IDENTIFICATION AND CHARACTERIZATION OF *SINORHIZOBIUM MELILOTI* GENES INVOLVED IN ESTABLISHING SYMBIOSIS WITH ITS HOST PLANT ALFALFA

by

Xue-Song Zhang

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy

The City University of New York

2006

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Biology in Satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Chair of Examining Committee

Dr. Haiping Cheng, Lehman College

Executive Officer
Dr. Gabriel L. Chappell

Dr. Gabriel O. Aisemberg, Lehman College

Dr. Edward J. Kennelly, Lehman College

Dr. David Calhoun, City College

Dr. David Bechhofer, Mount Sinai School Of Medicine

Supervising Committee

The City University of New York

ABSTRACT

CHARACTERIZATION OF *SINORHIZOBIUM MELILOTI* GENES INVOLVED IN ESTABLISHING SYMBIOSIS WITH ITS HOST PLANT ALFALFA

By

Xue-Song Zhang

Advisor: Dr. Hai-Ping Cheng

Soil bacterium *Sinorhizobium meliloti* establishes nitrogen-fixing symbiosis with its leguminous host plant, alfalfa, following a series of continuous signal exchanges. *S. meliloti* exopolysaccharide, succinoglycan, may be one such signal molecule that is essential for the symbiosis. Succinoglycan production is carried out by the products of *exoY* and other *exo* genes, and regulated by the ExoR protein, which has a putative glycosaminoglycan binding site. Site-directed mutagenesis analysis of this binding site carried out in this study suggests it may be important for the regulatory function of the ExoR protein. This is consistent with the finding that change of intracellular pool of glycosaminoglycan by mutagenizing the *glmS* gene or by supplying exogenous glucosamine could modulate the levels of succinoglycan production. All together, these findings raise the possibility that the regulatory function of the ExoR protein is modulated through its glycosaminoglycan binding site.

The complexity of alfalfa root structural changes during symbiosis and the amount of *S. meliloti* genes with unknown functions raised the possibility that more *S. meliloti* genes may function in the early stages of the symbiosis, which can not be easily identified using other methods. A positive functional screen of the entire *S. meliloti* genome for early symbiotic genes has been carried out using improved *In Vivo* Expression Technology (IVET). A group of 113 mostly previously unknown symbiotic genes or putative genes has been identified and 23 of them are inducible by alfalfa root exudates. Most of these genes have not been previously characterized and none of them belongs to the nodulation (*nod)* gene family. The identification of this group of alfalfa root exudates inducible *S. meliloti* genes suggests that the interactions in the early stages of the *S. meliloti* and alfalfa symbiosis could be more complex and that further characterization of these genes will lead to better understanding of the start of this symbiosis.

Taken together, this study provides new insights into *S. meliloti*-alfalfa interactions during the early stages of the symbiosis and lays the foundation for future studies of the symbiosis.

ACKNOWLEDGEMENTS

First my thanks must go to my supervisor Dr. Hai-Ping Cheng for his guidance and support throughout my studies. Without his excellent training and mentoring, it would not be possible for me to complete this project. The knowledge I gained from him, in both academic and non-academic matters, has been invaluable and will definitely be beneficial to my future career.

I would also like to express my gratitude to all members of my committee, Dr. Gabriel Aisemberg, Dr. Edward Kennelly, Dr. David Calhoun and Dr. David Bechhofer, for their valuable advice and support.

I would also like to thank members in Dr. Cheng's lab, particularly my wife, Haiyang Lu, for her many helpful suggestions and critical discussions.

I would also like to thank my parents and my wife for their unreserved love and encouragement throughout my studies, to which have been a source of inspiration and moral support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

A Short Review of Key Symbiotic Events and *Sinorhizobium meliloti* Symbiotic Genes

SYMBIOSIS BETWEEN *SINORHIZOBIUM MELILOTI* AND ALFALFA

The Gram-negative soil bacterium *Sinorhizobium meliloti* establishes symbiosis with its host plant, alfalfa (*Medicago sativa*), by eliciting and colonizing pink nitrogen-fixing root nodules (Figure 1-1). The establishment of the symbiosis requires a series of signal exchanges between the bacterium and its host (31, 32, 136, 138). The symbiosis starts when flavonoids (2-phenyl-1,4-benzopyrone derivatives), secreted by alfalfa plant roots, attract *S. meliloti* cells in the rhizosphere to the surface of host root hairs chemotactically and induce the expression of *S. meliloti* nodulation (*nod*) genes (107). The *nod* gene products mediate the synthesis of Nod factor, a lipochito-oligosaccharide signal molecule. Nod factor triggers the formation of curled root hairs, and stimulates root cell divisions, resulting in the formation of nodule primordium (137). Some of the root hairs form 360° coils, trapping bacterial cells in the middle (66, 137) (Figure 1-1).

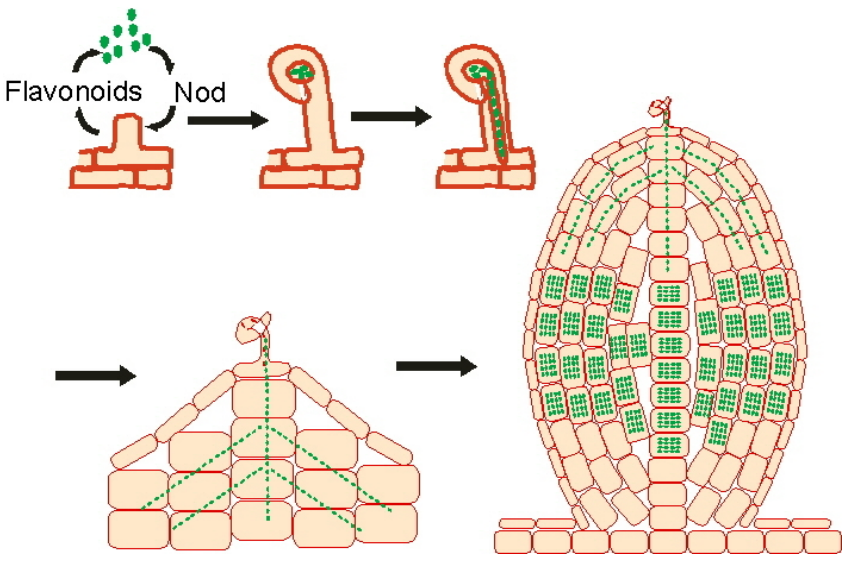

Figure 1-1. Schematic diagram of nodulation process, showing early stages of the symbiosis and nitrogen fixing nodules.

The trapped *S. meliloti* cells elicit the formation of tube-like structures, called infection threads, inside curled root hairs. The presence of *S. meliloti* exopolysaccharide, succinoglycan, is required for infection thread formation (46, 66, 84, 118, 126, 217). As the bacterial cells trapped in infection threads continue to multiply, infection threads elongate inside root hair cells, branch into multiple infection threads, penetrate layers of root cortical cells, enter the root cortex, and reach newly formed plant cells in the nodule primordium (82, 213, 216) (Figure 1-1).

S. meliloti cells inside the infection threads are released into the cytoplasm of the newly formed alfalfa root cells in an endocytic-like process (31). Surrounded by the host cell membrane, bacterial cells enlarge and differentiate into bacteroids inside root nodule cells, where they begin to convert atmospheric nitrogen into ammonia which serves as a nitrogen source for the plant (30, 31, 157, 218).

S. MELILOTI GENES INVOLVED IN ESTABLISHING SYMBIOSIS

The molecular mechanisms behind the symbiosis have been studied intensively only in the recent 20 years, although *S. meliloti*-alfalfa symbiosis has been studied continuously since Beijerinck first found that bacteria caused nodule formation in 1900's, (136). Many *S. meliloti* genes involved in different stages of the symbiosis establishment have been identified and well characterized. Among them, the early *nod* and *exo/exs* genes and the late *nif/fix* genes have been studied most intensively.

The *S. meliloti nod* genes are required for production of Nod factor, which is essential for nodulation

The *S. meliloti* nodulation (*nod*) genes are involved in the production of Nod factor, which is essential for nodulation. In response to alfalfa signal molecules, flavonoids, Nod factor is produced and released into rhizosphere by *S. meliloti* cells. Nod factor induces the curling of root hairs and division of root cortical cells, which leads to the formation of root nodules (168, 181, 185, 219). The *S. meliloti nod* genes are organized into several operons located on megaplasmid pSymA (6). The *nodABC* genes are required for the biosynthesis of the basic lipo-chitin oligosaccharide structure, the core of Nod factors, which is conserved among all rhizobia species. NodA, NodB and NodC function in oligomerization of N-acetyl glucosamine units, N-deacetylation of the oligosaccharide, and transfer of an acyl residue to the free amino group of the oligosaccharide, respectively (203). The host-specific NodF and NodE are essential for the biosynthesis of polyunsaturated fatty acid, which is transferred to oligosaccharide backbone (87). In addition to NodF and NodE, NodG is also involved in the biosynthesis of the polyunsaturated fatty acid (139). NodH, NodP and NodQ function together to modify the basic Nod factor with a sulfate residue, while NodL adds an acetyl to Nod factor (196). NodI and NodJ are involved in the secretion of Nod Factor (42).

S. meliloti has three different *nodD* genes, which encode transcription regulators that mediate sensing of different types of plant symbiotic signals in regulating *nod* gene expression (61, 111, 208, 230). NodD1 and NodD2 recognize different flavonoids and activate *nod* gene expression (105, 168), while NodD3 activates *nod* gene expression

independently of any plant-produced compounds (166). The expression of the *nodD3* gene is regulated by the *S. meliloti* SyrM protein (143, 208). The production of functional NodD1, NodD2 and NodD3 also requires the function of GroE and GroS chaperonin system (155, 230). The *nod* genes were reported to be expressed only at the early stages of nodulation, rather than in the later bacteroid stage (195, 198). The regulation mechanism of *S. meliloti nod* gene expression is not fully understood.

The *S. meliloti exo/exs* genes are required for the production of succinoglycan, which is essential for bacterial invasion to nodules

A set of *S. meliloti exo* genes on megaplasmid pSymB and chromosome are involved in the production of succinoglycan, the main *S. meliloti* exopolysaccharide (Figure 1-2). Succinoglycan is required for the successful invasion of alfalfa roots nodule by the bacterium (12, 93, 152, 183). It is a polymer of an octasaccharide subunit composed of a backbone of one galactose and three glucose residues, a side chain of four glucose residues with succinyl, acetyl, and pyruvyl modifications (127, 180, 183) (Figure 1-3). The octasaccharide repeating unit is assembled on a lipid carrier. The *exoY* gene encodes a glycosyl transferase, which adds the first sugar, galactose, to the lipid carrier as the first step of succinoglycan production. The loss of the *exoY* gene blocks succinoglycan production and bacterial invasion of nodule (183). ExoF is also required for the addition of the first galactose to the lipid carrier (183). ExoA, ExoL, ExoM, ExoO, ExoU and ExoW function in subsequent extension of the backbone and sidechain (12, 13, 183). ExoH, ExoZ, and ExoV are needed for the addition of succinyl, acetyl and pyruvyl

modification groups, respectively (12, 13, 124). ExoP, ExoQ, and ExoT are required for either polymerization or export of succinoglycan (13, 94, 97, 152, 183).

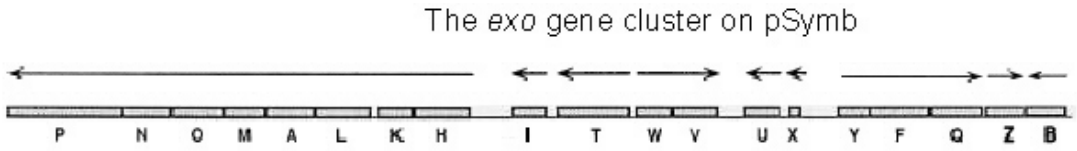

Figure 1-2. A diagram of *exo* gene cluster on megaplasmid pSymb (135, 183).

The *S. meliloti* ExoC, ExoN, and ExoB are essential for biosynthesis of sugar precursors, such as UDP-glucose and UDP-galactose, which supply several different pathways of polysaccharide biosynthesis including succinoglycan production pathway (33, 40, 215). And ExoK is involved in the breakdown of succinoglycan (94). The ExoK and PrsD/PrsE/ExsH proteins are components of independent degradative pathways which contribute to production of symbiotically active low-molecular-weight succinoglycan (231).

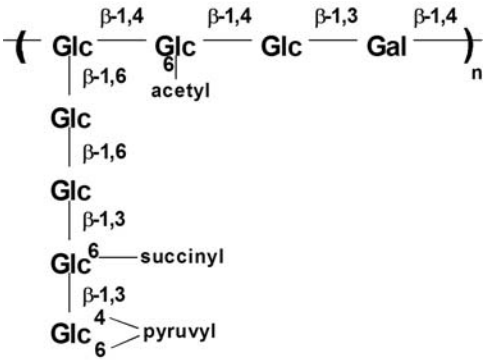

Figure 1-3. Chemical composition of succinoglycan of *S. meliloti* Rm1021 (183). Abbreviations: Glc, glucose; and Gal, galactose.

The regulation of succinoglycan production is very complicated, which involves *S. meliloti* ExoD, ExoX, ExsB, ExoR and ExoS proteins. Among them, ExoD is a membrane-associated protein needed for the bacterium to survive the alkaline conditions encountered during nodule invasion (124, 177). The loss of ExoD affects the distribution of high- and low-molecular-weight forms of succinoglycan depending on the growth medium, although the mechanism is still unclear (124, 177). ExoX inhibits the enzymic activities of ExoY through posttranscriptional modification (152, 175). The loss of ExoX leads to increased succinoglycan production (232). Moreover, ExsB may function as a negative regulator also affecting the *exo* genes expression at posttranscriptional levels (14). ExoR mediates the ammonia regulation of succinoglycan production (64). The loss of ExoR in the *exoR95*::Tn*5* mutant results in overexpression of the *exo* genes and overproduction of succinoglycan regardless the presence or absence of ammonia (176). ExoS is predicted to be a sensor of the ExoS/ChvI two-component regulatory system, regulating the expression of the *exo* genes (47). The *exoS96*::Tn*5* mutation results in the overproduction of succinoglycan. However, the *exoS96* mutant can still respond to ammonia in the same way as wild-type strain (64).

It is increasingly evident that succinoglycan plays important roles in the establishment of the symbiosis, especially at the stages of nodule invasion. When alfalfa plants are inoculated with succinoglycan-deficient mutants, nodulation process is blocked at the stages of initiation and elongation of infection threads (Figure 1-4) (46). EPS II (galactoglucan), (92, 96, 164) and K antigen, a capsular polysaccharide, (38, 71, 115, 164) can restore the defect in nodule invasion of succinoglycan-deficient mutant at a low

efficiency (164). EPS II consists of repeating galactose-glucose disaccharide subunits with acetyl and pyruvyl sidechains (106, 129). EPS II is only produced by some *S. meliloti* strains. The wild-type strain Rm1021 does not produce EPS II at detectable levels because of an insertion into the *expR* gene, which is required for EPS II production (15, 165). K antigen, a capsular polysaccharide, contains a high proportion of 2-deoxy-D-manno-2-octulosonic acid (Kdo) and is structurally analogous to one subgroup of K antigens of *Escherichia coli* (79, 184). K antigen is not produced by *S. meliloti* wild-type strain Rm1021 but is produced by other *S. meliloti* strains (82, 164).

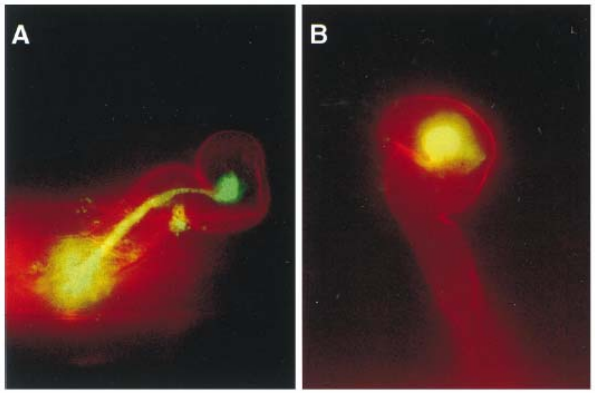

Figure 1-4. Fluorescence microscopy analyses of nodule invasion by GFP-expressing cells of *S. meliloti* (46). Picture A showed a colonized curled root hair with an extended infection thread filled with the wild-type Rm1021 bacterial cells. Picture B showed a colonized curled root hair with a fluorescent bacterial colony of succinoglycan-deficiency mutant *exoY210*, which suggests infection thread was blocked at the initiation stage (46).

S. meliloti nif/fix genes are required for bacteroids to perform nitrogen fixation inside the nodule cells

The successful nodule development and bacterial infection result in the establishment of root nodule cells filled with bacteroids, which express *nif* and *fix* genes to perform nitrogen fixation (75).

The *S. meliloti nifHDK* genes encode a nitrogenase, which is responsible for the reduction of molecular nitrogen (191). The nitrogenase is an enzymic complex of dinitrogenase and dinitrogenase reductase. The dinitrogenase is a $\alpha_2\beta_2$ heterotetramer. The α subunit is encoded by *nifD* gene, while the β subunit is encoded by *nifK* gene. The dinitrogenase reductase is a homodimer, which is encoded by *nifH* gene (88, 178). The nodules induced by *S. meliloti* with any mutation in one of the *nifHDK* genes cannot form N_2-fixing zone, and early nodule senescence is observed (108, 109, 218).

NifE, NifN and NifB are required for nitrogenase Fe-Mo cofactor biosynthesis (29, 58). NifS, NifW and NifX are also required for full nitrogenase activity, although the specific roles of them are unknown (29, 197). NifA, encoded by *nifA*, is a positive regulatory protein, which is required for transcription of all the other *nif* genes (19, 154). In *S. meliloti*, all known *nif* genes are located on pSymA, except for the *nifS* gene, which is on chromosome (86).

In addition to *nif* genes, *fix* genes are also required for the successful formation of functional root nodules (75). The *fix* gene products create an energy system that generates high levels of energy to support nitrogen fixation of the strict aerobic *S. meliloti* cells in the low-oxygen environment inside the nodules (43). Two operons, *fixNOQP* and *fixGHIS* are involved in the synthesis of a respiratory oxidase with high affinity for oxygen (75, 144, 145). The *fixK, fixL* and *fixJ* genes are regulatory genes (55, 202). FixL and FixJ genes encode the FixL/FixJ two-component regulatory system (55). FixL is an

oxygen-sensitive sensor that phosphorylates the FixJ transcription regulator in response to the dramatic drop in oxygen concentration within the nodule. Phosphorylated FixJ activates the expression of intermediary regulatory genes, *nifA* and *fixK* genes, which then activate the expression of the *nifHDKE* and *fixNOQP* genes, respectively (55).

Chapter II

Positive Screening of *Sinorhizobium meliloti* Genes Functioning During the Early Stages of the Symbiosis Establishment

INTRODUCTION

Sinorhizobium meliloti establishes nitrogen-fixing symbiosis inside alfalfa root nodules through a series of signal exchanges and structural changes. This symbiosis has been extensively studied over the years and large numbers of *S. meliloti* genes involved in this process have been identified and characterized (136). Given the complexity of *S. meliloti* genome and large numbers of unknown genes, there could be more bacterial symbiotic genes that remain undiscovered. Loss-of-function mutagenesis has been one of the major approaches used to identify the *S. meliloti* symbiotic genes (108, 124, 147, 183, 233). This approach is most effective for identifying genes that function as a single and essential factor for the symbiosis. However, loss-of-function mutagenesis has become a less effective approach in identifying new symbiotic genes as demonstrated in our recent discovery of only two symbiotic genes after screening ninety putative *S. meliloti* genes (140).

Positive screening approaches based on gene expression during symbiosis could be more effective in identifying symbiotic genes that have multiple homologs or those that are expressed only transiently during a particular stage of the symbiosis. Microarray and proteomics approaches have been used in recent years to identify symbiotic genes based on gene expression in bacteroids inside nodule cells (9, 11, 63). Microarray analysis, however, could be less effective in identifying genes expressed only in the invasion zone of nodules, which accounts for a very small portion of nodules. The genes that are only expressed in bacterial cells before the formation of nodules are even more difficult to monitor using these approaches because of the difficulties in collecting material.

Given our understanding and interests in identifying *S. meliloti* genes that function in the early stages of the nodulation, *In Vivo* Expression Technology (IVET) seems to be a more suitable approach. It has been used previously to identify *S. meliloti* genes that function in bacteroids (156). This technology was originally developed to identify bacterial pathogenic genes by identifying promoters that are activated during the invasion of host cells (142, 174). The key to the success of this approach is the use of reporter genes to screen for promoters that are activated at a particular developing stage. The *S. meliloti exoY* gene is required for the formation of infection threads, so it can be used as a perfect trap gene to screen for promoters that are active during the early stages of symbiosis.

Using an improved IVET approach, I have identified a total of 113 mostly previously unknown *S. meliloti* symbiotic genes functioning in the early stages of the symbiosis establishment.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

The bacterial strains and plasmids used in this project were listed in Table 2-1.

Table 2-1. Strains and plasmids (1)

Strain or plasmid	Relevant properties	References
E. coli		
DH5α	General purposed strain	(103)
MT616	MT607, pRK600, Cm^R	(74)
S. meliloti		
Rm1021	Wild-type, Sm^R	(126)
RmAR9007	Rm1021 *exoY-LacZ*, Gm^R	(117)
Plasmids		
pBBR1MCS3	Broad host range, Tc^R	(122)
pRK600	Helper plasmid, Cm^R	(74)
pMCS5	Cloning vector, Ap^R	(110)
pTR102	Carrying *parDE* locus, Tc^R	(224)
pHC77	Carrying *gfp* gene, Sp^R	(48)
pHC123	Promoter-trap vector, Tc^R	This work
pHC153	pHC123 with *exoY* native promoter, Tc^R	This work
pHC169	pHC123 with carrying fusion with s*mb21651* promoter, Tc^R	This work
pHC170	pHC123 with *cyaH* ORF, Tc^R	This work

Escherichia coli strains used in this study were grown in Luria-Bertani (LB) medium at 37°C. *Sinorhizobium meliloti* was grown in LB medium supplemented with 2.5mM $MgSO_4$ and 2.5mM $CaCl_2$ (LB/MC) at 30°C. Agar (1.5%) was used to solidify media. To examine succinoglycan production, calcofluor white M2R (CF) (Fluorescent Brightener 28, Sigma) was added to a final concentration of 0.02% in LB/MC agar, which was buffered to pH 7.4 with 10mM HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid) (93, 94, 126). The following antibiotics were used at the concentrations indicated: Ampicillin, 100μg/μl; chloramphenicol, 10μg/ml; gentamycin, 50μg/ml; kanamycin, 25μg/ml; neomycin, 200μg/ml; spectinomycin, 100μg/ml; streptomycin 500μg/ml; and tetracycline, 10μg/ml.

Construction of promoter trap vector

The promoter trap vector, pHC123, is a derivative of the plasmid pBBR1MCS3 (122) with an additional stability region, *exoY* and *gfp* genes, and restriction enzyme sites for cloning *S. meliloti* genomic DNA fragments. To construct this vector, a DNA fragment containing *E. coli parD* and *parE* genes was amplified by PCR using plasmid pTR102 (224) as template and the pair of primers *parDE*f*Xba*I and *parDE*r*Sac*II (Table 2-2) (Figure 2-1 PCR-1), and cloned into pBBR1MCS3, which resulted in pHC113. A DNA fragment containing transcriptional fusion of the *exoY* and *gfp* genes was first assembled in plasmid pMCS5 (110) and then cloned into pHC113. The 0.77Kb *Xho*I/*Nhe*I DNA fragment containing the *exoY* gene with its Shine-Dalgarno (SD) sequence and the intergenic region between the *exoY* gene and its downstream *exoF* gene was generated by PCR using *S. meliloti* Rm1021 genomic DNA as template and the pair of primers *exoY*f*Xho*I and *exoY*r*Nhe*I (Table 2-2) (Figure 2-1 PCR-2). The primer *exoY*f*Xho*I contains stop codons in all the three different reading frames. The 0.86Kb *Nhe*I/*Spe*I DNA fragment containing the *gfp* open reading frame (ORF) was generated by PCR using pHC77 (48) as template and the pair of primers *gfp*f*Nhe*I and *gfp*r*Spe*I (Table 2-2) (Figure 2-1 PCR-3), and then cloned into the *Nhe*I/*Spe*I site of pMCS5, which resulted in pHC108. The *gfp* gene used here contains the S65T mutation, which is also known as the super *gfp* gene (83). The 1.6kb fragment containing the *exoY* and *gfp* genes was obtained from the pMCS5 plasmid by PCR using the pair of primers *exoY*f*Xho*I and *gfp*r*Kpn*I (Table 2-2) (Figure 2-1 PCR-4), and then was subcloned into the unique *Xho*I/*Kpn*I

restriction site of plasmid pHC113 with stability region to make the promoter trap vector, pHC123.

Table 2-2. Oligo DNA primers used to construct promoter trap vector pHC123.

Name	Sequence of primers
*parDE*f*Xba*I	5'-GCTCTAGAAATGCGCCTAGCGCATT-3'
*parDE*r*Sac*II	5'-GCGTCCCCGCGGGGAAGTACGCCATCAGGACGTTG-3'
*exoY*f*Xho*I	5'-GCCTCGAGTAACGTAACGTAAATGGAGTCACCT-3'
*exoY*r*Nhe*I	5'-GCGCTAGCCATTTTTGGATATTTCCGTTC-3'
*gfp*f*Nhe*I	5'-GCGATGGCTAGCAAAGGAGAAGA-3'
*gfp*r*Spe*I	5'-GCGACTAGTATAGTTCCTCCTTTCAGCAA-3'
*gfp*r*Kpn*I	5'-CGGGGTACCCCGATAGTTCCTCCTTTCAGCAAAAAA-3'

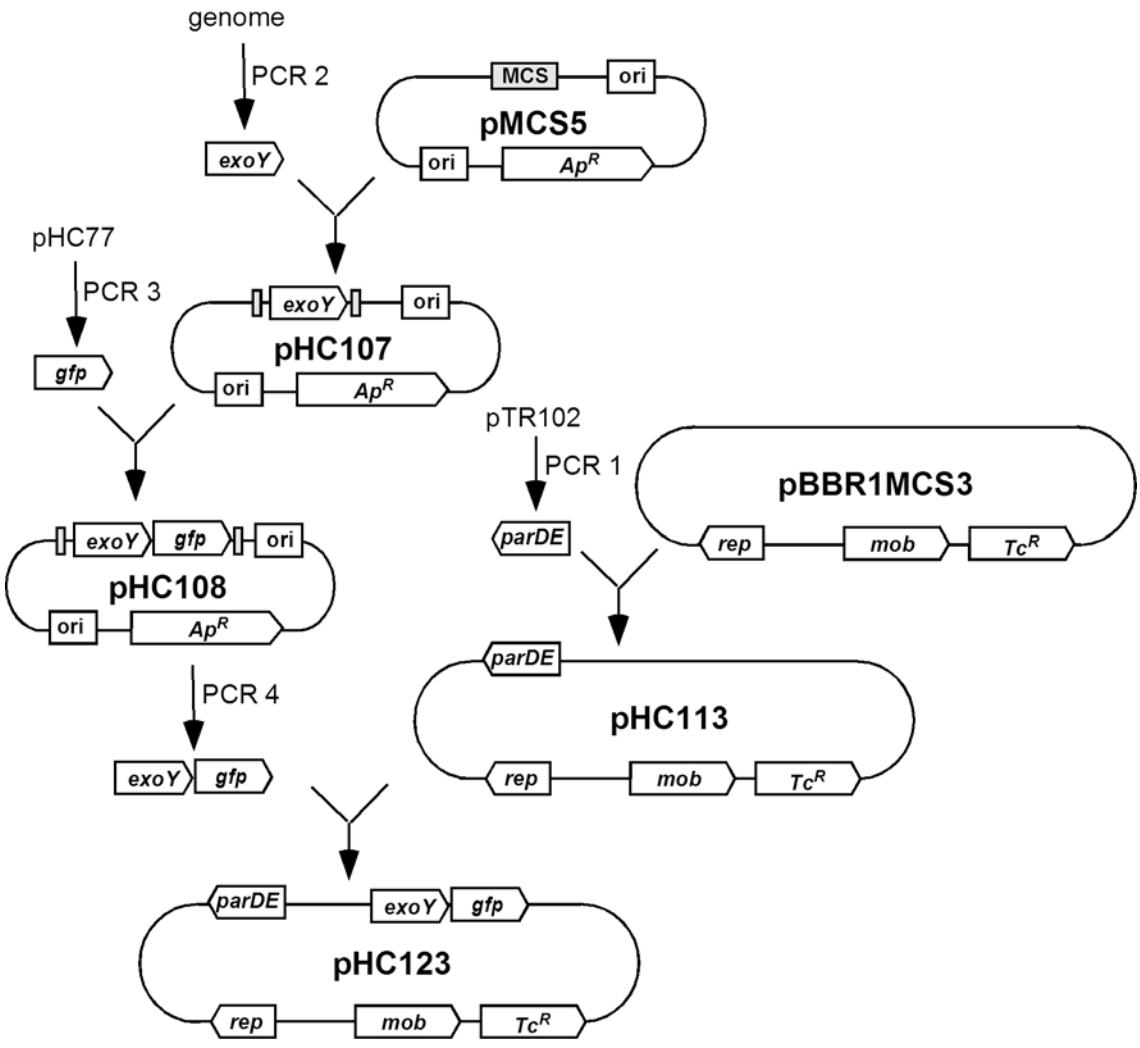

Figure 2-1. A diagram of construction of promoter trap vector pHC123.

Construction of promoter library

To construct promoter fusion library, *S. meliloti* Rm1021 genomic DNA was partially digested with restriction enzyme *Sau*3AI and separated based on size by agarose gel electrophoresis. The DNA fragments in 0.5-1Kb size range were recovered from agarose gel and cloned into the promoter trap vector, pHC123, using the partial filling-in technique as previously described (156). Briefly, the four-nucleotide overhang (GATC) of the *S. meliloti* genomic DNA fragments generated by restriction enzyme *Sau*3AI was partially filled in using dGTP, dATP, and klenow fragment, leaving a two-nucleotide "GA" overhang (Figure 2-2). Vector pHC123 was digested with restriction enzyme *Xho*I first, and the four-nucleotide overhang (TCGA) generated was partially filled in using dTTP, dCTP, and klenow fragment, leaving the "TC" overhang. Next, the resulted genomic DNA fragments and the promoter trap vectors were ligated together with T_4 DNA ligase. The ligation products were electroporated into *E. coli* DH5α following procedures recommended by the manufacture. A total of 7×10^4 independent transformants were collected and the DNA plasmids they conceived were transferred into *S. meliloti* RmAR9007 strain, the *exoY* null mutant (117), through conjugation using *E. coli* MT616 (pRK600) as the helper. A total of 3×10^5 colonies were collected and mixed in LB/MC with 5% DMSO and stored at -80°C for future analysis.

