATTCCGGGCACG
      |||||  ||||||||||| |||||  |||||  ||||||||||  ||||||||||||||||  |||||||||
ropA2 GAGCCTATGGAATACGTCCGCGTTTGCGATGCTTTCGGCATGGGCTACTTCTACATCCCGGGCACG
RopA2 E  P  M  E  Y  V  R  V  C  D  A  F  G  M  G  Y  F  Y  I  P  G  T

RopA1 E  T  C  L  K  I  G  G  F  I  R  I  Q  G  E  F  G  R  D  E  A  D
ropA1 GAAACCTGCCTCAAGATCGGCGGCTTCATCCGTATCCAGGGTGAATTCGGTCGTGACGAAGCTGAC
      ||||||||||||||||||||||||  |||||  | |||||  || |||||  ||||||  ||  || |
ropA2 GAAACCTGCCTCAAGATCGGCGGCTATATCCGCGTTCAGGGCGACTTCGGCCGTGAC-AA--TGTC
RopA2 E  T  C  L  K  I  G  G  Y  I  R  V  Q  G  D  F  G  R  D  N    V
```

```
RopA1 D  R  W  N  Q  E  D  F  G  G  Q  S  T  S  D  W  D  M  F  S  R  A
ropA1 GATCGCTGGAACCAGGAAGACTTCGGTGGTCAGTCCACGTCGGACTGGGACATGTTCTCCCGCGCC
       |||  ||  |||||      |||| || | |   |||||| || ||||||||||||||||||||
ropA2 AATC-ATCTGACCAGC-----TTCGATGATGACGGCACGTCCGATTGGGACATGTTCTCCCGCGCT
RopA2 N  H  L  T  S       F  D  D  D  G  T  S  D  W  D  M  F  S  R  A

RopA1 Y  I  S  I  D  A  K  S  D  T  E  Y  G  T  L  T  G  F  F  A  A  E
ropA1 TACATCTCGATCGACGCGAAGAGCGACACCGAATACGGCACGCTCACCGGCTTCTTCGCTGCTGAG
      ||||||||| ||||||||||||||| ||||| |||||||||||||||||||||||||| | |||
ropA2 TACATCTCGTTCGACGCGAAGAGCGATACCGAGTACGGCACGCTCACGGGCTTCTTCGCTTCCGAG
RopA2 Y  I  S  F  D  A  K  S  D  T  E  Y  G  T  L  T  G  F  F  A  S  E

RopA1 F  N  A  D  N  D  T  D           E  G  D  S  L  I  D  V  D  E  A
ropA1 TTCAACGCCGACAACGACACCGAC---------GAAGGCGACAGCCTGATCGACGTCGATGAAGCC
      |||||  |||| |  ||  | | |         |  | ||||||||  |||||||||| ||||||
ropA2 TTCAATTCCGATACGGATTCGGCCTCGGACGGCGGCGGCGACAGCCACATCGACGTCGACGAAGCC
RopA2 F  N  S  D  T  D  S  A  S  D  G  G  G  D  S  H  I  D  V  D  E  A

RopA1 Y  I  Q  L  G  G  F  K  A  G  F  F  Y  S  W  W  D  K  G  L  N  G
ropA1 TACATCCAGCTCGGCGGCTTCAAGGCCGGCTTCTTCTACAGCTGGTGGGATAAGGGCCTGAACGGC
      ||||||| |||||||||||||  |||||||||||||||||||||||||| ||||||||||||||||
ropA2 TACATCGAGCTCGGCGGCTTCCGGGCCGGCTTCTTCTACAGCTGGTGGGACAAGGGCCTGAACGGC
RopA2 Y  I  E  L  G  G  F  R  A  G  F  F  Y  S  W  W  D  K  G  L  N  G

RopA1 E  T  D  S  L  G  N  V  T  E  F  N  S  L  A  Y  I  Y  D  G  G  T
ropA1 GAAACCGACTCGCTCGGTAACGTCACCGAGTTCAACTCGCTCGCCTACATCTACGACGGCGGCACC
      ||||||||||  ||||  |   |||||||||||||||| ||  ||| | ||| ||||||||||
ropA2 GAAACCGACTCCATCGGCGAGAACACCGAGTTCAACTCGGTCCGCTATACCTATGACGGCGGCACT
RopA2 E  T  D  S  I  G  E  N  T  E  F  N  S  V  R  Y  T  Y  D  G  G  T

RopA1 F  Q  A  G  I  S  V  D  E  L  E  G  T  S  T  K  A  N  G  V  G  V
ropA1 TTCCAGGCCGGTATCTCGGTTGACGAACTCGAAGGCACGAGCACCAAGGCCAACGGCGTTGGCGTT
      |||||||  |||   |  | || ||||||||||| || ||    |||||| | |||||||| || ||
ropA2 TTCCAGGTCGGCGTTGCTGTCGACGAACTCGAGGGTACTTCGACCAAGCCGAACGGCGTCGGGGTC
RopA2 F  Q  V  G  V  A  V  D  E  L  E  G  T  S  T  K  P  N  G  V  G  V

RopA1 T  G  I  V  S  A  T  L  G  G  V  A  F  D  L  L  G  S  Y  D  T  E
ropA1 ACCGGTATCGTCTCCGCGACGCTCGGCGGCGTTGCTTTCGACTTGCTCGGCAGCTACGACACCGAG
         || ||||| || ||  |||||||||||||  |||||||| |||||||||||||||||| ||
ropA2 GAAGGCATCGTATCGGCTTCGCTCGGCGGCGTGACTTTCGACCTGCTCGGCAGCTACGACACGGAA
RopA2 E  G  I  V  S  A  S  L  G  G  V  T  F  D  L  L  G  S  Y  D  T  E

RopA1 W  E  E  G  A  V  R  G  L  L  S  A  D  L  G  P  G  T  L  Q  V  A
ropA1 TGGGAAGAAGGCGCTGTCCGCGGCTTGCTCTCGGCAGACCTCGGCCCGGGCACCCTTCAGGTCGCC
      |||||||||||||  |||| | | ||||||| || ||   ||| ||||||||| | |||||||||
ropA2 TGGGAAGAAGGCGCGATCCGTGCCCTGCTCTCCGCGGATGCCGGTCCGGGCACCTTCCAGGTCGCC
RopA2 W  E  E  G  A  I  R  A  L  L  S  A  D  A  G  P  G  T  F  Q  V  A

RopA1 G  I  W  A  S  N  P  N  A  Y  W  A  D  S  E  W  T  V  A  A  S  Y
ropA1 GGCATCTGGGCGTCCAACCCGAACGCTTACTGGGCTGACTCCGAGTGGACCGTTGCTGCGTCTTAC
      || |||||||| |||  | ||||||| ||||||  ||||||  |||||| |||||||||| || |||
ropA2 GGTATCTGGGCATCCGGCTCGAACGCATACTGGCATGACTCGGAGTGGGCCGTTGCTGCCTCCTAC
RopA2 G  I  W  A  S  G  S  N  A  Y  W  H  D  S  E  W  A  V  A  A  S  Y
```

```
RopA1 R  F  N  A  S  D  K  F  A  I  T  P  A  A  Q  Y  W  G  S  L     Q
ropA1 CGCTTCAACGCTTCCGACAAGTTCGCGATCACCCCGGCTGCTCAGTACTGGGGTAGCCTG---CAA
      ||||||||      ||||||| | |||||||| || |  ||||| ||||||||  |||||    ||
ropA2 CGCTTCAATGTGAGCGACAAGCTTGCGATCACGCCCGGCGCTCAATACTGGGGCGGCCTGAACGAA
RopA2 R  F  N  V  S  D  K  L  A  I  T  P  G  A  Q  Y  W  G  G  L  N  E

RopA1    D  S  N  T     S  F  G  N  D     D  Q  W  R  V  G  I  T  T  D
ropA1 ---GACTCGAACACC---AGCTTCGGTAACGAC---GATCAGTGGCGCGTTGGTATCACGACCGAC
         |||  |  |  |    ||||||  |||||    || ||||||||  || ||  | ||  |||||
ropA2 AGCGACAAGCTCGGCGGCCGCTTCGACAACGATGCGGACCAGTGGCGTGTCGGCCTTACCGCCGAC
RopA2 S  D  K  L  G  G  R  F  D  N  D  A  D  Q  W  R  V  G  L  T  A  D

RopA1 Y  D  I  T  E  G  L  A  T  R  F  T  I  N  Y  T  D  P  D  D  G  D
ropA1 TACGACATCACGGAAGGTCTCGCTACCCGCTTCACGATCAACTACACCGATCCGGATGATGGTGAC
      |||||||||| |||||||||  |||||||| |||||||||||| || |||    ||| |
ropA2 TACGACATCACCGAAGGTCTCGCGACCCGCTTCGCGATCAACTACACAGACCCGAGCAATGCTCCG
RopA2 Y  D  I  T  E  G  L  A  T  R  F  A  I  N  Y  T  D  P  S  N  A  P

RopA1 E  Y  V  S  G  F  L  R  L  Q  R  D  F  *
ropA1 GAGTACGTCAGCGGCTTCCTCCGCCTGCAGCGTGACTTCTAATCTGACCTGACCTCGGTCAGTGGA
      || | ||||||||| ||||| |||||||||| ||||||||||||||||
ropA2 GATTCCGTCAGCGGTTTCCTTCGCCTGCAGCGGGACTTCTAATCTGACCTAGAGCGGTATGTATTT
RopA2 D  S  V  S  G  F  L  R  L  Q  R  D  F  *
```

5.2 Supplemental Tables

Table S5.2-1. Strains and plasmids used in **§2.1**.

Strain	Description[a]	Reference
C017	*S. meliloti* N6B7 (Sm^R)	(285) and this study
B469	*S. meliloti* 74B17 (Sm^R)	(285) and this study
B800	*S. meliloti* M98 (Sm^R)	(285) and this study
C377	*S. meliloti* M256 (Sm^R)	(285) and this study
C017-GOC	C017 cured of accessory plasmid pHRC017	This study
B469-GOC	B469 cured of accessory plasmid pHRB469	This study
B800-GOC	B800 cured of accessory plasmid pHRB800	This study
C377-GOC	C377 cured of accessory plasmid pHRC377	This study
B001	DH5α harboring helper plasmid pRK600	(286)
C58	Wildtype *A. tumefaciens*; harbors two plasmids: pAtC58 and pTiC58	(287)
UBAPF2	*A. tumefaciens* C58 derivative cured of pAtC58 and pTiC58 (Rf^R)	(125)
C241	C017 with pJG461 integrated into pHRC017 (Sm^R, Nm^R)	This study
C382	C58 harboring pHRC017 from C241 (Rf^R, Nm^R),	This study
1021	*S. meliloti* SU47 (Sm^R); harbors no accessory plasmids	(288)
B464	*S. meliloti* N6B2 (Sm^R); harbors no accessory plasmids	(285) and this study
C285	*S. meliloti* 128A2 (Sm^R); harbors no accessory plasmids	(285) and this study
C131	*S. meliloti* NRG23 (Sm^R); harbors one accessory plasmid	(289)
C402	B469-GOC with pJG505 looped in (Sm^R, Tc^R), =B469-GOC*tet*	This study
C406	B464 with pJG505 looped in (Sm^R, Tc^R), =B464*tet*	This study

Plasmid	Description[a]	Reference
pRK600	Self-transmissible helper plasmid (Cm^R)	(290)
pJG194	2.2 kb mobilizable suicide vector (Km^R/Nm^R)	(123)
pJG461	Sequence upstream of pHR-C017 *acdS* cloned into pJG194 (Km^R/Nm^R)	This study
pJG463	Sequence upstream of pHR-B469 *acdS* cloned into pJG194 (Km^R/Nm^R)	This study
pJG476	Sequence upstream of pHR-B800 *acdS* cloned into pJG194 (Km^R/Nm^R)	This study
pJG499	Sequence upstream of pHR-C377 *traA* cloned into pJG194 (Km^R/Nm^R)	This study
pJG392	Small mobilizable suicide vector with pUC *oriV* and *tetR–tetA* (Tc^R)	This study
pJG505	pJG392 with *rhaK–icpA* intergenic region from 1021 (Tc^R)	This study

[a] Ap^R, ampicillin resistance; Cm^R, chloramphenicol resistance; Km^R/Nm^R, kanamycin/neomycin resistance; Rf^R, rifampicin resistance; Sm^R, streptomycin resistance; Tc^R, tetracycline resistance.

Table S5.2-2. Primers used in **§2.1**.

Primer	Sequence	Directionality	Purpose
oJG1127	CGCGGATCCAGTCGACGAGGACAGCCTG	Forward	Cloning of a region near *acdS* of pSMeSM11a
oJG1128	CGCTCTAGAAGATCGTCGATGCCACCTC	Reverse	Cloning of a region near *acdS* of pSMeSM11a
oMC014	CGCGGATCCGCTCAAGATTGACGACGATCG	Forward	Cloning of *rhaK–icpA* intergenic region of Rm1021
oMC015	CGCGGATCCGTGGATACCGGACCGAAACC	Reverse	Cloning of *rhaK–icpA* intergenic region of Rm1021
oMC073	ACAAATGGAAGGTGTTTCGC	Forward	Detection of *orf00007* of pHRC017
oMC074	ATCGCAGCGCTGATTAACTT	Reverse	Detection of *orf00007* of pHRC017
oMC081	TGTCGGCAATATCCTGATGA	Forward	Detection of *orf00087* of pHRC017
oMC082	ATAGGCGGGATAGGCGTAGT	Reverse	Detection of *orf00087* of pHRC018
oMC087	CGCGGATCCGGTATGCTCCGCTTTCGTGC	Forward	Cloning of a region near *traA* of pSMeSM11a
oMC088	CGCTCTAGACGGCTATGCAGCCCAAAACC	Reverse	Cloning of a region near *traA* of pSMeSM11a
oMC209	ATGACGTCGACCGTGACCG	Forward	Detection of *orf00016* of pHRC017
oMC210	TCAGTCGGCAGCATTCTTGC	Reverse	Detection of *orf00016* of pHRC017
oMC211	TGGAGACGGTCTGTTCATGG	Forward	Detection of *orf00104* of pHRC017
oMC212	AAAGCCGTGGCGGTAGAGG	Reverse	Detection of *orf00104* of pHRC017
oMC213	CCGATGCAGTCAAGGTGACG	Forward	Detection of *orf00163* of pHRC017
oMC214	CGTCCTTCTTCGGCAGTTCG	Reverse	Detection of *orf00163* of pHRC017
oMC215	CTATGGTTTGCTAGCCGTCG	Forward	Detection of *orf00227* of pHRC017
oMC216	CTTCCGCTGCTTGAAATAGTCC	Reverse	Detection of *orf00227* of pHRC017
oMC217	CCTGACGAATCTGCAGGACG	Forward	Detection of *orf00256* of pHRC017
oMC218	GATCGTGTCAGCGACCTTGG	Reverse	Detection of *orf00256* of pHRC017
oMC219	CCCTGGAAAGGTCGATCTCG	Forward	Detection of *orf00290* of pHRC017
oMC220	GCATATAGATGCCGTGCTCG	Reverse	Detection of *orf00290* of pHRC017
oMC221	CTGCAGGAAATGTCGTCTGG	Forward	Detection of *orf00365* of pHRC017
oMC222	GAGACGAGGAATATGCTGTTGG	Reverse	Detection of *orf00365* of pHRC017
oMC223	CCGTCTCACAACAAGAACAGC	Forward	Detection of *orf00407* of pHRC017
oMC224	AACGTGGTGCTTTCGCTTCC	Reverse	Detection of *orf00407* of pHRC017
oMC225	ATTCCTTGATCTGCCGGAGG	Forward	Detection of *orf00437* of pHRC017
oMC226	CCGAGTTCCGGTTGCTTCTC	Reverse	Detection of *orf00437* of pHRC017
oMC227	CGGATTCCTGCTGCTCATCC	Forward	Detection of *orf00479* of pHRC017
oMC228	CGCAACCAAGAGAGCGATCG	Reverse	Detection of *orf00479* of pHRC017
oMC231	ATTGTCGGTGATGAGATCGTGC	Forward	Detection of *orf00567* of pHRC017
oMC232	TTGGTTCTCATGCCTTCCTGG	Reverse	Detection of *orf00567* of pHRC017
oMC233	ACGCTCAGACTACAGCAATGG	Forward	Detection of *orf00623* of pHRC017
oMC234	AGCTGGGCTCGTTGAACTGG	Reverse	Detection of *orf00623* of pHRC017
oMC235	CGAATGGAGTCTCCTGCTCC	Forward	Detection of *orf00639* of pHRC017
oMC236	TAGTTGGCTAAGGTCTGCAGC	Reverse	Detection of *orf00639* of pHRC017
oMC241	GTTCTGGCGTACCGAGAACG	Forward	Detection of *orf00528* of pHRC017

(Table S5.2 2 continued)

Primer	Sequence	Directionality	Purpose
oMC242	GCGACAGCGTCTTGGATACG	Reverse	Detection of *orf00528* of pHRC017
oMC243	CACTTCAAGGAGGTGATCAGC	Forward	Detection of *orf00671* of pHRC017
oMC244	GGAATTCGAGCGCATCGTCG	Reverse	Detection of *orf00671* of pHRC017
oMC245	CCGTGATGCAACTGACGACC	Forward	Detection of *orf00689* of pHRC017
oMC246	GAACACCTTCTGGAAGATGTCC	Reverse	Detection of *orf00689* of pHRC017
oMC247	GGAGTTGTTCGAGACGCTGG	Forward	Detection of *orf00052* of pHRC017
oMC248	GATGGTCACGTCGGTCCTC	Reverse	Detection of *orf00052* of pHRC017
oMC249	GCGTGGATTGGTTTGCAAGC	Forward	Detection of *orf00066* of pHRC017
oMC250	CACGGCAAGTCCAAGTGTAC	Reverse	Detection of *orf00066* of pHRC017
oMC251	GAAGGCGATCGACTTTGAGC	Forward	Detection of *orf00136* of pHRC017
oMC252	GTGCAGCTTGACATCGAACG	Reverse	Detection of *orf00136* of pHRC017
oMC253	CGATTTCGGGAGACCAGTGG	Forward	Detection of *orf00198* of pHRC017
oMC254	CAGACCGCCATCACCTATCG	Reverse	Detection of *orf00198* of pHRC017
oMC255	CGGTCGTAGTTGACTTCTGG	Forward	Detection of *orf00276* of pHRC017
oMC256	CCTTGGTGTCGGCTACTTCC	Reverse	Detection of *orf00276* of pHRC017
oMC257	CAGACCTTCCACTCAACTCC	Forward	Detection of *orf00318* of pHRC017
oMC258	CTATTTCGACCTGGACCTTTGG	Reverse	Detection of *orf00318* of pHRC017
oMC259	CGAGTTTCATACTCGGCTTGC	Forward	Detection of *orf00398* of pHRC017
oMC260	CGAGAAGATCAGCGTCCACG	Reverse	Detection of *orf00389* of pHRC017
oMC261	CCAGCTCATTGATGGCTTTGC	Forward	Detection of *orf00429* of pHRC017
oMC262	GCTGGATAGCGTCCCAGTAG	Reverse	Detection of *orf00429* of pHRC017
oMC263	GTGAGACATTAGGCGTCGTC	Forward	Detection of *orf00460* of pHRC017
oMC264	GCATGTCTCTGCAACAACACC	Reverse	Detection of *orf00460* of pHRC017
oMC265	GGTCTTCGATGTCGTTTTCACG	Forward	Detection of *orf00499* of pHRC017
oMC266	GTTGAAGCATCCGCCTCATC	Reverse	Detection of *orf00499* of pHRC017
oMC267	CACCTGGGTTTCCAATGGTG	Forward	Detection of *orf00550* of pHRC017
oMC268	GTTTCTCAGCCACGGCATAG	Reverse	Detection of *orf00550* of pHRC017
oMC269	CTACGACACGCGATGCATCG	Forward	Detection of *orf00592* of pHRC017
oMC270	GCAATGGAAATCGCGTCAACG	Reverse	Detection of *orf00592* of pHRC017
oMC271	CTCTTTTCGCCAAACGCCAG	Forward	Detection of *orf00001* of pHRC017
oMC272	GCTCAAACTCGTGAAGCTCC	Reverse	Detection of *orf00001* of pHRC017
oMC273	GCGACCACCTTAGATTGAAGG	Forward	Detection of *orf00025* of pHRC017
oMC274	CAAGAGCGATATGCGTTCTGG	Reverse	Detection of *orf00025* of pHRC017
oMC275	TGCTCGCAAAGAACGACCTG	Forward	Detection of *orf00070* of pHRC017
oMC276	GATCTTGTTCTGCATGCCAACG	Reverse	Detection of *orf00070* of pHRC017
oMC277	TACCAATCTTATCGGCTGTTCC	Forward	Detection of *orf00151* of pHRC017
oMC278	GCAAATTCATCACCGGCAAGC	Reverse	Detection of *orf00151* of pHRC017
oMC279	CAGCGCGGATGTCTCTAACC	Forward	Detection of *orf00172* of pHRC017

(Table S5.2 2 continued)

Primer	Sequence	Directionality	Purpose
oMC280	CTTGGTTTTCGACGTAACCACG	Reverse	Detection of *orf00172* of pHRC017
oMC281	CTCATGAAACTTCTGCAGCGTG	Forward	Detection of *orf00223* of pHRC017
oMC282	GCTCGCTTCTTCGCTTTTCC	Reverse	Detection of *orf00223* of pHRC017
oMC283	CGCAGACAAGGCTCTGTGCT	Forward	Detection of *orf00232* of pHRC017
oMC284	GAAGCCCACCGAAATCTTTTGC	Reverse	Detection of *orf00232* of pHRC017
oMC285	CAGGACCTCAGTAAGTTGTTGC	Forward	Detection of *orf00269* of pHRC017
oMC286	CCTTTCACACTCTTCAGGACG	Reverse	Detection of *orf00269* of pHRC017
oMC287	TGGTCGCCTATCATCGGAAC	Forward	Detection of *orf00281* of pHRC017
oMC288	GTCAGAGTTGAATGGCACACG	Reverse	Detection of *orf00281* of pHRC017
oMC289	CAGATGGTCGAGAACATCTCTG	Forward	Detection of *orf00636* of pHRC017
oMC290	CTCAGTTTCTCTGGCAGTAAGG	Reverse	Detection of *orf00636* of pHRC017
oMC291	GTTGCGCGTGAAAGACTATCG	Forward	Detection of *orf00717* of pHRC017
oMC292	GGTTCATCGCGTTCTCAATCG	Reverse	Detection of *orf00717* of pHRC017

Table S5.2-3. Plants used in **§2.1**.

Plant	Accession Number	Location Origin	Source/Reference
Medicago arabica	PI 495200	France	USDA-ARS[a]
Medicago constricta	PI 495240	Greece	USDA-ARS
Medicago disciformis	PI 487317	Bulgaria	USDA-ARS
Medicago italica	PI 384640	Morocco	USDA-ARS
Medicago lesinsii	PI 534233	Israel	USDA-ARS
Medicago lupulina	PI 211605	Afghanistan	USDA-ARS
Medicago praecox	PI 495429	Greece	USDA-ARS
Medicago rugosa	PI 368962	Italy	USDA-ARS
Medicago tenoreana	PI 499161	Italy	USDA-ARS
Medicago truncatula cv. Jemalong A17	—	Australia	(291)
Medicago truncatula cv. Jemalong A20	—	Australia	(291)

[a] USDA ARS, United States Department of Agriculture Agricultural Research Service, U.S. annual Medicago species core collection (http://bldg6.arsusda.gov/pberkum/Public/sarl/bauchan/core2.html).

Table S5.2-4. Symbiosis phenotypes for various *Medicago–Sinorhizobium* pairs.

Orig. Isolate[a]	Source Host[b]	Geographic Origin[b]	Lab Name	*Medicago* Host Plant[c,d]										
				LU	IT	CO	DI	PR	A17	A20	TE	RU	AR	LE
Ve8	*sativa*	France	B639	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M249	*truncatula*	Jordan	B645	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{+}	Fix^{-}	Fix$^{+/-}$
102F28	*sativa*	USA	B646	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M207	*rigidula*	Turkey	B427	Fix^{-}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
19A4	*sativa*	Pakistan	B474	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M241	*rotata*	Jordan	B429	Fix^{-}	Nod$^{+/-}$	Fix^{-}	Nod^{-}	Nod^{-}	Nod$^{+/-}$	NT	Fix^{-}	Fix^{-}	Nod$^{+/-}$	Fix^{-}
19A9	*sativa*	Pakistan	B460	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
Sa-10	*sativa*	France	B473	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
102F85	*sativa*	Canada	B468	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
M243	*rotata*	Jordan	B634	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}
102F82	*sativa*	Canada	B432	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
74B17	*sativa*	Pakistan	B469	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}
19A5	*sativa*	Pakistan	B640	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
N6B7	*falcata*	Nepal	C017	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
102F34	*sativa*	USA	B475	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Nod$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
N6B11	*falcata*	Nepal	B476	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M24	*rigidula*	Syria	B434	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Nod$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
15A1	*sativa*	Pakistan	B466	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
M119	unspecified	Syria	B436	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M210	*noeana*	Turkey	B641	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
CC 2003	*sativa*	Australia	B437	Fix^{-}	Fix^{-}	Fix$^{+/-}$	Fix^{-}	Nod^{-}	Nod^{-}	NT	Nod^{-}	Nod^{-}	Nod^{-}	Nod^{-}
74B3	*sativa*	Pakistan	B461	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Nod^{-}	Fix^{-}	Nod^{-}
M95	*rotata*	Syria	B471	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix$^{+/-}$	Fix^{-}	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Fix^{-}
M270	*truncatula*	Jordan	B438	Fix^{-}	Fix^{-}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M56	*rotata*	Syria	B465	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
N6B13	*falcata*	Nepal	B470	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Nod^{-}	Fix^{-}	Nod^{-}
N6B1	*falcata*	Nepal	B477	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M262	*rigidula*	Jordan	B439	Fix^{-}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
19A18	*sativa*	Pakistan	B467	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
N4A7	*sativa*	Nepal	B463	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Nod^{-}
19A8	*sativa*	Pakistan	B462	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
N6B2	*falcata*	Nepal	B464	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Nod^{-}
M29	*rigidula*	Syria	B647	Fix^{+}	Fix^{+}	NT	Fix^{-}	Fix^{-}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}

(Table S5.2 4 continued)

Orig. Isolate[a]	Source Host[b]	Geographic Origin[b]	Lab Name	*Medicago* Host Plant[c,d]										
				LU	IT	CO	DI	PR	A17	A20	TE	RU	AR	LE
M193	*rigidula*	Turkey	B642	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
M268	*orbicularis*	Jordan	B643	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Nod^{-}
M272	*rigidula*	Jordan	B637	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M256	*polymorpha*	Jordan	C377	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Nod^{-}	Nod^{-}	Fix^{-}	Nod^{-}
17B6	*sativa*	Pakistan	B635	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{+}	Nod^{-}	Fix^{-}	Nod^{-}
M294	*polymorpha*	Jordan	B636	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Nod^{-}	Fix^{-}	Nod^{-}
M273	*truncatula*	Jordan	B649	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M245	*orbicularis*	Jordan	B638	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M156	*rigidula*	Syria	B447	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Nod^{-}
M247	*orbicularis*	Jordan	B644	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M246	*rotata*	Jordan	B801	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M48	*rigidula*	Syria	B448	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Nod^{-}
M181	unspecified	Syria	B449	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M124	unspecified	Syria	B301	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M76	*rigidula*	Syria	B450	Fix^{-}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M257	*rotata*	Jordan	–	Fix^{-}	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M98	*rotata*	Syria	B800	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
N4A12	*sativa*	Nepal	B799	Fix^{+}	Fix^{+}	Fix^{+}	Nod^{-}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Nod^{-}	Fix^{-}	Fix^{-}
74B15	*sativa*	Pakistan	C134	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{-}	Fix^{-}	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
74B4	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
74B12	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}
M30	*polymorpha*	Syria	B302	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M259	unspecified	Jordan	B318	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Nod^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M286	*rotata*	Jordan	B319	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M263	unspecified	Jordan	–	Fix$^{+/-}$	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Nod^{-}	Nod^{-}	Fix^{-}	Fix^{-}
56A6	*sativa*	Pakistan	C023	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Fix^{+}	NT	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}
56A16	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{+}	Fix^{+}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
56A14	*sativa*	Pakistan	B340	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Nod^{-}
128A2	*sativa*	Pakistan	C285	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	Nod^{-}	Fix^{-}	Nod^{-}
15B4	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Fix^{-}
15A6	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix$^{+/-}$	Fix^{-}	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Nod^{-}
128A3	*sativa*	Pakistan	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Nod^{-}	Fix^{-}	Fix^{-}
M10	*blancheana*	Syria	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Nod^{-}	Nod^{-}	Nod^{-}
S33	*sativa*	USA	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix$^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}

(Table S5.2 4 continued)

Orig. Isolate[a]	Source Host[b]	Geographic Origin[b]	Lab Name	*Medicago* Host Plant[c,d]										
				LU	**IT**	**CO**	**DI**	**PR**	**A17**	**A20**	**TE**	**RU**	**AR**	**LE**
102F51	*sativa*	USA	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	$Fix^{+/-}$	Fix^{-}	$Fix^{+/-}$	Fix^{-}
CC 2013	*sativa*	Australia	–	Fix^{-}	NT	$Fix^{+/-}$	Fix^{-}	Nod^{-}	Fix^{-}	NT	Nod^{-}	Nod^{-}	Fix^{-}	Nod^{-}
M22	*polymorpha*	Syria	**B339**	Fix^{+}	Fix^{+}	Fix^{+}	$Fix^{+/-}$	Fix^{+}	Fix^{+}	NT	Fix^{+}	Fix^{-}	Fix^{+}	Fix^{-}
M7	*orbicularis*	Syria	–	$Fix^{+/-}$	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{-}	Fix^{+}	Fix^{+}	Nod^{-}
M161	*noeana*	Syria	–	$Fix^{+/-}$	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{+}
M158	*noeana*	Syria	–	Fix^{+}	Fix^{+}	Fix^{+}	$Fix^{+/-}$	Fix^{+}	Fix^{+}	NT	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}
M3	*orbicularis*	Syria	–	$Fix^{+/-}$	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Nod^{-}	Fix^{+}	Fix^{+}	Fix^{+}
M254	*rotata*	Jordan	–	Fix^{-}	Fix^{-}	Fix^{+}	Fix^{-}	Fix^{-}	Nod^{-}	NT	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M2	*blancheana*	Syria	–	Fix^{-}	Nod^{-}	Nod^{-}	Nod^{-}	Nod^{-}	Nod^{-}	NT	Nod^{-}	Nod^{-}	Nod^{-}	Nod^{-}
M285	*truncatula*	Jordan	**B303**	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}
M75	*radiata*	Syria	–	Nod^{-}	Nod^{-}	Nod^{-}	Fix^{-}	Nod^{-}	Nod^{-}	NT	Nod^{-}	Nod^{-}	Nod^{-}	Nod^{-}
M280	*orbicularis*	Jordan	–	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M58	*rotata*	Jordan	**B284**	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{-}	NT	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}
M278	*rotata*	Jordan	–	Fix^{+}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	Fix^{-}	NT	Fix^{-}	Nod^{-}	Fix^{-}	Fix^{-}
M1	*orbicularis*	Syria	**B264**	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}
M205	*truncatula*	Turkey	–	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}	NT	Fix^{+}	Fix^{+}	Fix^{+}	Fix^{+}

[a] Strains in bold text are classified as *S. medicae*; all remaining strains are *S. meliloti* (122). [b] Detailed information about host and geographic origins for these strains is from (303). [c] Host plant abbreviations: LU, *M. lupulina*; IT, *M. italica*; CO, *M. constricta*; DI, *M. disciformis*; PR, *M. praecox*; A17, *M. truncatula* cv. A17; A20, *M. truncatula* cv. A20; TE, *M. tenoreana*; RU, *M. rugosa*; AR, *M. arabica*; LE, *M. lesinsii*. [d] Fix+, effective N-fixing pairs; Fix+/–, less effective N-fixing pairs; Fix–, abortively nodulating pairs; Nod+/–, pairs that result in barely perceptible nodulation; Nod– pairs that result in no visible nodulation. NT = not tested.