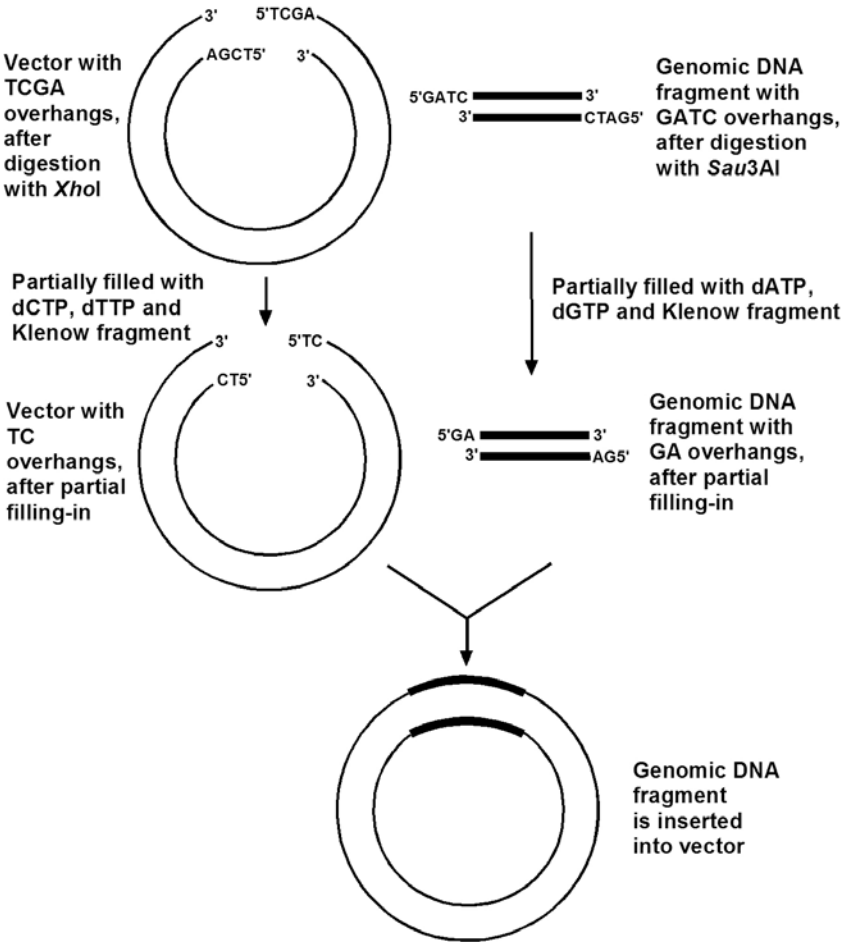

Figure 2-2. Strategy of cloning genomic DNA fragments into the promoter trap vector using partial filling-in techniques.

Determination of GFP fluorescence intensities of bacterial colonies

The GFP fluorescence intensities of bacterial colonies were examined under Zeiss fluorescence microscope, M^2 Bio Quad System at ×100 magnification. Images were obtained by using a filter set consisting of an exciter band pass filter (470 nm +/- 40 nm), a 495 nm long pass dichroic mirror, and an emission band pass filter (525 nm +/- 50 nm). Images were captured by software provided by Optronics using digital camera: Magnafire Model S99802 by Optronics, Goleta, CA installed on PC.

Determination of specific GFP fluorescence intensity of bacterial culture

Liquid culture of each *S. meliloti* strain were centrifuged, and resuspended in sterile phosphate-buffered saline (PBS, pH 7.4) to an OD_{600} of about 0.1. Equal volume of each suspension was transferred to wells of a transparent 96-well plate and a black 96-well plate. Cell suspensions in transparent 96-well plate were used to determine the cell density using an absorbance microplate reader (Spectra Max 340PC, Molecular Device, Sunnyvale, CA, USA), and suspensions in black 96-well plate were used to determine the intensities of GFP fluorescence using a fluorescence microplate reader (Spectra Max Germini XS, Molecular Device, Sunnyvale, CA, USA). The GFP fluorescence intensity of each bacterial suspension was normalized to its cell density to generate the specific GFP fluorescence intensity.

Removal of cells carrying housekeeping promoters through FACS sorting

RmAR9007 cells carrying the promoter library were collected from fresh LB/MC liquid culture, washed with PBS (pH 7.4) at 4°C, and resuspended in 1 ml ice-cold PBS (214). Before sorting, cell suspensions were diluted to a concentration of 2×10^6 cell/ml in ice-cold phosphate-buffered saline. Cells were separated into three groups by FACS (Moflo-MLS, Cytomation, Fort Collins, CO.) at Albert Einstein College of Medicine based on GFP fluorescence intensities of individual cells. RmAR9007 (pHC153) cells, which expresses both *exoY* and *gfp* genes from native *exoY* promoter region, were used as the standard to separate RmAR9007 cells carrying the promoter library. The cells of RmAR9007 carrying the promoter library showing GFP fluorescence intensities similar

to those of the cells of RmAR9007 (pHC153) were collected as dim sublibrary. The cells of RmAR9007 carrying the promoter library showing higher or lower levels of GFP fluorescence intensities were collected as bright and dark sublibraries, respectively. The cells of each sublibrary were amplified with overnight growth at 30°C in LB/MC liquid with gentamycin and tetracycline, and stored at -80°C for future screening.

Plant growth conditions

Alfalfa plants growing in hydroponic system with nitrogen-free Jensen medium (126) were used to screen *S. meliloti* RmAR9007 cells containing the dark sublibrary. Alfalfa seeds were surface-sterilized in 50% Clorox bleach and washed with sterile distilled water as previously described (126). The surface-sterilized seeds were spread on wet cheesecloth suspended with aluminum screen in 11×11×9 mm^3 containers with 200 ml nitrogen-free Jensen liquid medium. Two ends of the cheesecloth were submerged in liquid medium to ensure that the cheesecloth on the aluminum screen would remain wet. The containers with seeds were kept in dark at 26°C for 40 hours for seed germination. The germinated seedlings were then inoculated with 1 ml bacterial suspension which was collected from fresh bacterial culture in LB/MC medium, washed and diluted with sterile distilled water to an OD_{600} of 0.3. The seedlings were incubated with light at 26°C, and checked for pink nitrogen-fixing nodules after four-week incubation.

Alfalfa plants growing on Petri dishes were used to determine the ability of individual bacterial strains to establish symbiosis with alfalfa (126). Alfalfa seeds were surface-sterilized, allowed to germinate on 0.7% agar plate in dark at 26°C for 40 hours. The

seedlings were transferred to 9×9 mm^2 square Petri dishes each containing 30 ml Jensen agar with 6 seedlings per plate. Then the seedlings on each plate were inoculated with 1 ml *S. meliloti* cell suspension (OD_{600} 0.03), and incubated with light at 26°C for four weeks. The alfalfa plants in the Petri dishes were examined for height, color, and the number of nodules.

Recovery of bacterial cells from root nodules

Bacterial cells were recovered from alfalfa root nodules as previously described (126). Briefly, nodules were removed from roots, surface-sterilized with 2-minute incubation in 50% Clorox bleach, washed six times in sterile distilled water, and crushed inside the wells of 96-well plates in 100µl LB/MC liquid medium with 5.4% glucose. This suspension was then diluted 1:100 in the same medium and plated on LB/MC agar plates with antibiotics for bacterial colonies.

Qualitative detection of succinoglycan production

The levels of succinoglycan production by *S. meliloti* strains were estimated as previously described (126). Briefly, bacterial cells were plated on LB/MC/CF agar medium and incubated for 2 days. The calcofluor white M2R (Fluorescent Brightener 28, Sigma) in the medium will bind to succinoglycan and emit white fluorescence light under ultraviolet light UV light. The levels of succinoglycan produced by different strains were estimated based on the intensities of calcofluor fluorescence.

Sequence analysis of trapped *S. meliloti* genomic DNA

Plasmid DNA was isolated from *S. meliloti* cells using a QIAprep Spin Miniprep protocol (Qiagen) following manufacturer's instructions. The plasmids were electroporated into *E. coli* DH5α. The plasmids purified from *E. coli* DH5α (using QIAprep Miniprep) were sequenced by Cancer Research Center DNA sequencing facility at the University of Chicago, using sequencing primer 5'-CGA ACT GGC CGA GCG CGT C-3', which anneals to the *exoY* ORF. The upstream end of the trapped genomic DNA fragment was estimated based on the size of the genomic DNA fragment. The DNA sequences of the trapped genomic DNA fragments were analyzed using the *S. meliloti* strain Rm1021 genome project website (http://bioinfo.genopole-toulouse.prd.fr).

RESULTS

Construction of a promoter library covering the entire *S. meliloti* genome

To screen *S. meliloti* symbiotic genes in the early stages of the symbiosis establishment, a promoter library in *exoY* null mutant was constructed, cleared of constitutively active promoters using FACS, and then screened for symbiotically active promoters in nodulation assay (Figure 2-3).

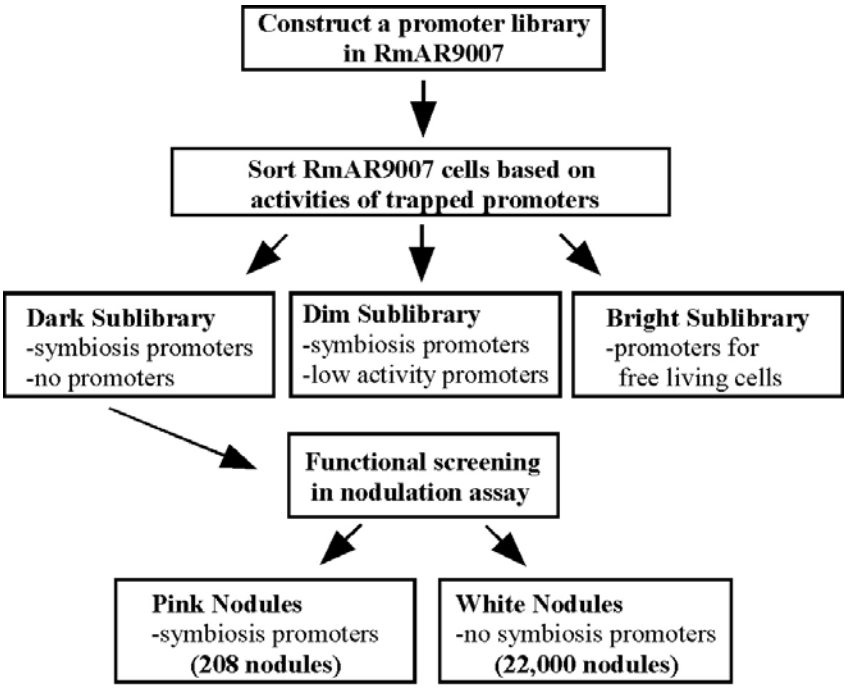

Figure 2-3. Schematic diagram of the screen for the *S. meliloti* genes functioning during the early stages of the symbiosis establishment.

A promoter trap vector that can be stably maintained in *S. meliloti* cells without antibiotic selection was constructed because the use of antibiotics would interfere with the growth of alfalfa plant. The vector pHC123 was stabilized by the presence of the *E. coli parD* and *parE* genes in the plasmid (46). This promoter trap vector contains the promoterless

S. meliloti exoY gene followed by the *gfp* gene replacing the native *exoF* ORF (Figure 2-4). Both *exoY* and *gfp* genes are transcribed in the same transcript but translated into two independent proteins. The specific GFP fluorescence intensities of bacterial cells, which reflect the amount of the intracellular GFP and ExoY proteins, can be used to monitor the levels of the *exoY* gene expression.

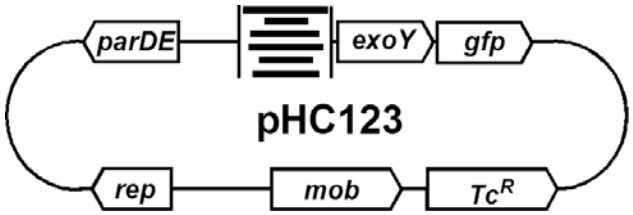

Figure 2-4. Schematic representation of the promoter trap vector showing relative positions of key genes including *exoY, gfp,* and *parDE*, and the Rm1021 genomic DNA fragments in front of the *exoY* gene.

To test the feasibility of this approach, the native *exoY* promoter region was cloned in front of the *exoY* gene in the promoter trap vector to form plasmid pHC153, which was moved into *S. meliloti* RmAR9007, a nonpolar *exoY* null mutant that does not have succinoglycan production. The strain RmAR9007 (pHC153) produced succinoglycan, suggesting that the *exoY* and *gfp* genes on this plasmid were expressed from the native *exoY* promoters (Figure 2-5). More importantly, RmAR9007 (pHC153) was also able to establish nitrogen-fixing symbiosis with alfalfa (Figure 2-9), whereas RmAR9007 (pHC123), the control strain which carries the promoter trap vector, was not (data not shown). This confirmed that the expression of the *S. meliloti exoY* gene from plasmid pHC153 can suppress the symbiotic deficiency of RmAR9007, and that plasmid pHC123 can be used to trap promoters that may direct *exoY* gene expression during the early stages of the symbiosis.

Figure 2-5. *S. meliloti* strains with different promoter fusions showing the different intensities of calcofluor fluorescence, which reflects the levels of succinoglycan production by the strains. RmAR9007, *exoY* null mutant, does not produce succinoglycan. RmAR9007 (pHC153) with *exoY* native promoter fusion produces succinoglycan at a low level. RmAR9007 (pHC169) with trapped constitutively active promoter fusion produces succinoglycan at a high level. And RmAR9007 (pHC170) with trapped non-promoter region fusion does not produce succinoglycan.

To construct a promoter library for a more effective promoter screening, a collection of small-size (0.5-1 Kb) *S. meliloti* genomic DNA fragments was cloned in front of the *exoY* gene in pHC123. The constructed plasmids were transformed into *E. coli* DH5α by electroporation. To verify the size of trapped genomic DNA, plasmids were purified individually from 50 random transformants and further analyzed by restriction digestion. Results showed that the average size of the trapped genomic DNA fragments was 0.75kb. Based on the formula N= ln (1-P)/ln (1-f) (N: size of minimum library, P: with a probability of 95%, f: fraction of the size of cloned fragment to total size of genome 0.75kb/6700kb) (192), the minimum size of the library should be 3×10^4 for trapping any promoter with a probability of 95%. Considering that promoters may be trapped in either direction, at least 6×10^4 independent transformants are needed to screen all possible promoters in *S. meliloti* genome. A total of 7×10^4 independent *E. coli* DH5α transformants were actually collected as the promoter library. The plasmids were

conjugally transferred into *S. meliloti* RmAR9007 (see methods), and a total of 3×10^5 independent conjugants were collected to ensure the complete transfer of the promoter library.

RmAR9007 cells carrying the promoter library formed colonies with different intensities of GFP fluorescence (Figure 2-6). About 10% of the colonies exhibited high levels of GFP fluorescence and 90% displayed low levels or no GFP fluorescence. Two plasmids, named pHC169 and pHC170, were isolated from two randomly picked colonies with high and no GFP fluorescence, respectively. Sequencing results showed that plasmid pHC169 trapped the promoter region of the *smb21651* putative ORF, and that plasmid pHC170 trapped a region from the middle of the *S. meliloti cyaH* gene.

The two plasmids were moved into strain RmAR9007 through conjugation. When streaked on LB/MC plate with calcofluor, RmAR9007 (pHC169) showed succinoglycan production while RmAR9007 (pHC170) did not (Figure 2-5). This confirmed the fusion on pHC169 carries the promoter that directed the *exoY* expression, which complemented *exoY* mutation of RmAR9007 strain.

In nodulation assays, plants inoculated with Rm1021 (the wild-type strain) were tall and green with 4-12 pink Fix^+ nodules per plant as expected after four weeks. The plants inoculated with RmAR9007 and RmAR9007 (pHC170) were yellow and short with 8-25 white Fix^- nodules per plant. In contrast, the plants inoculated with RmAR9007

(pHC169) were tall and green with pink Fix^+ nodules like those inoculated with the wild type (Figure 2-7).

These findings demonstrated that the nodulation assay could be used to screen the promoter library for the promoters that are active during the symbiosis. The presence of these promoters, such as the *exoY* promoter, in the promoter trap vector is essential for complementing the genomic *exoY* null mutation and for successful symbiosis. But the nonsymbiotic promoters that are highly active in free-living cells, such as the one trapped in plasmid pHC169, could remain active during the early stages of the symbiosis establishment so that these promoters need to be removed from the library in order to screen for symbiotic promoters.

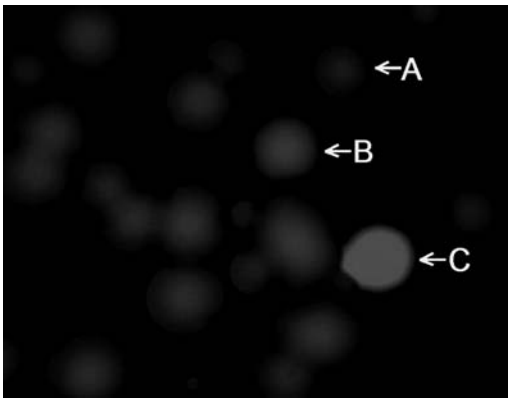

Figure 2-6. A picture of colonies of *S. meliloti* RmAR9007 strain carrying the unsorted promoter library showing different intensities of GFP fluorescence. Colony C was formed by cells carrying plasmid pHC169 and colony A was formed by cells carrying plasmid pHC170. Colony B was formed by cells showing low level of GFP fluorescence, which was not further analyzed.

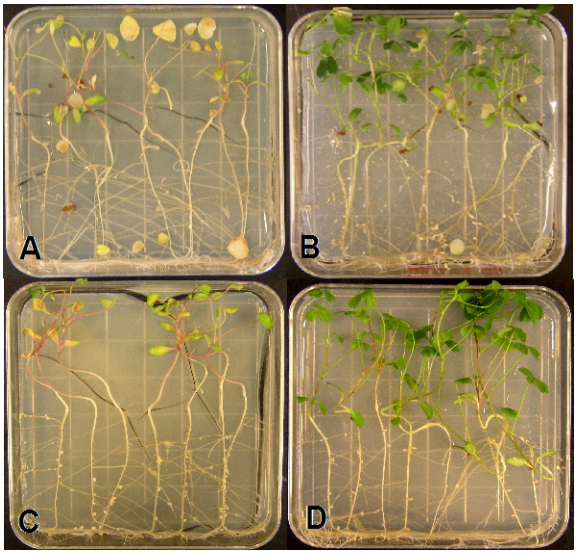

Figure 2-7. Pictures of four-week old alfalfa plants inoculated with different *S. meliloti* stains. A) no bacterial control, B) Rm1021, C) RmAR9007 (pHC170), and D) RmAR9007 (pHC169).

All together, these findings suggest that a promoter library has been constructed and that it can be used to identify promoters based on their abilities to express the *exoY* gene and suppress the symbiotic defect of the *exoY* null mutant.

Removal of promoters that are active in free-living cells from the promoter library

The promoter library contains *S. meliloti* promoters of housekeeping genes that are expressed both in free-living conditions and during symbiosis. The promoter of those genes could express the *exoY* gene and create false positive nodule isolates. In order to remove those promoters from the promoter library, RmAR9007 cells carrying the promoter library were separated by FACS into three populations, bright, dim, and dark, using the intensities of GFP fluorescence of RmAR9007 (pHC153) cells as a standard (Figure 2-8). The bright, dim, and dark sublibrary accounted for 4%, 25%, and 71% of the total population, respectively.

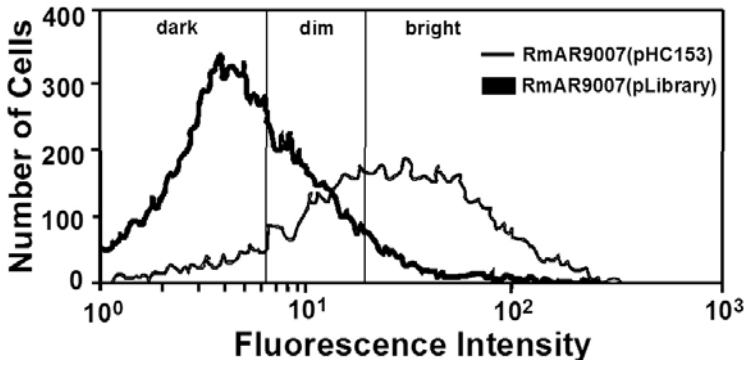

Figure 2-8. FACS profiles of intensities of GFP fluorescence of RmAR9007 (pHC153) and RmAR9007 carrying the entire promoter library. The activities of the *exoY* promoter in free-living RmAR9007 (pHC153) cells, which is reflected in the levels of GFP fluorescence intensities of the cells, were used as a guide to determine the levels of GFP fluorescence to separate the cells of RmAR9007 carrying the promoter library into dark, dim, and bright populations. The cells with fluorescence intensities less than 6.3 were collected as the dark sublibrary. The cells with fluorescence intensities between 6.3 and 20.6 were collected as dim sublibrary. The cells with fluorescence intensities above 20.6 were collected as bright sublibrary.

The effectiveness of the FACS sorting was further examined by plating cells of the dark and dim sublibraries on LB/MC agar media because the activities of the same promoter can vary in different cells (123, 163). The intensity of GFP fluorescence of any given bacterial colony reflects the average activities of the trapped promoter in different cells. Colonies formed by cells from the dark sublibrary were all dark or close to dark under fluorescence microscope, suggesting that the promoter library carried by most of cells in this sublibrary did not contain promoters constitutively active in free-living conditions. A small portion of colonies formed by cells from the dim sublibrary, however, exhibited high levels of GFP fluorescence intensities, suggesting that the dim sublibrary contained the promoters that are highly active in free-living conditions. In an effort to remove these active promoters, the entire dim sublibrary was plated on solid media for single colonies. After the GFP bright colonies were removed from the plates individually, the remaining colonies of the dim sublibrary were collected, pooled and plated on the LB/MC solid media again. However, when reexamined under fluorescence microscope, a similar

percentage of colonies in the dim sublibrary were found to exhibit high levels of GFP fluorescence intensities again. Since this small portion of housekeeping promoters in the dim sublibrary could not be easily removed, the dim sublibrary was not further screened for the symbiotic promoters that are expressed at low levels in free-living cells but will be highly induced during symbiosis.

Isolation of *S. meliloti* symbiotic promoters by functional complementation

Using a specially designed hydroponic system, a total of 2,400 alfalfa plants were used to screen the dark sublibrary. Most of the alfalfa plants were yellow and short while a few were slightly green after four-week growth (Figure 2-9). Some of the plants had one or two pink nodules, and few plants had more than two pink nodules. The alfalfa plants that had one or more pink nodules did appear slightly greener than those that did not have any pink nodules on them. After screening around 22,000 root nodules, a total of 208 pink nodules were identified. These nodules were crushed to generate 208 nodule isolates.

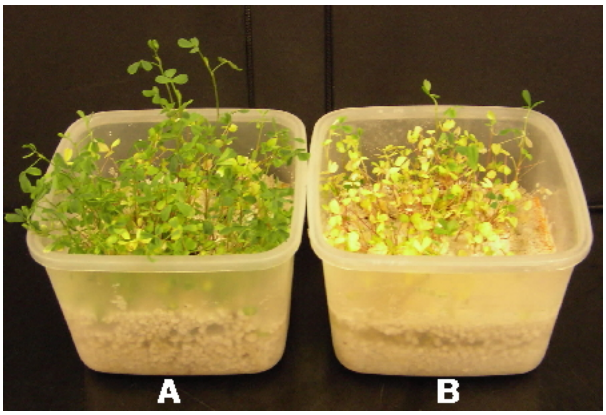

Figure 2-9. A picture of alfalfa plants inoculated with *S. meliloti* RmAR9007 (pHC153) and with RmAR9007 carrying the dark sublibrary. Alfalfa plants in tray A were inoculated with RmAR9007 (pHC153) and they were green and healthy. The RmAR9007 cells carrying the dark sublibrary were used to screen for symbiotically active promoters in tray B. Pink nodules were isolated from the few green plants in tray B.

Each of the 208 nodule isolates was used to inoculate 6 alfalfa plants, all of which became green and tall with 55-100% pink nodules four weeks later. Alfalfa plants inoculated with the negative control, RmAR9007 (pHC123), were almost dead with only white nodules. Alfalfa plants inoculated with the positive controls Rm1021 and RmAR9007 (pHC153) were tall and green with 100% pink nodules. These results suggest that these 208 nodule isolates indeed carried the promoters that are active during the early stages of the symbiosis.

Analysis of the specific GFP intensity of each nodule isolate showed that the promoter activities of 183 out of 208 nodule isolates were lower than that of RmAR9007 (pHC153) strains, suggesting that the activities of these promoters were lower than the native *exoY* promoter in free-living cells. All together, these findings suggest that these 183 nodule isolates most likely have putative promoters that are only active during early stages of nodulation.

Identification of *S. meliloti* symbiotic genes based on trapped promoters

DNA sequencing of the trapped genomic DNA fragments in these 183 nodule isolates showed that the DNA fragments originated from 113 different regions of the *S. meliloti* genome including the chromosome and both megaplasmids. The genes were listed in Table 2-3. They include genes functioning in cell motility or chemotaxis, stress response and adaptation, nutrient uptake, central intracellular metabolism, transcription regulation and unknown functions (Table 2-3). The putative promoter which directs the each of the identified genes were predicted using prokaryotic promoter prediction software online

and listed in Table 2-4. The identified genes were analyzed based on previous reports and by homology. Most of them have not been characterized before in *S. meliloti*. A detailed description of these identified genes has been included in Appendix and some examples of their homologs were listed in Table 2-5. These identified genes were compared with microarray data of previous reports (9, 11), results showed that most of them have not been previously described in microarray analysis of gene expression in bacteroids, because the level of the expression of these genes did not change more than two folds after bacterial cells were released into plant cells (Table 2-6).

Table 2-3. List of genes expressed during early stages of nodulation

Gene or ORF	Functions or putative functions
*lsrA (smc00037)**	LysR type transcriptional regulatory protein
dgkA (smc04213)	Diacylglycerol kinase
exsA (smb20941)	Msba-like saccharide exporting ABC transporter protein, involved in exopolysaccharide EPSII secretion
flaC (smc03040)	Flagellin protein, involved in bacterial mobility chemotaxis
fliF (smc03014)	Transmembrane flagellar m-ring protein, involved in bacterial mobility chemotaxis
degP1 (smc02365)	Protease precursor, involved in degradation of proteins, peptides or glycopeptides
ppdK (smc00025)	Pyruvate phosphate dikinase, involved in gluconeogenesis
thuF (smb20326)	Trehalose/maltose transporter permease
thuG (smb20327)	Trehalose/maltose transporter permease
thuK (smb20328)	Trehalose/maltose transporter ATP-binding protein, respectively
glxD (smc02612)	Glutamate synthase large subunit-like protein Oxidoreductase, involved in small molecule metabolism
ppe (smb20195)	Putative pentose-5-phosphate-3-epimerase protein, involved in small molecule metabolism
sma0340 (wrbA2)	Putative trp-repressor binding protein, regulator
smb21662 (exoI2)	Putative periplasmic protein
smc00234 (ppiD)	Putative peptidyl-prolyl cis-trans isomerase
smc00289 (cspA5)	Putative cold shock-like transcription regulator protein, involved in bacterial adaptation to atypical conditions
smc00288	Signal peptide protein
smc00643 (purA)	Putative adenylosuccinate synthetase IMP-aspartate ligase, involved in synthesis of purine ribonucleotide
smc00644	Hypothetical protein
smc00695 (aroK)	Putative shikimate kinase
smc00696 (aroB)	Putative transmembrane 3-dehydroquinate synthase
smc00697	Hypothetical transmembrane protein
smc00811 (FtsJ)	Probable cell division protein
smc00812	Hypothetical protein
smc01512 (hmuT)	Hemin binding periplasmic protein, putative hemin transport system
smc01511 (hmuU)	Hemin transport system permease
smc01510 (hmuV)	Hemin transport system ATP-binding protein
smc02426	Conserved hypothetical protein
smc02427 (hyuE)	Putative hydantoin racemase, involved in amino acid biosynthesis
smc02653 (lepB)	Putative probable transmembrane signal peptidase I, involved in cell processes or transport of small molecules
smc02737 (opuC)	ABC-type glycine betaine transporter
smc03806 (glnK)	Probable nitrogen regulatory protein PII, involved in nitrogen metabolism
smc03807 (amtB)	Putative ammonium transporter, involved in nitrogen metabolism
smc03927 (nodN2)	Putative nodulation protein, involved in nodulation
smc03928	Conserved hypothetical protein
smc04047 (azu2)	Putative copper-binding protein Azu2, involved in electron transport
smc04048	Putative cytochrome C protein, involved in electron transport
smc04234 (csp4)	Putative cold shock-like transcription regulator protein, involved in bacterial adaptation to atypical conditions

***: For previously characterized genes, the names of the genes are listed with the serial numbers of the genes designated by the Sequencing Project listed in the parentheses. For previously unknown genes, the serial numbers of the genes are listed with the names of the homologs proposed by genome sequencing projects listed in parentheses.**

Table 2-3. List of genes expressed during early stages of nodulation (continues)

Gene or ORF	Functions or putative functions
smc00880	Putative oxidoreductase
smc00881 (dgoK1)	2-dehydro-3-deoxygalactonokinase, involved in small molecule metabolism
smc00882	KHG/KDPG aldolase, involved in small molecule metabolism
smc00883	Conserved hypothetical protein
smc01522	Hypothetical protein
smc01523 (emrE)	Putative transmembrane multi-drug transporter
smc00072	Putative peroxiredoxin protein
smc00371	Conserved hypothetical protein
smc00570	Putative oxidoreductase
smc00722	Hypothetical transmembrane protein
smc00856	Hypothetical transmembrane protein
smc00976	Putative an enoyl-coA hydratase, involved in fatty acid metabolism
smc00977	Putative acyl-coA dehydrogenase, involved in fatty acid degradation
smc01214	Putative zinc-containing alcohol dehydrogenase protein, involves in small molecule metabolism
smc01488	Hypothetical transmembrane protein
smc01571	Putative oxidoreductase
smc01572	Putative oxidoreductase
smc01577	Conserved hypothetical protein
smc01731	Hypothetical protein
smc01759	Conserved hypothetical protein
smc01817	Putative LysR-family transcriptional regulator
smc01945	Putative transcription regulator
smc01957	Conserved hypothetical protein
smc01988	Hypothetical protein
smc02151	Putative virulence associated protein homolog
smc02171	Putative periplasmic binding protein of an ABC transporter, typically involved in the transport of small molecules
smc02381	Conserved hypothetical protein
smc02616	Hypothetical transmembrane permease
smc02617	Hypothetical transmembrane protein
smc02618	Hypothetical transmembrane protein
smc02619	Hypothetical transmembrane protein
smc02817	Hypothetical transmembrane protein
smc02818	Conserved hypothetical protein
smc02987	Conserved hypothetical protein
smc02988	Conserved hypothetical protein
smc03130	Hypothetical protein
smc03746	Hypothetical protein
smc03767	Conserved hypothetical protein
smc03844	Conserved hypothetical protein
smc03890	Putative transcription regulator
smc04054	Conserved hypothetical protein
smc04194	Hypothetical transmembrane protein
smc04311	Hypothetical transmembrane protein
smc04391	Putative oxidoreductase, involved in small molecule metabolism
smc04439	Putative glycine betaine transport ATP-binding protein
sma0126	Putative cold shock protein, regulator
sma0232	Hypothetical protein
sma0254	Hypothetical protein
sma0447	Conserved hypothetical protein
sma0537	Hypothetical protein
sma0538	Hypothetical protein