Table S5.2-5. Performance of various engineered strains on *M. truncatula* A20.

Strain	*M. t.* A20[a]	Strain	*M. t.* A20[a]
B469	Fix– *		
B469-GOC*tet*	Fix+	Rm1021*tet*	Fix+
B469-GOC*tet*/pB469	Fix– *	Rm1021*tet*/pB469	Fix+
B469-GOC*tet*/pC017	Fix– *	Rm1021*tet*/pC017	Fix–
B469-GOC*tet*/pB800	Fix– *	Rm1021*tet*/pB800	Fix+
B469-GOC*tet*/pC377	Fix– *	Rm1021*tet*/pC377	Fix+
C017	Fix– *		
C017-GOC*tet*	Fix+	B464*tet*	Fix+
C017-GOC*tet*/pB469	NT[b]	B464*tet*/pB469	Fix–
C017-GOC*tet*/pC017	NT	B464*tet*/pC017	Fix– *
C017-GOC*tet*/pB800	NT	B464*tet*/pB800	Fix–
C017-GOC*tet*/pC377	NT	B464*tet*/pC377	Fix–
B800	Fix– *		
B800-GOC*tet*	Fix+	C285*tet*	Fix+
B800-GOC*tet*/pB469	Fix+	C285*tet*/pB469	Fix– *
B800-GOC*tet*/pC017	Fix+	C285*tet*/pC017	Fix– *
B800-GOC*tet*/pB800	Fix–	C285*tet*/pB800	Fix–
B800-GOC*tet*/pC377	Fix+	C285*tet*/pC377	Fix–
C377	Fix– *		
C377-GOC*tet*	Fix+	C131*tet*	Fix+
C377-GOC*tet*/pB469	Fix+	C131*tet*/pB469	Fix+
C377-GOC*tet*/pC017	Fix–	C131*tet*/pC017	Fix–
C377-GOC*tet*/pB800	Fix+	C131*tet*/pB800	Fix– *
C377-GOC*tet*/pC377	Fix–	C131*tet*/pC377	Fix–
		None	Nod–

[a] Symbiotic phenotypes were scored approximately 30 days post-inoculation (dpi). Fix^+, effective N-fixing pairs; Fix^-, abortively nodulating pairs; Fix^- *, pairs exhibiting suppressible incompatibility (the GOC phenomenon).
[b] NT, not tested; for reasons as yet unknown we were unable to generate these strains.

Table S5.2-6. PCR targets for comparative analysis of HR plasmids. ORF designations refer to the annotated pHRC017 plasmid. PCR tests were performed using lysates from *A. tumefaciens* strain UBAPF2 harboring none (UBAPF2) or one of the four plasmids shown. Presence or absence of target DNA is indicated with (+) or (–), respectively.

target	putative function	UBAPF2	pHRC017	pHRB469	pHRB800	pHRC377
orf00001	plasmid partitioning protein	–	+	+	–	–
orf00007	plasmid partitioning protein	–	+	+	–	–
orf00016	prevent-host-death family protein	–	+	+	+	+
orf00025	hypothetical protein	–	+	–	–	–
orf00054	cytochrome P450 monooxygenase protein	–	+	+	+	–
orf00066	magnesium and cobalt transport protein	–	+	+	+	–
orf00070	LuxR family transcriptional regulator	–	+	+	+	–
orf00087	1-aminocyclopropane-1-carboxylate deaminase	–	+	+	+	–
orf00104	thiamine biosynthesis oxidoreductase	–	+	+	+	–
orf00136	NAD-dependent formate dehydrogenase	–	+	+	–	–
orf00151	diguanylate cyclase/phosphodiesterase (GGDEF)	–	+	+	–	–
orf00163	chaperonin	–	+	+	+	+
orf00172	acyl carrier protein	–	+	+	+	+
orf00198	XRE family transcriptional regulator	–	+	+	–	–
orf00223	MutT/NUDIX family NTP pyrophosphohydrolase	–	+	+	+	+
orf00227	sodium–alanine/glycine symporter family protein	–	+	+	+	+
orf00232	outer membrane protease	–	+	+	+	–
orf00256	HipA-like protein	–	+	+	–	+
orf00269	major facilitator superfamily transporter	–	+	+	–	+
orf00276	thioredoxin	–	+	+	+	+
orf00281	necrosis- and ethylene-inducing protein	–	+	+	–	–
orf00290	aromatic amino acid aminotransferase	–	+	+	–	+
orf00318	AAA ATPase central domain protein	–	+	+	–	+
orf00365	conjugative transfer relaxase	–	+	–	–	–
orf00389	Xaa-Pro metalloaminopeptidase	–	+	–	–	+
orf00407	X-Pro dipeptidyl-peptidase domain protein	–	+	–	–	–
orf00429	aminotransferase class III	–	+	–	–	–
orf00444	glutamate dehydrogenase	–	+	–	–	+
orf00460	oligopeptide ABC transporter, ATPase component	–	+	–	–	–

orf00479	polyamine ABC transporter, permease component	–	+	–	–	–
orf00499	MatE efflux family protein	–	+	–	–	–
orf00528	GntR family transcriptional regulator	–	+	+	–	–
orf00550	succinoglycan biosynthesis protein	–	+	+	–	–
orf00567	type IV secretion/conjugal transfer protein	–	+	+	–	–
orf00592	amidohyrolase	–	+	–	–	–
orf00623	ligase/carboxylase	–	+	–	–	–
orf00636	LuxR family transcriptional regulator	–	+	+	–	–
orf00639	redoxin	–	+	+	–	+
orf00671	glutamate synthase, large subunit protein	–	+	+	–	+
orf00689	NADP-dependent isocitrate dehydrogenase	–	+	+	–	+
orf00717	plasmid partitioning protein	–	+	+	–	–

Table S5.2-7. Strains, plasmids, and bacteriophages used in **§2.2**.

Strain	Description[a]	Reference
B001	DH5α harboring helper plasmid pRK600; Cm^R	(286)
C017	*S. meliloti* N6B7; Sm^R	(110)
B469	*S. meliloti* 74B17; Sm^R	(110)
B800	*S. meliloti* M98; Sm^R	(110)
B464	*S. meliloti* N6B2; Sm^R	(110)
C285	*S. meliloti* 128A2; Sm^R	(110)
Rm1021	*S. meliloti* SU47; Sm^R	(288)
C58	Wildtype *A. tumefaciens* biovar 1; harbors two accessory plasmids: pAtC58 and pTiC58	(287)
UBAPF2	*A. tumefaciens* LBA290 cured of pAtC58 and pTiC58; Rf^R	(125)
C406	B464 with pJG505 looped into the symbiotically-neutral *rhaK–icpA* intergenic region; Sm^R, Tc^R	(110)
C407	C285 with pJG505 looped into the symbiotically-neutral *rhaK–icpA* intergenic region; Sm^R, Tc^R	(110)
B171	*exoY*::Tn*5*-110 (formerly R2D1; an EPS I-defective symbiotic mutant); Sm^R, Nm^R	(286)
C270	UBAPF2 harboring pSmeN6B7a::Tn*5*-310 from C017; Rf^R, Nm^R	This study
C272	UBAPF2 harboring pSmeN6B7b::Tn*5*-310 from C017; Rf^R, Nm^R	This study
C274	UBAPF2 harboring pSme74B17a::Tn*5*-310 from B469; Rf^R, Nm^R	This study
C382	UBAPF2 harboring pHRC017; Rf^R, Nm^R	This study
C699	UBAPF2 harboring pHRC017 ΔQ1; Rf^R, Nm^R	This study
C650	UBAPF2 harboring pHRC017 ΔQ2; Rf^R, Nm^R	This study
C654	UBAPF2 harboring pHRC017 ΔQ1Q2; Rf^R, Nm^R	This study
C251	UBAPF2 harboring pHRB469 Rf^R, Nm^R	This study
C645	UBAPF2 harboring pHRB469 Δ'Q1' Rf^R, Nm^R	This study
C659	UBAPF2 harboring pHRB469 Δ'Q2' Rf^R, Nm^R	This study
C655	UBAPF2 harboring pHRB469 Δ1Q1Q2' Rf^R, Nm^R	This study
47-1	B800 with pHRB800 ΔX	This study
C436	C406 harboring pHRC017; Sm^R, Nm^R, Tc^R	(110)
C703	C406 harboring pHRC017 ΔQ1; Sm^R, Nm^R, Tc^R	This study
C710	C407 harboring pHRC017 ΔQ2; Sm^R, Nm^R, Tc^R	This study
C708	C407 harboring pHRC017 ΔQ1Q2; Sm^R, Nm^R, Tc^R	This study
C434	C406 harboring pHRB469; Sm^R, Nm^R, Tc^R	(110)
C652	C406 harboring pHRB469 Δ'Q1'; Sm^R, Nm^R, Tc^R	This study
C663	C406 harboring pHRB469 Δ'Q2'; Sm^R, Nm^R, Tc^R	This study
C660	C406 harboring pHRB469 Δ'Q1Q2'; Sm^R, Nm^R, Tc^R	This study
C628	*S. fredii* NGR234; Sm^R, Rf^R, Sp^R	(292, 293) and this study

C475	*S. fredii* USDA 257; Sm^R	(294) and this study
C591	B001 + pJG528; Cm^R, Km^R	This study
C657	UBAPF2 harboring pNGR234a; Rf^R, Nm^R	This study
C568	UBAPF2 harboring pUSDA257a; Rf^R, Nm^R	This study

Plasmid	**Description**[a]	**Reference**
pRK600	Self-transmissible helper plasmid; Cm^R	(290)
pRK7813	RK2 derivative carrying pUC9 polylinker and λ *cos* site, highly unstable in *S. meliloti*; Tc^R	(224)
pJG109	Transposon delivery vector; Km^R/Nm^R	This study
pKD46	$SC101^{ts}$ *oriV*; Ap^R	(295)
pCDF	CloDF13 *oriV*; Sm^R	Novagen
pJG194	2.2 kb mobilizable suicide vector; Km^R/Nm^R	(123)
pJG310	Transposon delivery vector; Km^R/Nm^R	This study
pJG505	pJG392 with *rhaK–icpA* intergenic region from Rm1021 cloned into BamHI site; Tc^R	(110)
pJG520	Transposon delivery vector; Gm^R	This study
pJG563	pJG194 with *loxP* cloned into XhoI/HindIII	This study
pJG565	pCDF *oriV*, RP4 *oriT*, *loxP*; Gm^R	This study
pJG577	*cre* cloned into BamHI site of pRK7813; Tc^R	This study
pJG568	A fragment of *SMb20931* of Rm1021 cloned into pJG563	This study
pJG569	A fragment of *thiD* of Rm1021 cloned into pJG565	This study
pJG572	A fragment of *orf00007* of pHRC017 cloned into pJG563	This study
pJG573	A fragment of *orf00202* of pHRC017 cloned into pJG565	This study
pJG574	The *orf00200–orf00201* intergenic region of pHRC017 cloned into pJG563	This study
pJG575	A fragment of *orf00404* of pHRC017 cloned into pJG565	This study
pJG527	*y4xK–y4xL* intergenic region of C475 cloned into the BamHI/XbaI sites of pJG194	This study
pJG528	*y4xK–nopL* intergenic region of C628 cloned into the BamHI/XbaI sites of pJG194	This study

Bacteriophage	**Description**	**Reference**
ΦN3	*S. meliloti* transducing phage	(216)

Ap^R, ampicillin resistance; Cm^R, chloramphenicol resistance; Km^R kanamycin resistance; Nm^R, neomycin resistance; Rf^R, rifampicin resistance; Sm^R, streptomycin resistance; Tc^R, tetracycline resistance.

Table S5.2-8. Primers used in **§2.2**.