Table 2-3. List of genes expressed during early stages of nodulation (continues)

Gene or ORF	Functions or putative functions
sma0945	Hypothetical protein
sma0969	Putative response regulator of two-component system
sma1124	Hypothetical protein
sma1740	Conserved hypothetical protein
sma1760	Hypothetical protein
sma1761	Putative aminotransferase, involved in small molecules metabolism
sma1898	Hypothetical protein
sma1921	Hypothetical protein
sma1954	Putative LysR-family transcriptional regulator
sma2049	Putative LacI-family transcription regulator, respond to the presence of small effector molecules
sma2199	Putative periplasmic solute-binding ABC transporter, involved in transport of small molecules
sma2235	Hypothetical protein
sma2239	Conserved hypothetical protein
sma2241	Hypothetical protein
smb20156	Hypothetical ABC transporter ATP-binding protein
smb20162	Putative *luxR* family transcription factor
smb20179	Hypothetical signal peptide protein
smb20180	Hypothetical signal peptide protein
smb20243	Putative glycosyltransferase protein, involved in surface polysaccharides or antigens synthesis
smb20278	Hypothetical protein
smb20399	Hypothetical protein
smb20450	Putative regulatory protein
smb20451	Hypothetical protein
smb20452	Hypothetical protein
smb20463	Hypothetical protein
smb20476	Putative periplasmic dipeptide-binding ABC transporter protein
smb20498	Conserved hypothetical protein
smb20504	Putative ABC transporter periplasmic sugar-binding protein
smb20590	Hypothetical protein
smb20591	Hypothetical protein
smb20600	Hypothetical protein
smb20724	Conserved hypothetical exported protein
smb20725	Conserved Hypothetical transmembrane protein
smb20861	Putative oxidoreductase FAD flavoprotein, involved in small molecules metabolism
smb20888	Conserved hypothetical protein
smb20909	Hypothetical protein
smb20910	Hypothetical protein
smb21157	Putative translation initiation inhibitor
smb21158	Putative transcriptional regulator, involved in sugar phosphate metabolism
smb21202	Hypothetical protein
smb21259	Hypothetical exported protein precursor
smb21333	Hypothetical protein
smb21397	Hypothetical calcium-binding protein
smb21400	Hypothetical protein
smb21401	Hypothetical protein
smb21402	Hypothetical calcium-binding protein
smb21468	Conserved hypothetical membrane protein
smb21595	Putative sugar-uptake ABC transporter periplasmic solute-binding protein precursor

Table 2-3. List of genes expressed during early stages of nodulation (continues)

Gene or ORF	Functions or putative functions
smc01203R	Hypothetical protein
smc02230R	Hypothetical protein
smc02773R	Hypothetical protein
smc03181R	Hypothetical protein
smc03205R	Hypothetical protein
smc04135R	Hypothetical protein
sma1252R	Hypothetical protein
sma1523R	Hypothetical protein
smb20463S1	Hypothetical protein
smb20463S2	Hypothetical protein
smb20276R	Hypothetical protein
smb20714R	Hypothetical protein
smb21066R	Hypothetical protein
smb21573R	Hypothetical protein

Table 2-4. List of predicted putative promoters trapped in the fusions that direct the newly identified symbiotic genes expression

Gene or ORF	Score(s) of the putative promoter(s) assigned by promoter prediction software*
lsrA (smc00037)	0.40
dgkA (smc04213)	0.90
exsA (smb20941)	0.95, 0.89
flaC (smc03040)	0.75
fliF (smc03014)	0.92
degP1 (smc02365)	0.69, 0.54
ppdK (smc00025)	0.44, 0.42
thuF (smb20326)	0.65
thuG (smb20327)	0.65
thuK (smb20328)	0.65
glxD (smc02612)	0.83, 0.81, 0.75
ppe (smb20195)	0.90, 0.80
sma0340 (wrbA2)	0.94, 0.59
smb21662 (exoI2)	0.89, 0.72, 0.52
smc00234 (ppiD)	0.80
smc00289 (cspA5)	0.92
smc00288	0.92
smc00643 (purA)	0.96
smc00644	0.96
smc00695 (aroK)	0.95
smc00696 (aroB)	0.95
smc00697	0.95
smc00811 (FtsJ)	0.45
smc00812	0.45
smc01512 (hmuT)	0.70
smc01511 (hmuU)	0.70
smc01510 (hmuV)	0.70
smc02426	0.76, 0.71, 0.64
smc02427 (hyuE)	0.76, 0.71, 0.64
smc02653 (lepB)	0.89
smc02737 (opuC)	0.48, 0.47
smc03806 (glnK)	0.88
smc03807 (amtB)	0.88
smc03927 (nodN2)	0.48, 0.41
smc03928	0.48, 0.41
smc04047 (azu2)	0.33
smc04048	0.33
smc04234 (csp4)	0.40, 0.35
smc00880	0.95, 0.88, 0.79
smc00881 (dgoK1)	0.95, 0.88, 0.79
smc00882	0.95, 0.88, 0.79
smc00883	0.95, 0.88, 0.79
smc01522	0.83, 0.51
smc01523 (emrE)	0.83, 0.51
smc00072	0.87, 0.58
smc00371	0.22
smc00570	0.34
smc00722	0.22, 0.21
smc00856	0.45

*** Promoter is predicted using the prokaryotic promoter predicting software on line at the webpage: (http://www.fruitfly.org).**

Table 2-4. List of predicted putative promoters trapped in the fusions that direct the newly identified symbiotic genes expression (continues)

Gene or ORF	Score(s) of the putative promoter(s) assigned by promoter prediction software*
smc00976	0.94, 0.67
smc00977	0.94, 0.67
smc01214	0.98, 0.89
smc01488	0.84
smc01571	0.65, 0.48
smc01572	0.65, 0.48
smc01577	0.65, 0.48
smc01731	0.56
smc01759	0.97, 0.86
smc01817	0.96, 0.84, 0.80, 0.69
smc01945	0.92, 0.73
smc01957	0.55
smc01988	0.61
smc02151	0.73, 0.57
smc02171	0.90, 0.70
smc02381	0.97
smc02616	0.99
smc02617	0.99
smc02618	0.99
smc02619	0.99
smc02817	0.98, 0.60
smc02818	0.98, 0.60
smc02987	0.91
smc02988	0.91
smc03130	0.82
smc03746	0.96, 0.89
smc03767	0.91, 0.87
smc03844	0.65, 0.60
smc03890	0.85
smc04054	0.80, 0.53
smc04194	0.87, 0.57, 0.53
smc04311	0.93
smc04391	0.60
smc04439	0.78
sma0126	0.96
sma0232	0.50, 0.40
sma0254	0.91, 0.65
sma0447	0.91
sma0537	0.76
sma0538	0.76
sma0945	0.75, 0.72, 0.42
sma0969	0.98, 0.98
sma1124	0.63, 0.45
sma1740	0.43, 0.37
sma1760	0.27
sma1761	0.27
sma1898	0.34, 0.31
sma1921	0.86, 0.82, 0.63
sma1954	0.78, 0.76, 0.70
sma2049	0.89
sma2199	0.68, 0.51
sma2235	0.54

Table 2-4. List of predicted putative promoters trapped in the fusions that direct the newly identified symbiotic genes expression (continues)

Gene or ORF	Score(s) of the putative promoter(s) assigned by promoter prediction software*
sma2239	0.74, 0.56
sma2241	0.74, 0.56
smb20156	0.78, 0.58
smb20162	0.97, 0.69
smb20179	0.63, 0.60
smb20180	0.63, 0.60
smb20243	0.80
smb20278	0.96, 0.63
smb20399	0.90, 0.85
smb20450	0.91, 0.82, 0.67, 0.61
smb20451	0.91, 0.82, 0.67, 0.61
smb20452	0.91, 0.82, 0.67, 0.61
smb20463	0.98, 0.46
smb20476	0.72, 0.64
smb20498	0.37
smb20504	0.94, 0.81
smb20590	0.92
smb20591	0.92
smb20600	0.91
smb20724	0.99, 0.54
smb20725	0.99, 0.54
smb20861	0.79
smb20888	0.55
smb20909	0.91, 0.83, 0.64
smb20910	0.91, 0.83, 0.64
smb21157	0.96, 0.83
smb21158	0.96, 0.83
smb21202	0.47
smb21259	0.81
smb21333	0.58
smb21397	0.66, 0.44
smb21400	0.86, 0.52
smb21401	0.86, 0.52
smb21402	0.97
smb21468	0.80
smb21595	0.80, 0.70
smc01203R	0.38
smc02230R	0.41, 0.37
smc02773R	0.97, 0.97
smc03181R	0.40, 0.38, 0.37
smc03205R	0.90, 0.64, 0.62
smc04135R	0.60, 0.53
sma1252R	0.87, 0.77
sma1523R	0.80
smb20463S1	0.98, 0.46
smb20463S2	0.98, 0.46
smb20276R	0.91, 0.74
smb20714R	0.94, 0.87, 0.82
smb21066R	0.99, 0.98, 0.97, 0.80, 0.77
smb21573R	0.32

Table 2-5. The examples of homologs of the putative proteins encoded by the new discovered putative symbiotic genes

Gene or ORF	Examples of the hypothetical protein's homologs in *S. meliloti* or in other bacterial species
lsrA (smc00037)	*R. etli* CFN 42 YP_468687, 68%*; *M. loti* MAFF303099 BAB53509, 54%
dgkA (smc04213)	*Pseudomonas denitrificans* DgkA, 81%; *E. coli* K12 DgkA, 37%
exsA (smb20941)	*S. meliloti* Smc00186, 39%
flaC (smc03040)	*S. meliloti* FlaD, 98%; *S. meliloti* FlaB, 49%; *S. meliloti* FlaB, 48%;
fliF (smc03014)	*B. melitensis biovar Abortus* FliF, 50%
degP1 (smc02365)	*A. tumefaciens* C58 Dop, 69%
ppdK (smc00025)	*A. tumefaciens* C58 PpdK, 79%
thuF (smb20326)	*A. tumefaciens* PalF, 79%
thuG (smb20327)	*S. meliloti* Smc04028, 28%
thuK (smb20328)	*S. meliloti* Smc02869, 59%
glxD (smc02612)	*B. sp.* BTAi1 GlxD, 84%
ppe (smb20195)	*S. meliloti* Rpe, 65%
sma0340 (wrbA2)	*E. coli* WrbA, 64%
smb21662 (exoI2)	*S. meliloti* ExoI, 59%
smc00234 (ppiD)	*R. etli* CFN 42 PpiD, 59%
smc00289 (cspA5)	*A. tumefaciens* C58 CspA, 82%
smc00288	*A. tumefaciens* C58 Atu1804, 61%
smc00643 (purA)	*R. etli* CFN 42 PurA, 91%
smc00644	*A. tumefaciens* Atu2446, 38%; *M. loti* Mll3872, 35%; *B. suis* Br1681, 44%
smc00695 (aroK)	*R. etli* AroK, 84%
smc00696 (aroB)	*R. etli* AroB, 81%
smc00697	*E. coli* YfjD, 36%
smc00811 (FtsJ)	*E. coli* FtsJ, 42%
smc00812	*B. japonicum* USDA 110 Blr3793, 38%
smc01512 (hmuT)	*R. etli* CFN 42 HmuT, 62%
smc01511 (hmuU)	*R. leguminosarum* HmuU, 71%
smc01510 (hmuV)	*R. etli* CFN 42 HmuV, 61%
smc02426	*Desulfuromonas acetoxidans* Dsm684, 40%
smc02427 (hyuE)	*A. tumefaciens* C58 Huy, 82%
smc02653 (lepB)	*B. japonicum* SipS, 46%
smc02737 (opuC)	*A. tumefaciens* C58 OpuC, 79%
smc03806 (glnK)	*R. etli* GlnK, 92%
smc03807 (amtB)	*R. etli* AmtB, 61%
smc03927 (nodN2)	*S. meliloti* NodN, 64%
smc03928	*Rhodobacter sphaeroides* Rsph03002781, 28%
smc04047 (azu2)	*S. meliloti* Azu1, 44%
smc04048	*S. meliloti* Smb21368, 26%
smc04234 (csp4)	*R. etli* Csp4, 78%
smc00880	*S. meliloti* Smc00268, 35%; *S. meliloti* Smc02037, 32%; *S. meliloti* Smc02041, 33%
smc00881 (dgoK1)	*R. etli* DgoK, 63%
smc00882	*A. tumefaciens* C58 KdgA, 72%
smc00883	*A. tumefaciens* C58 Atu0702, 63%
smc01522	*A. tumefaciens* Atu2320, 42%
smc01523 (emrE)	*E. coli* QacEdelta1 multidrug exporter QacEdelta1, 51%
smc00072	*S. meliloti* Smc01834, 37%
smc00371	*Parachlamydia sp.* UWE25 YciF, 47%

***: The percentage of the identical amino acids shared with target protein.**

Table 2-5. The examples of homologs of the putative proteins encoded by the new discovered putative symbiotic genes (continues)

Gene or ORF	Examples of the hypothetical protein's homologs in *S. meliloti* or in other bacterial species
smc00570	*S. meliloti* MocA, 32%
smc00722	*Ralstonia eutropha* putative chemotaxis sensory transducer Reut_A2620, 25%
smc00856	*A. tumefaciens* str. C58 Atu0638, 58%; *M. loti* MAFF303099, Msl7423, 50%
smc00976	*A. tumefaciens* str. C58 Atu0733, 68%
smc00977	*S. meliloti* Smc02229, 44%
smc01214	*S. meliloti* Tdh, 30%
smc01488	*B. melitensis* BMEI1461, 41%
smc01571	*S. meliloti* Smb20871, 42%
smc01572	*S. meliloti* Sma0237, 30%
smc01577	*Bacillus anthracis* str. Sterne BAS1707, 27%
smc01731	*A. tumefaciens* DapE, 76%
smc01759	*B. japonicum* USDA 110 Bsl1405, 47%
smc01817	*S. meliloti* Sma0039, 44%; *S. meliloti* Sma2027, 35%; *S. meliloti* Sma2107, 29%
smc01945	*Bacillus licheniformis* OhrR, 47%
smc01957	*Streptomyces coelicolor* Sco2105, 28%
smc01988	*Anaeromyxobacter dehalogenans* hypothetical spermine synthase AdehDRAFT_0174, 28%
smc02151	*Dichelobacter nodosus* virulence associated protein VapA, 50%
smc02171	*A. tumefaciens* C58 FrcB, 85%; *M. loti* MAFF303099 Mlr7582, 80%
smc02381	*B. japonicum* Bll5370, 62%
smc02616	*M. loti* MAFF303099 Mll7315, 68%
smc02617	*M. loti* MAFF303099 Msl7316, 49%
smc02618	*B. melitensis* lysophospholipase L2, 47%
smc02619	*M. loti* MAFF303099 Mll7319, 44%
smc02817	*R. etli* RHE_CH02876, 66%
smc02818	*A. tumefaciens* C58 Atu2126, 51%
smc02987	*M. sp*. BNC1 PilT, 28%; *S. meliloti* Smc02715, 46%
smc02988	*S. meliloti* Smc02716, 44%
smc03130	*M. sp*. BNC1 MesoDRAFT_0565, 77%; *A. tumefaciens* str. C58 hypothetical protein Atu3565, 71%
smc03746	*Pseudomonas fluorescens* glycine cleavage system T protein GcvT, 35%
smc03767	*E. coli* probable arylsulfatase regulator protein AslB, 25%
smc03844	*Vibrio vulnificus* CMCP6 hypothetical signal transduction histidine kinase VV12578, 44%
smc03890	*S. meliloti* Smb21533, 25%; *S. meliloti* Smc02521, 23%
smc04054	*M. sp*. BNC1 putative glyoxalase MesoDRAFT_4496, 52%
smc04194	N/A
smc04311	*S. meliloti* Smc02378, 57%; *S. meliloti* Smb21572, 52%
smc04391	*B. abortus* BAB1_2043 70%
smc04439	*S. meliloti* OpuA,46%
sma0126	Rhizobium sp. NGR234 CspA, 75%
sma0232	*M. loti* MAFF303099 Mll3993, 29%
sma0254	*A. tumefaciens* C58 AGR_L_2570p, 91%; *B. japonicum* USDA 110 Bll4163, 69%; *M. loti* MAFF303099 hypothetical protein Mlr0494, 62%

Table 2-5. The examples of homologs of the putative proteins encoded by the new discovered putative symbiotic genes (continues)

Gene or ORF	Examples of the hypothetical protein's homologs in *S. meliloti* or in other bacterial species
sma0447	*S. meliloti* Smc02160, 22%; *A. tumefaciens* str. C58 Atu1739, 32%; *B. japonicum* Blr8064, 32%
sma0537	*Symbiobacterium thermophilum* IAM 14863 putative peptidase STH913, 34%
sma0538	*M. loti* MAFF303099 hypothetical protein Mlr0200, 27%
sma0945	*R. eutropha* hypothetical permease Reut_A1129, 27%
sma0969	*S. meliloti* phosphate regulon transcriptional regulatory protein PhoB, 28%
sma1124	*M. sp.* BNC1 MesoDRAFT_2732, 43%
sma1740	*M. sp*. BNC1 MesoDRAFT_0445, 36%; *A. tumefaciens* str. C58 Atu2454, 34%
sma1760	*S. meliloti* Smb20973, 32%; *S. meliloti* Smc00677, 25%; *S. meliloti* Smc00675, 28%
sma1761	*Rhizobium sp*. NGR234 RNGR00456, 55%
sma1898	*Clostridium acetobutylicum* ATCC 824 Stress-induced protein OsmC, 25%
sma1921	*Corynebacterium jeikeium* peptide methionine sulfoxide reductase MsrA, 29%
sma1954	*Pseudomonas aeruginosa* MexT, 33%
sma2049	*S. meliloti* Smc02022, 36%; *S. meliloti* Smb20817, 28%
sma2199	*S. meliloti* Smc03864, 31%; *S. meliloti* Smb20976, 27%
sma2235	*Bacillus anthracis* str. A2012 Bant_01002356, 22%
sma2239	*M. loti* MAFF303099 Mlr7981, 62%
sma2241	*P. aeruginosa* aromatic-amino-acid aminotransferase TyrB, 26%
smb20156	*A. tumefaciens* str. C58 UgpC, 42%
smb20162	*Paracoccus denitrificans* FlhR 41%; *R. sp*. NGR234 RngR00478 39%; *B. japonicum* USDA 110 putative two-component response regulator Bll0331, 36%
smb20179	*P. denitrificans* PdenDRAFT_2860, 75%
smb20180	*Synechococcus elongatus* PCC6301 blue-copper-protein-like protein Syc1570_d, 26%
smb20243	*P. aeruginosa* PAO1 glycosyltransferase WbpY, 32%
smb20278	*Burkholderia ambifaria* peroxidase-related alkylhydroperoxidase AhpD, 43%
smb20399	*M. loti* MAFF303099 Msr6861, 62%
smb20450	*R. sp.* NGR234 sigma factor SigB regulation protein RsbU, 86%; *B.subtilis* RsbU, 26%
smb20451	*M. loti* MAFF303099 Mll6699, 60%
smb20452	*M. loti* MAFF303099 Mll6698, 63%
smb20463	*S. meliloti* Smc02017, 42%
smb20476	*Bacillus firmus* dipeptide transporter protein DppA, 33%*; E. coli* DppA, 32%
smb20498	*Silicibacter sp.* M1040 deoxyribose-phosphate aldolase RoseDRAFT_3695, 74% ; *A. tumefaciens* C58 aldolase Atu3875, 74%
smb20504	*S. meliloti* Smc02774, 28%; *S. meliloti* Smb20484, 27%
smb20590	*M. loti* MAFF303099 acylphosphatase Mll7735, 58%
smb20591	*Magnetospirillum magnetotacticum* MS-1 FAD/FMN-containing dehydrogenases Magn03005788, 40%
smb20600	*Bartonella quintana* Electron transfer flavoprotein alpha-subunit BruAb2_0445, 29%
smb20724	*S. meliloti* Sma1927, 33%; *S meliloti* Smc00499, 27%
smb20725	*S. meliloti* Smb20770, 29%
smb20861	*R. sp.* NGR234 RNGR00362, 52%
smb20888	*M. loti* MAFF303099 Mll0715, 57%
smb20909	N/A
smb20910	*S. meliloti* Smc02316, 78%
smb21157	*S. meliloti* Smb21229, 70%; *S. meliloti* Smc03108, 63%
smb21158	*M. loti MAFF303099* Mlr4153, 40%

Table 2-5. The examples of homologs of the putative proteins encoded by the new discovered putative symbiotic genes (continues)

Gene or ORF	Examples of the hypothetical protein's homologs in *S. meliloti* or in other bacterial species
smb21202	*A. tumefaciens* C58 Atu2453, 77%; *M. loti* MAFF303099 MesoDRAFT_4275, 75%
smb21259	*Nitrobacter hamburgensis* Fructose-6-phosphate phosphoketolase NhamDRAFT_2943, 26%
smb21333	N/A
smb21397	*S. meliloti* Smb20838, 51%; *P. fluorescens* protease PrtA, 48%
smb21400	*S. meliloti* Sma1797, 64%; *S. meliloti* Sma0450, 63%; *S. meliloti* Smc01186, 50%; *S. meliloti* SMc00284, 50%
smb21401	N/A
smb21402	*S. meliloti* ExpE1, 35%; *S. meliloti* Smb21229, 70%; *S. meliloti* Smb20079, 29%
smb21468	*Caulobacter crescentus* sulfate ABC transporter permease CC1597, 21%
smb21595	*A. tumefaciens* C58 Atu4361, 76%
smc01203R1	*Burkholderia cenocepacia* putative short-chain dehydrogenase Bcen2424DRAFT_5053, 31%;
smc01203R2	*B. pseudomallei* putative non-ribosomal peptide synthase BPSS1197, 35%; *B. japonicum* Blr4533, 32%
smc02230R	N/A
smc02773R1	*Kineococcus radiotolerans* TrkA-C protein SRS30216, 28%; *Chlorobium phaeobacteroides* heat shock protein HslU, 26%; *M. loti* MAFF303099 Mll4969, 32%
smc03181R1	*Xanthomonas axonopodis* XAC3375, 36%
smc03181R2	*Actinobacteria Frankia sp.* EAN1pec hypothetical protein Franean1DRAFT_7093, 33%;
smc03205R1	*Mesorhizobium sp.* BNC1 MesoDRAFT_3624, 25%; *Methanocaldococcus jannaschii* DSM 2661 MJ0044, 31%
smc04135R	*Pseudomonas syringae pv. syringae* B ATP-dependent DNA helicase RecQ, 30%
sma1252R1	*Pseudomonas oleovorans* PhaF protein, 31%
sma1252R2	*S. pyogenes* collagen-like surface protein SclB, 28%
sma1523R	N/A
smb20463S1	*K. radiotolerans* SRS30216's putative short-chain dehydrogenase protein, KradDRAFT_0730, 39%; *Rhodoferax ferrireducens* extracellular solute-binding protein, RferDRAFT_0763, 36%
smb20463S2	N/A
smb20276R	*K. radiotolerans* regulator DeoR protein *KradDRAFT_1350*, 34%; *A. dehalogenans* AdehDRAFT_1169, 39%
smb20714R1	*K. radiotolerans* SRS30216 glycosyl transferase KradDRAFT_2460, 32%; *Pediococcus pentosaceus* ATCC 25745 alpha-acetolactate decarboxylase PpenA01001500, 30%
smb21066R1	*B. subtilis* hypothetical tripeptidase homolog YqjE protein, 31%
smb21573R1	*Mycobacterium leprae* mycocerosic acid synthase MasA, 36%

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
lsrA (smc00037)	Yes	Yes	
dgkA (smc04213)	Yes	Yes	
exsA (smb20941)	Yes	Yes	
flaC (smc03040)	< 1.95 times	Decreased; ≥2 times*	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn*5*) nodule bacteria
fliF (smc03014)	Yes	Yes	Decreased with nodD3 overexpressed; down-regulated by phosphate starvation
degP1 (smc02365)	Yes	Yes	
ppdK (smc00025)	Yes	Yes	
thuF (smb20326)	Yes	Yes	
thuG (smb20327)	Yes	Yes	
thuK (smb20328)	Yes	Yes	
glxD (smc02612)	Increased; ≥ 1.95 times	< 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn*5*) nodule bacteria;
ppe (smb20195)	Yes	Yes	
sma0340 (wrbA2)	Yes	Yes	
smb21662 (exoI2)	Yes	Yes	
smc00234 (ppiD)	Decreased; ≥ 1.95 times	Decreased; ≥ 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium; decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00289 (cspA5)	Decreased; ≥ 1.95 times	Decreased; ≥ 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium
smc00288	Yes	Yes	
smc00643 (purA)	Decreased; ≥ 1.95 times	< 2 times	
smc00644	Decreased; ≥ 1.95 times	< 2 times	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00695 (aroK)	Yes	Yes	
smc00696 (aroB)	Decreased; ≥ 1.95 times	< 2 times	
smc00697	Yes	Yes	
smc00811 (FtsJ)	Yes	Yes	
smc00812	Yes	Yes	
smc01512 (hmuT)	Yes	Yes	Increased in minimal M9 compared to rich TY medium

*** When changes of expression are greater than minimum detection levels, details of the results are presented.**

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
smc01511 (hmuU)	Yes	Yes	
smc01510 (hmuV)	Yes	Yes	
smc02426	Yes	Yes	
smc02427 (hyuE)	Yes	Yes	
smc02653 (lepB)	Increased; ≥ 1.95 times	< 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn*5*) nodule bacteria
smc02737 (opuC)	Decreased; ≥ 1.95 times	Decreased; ≥ 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium
smc03806 (glnK)	Yes	Yes	
smc03807 (amtB)	Yes	Yes	Down-regulated by phosphate starvation
smc03927 (nodN2)	Decreased; ≥ 1.95 times	< 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium
smc03928	Yes	Yes	
smc04047 (azu2)	Yes	Yes	Increased in M9 glucose medium (compared to M9 succinate)
smc04048	Yes	Yes	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc04234 (csp4)	< 1.95 times	Decreased; ≥ 2 times	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00880	Yes	Yes	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00881 (*dgoK1)*	Decreased; ≥ 1.95 times	< 2 times	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00882	Decreased; ≥ 1.95 times	< 2 times	
smc00883	Decreased; ≥ 1.95 times	< 2 times	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc01522	Variable genes (average signal/standard deviation ⩽3)	< 2 times	Increased with dimethyl formamide (DMF)

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
smc01523 (emrE)	Yes	Yes	Decreased in VO2675 (fixJ::Tn*5*) nodule bacteria compared to VO2675 grown in TY medium
smc00072	< 1.95 times	Decreased; ≥ 2 times	Up-regulated by phosphate starvation
smc00371	Yes	Yes	Increased in M9 glucose medium (compared to M9 succinate)
smc00570	Decreased; ≥ 1.95 times	Decreased; ≥ 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium
smc00722	Yes	Yes	Up-regulated by phosphate starvation
smc00856	Yes	Yes	
smc00976	Yes	Yes	
smc00977	Yes	Yes	
smc01214	Yes	Yes	
smc01488	Increased; ≥ 1.95 times	< 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn*5*) nodule bacteria
smc01571	Yes	Yes	
smc01572	Yes	Yes	
smc01577	Yes	Yes	
smc01731	Decreased; ≥ 1.95 times	< 2 times	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
smc01759	Increased; ≥ 1.95 times	< 2 times	Increased in minimal M9 compared to rich TY medium; Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
smc01817	Yes	Yes	
smc01945	Yes	Yes	
smc01957	Increased; ≥ 1.95 times	< 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
smc01988	Yes	Yes	
smc02151	< 1.95 times	Decreased; ≥ 2 times	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium; decreased in bacteroids compared to in VMM medium

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
smc02171	Detected, no conclusive changes	< 2 times	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium; up-regulated by phosphate starvation
smc02381	Yes	Yes	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
smc02616	Yes	Yes	
smc02617	Yes	Yes	
smc02618	Yes	Yes	
smc02619	Yes	Yes	
smc02817	Yes	Yes	
smc02818	Decreased; ≥ 1.95 times	< 2 times	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
smc02987	Yes	Yes	Increased with nodD3 overexpressed
smc02988	Decreased; ≥ 1.95 times	< 2 times	Increased with nodD3 overexpressed; Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
smc03130	Decreased; ≥ 1.95 times	Decreased; ≥ 2 times	Decreased in Rm1021 bacteroids compared to Rm1021 grown in TY medium; Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium; Decreased in microoxic condition compared to oxic condition; Up-regulated by phosphate starvation
smc03746	Yes	Yes	Decreased with nodD3 overexpressed; Down-regulated by phosphate starvation
smc03767	Yes	Yes	
smc03844	Yes	Yes	Increased in minimal M9 compared to rich TY medium
smc03890	< 1.95 times	Increased; ≥ 2 times	

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
smc04054	Increased; ≥ 1.95 times	< 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
smc04194	Increased; ≥ 1.95 times	< 2 times	Increased in minimal M9 compared to rich TY medium; increased in microoxic condition compared to oxic condition; up-regulated by phosphate starvation
smc04311	Yes	Yes	
smc04391	Yes	Yes	
smc04439	Yes	Yes	Increased in minimal M9 compared to rich TY medium
sma0126	< 1.95 times	Decreased; ≥ 2 times	Decreased with luteolin; Down-regulated by phosphate starvation
sma0232	Increased; ≥ 1.95 times	Increased; ≥ 2 times	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
sma0254	< 1.95 times	Decreased; ≥ 2 times	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
sma0447	Yes	Yes	
sma0537	Yes	Yes	
sma0538	Yes	Yes	
sma0945	Yes	Yes	
sma0969	Yes	Yes	
sma1124	Yes	Yes	Up-regulated in micro- oxic conditions compared to oxic conditions
sma1740	Yes	Yes	Increased in minimal M9 compared to rich TY medium
sma1760	Yes	Yes	
sma1761	Yes	Yes	Down-regulated by phosphate starvation
sma1898	Yes	Yes	
sma1921	Increased; ≥ 1.95 times	< 2 times	
sma1954	Yes	Yes	
sma2049	Yes	Yes	
sma2199	< 1.95 times	Decreased; ≥ 2 times	
sma2235	Yes	Yes	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
sma2239	< 1.95 times	Decreased; ≥ 2 times	Decreased in bacteroids compared to in VMM medium
sma2241	Yes	Yes	Decreased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
smb20156	Yes	Yes	
smb20162	Yes	Yes	Increased in M9 glucose medium (compared to M9 succinate)
smb20179	Yes	Yes	
smb20180	Yes	Yes	
smb20243	Yes	Yes	
smb20278	Yes	Yes	
smb20399	Yes	Yes	
smb20450	Yes	Yes	
smb20451	Yes	Yes	
smb20452	Yes	Yes	
smb20463	Yes	Yes	
smb20476	Yes	Yes	Increased in minimal M9 compared to rich TY medium
smb20498	Yes	Yes	
smb20504	Yes	Yes	
smb20590	Yes	Yes	
smb20591	Yes	Yes	
smb20600	< 1.95 times	Decreased; ≥ 2 times	Increased in minimal M9 compared to rich TY medium; increased in microoxic condition compared to oxic condition; up-regulated by phosphate starvation
smb20724	Yes	Yes	Increased in M9 glucose medium (compared to M9 succinate)
smb20725	Yes	Yes	Increased in M9 glucose medium (compared to M9 succinate)
smb20861	Yes	Yes	Down-regulated by phosphate starvation
smb20888	Yes	Yes	Down-regulated by phosphate starvation
smb20909	Yes	Yes	
smb20910	Yes	Yes	Increased in minimal M9 compared to rich TY medium
smb21157	Yes	Yes	
smb21158	Yes	Yes	Increased in Rm1021 bacteroids compared to Rm1021 grown in TY medium

Table 2-6. The expression levels of new discovered putative symbiotic genes as described in previous microarray analysis (continues)

Gene or ORF	Changes of gene expression inside nodules less than minimum detection level (1.95 times) based on Barnett's report (9)	Changes of gene expression inside nodules less than minimum detection level (2 times) based on Becker's report (11)	Changes of gene expression in different conditions based on microarray (9, 11)
smb21202	Yes	Yes	
smb21259	Yes	Yes	
smb21333	Increased; ≥ 1.95 times	< 2 times	
smb21397	Yes	Yes	Increased in Rm1021 bacteroids compared to VO2675 (fixJ::Tn5) nodule bacteria
smb21400	Increased; ≥ 1.95 times	Increased; ≥ 2 times	Increased in VO2675 (fixJ::Tn5) nodule bacteria compared to VO2675 grown in TY medium
smb21401	Increased; ≥ 1.95 times	< 2 times	
smb21402	Yes	Yes	
smb21468	Yes	Yes	
smb21595	Yes	Yes	
smc01203R	N/A	N/A	
smc02230R	N/A	N/A	
smc02773R	N/A	N/A	
smc03181R	N/A	N/A	
smc03205R	N/A	N/A	
smc04135R	N/A	N/A	
sma1252R	N/A	N/A	
sma1523R	N/A	N/A	
smb20463S1	N/A	N/A	
smb20463S2	N/A	N/A	
smb20276R	N/A	N/A	
smb20714R	N/A	N/A	
smb21066R	N/A	N/A	
smb21573R	N/A	N/A	

DISCUSSION

Development of an improved IVET approach for screening *S. meliloti* symbiotic genes

The complexity of *S. meliloti*-alfalfa symbiosis suggests that more *S. meliloti* genes could be involved in the symbiosis. Some of the symbiotic genes may function as a single essential factor, which could have already been identified through saturated loss-of-function mutagenesis. Some of the symbiotic genes could be activated and remain highly expressed throughout the symbiosis. Those genes should be identified in microarray and proteomic analyses, which have identified at least 350 symbiotic genes that are highly expressed in bacteroids (9). Some of the symbiotic genes, however, may have multiple functional homologs and may be expressed only very briefly during the early stages of the symbiosis. It would be difficult to collect enough materials to screen for these genes by microarray or proteomic analysis in a genome-wide screen.

To carry out a thorough genome-wide search for genes that are expressed at the early stages of the symbiosis, I have improved the *In Vivo* Expression Technology (IVET) (142). The key modification I made was to monitor the activity of trapped promoters in free-living cells by the expression of the *gfp* gene, which was added after the reporter gene *exoY*. The use of *gfp* made it possible to prescreen the library to clear out the non-symbiotic promoters (see below). This modification also ensures a more effective screen of a large size library utilizing FACS.

The use of shorter genomic DNA fragments in the construction of the promoter library dramatically increased the size of the promoter library. Due to the low transformation efficiency of *S. meliloti* cells, the promoter library was constructed in *E. coli* cells first and then moved into *S. meliloti* cells by conjugation. The promoter library was checked for its size, the rate of duplication, and the promoter activities of individual clones in the library to ensure that the promoter library had a 95% coverage of *S. meliloti* genome.

The large size of the promoter library increased the number of the clones carrying *S. meliloti* promoters of housekeeping genes that would interfere with the screening for symbiotic promoters. These promoters had to be removed from the promoter library before screening. Anticipating this difficult task, I have included the *gfp* gene in the same transcript following the *exoY* gene in the promoter trap vector to monitor the activities of trapped promoters. Housekeeping promoters will be cleared out based on the level of GFP fluorescence before a functional screen is carried out for promoters that are activated at the early stages of the symbiosis. This is different from the method of Differential Fluorescence Induction (DFI) which uses the expression of the *gfp* gene alone to select for promoters activated under conditions of interest (174).

To remove the promoters of housekeeping genes, a particular level of gene expression has to be determined. It also has to be taken into consideration that some of the symbiotic genes might express at low levels in free-living cells and at high levels during symbiosis. The *exoY* gene is expressed at low levels in free-living cells, which results in low levels of succinoglycan production (47). Identifying promoters that have low activities in free-

living cells, such as the *exoY* promoter, could also lead to the isolation of symbiotic promoters that are up-regulated during symbiosis. Therefore, the level of gene expression under the *exoY* native promoter was used as a standard to sort library into dark, dim and bright sublibraries (although only the dark library was used in the later screen).

A large number of nodules were screened in this study. Because bacterial invasion to plant root hair cell is started typically by one or only a few bacterial cells trapped in one curved root hair, bacterial cells in each root nodule are derived from one or few bacterial cells. Thus, a large number of nodules have to be screened to cover all of the trapped *S. meliloti* promoters in the dark sublibrary. I developed a hydroponic system to incubate a large amount of alfalfa seedlings. A total of 22,000 alfalfa root nodules were screened, which is five times the minimum requirement to cover the dark sublibrary. Analysis of fusions carried by the 208 pink nodule isolates led to the discovery of 113 mostly previously unknown symbiotic genes and putative symbiotic genes (Table 2-3). This suggests the improved IVET approach is effective in identifying symbiotic genes involved in the *S. meliloti*-alfalfa symbiosis.

Comparison of *S. meliloti* genes identified using this improved IVET approach with previous microarray analysis

A total of 113 different putative *S. meliloti* promoter fusions were identified through this improved IVET approach (Table 2-5). Surprisingly, none of the known *nod* gene was among them. One explanation could be that *nod* genes are expressed too early to express the *exoY* gene at the right time.

Among these identified genes or putative genes, *glxD*, *lepB*, *smc01488*, *smc01759*, *smc01957*, *smc04054*, *smc04194, sma0232*, *sma1921*, *smb21333*, *smb21400* and *smb21401* were described as genes showing elevated expression levels in bacteroids inside root nodules compared to that in free-living cells in Barnett's microarray analysis (9), while *smc03890, sma0232* and *smb21400* were described as genes showing elevated expression levels in bacteroids in Becker's microarray analysis (11). However, little is known about their roles at different stages of the *S. meliloti*-alfalfa symbiosis. Among these identified genes or putative genes, *ppiD, cspA5, smc00063 (purA), smc00064, smc00695 (aroB), smc02737 (opuC), smc03927 (nodN2), smc00881 (dgoK1), smc00882, smc00883, smc00570, smc01731, smc02818, smc02988, smc03130* and *smc00644* were described as genes showing decreased expression levels in bacteroids inside root nodules compared to that in free-living cells in Barnett's microarray analysis (9), while *flaC*, *smc04234 (csp4), smc00072, smc01251, sma0126, sma0254, sma2199, sma2239, smb20600, ppiD, cspA5, smc02737 (opuC), smc00570* and *smc03130* were described as genes showing decreased expression levels in bacteroids inside root nodules in Becker's microarray analysis (11). The rest of the genes were not described in Barnett's and Becker's microarray analysis because the differences between gene expression levels in bacteroids and in free-living cells were less than two fold (9, 11).

The bacterial genes identified by Barnett's or Becker's microarray analyses are the genes induced or repressed in the late stages of symbiosis establishment, since total RNA prepared for microarray was extracted from mature nodules. The genes that are identified

using this improved IVET approach are the genes induced in the early stages of symbiosis establishment, because the *exoY* gene was used as the reporter gene for identifying promoters that are activated during infection thread initiation and elongation. These genes could be down-regulated after bacterial cells are released into the cytoplasm of nodules cells and differentiate into bacteroids. This also could account for the fact that most genes identified in my study were not described by the previously published microarray analyses.

Possible roles of the identified genes during the symbiosis establishment

Analysis of the identified putative genes based on previous reports and homology prediction showed that these putative genes might play roles in different aspects of *S. meliloti*-alfalfa plant interactions, including the bacterial motility and chemotaxis, stress response and adaptation, transcription regulation, nutrient uptake, central intracellular metabolism and unknown functions.

The identification of two *S. meliloti* flagellum synthesis genes, *flaC* and *fliF*, might increase our understanding of the regulation of flagellum production and motility in the *S. meliloti*-alfalfa symbiosis. The homology of *S. meliloti flaC* gene in *A. tumefaciens* was reported to be regulated in response to host plant stimuli (49), while the homolog of the same gene in *Vibrio parahaemolyticus* was reported to be controlled in a surface-dependent manner (119). The *fliF* gene was previous identified as an infection-inducible gene during colonization of sugar beet (*Beta vulgaris altissima*) in rhizosphere by *Pseudomonas fluorescens* (85) and during *Streptococcus suit* infection of piglets (201)

using similar positive screening strategies. These findings raise the possibility that temporal up-regulation of flagellum synthesis is related to flagellum-driven chemotaxis which attracts bacterial cells towards host plant surfaces at the very early stages of symbiosis establishment.

Some *S. meliloti* genes identified in this study are involved in adaptation to environmental stresses. Among them, the *opuC* gene encodes a glycine betaine transporter and the putative *smc04439* gene encodes a glycine betaine transport ATP-binding protein. Glycine betaine was reported to relate to osmotic stress tolerance of *S. meliloti* and many other members of the family Rhizobiaceae (22, 32, 210). The identification of the two genes suggests they might play roles in bacterial adaptation to the physical environment in plant cells. The *dgkA* gene, whose gene product is involved in cyclic β-1,2-glucan synthesis, was identified. The presence of β-1,2-glucan is essential for the establishment of the *S. meliloti*-alfalfa symbiosis (204). The cyclic β-glucans of *Rhizobium leguminosarum* were reported to be required for the bacterial infection of plant roots as suppressor to host defense response (174). My further research showed that the *dgkA* null mutant remains symbiotically active, suggesting that the *dgkA* gene is not a symbiotically essential gene. It is possible that the function of the DgkA protein can be provided by other proteins in bacterial cells. Another identified putative *smc00072* gene is predicted to encode a peroxiredoxin protein which might be involved in *S. meliloti* defense against oxidative stress under microaerobic conditions in bacteroids. A putative transmembrane multi-drug transporter gene *emr* was identified in this research, suggesting that some bacterial detoxification system might be upregulated in interactions

between *S. meliloti* and alfalfa. This is consistent with the previous report that the multi-drug efflux pump gene *acrF* was identified as an infection-inducible gene during *P. fluorescens* infection of its host plant (85).

More than 10 *S. meliloti* transcription regulators were identified in this functional screening, which belong to different families, such as LysR-, LuxR-, LacI-, cold-shock protein and two-component regulatory system, or even unclassified. Such a finding indicates that in responding to environmental challenges and perhaps to plant signals, *S. meliloti* cells will perform complex regulation at the early stages of the symbiosis establishment. Confirmation of the symbiosis function of these transcription factors by constructing null mutations and seeking their target genes will further reveal interactions between *S. meliloti* and alfalfa at specific stages of the symbiosis establishment.

Some *S. meliloti* genes identified are involved in different nutrient uptake systems. The identification of *S. meliloti* hemin transporter system *hmuTUV* genes suggests that bacterial cells might need to regulate iron transport system following their environmental changes from rhizosphere to infection thread. This is consistent with previous report of upregulation of *hmuT* and *hmuU* genes in plant pathogen *Erwinia chrysanthemi* during infection of spinach plant (227). Since iron is essential for microbial survival, the plant-associated bacteria may have to regulate iron uptake accurately in different environments. The identification of *S. meliloti* trehalose/maltose uptake system *thuFGK* genes is consistent with the previous report that uptake of disaccharides is critical for *S. meliloti* colonization of alfalfa roots (114). The putative sugar-uptake transporter genes *smb20504*

and *smb21596* were also identified. In addition, four ABC transporter genes, *sma2199, smb20156, smb20476* and *smc02171*, which are involved in transporting unknown substrates were also in the list. Further characterization of these putative genes and identification of their substrates may reveal new symbiotic phenotypes at certain early stages of the symbiosis establishment.

Some *S. meliloti* genes involved in various metabolic processes were identified. The identified *glnK* and *amtB* genes are involved in nitrogen metabolism. The *ppdK* gene is involved in gluconeogenesis. The *glxD*, *hyuE*, *aroB* and *aroK* genes are involved in amino acid biosynthesis. The genes *azu2* and *smc04048* encoding a copper-binding protein and a cytochrome C protein, respectively, are involved in electron transport. The genes *smc00976* and *smc00977* are predicted to be involved in fatty acid degradation. The *ppe* gene product is predicted to be involved in Calvin cycle. The *sma1761, smc00881, smc00882, smb20861, smc04391* are predicted to be involved in small molecule metabolism. The *purA* gene in synthesis of purine ribonucleotide, the *ppiD* gene involved in protein folding, and the *degP1* gene involved in protein degradation were also identified. All these findings suggest profound changes in intracellular metabolism of *S. meliloti* cells during early symbiosis establishment with alfalfa plant.

Interestingly, more than 10 putative genes which were identified based on trapped symbiotically active promoters overlap the previously identified genes or the putative open reading frames (defined by the *S. meliloti* Rm1021 Genome Project) in a reversed direction. There were similar reports by different labs when using IVET technology to

identify host-induced genes in different animal- and plant- pathogen bacteria (37, 172, 174, 221). It is unclear whether the putative genes encode functional proteins which may play roles during symbiosis. But it is possible that the trapped promoters might promote synthesis of small noncoding RNA molecules, which may down-regulate overlapped gene expression (174, 200). However, little is revealed about the antisense transcript concept in these bacteria.

In summary, a large group of previous unknown symbiotic genes has been identified through the use of the improved IVET approach, which suggests that more *S. meliloti* genes could be involved in the early stages of the symbiosis establishment. Further characterization of these symbiotic genes will lead to a better understanding of the roles of *S. meliloti* in the symbiosis.

Chapter III

Identification of Root-Exudate Induced *Sinorhizobium meliloti* Genes

INTRODUCTION

The signal exchanges between *S. meliloti* and its host plant at specific stages are essential for the successful symbiosis establishment. The best known class of plant symbiotic signals required for early symbiosis establishment is flavonoids, which induce the expression of bacterial nodulation (*nod*) genes (35, 76, 166). Flavonoids are polycyclic aromatic compounds, 2—phenyl-1,4-benzopyrone derivatives. They are released by plant roots into rhizosphere and could be accumulated at high levels. Two alfalfa-produced flavonoids, apigenin and luteolin (Figure 3-1), can induce *S. meliloti* Rm1021 *nod* gene expression (16, 78, 166, 167). Flavonoids are detected by bacterial NodD proteins, which bind to *nod* box in the *nod* gene promoters and activate the *nod* gene expression (76, 77, 121). The expression of *nod* genes leads to the synthesis of Nod factors, which in turn induce specific responses in the host plant, including deformation and curling of root hairs, and root cortical cell division (137, 153).

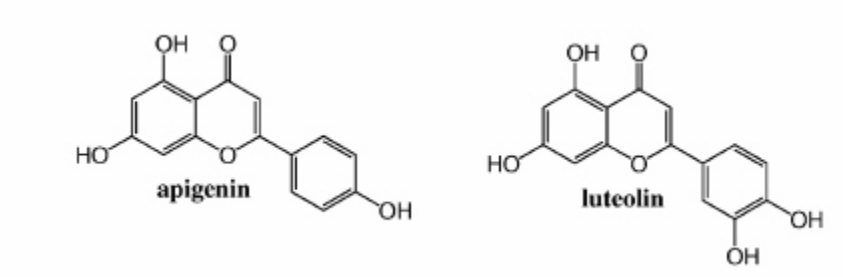

Figure 3-1. Chemical composition of two main flavonoids, apigenin and luteolin, which induce *S. meliloti nod* gene expression (168).

In addition to flavonoids, plant roots also secrete high levels of other organic compounds which could account for 20% of all fixed carbon in the rhizosphere (29, 146). They

include low-molecular-weight compounds such as amino acids, sugars, organic acids, aromatics, vitamins, purines, nucleosides, and high-molecular-weight compounds such as polysaccharides and proteins (29, 146). Some of these compounds could serve as symbiotic signals.

We have identified a set of 113 mostly previously unknown *S. meliloti* symbiotic genes functioning at the early stages of the symbiosis establishment using the improved IVET approach. These genes are predicted to perform different functions and none of them belong to previously known *nod* genes. Some of these genes could be involved in signal exchanges with alfalfa, and the expression of these genes might be activated by plant signals.

In this study, the symbiotic promoter activities of nodule isolates were evaluated by measuring the specific GFP fluorescence intensity of each strain in the presence and absence of alfalfa root exudates. These findings provide further evidence for the complexity of the symbiosis and open new avenues to further analysis of the early interactions between the bacterium and its host.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

The bacterial strains and plasmids used in this project were listed in Table 3-1.

Table 3-1. Strains and plasmids

Strain or plasmid	Relevant properties	References
E. coli		
DH5α	General purposed strain	(103)
MT616	MT607, pRK600, Cm^R	(74)
S. meliloti		
Rm1021	Wild type, Sm^R	(126)
RmAR9007	Rm1021 *exoY-LacZ*, Gm^R	(117)
Rm7095	Rm1021 *exoR95*::Tn*5*, Nm^R	(64)
Rm7210	Rm1021 *exoY210*::Tn*5*, Nm^R	(126)
SmHC10	Rm1021, *dgkAΩ*135, Sm^R, Nm^R	This work
SmHC11	Rm1021, *smb20195Ω24*, Sm^R, Nm^R	This work
SmHC12	Rm1021, *smc02171Ω151*, Sm^R, Nm^R	This work
SmHC13	Rm1021, *smc02773Ω1453*, Sm^R, Nm^R	This work
SmHC14	Rm1021, *smc03205Ω10*, Sm^R, Nm^R	This work
Plasmids		
pK19mob2omegaHMB	Suicide vector, Km^R	(140)
pRK600	Helper plasmid, Cm^R	(74)
pHC60	Constitutively expressing GFP, Tc^R	(46)
pHC93	Expression vector, pBBR1MCS3, Tc^R	(122)
pHC172	pMB393, carrying *PnodABC-gfp* fusion, Sp^R	This lab
pHC183	Suicide plasmid carrying *smb20195* homology region, Km^R	This work
pHC184	Suicide plasmid carrying *smc02773* homology region, Km^R	This work
pHC185	Suicide plasmid carrying *dgkA* homology region, Km^R	This work
pHC186	Suicide plasmid carrying *smc02171* homology region, Km^R	This work
pHC199	Suicide plasmid carrying *smc03205* homology region, Km^R	This work
Phage		
ϕM12	Generalized transducing phage for *S. meliloti*	(73)

Escherichia coli strains used in this study were grown in Luria-Bertani (LB) medium at 37°C, *S. meliloti* strains were grown in LB medium supplemented with 2.5mM $MgSO_4$ and 2.5mM $CaCl_2$ (LB/MC) at 30°C. Agar (1.5%) was used to solidify media. To examine succinoglycan production, calcofluor white M2R (Fluorescent Brightener 28, Sigma) was added to a final concentration of 0.02% in LB/MC agar, which was buffered

to pH 7.4 with 10mM HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid) (93, 94, 126). To examine cell motility, *S. meliloti* was spotted on swarming plate, which was made of 1/10 strength LB/MC media with 0.3% agar (223). The following antibiotics were used at the concentrations indicated: Ampicillin, 100μg/μl; chloramphenicol, 10μg/ml; gentamycin, 50μg/ml; kanamycin, 25μg/ml; neomycin, 200μg/ml; spectinomycin, 100μg/ml; streptomycin 500μg/ml; and tetracycline, 10μg/ml.

Plant root exudates preparation

Alfalfa seedlings were grown in hydroponic system described in Chapter II to prepare exudates. Briefly, about 100 alfalfa seeds were surface-sterilized, and then spread on wet cheesecloth suspended with aluminum screen in a 11×11×9 mm^3 container with 200 ml nitrogen-free Jensen liquid medium. Two ends of the cheesecloth were submerged in liquid medium to ensure the cheesecloth on the aluminum screen remain wet. The containers with seeds were kept in dark at 26°C for 40 hours for seed germination, and then exposed to light. The liquid medium in each container was inoculated with 1 ml *S. meliloti* Rm1021 bacterial suspension (OD_{600} 0.3). On day 5 after inoculation, 100 ml liquid Jensen medium with bacterial cells and root exudates was removed from each container. Bacterial cells and insoluble materials were removed from the liquid medium by centrifugation (13,000rpm, 15 min, 4°C). The supernatant was further purified by passing through a sterilization filter with pore size of 0.22μm and used as crude root exudates.

Determination of specific GFP fluorescence intensity of bacterial cultures

To determine the specific GFP fluorescence intensity of a given bacterial culture, cell suspensions were transferred to wells of a transparent 96-well plate and a black 96-well plate in equal amounts. The cell suspensions in transparent 96-well plate were used to determine the cell density (OD_{600}) using an absorbance microplate reader (Spectra Max 340PC, Molecular Device, Sunnyvale, CA, USA), and the cell suspensions in black 96-well plate were used to determine the intensities of GFP fluorescence using a fluorescence microplate reader (Spectra Max Germini XS, Molecular Device, Sunnyvale, CA, USA). The GFP fluorescence intensity of each of the cell suspension was normalized to its cell density to generate the specific GFP fluorescence intensity of bacterial cultures.

Root exudates induction assay

To determine the effect of root exudates on gene expression, bacterial cells were collected from fresh LB/MC culture, washed with sterile distilled water, and resuspended in alfalfa root exudates to a concentration of an OD_{600} of 0.1. After 18-hour incubation at 30°C, specific GFP fluorescence intensity of bacterial cells in the root exudates was determined. Bacterial cells were also incubated in the supernatant of liquid Jensen medium and in the supernatant of liquid Jensen medium with alfalfa flavonoid apigenin (2.5μg/ml).

Targeted mutagenesis of *S. meliloti* genes

The *S. meliloti dgkA* gene and putative *smb20195, smc02171, smc02773R,* and *smc03205R* genes were mutagenized by plasmid insertion. A suicide plasmid was

constructed specifically by cloning a 0.3-1 Kb fragment from the middle of the targeted gene into vector pK19mob2ΩHMB (140). The inserted DNA fragment will mediate the integration of the suicide plasmid into genome through homologous recombination (Figure 3-2).

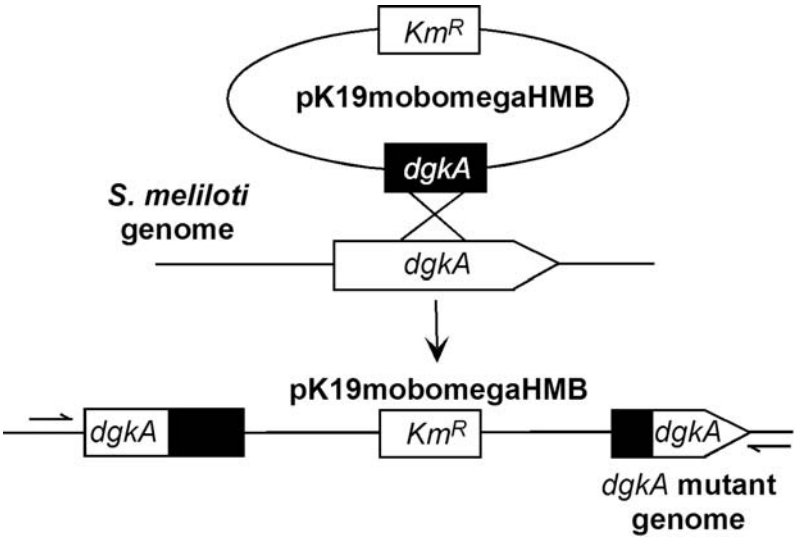

Figure 3-2. Schematic representation of the location of the genomic and plasmid-borne genes before and after integration of the plasmid into the genome.

To mutagenize the *dgkA* gene, a DNA fragment containing +135bp to +362bp region of the *dgkA* open reading frame was obtained by PCR using *S. meliloti* Rm1021 genomic DNA as template and the pair of primers: *dgkA*f135*Hind*III and *dgkA*r362*Bsr*GI (Table 3-2). The PCR fragment was digested with restriction enzymes *Hind*III and *BsrG*I, and cloned into the unique *Hind*III/*BsrG*I restriction site of suicide plasmid vector pK19mob2ΩHMB, resulting in plasmid pHC185 (Figure 3-2 and Figure 3-3A). Plasmid pHC185 was electroporated in *E. coli* DH5α, moved into *S. meliloti* Rm1021 using helper strain *E. coli* MT616 (pRK600) in triparental mating. Single colonies, formed by

cells carrying the mutagenized *dgkA* gene, were selected on LB/MC agar plate containing streptomycin and neomycin. To ensure the mutant has only one single insertion of the suicide plasmid in the genome, the inserted plasmid was transduced into wild-type strain to reconstruct the mutant using phage ϕM12. To confirm that the *dgkA* open reading frame is disrupted by the plasmid insertion, PCR was performed using genome of the *dgkA* mutant as template and the pair of primers *dgkA*f*Nde*I and *dgkA*r*Kpn*I (Table 3-2). The PCR cycles were performed as 3 min at 95°C, 30 sec at 56°C, and 1 min at 72°C for 35 cycles. As a control, PCR was performed using the genome of Rm1021 with the complete *dgkA* open reading frame as template in the same conditions.

Table 3-2. Oligo DNA primers used to construct and confirm mutations.

Name	Sequence of primers
Oligo DNA primers used to construct mutations	
*dgkA*f135*Hind*III	5'-CCCAAGCTTCAGGCACGAACTCATC-3'
*dgkA*r362*Bsr*GI	5'-GGCTGTACAGGCGAGAATGAGGCAA-3'
*smc02171*f151*Hind*III	5'-CCCAAGCTTGCCAAGGCCAAGGAACTG-3'
*smc02171*r548*Bsr*GI	5'-GGCTGTACACCGAAACCGATCATGAAG-3'
*smc02773*Rf1453*Hind*III	5'-CCCAAGCTTCCAGTTCATAGACCTCCG-3'
*smc02773*Rr377*Bsr*GI	5'-GGCTGTACATCCTGGGCGACGAGATGAA-3'
*smc03205*Rf10*Hind*III	5'-CCCAAGCTTTATGTGCTGACCGTCAGC-3'
*smc03205*Rr425*Bsr*GI	5'-GGCTGTACAGTCACCTTGATGTGGTGG-3'
*smb20195*Rf24*Hind*III	5'-CCCAAGCTTGCCTTCGGTTCTGTCGGC-3'
*smb20195*Rr339*Bsr*GI	5'-GGCTGTACACGCCTTGCAGCCGACGTC-3'
Oligo DNA primers used to confirm mutations	
*dgkA*fNdeI	5'-TTCCATATGCGGGGCGGCTTCTC-3'
*dgkA*rKpnI	5'-GGGGTACCTATCAGCCCAGGAAGCGAGCGA-3'
p*smc02171*f	5'-CTTGATTGACCATCGGAAGCG-3'
p*smc02171*r	5'-CCTTGTCAAAGACCTCGGCACT-3'
p*smc02773*Rf	5'-TCGCCTTCAGGAGAACAGTGCA-3'
p*smc02773*Rr	5'-CGCTTCCTCCCTGCGTCCTT-3'
p*smc03205*Rf	5'-CGTGGGAACACCTCAAAACGAGA-3'
p*smc03205*Rr	5'-CGGCGCTGTAATGCCAATGTT-3'
p*smb20195*Rf	5'-CGCCACGGTCGAACCCAGTCA-3'
p*smb20195*Rr	5'-CGGGCCAGCCGAAATCACAAA-3'

Four other putative symbiotic genes, *smc02171, smc02773R, smc03205R* and *smb20195,* have been similarly mutagenized (Table 3-2) (Figure 3-3 B to E). To mutagenize *smc02171,* suicide plasmid pHC186 was constructed, which carries the homologous DNA fragment of +151bp to +548bp region of the *smc02171* open reading frame obtained by PCR using the pair of primers *smc02171*f151*Hind*III and *smc02171*r548*Bsr*GI. To mutagenize *smc02773R,* suicide plasmid pHC184 was constructed, which carries the homologous DNA fragment of +377bp to +1453bp region of the *smc02773* open reading frame obtained by PCR using the pair of primers *smc02773R*f1453*Hind*III and *smc02773R*r377*Bsr*GI. To mutagenize *smc03205R,* a suicide plasmid pHC199 was constructed, which carries the homologous DNA fragment of +10bp to +425bp region of *smc03205* open reading frame obtained by PCR using the pair of primers *smc03205R*f10*Hind*III and *smc03205R*r425*Bsr*GI. To mutagenize *smb20195,* suicide plasmid pHC183 was constructed, which carries the homologous DNA fragment of +24bp to +339bp region of *smb20195* open reading frame obtained by PCR using primers *smb20195R*f24*Hind*III and *smb20195R*r339*Bsr*GI. PCR was performed to confirm that the targeted gene is disrupted by the plasmid insertion in each mutant.

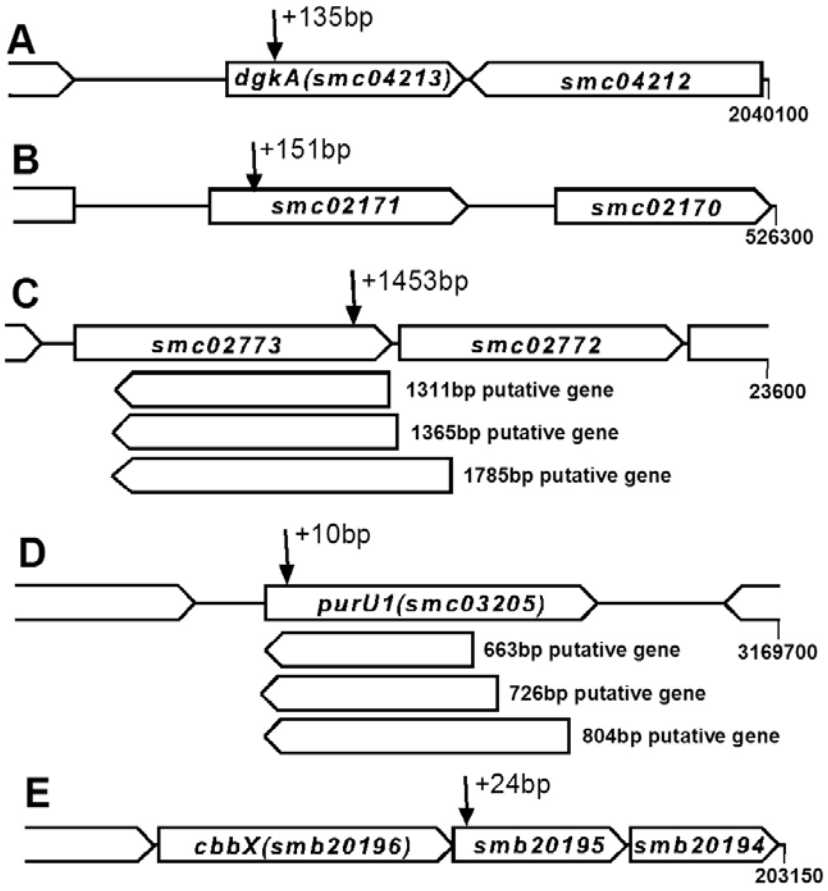

Figure 3-3. A diagram of the constructed mutations showing the suicide plasmid insertion position in the open reading frame(s) of *dgkA* (A), *smc02171* (B), *smc03205* and *smc03205*Rs (C), *smc02773* and *smc02773*Rs (D), and *smb20195* (E), respectively. The number next the arrow represents the position of the nucleotide of the open reading frame, after which the plasmid was inserted.