Name	Sequence[a]	Directionality	Purpose
oJG664	CAGTTTACTTTGCAGGGCTTCC	Forward	Sequence verification of pJG194 inserts
oJG1243	TGCGAAAAAGGATGGATATACCG	Reverse	Sequence verification of pJG194 inserts
oJG705	GACGCCTAGGTTTGTGCCAATACCAGTAG	Forward	Clone SC101^{ts} *oriV* from pKD46
oJG706	GACGCCTAGGTGAAGATCCTTTTTGATAATC	Reverse	Clone SC101^{ts} *oriV* from pKD46
oJG707	GACGCCTAGGTACCTCAGATCTTGATCCC	Forward	Clone Tn*5*–Km^{R} from pJG109
oJG708	GACGCCTAGGGAGCTCTGTCTCTTATACAC	Reverse	Clone Tn*5*–Km^{R} from pJG109
oJG726	CGCGGATCCTCTGGGCTGCCCTTCCTG	Forward	Clone *oriT*
oJG727	CGCTCTAGACTCCCTGCTTCGGGGTCATT	Reverse	Clone *oriT*
oJG1350	CGCAAGCTTGGAGTACACCATGTCCAATTTACTGAC	Forward	Clone *cre* into pRK7813
oJG1320	CGCGAATTCCGGTTCGCTTGCTGTGG	Reverse	Clone *cre* into pRK7813
oJG1321	TCCGGCTCGTATGTTGTGC	Forward	Sequence verification of pRK7813 inserts
oJG1322	TGCTGCAAGGCGATTAAGTTGG	Reverse	Sequence verification of pRK7813 inserts
oMC099	CGCGGATCCCGGCAGATCCTGGTGTTTGC	Forward	Clone the *y4xK–y4xL* intergenic region of USDA 257
oMC100	CGCTCTAGACCTCACCATTGGCTCGATCG	Reverse	Clone the *y4xK–y4xL* intergenic region of USDA 257
oMC101	CAATCCAGGCATTGCCAATCTGG	Forward	Loop-in verification of pJG527
oMC102	GAGGACCCTGATCAAGCTGG	Reverse	Loop-in verification of pJG527
oMC103	CGCGGATCCGCATGAACGCGACTGTTCTCG	Forward	Clone the *y4xK–nopL* intergenic region of NRG234
oMC104	CGCTCTAGAGGGCTTTCCTGCAGAGTAGG	Reverse	Clone the *y4xK–nopL* intergenic region of NRG234
oMC105	CCAGTTACCACCTACAGGACC	Forward	Loop-in verification of pJG528
oMC106	GTCATTCGCATCAGGCCATCG	Reverse	Loop-in verification of pJG528
oMC119	TCGAG**ATAACTTCGTATAGCATACATTATACGAAGTTAT**A	Top	Construct *loxP*
oMC120	AGCTT**ATAACTTCGTATAATGTATGCTATACGAAGTTAT**C	Bottom	Construct *loxP*
oMC121	CTATGGAAAAACGCCAGCAACG	Forward	Sequence verification of *loxP* insert into pJG194
oMC122	CCTCGTGCTTTACGGTATCG	Reverse	Sequence verification of *loxP* insert into pJG194
oMC123	CGCCCTAGGCTCGAGGTCGACGGTATCG	Forward	Clone Gm^{R}–*oriT* from pJG520
oMC124	CGCGCATGCTGGGCTGCCCTTCCTGG	Reverse	Clone Gm^{R}–*oriT* from pJG520
oMC125	CGCGCATGC**ATAACTTCGTATAGCATACATTATACGAAGT TAT**TCACTCGGTCGCTACGCTC	Forward	Clone CloDF13 *oriV* from pCDF
oMC126	CGCCCTAGGCGGTTCAGTAGAAAAGATCAAAG	Reverse	Clone CloDF13 *oriV* from pCDF
oMC127	GTTACCACGGTTAAGCAGTTCC	Forward	Sequence verification of Gm^{R}–*oriT*–pCDF ligation
oMC128	GTTGCTGCTGCGTAACATCG	Reverse	Sequence verification of Gm^{R}–*oriT*–pCDF ligation
oMC129	TGTCCCTTATTCGCACCTGG	Forward	Sequence verification of Gm^{R}–*oriT*–*loxP*–pCDF ligation
oMC130	CCTGCTGTTTTGCCTCACATG	Reverse	Sequence verification of Gm^{R}–*oriT*–*loxP*–pCDF ligation

(Table S5.2 8 continued)

Name	Sequence[a]	Directionality	Purpose
oMC131	CATAGCCGAATAGCCTCTCC	Reverse	PCR verification off pJG563 and its derivatives
oMC132	CGTTCTGCCCAAGTTTGAGC	Reverse	PCR verification off pJG565 and its derivatives
oMC133	CGCTCTAGACTGGGGACACATCGAGGAG	Forward	Clone of a fragment of *orf00007* into pJG563
oMC134	CGCGGATCCGGACGCTACGAAGCCTTACG	Reverse	Clone of a fragment of *orf00007* into pJG563
oMC135	CGCTCTAGAGTCCTTATCACAAGCACCAACG	Forward	Clone *orf00202* into pJG565
oMC136	CGCGGATCCGCTCATTGAAGTCTCCTGGTTG	Reverse	Clone *orf00202* into pJG565
oMC137	GCCAACGGTAAGCGCTACG	Forward	Detection of loop-in of pJG572 (with oMC131)
oMC138	TCGGCAAAGGCTCTGAAATGC	Forward	Detection of loop-in of pJG573 (with oMC132)
oMC139	CGCTCTAGACGAATCCCTTGATGTGCTCC	Forward	Clone of *orf00200–orf00201* into pJG563
oMC140	CGCGGATCCGTTGGTGCTTGTGATAAGGAC	Reverse	Clone of *orf00200–orf00201* into pJG563
oMC141	CGCTCTAGACTTCGGTGAGCTCATTTGACC	Forward	Clone of a fragment of *orf00404* into pJG565
oMC142	CGCGGATCCGTTTGAAGGGCGTCGAACTG	Reverse	Clone of a fragment of *orf00404* into pJG565
oMC143	GGTTTCGACGAGTTTGCTCG	Forward	Detection of loop-in of pJG574 (with oMC131)
oMC144	CTCCAACCTGGTGGAAAACG	Forward	Detection of loop-in of pJG575 (with oMC132)
oMC209	ATGACGTCGACCGTGACCG	Forward	Check for *orf00016* of pHRC017
oMC210	TCAGTCGGCAGCATTCTTGC	Reverse	Check for *orf00016* of pHRC017
oMC211	TGGAGACGGTCTGTTCATGG	Forward	Check for *orf00104* of pHRC017
oMC212	AAAGCCGTGGCGGTAGAGG	Reverse	Check for *orf00104* of pHRC017
oMC213	CCGATGCAGTCAAGGTGACG	Forward	Check for *orf00163* of pHRC017
oMC214	CGTCCTTCTTCGGCAGTTCG	Reverse	Check for *orf00163* of pHRC017
oMC215	CTATGGTTTGCTAGCCGTCG	Forward	Check for *orf00227* of pHRC017
oMC216	CTTCCGCTGCTTGAAATAGTCC	Reverse	Check for *orf00227* of pHRC017
oMC217	CCTGACGAATCTGCAGGACG	Forward	Check for *orf00256* of pHRC017
oMC218	GATCGTGTCAGCGACCTTGG	Reverse	Check for *orf00256* of pHRC017
oMC231	ATTGTCGGTGATGAGATCGTGC	Forward	Check for *orf00567* of pHRC017
oMC232	TTGGTTCTCATGCCTTCCTGG	Reverse	Check for *orf00567* of pHRC017
oMC241	GTTCTGGCGTACCGAGAACG	Forward	Check for *orf00528* of pHRC017
oMC242	GCGACAGCGTCTTGGATACG	Reverse	Check for *orf00528* of pHRC017
oMC245	CCGTGATGCAACTGACGACC	Forward	Check for *orf00689* of pHRC017
oMC246	GAACACCTTCTGGAAGATGTCC	Reverse	Check for *orf00689* of pHRC017
oMC247	GGAGTTGTTCGAGACGCTGG	Forward	Check for *orf00054* of pHRC017
oMC248	GATGGTCACGTCGGTCCTC	Reverse	Check for *orf00054* of pHRC017
oMC249	GCGTGGATTGGTTTGCAAGC	Forward	Check for *orf00066* of pHRC017
oMC250	CACGGCAAGTCCAAGTGTAC	Reverse	Check for *orf00066* of pHRC017

(Table S5.2 8 continued)

Name	Sequence[a]	Directionality	Purpose
oMC257	CAGACCTTCCACTCAACTCC	Forward	Check for *orf00318* of pHRC017
oMC258	CTATTTCGACCTGGACCTTTGG	Reverse	Check for *orf00318* of pHRC017
oMC259	CGAGTTTCATACTCGGCTTGC	Forward	Check for *orf00389* of pHRC017
oMC260	CGAGAAGATCAGCGTCCACG	Reverse	Check for *orf00389* of pHRC017
oMC275	TGCTCGCAAAGAACGACCTG	Forward	Check for *orf00070* of pHRC017
oMC276	GATCTTGTTCTGCATGCCAACG	Reverse	Check for *orf00070* of pHRC017
oMC279	CAGCGCGGATGTCTCTAACC	Forward	Check for *orf00172* of pHRC017
oMC280	CTTGGTTTTCGACGTAACCACG	Reverse	Check for *orf00172* of pHRC017
oMC281	CTCATGAAACTTCTGCAGCGTG	Forward	Check for *orf00223* of pHRC017
oMC282	GCTCGCTTCTTCGCTTTTCC	Reverse	Check for *orf00223* of pHRC017
oMC283	CGCAGACAAGGCTCTGTGCT	Forward	Check for *orf00232* of pHRC017
oMC284	GAAGCCCACCGAAATCTTTTGC	Reverse	Check for *orf00232* of pHRC017
oMC285	CAGGACCTCAGTAAGTTGTTGC	Forward	Check for *orf00269* of pHRC017
oMC286	CCTTTCACACTCTTCAGGACG	Reverse	Check for *orf00269* of pHRC017
oMC287	TGGTCGCCTATCATCGGAAC	Forward	Check for *orf00281* of pHRC017
oMC288	GTCAGAGTTGAATGGCACACG	Reverse	Check for *orf00281* of pHRC017
oMC289	CAGATGGTCGAGAACATCTCTG	Forward	Check for *orf00636* of pHRC017
oMC290	CTCAGTTTCTCTGGCAGTAAGG	Reverse	Check for *orf00636* of pHRC017
oMC293	CTTCGACTATGTCGTCGTCG	Forward	Check primer inside *exo* deletion (*exoP–thiD*)
oMC294	CAGGCGGAAGCGATCATGC	Reverse	Check primer inside *exo* deletion (*exoP–thiD*)
oMC308	CGCTCTAGAGCAAGATCTGCACCATCGAG	Forward	Clone a fragment of *SMb20931* into pJG563
oMC309	CGCGGATCCGTCTCCTTGGTGCACTCGTC	Reverse	Clone a fragment of *SMb20931* into pJG563
oMC310	CGCTCTAGAGCTGCCATTAACGCAATGACC	Forward	Clone a fragment of *thiD* into pJG565
oMC311	CGCGGATCCACAGACCTGTTCTTCGACG	Reverse	Clone a fragment of *thiD* into pJG565
oMC312	CCAGGACATGCTTGGCAAGG	Forward	Detection of loop-in of pJG568 (with oMC131)
oMC007	GGCAAATACAGCGACTTCGACG	Forward	Detection of loop-in of pJG569 (with oMC132)
oMC322	GCTCGATCTCCAGCTTCATC	Forward	Detect pNGR234b
oMC323	GATGCGCTTCAGGATCAACG	Reverse	Detect pNGR234b
oMC324	GCTCCTAGTTTCTTAGTTCGCC	Forward	Detect pNGR234a/pUSDA257a
oMC325	CAGCCAATTGAGCACGTCAC	Reverse	Detect pNGR234a/pUSDA257a
oMC371	CGTGTCCGCTTGTTTCAAGG	Forward	Check primer across *exo* deletion
oMC372	GGACCTCAAGACGTTTTCAGC	Reverse	Check primer across *exo* deletion

[a]Restriction sites are underlined; *loxP* sites are in bold.

Table S5.2-9. Predicted toxins and antitoxins on pSymB of *S. meliloti* Rm1021. Shading indicates singles/pairs/triplets.

Name	Type	Strand	Location (nt)
'SMb20028.5'	antitoxin	+	40548–40795
SMb20062	antitoxin	+	72518–72766
SMb20063	toxin	+	72756–73052
SMb22004	toxin	–	132946–133338
SMb20121	antitoxin	–	133283–133558
SMb20411	toxin	–	426694–426942
SMb20412	toxin	–	426973–427137
SMb20413	antitoxin	–	427134–427361
SMb22020	antitoxin	–	507618–508478
SMb21035	antitoxin	–	653580–653798
SMb21117	antitoxin	–	758135–758593
SMb21127	toxin	+	766495–766866
SMb21128	antitoxin	+	766873–767427
SMb21666	antitoxin	–	955649–955867
SMb21007	toxin	–	1241319–1241738
SMb21008	antitoxin	–	1241735–1242067
SMb21475	antitoxin	+	1400396–1400599
SMb21476	toxin	+	1400599–1400946
SMb21509	antitoxin	+	1434924–1435552
SMb21510	toxin	+	1435168–1435293
SMb21511	toxin	+	1435307–1435552
SMb20695	toxin	–	1495765–1496271
SMb20696	antitoxin	–	1496268–1496753
SMb20607	toxin	–	1621026–1621490
SMb20608	antitoxin	–	1621523–1621861
SMb20627	toxin	–	1647906–1648214
SMb20628	antitoxin	–	1648211–1648480
SMb20629	antitoxin	–	1648566–1648844
SMb21651	toxin	–	1682468–1682881
SMb22018	antitoxin	–	1682878–1683126

[a]Names in *italics* are annotated genes in *S. meliloti* Rm1021; names in 'quotes' were predicted by (149) (names suggested by me).

Table S5.2-10. Predicted toxins and antitoxins on pSymA of *S. meliloti* Rm1021. Shading indicates singles/pairs/triplets.

Name[a]	Type	Strand	Location (nt)
SMa5001	toxin	+	58298–58612
SMa0191	toxin	–	105799–106329
SMa0193	antitoxin	–	106326–106652
SMa0285	toxin	–	159165–159389
SMa0286	antitoxin	–	159590–160051
Rorf_2431[b]	antitoxin	–	192666–192965
Rorf_2432[b]	toxin	–	192997–193278
Rorf_3101	antitoxin	+	242901–243155
SMa0453	toxin	+	243155–243562
SMa0471	antitoxin	+	258225–258533
SMa0473	toxin	+	258530–258823
SMa0545	toxin	–	291198–291638
SMa0548	antitoxin	–	291638–291877
Rorf_3914	antitoxin	+	304285–304560
SMa0572	toxin	+	304557–304880
SMa0592	toxin	+	315280–316452
SMa0594	antitoxin	+	316428–317459
SMa0917	antitoxin	+	509885–510211
Rorf_6690	antitoxin	+	510391–510750
SMa5006	toxin	+	510778–511137
ntrR2	toxin	–	546465–546884
Rorf_6690	antitoxin	–	546881–547150
SMa5007	antitoxin	–	550344–550724
Rorf_7221	toxin	–	550742–551053
SMa1076	antitoxin	–	587855–588439
SMa5008	toxin	–	750238–750459
Rorf_10090	toxin	–	776787–777155
SMa1413	antitoxin	–	777118–777393
‘SMa1449’	antitoxin	–	798865–799065
SMa1455	toxin	–	802324–802842
SMa1456	antitoxin	–	802839–803228
‘SMa1633’[b]	antitoxin	+	909138–909366
‘SMa1634’[b]	toxin	+	909363–909721
‘SMa1707’	toxin	–	962237–962437
SMa1706	antitoxin	–	962812–963234
SMa1770	toxin	–	1005927–1006100
Rorf_12944[b]	antitoxin	–	1006097–1006644
‘SMa1923’	toxin	–	1094579–1094827
SMa1924	antitoxin	–	1095016–1095690
‘SMa1989’	antitoxin	+	1128526–1128780
SMa1990	toxin	+	1128777–1129133
SMa2105	toxin	+	1187433–1188653
‘SMa2150’[b]	toxin	+	1214674–1215040

SMa2151	antitoxin	+	1215030–1215353
'SMa2230'	antitoxin	+	1248848–1249048
SMa2231	toxin	+	1249045–1249428
Rorf_16392	antitoxin	+	1262737–1263027
SMa2253	toxin	+	1263024–1263401
SMa2255	antitoxin	+	1263404–1263742
SMa2273	toxin	–	1268200–1268550
SMa2275	antitoxin	–	1268531–1268806
SMa2279	toxin	–	1269640–1270122
SMa2281	antitoxin	–	1270119–1270451
'SMa2320'	antitoxin	–	1292672–1292911

[a]Names in *italics* are annotated genes in *S. meliloti* Rm1021; names beginning with Rorf were predicted by (148); names in 'quotes' were predicted by (149) (names suggested by me).
[b]Likely pseudogenes.

Table S5.2-11. Strains and plants used in **§2.3**.

Strain	Description[a]	Reference
C017	*S. meliloti* N6B7; Sm^R	(110)
B469	*S. meliloti* 74B17; Sm^R	(110)
B800	*S. meliloti* M98; Sm^R	(110)
C377	*S. meliloti* M256; Sm^R	(110)
B464	*S. meliloti* N6B2; Sm^R	(110)
C285	*S. meliloti* 128A2; Sm^R	(110)
Rm1021	*S. meliloti* SU47; Sm^R	(288)

Plant	Accession Number/Ecotype	Location Origin	Source/Reference
Medicago praecox	PI 495429	Greece	USDA-ARS[b]
Medicago sativa	GT13R+	–	ABI, Nampa, ID

[a]Sm^R, streptomycin resistance. [b]United States Department of Agriculture-Agricultural Research Service, U.S. annual *Medicago* species core collection (http://bldg6.arsusda.gov/pberkum/Public/sarl/bauchan/core2.html).

Table S5.2-12. Primers used in **§2.3**.