Plant nodulation assays

Alfalfa plants growing on Petri dishes were used to determine the ability of individual bacterial strains to establish symbiosis with alfalfa as previously described (126). Alfalfa seeds were surface-sterilized in 50% Clorox bleach and washed with sterile distilled water. Seeds were allowed to germinate on 0.7% agar plate in dark at 26°C for 40 hours. The germinated seedlings were transferred to 9×9 mm^2 square Petri dishes containing Jensen solid media with 8 seedlings per plate and inoculated with 1 ml *S. meliloti* cell suspension, which was collected from fresh bacterial culture in LB/MC medium, washed

and diluted with sterile distilled water to an OD_{600} of 0.03. The inoculated alfalfa seedlings were incubated with light at 26°C for four weeks.

Determination of the efficiency of invasion of root hairs

Alfalfa seeds were surface-sterilized and germinated on 0.7% agar plate in the dark at 26°C for 40 hours. Two seedlings 1 to 2 cm long were placed on top of microscope slides, which were covered with a layer of 2 ml of Jensen solid media in advance. The bacterial suspension of each *S. meliloti* strain carrying plasmid pHC60 was collected from fresh bacterial culture in LB/MC liquid medium, washed and diluted with sterile distilled water. Total 0.1 ml bacterial suspension (OD_{600} 0.03) was spread evenly on the roots of seedlings on each slide. Root region of seedlings was covered by a piece of dialysis membrane (12,000 to 14,000 molecular weight cutoff). The microscope slides with the seedlings were placed in 50 ml culture tubes containing 5 ml liquid Jensen medium, covered loosely with Saran Wrap, and incubated with light at 26°C for two weeks.

To evaluate root hair invasion efficiency, the alfalfa seedlings on the slides were examined for the number of colonized curled root hairs, initiated infection threads, extended infection threads and GFP-bright or GFP-dark nodules from Day 1 to Day 12 under Zeiss fluorescence microscope, M^2 Bio Quad System at x100 magnification. Images were obtained by using a filter set consisting of an exciter band pass filter (470 nm +/- 40 nm), a 495 nm long pass dichroic mirror, and an emission band pass filter

(525nm +/- 50nm). Images were captured by software provided by Optronics using digital camera: Magnafire Model S99802 by Optronics, Goleta, CA installed on PC.

Determination of the motility of bacterial cells

Bacterial motility was examined using soft agar plates as previously described (223). Briefly, the fresh bacterial cultures (2µl) were spotted on swarming plate, which was made of 1/10 strength LB/MC medium with 0.3% agar. The plate was sealed with Parafilm and incubated for 2 days at 30°C to determine colony size. Wild-type strain Rm1021 and slow-swarming strain Rm7095 were loaded on the same plate as positive and negative controls, respectively.

RESULTS

Symbiotic genes activated by root exudates

The expression of the 113 mostly previously unknown *S. meliloti* genes that function specifically in the early stages of symbiosis could be activated by plant signals, which would most likely come from alfalfa root exudates. The specific GFP fluorescence intensities of *S. meliloti* RmAR9007 strains expressing the *exoY* and *gfp* genes from different promoters were determined to measure the activities of the trapped promoters. As a positive control, the presence of root exudates caused a 116% increase in the specific GFP fluorescence intensity of the cells that express GFP from the *nodA* gene promoter (Figure 3-4 A). The levels of the specific GFP fluorescence intensities of the other two negative control strains carrying constitutive *smb21651* promoter and non-promoter region, respectively, showed no change in the presence of alfalfa root exudates (Figure 3-4 C and D). These finding suggest that the specific GFP fluorescence intensity could be used to identify root-exudates induced genes. The specific GFP fluorescence intensities of strains carrying the 113 different promoter fusions were evaluated in the presence and absence of alfalfa root exudates. The results showed that the activities of only 23 out of the 113 trapped promoters were elevated more than 30% in the presence of alfalfa root exudates. The activity of the *exoY* native promoter did not change significantly in the presence of alfalfa root exudates (Figure 3-4 B), which suggests *exoY* expression is not induced by root exudates.

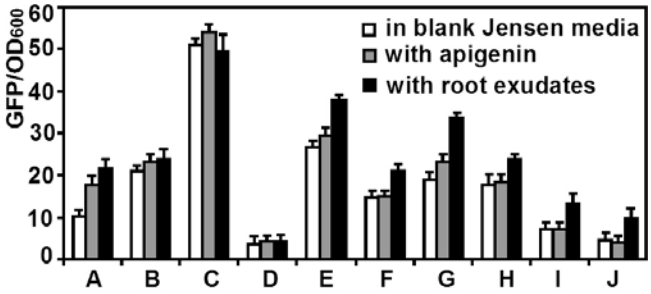

Figure 3-4. A diagram of the activities of the symbiosis promoters isolated from nodule isolates in the absence or presence of alfalfa root exudates or apigenin. A: *Pnod* activities, B: *PexoY* activities, C: *Psmb21651*, D: non-promoter region. E-J: The nodule isolates, carrying putative promoter fusions of *lsrA*, *dgkA*, *smb20195 (ppiD)*, *smc02171*, *smc02773R* and *smc03205R* genes respectively, were induced by root exudates but not by apigenin.

Table 3-3. The values of the activities of the symbiosis promoters isolated from nodule isolates in the absence or presence of alfalfa root exudates or apigenin and the T tests.

	Promoter activity in Jensen (GFP/OD600 J0)	Promoter activity in Jensen with apigenin (GFP/OD600 Ja)	Promoter activity in root exudates (GFP/OD600 Jr)	T value of (GFP/OD600 Ja-GFP/OD600 J0)	T value of (GFP/OD600 Jr-GFP/OD600 J0)
p*nod*	10.052±1.237	17.709±1.725	21.728±1.732	5.102	7.758
p*exoY*	21.060±0.864	22.963±1.566	23.642±2.244	1.505	1.518
p+	51.091±1.258	54.099±1.264	49.413±3.666	2.386	-0.612
p-	3.657±0.482	4.137±0.537	4.175±0.753	0.942	0.820
p*lsrA*	26.625±1.362	29.530±1.267	37.669±1.078	2.209	8.991
p*dgkA*	14.558±1.236	14.875±0.876	21.038±1.226	0.296	5.265
p*cbb-ppe*	18.770±1.632	23.218±1.483	33.573±1.010	2.852	10.906
p*smc02171*	17.715±2.228	18.361±1.313	23.773±0.641	0.353	3.696
p*smc02773R*	7.118±1.236	7.084±1.240	13.165±1.988	-0.027	3.654
p*purU1R*	4.576±1.283	3.825±0.526	9.744±2.063	-0.765	3.009

Symbiotic genes inducible by apigenin

To determine whether the root-exudate-activated 23 promoters are induced by apigenin, which is a known inducer for the *nod* gene expression, the specific GFP fluorescence intensities of the strains carrying the 23 trapped promoter fusions were further evaluated in the presence or absence of apigenin. The presence of apigenin (2.5μg/ml) resulted in a 76% increase in the specific GFP fluorescence intensity of the cells that express GFP from *nodA* promoter, but no increase for cells that carry constitutive *smb21651* promoter or non-promoter region in their promoter trap plasmid (Figure 3-4). The specific GFP

fluorescence intensity of cells carrying the *exoY* native promoter fusion did not change significantly in the presence of apigenin (Figure 3-4). The expression of 17 trapped promoters out of the 23 were activated by apigenin (data not shown), while the activities of the other 6 did not significantly change in the presence of apigenin (Figure 3-4).

Among the 17 genes whose promoters are inducible by both of apigenin and plant root exudates, 5 of them, *lepB, glxD, ppiD*, *aroK* and *hmuT,* were named based on their homologies to previously characterized genes in other bacteria, but little is known about their roles in the *S. meliloti*-alfalfa symbiosis. All five genes appear to be unrelated and located at different loci of the chromosome. In a previous microarray analysis (9), the *lepB* and *glxD* gene have been found to have elevated levels of expression in bacteroids inside root nodules, while the *ppiD* gene was expressed at lower levels in bacteroids. The other two, *aroK* and *hmuT*, were not described in that microarray study. For the remaining 12 putative genes, the expression of 3 putative genes, *smb21333, smb21400,* and *smc04194,* was reported to be increased in bacteroids, while the expression of *smc03130* was decreased in bacteroids (9).

Symbiotic genes uninducible by apigenin

The specific GFP fluorescence intensities of cells carrying the putative promoters of genes *lsrA* and *dgkA* and putative genes *smb20195 (ppe), smc02171, smc02773R* and *smc03205R* increased by 41%, 48%, 78%, 34%, 84% and 112%, respectively, in the presence of alfalfa root exudates. However, none of them showed significant difference when induced with apigenin alone (Figure 3-4), or luteolin (data not shown). Among the

6 root-exudates inducible symbiotic genes, only *lsrA* and *dgkA* have been previously characterized. The *lsrA* gene is predicted to encode a LysR type transcriptional regulatory protein and has recently been identified as a symbiotic gene in a loss-of-function mutagenesis screen by our lab (140) (Figure 3-3). My new finding here that the expression of the *lsrA* gene is activated by root exudates is consistent with the fact that the presence of this gene is essential for the early stages of nodule development. The *dgkA* gene encodes a diacylglycerol kinase required for recycling L-diacylglycerol generated during cyclic β-1,2-glucan synthesis (150). The presence of cyclic β-1,2-glucan is required for the successful nodulation of alfalfa by *S. meliloti* (27). However, it is not clear whether the *dgkA* gene is essential for the symbiosis. The putative *smb20195 (ppe)* gene is predicted to encode a pentose phosphate epimerase. *smc02171* is predicted to encode a putative periplasmic binding protein of an ABC transporter, which is typically involved in the transport of small molecules. *smc02773R* could be any of the three putative open reading frames, *smc02773R1*, *smc02773R2 and smc02773R3*, overlapping the putative open reading frame *smc02773* in opposite direction. Similarly, *smc03205R* is any of the three putative open reading frame, *smc03205R1*, *smc03205R2* and *smc03205R3,* overlapping gene *purU1* (*smc03205*) in opposite direction. Although homologs of *smc02773R* and *smc03205R* are found in other bacteria, the functions of the two putative genes cannot be predicted.

Mutagenesis of apigenin uninducible genes

The five apigenin uninducible symbiotic genes, *dgkA, smb20195 (ppe), smc02171, smc02773R* and *smc03205R*, were mutagenized individually using plasmid insertion.

Each plasmid insertion in genome of the mutants was first transduced into *S. meliloti* Rm1021 (wild-type) background to ensure there is only one mutation in the genome. The insertions in the targeted genes were confirmed by PCR as described in detail in the method section.

The constructed mutant strains were named as SmHC10, SmHC11, SmHC12, SmHC13, and SmHC14, which carry plasmid-insertion mutations in the *dgkA, smb20195 (ppe), smc02171, smc02773R* and *smc03205R* genes, respectively (Table 3-1) (Figure 3-2). Since the putative *smc02773R* ORF overlaps with the putative *smc02773* ORF (also see description in Chapter II), the constructed mutant SmHC13 carries the plasmid-insertion mutation in both putative *smc02773* and *smc02773R* ORFs. Similarly, *S. meliloti* mutant SmHC14 carries the mutation in both *smc03205* and *smc03205R* ORFs due to the putative *smc03205R* overlaps with *smc03205* ORF.

Free-living phenotypes of the constructed mutants

Considering that some of these genes are predicted to be involved in cellular metabolism, all of the 5 mutants, SmHC10, SmHC11, SmHC12, SmHC13, and SmHC14 were examined for growth rate in both nutrient-rich medium LB/MC and minimal medium M9. None of them showed observable difference from wild-type strain Rm1021 (data not shown).

Since motility of *S. meliloti* cells affects nodulation, swarming tests were done for each of the five mutants. After 2-day incubation, the wild-type control strain Rm1021 produced a

colony with a diameter of about 2.0 cm on swarming plate, while on the same plate the negative control Rm7095 produced a colony with a diameter less than 0.5 cm. All five mutants produced swarming colonies similar to the size of Rm1021, which suggests that these five genes are not likely to be essential for *S. meliloti* mobility.

Succinoglycan production by each of the five mutants was also evaluated by its calcofluor fluorescence intensity on LB/MC/CF plate. No apparent difference was observed between these mutants and wild-type strain. This suggests these genes do not affect succinoglycan biosynthesis.

Symbiotic phenotypes of the constructed mutants

Each of the constructed plasmid-insertion mutants, SmHC10, SmHC11, SmHC12, SmHC13, and SmHC14 were used to inoculate alfalfa seedlings growing on plates. After four-week incubation, the alfalfa plants inoculated with Rm1021 (wild-type) were tall and green with pink nodules. The plants that were not inoculated with any *S. meliloti* cells turned yellow and with no root nodules. The plants inoculated with each of the five mutants were as green and healthy as those inoculated with the wild-type control. The percentages of pink nodules on plants inoculated with these mutants were from 94% to 100%, about the same as the wild-type strain.

The two mutants, SmHC10 and SmHC13, were further examined for their root hair invasion efficiency, because few white and gray nodules were found on the plants inoculated with these two mutants. To monitor the efficiencies of root hair invasion of the

two mutants, plasmid pHC60, which constitutively expresses the *gfp* gene, was moved into the two mutants. *S. meliloti* cells carrying plasmid pHC60 will give visible GFP fluorescence constitutively, which can be used to visualize *S. meliloti* cells in alfalfa root hairs under fluorescence microscope (46). As controls, the plants inoculated with Rm1021 formed colonized curled root hairs and had both initiated and extended infection threads, while the plants inoculated with Rm7210 (the *exoY* mutant) formed colonized curled root hairs but no initiated infection thread or extended infection thread (Figure 3-5). These are consistent with previous report (46). The plants inoculated with SmHC10 and SmHC13 exhibited the colonized curled root hairs, the initiated and extended infection threads at the levels similar to those plants inoculated with wild-type strain (Figure 3-5). These findings confirmed that the loss of the *dgkA* and *smc02773* (*smc02773R*) genes did not affect root hair invasion efficiencies by the bacteria.

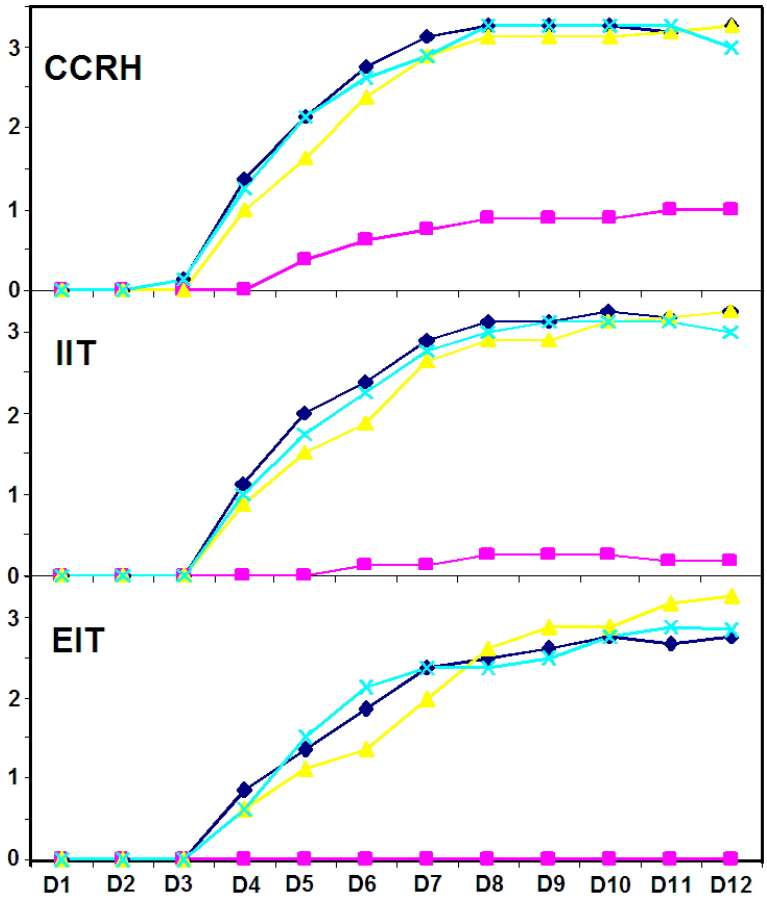

Figure 3-5. Infection thread formation of alfalfa seedlings inoculated with mutants, *dgkA* (—▲—) and *smc02773R* (—×—), or with control strains wild-type Rm1021 (—♦—) and mutant *exoY210* (—■—). Eight seedlings per treatment were used for each bacterial strain. The average number of colonized curled root hairs (CCRH), initiated infection threads (IIT) and extended infection threads was recorded from Day 1 to Day 12.

DISCUSSION

The exchanges of signal molecules between *S. meliloti* and alfalfa plant are required for the symbiosis establishment (29, 136). There could be additional unknown plant signals and bacterial genes required for the symbiosis, which is confirmed by the identification of 113 putative promoters using the improved IVET approach in my study. I have determined the promoter activities of the 113 different promoter fusions in the presence and absence of alfalfa root exudates. My results showed that 23 *S. meliloti* putative promoters among them were activated by alfalfa root exudates, which suggest that the genes expressed from these promoters could function at the very early stages of the symbiosis.

Induction test results showed 17 out of the 23 root-exudates inducible promoters are also induced by the *nod* gene inducer flavonoid apigenin. These genes or putative genes encode transcription factors, kinase, oxidase, transmembrane proteins and hypothetical proteins with unknown functions. These findings suggest that the interactions between *S. meliloti* and alfalfa at the early stages of the symbiosis establishment are much more complicated than previously thought. Although little is known about their roles in the symbiosis establishment so far, further analysis of these genes should result in better understanding of the *S. meliloti*-alfalfa symbiosis.

The other 6 out of the 23 root-exudates induced promoters did not respond to apigenin or luteolin. One of these six genes is *lsrA* gene. It has been identified as an essential symbiotic gene, because alfalfa inoculated with the *lsrA* null mutant developed only

white nodules (140). My finding that the expression of the *lsrA* gene is inducible by root exudates in the current study may provide new leads in characterizing the symbiotic function of the LsrA protein. The *lsrA* gene is likely expressed very early during symbiosis. The expression of the *lsrA* gene should be further analyzed for its symbiotic function and for its possible links to the succinoglycan production genes.

Another gene that has been characterized previously is the *dgkA* gene. It has long been linked to symbiosis through its role in the biosynthesis of cyclic 1,2 β-glucan (150), but the symbiotic phenotypes of *dgkA* null mutant have not been reported. Cyclic 1,2 β-glucan is essential for legume nodulation, which may function during several stages of plant infection (27). My finding that the *dgkA* mutant remains symbiotically active suggests that the *dgkA* gene is not essential for symbiosis. The enzymatic function of the DgkA protein could potentially be replaced by other diacylglycerol kinases in the cells.

The other four of the six genes are putative genes. The mutations in these genes did not affect their symbiosis efficiently, suggesting that none of the four genes is essential for symbiosis. It is possible that the functions of these gene products could potentially be replaced by other proteins.

The induction tests showed that six symbiotic genes were inducible by yet unknown components of root exudates rather than the known plant signal flavonoids, apigenin or luteolin. The component complexity of root exudates increases the possibility that other plant signals may exist, which induce bacterial symbiotic gene expression. To identify the

unknown plant signals in root exudates, the components of root exudates will be fractionated, and the induction activities of each fraction will be determined. This is currently being carried out jointly by Dr. Kennelly's lab and our lab. This study could lead to identification of new inducers in root exudates, which might reveal new mechanisms of the interactions between *S. meliloti* and alfalfa during symbiosis establishment.

A preliminary experiment was performed to better understand the plant signals in the root exudates inducing the *dgkA* gene. Plant root exudates were prepared as described in Method section from Day 1 to Day 6. The results of the first such analysis showed that the *dgkA* promoter activities were upregulated by the root exudates collected from Day 1 to Day 3, and were further upregulated by the root exudates collected from Day 4 to Day 6 (Figure 3-5). This *dgkA* expression pattern could be the result of a continuous accumulation of plant signals. Alternatively, there could be different signal molecules in root exudates upregulating the *dgkA* gene expression at the different stages of the symbiosis. For example, some signals appear in root exudates since Day 1, which upregulate the *dgkA* expression, while additional signals appear later in root exudates following the communication between bacterial cells and plant roots, which further upregulate the *dgkA* expression. In future experiments, root exudates by plant without inoculation of bacteria will be used as a second control to determine whether the unknown plant signals are produced specifically after communication with *S. meliloti* or are produced constitutively by plants even in the absence of bacteria. Further identification of the plant signals inducing the *dgkA* or the other symbiotic genes may

reveal new mechanisms of the interactions between *S. meliloti* and alfalfa during symbiosis establishment.

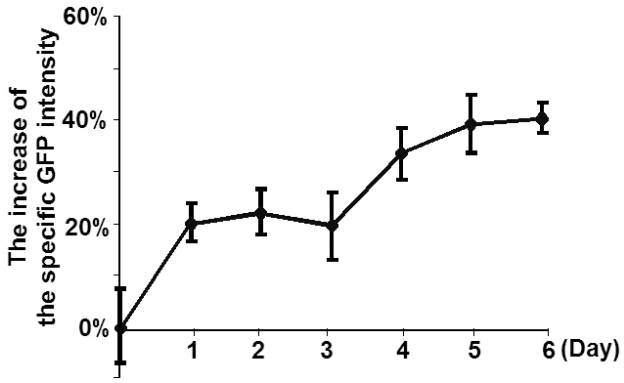

Figure 3-6. A diagram of induction effects of root exudates on the *dgkA* promoter activities. The *dgkA* promoter activity in the absence of any root exudates was set as a blank control. The increase of the *dgkA* promoter activities in the presence of root exudates after inoculation from Day 1 to Day6 compared to that in the absence of any root exudates.

Chapter IV

The role of the Putative Glycosaminoglycan Binding Site in Modulating the Function of the *Sinorhizobium meliloti* ExoR Protein

INTRODUCTION

Succinoglycan is required for the invasion of alfalfa by *S. meliloti* (46). Specifically, it is required for the initiation and elongation of infection threads. It might play important roles in the attachment of the bacteria to the plant root, suppression of host defense responses, and activation of enzymes involved in infection thread growth or root hair curling (21, 79, 222). It is also possible that succinoglycan might act as a signal and induce plant responses leading to the formation of infection threads (10, 46).

S. meliloti succinoglycan production is regulated both in free-living conditions and during symbiosis. In free-living cells, succinoglycan production can be affected by concentrations of ammonia, phosphate and sulfate (62, 126, 148); while during symbiosis, it might be turned off after bacterial cells are released into nodule primordium but before they differentiate into bacteroids (182).

The *S. meliloti* ExoR protein and the ExoS/ChvI two-component regulatory system are key regulators of succinoglycan production in free-living *S. meliloti* cells. ExoR suppresses succinoglycan production by negatively regulating the *exo* genes expression (47, 64, 162, 176). The loss-of-function mutation, *exoR95*::Tn*5*, results in the dramatic increase of succinoglycan production, which suggests ExoR is a negative regulator. ExoR regulates the *exo* genes expression at the level of transcription, because higher levels of mRNA of the *exo* genes were observed in the *exoR95*::Tn*5* mutant.

Agrobacterium tumefaciens also produces succinoglycan, which is its major acid exopolysaccharide and can be detected by calcofluor fluorescence on plate containing calcofluor (39). The amount of succinoglycan produced by *A. tumefaciens* is much higher than *S. meliloti* in free-living conditions (Figure 4-1). The *A. tumefaciens exo* genes involved in its succinoglycan production are also analogous to the *S. meliloti exo* genes. The *A. tumefaciens exo* genes mutation also can abolish its succinoglycan production, which generates calcofluor-dark phenotype on plate containing calcofluor (39). The putative *A. tumefaciens* ExoR protein shares 73% identity with *S. meliloti* ExoR, but the role of *A. tumefaciens* ExoR protein in regulating succinoglycan production has not been characterized yet.

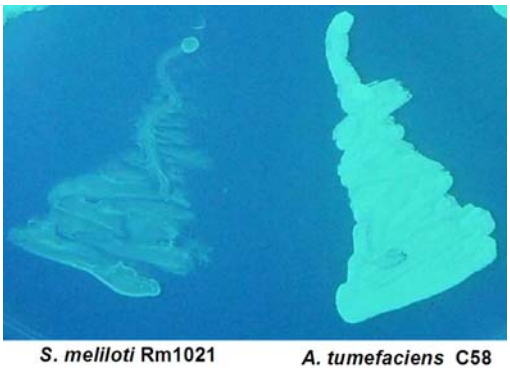

Figure 4-1. A picture of calcofluor fluorescence of *S. meliloti* and *A. tumefaciens* strains on LB/MC/CF plate. *S. meliloti* wild-type Rm1021 showed a low level of calcofluor fluorescence, while *A. tumefaciens* wild-type At C58 showed a high level of calcofluor fluorescence.

A careful analysis of the *S. meliloti* ExoR protein sequence showed that it has a putative glycosaminoglycan binding site. Using site-directed mutagenesis, the role of this putative binding site in modulating the regulatory function of the *S. meliloti* ExoR protein has been examined in this study.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

The bacterial strains and plasmids used in this project were listed in Table 4-1.

Table 4-1. Strains and plasmids

Strain or plasmid	Relevant properties	References
E. coli		
DH5α	General purposed strain	(103)
MT616	MT607, pRK600, Cm^R	(74)
S. meliloti		
Rm1021	Wild type, Sm^R	(126)
Rm7095	Rm1021 *exoR95*::Tn*5*, Nm^R	(64)
Rm7096	Rm1021 *exoS96*::Tn*5*, Nm^R	(64)
Rm7210	Rm1021 *exoY210*::Tn*5*, Nm^R	(126)
Rm8395	Rm1021 *exoR395*::Tn*5-233*, Sp^R	(229)
Rm8396	Rm1021 *exoS396*::Tn*5-233*, Sp^R	(229)
SmHC17	Rm1021 *glmS10*::Tn*5*, Nm^R	This work
SmHC18	Rm8395 *glmS10*::Tn*5*, Sp^R, Nm^R	This work
SmHC19	Rm8396 *glmS10*::Tn*5*, Sp^R, Nm^R	This work
A. tumefaciens		
C58	Wild type	(70)
Plasmids		
pHC93	Expression vector, pBBR1MCS3, Tc^R	(122)
PHC189	pHC93, constitutively expressing *S. meliloti glmS,* Tc^R	This work
PHC194	pHC93, constitutively expressing *S. meliloti exoR* (S173A), Tc^R	This work
PHC195	pHC93, constitutively expressing *A. tumefaciens exoR,* Tc^R	This work
PHC196	pHC93, constitutively expressing *A. tumefaciens exoR*(A173S)„ Tc^R	This work
PHC197	pHC93, constitutively expressing *S. meliloti exoR* (G174D), Tc^R	This work
pHC518	pHC93, constitutively expressing *S. meliloti exoR*, Tc^R	This work
pRK600	Helper plasmid, Cm^R	(74)
Phage		
ϕM12	Generalized transducing phage for *S. meliloti*	(73)

Escherichia coli strains used in this study were grown in Luria-Bertani (LB) medium at 37°C. *Agrobacterium tumefaciens* strains were grown in LB medium at 28°C. *Sinorhizobium meliloti* was grown in LB medium supplemented with 2.5mM $MgSO_4$ and 2.5mM $CaCl_2$ (LB/MC) at 30°C. Agar (1.5%) was used to solidify media. To examine succinoglycan production, calcofluor white M2R (Fluorescent Brightener 28, Sigma) was

added to a final concentration of 0.02% in LB/MC agar, which was buffered to pH 7.4 with 10mM HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid) (126). To quantify succinoglycan production with anthrone-H_2SO_4 method, bacteria were grown in minimal Z-MGS liquid media (140). The following antibiotics were used at the concentrations indicated: Ampicillin, 100μg/μl; chloramphenicol, 10μg/ml; gentamycin, 50μg/ml; kanamycin, 25μg/ml; neomycin, 200μg/ml; spectinomycin, 100μg/ml; streptomycin 500μg/ml; and tetracycline, 10μg/ml.

Construction of plasmids expressing the *S. meliloti exoR* and *A. tumefaciens exoR* genes

A 0.8kb DNA fragment containing the *S. meliloti exoR* ORF was generated by PCR using the genomic DNA of the wild-type *S. meliloti* Rm1021 as template and the pair of the primers *exoR*f*Nde*I and *exoR*r*Kpn*I (Table 4-2), and then cloned into the unique *Nde*I/*Kpn*I restriction site of vector pHC93, resulting in plasmid pHC518, on which the *S. meliloti exoR* gene will be expressed from the constitutive *lac* promoter (Figure 4-2).

Similarly, the *A. tumefaciens exoR* gene was cloned into vector pHC93 to make pHC195. The genomic DNA of the wild-type *A. tumefaciens* C58 and the pair of the primers *AtexoR*f*Nde*I and *AtexoR*r*Kpn*I were used in PCR reaction to generate the DNA fragment for cloning the *A. tumefaciens exoR* gene (Table 4-2).

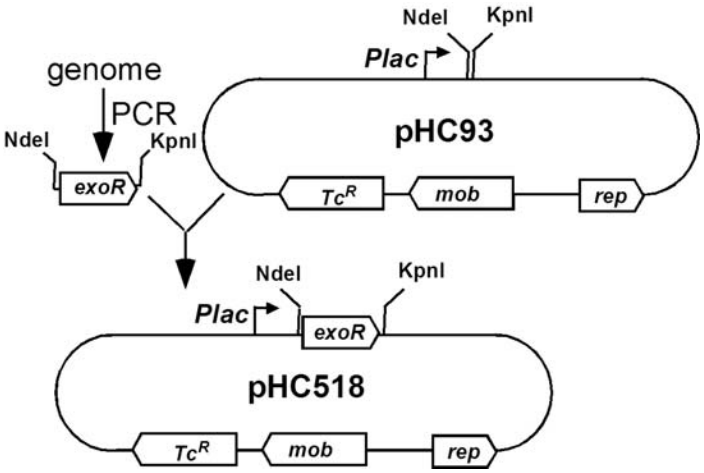

Figure 4-2. A diagram of construction of plasmid pHC518 to express *S. meliloti exoR* gene from *lac* promoter.