Name	Sequence	Direction	Purpose
oJG1035	ACTCCTACGGGAGGCAGCAGT	Forward	16S rDNA primer
oJG1036	TACGGTTACCTTGTTACGACTT	Reverse	16S rDNA primer
oMC035	GGATTGCAAATGGCTGAGG	Forward	Sequencing of *nodH–nodFE* intergenic region (pSymA)
oMC036	CTTCGCGCATCCATTTCCAG	Reverse	Sequencing of *nodH–nodFE* intergenic region (pSymA)
oMC037	TCTTTGCGACCCTCGGTCTTG	Forward	Sequencing of *SMa2075–SMa2077* intergenic region (pSymA)
oMC038	GATCGCCGACCTGATCAAGG	Reverse	Sequencing of *SMa2075–SMa2077* intergenic region (pSymA)
oMC039	GGGATAAAAAACGGCGAACTGG	Forward	Sequencing of *exoY–exoX* intergenic region (pSymB)
oMC040	ATCAGGGTCTGGATCGCCGT	Reverse	Sequencing of *exoY–exoX* intergenic region (pSymB)
oMC041	CTCACGAGAATGTCGATATCC	Forward	Sequencing of *thuR–thuE* intergenic region (pSymB)
oMC042	CTTCTCGAACCTGGCGATCTG	Reverse	Sequencing of *thuR–thuE* intergenic region (pSymB)
oMC043	CAAGGCATGCACGCCCTATGA	Forward	Sequencing of *rkpA–rkpU* intergenic region (Chromosome)
oMC044	TTGATCGTCTCGCTGACGAG	Reverse	Sequencing of *rkpA–rkpU* intergenic region (Chromosome)
oMC045	CGGATGAAGTCCATGAAGGTG	Forward	Sequencing of *fumC–SMc00150* intergenic region (Chromosome)
oMC046	CCTCGTTCGCATTCATGTTCG	Reverse	Sequencing of *fumC–SMc00150* intergenic region (Chromosome)
oMC091	ATGTCACCACAAACAGAAAC	Forward	Universal *rbcL* primer (184)
oMC092	TCGCATGTACCTGCAGTAGC	Reverse	Universal *rbcL* primer (184)
oMC093	CCTTATCATTTAGAGGAAGGAG	Forward	Universal ITS primer (184)
oMC094	TCCTCCGCTTATTGATATGC	Reverse	Universal ITS primer (184)
oMC095	GTTATGCATGAACGTAATGCTC	Forward	Universal *psbA–trnH* primer (184)
oMC096	CGCGCATGGTGGATTCACAATCC	Reverse	Universal *psbA–trnH* primer (184)
oMC195	GTCGCATGATCGTGTTGGTC	Forward	Sequencing of *fixA–SMa0824* intergenic region (pSymA)
oMC196	GATAGAAGGGGCCTGCAATAAG	Reverse	Sequencing of *fixA–SMa0824* intergenic region (pSymA)
oMC197	GACGACGTGGTGGTTGTCG	Forward	Sequencing of *ndvB–SMc04822* intergenic region (chromosome)
oMC198	CTGCTGCATCTGGCGACG	Reverse	Sequencing of *ndvB–SMc04822* intergenic region (chromosome)
oMC199	GACGGCACTTTATCCACAGG	Forward	Sequencing of *ntrC–ntrY* intergenic region (chromosome)
oMC200	GTACGCAGCGAGAACCAGC	Reverse	Sequencing of *ntrC–ntrY* intergenic region (chromosome)
oMC201	AGAGATCGACCCAGACGGTC	Forward	Sequencing of *fixG–fixP1* intergenic region (pSymA)
oMC202	CGACATGCTGACACCAGAGC	Reverse	Sequencing of *fixG–fixP1* intergenic region (pSymA)
oMC203	CCAGAACCAGTCGATCGACG	Forward	Sequencing of *cycL–degP1* intergenic region (chromosome)
oMC204	TGAAGAAGCCGGAGCCTTGC	Reverse	Sequencing of *cycL–degP1* intergenic region (chromosome)
oMC205	TCGGTGACGATGGGACACTG	Forward	Sequencing of *hemA–SMc03103* intergenic region (chromosome)
oMC206	TCCATCGCCAACAACGATCC	Reverse	Sequencing of *hemA–SMc03103* intergenic region (chromosome)
oMC207	GCATAAAGAAGCGTCACGACG	Forward	Sequencing of *wgdA–wgcA* intergenic region (pSymB)

oMC208	CGGTGTACGAATGCTACATGC	Reverse	Sequencing of *wgdA–wgcA* intergenic region (pSymB)
oMC293	CTTCGACTATGTCGTCGTCG	Forward	AFLP between *S. meliloti* and *S. medicae* at *exoP–thiD*
oMC294	CAGGCGGAAGCGATCATGC	Reverse	AFLP between *S. meliloti* and *S. medicae* at *exoP–thiD*
oMC295	ATACCGTCTGGACGGTTTGTC	Forward	AFLP between *S. meliloti* and *S. medicae* at *SMc01522*–[*ntrP*]–*ntrR1*
oMC296	CAGCCGGTCAGATATTCGTAG	Reverse	AFLP between *S. meliloti* and *S. medicae* at *SMc01522*–[*ntrP*]–*ntrR1*
oMC297	CCATATCGCGACCGTTTTAAGC	Forward	AFLP between *S. meliloti* and *S. medicae* at *pyc–SMc03896*
oMC298	CACCGATCCGGAGATCATGG	Reverse	AFLP between *S. meliloti* and *S. medicae* at *pyc–SMc03896*

Table S5.2-13. Characteristics of sinorhizobia collected in Utah in the summer of 2010.

Strain	Plant	Collection Site	Date Collected	Species[a]	Growth on LB[b]
U001	*Medicago sativa*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U003	*Melilotus officinalis*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U005	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	–
U006	*Melilotus officinalis*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U007	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U008	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U009	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U010	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	–
U011	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+
U012	*Melilotus alba*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U013	*Medicago sativa*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U014	*Medicago sativa*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U015	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+
U016	*Melilotus officinalis*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U017	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U018	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U019	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U020	*Melilotus officinalis*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U022	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	+/–
U023	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	+/–
U024	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	+/–
U025	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	+/–
U026	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U027	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U028	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	+/–
U029	*Medicago lupulina*	Canyon Glen Park	15-Jul-10	*Sinorhizobium medicae*	–

(Table S5.2-13 continued)

Strain	Plant	Collection Site	Date Collected	Species[a]	Growth on LB[b]
U030	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U031	*Melilotus officinalis*	Canyon Glen Park	15-Jul-10	*Sinorhizobium meliloti*	+
U032	*Melilotus officinalis*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U033	*Melilotus alba*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U034	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+/–
U035	*Melilotus alba*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U036	*Melilotus officinalis*	Logan River Trail	24-Jul-10	*Sinorhizobium meliloti*	+
U037	*Medicago lupulina*	Logan River Trail	24-Jul-10	*Sinorhizobium meliloti*	+
U038	*Medicago lupulina*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U039	*Medicago lupulina*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U040	*Medicago lupulina*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U041	*Melilotus officinalis*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U042	*Melilotus officinalis*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U043	*Medicago sativa*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U044	*Medicago lupulina*	Willard, Utah	24-Jul-10	*Sinorhizobium meliloti*	+
U045	*Melilotus officinalis*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U046	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+/–
U047	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U048	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium meliloti*	+
U054	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	–
U055	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	–
U057	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+
U058	*Medicago lupulina*	Provo River Trail	15-Jul-10	*Sinorhizobium medicae*	+/–
U059	*Melilotus officinalis*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U060	*Melilotus officinalis*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U061	*Melilotus officinalis*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U062	*Melilotus alba*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+

(Table S5.2-13 continued)

Strain	Plant	Collection Site	Date Collected	Species[a]	Growth on LB[b]
U063	*Melilotus officinalis*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U067	*Melilotus alba*	Provo River Trail	5-Aug-10	*Sinorhizobium meliloti*	+
U068	*Melilotus officinalis*	Provo River Trail	5-Aug-10	*Sinorhizobium meliloti*	+
U069	*Melilotus officinalis*	Provo River Trail	5-Aug-10	*Sinorhizobium meliloti*	+
U070	*Melilotus alba*	Provo River Trail	5-Aug-10	*Sinorhizobium meliloti*	+
U072	*Medicago lupulina*	City Creek Canyon	31-Jul-10	*Sinorhizobium medicae*	+/–
U074	*Melilotus alba*	City Creek Canyon	31-Jul-10	*Sinorhizobium meliloti*	+
U076	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U077	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U078	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U079	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U081	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U082	*Melilotus alba*	Canyon on 300 South, Provo	11-Aug-10	*Sinorhizobium meliloti*	+
U089	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U092	*Medicago sativa*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U093	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U094	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U096	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U097	*Medicago sativa*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U100	*Melilotus officinalis*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U101	*Medicago lupulina*	Red Rock Lakes, Montana	23-Aug-10	*Sinorhizobium medicae*	+
U102	*Melilotus officinalis*	West Yellowstone, Montana	23-Aug-10	*Sinorhizobium meliloti*	+
U104	*Medicago lupulina*	Northwest Wyoming	24-Aug-10	*Sinorhizobium meliloti*	+
U108	*Melilotus officinalis*	Northwest Wyoming	25-Aug-10	*Sinorhizobium meliloti*	+
U110	*Medicago lupulina*	Northwest Wyoming	25-Aug-10	*Sinorhizobium medicae*	+
U111	*Medicago lupulina*	Provo Center Street, west of I-15	14-Jul-09	*Sinorhizobium meliloti*	+/–
U112	*Melilotus officinalis*	Provo River Trail	14-Jul-09	*Sinorhizobium meliloti*	+

(Table S5.2-13 continued)

Strain	Plant	Collection Site	Date Collected	Species[a]	Growth on LB[b]
U113	*Medicago lupulina*	Provo River Trail	14-Jul-09	*Sinorhizobium medicae*	–
U114	NA[b]	NA[c]	NA[b]	*Sinorhizobium meliloti*	+
U115	NA[b]	NA[c]	NA[b]	*Sinorhizobium meliloti*	+
U116	NA[b]	NA[c]	NA[b]	*Sinorhizobium meliloti*	+
U117	NA[b]	NA[c]	NA[b]	*Sinorhizobium meliloti*	+
U118	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U120	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+
U121	*Melilotus alba*	1390 North, Geneva Road	1-Sep-10	*Sinorhizobium meliloti*	+

[a]As determined by AFLP and/or 16S rDNA sequencing. [b]+ = grew well; +/– = grew poorly; – = did not grow. [c] NA = not available (this data was lost).

Table S5.2-14. Characteristics of sinorhizobia collected in Utah in the summer of 2012.

Strain	ID[a]	Plant	Collection Site	Date Collected
W001	P1a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W002	P1b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W003	P1c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W004	P1d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W005	P2a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W006	P2b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W007	P2c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W008	P2d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W009	P3a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W010	P3b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W011	P3c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W012	P3d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W013	P4a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W014	P4b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W015	P4c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W016	P4d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W017	P5a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W018	P5b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W019	P5c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W020	P5d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W021	P6a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W022	P6b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W023	P6c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W024	P6d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W025	P7a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W026	P7b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W027	P7c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W028	P7d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W029	P8a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W030	P8b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W031	P8c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W032	P8d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W033	P9a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W034	P9b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W035	P9c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W036	P9d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W037	P10a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W038	P10b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W039	P10c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W040	P10d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W041	P11a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W042	P11b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W043	P11c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W044	P11d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W045	P12a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W046	P12b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W047	P12c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W048	P12d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W049	P13a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W050	P13b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W051	P13c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W052	P13d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W053	P14a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W054	P14b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W055	P14c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W056	P14d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W057	P15a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W058	P15b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W059	P15c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W060	P15d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W061	P16a	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W062	P16b	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W063	P16c	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W064	P16d	*Medicago lupulina*	800 N. University Avenue, Provo	12-Jun-12
W065	DN1a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W066	DN1b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W067	DN1c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W068	DN1d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W069	DN2a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W070	DN2b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W071	DN2c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W072	DN2d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W073	DN3a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W074	DN3b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W075	DN3c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W076	DN3d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W077	DN4a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W078	DN4b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W079	DN4c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W080	DN4d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W081	DE1a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W082	DE1b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W083	DE1c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W084	DE1d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W085	DE2a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W086	DE2b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W087	DE2c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W088	DE2d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W089	DE3a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W090	DE3b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W091	DE3c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W092	DE3d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W093	DE4a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W094	DE4b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W095	DE4c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W096	DE4d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W097	DS1a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W098	DS1b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W099	DS1c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W100	DS1d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W101	DS2a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W102	DS2b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W103	DS2c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W104	DS2d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W105	DS3a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W106	DS3b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W107	DS3c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W108	DS3d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W109	DS4a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W110	DS4b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W111	DS4c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W112	DS4d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W113	DW1a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W114	DW1b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W115	DW1c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W116	DW1d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W117	DW2a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W118	DW2b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W119	DW2c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W120	DW2d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W121	DW3a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W122	DW3b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W123	DW3c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W124	DW3d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W125	DW4a	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W126	DW4b	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W127	DW4c	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W128	DW4d	*Medicago sativa*	Diamond Fork Canyon	30-Apr-12
W129	VN1a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W130	VN1b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W131	VN1c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W132	VN1d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W133	VN2a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W134	VN2b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W135	VN2c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W136	VN2d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W137	VN3a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W138	VN3b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W139	VN3c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W140	VN3d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W141	VN4a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W142	VN4b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W143	VN4c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W144	VN4d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W145	VE1a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W146	VE1b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W147	VE1c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W148	VE1d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W149	VE2a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W150	VE2b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W151	VE2c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W152	VE2d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W153	VE3a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W154	VE3b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W155	VE3c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W156	VE3d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W157	VE4a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W158	VE4b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W159	VE4c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W160	VE4d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W161	VS1a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W162	VS1b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W163	VS1c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W164	VS1d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W165	VS2a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W166	VS2b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W167	VS2c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W168	VS2d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W169	VS3a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W170	VS3b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W171	VS3c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W172	VS3d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W173	VS4a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W174	VS4b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W175	VS4c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W176	VS4d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W177	VW1a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W178	VW1b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W179	VW1c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W180	VW1d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W181	VW2a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W182	VW2b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W183	VW2c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W184	VW2d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W185	VW3a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W186	VW3b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W187	VW3c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W188	VW3d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12

(Table S5.2-14 continued)

Strain	ID[a]	Plant	Collection Site	Date Collected
W189	VW4a	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W190	VW4b	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W191	VW4c	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12
W192	VW4d	*Medicago sativa*	Spring Hollow Road, Provo Canyon	1-May-12

[a]The first letter indicates collection site (P = 800 N. University Avenue, Provo; D = Diamond Fork Canyon; V = Spring Hollow Road, Provo Canyon), the second letter indicates which corner of the 1 m^2 plot the sample came from (N = north, E = east, S = south, W = west), the number indicates which plant the isolate came from (1 through 4), and the final letter indicates which nodule the isolate came from (a through d).

Table S5.2-15. Accessory plasmids of Utah sinorhizobia.