Table 4-2. Oligo DNA primers used for gene constructions

Name	Sequence of primers
*exoR*f*Nde*I	5'-GGCATATGCGCGCGGGTGAATTGAA-3'
*exoR*r*Kpn*I	5'-GGGGTACCTATCAATCGTCGTCGTTCTGC-3'
*AtexoR*f*Nde*I	5'-AGCCATATGCTGAAATGTGAAGCCAAC-3'
*AtexoR*r*Kpn*I	5'-GGGGTACCTCAATCCGGATCGTTGAA-3'
*exoR*rS173A	5'-TCCGCCCTCGCCCGCAAGCAGCATGCG-3'
*exoR*fS173A	5'-GCGCATGCTGCTTGCGGGCGAGGGCG-3'
seq*exoR*	5'-GGCATATGCGCGCGGGTGAATTGAA-3'
*AtexoR*rA170S	5'-CACCCTCGCCCGAGAGGATCATCTCCG-3'
*AtexoR*fA170S	5'-GGAGATGATCCTCTCGGGCGAGGGTGG-3'
seq*AtexoR*	5'-AGCCATATGCTGAAATGTGAAGCCAAC-3'
*glmS*f*Nde*I	5'- TTCCATATGTGCGGGATTGTTGGCA -3'
*glmS*r*Kpn*I	5'- GGGGTACCTATCATTCGACCGTCACCGA -3'

Site-directed mutagenesis of the *S. meliloti exoR* and *A. tumefaciens exoR* genes

To evaluate the effects of the putative glycosaminoglycan binding site on the function of the *S. meliloti* ExoR protein, the serine residue (S) at position 173 was replaced with alanine (A) using site-directed mutagenesis (Figure 4-3). A 0.53kb DNA fragment containing the first part of the *S. meliloti exoR* ORF (+1bp to +531bp) was generated by PCR using the genomic DNA of Rm1021 as template and the pair of the primers

*exoR*f*Nde*I and *exoR*rS173A (Table 4-2). The S173A site mutation of *exoR* ORF in this 0.53kb DNA fragment was introduced by primer *exoR*rS173A. A 0.30kp DNA fragment containing the second part of the *S. meliloti exoR* ORF (+504bp to +807bp) was generated by PCR using the genomic DNA of Rm1021 as template and the pair of the primers *exoR*fS173A and *exoR*r*Kpn*I (Table 4-2). The S173A site mutation of *exoR* ORF in this 0.30kb DNA fragment was introduced by primer *exoR*fS173A. The 0.53kb and 0.30kb DNA fragments have a 28 basepair overlap including the S173A site. Both DNA fragments were purified with QIAquick Gel Purification Kit (Qiagen Inc., Valencia, CA, USA) to remove excess primers. A 0.8kb DNA fragment containing the complete *S. meliloti exoR* ORF (+1 to +807bp) with the S173A site mutation was generated using the mixture of the 0.53kb and 0.30kb DNA fragments as template and the pair of primers *exoR*f*Nde*I and *exoR*r*Kpn*I. After double digestion with restriction enzymes *Nde*I and *Kpn*I, this 0.8kb fragment was cloned into the unique *Nde*I/*Kpn*I restriction site of vector pHC93, resulting in plasmid pHC194, on which the *S. meliloti exoR* (S173A) gene is expressed from *lac* promoter. The complete *S. meliloti exoR* (S173A) gene sequence was confirmed by sequencing using primer seq*exoR* at Cornell University DNA Sequencing Center (Table 4-2).

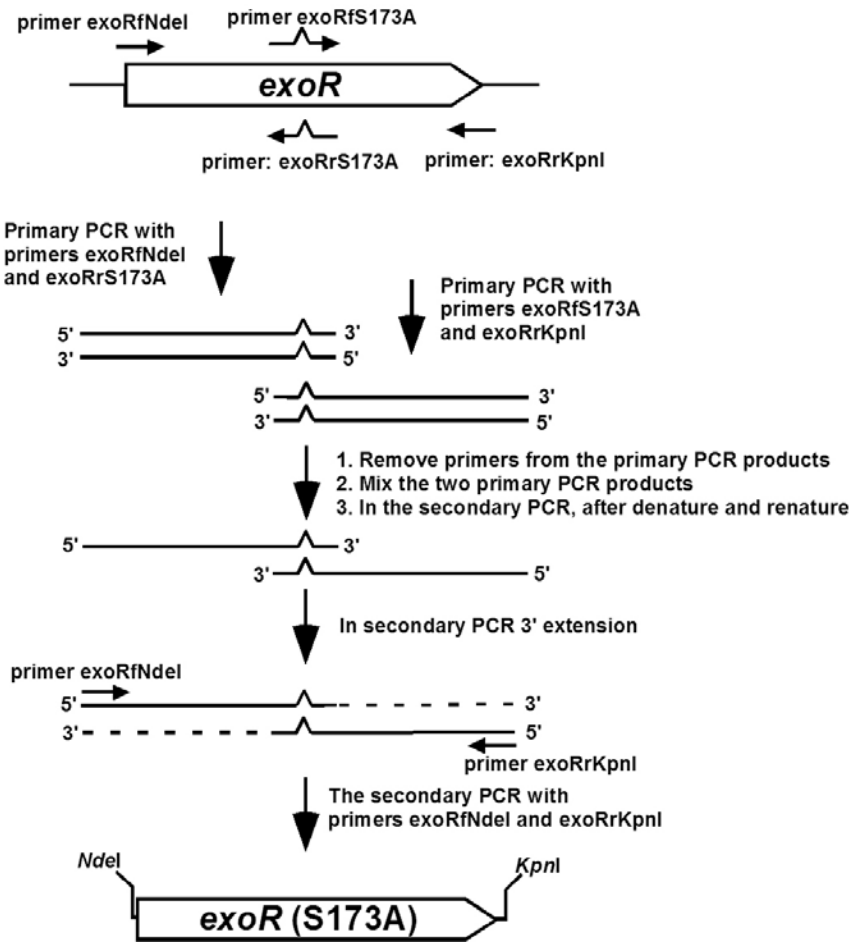

Figure 4-3. A diagram of construction of site-directed mutagenesis of *exoR* gene using recombinant PCR.

To add a putative glycosaminoglycan binding site on the *A. tumefaciens* ExoR protein, a similar site-directed mutagenesis method was used to replace the alanine residue at position 170 with serine. A 0.52kb DNA fragment containing the first part of the *A. tumefaciens exoR* ORF (+1bp to +520bp) was generated by PCR using the genomic DNA of the wild-type *A. tumefaciens* C58 as template and the pair of the primers *AtexoR*f*Nde*I and *AtexoR*rA170S (Table 4-2). The A170S site mutation of the *A. tumefaciens exoR* ORF in this 0.52kb DNA fragment was introduced by primer *AtexoR*rA170S. A 0.30kp DNA fragment containing the second part of the *A. tumefaciens exoR* ORF (+495bp to

+795bp) was generated by PCR using the genomic DNA of the *A. tumefaciens* C58 as template and the pair of the primers *AtexoR*fA170S and *AtexoR*r*Kpn*I (Table 4-2). The A170S site mutation of the *A. tumefaciens exoR* ORF in this 0.30kb DNA fragment was introduced by primer *AtexoR*fA170S. The 0.52kb and 0.30kb DNA fragments have a 26 basepair overlap including the A170S site. Both DNA fragments were purified with QIAquick Gel Purification Kit (Qiagen Inc., Valencia, CA, USA) to remove excess primers. A 0.8kb DNA fragment containing the complete *A. tumefaciens exoR* ORF (+1 to +795bp) with the A170S site mutation was generated using the mixture of the 0.52kb and 0.30kb DNA fragments as template and the pair of primers *AtexoR*f*Nde*I and *AtexoR*r*Kpn*I. After double digestion with restriction enzymes *Nde*I and *Kpn*I, this 0.8kb DNA fragment was cloned into the unique *Nde*I/*Kpn*I site in vector pHC93, resulting in plasmid pHC196, on which the *A. tumefaciens exoR* (A170S) gene is expressed from *lac* promoter. The complete *A. tumefaciens exoR* (A170S) gene sequence was confirmed by sequencing using primer seq*AtexoR* at Cornell University DNA Sequencing Center (Table 4-2).

All plasmids were electroporated into *E. coli* DH5α and then moved into *S. meliloti* Rm1021 or Rm7095 (the *exoR95* mutant) by conjugation. To move these plasmids in *A. tumefaciens,* plasmids were purified from *E. coli* DH5α and then electroporated into *A. tumefaciens* C58 directly.

Construction of plasmid expressing the *glmS* gene

The *S. meliloti glmS* gene was cloned into vector pHC93 and expressed from the constitutive *lac* promoter. A 1.9kb DNA fragment containing the complete *glmS* ORF was generated by PCR using the *S. meliloti* Rm1021 genomic DNA as template and the pair of primers *glmS*f*Nde*I and *glmS*r*Kpn*I (Table 4-2), and then cloned into the unique *Nde*I/*Kpn*I restriction site of pHC93, resulting in plasmid pHC189, on which the *glmS* gene is expressed from *lac* promoter.

Construction of double mutants

To facilitate the construction of double mutants, *S. meliloti* strain Rm8395 and Rm8396 were used, which carries the *exoR395*::Tn*5-233* (Sp^R) and *exoS396*::Tn*5-233* (Sp^R) mutations, respectively. They are the same as the *exoR95*::Tn*5* and *exoS96*::Tn*5* mutants with the exception of the antibiotic markers (64, 229).

Double mutations were constructed by using phage ϕM12-mediated transduction (73). SmHC17 carrying the *glmS10*::Tn*5* mutation was lysed with bacteriophage ϕM12 first and then transduced into Rm8395. The Rm8395 carrying the *glmS10*::Tn*5* mutation was selected on LB plate containing neomycin. The constructed *glmS10*::Tn*5 exoR395*::Tn*5-233* double mutant was named as SmHC18. Similarly, the *glmS10*::Tn*5* mutation was transduced into Rm8396, which resulted in the *glmS10*::Tn*5 exoS396*::Tn*5-233* double mutant named as SmHC19.

Succinoglycan production measurement assay

To detection the levels of succinoglycan production, *S. meliloti* cells were collected from fresh LB/MC liquid culture, washed with Z-MGS medium and then resuspended in 50 ml Z-MGS with streptomycin (200μg/ml) to an OD_{600} of 0.1. The cultures were incubated at 30°C for 5 days with shaking (120 rpm), and sampled (1 ml culture) every 24 hours to monitor OD_{600} and collect succinoglycan. To prepare succinoglycan, 0.2 ml of bacterial sample was diluted with 0.8 ml sterile distilled water, and then centrifuged at 13,000 rpm for 10 min at 4°C. The supernatant containing succinoglycan was transferred to a new 1.5 ml centrifuge tube and kept at -20°C for future analysis.

The amount of succinoglycan from each bacterial sample was determined using anthrone-H_2SO_4 method as described before (134). Briefly, each succinoglycan sample (0.4ml) was mixed with 0.1 ml fresh anthrone (2% in ethyl acetate) in clean 10 ml glass test tube. A total of 1 ml undiluted H_2SO_4 was added into the solution slowly. The reaction solution was let stand for 10 min to cool to room temperature before 100 μl solution was transferred into a transparent 96-well plate to determine OD_{620} using an absorbance microplate reader (Spectra Max 340PC, Molecular Device, Sunnyvale, CA, USA). The value of OD_{620} of each sample was normalized to its corresponding cell density to generate the specific succinoglycan production value. Z-MGS medium was used as blank control in this assay. All experiments were repeated for three times independently.

RESULTS

Identification of the putative glycosaminoglycan binding site in the *S. meliloti* ExoR protein

A putative glycosaminoglycan binding site at residue positions from 173 to 176 (Ser-Gly-Glu-Gly) was identified by an analysis of the *S. meliloti* ExoR protein sequence using softwares in Biology Workbench Program (http://workbench.sdsc.edu) (the upper sequence in box, Figure 4-4). It matches the consensus sequence of the glycosaminoglycan binding site Ser-Gly-Xaa-Gly, where Xaa is any amino acid. Glycosaminoglycan attaches to the core protein through a xyloside residue which is linked to the serine residue of tetrapeptide Ser-Gly-Xaa-Gly of the core protein (24, 187). Substitution of the serine residue in this tetrapeptide by any other residue would completely eliminate the glycosaminoglycan binding activity (24).

The *S. meliloti* ExoR protein shares high identity with the putative *A. tumefaciens* ExoR protein. However, no glycosaminoglycan binding site was predicted by analysis of the *A. tumefaciens* ExoR protein sequence using the same software. The amino acid sequence of *A. tumefaciens* ExoR protein corresponding to *S. meliloti* ExoR Ser-Gly-Xaa-Gly (residue 173-176) is Ala-Gly-Xaa-Gly (residue 170-173) (the lower sequence in box, Figure 4-4), which misses the most critical serine residue.

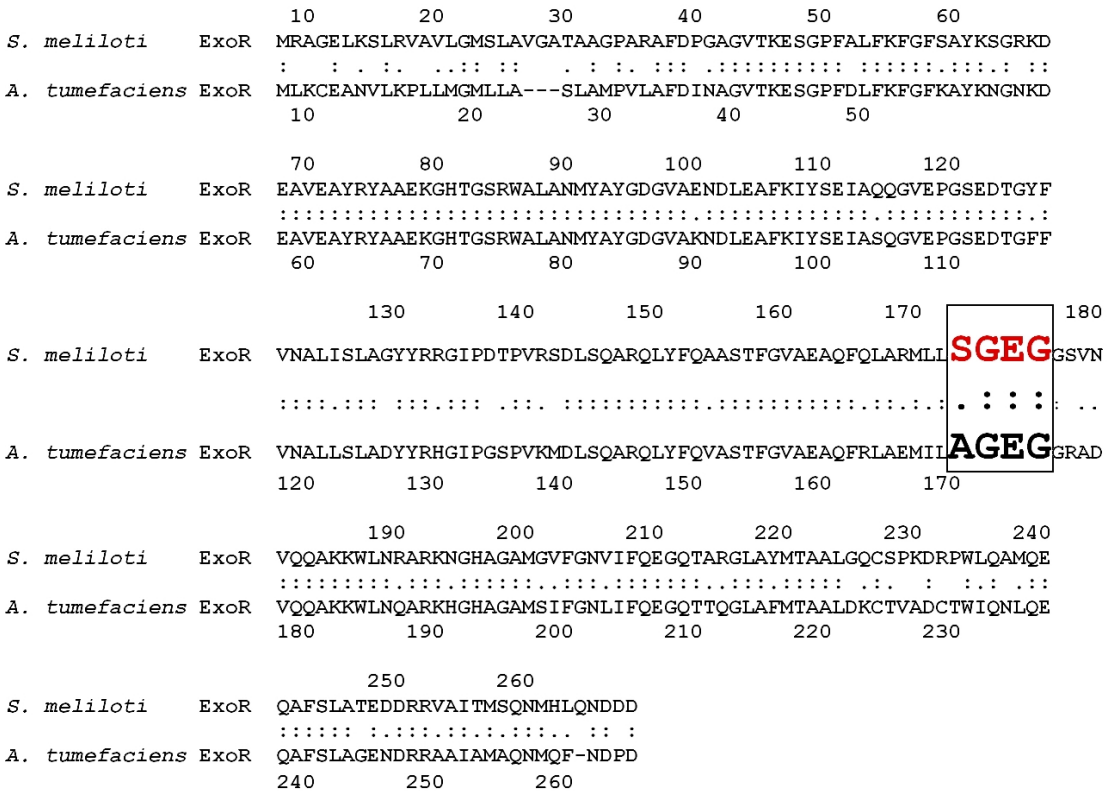

Figure 4-4. Alignment of the amino acid sequences of the *S. meliloti* ExoR and *A. tumefaciens* ExoR proteins. Identical amino acids are shown by two dots. Conservative amino acid substitutions are shown by one dot. Box shows the glycosaminoglycan binding site (SGXG).

Site-directed mutagenesis of the putative glycosaminoglycan binding site of the *S. meliloti* and *A. tumefaciens* ExoR proteins

To evaluate the importance of the putative glycosaminoglycan binding site (SGXG) on the function the *S. meliloti* ExoR protein, the serine residue at 173 was replaced by alanine using site-directed mutagenesis. The constructed *S. meliloti* $ExoR_{S173A}$ protein might lose the glycosaminoglycan binding site (if it is real) due to the importance of the serine residue in the binding site. Expression plasmid pHC518 and pHC194, which carries wild-type *exoR* ORF and $exoR_{S173A}$ ORF, respectively, were constructed and conjugated into *S. meliloti* Rm1021 and Rm7095.

As a control, the residue alanine at position 170 of the *A. tumefaciens* ExoR protein was replaced by serine to generate a SGXG glycosaminoglycan binding site at the corresponding position using site-directed mutagenesis. The *A. tumefaciens exoR* and $exoR_{A170S}$ genes were cloned similarly as those of *S. meliloti*, forming plasmid pHC195 and plasmid pHC196, respectively, and introduced into *S. meliloti* Rm1021 and Rm7095.

The loss of the putative glycosaminoglycan binding site affects the function of the *S. meliloti* ExoR protein

Succinoglycan production by *S. meliloti* cells expressing the *S. meliloti* ExoR and $exoR_{S173A}$ proteins was first examined on LB/MC/CF plate and then quantified using anthrone-H_2SO_4 assay. As shown in Figure 4-5, the amount of succinoglycan produced by Rm1021 (wild-type strain) and Rm7095 (the *exoR95* mutant) were 33.3±10.2 and 264.4±25.8, respectively, which are consistent with previous report (64). Rm7095 (pHC518) that expresses the *S. meliloti exoR* gene on plasmid pHC518 produced succinoglycan at levels similar to wild-type. The amount of succinoglycan produced by Rm7095 (pHC194) that expresses the *S. meliloti* $exoR_{S173A}$ was 112.7±17.3. This was about 3-fold of that produced by the *exoR95* mutant expressing the *S. meliloti exoR* gene, but still significantly lower than that of the *exoR95* mutant. As shown in Figure 4-6, while the presence of multiple copies of the *S. meliloti exoR* gene greatly reduced succinoglycan production in wild-type strain, the presence of multiple copies of the *S. meliloti* $exoR_{S173A}$ gene only slightly reduced succinoglycan production in Rm1021.

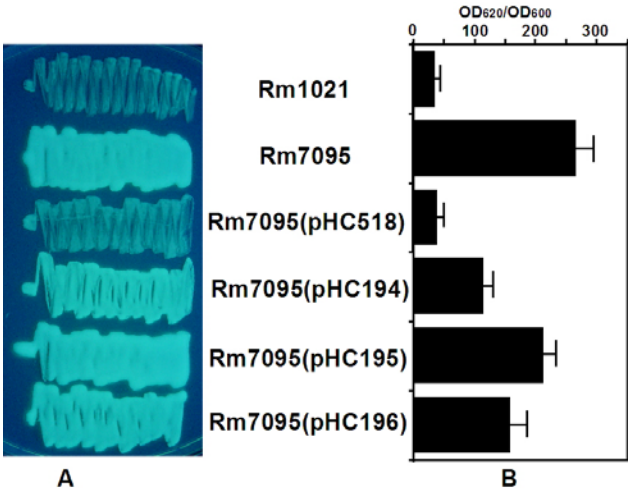

Figure 4-5. A picture of calcofluor fluorescence of *S. meliloti* strains on LB/MC/CF plate to evaluate succinoglycan production (A) and a diagram of the specific succinoglycan production by *S. meliloti* strains (B). From top to bottom: wild-type Rm1021; the *exoR95* mutant Rm7095; Rm7095 (pHC518) expressing the *S. meliloti exoR* gene; Rm7095 (pHC194) expressing the *S. meliloti* $exoR_{S173A}$ gene; Rm7095 (pHC195) expressing the *A. tumefaciens exoR* gene; and Rm7095 (pHC196) expressing the *A. tumefaciens* $exoR_{A170S}$ gene.

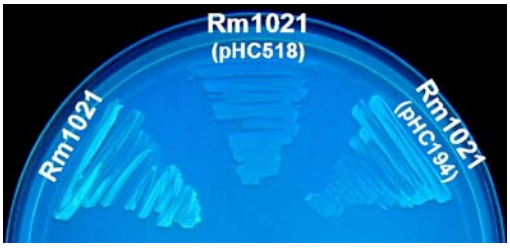

Figure 4-6. A picture of calcofluor fluorescence of *S. meliloti* strains on LB/MC/CF plate. Kodak 1D analysis of calcofluor fluorescence intensities indicated the value of Rm1021 is 43.9±7.7, the value of Rm1021 (pHC518) expressing the *S. meliloti exoR* gene is 10.8±1.4, and the value of Rm1021 (pHC194) expressing the *S. meliloti* $exoR_{S173A}$ gene is 28.5±3.1.

These data clearly indicate that the S173A mutation reduced function of the *S. meliloti* ExoR protein, but did not abolish its function completely. Considering that S173A site mutation abolishes the putative glycosaminoglycan binding site, these results suggest that the binding site may be important but not essential for the *S. meliloti* ExoR protein function.

The amount of succinoglycan produced by *S. meliloti* Rm7095 (pHC195) that expresses the *A. tumefaciens exoR* gene was 212.1±20.1, which was 20% lower than that of the *S. meliloti exoR95* mutant. The amount of succinoglycan produced by Rm7095 (pHC196) that expresses the *A. tumefaciens exoR*$_{A170S}$ gene was 156.8±28.3, which was 40% lower than that of the *S. meliloti exoR95* mutant. That the presence of the *A. tumefaciens* ExoR$_{A170S}$ protein suppressed succinoglycan overproduction in the *S. meliloti exoR95* mutant background at a greater level suggests that the addition of the glycosaminoglycan binding site might contribute to the regulatory function of the *A. tumefaciens* ExoR protein in *S. meliloti* cells.

Reduction of intracellular pool of glucosamine resulting in succinoglycan overproduction

GlmS negatively regulates S. meliloti succinoglycan production

The *S. meliloti glmS10*::Tn*5* mutant which was isolated in our lab was examined for succinoglycan production. *S. meliloti* cells were streaked on LB/MC/CF plate and the calcofluor fluorescence intensities were quantified using Kodak 1-D analysis software (Figure 4-7). As controls, the wild-type strain Rm1021 showed low level of fluorescence intensities; the *exoY210* mutant showed the lowest fluorescence intensities due to no succinoglycan production; the *exoR95* mutant gave the highest fluorescence level. These data are consistent with previous report (64, 152). The level of calcofluor fluorescence intensities of the *glmS10* mutant was higher than Rm1021 but lower than the *exoR95* mutant, indicating that loss of the *glmS* gene increased succinoglycan production. The expression of the *glmS* gene on plasmid pHC189 reduced the succinoglycan production

by the *glmS10* mutant, suggesting the *S. meliloti* GlmS protein is most likely a negative regulator of succinoglycan production.

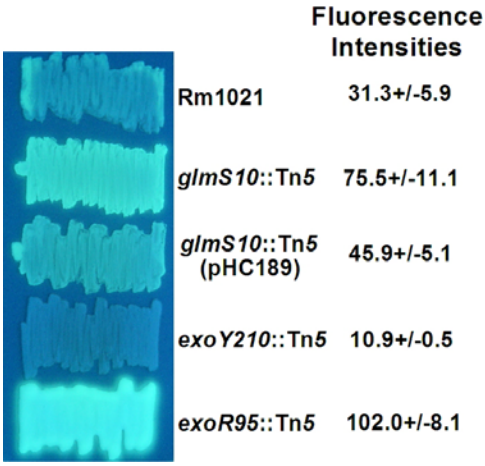

Figure 4-7. A picture of calcofluor fluorescence of *S. meliloti* strains, Rm1021, *glmS*::Tn*5*, *glmS*::Tn*5* (pHC189), *exoY210*::Tn*5*, *exoR95*::Tn*5* on LB/MC/CF plate. 1-D analysis indicated the values of the fluorescence intensities.

Glucosamine suppresses the phenotype of the glmS10::Tn5 mutation

The *S. meliloti* GlmS protein is a glucosamine synthase. It catalyzes D-glucosamine-6-phosphate synthesis, which is the first committed reaction in biosynthesis of glucosamine, an essential substrate of most kinds of glycosaminoglycan (3, 59). Since the GlmS protein might negatively regulate succinoglycan production, the intracellular concentration of glucosamine may be related to the levels of succinoglycan production. To test this possibility, succinoglycan productions by *S. meliloti* strains were evaluated by calcofluor fluorescence intensities on LB/MC/CF plates with or without exogenous glucosamine (10mM).

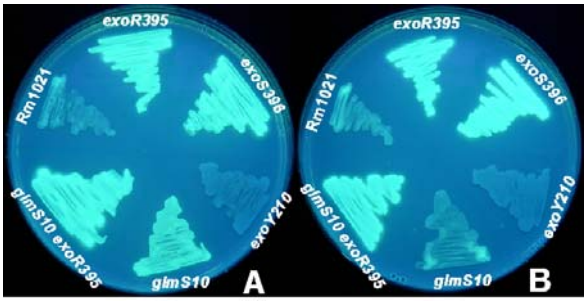

Figure 4-8. A picture of calcofluor fluorescence of *S. meliloti* strains on LB/MC/CF plate with or without exogenous glucosamine. Plate A: LB/MC/CF plate. Plate B: LB/MC/CF plate containing 10mM glucosamine.

As shown in Figure 4-8, succinoglycan production by wild-type strain Rm1021, the *exoR395* mutant or the *exoS396* mutant did not change regardless of the presence or absence of exogenous glucosamine. In the absence of exogenous glucosamine, the *glmS10* mutant produced more succinoglycan than wild-type Rm1021, but less than the *exoR395* and *exoS396* mutants, while in the presence of 10 mM exogenous glucosamine, the succinoglycan production by the *glmS10* mutant was significantly reduced. The *glmS10 exoR395* double mutant had the same succinoglycan overproduction phenotype as the *exoR395* mutant in both media.

DISCUSSION

Glycosaminoglycans (GAG) are polymers of repeating disaccharides. Within the disaccharides, the sugars tend to be modified, with amino groups, acidic groups, sulfated hydroxyl groups, etc. One or more GAG chains covalently links to a core protein to form proteoglycan (120). Glycosaminoglycan attaches to the core protein through a xyloside residue which is covalently linked to the serine residue of tetrapeptide Ser-Gly-Xaa-Gly (24, 187), the consensus sequence recognized by xylosyltransferase. The significance of this sequence as glycosaminoglycan binding site was confirmed by the fact that different synthetic peptides containing this Ser-Gly-Xaa-Gly sequence were efficient to serve as biosynthetic acceptors for xylosyltransferase-catalyzed linkage of xyloside to peptides (24, 186). The serine residue is the most critical of the invariant residues, even substitution by threonine at this position would eliminate the glycosaminoglycan binding activity completely (24).

The analysis of primary sequence of *S. meliloti* ExoR showed that the *S. meliloti* ExoR protein contains the tetrapeptide Ser-Gly-Xaa-Gly at positions 173-176. The secondary structure prediction of the *S. meliloti* ExoR protein suggests the tetrapeptide Ser-Gly-Xaa-Gly may be located in a coil region between two alpha-helices. The tertiary structure prediction of the *S. meliloti* ExoR protein suggests the tetrapeptide may be displayed on the surface of the protein (Analysed by softwares on http://www.expasy.org). These analyses raised the possibility that the tetrapeptide Ser-Gly-Xaa-Gly on the *S. meliloti* ExoR protein may be a glycosaminoglycan binding site.

The *S. meliloti* ExoR protein is thought to mediate ammonia regulation of the *exo* genes expression and succinoglycan production. This behavior is quite different from that of its close homologs, the *Rhizobium leguminosarum biovar viciae* ExoR protein, which does not have the putative glycosaminoglycan binding site (64, 179). The *S. meliloti* ExoR protein shares 73% amino acid identity with the *A. tumefaciens* ExoR protein on the primary sequence. However, neither the *A. tumefaciens* ExoR protein has the glycosaminoglycan binding site. To better understand the regulatory mechanisms of succinoglycan by the ExoR, the role of the binding site in modulating the function of the *S. meliloti* ExoR protein was extensively analyzed through site-directed mutagenesis.

The site-mutagenized *S. meliloti* ExoR protein, $ExoR_{S173A}$, significantly disabled its function in suppressing succinoglycan production, suggesting that the binding site is important for the function of *S. meliloti* ExoR protein. But the fact that *S. meliloti* $ExoR_{S173A}$ still suppresses succinoglycan production in both wild-type strain Rm1021 and the *exoR95* mutant at certain levels indicates that *S. meliloti* ExoR protein may have other domain(s) to perform its regulatory function.

These findings and that the GlmS protein negatively regulates succinoglycan production possibly through affecting intracellular pool of glucosamine, have led to the proposal for a new regulation model (Figure 4-9). We suggest that the GlmS protein regulates the intracellular pool of glucosamine, which indirectly controls the levels of glycosaminoglycans. The binding of the ExoR protein with glycosaminoglycans derivative is important for the activity of ExoR, which affects the *exo* genes expression

and succinoglycan production. But a lot more detailed analyses are required to support this model.

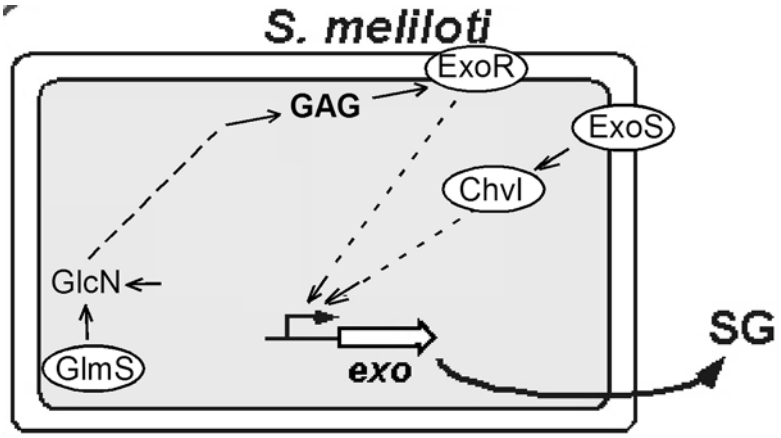

Figure 4-9. The proposed model that glucosamine regulates succinoglycan production through affecting ExoR activity. GlmS controls glucosamine (GlcN) biosynthesis. As an essential substrate, GlcN affects glycosaminoglycan (GAG) biosynthesis. The GAG binding is important for the activity of the ExoR protein, which affects the *exo* genes expression and succinoglycan production.

Since very few homologs of *S. meliloti* ExoR have been studied, it is difficult to predict ExoR regulatory function based on known knowledge. But the finding of SEL-1 domain on *S. meliloti* ExoR protein and its homologs suggests ExoR may directly bind to other protein(s) to perform its regulatory function. One direction of the future research is to find the protein that may interact with the ExoR by using the two hybrid screening system specially designed to identify the partner of membrane protein (170).

Chapter V

Both *Sinorhizobium meliloti exoY* Promoters are Active During Symbiosis

INTRODUCTION

The presence of succinoglycan is required for the formation of infection threads, a critical step to establish nitrogen-fixing symbiosis. The amount of succinoglycan is also critical. The overproduction of succinoglycan caused by mutations in the *exoR* and *exoS* genes reduces the ability of *S. meliloti* to elicit infection thread formation (46, 47). After successful invasion of plant cells inside root nodules, the succinoglycan production is down regulated (182).

While the regulatory mechanism of succinoglycan production is not clear, the expression of the *exoY* gene appears to be a critical part of it. The succinoglycan production by *S. meliloti* is most likely regulated by its ExoR protein and the ExoS/ChvI regulatory system through the expression of the *exoY* gene (18, 46, 47). The *exoY* gene is the first gene in the *exoYFQ* operon and encodes a glycosyl transferase that carries out the first step of succinoglycan production (18, 93, 94, 183). Loss of the *exoY* gene blocks succinoglycan production completely (126).

To better understand the regulation of the *exoY* gene expression, a study to identify the *exoY* promoters by using nested deletions was carried out in our lab, showing that the *exoY* gene is expressed from two different promoters in free-living cells (48). To answer the crucial question which one of the two or both exoY promoters are active during symbiosis, a functional complementation system was established to analyze the activities of the two *exoY* promoters in free-living conditions and during symbiosis individually.

MATERIALS AND METHODS

Strains and growth media

Bacterial strains and plasmids used in this study were listed in Table 5-1.

Table 5-1. Strains and plasmids

Strain and plasmid	Relevant properties	References
E. coli		
DH5α	Host for cloning	(103)
MT616	MT607, pRK600, Cm^R	(74)
S. meliloti		
Rm1021	Wild-type, Sm^R	(125)
RmAR9007	Rm1021 *exoYΩLacZ*, Gm^R	(117)
Rm7095	*exoR95*::Tn*5*, Sm^R, Nm^R	(64)
Rm7096	*exoS96*::Tn*5*, Sm^R, Nm^R	(64)
SmHC15	*exoR95 exoYΩLacZ,* Sm^R, Nm^R, Gm^R	This work
SmHC16	*exoS96 exoYΩLacZ,* Sm^R, Nm^R, Gm^R	This work
Plasmids		
pHC77	pMB393, carrying *exoX-exoY* intergenic region and *exoY*::*gfp* fusion, Sp^R	(48)
pRK600	Helper plasmid, Cm^R	(74)
pPHU231	Derivative of pRK290, low-copy number, broad-host range, Tc^R	(68)
pMB393	Cloning vector, Sp^R	(83)
pHC153	Carrying *exoY* native promoter region, *exoY gfp* fusion, Tc^R	This work
pHC172	Carrying *Pnod and gfp* fusion, Sp^R	This work
pHC178	Carrying P_{exoYup} and $P_{exoYdown}$, *exoY, gfp* fusion, Tc^R	This work
pHC181	Carrying $P_{exoYdown}$, *exoY, gfp* fusion, Tc^R	This work
pHC182	Carrying *exoY, gfp* genes without promoter, Tc^R	This work
pHC191	Carrying P_{exoYup}, *exoY, gfp* fusion, Tc^R	This work
Phage		
ϕM12	Generalized transducing phage for *S. meliloti*	(73)

Escherichia coli strains used in this study were grown in Luria-Bertani (LB) medium at 37°C. *Sinorhizobium meliloti* was grown in LB medium supplemented with 2.5mM $MgSO_4$ and 2.5mM $CaCl_2$ (LB/MC) at 30°C. Agar (1.5%) was used to solidify media. To examine succinoglycan production, calcofluor white M2R (Fluorescent Brightener 28, Sigma) was added to a final concentration of 0.02% in LB/MC agar, which was buffered to pH 7.4 with 10mM HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid) (126). To prepare cells for the *exoY* promoter activity analysis, *S. meliloti* strains were

incubated in Z-MGS liquid medium (126). The following antibiotics were used at the concentrations indicated: Ampicillin, 100μg/μl; chloramphenicol, 10μg/ml; gentamycin, 50μg/ml; kanamycin, 25μg/ml; neomycin, 200μg/ml; spectinomycin, 100μg/ml; streptomycin 500μg/ml; and tetracycline, 10μg/ml.

Construction of double mutants

Double mutations were constructed by using phage ϕM12-mediated transduction (73). Briefly, RmAR9007 carrying the *exoY/lacZ* mutation was lysed with bacteriophage ϕM12 and then transduced into Rm7095 (the *exoR95*::Tn*5* mutant). The Rm7095 carrying *exoY/lacZ* mutation was selected on LB plate containing gentamycin. The constructed *exoY/lacZ* and *exoR95*::Tn*5* double mutant was named as SmHC15. Similarly, the *exoY/lacZ* mutation was transduced into Rm7096 (the *exoR96*::Tn*5* mutant), which resulted in the *exoY/lacZ exoR96*::Tn*5* double mutant named as SmHC16.

Plasmid Construction

To determine activities of the two *exoY* promoters individually during free-living and symbiotic stages, a set of DNA fragments with or without the individual *exoY* promoters were fused with the *exoY* and *gfp* genes and cloned into a low-copy-number plasmid pPHU231 (Figure 5-1).

To construct a plasmid carrying a promoterless *exoY* and *gfp* genes, a 1.7kb DNA fragment containing ORFs of both genes was generated by PCR using plasmid pHC153 as the template and the pair of primers Y182f and GFPr (see Table 5-2). Plasmid pHC153

carries the almost entire *exoX-exoY* intergenic region and the *exoY* ORF fused with *gfp* ORF, which replaces the native *exoF* ORF. An *Xho*I and *Kpn*I restriction site were added to the 1.7kb DNA fragment through the use of primers in PCR. The DNA fragment was purified using QIAquick PCR purification kit (Qiagen Inc. Valencia, CA, USA), digested with restriction enzymes *Xho*I and *Kpn*I, and cloned into the unique *Xho*I/*Kpn*I restriction site of pPHU231 to form plasmid pHC182. Neither *exoY* nor *gfp* is expressed by any promoters on this plasmid.

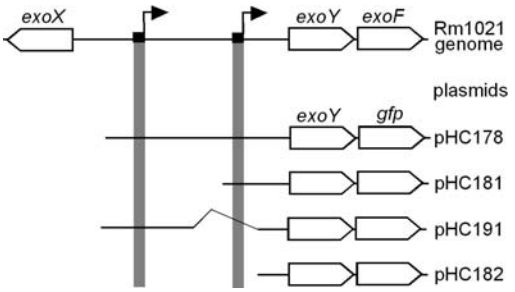

Figure 5-1. A diagram of construction of a set of fragments with or without the individual *exoY* promoters fused with the *exoY* and *gfp* genes.

Table 5-2. Oligo DNA primers used to clone and express the *exoY* and g*fp* genes under the different *exoY* promoters in the low-copy-number plasmid pPHU231.

Name	Sequence of primers
Y178f	5'-ATACCGCTCGAGCGGGGCCACTATATTAGCGCCC-3'
Y181f	5'-ATACCGCTCGAGCGGGCCATCATTCCGCCTTCA-3'
Y182f	5'-ATACCGCTCGAGCGGATGAGCCCGCGTCCCAC-3'
Y191f	5'-GCTCTAGAGCGATTGGGAACCTCATA-3'
Y191r	5'-CCGCTCGAGCAGGGGTTAGGATAGTTC-3'
GFPr	5'-CGGGGTACCCCGATAGTTCCTCCTTTCAGCAAAAAA-3'

To construct the fusion of both upstream and downstream promoters of the *exoY* gene and both the *exoY* and *gfp* genes, a 2.3kb DNA fragment covering the –705 to –1 region upstream of the *exoY* gene and the *exoY* and *gfp* genes was generated by PCR using plasmid pHC153 as the template and the pair of primers Y178f and GFPr (see Table 5-2).

An *Xho*I restriction site and a *Kpn*I restriction site were similarly added to the fragment through the use of the primers. The 2.3kb DNA fragment was digested and then cloned into the unique *Xho*I/*Kpn*I restriction site of pPHU231 to form plasmid pHC178. On this plasmid both the *exoY* and *gfp* genes can be expressed from the two *exoY* promoters in front of the *exoY* gene.

To construct the fusion of the *exoY* downstream promoter, $P_{exoYdown}$, and both the *exoY* and *gfp* genes, a 1.7kb DNA fragment covering the -104 to –1 region upstream of the *exoY* gene and the *exoY* and *gfp* genes was generated by PCR using plasmid pHC153 as the template and the pair of primers Y181f and GFPr (see Table 5-2). An *Xho*I restriction site and a *Kpn*I restriction site were added to the fragment through the use of the primers. The 1.7kb DNA fragment was digested and cloned into the *Xho*I/*Kpn*I restriction site of pPHU231 to form plasmid pHC181. On this plasmid, both the *exoY* and *gfp* genes can only be expressed from the *exoY* downstream promoter $P_{exoYdown}$ in front of the *exoY* gene.

To construct the fusion of the *exoY* upstream promoter, P_{exoYup}, and both the *exoY* and *gfp* genes, a 0.6kb DNA fragment covering the –749 to –190 region upstream of the *exoY* gene was generated by PCR using the pair of primers Y191f and Y191r (see Table 5-2), and plasmid pHC77, which carries the entire *exoX-exoY* intergenic region, as template. An *Xba*I restriction site was added to the –749 end of the fragment and an *Xho*I restriction site was added to the –190 end of the fragment. The 0.6kb DNA fragment was digested and cloned into the unique *Xba*I/*Xho*I restriction site of plasmid pHC182 to

generate plasmid pHC191. On this plasmid both the *exoY* and *gfp* genes can only be expressed from the *exoY* upstream promoter P_{exoYup} in front of the *exoY* gene.

All these plasmids were constructed in *E. coli* DH5α and then moved into *S. meliloti* strains through triparental mating with *E. coli* MT616 (pRK600) as the helper.

Nodulation assays on plates

Assays of nodulating alfalfa on plates were carried out as previously described (126). Briefly, bacterial cells of different strains were collected from fresh LB/MC cultures, washed and diluted with sterile distilled water to an OD_{600} of 0.03. Alfalfa seeds were surface-sterilized and germinated in dark at 26°C for 40 hours. Eight seedlings were placed on each square Petri dish with solid Jensen medium. The seedlings were inoculated with *S. meliloti* by flooding the roots with 1 ml bacterial cell suspension (OD_{600} 0.03). The plates were wrapped on both sides and bottom with aluminum foil and placed in plant growth chamber under light at 26°C. Plants were examined after three weeks for the number, color, and positions of roots nodules to determine the overall symbiosis efficiency. The stems and leaves of each of the plants might be separated from their roots and weighed when needed.

Nodule profile

Profiles of nodules along main alfalfa roots were determined as previously described (36). Briefly, nodules formed on the main roots of alfalfa plants were examined for the color, number and their distance from the stem-root junction at the end of three weeks.

Determination of the specific GFP fluorescence intensity of bacterial cultures

To determine the specific GFP fluorescence intensity, *S. meliloti* cells were collected from fresh Z-MGS liquid culture, washed and diluted with sterile PBS buffer (pH 7.4) to an OD_{600} of 0.1. Each suspension was transferred to wells of a transparent 96-well plate and a black 96-well plate in equal amounts. Cell suspensions in transparent 96-well plate were used to determine the cell density using an absorbance microplate reader (Spectra Max 340PC, Molecular Device, Sunnyvale, CA, USA), and suspensions in black 96-well plate were used to determine the intensity of GFP fluorescence using a fluorescence microplate reader (Spectra Max Germini XS, Molecular Device, Sunnyvale, CA, USA). The GFP fluorescence intensity of each bacterial culture was normalized to its cell density to generate the specific GFP fluorescence intensity. All tests were performed in triplicate.

Plant root exudates preparation

Alfalfa seedlings were grown in hydroponic system described in Chapter II to prepare exudates. Briefly, about 100 alfalfa seeds were surface-sterilized, and then spread on wet cheesecloth suspended with aluminum screen in a $11\times11\times9$ mm^3 container with 200 ml nitrogen-free Jensen liquid medium. Two ends of the cheesecloth were submerged in liquid medium to ensure the cheesecloth on the aluminum screen remain wet. The containers with seeds were kept in dark at 26°C for 40 hours for seed germination, and then exposed to light. The liquid medium in each container was inoculated with 1 ml *S. meliloti* Rm1021 bacterial suspension (OD_{600} 0.3). On day 5 after inoculation, 100 ml

liquid Jensen medium with bacterial cells and root exudates was removed from each container. Bacterial cells and insoluble materials were removed from the liquid medium by centrifugation (13,000rpm, 15 min, 4°C). The supernatant was further purified by passing through a sterilization filter with pore size of 0.22μm and used as crude root exudates.

Root exudates and flavonoid induction assay

To determine the induction effect of root exudates on each of the two *exoY* promoters, the cells of RmAR9007 carrying the set of plasmids were collected from fresh LB/MC culture, washed with sterile distilled water, and resuspended in alfalfa root exudates to a concentration of an OD_{600} of 0.1. After 18-hour incubation at 30°C, the specific GFP fluorescence intensity of bacterial cells in the root exudates was determined.

Similarly, to determine the effect of flavonoids on each of the two *exoY* promoter, the cells of RmAR9007 carrying the same set of plasmids were resuspended and incubated in the supernatant of liquid Jensen medium with alfalfa flavonoid apigenin or luteolin (2.5μg/ml).

The cells of each bacterial strain were also incubated in the supernatant of liquid Jensen medium without root exudates or flavonoids as the blank control for the root exudates and flavonoid induction test.

RESULTS AND DISCUSSION

Both the *exoY* promoters are active in free-living cells

To determine promoter activities of the two *exoY* promoters individually in free-living cells, a set of four plasmids, pHC178, pHC181, pHC191 and pHC192, carrying the *exoY* and *gfp* genes fused to the different *exoY* promoters was constructed and conjugated into the wild-type strain Rm1021, the *exoR95* and *exoS96* mutants. The activities of the *exoY* promoters were determined by measuring the levels of specific GFP fluorescence intensities of the cells carrying different plasmids (Figure 5-2).

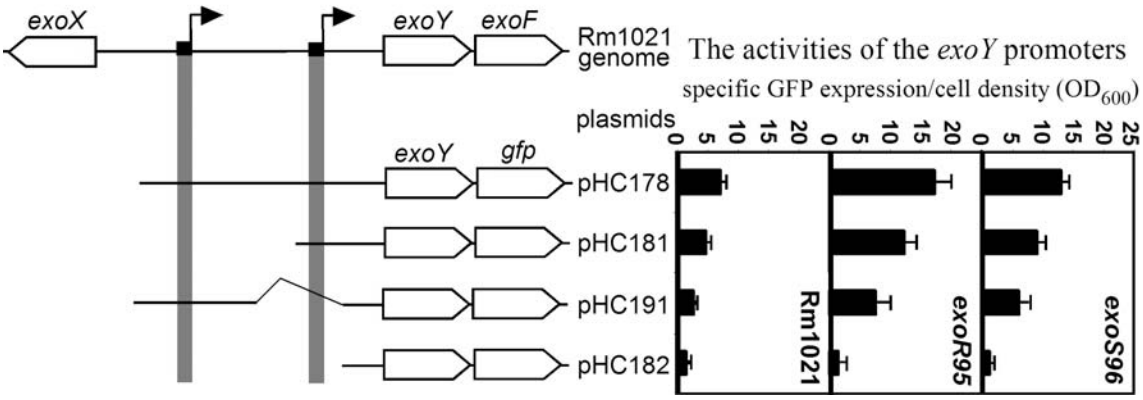

Figure 5-2. Schematic representation of the *exoY* promoters and their activities in different genetic backgrounds. A line graph shows the *exoYF* and *exoX* genes region. The two gray vertical bars represent the relative position of the two *exoY* promoters. The different promoter activities of the set of plasmids in the wild-type Rm1021, the *exoR95* and *exoS96* mutants are shown in three different bar graphs.

In the wild-type background, the specific GFP fluorescence intensities of cells carrying the *exoY* and *gfp* genes fused to no promoter, the P_{exoYup} promoter, the $P_{exoYdown}$ promoter, and both P_{exoYup} and $P_{exoYdown}$ promoters were 1.51±0.41, 2.73±0.30, 4.77±0.91, and 7.13±0.78, respectively. These results suggest that in free-living wild-type cells, both

exoY promoters are active and the $P_{exoYdown}$ promoter is stronger than the P_{exoYup} promoter.

In the *exoR95* mutant background, the specific GFP fluorescence intensities of the cells carrying the *exoY* and *gfp* genes fused to no promoter, the P_{exoYup} promoter, the $P_{exoYdown}$ promoter, and both P_{exoYup} and $P_{exoYdown}$ promoters were 1.58±1.13, 7.71±2.17, 12.29±2.10, and 17.36±2.78, respectively. In the *exoS96* mutant background, the set of the corresponding values were 1.28±0.18, 6.14±1.66, 9.12±1.17 and 13.09±1.16.

These results suggest that both promoters are more active in the *exoR95* and the *exoS96* mutant backgrounds than in wild-type cells, and the $P_{exoYdown}$ promoter is still stronger than the P_{exoYup} promoter in both mutant backgrounds. Since the DNA sequences of the two *exoY* promoters are very different, these results also suggest that the ExoR protein and the ExoS/ChvI regulatory system most likely regulate the expression of the *exoY* gene indirectly. The effect of ExoR on the expression of the *exoY* gene is stronger than that of ExoS/ChvI on it. But it is not clear at this time whether the ExoR protein and ExoS/ChvI regulatory system belong to the same signal transduction pathway or different pathways.

The expression of the *exoY* gene from either promoters allows complement an *exoY* null mutant to synthesize succinoglycan

To determine whether the expression levels of the *exoY* gene under different promoters are sufficient to complement the *exoY* null mutation in strain RmAR9007, the same set of

plasmids carrying the *exoY* and *gfp* genes fused to the different *exoY* promoters were conjugally introduced into RmAR9007 individually. When streaked on LB/MC with calcofluor plate, different strains showed different calcofluor fluorescence intensities reflecting their levels of succinoglycan production. As shown in Figure 5-3, RmAR9007 (pHC182) (without either *exoY* promoters) was dark, indicating that no succinoglycan was produced. RmAR9007 (pHC191) (with the P_{exoYup} promoter) were slightly brighter than RmAR9007 (pHC182), but darker than RmAR9007 (pHC181) (with the $P_{exoYdown}$ promoter), indicating that RmAR9007 (pHC181) produced more succinoglycan than RmAR9007 (pHC191). RmAR9007 (pHC178) (with both promoters) showed the highest levels of succinoglycan production.

To further compare the *exoY* promoter activities in different mutant backgrounds, the same set of plasmids was also conjugated individually into strain SmHC15, the *exoY/lacZ exoR395* double mutant, and SmHC16, the *exoY/lacZ exoS396* double mutant. Neither SmHC15 nor SmHC16 produce succinoglycan due to the *exoY* null mutation. As negative controls, both SmHC15 (pHC182) and SmHC16 (pHC182) were dark on calcofluor media because pHC182 dose not carry either *exoY* promoters. Again, the presence of either promoter on plasmid is able to complement the *exoY* null mutation to produce succinoglycan, although at different levels (Figure 5-3).

These results confirmed again that in free-living conditions, the *exoY* gene is expressed from both promoters. The expression profiles obtained in different backgrounds are consistent with the previous results (in the previous section). The expression of *exoY*

from either *exoY* promoter carried on the plasmid is able to complement the *exoY* null mutation to produce succinoglycan.

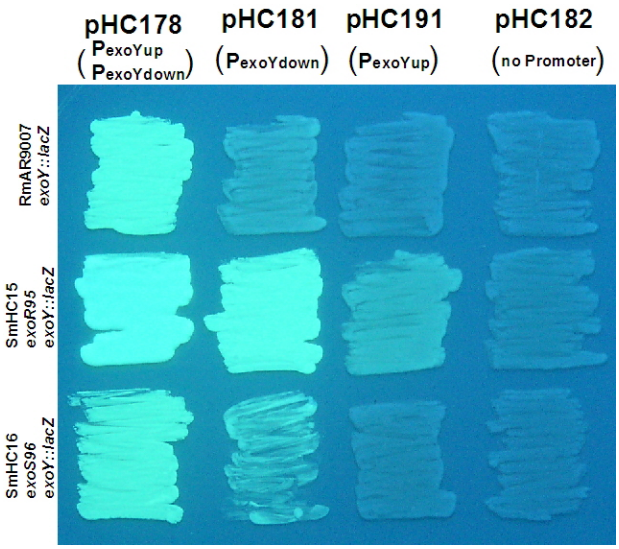

Figure 5-3. A picture of calcofluor fluorescence of *S. meliloti* strains in the complementary test on LB/MC/CF plate.

The fact that the lack of succinoglycan production by the RmAR9007 strain due to the loss of the *exoY* gene can be restored by expressing the *exoY* gene on a plasmid and that the amount of succinoglycan produced reflects the levels of the *exoY* promoter activities make it possible to examine the activities of the *exoY* promoters during symbiosis based on the production of succinoglycan.

Both *exoY* promoters are active during symbiosis

To determine promoter activities of the two *exoY* promoters individually during symbiosis, the cells of RmAR9007 carrying the set of plasmids with the *exoY* and *gfp* genes fused to the different *exoY* promoters were used to inoculate alfalfa seedlings. As shown in Table 5-3, the alfalfa plants inoculated with Rm1021 (wild-type strain) were

very healthy. The average weight per plant was 82 mg and the average number of pink nodules per plant was 6.1. In contrast, plants inoculated with RmAR9007 (pPHU231) or RmAR9007 (pHC182) had only white nodules and did not grow after exhausting nutrients from the seeds. The alfalfa plants inoculated with RmAR9007 (pHC191) grew well with an average of 4 pink nodules per plant. These pink nodules account for 44% percent of the total nodules, which reflects a low efficiency of nodule invasion. The alfalfa plants inoculated with RmAR9007 (pHC181) grew better than plants with RmAR9007 (pHC191), with an average of 5.6 pink nodules per plant, suggesting the efficiency of nodule invasion was higher by this strain. The alfalfa plants inoculated with RmAR9007 (pHC178) grew as well as those inoculated with the wild-type control, and also have similar number of pink nodules. These results demonstrated that the expression of the *exoY* gene from either promoter can complement the nodulation deficiency of an *exoY* null mutant at different levels.

Table 5-3. Nodulation efficiency of the strains carrying the *exoY* and *gfp* genes from the two different *exoY* promoters*

Strains	Pink Nodules (%)	No. Pink Nodule/Plant**	Wet Weight of Stem & Leaves/plant (mg)***
Rm1021	100	6.1	82.2
RmAR9007 (pPHU231)	0	0	24.6
RmAR9007 (pHC178)	98	6.5	84.8
RmAR9007 (pHC181)	78	5.6	68.3
RmAR9007 (pHC191)	44	4	52.6
RmAR9007 (pHC182)	0	0	25.5

*Alfalfa plants were examined for nodules and total wet weight at the end of three weeks.
**The number of pink nodules was average of eight plants for each bacterial strain.
***The wet weight of stems and the leaves were average of eight plants for each of the strains.

To further explore the difference between the expressions of the two *exoY* promoters during nodulation, profiles of nodules along main alfalfa root were determined (Figure 5-

4). On alfalfa plants inoculated with Rm1021, all nodules formed on main root were pink. The nodule number increased gradually to 7 per eight plants at regions 40 to 45 mm (millimeter) from the stem-root junction, and then decreased to negligible in region 55 to 60 mm. The nodules induced by RmAR9007 and RmAR9007 (pPHU231) were all white, and the number of the nodules peaked and decreased further away from the stem-root junction compared to the wild-type control. This is consistent with the notion that plants allow more nodules to develop when starved for nitrogen (36). The nodule profiles on plants inoculated with RmAR9007 (pHC178) are very similar to those induced by wild-type strain Rm1021. RmAR9007 (pHC181) elicited mostly pink (82%) nodules on plants, and the nodule number peaked and declined at regions slightly further away from the stem-root junction. Plants inoculated with RmAR9007 (pHC191) had a lot of white nodules (58%) which were spread evenly along the roots. Fewer nodules were formed at regions closer to the stem-root junction.

All together, these data suggest that both the *exoY* promoters are active during symbiosis, and strains expressing the *exoY* gene from either promoter were less efficient in nodulating alfalfa than the strains expressing the *exoY* gene from both promoters. The $P_{exoYdown}$ promoter was found again to be more active than the P_{exoYup} promoter during symbiosis. The distributions of the pink nodules along the main roots inoculated by bacterial strains carrying either promoter are similar, suggesting that both P_{exoYup} and $P_{exoYdown}$ promoters may function at the similar stage of nodulation and it is possible that the two promoters are similarly regulated.

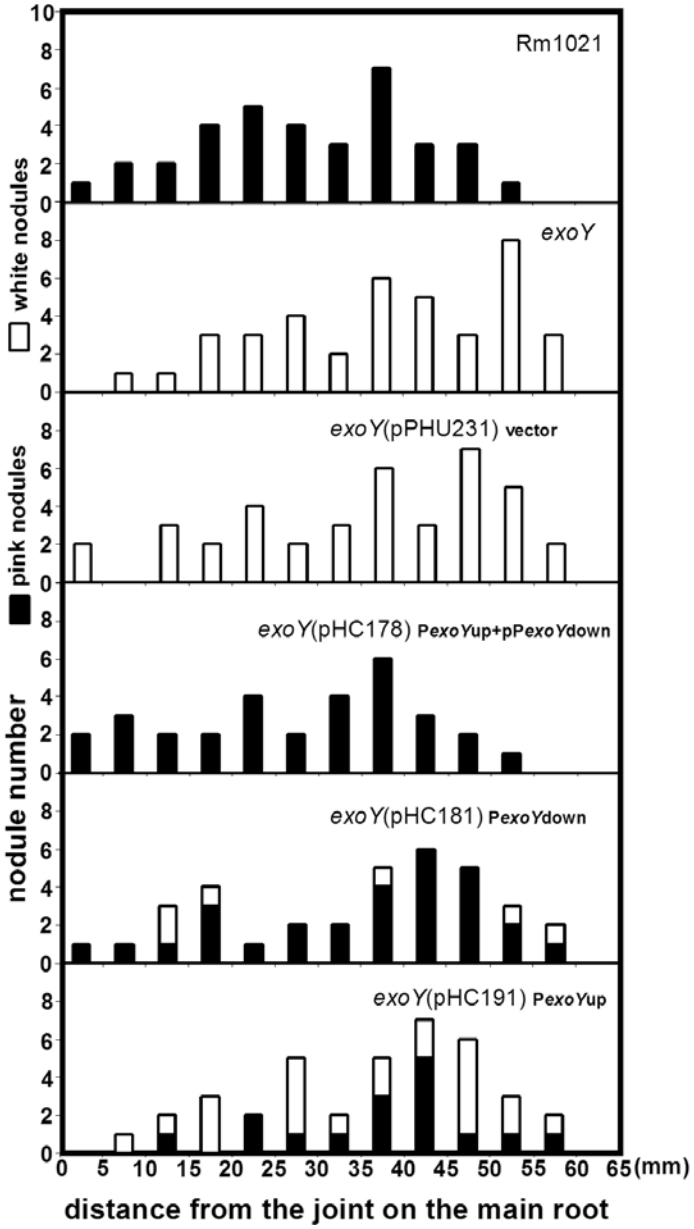

Figure 5-4. A diagram of symbiosis efficiency of strains expressing the *exoY* gene from the different *exoY* promoters. Eight alfalfa plants were inoculated with each one of the bacterial strains. The color of nodules and their distance from the juncture of root and stem was determined after three weeks. The total number of nodule from 5 mm section on each of the eight plants were combined and plotted. The black portion of the bar represents the number of pink nodules and the white portion of the bar represents the number of white nodules.

Neither *exoY* promoters are alfalfa root-exudate inducible

To examine whether one of the two *exoY* promoters or both promoters are induced by flavonoids or root exudates during symbiosis, the specific GFP fluorescence intensities of

RmAR9007 carrying the set of plasmids were determined in presence and absence of root exudates or selected flavonoids (Figure 5-5). However, none of these strains showed apparent difference in specific GFP fluorescence intensities after treated with them. These findings suggest that neither *exoY* promoters were inducible by apigenin, luteolin or alfalfa root exudates.

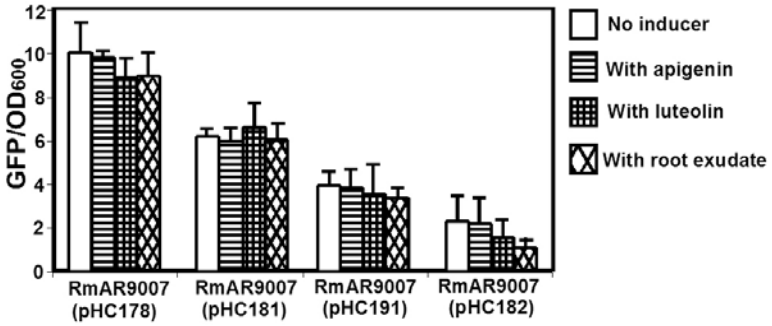

Figure 5-5. A diagram of the activities of the two *exoY* promoters in the presence of apigenin, luteolin or root exudates or in the absence of any of them. RmAR9007 (pHC178) was used to determine $P_{exoYdown}$ and P_{exoYup} activity, RmAR9007 (pHC181) was used to determine $P_{exoYdown}$ activity, RmAR9007 (pHC191) was used to determine P_{exoYup} activity and RmAR9007 (pHC182) was used as negative control to show non-promoter activity.

These findings also raise the possibility that other signals or changes such as direct contacting with seedling root or root hair may regulate the activity of either of the *exoY* promoters to increase bacterial succinoglycan production at early stages during symbiosis establishment.

In summary, the symbiotically important *exoY* gene is indeed expressed from the two different promoters, $P_{exoYdown}$ and P_{exoYup}, in free-living conditions and during symbiosis. Both promoters are similarly regulated by the ExoR protein and the ExoS/ChvI two-component regulatory system. The two promoters appear to function at similar stages

during nodulation. The $P_{exoYdown}$ promoter is more active than the P_{exoYup} promoter both in free-living conditions and during symbiosis. Further studies are needed to characterize the expression of the *exoY* genes during symbiosis and the roles of the exopolysaccharide succinoglycan in the interactions between the bacterium and its plant host alfalfa.

Chapter VI

Identification of *Sinorhizobium meliloti* Genes Involved in the ExoS/ChvI Pathway to Regulate Succinoglycan Biosynthesis

INTRODUCTION

S. meliloti succinoglycan biosynthesis is regulated in part by the ExoR and ExoS proteins. The *exoR95*::Tn*5* mutant and *exoS96*::Tn*5* mutant overproduce succinoglycan by increasing the transcriptional level of the *exo* genes (64, 176). The *exoR95*::Tn*5* mutant overproduces succinoglycan regardless of the presence or absence of ammonia in the medium, whereas the *exoS96*::Tn*5* mutant still responds to ammonia and produces higher amount of succinoglycan in its absence. Thus, the ExoR protein was thought to mediate sensing of ammonia in regulating succinoglycan production.

The ExoS protein encodes the histidine kinase sensor protein of the *S. meliloti* ExoS/ChvI two-component regulatory system (47). The *S. meliloti chvI* gene was located just upstream of the *exoS* gene, and ChvI serves as the response regulator of ExoS protein. The ExoS/ChvI two-component regulatory system might be involved in sensing environmental signals that influence succinoglycan production (47). The *exoS* gene is homologous to the *Agrobacterium tumefaciens chvG* gene, which encodes the histidine kinase sensor protein of the ChvG/ChvI two-component regulatory system and is involved in the regulation of acid-inducible genes (44, 131).

The *exoS96*::Tn*5* allele encodes an N-terminal truncated derivative of ExoS in the *exoS96*::Tn*5* mutant. The shortened transmembrane domain of ExoS96 was thought to keep the protein itself in a constitutively active form, leading to an increase in the phosphorylation of the regulator ChvI. This change results in the overexpression of the e*xo* genes and the subsequent overproduction of succinoglycan (47). The *S. meliloti*

ExoS/ChvI two-component regulatory system also appears to be involved in flagellum biosynthesis through regulating flagellum protein gene expression, although there is no sequence similarity between the *fla* gene and *exo* genes (229). This suggests that ExoS regulates the *fla* and *exo* genes indirectly and there are other elements downstream of the ExoS/ChvI two-component regulatory system which may specifically regulate *fla* or *exo* genes expression.