Strain	Plant	Collection Site	16S rDNA	# Plasmids	*repA*[a]	*traA*[a]	*acdS*[a]
U001	*Medicago sativa*	Provo River Trail	*Sinorhizobium meliloti*	1	–	+	–
U003	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U005	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U006	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	1	+	–	–
U007	*Medicago lupulina*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U008	*Medicago lupulina*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U009	*Medicago lupulina*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U010	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	1	–	–	+
U011	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	1	–	–	+
U012	*Melilotus alba*	Provo River Trail	*Sinorhizobium meliloti*	2	NT	–	–
U013	*Medicago sativa*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U014	*Medicago sativa*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U015	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	1	–	–	–
U016	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	3	NT	+	+
U017	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	0	NT	NT	NT
U018	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	2	NT	–	+
U019	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	1	–	–	–
U020	*Melilotus officinalis*	City Creek Canyon	*Sinorhizobium meliloti*	1	NT	NT	+
U022	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	0	NT	NT	NT
U023	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	0	NT	NT	NT
U024	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	1	–	–	+
U025	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	0	NT	NT	NT
U026	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	1	+	–	–
U027	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	0	NT	NT	NT
U028	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	0	NT	NT	NT
U029	*Medicago lupulina*	Canyon Glen Park	*Sinorhizobium medicae*	0	NT	NT	NT

(Table S5.2-15 continued)

Strain	Plant	Collection Site	16S rDNA	# Plasmids	*repA*[a]	*traA*[a]	*acdS*[a]
U030	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	1	–	–	+
U031	*Melilotus officinalis*	Canyon Glen Park	*Sinorhizobium meliloti*	1	+	+	+
U032	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U033	*Melilotus alba*	Provo River Trail	*Sinorhizobium meliloti*	1	–	+	+
U034	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U035	*Melilotus alba*	Provo River Trail	*Sinorhizobium meliloti*	1	–	–	+
U036	*Melilotus officinalis*	Logan River Trail	*Sinorhizobium meliloti*	2	NT	+	+
U037	*Medicago lupulina*	Logan River Trail	*Sinorhizobium meliloti*	1	–	+	+
U038	*Medicago lupulina*	Willard, Utah	*Sinorhizobium meliloti*	1	–	+	+
U039	*Medicago lupulina*	Willard, Utah	*Sinorhizobium meliloti*	1	+	–	–
U040	*Medicago lupulina*	Willard, Utah	*Sinorhizobium meliloti*	1	–	+	+
U041	*Melilotus officinalis*	Willard, Utah	*Sinorhizobium meliloti*	1	NT	+	+
U042	*Melilotus officinalis*	Willard, Utah	*Sinorhizobium meliloti*	0	NT	NT	NT
U043	*Medicago sativa*	Willard, Utah	*Sinorhizobium meliloti*	0	NT	NT	NT
U044	*Medicago lupulina*	Willard, Utah	*Sinorhizobium meliloti*	1	NT	–	–
U045	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	1	NT	+	+
U046	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U047	*Medicago lupulina*	Provo River Trail	*Sinorhizobium meliloti*	1	NT	+	–
U048	*Medicago lupulina*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U054	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	2	NT	+	–
U055	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U057	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U058	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	0	NT	NT	NT
U059	*Melilotus officinalis*	City Creek Canyon	*Sinorhizobium meliloti*	3	NT	+	+
U060	*Melilotus officinalis*	City Creek Canyon	*Sinorhizobium meliloti*	3	NT	+	+
U061	*Melilotus officinalis*	City Creek Canyon	*Sinorhizobium meliloti*	1	NT	–	+
U062	*Melilotus alba*	City Creek Canyon	*Sinorhizobium meliloti*	2	NT	+	+

(Table S5.2-15 continued)

Strain	Plant	Collection Site	16S rDNA	# Plasmids	*repA*[a]	*traA*[a]	*acdS*[a]
U063	*Melilotus officinalis*	City Creek Canyon	*Sinorhizobium meliloti*	0	NT	NT	NT
U067	*Melilotus alba*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U068	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U069	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	1	NT	+	+
U070	*Melilotus alba*	Provo River Trail	*Sinorhizobium meliloti*	0	NT	NT	NT
U072	*Medicago lupulina*	City Creek Canyon	*Sinorhizobium medicae*	0	NT	NT	NT
U074	*Melilotus alba*	City Creek Canyon	*Sinorhizobium meliloti*	0	NT	NT	NT
U076	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	0	NT	NT	NT
U077	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	0	NT	NT	NT
U078	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	0	NT	NT	NT
U079	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	0	NT	NT	NT
U081	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	0	NT	NT	NT
U082	*Melilotus alba*	Canyon on 300 South, Provo	*Sinorhizobium meliloti*	1	NT	+	+
U089	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	2	NT	–	+
U092	*Medicago sativa*	1390 North, Geneva Road	*Sinorhizobium meliloti*	1	NT	–	–
U093	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	2	NT	–	+
U094	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	0	NT	NT	NT
U096	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	1	NT	–	+
U097	*Medicago sativa*	1390 North, Geneva Road	*Sinorhizobium meliloti*	0	NT	NT	NT
U100	*Melilotus officinalis*	1390 North, Geneva Road	*Sinorhizobium meliloti*	1	NT	+	+
U101	*Medicago lupulina*	Red Rock Lakes, Montana	*Sinorhizobium medicae*	1	NT	–	+
U102	*Melilotus officinalis*	West Yellowstone, Montana	*Sinorhizobium meliloti*	1	NT	–	+
U104	*Medicago lupulina*	Northwest Wyoming	*Sinorhizobium meliloti*	3	NT	+	+
U108	*Melilotus officinalis*	Northwest Wyoming	*Sinorhizobium meliloti*	2	NT	–	+
U110	*Medicago lupulina*	Northwest Wyoming	*Sinorhizobium medicae*	3	NT	–	+
U111	*Medicago lupulina*	Provo Center Street, west of I-15	*Sinorhizobium meliloti*	1	NT	–	–
U112	*Melilotus officinalis*	Provo River Trail	*Sinorhizobium meliloti*	1	NT	+	+

(Table S5.2-15 continued)

Strain	Plant	Collection Site	16S rDNA	# Plasmids	*repA*[a]	*traA*[a]	*acdS*[a]
U113	*Medicago lupulina*	Provo River Trail	*Sinorhizobium medicae*	1	NT	–	–
U114	NA[b]	NA[b]	*Sinorhizobium meliloti*	3	NT	–	–
U115	NA[b]	NA[b]	*Sinorhizobium meliloti*	0	NT	NT	NT
U116	NA[b]	NA[b]	*Sinorhizobium meliloti*	1	NT	–	–
U117	NA[b]	NA[b]	*Sinorhizobium meliloti*	0	NT	NT	NT
U118	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	0	NT	NT	NT
U120	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	0	NT	NT	NT
U121	*Melilotus alba*	1390 North, Geneva Road	*Sinorhizobium meliloti*	0	NT	NT	NT

[a]PCR-based check for the presence of the *repA*, *traA*, or *acdS* gene. + = gene detected, – = gene not detected, NT = not tested. [b] NA = not available (this data was lost).

Table S5.2-16. Host ranges of tested Utah sinorhizobia.

Strain	Plant[a]	Collection Site	16S rDNA	# Plasmids	*repA*[b]	*traA*[b]	*acdS*[b]	*M. sativa*[c]	*M. praecox*[c]
U001	SA	Provo River Trail	*S. meliloti*	1	–	+	–	Fix+	Fix+
U003	OF	Provo River Trail	*S. meliloti*	0	NT	NT	NT	Fix+	Fix+
U006	OF	Provo River Trail	*S. meliloti*	1	+	–	–	Fix+	Fix+
U010	LU	Provo River Trail	*S. medicae*	1	–	–	+	Fix+	Fix+
U011	LU	Provo River Trail	*S. medicae*	1	–	–	+	Fix+	Fix+
U012	AL	Provo River Trail	*S. meliloti*	2	NT	–	–	Fix+	Fix+
U015	LU	Provo River Trail	*S. medicae*	1	–	–	–	Fix+	Fix+
U016	OF	Provo River Trail	*S. meliloti*	3	NT	+	+	Fix+	Fix–
U018	OF	Canyon Glen Park	*S. meliloti*	2	NT	–	+	Fix+	Fix+
U019	OF	Canyon Glen Park	*S. meliloti*	1	–	–	–	Fix+	Fix+
U020	OF	City Creek Canyon	*S. meliloti*	1	NT	NT	+	Fix+	Fix+
U022	LU	Canyon Glen Park	*S. medicae*	0	NT	NT	NT	Fix+	Fix+
U024	LU	Canyon Glen Park	*S. medicae*	1	–	–	+	Fix+	Fix+
U026	OF	Canyon Glen Park	*S. meliloti*	1	+	–	–	Fix+	Fix+
U030	OF	Canyon Glen Park	*S. meliloti*	1	–	–	+	Fix+	Fix+
U031	OF	Canyon Glen Park	*S. meliloti*	1	+	+	+	Fix+	Fix+
U033	AL	Provo River Trail	*S. meliloti*	1	–	+	+	Fix+	Fix+
U035	AL	Provo River Trail	*S. meliloti*	1	–	–	+	Fix+	Fix+
U036	OF	Logan River Trail	*S. meliloti*	2	NT	+	+	Fix+	Fix–
U037	LU	Logan River Trail	*S. meliloti*	1	–	+	+	Fix+	Fix+
U038	LU	Willard, Utah	*S. meliloti*	1	–	+	+	Fix+	Fix+
U039	LU	Willard, Utah	*S. meliloti*	1	+	–	–	Fix+	Fix+
U040	LU	Willard, Utah	*S. meliloti*	1	–	+	+	Fix+	Fix+
U041	OF	Willard, Utah	*S. meliloti*	1	NT	+	+	Fix+	Fix+
U044	LU	Willard, Utah	*S. meliloti*	1	NT	–	–	Fix+	Fix+
U045	OF	Provo River Trail	*S. meliloti*	1	NT	+	+	Fix+	Fix+

U047	LU	Provo River Trail	*S. meliloti*	1	NT	+	–	Fix+	Fix+
U054	LU	Provo River Trail	*S. medicae*	2	NT	+	–	Fix–	Fix–
U057	LU	Provo River Trail	*S. medicae*	0	NT	NT	NT	Fix+	Fix+
U059	OF	City Creek Canyon	*S. meliloti*	3	NT	+	+	Fix+	Fix+
U060	OF	City Creek Canyon	*S. meliloti*	3	NT	+	+	Fix+	Fix+
U061	OF	City Creek Canyon	*S. meliloti*	1	NT	–	+	Fix+	Fix+
U062	AL	City Creek Canyon	*S. meliloti*	2	NT	+	+	Fix+	Fix+
U069	OF	Provo River Trail	*S. meliloti*	1	NT	+	+	Fix+	Fix+
U079	AL	Canyon on 300 South, Provo	*S. meliloti*	0	NT	NT	NT	Fix+	Fix+
U082	AL	Canyon on 300 South, Provo	*S. meliloti*	1	NT	+	+	Fix+	Fix+
U089	AL	1390 North, Geneva Road	*S. meliloti*	2	NT	–	+	Fix+	Fix+
U092	SA	1390 North, Geneva Road	*S. meliloti*	1	NT	–	–	Fix+	Fix+
U093	AL	1390 North, Geneva Road	*S. meliloti*	2	NT	–	+	Fix+	Fix+
U096	AL	1390 North, Geneva Road	*S. meliloti*	1	NT	–	+	Fix+	Fix+
U100	OF	1390 North, Geneva Road	*S. meliloti*	1	NT	+	+	Fix+	Fix+
U101	LU	Red Rock Lakes, Montana	*S. medicae*	1	NT	–	+	Fix+	Fix+
U102	OF	West Yellowstone, Montana	*S. meliloti*	1	NT	–	+	Fix+	Fix+
U104	LU	Northwest Wyoming	*S. meliloti*	3	NT	+	+	Fix+	Fix+
U108	OF	Northwest Wyoming	*S. meliloti*	2	NT	–	+	Fix+	Fix–
U110	LU	Northwest Wyoming	*S. medicae*	3	NT	–	+	Fix+	Fix+
U111	LU	Center Street, west of I-15	*S. meliloti*	1	NT	–	–	Fix+	Fix+
U112	OF	Provo River Trail	*S. meliloti*	1	NT	+	+	Fix+	Fix+
U113	LU	Provo River Trail	*S. medicae*	1	NT	–	–	Fix+	Fix+
U114	NA[d]	NA[d]	*S. meliloti*	3	NT	–	–	Fix+	Fix+
U116	NA[d]	NA[d]	*S. meliloti*	1	NT	–	–	Fix+	Fix+

[a]AL = *Melilotus alba*; LU = *Medicago lupulina*; OF = *Melilotus officinalis*; SA = *Medicago sativa*. [b]PCR-based check for the presence of the *repA*, *traA*, or *acdS* gene. + = gene detected, – = gene not detected, NT = not tested. [c]Symbiotic phenotypes were scored approximately 30 days post-inoculation (dpi). Fix+, effective N-fixing pairs; Fix–, abortively nodulating pairs. [d]NA = not available (this data was lost).

Table S5.2-17. Pairwise comparisons of different *S. meliloti* strains for the different loci used in the phylogenetic analysis. Boxes are color-coded according to the number of differences between the two strains: purple, 0–2 differences; blue, 3–5 differences; cyan, 6–8 differences; green, 9–11 differences; yellow, 12–14 differences; orange, 15–20 differences; pink, 21+ differences; X = at least one (*) of the pair being compared doesn't amplify at this locus. B100 = *S. meliloti* Rm1021.

Chromosomal Loci

***rkpA–rkpU* (*rkp*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	8	0					
C017	8	0	0				
B800	9	3	3	0			
C377	8	0	0	3	0		
B464	0	8	8	9	8	0	
C285	9	3	3	0	3	9	0

***fumC–SMc00150* (*cgm*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	0	0	0				
B800	0	0	0	0			
C377	1	1	1	1	0		
B464	0	0	0	0	1	0	
C285	4	4	4	4	5	4	0

***ndvB–SMc04822* (*ndv*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	0	0	0				
B800	0	0	0	0			
C377	0	0	0	0	0		
B464	0	0	0	0	0	0	
C285	7	7	7	7	7	7	0

***ntrC–ntrY* (*ntr*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	0	0	0				
B800	7	7	7	0			
C377	9	9	9	6	0		
B464	0	0	0	7	9	0	
C285	8	8	8	1	7	8	0

***cycL–degP1* (*cyc*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	0	0	0				
B800	0	0	0	0			
C377	0	0	0	0	0		
B464	0	0	0	0	0	0	
C285	5	5	5	5	5	5	0

***hemA–SMc03103*(*hem*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	0	0	0				
B800	0	0	0	0			
C377	0	0	0	0	0		
B464	0	0	0	0	0	0	
C285	4	4	4	4	4	4	0

pSymB Loci

***exoY–exoX* (*exo*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	7	0					
C017	6	3	0				
B800	6	3	2	0			
C377	7	0	2	3	0		
B464	5	2	1	1	2	0	
C285	18	21	20	20	20	19	0

***thuR–thuE* (*thu*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	8	0					
C017	0	8	0				
B800	0	8	0	0			
C377	13	9	13	13	0		
B464	43	42	43	43	45	0	
C285	5	5	5	5	10	42	0

***wgdA–wgcA* (*wg-*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	9	0					
C017	9	4	0				
B800	11	20	20	0			
C377	8	1	3	19	0		
B464	8	1	3	19	0	0	
C285	7	6	6	16	5	5	0

pSymA Loci

***fixG–fixP1* (*fix*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	0	0					
C017	1	1	0				
B800	3	3	2	0			
C377	2	2	1	1	0		
B464	4	4	5	5	4	0	
C285	62	62	61	61	60	64	0

***fixA–SMa0824* (*nif*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	11	0					
C017	0	11	0				
B800	48	48	48	0			
C377	44	37	44	18	0		
B464	0	11	0	50	44	0	
C285	12	2	12	47	37	12	0

***nodH–nodFE* (*nod*)**

	B100	B469	C017	B800	C377	B464	C285
B100	0						
B469	3	0					
C017	0	3	0				
B800	10	9	10	0			
C377	6	5	6	4	0		
B464	0	3	0	10	6	0	
C285	4	1	4	10	6	4	0

***SMa2075–SMa2077* (anti-*nod*)**

	B100	B469	C017	B800*	C377*	B464	C285*
B100	0						
B469	11	0					
C017	0	11	0				
B800*	X	X	X	X			
C377*	X	X	X	X	X		
B464	3	14	3	X	X	0	
C285*	X	X	X	X	X	X	X

Table S5.2-18. Other rhizobia collected as part of the SymLC project.

Strain	Host Plant	Collection Site	Date Collected	16S rDNA	Growth[a]		# Plasmids	PCR[b]	
					LB	TY		*traA*	*acdS*
U075	*Colutea arborescens*[c]	Willard Peak Access Road, Utah	31-Jul-10	*Rhizobium pisi*	–	+	6	–	+
U083	*Trifolium repens*	Butterfly Lake in the Uintahs, Utah	14-Aug-10	*Rhizobium leguminosarum*	–	+	1	–	–
U084	*Trifolium repens*	Butterfly Lake in the Uintahs, Utah	14-Aug-10	*Rhizobium leguminosarum*	–	+	1	–	–
U085	*Trifolium repens*	Butterfly Lake in the Uintahs, Utah	14-Aug-10	*Rhizobium leguminosarum*	–	+	1	–	–
U086	*Trifolium repens*	Butterfly Lake in the Uintahs, Utah	14-Aug-10	*Rhizobium leguminosarum*	–	+	1	–	–
U087	*Trifolium repens*	Butterfly Lake in the Uintahs, Utah	14-Aug-10	*Rhizobium leguminosarum*	–	+	1	–	–
U095	*Astragalus alpinus*	Northwestern Wyoming	24-Aug-10	*Mesorhizobium loti*	–	+	5	–	–
U098	*Astragalus alpinus*	Northwestern Wyoming	24-Aug-10	*Mesorhizobium loti*	–	+	6	–	–
U103	*Trifolium pratense*	Northwestern Wyoming	24-Aug-10	*Rhizobium leguminosarum*	–	+	2	–	+
U105	*Trifolium hybridum*	Northwestern Wyoming	24-Aug-10	*Rhizobium leguminosarum*	–	+	2	–	+
U106	*Trifolium pratense*	Northwestern Wyoming	24-Aug-10	*Rhizobium leguminosarum*	–	+	2	–	+
U107	*Trifolium pratense*	Northwestern Wyoming	24-Aug-10	*Rhizobium leguminosarum*	–	+	3	–	+
U109	*Trifolium hybridum*	Northwestern Wyoming	25-Aug-10	*Rhizobium leguminosarum*	–	+	5	–	–
U122	*Vicia americana*	Spring Hollow Road	1-May-12	*Rhizobium pisi*	–	+/–[g]	3	NT	NT
U123	*Vicia americana*	Spring Hollow Road	1-May-12	*Bosea* sp.[e]	–	+/–[g]	0	NT	NT
U124	*Securigera varia*	Cold Spring Harbor, New York	29-Aug-12	*Mesorhizobium* sp.[f]	–	+/–[g]	1	NT	NT
U125	*Lotus corniculatus*	Cold Spring Harbor, New York	29-Aug-12	*Mesorhizobium amorphae*	–	+/–[g]	1	NT	NT
U126	*Sesbania* sp.[d]	Round Rock, Texas	6-Sep-12	*Rhizobium huautlense*	–	+	3	NT	NT

[a]+ = grew well; +/– = grew poorly; – = did not grow.

[b]PCR-based check for the presence of the *repA*, *traA*, or *acdS* gene. + = gene detected, – = gene not detected, NT = not tested.

[c]DNA was not recovered from this host so identification was based on memory of morphological characteristics and BONAP and is therefore uncertain.

[d]Sequence analysis in consultation with BONAP could not distinguish between *Sesbania drummondii* or *Sesbania vesicaria*.

[e]The type strain with the closest S_ab score (296) was *Bosea massiliensis* 63287^T, with a score of 0.955.

[f]The 16S rDNA sequence of strain U124 had a single nucleotide difference from the type strains of four *Mesorhizobium* species: *M. tianshanense*, *M. tariemense*, *M. gobiense*, and *M. metallidurans*.

[g]These strains grew best on YEM agar plates.

Table S5.2-19. Homologs of *ropA* in sequenced genomes of the Order Rhizobiales.

Species Gene	GenBank Accession
Agrobacterium radiobacter K84	
Arad_1337	YP_002543723.1
Arad_1575	YP_002543908.1
Arad_1578	YP_002543909.1
Arad_1777	YP_002544056.1
Arad_2103	YP_002544299.1
Arad_3620	YP_002545457.1
Arad_4722	YP_002546302.1
Arad_8343[a]	YP_002541532.1
Arad_8731[a]	YP_002541532.1
Agrobacterium sp. H13-3	
AGROH133_05007	YP_004278265.1
AGROH133_05009	YP_004278266.1
AGROH133_09570[a]	YP_004442978.1
Agrobacterium tumefaciens C58	
Atu1020	NP_354043.1
Atu1021	NP_354045.1
Atu4693[a]	NP_355972.1
Agrobacterium vitis S4	
Avi_1427	YP_002549025.1
Avi_1428	YP_002549026.1
Azorhizobium caulinodans ORS 571	
AZC_1213	YP_001524129.1
AZC_3535	YP_001526451.1
Bartonella bacilliformis KC583	
BARBAKC583_0447	YP_988765.1
Bartonella clarridgeiae 73	
BARCL_0408	YP_004158675.1
Bartonella grahamii as4aup	
Bgr_16380	YP_002972477.1
Bartonella henselae Houston-1	
BH12500	YP_034002.1
Bartonella quintana Toulouse	
BQ09890	YP_032578.1
Bartonella tribocorum CIP 105476	
BT_1902	YP_001610159.1
Beijerinckia indica subsp. *indica* ATCC 9039	
Bind_0628	YP_001831767.1
Bind_1873	YP_001832986.1
Bind_2544	YP_001833635.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Bradyrhizobium japonicum USDA 6		
	BJ6T_46390	YP_005609493.1
	BJ6T_47020	YP_005609555.1
	BJ6T_79340	YP_005612768.1
Bradyrhizobium japonicum USDA 110		
	bll4983	NP_771623.1
	bll5076	NP_771716.1
	bll6888	NP_773528.1
Bradyrhizobium sp. BTAi1		
	BBta_3063	YP_001239088.1
	BBta_4696	YP_001240630.1
	BBta_4698	YP_001240631.1
	BBta_4701	YP_001240634.1
Bradyrhizobium sp. ORS278		
	BRADO4478	YP_001206439.1
	BRADO4479	YP_001206440.1
Bradyrhizobium sp. S23321		
	S23_13850	YP_005448714.1
	S23_31370	YP_005450454.1
	S23_32010	YP_005450517.1
Brucella abortus A13334		
	BAA13334_I02932	YP_005152206.1
	BAA13334_I02934	YP_005152207.1
Brucella abortus bv. 1 str. 9-941		
	BruAb1_0655	YP_221391.1
	BruAb1_0657	YP_221393.1
Brucella abortus S19		
	BAbS19_I06180	YP_001934623.1
	BAbS19_I06190	YP_001934624.1
Brucella canis ATCC 23365		
	BCAN_A0651	YP_001592497.1
	BCAN_A0653	YP_001592498.1
Brucella canis HSK A52141		
	BCA52141_I0694	YP_005115048.1
	BCA52141_I0695	YP_005115049.1
Brucella melitensis ATCC 23457		
	BMEA_A0675	YP_002732394.1
	BMEA_A0677	YP_002732395.1
Brucella melitensis bv. 1 str. 16M		
	BMEI1305	NP_540222.1
	BMEI1306	NP_540223.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Brucella melitensis bv. *abortus* 2308		
	BAB1_0659	YP_414101.1
	BAB1_0660	YP_414102.1
Brucella melitensis M5-90		
	BM590_A0655	YP_005600103.1
	BM590_A0656	YP_005600104.1
Brucella melitensis M28		
	BM28_A0650	YP_005596735.1
	BM28_A0651	YP_005596736.1
Brucella melitensis NI		
	BMNI_I0640[b]	YP_005603448.1
	BMNI_I0641[b]	YP_005603449.1
	BMNI_I0642[b]	YP_005603450.1
	BMNI_I0643[b]	YP_005603451.1
Brucella microti CCM 4915		
	BMI_I636	YP_003106584.1
	BMI_I639	YP_003106586.1
Brucella ovis ATCC 25840		
	BOV_0632[b]	5201197 (gene ID)[c]
	BOV_0634	YP_001258626.1|
Brucella pinnipedialis B2-94		
	BPI_I673	YP_004755738.1
	BPI_I675	YP_004755739.1
Brucella suis 1330		
	BS1330_I0633	YP_005615475.1
	BS1330_I0635	YP_005615477.1
Brucella suis ATCC 23445		
	BSUIS_A0665[b]	YP_001627314.1
	BSUIS_A0667	YP_001627315.1
Brucella suis VBI22		
	BSVBI22_A0633	YP_005154323.1
	BSVBI22_A0635	YP_005154325.1
Candidatus Liberibacter asiaticus str. psy62		
	CLIBASIA_00995	YP_003064727.1
Candidatus Liberibacter solanacearum CLso-ZC1		
	CKC_02565	YP_004062750.1
Chelativorans sp. BNC1		
	Meso_0939	YP_673501.1
	Meso_2915	YP_675453.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Mesorhizobium ciceri sv. *biserrulae* WSM1271		
	Mesci_2044	YP_004141247.1
	Mesci_4818	YP_004143977.1
	Mesci_4837	YP_004143996.1
	Mesci_4839	YP_004143998.1
Mesorhizobium loti MAFF303099		
	mll4029	NP_104998.1
	mll6389	NP_106908.1
	mll7738	NP_107992.1
	mlr7740	NP_107994.1
	mlr7768	NP_108014.1
Mesorhizobium opportunistum WSM2075		
	Mesop_5317	YP_004613827.1
	Mesop_5339	YP_004613849.1
	Mesop_5341	YP_004613851.1
Methylobacterium chloromethanicum CM4		
	Mchl_3429	YP_002422177.1
	Mchl_3510	YP_002422258.1
	Mchl_4328	YP_002423032.1
Methylobacterium extorquens AM1		
	MexAM1_META1p3321	YP_002964337.1
	MexAM1_META1p3398	YP_002964412.1
	MexAM1_META1p4345	YP_002965255.1
Methylobacterium extorquens DM4		
	METDI3973	YP_003069453.1
	METDI4952	YP_003070380.1
Methylobacterium extorquens PA1		
	Mext_3110	YP_001640569.1
	Mext_3186	YP_001640644.1
	Mext_3960	YP_001641402.1
Methylobacterium nodulans ORS 2060		
	Mnod_1311	YP_002496614.1
	Mnod_1479	YP_002496774.1
Methylobacterium populi BJ001		
	Mpop_0264	YP_001922984.1
	Mpop_3381	YP_001926067.1
	Mpop_4437	YP_001927071.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Methylobacterium radiotolerans JCM 2831		
	Mrad2831_0221	YP_001752930.1
	Mrad2831_4747	YP_001757391.1
	Mrad2831_4954	YP_001757596.1
	Mrad2831_5088	YP_001757728.1
	Mrad2831_5089	YP_001757729.1
Methylobacterium sp. 4-46		
	M446_0703	YP_001767697.1
	M446_0925	YP_001767908.1
	M446_3475	YP_001770294.1
Methylocella silvestris BL2		
	Msil_0734	YP_002361067.1
Nitrobacter hamburgensis X14		
	Nham_2265	YP_577517.1
	Nham_2266	YP_577518.1
	Nham_2957	YP_578182.1
Nitrobacter winogradskyi Nb-255		
	Nwi_0664	YP_317282.1
	Nwi_1933	YP_318545.1
Ochrobactrum anthropi ATCC 49188		
	Oant_0447	YP_001369007.1
	Oant_2221	YP_001370766.1
	Oant_2645	YP_001371187.1
	Oant_2646	YP_001371188.1
	Oant_3597[a]	YP_001372132.1
Oligotropha carboxidovorans OM4		
	OCA4_c21880	YP_005951196.1
Oligotropha carboxidovorans OM5		
	OCAR_5827	YP_002288817.1
Pelagibacterium halotolerans B2		
	KKY_1090	YP_004898873.1
Polymorphum gilvum SL003B-26A1		
	SL003B_1886	YP_004303614.1
	SL003B_3435	YP_004305161.1
Pseudovibrio sp. FO-BEG1		
	PSE_1509	YP_005080043.1
	PSE_3622	YP_005082148.1
	PSE_3623	YP_005082149.1
	PSE_3624	YP_005082150.1
	PSE_3626	YP_005082152.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Rhizobium etli CFN 42		
	RHE_CH01349	YP_468879.1
	RHE_CH02437	YP_469941.1
	RHE_CH03578	YP_471060.1
	RHE_PE00260[d]	YP_472422.1
Rhizobium etli CIAT 652		
	RHECIAT_CH0001446	YP_001977603.1
	RHECIAT_CH0002537	YP_001978667.1
	RHECIAT_CH0003841	YP_001979956.1
	RHECIAT_PA0000205[d]	YP_001985812.1
Rhizobium leguminosarum bv. *trifolii* WSM1325		
	Rleg_1139	YP_002974973.1
	Rleg_2312	YP_002976126.1
	Rleg_6754	YP_002984755.1
Rhizobium leguminosarum bv. *trifolii* WSM2304		
	Rleg2_0986	YP_002280506.1
	Rleg2_2075	YP_002281585.1
	Rleg2_3326	YP_002282819.1
	Rleg2_5842[d]	YP_002278115.1
Rhizobium leguminosarum bv. *viciae* 3841		
	pRL110375[d]	YP_771409.1
	RL1499	YP_767103.1
	RL2775	YP_768359.1
Rhodopseudomonas palustris BisA53		
	RPE_2319	YP_781240.1
	RPE_2632	YP_781549.1
Rhodopseudomonas palustris BisB5		
	RPD_0323	YP_567462.1
	RPD_2414	YP_569545.1
	RPD_3281	YP_570406.1
Rhodopseudomonas palustris BisB18		
	RPC_2891	YP_532758.1
	RPC_3136	YP_532997.1
Rhodopseudomonas palustris CGA009		
	RPA0538	NP_945891.1
	RPA2419	NP_947761.1
	RPA3423	NP_948762.1
Rhodopseudomonas palustris DX-1		
	Rpdx1_1954	YP_004108298.1
	Rpdx1_3090	YP_004109403.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Rhodopseudomonas palustris HaA2		
	RPB_0502	YP_484124.1
	RPB_2143	YP_485760.1
	RPB_3036	YP_486649.1
Rhodopseudomonas palustris TIE-1		
	Rpal_0539	YP_001989574.1
	Rpal_2696	YP_001991680.1
	Rpal_3909	YP_001992882.1
Sinorhizobium fredii HH103		
	SFHH103_00747	YP_005188071.1
	SFHH103_00750	YP_005188074.1
Sinorhizobium fredii NGR234		
	NGR_a03720[d]	NP_443854.1
	NGR_c08140	YP_002825359.1
	NGR_c08170[b]	YP_002825361.1
Sinorhizobium fredii USDA 257		
	USDA257_c07880	YP_006396137.1
Sinorhizobium medicae WSM419		
	Smed_0668	YP_001326359.1
	Smed_0670	YP_001326361.1
Sinorhizobium meliloti AK83		
	Sinme_0848	YP_004548218.1
	Sinme_0850	YP_004548219.1
Sinorhizobium meliloti BL225C		
	SinmeB_0676	YP_005712855.1
	SinmeB_0678	YP_005712857.1
Sinorhizobium meliloti GR4		
	C770_GR4Chr1062	AGA06023.1
	C770_GR4Chr1064	AGA06025.1
Sinorhizobium meliloti Rm41		
	BN406_00789	YP_006839660.1
	BN406_00791	YP_006839662.1
Sinorhizobium meliloti Rm1021		
	SMc02396	NP_385158.1
	SMc02400	NP_385162.1
Sinorhizobium meliloti SM11		
	SM11_chr0717	YP_005719256.1
	SM11_chr0719	YP_005719258.1
Starkeya novella DSM 506		
	Snov_1099	YP_003693038.1
	Snov_4077	YP_003695966.1

(Table S5.2-19 continued)

Species	Gene	GenBank Accession
Stenotrophomonas maltophilia D457		
	SMD_2579	YP_006185290.1
Stenotrophomonas maltophilia JV3		
	BurJV3_2449	YP_004792996.1
Stenotrophomonas maltophilia K279a		
	Smlt2944	YP_001972693.1
Stenotrophomonas maltophilia R551-3		
	Smal_2395	YP_002028781.1
Xanthobacter autotrophicus Py2		
	Xaut_2286	YP_001417186.1
	Xaut_3823	YP_001418704.1
	Xaut_3824	YP_001418705.1

[a]These homologs of *ropA* are located on a second chromosome. [b]These homologs of *ropA* are pseudogenes. [c]Translated from the DNA sequence. [d]These homologs of *ropA* are encoded on an accessory plasmid.

Table S5.2-20. Strains, plasmids, and bacteriophages used in **§2.4**.

Strain	Relevant Characteristics	Source or Reference
DH5α	*E. coli* cloning strain	(297)
B001	DH5α harboring helper plasmid pRK600	(286)
Rm1021	*S. meliloti* SU47 Sm^r (progenitor to strains listed below)	(288)
B199	*lpsB*::Tn*5*-110 (formerly R5D6; an LPS-defective symbiotic mutant); Sm^r, Nm^r	(286)
B912	Rm1021 *ropA1*^{G129D}	This study
B920	Rm1021 *ropA1*^{G84D}	This study
B955	Rm1021 *ropA1*^{G84A}	This study
B956	Rm1021 *ropA1*^{G84V}	This study
B957	Rm1021 *ropA1*^{G84R}	This study
B958	Rm1021 *ropA1*$^{\Delta A122–N124}$	This study
B959	Rm1021 *ropA1*$^{\Delta G203–V204}$	This study
B961	Rm1021 *ropA1*^{S87Y}	This study
B962	Rm1021 *ropA1*^{S87F}	This study
B970	Rm1021 *ropA1*$^{205::GV}$	This study
B971	Rm1021 *ropA1*^{D134Y}	This study
B972	Rm1021 *ropA1*$^{\Delta N124–D125}$	This study
B973	Rm1021 *ropA1*$^{126::ND}$	This study
B974	Rm1021 *ropA1*^{A199V}	This study
C540	Rm1021 *ropA1*^{S89P}	This study
C551	Rm1021 *ropA1*$^{\Delta V204–T205}$	This study
C566	Rm1021 *ropA1*$^{\Delta N121–D123}$	This study
C617	*ropA2*::pJG584	This study
C069	B955 + pJG396	This study
C070	B955 + pRF771	This study
C077	B959 + pJG396	This study
C078	B959 + pRF771	This study
C089	B972 + pJG396	This study
C090	B972 + pRF771	This study
C097	Rm1021 + pJG396	This study

C098	Rm1021 + pRF771	This study
Plasmid		
pRF771	Empty vector for P_{trp} transcriptional fusions; Tc^r	(298)
pRK600	Self-transmissible helper plasmid; Cm^r	(290)
pRK7813	RK2 derivative carrying pUC9 polylinker and λ *cos* site, highly unstable in *S. meliloti*; Tc^r	(224)
pJG110	Transposon delivery vector; Km/Nm^r, Ap^r	(286)
pJG194	2.2-kb mobilizable suicide vector; Km/Nm^r	(123)
pJG396	wildtype *ropA1* cloned into pRK771; Tc^r	This study
pJG581	A 367-bp internal fragment of *ropA1* cloned into pJG194	This study
pJG582	A 334-bp internal fragment of *hisC4* cloned into pJG194	This study
pJG583	A 405-bp fragment upstream of *ropA1* cloned into pJG194	This study
pJG584	A 314-bp internal fragment of *ropA2* cloned into pJG194	This study
pJG624	A 320-bp internal fragment of *hisC4* cloned into pJG194	This study
pJG627	A 330-bp fragment upstream of *ropA1* cloned into pJG194	This study
pJG628	A 319-bp internal fragment of *ropA1* cloned into pJG194	This study
pJG629	A 291-bp internal fragment of *ropA1* cloned into pJG194	This study
pJG630	A 333-bp internal fragment of *SMc02397* cloned into pJG194	This study
pJG631	A 330-bp internal fragment of *ropA2* cloned into pJG194	This study
Bacteriophage		
ΦM1	*S. meliloti* lytic phage	(215)
ΦM5	*S. meliloti* lytic phage	(215)
ΦM6	*S. meliloti* lytic phage	(215)
ΦM7	*S. meliloti* lytic phage isolated from an alfalfa field	(215)
ΦM9	*S. meliloti* lytic phage isolated from a commercial inoculant	(215)
ΦM10	*S. meliloti* lytic phage isolated from a commercial inoculant	(215)
ΦM12	*S. meliloti* lytic phage isolated from a commercial inoculant	(215)
ΦM14	*S. meliloti* lytic phage isolated from a commercial inoculant	(215)
ΦM19	*S. meliloti* lytic phage	(215)
ΦN3	*S. meliloti* lytic phage isolated from an alfalfa field	(216)

Ap^r, ampicillin resistance; Cm^r, chloramphenicol resistance; Km^r, kanamycin resistance; Nm^r, neomycin resistance; Sm^r, streptomycin resistance; Tc^r, tetracycline resistance.

Table S5.2-21. Primers used in **§2.4**.

Name	Sequence	Direction	Purpose
oJG664	CAGTTTACTTTGCAGGGCTTCC	Forward	Sequence verification of pJG194 inserts
oJG1243	TGCGAAAAAGGATGGATATACCG	Reverse	Sequence verification of pJG194 inserts
oJG524	GGTGGCGCACTTCCTGATAGC	Forward	Sequence verification of pRF771 inserts
oJG525	CGTTATCAGAACCGCCCAGACC	Reverse	Sequence verification of pRF771 inserts
oMC023	CGCTCTAGACCCAGACCCGTTTGAAACTTTTG	Forward	Clone *ropA1* into pRF771
oMC024	CGCGGATCCGTAGCCATACTCCAGAAAAGAG	Reverse	Clone *ropA1* into pRF771
oMC029	CGAAAGCCTACGATCACAGG	Forward	Sequencing of *ropA1* mutants
oMC030	CGAAGAAGAGGTGCTGTTCC	Reverse	Sequencing of *ropA1* mutants
oMC303	CGCGGATCCTGAAGCCTACATCCAGCTCG	Forward	Clone a 367-bp fragment of *ropA1* into pJG194
oMC304	CGCTCTAGAGTAAGCGTTCGGGTTGGACG	Reverse	Clone a 367-bp fragment of *ropA1* into pJG194
oMC305	CTGGAACCAGGAAGACTTCG	Forward	Detection of integration of pJG581
oMC314	CGCGGATCCGAAGATCTCGAAGGACTGCTC	Forward	Clone a 334-bp fragment of *hisC4* into pJG194
oMC315	CGCTCTAGAGATTGCGGATCTTGTCGAAGG	Reverse	Clone a 334-bp fragment of *hisC4* into pJG194
oMC316	CGCGGATCCCATGGCTTCCGCAAGGACC	Forward	Clone a 405-bp fragment upstream of *ropA1* into pJG194
oMC317	CGCTCTAGACTTGATGTTCATTTCTGACCTCC	Reverse	Clone a 405-bp fragment upstream of *ropA1* into pJG194
oMC318	CGCGGATCCGTTCAATTCCGATACGGATTCG	Forward	Clone a 314-bp fragment of *ropA2* into pJG194
oMC319	CGCTCTAGACGAGCAGGTCGAAAGTCACG	Reverse	Clone a 314-bp fragment of *ropA2* into pJG194
oMC320	CGCAAGCTTGAAGGTCCGAAGCCAGTCG	Forward	Detection of integration of pJG583
oMC326	CCAATATCGCCATCGGAGAG	Forward	Detection of integration of pJG582
oMC345	CGCGGATCCAAGATTGCGGCACGCATCG	Forward	Clone a 320-bp fragment of *hisC4* into pJG194
oMC346	CGCTCTAGACATAGGGTACCGTGACCAGC	Reverse	Clone a 320-bp fragment of *hisC4* into pJG194
oMC347	AACGTCACAACGCCAAGTGC	Forward	Detection of integration of pJG624
oMC354	CGCGGATCCAACGATGGGCATATGTACC	Forward	Clone a 330-bp fragment upstream of *ropA1* into pJG194
oMC355	CGCTCTAGAGGATAAAACCGGGCAAGAGC	Reverse	Clone a 330-bp fragment upstream of *ropA1* into pJG194
oMC356	TGACGCGGATCGAATGCAGC	Forward	Detection of integration of pJG627
oMC357	CGCGGATCCGAGCCCATGGAATACGTTCG	Forward	Clone a 319-bp fragment of *ropA1* into pJG194
oMC358	CGCTCTAGACTTCATCGACGTCGATCAGG	Reverse	Clone a 319-bp fragment of *ropA1* into pJG194
oMC359	GAAGCAAGGGCGGTTGATCG	Forward	Detection of integration of pJG628

oMC360	CGCGGATCCAACCCGAACGCTTACTGG	Forward	Clone a 291-bp fragment of *ropA1* into pJG194
oMC361	CGCTCTAGATCAGGTCAGATTAGAAGTCACG	Reverse	Clone a 291-bp fragment of *ropA1* into pJG194
oMC362	GCTCGCCTACATCTACGACG	Forward	Detection of integration of pJG629
oMC363	CGCGGATCCGACCATCAACAGGAAGATGG	Forward	Clone a fragment of *SMc02397* into pJG194
oMC364	CGCTCTAGACTTTTGCTCTCACCGTAAGCG	Reverse	Clone a fragment of *SMc02397* into pJG194
oMC365	GTCAAGGAGACCACGCTTGC	Forward	Detection of integration of pJG630
oMC366	CGCGGATCCGCAGCTACGACACGGAATGG	Forward	Clone a second fragment of *SMc02400* into pJG194
oMC367	CGCTCTAGACTGTGTAGTTGATCGCGAAGC	Reverse	Clone a second fragment of *SMc02400* into pJG194
oMC368	GCTTCTTCTACAGCTGGTGG	Forward	Detection of integration of pJG631
oMC369	TTTGCGATGCTTTCGGCATGG	Forward	Detection of integration of pJG581
oMC370	CAAGATCGGCGGCTTCATCC	Forward	Detection of integration of pJG584

Table S5.2-22. The phages in our collection exhibit variable host ranges. Plaquing ability of phages was scored as follows: SS, very susceptible (>75% transparent); S, susceptible (50%–75% transparent); R, resistant (25%–75% transparent); RR, very resistant (<25% transparent).

Isolate[a]	ΦM1	ΦM5	ΦM6	ΦM7	ΦM9	ΦM10	ΦM12	ΦM14	ΦM19	ΦN3
Ve8	R	S	R	SS	SS	RR	S	S	SS	SS
M249	RR	S	RR	RR	SS	SS	RR	SS	RR	RR
102F28	RR	R	RR	R	R	RR	S	R	S	RR
M207	R	R	R	S	RR	R	RR	R	S	R
19A4	R	SS	R	S	SS	R	S	S	S	S
128A7	S	R	S	SS	RR	R[b]	R	R[b]	SS	S
M241	S	S	S	SS	RR	R	SS	R	SS	R
19A9	S	SS	S	SS	SS	R	SS	S	SS	S
Sa-10	R	R	R	S	SS	RR	S	S	SS	SS
102F85	RR	R	RR	R	SS	R	RR	SS	R	RR
M243	RR	R	RR	RR	S	R	RR	SS	R	RR
102F82	R	S	R	SS	SS	RR	S	S	SS	S
74B17	S	S	S	S	RR	S	R	S	S	R
19A5	S	SS	S	SS	SS	R	SS	S	SS	S
N6B7	R	R	R	SS	R[b]	R	R	SS	SS	S
102F34	RR	RR	RR	RR	R	RR	RR	R	R	RR
N6B11	R	S	R	SS	SS	RR	SS	S	SS	SS
M24	S	R	S	R	SS	S	RR	SS	R	R
15A1	S	S	S	S	RR	RR	RR	R	S	S
M119	R	R	R	R	SS	S	RR	S	R	S
M210	RR	R	RR	RR	R	R	RR	R	RR	RR
CC 2003	R	R	R	SS	S	S	RR	SS	SS	S
74B3	S	R	S	R	RR	RR	R	RR	S	R
M95	S	S	S	SS	RR	S	S	S	SS	SS
M270	SS	SS	SS	R[b]	SS	S	SS	SS	SS	R
M56	R	S	R	SS	SS	R	SS	S	SS	S
N6B13	SS	SS	SS	R	RR	RR	RR	RR	R	RR
N6B1	R	S	R	SS	SS	RR	S	S	SS	SS
M262	S	R	R	R	SS	S	RR	SS	S	S
19A18	R	SS	SS	SS	RR	RR	SS	RR	SS	SS
Rm41	RR	R	RR	R	SS	RR	RR	SS	S	RR
N4A7	SS	SS	SS	SS	S	RR	SS	S	SS	S
19A8	R	SS	R	S	SS	R	SS	S	SS	S
N6B2	R	SS	R	SS	RR	RR	R	R	SS	R
M29	RR	R	RR	RR	SS	S	RR	SS	RR	RR
M193	RR	R	RR	RR	S	S	RR	SS	RR	RR
M268	RR	R	RR	RR	SS	SS	RR	SS	RR	RR
M272	RR	RR	RR	RR	R	R	RR	R	RR	RR
M256	RR	S	RR	RR	SS	RR	RR	S	RR	RR
17B6	R	R	R	R	R	R	RR	S	R	R
M294	RR	R	RR	RR	R	R	RR	R	RR	RR

(Table S5.2-22 continued)

Isolate[a]	ΦM1	ΦM5	ΦM6	ΦM7	ΦM9	ΦM10	ΦM12	ΦM14	ΦM19	ΦN3
M273	RR	R	RR	RR	RR	R	RR	R	RR	RR
M245	RR	R	RR	RR	RR	S	RR	S	RR	RR
M156	RR	R	RR	RR	S	S	RR	S	RR	RR
M247	RR	SS	R[b]	S	S	RR	SS	S	SS	RR
M246	R	SS	SS	SS	RR	SS	RR	RR	SS	SS
M48	R	SS	R	S	RR	RR	RR	R	S	S
M181	S	S	S	S	R	R	S	S	SS	S
M124	R	S	R	SS	S	S	SS	S	SS	SS
M76	R	R	R	S	SS	R[b]	S	S	S	S
M257	S	S	S	S	SS	RR	S	SS	S	S
M98	S	S	S	SS	RR	R	S	R	SS	SS
N4A12	RR	S	RR	RR	S	RR	RR	R	RR	RR
74B15	RR	SS	R[b]	RR	S	S	RR	S	RR	RR
74B4	RR	SS	S	RR	S	R	RR	R	RR	RR
74B12	RR	R	RR	RR	SS	SS	RR	SS	RR	RR
M30	RR	R	RR	R	R	RR	RR	R	RR	RR
M259	RR	S	RR	RR	S	R[b]	RR	S	RR	RR
M286	RR	R	RR	RR	S	R	R	R[b]	RR	RR
M263	RR	SS	S	S	R	RR	S	RR	S	RR
56A7	S	SS	S	SS	S	R	S	S	SS	SS
56A6	RR	S	RR	RR	RR	R	RR	R	R	S
56A16	R	R	R	RR	RR	RR	RR	R	RR	RR
56A14	S	SS	S	SS	S	R	S	S	SS	S
128A2	S	R	R	R	SS	S	RR	SS	R	R
15B4	S	R	S	R	SS	S	RR	SS	R	S
15A6	S	R	S	R	S	S	RR	S	R	R
128A3	S	S	R	R	S	R	R	S	S	R
M10	S	S	S	SS	S	R[b]	RR	S	SS	SS
S33	RR	R	RR	RR	RR	RR	RR	RR	R	RR
102F51	S	SS	S	SS	S	R[b]	RR	SS	SS	S
CC 2013	RR	R	RR	RR	RR	RR	RR	RR	RR	RR

[a]Strains from (122). [b]Cases where phage mutants were observed to overcome bacterial resistance.

Table S5.2-23. PROVEAN predictions for the effect of mutations on *ropA1* function.

Variant	Score	Prediction
G84A	–1.135	Neutral
G84D	–1.054	Neutral
G84R	–1.770	Neutral
G84V	–1.334	Neutral
S87F	–1.049	Neutral
S87Y	–0.947	Neutral
S89P	–2.738	Deleterious
ΔN121–D123	–10.086	Deleterious
ΔA122–N124	–10.086	Deleterious
ΔN124–D125	–5.000	Deleterious
+126ND127	–2.404	Neutral
G129D	–1.063	Neutral
D134Y	–0.443	Neutral
A199V	–1.580	Neutral
ΔG203–V204	–7.540	Deleterious
+205GV206	–3.161	Deleterious
ΔV204–T205	–6.376	Deleterious

[a]Cutoff = –2.5.