In this project, I reported my efforts to identify elements related to the ExoS/ChvI signal transduction pathway in regulating succinoglycan.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

The bacterial strains and plasmids used in this project were listed in Table 6-1.

Table 6-1. Strains and plasmids

Strain and plasmid	Relevant properties	References
E. coli		
DH5α	General purposed strain	(103)
MM294	Pro-82 thi-1 endA1 hsdR17 supE44, pRK602, Km^R	(126)
MT616	MT607, pRK600, Cm^R	(74)
S. meliloti		
Rm1021	Wild-type, Sm^R	(126)
Rm5003	Rm1021 *ntrC*::Tn*5*, Nm^R	(209)
Rm7210	Rm1021 *exoY210*::Tn*5*, Nm^R	(126)
Rm8395	Rm1021 *exoR395*::Tn*5-233*, Sp^R	(229)
Rm8396	Rm1021 *exoS396*::Tn*5-233*, Sp^R	(229)
SmHC16	Rm7096 *exoYΩLacZ,* Sm^R, Nm^R, Gm^R	This work
STB1	Rm8396 *sma1541*::Tn*5*, Sp^R, Nm^R	This work
STJ48	Rm8396 *smc01135*::Tn*5*, Sp^R, Nm^R	This work
STJ53	Rm8396 *smc00248*::Tn*5*, Sp^R, Nm^R	This work
Plasmids		
pEX312	*exoAHF*-complementing plasmid, Tc^R	(135)
pRK600	Helper plasmid, Cm^R	(74)
pRK602	pRK600, containing Tn*5*, Nm^R	(74)
Phage		
ϕM12	Generalized transducing phage for *S. meliloti*	(73)

Escherichia coli strains used in this study were grown in LB medium at 37°C. *S. meliloti* strains were grown in LB medium supplemented with 2.5mM $MgSO_4$ and 2.5mM $CaCl_2$ (LB/MC) at 30°C. Agar (1.5%) was used to solidify media. To examine succinoglycan production, calcofluor white M2R (Fluorescent Brightener 28, Sigma) was added to a final concentration of 0.02% in LB/MC agar, which was buffered to pH 7.4 with 10mM HEPES (N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid) (126). The following antibiotics were used at the concentrations indicated: chloramphenicol, 10μg/ml;

kanamycin, 25μg/ml; neomycin, 200μg/ml; spectinomycin, 100μg/ml; streptomycin 500μg/ml; and tetracycline, 10μg/ml.

Transposon Tn*5* mutagenesis

Transposon Tn*5* insertions in *S. meliloti* were performed by triparental conjugation (126). Briefly, by using *E. coli* MT616 (pRK600) as a helper, *E. coli* strain MM294 (pRK602) was mated with *S. meliloti* strain Rm8396 (pEX312). Tn*5* mutants of *S. meliloti* Rm8396 (pEX312) were selected on plates with the appropriate antibiotics, neomycin, tetracycline and streptomycin.

Screen for Tn*5* mutants of *S. meliloti* Rm8396 (pEX312) that exhibit altered succinoglycan phenotypes

Transposon Tn*5* insertion mutants of *S. meliloti* Rm8396 (pEX312) that suppress succinoglycan-overproduction phenotype of the *exoS396*::Tn*5-233* mutation and appear as calcofluor-dim or dark colonies under UV light were picked from the selective LB/MC/CF plates. Mutants with other altered succinoglycan phenotypes were also picked from the plates, such as colonies exhibiting different calcofluor fluoresces color or more mucoid colonies. All the isolates were purified by streaking on new selective plates to obtain single colonies. To ensure there was no secondary mutation, the isolated mutants were transduced by ϕM12 as described before (73). All transductants were evaluated again on LB/MC/CF plate under UV light.

Mapping the positions of the Tn*5* insertion in the *S. meliloti* genome

The positions of the Tn*5* insertion in the genome of the isolated mutants were located with a PCR-based mapping method (Figure 6-1).

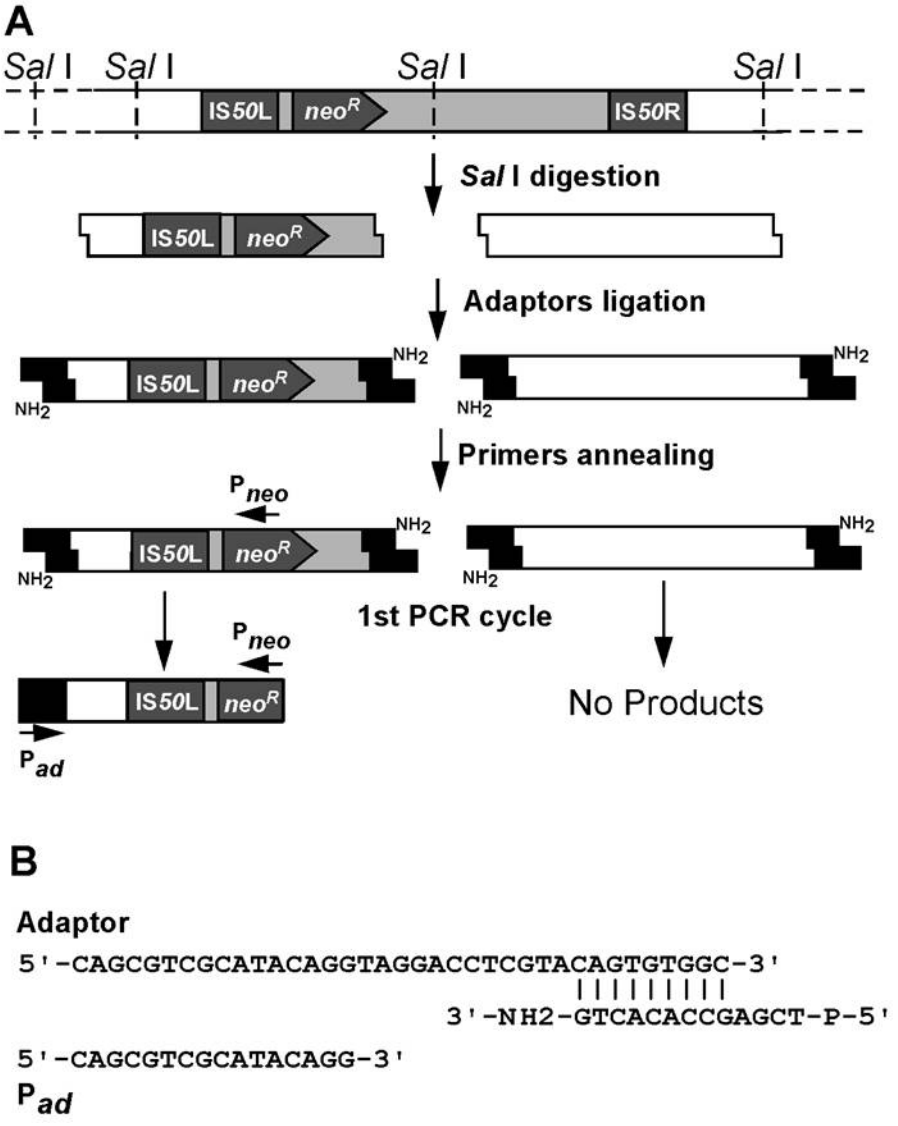

Figure 6-1. Schematic representation of the strategy of direct mapping of transposon Tn*5* insertion in genome (panel A). The sequence alignment of the two oligos of the adaptor and primer Pad (panel B).

The genomic DNA of each *S. meliloti* Tn*5* mutant was extracted using Wizard genomic DNA preparation kit (Promega Corp., Madison, WI, USA) and digested with restriction

enzyme *Sal*I (New England BioLabs Inc., Beverly, MA, USA). The digested DNA was recovered using QIAquick PCR purification kit (Qiagen Inc., Valencia, CA, USA), and then resuspended in sterile type-III water to a concentration of 0.1 μg/μl.

An adaptor was constructed by annealing a long oligo 5'-CAG CGT CGC ATA CAG GTA GGA CCT CGT ACA GTG TGG C-3' and a short oligo DNA 5'-P-TCG AGC CAC ACT G-NH2-3' (Invitrogen Corp., Frederick, MD, USA). To prepare the adaptor, the two oligonucleotides were suspended into sterile type-III water at a concentration of 0.25 nmol/μl and mixed (20 μl each) in 0.3 ml PCR tube. The solution was heated at 92°C for 2 min, and then cooled down slowly at 3°C per 1 min to 24°C using a PCR Mastercycler Gradient by Eppendorf Scientific (Eaton Corp., Cleveland, OH, USA).

The adaptor was ligated to restriction enzyme *Sal*I generated genomic DNA fragments in a 30 μl ligation mixture containing 20 μl (2 μg) *Sal*I digested genomic DNA fragments, 1.5 μl adaptors (0.15 nmol), 2 μl T4 ligase (800 units), and 3 μl 10×T4 ligase buffer. The ligation reaction was carried out at 16°C for 12 hours and terminated at 65°C for 10 min. The DNA in the ligation mixture was recovered using QIAquick PCR purification kit and resuspended in 30 μl sterile type-III water.

The juncture of *S. meliloti* genomic DNA and Tn*5* was amplified by PCR using adaptor capped genomic DNA as template and the pair of primers, the transposon primer P_{neo} 5'-GAG CAG CCG ATT GTC TGT TGT G-3' and the adaptor primer P_{ad} 5'-CAG CGT CGC ATA CAG GTA GG-3'. The PCR reactions were carried out in 50 μl mixtures

containing 5 µl 10×Taq polymerase buffer, 1 µl diluted (1:10) ligation product, 0.75 µl of each primers (10 pmol/µl), 0.5 µl Taq DNA polymerase (New England BioLabs Inc., Beverly, MA, USA), 1 µl dNTPs (10 mM each) and 41 µl sterile type-III water. The PCR reactions were performed as following: 94°C for 3 min to denature the DNA, and then 94°C for 1 min, 70°C for 30 sec, 72°C for 4 min in the first cycle. In the subsequent cycles, the annealing temperature is gradually decreased 1°C every two cycles, until it reached to 55°C, and continued for another 20 more cycles using Mastercycler gradient by Eppendorf Scientific. PCR products were recovered using QIAquick PCR purification kit, resuspended in 40 µl sterile type-III water, and examined on a 1% agarose gel.

The exact Tn*5* insertion site was determined by sequencing the PCR product using sequencing primer 5'-CCT CTG ATG AGA TGG TTA TTG A-3', which anneals to the IS*50*L region of Tn*5*. DNA sequencing was performed by DNA Sequencing Center at Cornell University. The sequence results were analyzed using BLASTN program on *S. meliloti* strain Rm1021 genome project website (http://bioinfo.genopole-toulouse.prd.fr).

RESULTS AND DISCUSSION

Development of a PCR-walking method to map Tn*5* insertion in *S. meliloti* genome

To quickly identify *S. meliloti* genes interrupted by transposon Tn*5*, a PCR based mapping strategy was developed from a previously reported PCR-walking method used to map the sites of T-DNA insertions in plant genome (5). The key to the success of this method is the use of a special adaptor, which produces a primer binding site only after the first cycle of PCR reaction was carried out. Therefore, it allows specific amplification of the juncture region between transposon and adjacent genomic DNA. The interrupted genes can be identified based on DNA sequence at the juncture.

Genomic DNA of *S. meliloti* Tn*5* mutants was first digested thoroughly with the sticky-end producing restriction enzyme *Sal*I, and then ligated to adaptors. The adaptor has a 5'-phosphorylated compatible cohesive end for ligating to the overhanging sticky end of genomic DNA fragments produced by *Sal*I enzyme (Figure 6-1 B).

There are 4514 *Sal*I restriction sites in the 6.7 Mb *S. meliloti* genome. The average length of restriction fragments is about 1.5 Kb. Transposon Tn*5* contains a unique *Sal*I restriction site, which is located at 2.7kb to the IS*50*L end (Figure 6-1 A). The size of the unique fragment that carries the left part of Tn*5* and adjacent genomic DNA could range from 2.7 to 5 Kb, which should be able to be amplified by PCR efficiently.

In the first cycle of PCR reaction, transposon primer P_{neo} anneals to the neo^R gene in Tn*5* and a new strand with 3' end complementary to the long overhanging strand of the

adaptor will be produced. The *neo*^{R} gene sequence was chosen as the transposon primer binding site so that only the targeted fragment can be amplified in the first PCR. No genomic DNA fragments other than this will be amplified due to lack of primer binding site. The amine (-NH_2) modification of the 3' end of the shorter strand of the adaptor blocks its extension and further ensures that the complementary strand of the long overhanging strand is extended only from the transposon primer. In the second and subsequent PCR cycles, the adaptor primer was able to anneal to adaptor region and allowed the juncture region to be amplified.

To test this mapping method, a previously characterized *S. meliloti ntrC*::Tn*5* mutant was used (209). After PCR, one single band at a size of little more than 4kb was detected on agarose gel (Figure 6-2). Further sequencing of the PCR product revealed that Tn*5* is inserted between 861st and 862nd base pairs in the *ntrC* ORF. This was the first direct mapping of the site of Tn*5* insertion in this mutant and the result was consistent with the original prediction using the restriction enzyme analysis. This suggests this PCR-based mapping method would be efficient to identify the genes interrupted by transposon Tn*5* in this screen.

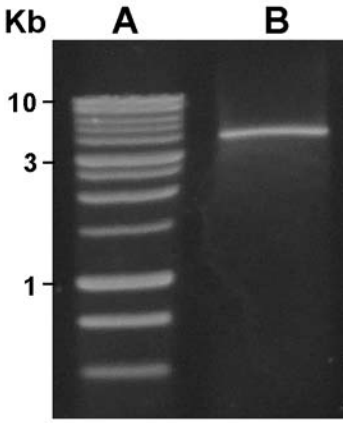

Figure 6-2. The picture of an agarose gel showing molecular weight standard (lane A) and the DNA fragment containing the juncture of Tn*5* and adjacent *ntrC* gene generated from PCR-walking (lane B).

Screening for Tn*5* mutants of *S. meliloti* Rm8396 (pEX312) that suppress succinoglycan-overproduction phenotype of *exoS396*::Tn*5-233* mutation

To identify the unknown genes involved in the regulation of succinoglycan production, especially those related to ExoS/ChvI regulation pathway, *S. meliloti* cells with the *exoS396*::Tn*5-233* mutation background were mutagenized with transposon Tn*5* and then screened for mutants suppressing the succinoglycan-overproduction phenotype of the *exoS396* mutation. The *exoS396*::Tn*5-233* mutation is a mutation generated by the replacing of transposon Tn*5* in the *exoS96*::Tn*5* mutant with transposon Tn*5-233* (47, 64, 229).

S. meliloti has 22 genes involved in succinoglycan production and it may have a few more regulatory genes in addition to the *exoR*, *exoS* and *chvI* genes. Transposon insertion in any one of the genes may block succinoglycan production; therefore, succinoglycan deficient transposon mutants would most likely be *exo* mutants. To reduce the chance of

picking up *exo* mutants, a second set of *exo* genes was introduced into strain Rm8396 (the *exoS396*::Tn*5-233* mutant) on plasmid pEX312 before Tn*5* mutagenesis. To test its effect on succinoglycan production by an *exo* mutant, plasmid pEX312 was conjugally introduced into Rm8396 and SmHC16. SmHC16 is an *exoS396 exoY/lacZ* double mutant and does not produce succinoglycan. When streaked on LB/MC calcofluor plate, both Rm8396 (pEX312) and SmHC16 (pEX312) showed the same high levels of calcofluor fluorescence intensity (data not shown). This result confirmed that the presence of plasmid pEX312 would complement mutations in the *exo* genes cluster and avoid re-identifying the *exo* genes in the screen.

The genes that were expected to be identified in this screen could come from the following three groups. Group I includes the possible genes that are in the ExoS/ChvI regulatory pathway and are regulated by the ExoS/ChvI two-component regulatory system. Mutations of these genes may suppress succinoglycan overproduction by blocking the effect of upregulation of *exo* genes expression by the ExoS396 protein. Group II includes the possible genes that are involved in positively regulating the *exoS* expression. Mutations of these genes may suppress succinoglycan overproduction by reducing or shutting down the *exoS396* expression. Group III includes the possible genes that are on other unknown regulatory pathways to regulate succinoglycan production independent of ExoS. The starting strain Rm8396 (pEX312) overproduces succinoglycan and will be very bright on calcofluor plate. Tn*5* insertions in any of the genes from those three groups will most likely reduce succinoglycan production by Rm8396 (pEX312) and result in dim colonies on LB/MC/CF plates when observed under UV light.

A total 63 calcofluor fluorescence dim colonies were isolated from 4×10^4 colonies of individual Tn*5*-generated mutants of Rm8396 (pEX312) (Figure 6-3). All 63 isolates were purified and confirmed for their ability to reduce succinoglycan production at different levels (part of the data is shown in Figure 6-4).

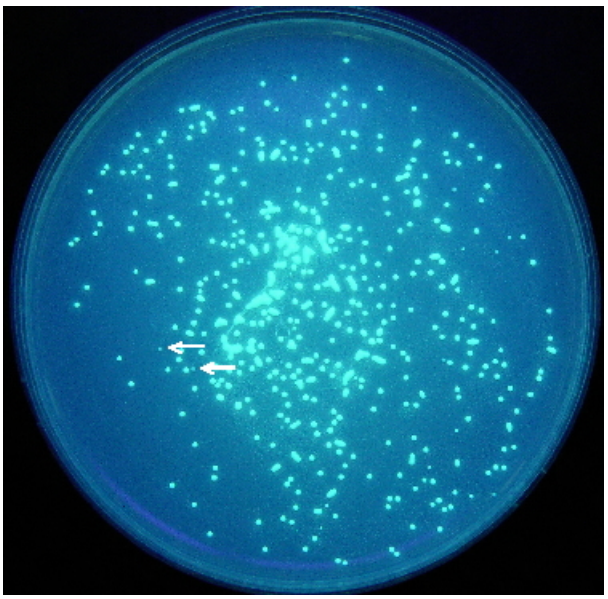

Figure 6-3. A picture of colonies of Tn*5*-generated mutants showing calcofluor fluorescence on selective LB/MC/CF plates. Occasionally dim colonies were observed (see colonies pointed by the arrows), which are composed of Tn*5*-generated mutants suppressing succinoglycan overproduction.

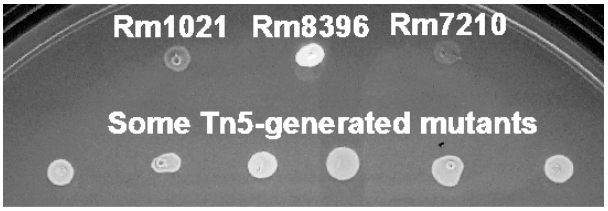

Figure 6-4. A picture of the purified Tn*5*-generated mutant cultures exhibiting different levels of calcofluor fluorescence under UV light.

Unexpectedly, when the Tn*5*-generated mutations from each of the 63 isolates were transduced into Rm8396, all transductants showed the same levels of calcofluor fluorescence intensity as Rm8396 (data not shown). Therefore, the succinoglycan

overproduction-suppressing phenotypes obtained by the 63 isolates were not the result of single Tn*5* insertion in the genome, and the isolates could carry spontaneous mutations or even multiple Tn*5* insertions. These results also suggest that there could be only few genes exist which belong to those three groups and chances to find them are very low. Another explanation is that those genes could have very important functions and disruption of any may be fatal to bacterial cells.

Preliminary characterization of Tn*5* mutants of *S. meliloti* Rm8396 (pEX312) that exhibit altered succinoglycan phenotypes

In addition to the 63 succinoglycan-reducing isolates, several isolates with different phenotypes have also been obtained in the screen. More importantly, transduction tests confirmed that the phenotypes of these isolates were the result of single Tn*5* insertion mutation. These mutants have been further characterized.

Two Tn*5*-generated mutants of Rm8396 (pEX312) showed no calcofluor fluorescence on LB/MC/CF plate under UV light. This suggests that each mutation shuts down succinoglycan production in the *exoS396* mutant background. Because the presence of an additional copy of the *exo* genes (provided by pEX312) in both mutants did not restore succinoglycan production, it is very likely that the mutations come from other unknown loci rather than the *exo* genes in genome. When transduced into wild-type strain Rm1021 background, each mutation led to calcofluor fluorescence dark phenotype. Thus, the Tn*5*-generated mutations might shut down succinoglycan production independent of the

exoS396 mutation. Further identification of the two mutations by mapping the Tn*5* positions might reveal new pathways controlling succinoglycan production.

Nine Tn*5*-generated mutants that showed remarkably higher levels of calcofluor fluorescence intensity even than Rm8396 (pEX312) were isolated. All of them were transduced into Rm8396 and Rm1021. Based on their calcofluor fluorescence phenotypes, Rm8396 carrying each mutation produced more succinoglycan than Rm8396, and Rm1021 carrying each mutation produced more succinoglycan than Rm1021 but less than Rm8396 (data not shown). This suggests that the nine Tn*5*-generated mutations cause upregulation of succinoglycan production independent of the *exoS396* mutation.

Isolate STJ53 is one of the nine Tn*5*-generated mutants. Mapping result revealed that Tn*5* inserted in the ORF of the putative gene *smc00248* (*ccsA*) (Figure 6-5 C). The gene product CcsA is predicted to encode a cytochrome c-type biogenesis protein, which is involved in small molecule metabolism, energy transfer and electron transport. This finding increased the possibility that the succinoglycan production can be regulated at the metabolism levels. Further identification of the other eight mutations may provide more information.

Two Tn*5*-generated mutants, STB1 and STB2, were isolated because of their bright blue calcofluor fluorescence phenotype which is apparent different from the common white calcofluor fluorescence of most *S. meliloti* cells. This phenotype suggests the

succinoglycan structure could be modified by the two Tn*5*-generated mutations in the two mutants. It was also noticed that the colonies of both mutants were more mucoid than that of the *exoR* or *exoS* mutant. Each mutation was successfully transduced into Rm8396 background and kept the same phenotype as the original isolates. However, when transduced into Rm1021 background, transductants of both lost the bright blue calcofluor fluorescence phenotype and did not show apparent difference from Rm1021. These results suggest that the *exoS396* mutation might relate to the new phenotypes of the two Tn*5*-generated mutants. Tn*5*-generated mutation in the isolate STB1 was identified using the One-step PCR method. Mapping result revealed the Tn*5* inserted in the ORF of the putative gene *sma1541* (Figure 6-5 B). The putative *sma1541* is predicted to encode an oxidoreductase, which is involved in small molecule metabolism. There is no report on its involvement in exopolysaccharide biosynthesis. Further analysis of structure of the exopolysaccharide produced by this *sma1541*::Tn*5 exoS396*::Tn*5-233* double mutant by NMR and HPLC may reveal the nature of this blue calcofluor fluorescence producing exopolysaccharide. It is also possible that a different polysaccharide has been produced by this double mutant strain.

Another transducible mutant STJ48 isolated in this screen formed fluorescence-halo lacking colonies on LB/MC/CF plate under UV light. Mapping result revealed that the Tn*5* inserted in the ORF of the putative gene *smc01135* (Figure 6-5 A). This putative gene is predicted to encode protease IV, which is a transmembrane protein involved in degradation of proteins, peptides and glycopeptides. The phenotype of this STJ48 strain is similar to that of the *S. meliloti exoH* mutant, which was reported lacking the halo of

fluorescence around its colonies. The *exoH* mutant produces succinoglycan without succinyl moiety that is normally present and fails to gain entry into alfalfa root nodules (125, 135). It is possible that the *smc011135*::Tn*5* mutation also affects modifications of succinoglycan. However, when the mutation was transduced into Rm1021, the transductants did not show apparent difference from the wild-type strain. It seems that fluorescence-halo lacking phenotype can only be observable in the *exoS* mutation background. Nevertheless, nodulation assay will confirm whether the exopolysaccharide produced in *smc011135*::Tn*5* mutant functions during the infection thread formation. Further analysis of the exopolysaccharide produced by STJ48 mutant using NMR and HPLC might still be helpful to reveal the nature of the fluorescence-halo lacking phenotype.

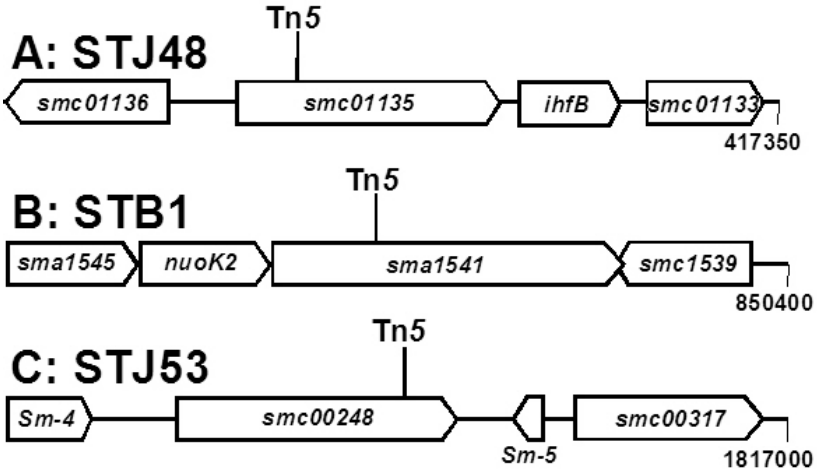

Figure 6-5. Schematic representation of transposon Tn*5* insertion of mutant STJ48 (A), STB1 (B) and STJ53 (C).

Taken together, characterization of the genes identified in this transposon Tn*5* mutagenesis screen may provide more information about the succinoglycan production and the ExoS/ChvI regulatory pathway that regulates succinoglycan production.

APPENDIX

The 113 identified mostly previously unknown *S. meliloti* genes

The *S. meliloti* genomic DNA fragment carried by the promoter trap vector in the nodule isolate J052 covers the upstream region of the *lsrA (smc00037)* gene on the chromosome. One putative promoter in this region was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org), which scored 0.4. A score of 1.0 is an ideal promoter. The *lsrA* gene is predicted to encode a LysR type transcriptional regulatory protein and it has recently been identified as a symbiotic gene by our lab in a loss-of-function mutagenesis screen (140). The presence of the *lsrA* gene is essential for the early stages of nodule development (140). The fact that the promoter of the *lsrA* gene was isolated in my current promoter trap screen suggests that this screen is quite effective.

Three nodule isolates, J003, J046 and J204, carry the identical *S. meliloti* genomic DNA fragments covering the upstream region of the *dgkA (smc04213)* gene. One putative promoter scoring 0.9 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). This gene encodes a diacylglycerol kinase required for recycling L-diacylglycerol generated during cyclic β-1,2-glucan synthesis (150). The presence of cyclic β-1,2-glucan is required for the successful nodulation of alfalfa by *S. meliloti* (27). The possible roles of the *dgkA* gene in symbiosis were further explored through the construction of a *dgkA* null mutant, which is described in the Chapter III.

The nodule isolates, J001 and J009, carry the exact same *S. meliloti* genomic DNA fragment covering the upstream region of the *exsA* gene (14). Two putative promoters were predicted. The putatively upstream promoter scored 0.89, while the putative downstream promoter scored 0.95 using the prokaryotic promoter predicting software (http://www.fruitfly.org). The *exsA* (*smb20941*) locates at one end of the *exs* gene cluster and is adjacent to the *exo* loci on the megaplasmid pSymB. The *exo*/*exs* gene clusters are involved in biosynthesis of exopolysaccharide EPS I (succinoglycan), which is essential for symbiosis establishment. This *S. meliloti exsA* gene codes a Msba-like saccharide exporting ABC transporter protein ExsA, which was reported to be involved in succinoglycan secretion (14, 188). Mutations of *S. meliloti exsA* gene were reported to result in a decreased ratio of HMW to LMW succinoglycan. The closest homologue of *exsA* gene in *S. meliloti* is *ndvA* gene, which is involved in the transport of cyclic β-(1,2)-glucan and is specially expressed in infection threads (14, 28, 67). The identification of *exsA* gene in my current promoter trap screen confirmed the importance of bacterial exopolysaccharides at the early stages of the symbiosis between *S. meliloti* and alfalfa. This also suggests this screen is effective.

The nodule isolates, J095 and J096, carry the exact same *S. meliloti* genomic DNA fragment covering the upstream region of the *flaC (smc03040)* gene. One putative promoter scoring 0.75 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). This gene codes a flagellin protein FlaC. FlaC, like FlaB and FlaD, is one of the secondary flagellin proteins. The principal flagellin protein FlaA and at least one secondary Fla protein are required to assemble a

functional flagellar filament (194). Flagellum synthesis is involved in bacterial mobility chemotaxis in environment. The adjustment of synthesis complex flagellar filaments results in different modes of directional control of swimming cells, which might be important for *S. meliloti* cells to mediate swarming motility and adhesion to host surface in response to the specific condition. These suggest *S. meliloti* accurately regulates cell motility genes at different stages of the symbiosis to adaptor environment changes.

The nodule isolate, J070, carries the *S. meliloti* genomic DNA fragment covering the upstream region of the *fliF (smc03014)* gene. One putative promoter scoring 0.92 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). This gene encodes a flagellar M-ring transmembrane protein FliF. The *S. meliloti fliF* gene has homologs in pathogen bacteria *Brucella melitensis* and *Listeria monocytogenes* (20, 128). *B. melitensis fliF* gene was identified during the acute phase of infection in mice by using signature-tagged mutagenesis (128). Mutants of *B. melitensis fliF* gene results in an attenuated in vivo, which suggests *fliF* gene plays an important role in *B. melitensis* pathogenetic virulence. *L. monocytogenes fliF* gene mutations were reported to impair the bacterial adhesion and entry into nonphagocytic epithelial cells (20). These findings raised the possibility of *S. meliloti* FliF plays important roles in the early stages of colonizing and host alfalfa plant during the symbiosis.

The nodule isolate, J063, carries the *S. meliloti* genomic DNA fragment covering the upstream region of the *degP1* gene. Two putative promoters, scoring 0.69 and 0.54 respectively, were predicted in this region using the prokaryotic promoter predicting

software. The *degP1 (smc02365)* gene locates downstream of the *S. meliloti* symbiosis-essential *cycHJKL* operon (91). DegP1 is a protease precursor, which is involved in degradation of proteins, peptides or glycopeptides. The function of DegP protein is important for the virulence of some pathogen bacteria, such as *B. abortus, Streptococcus pyogenes,* and *Yersinia pestis* (69, 141, 226). *S. meliloti degP* mutations alter the expression of the symbiotic gene *bacA*, which is required for differentiation of *S. meliloti* nitrogen-fixing bacteroids (91). These suggest that DegP might be involved in symbiosis establishment. In the other respect, there are 4 copies of *degP* genes on *S. meliloti* genome. When *S. meliloti* carrying *DegP1* mutation was inoculated with alfalfa plant, pink Fix^+ nodules were elicited on plants that showed visually indistinguishable from those induced by the $degP1^+$ parent strain (91). The identification of the *degP1* gene using this improved IVET approach suggests that this positive screening is efficient to identify redundant symbiotic genes than traditional mutagenesis techniques.

The nodule isolates, J044, J060, J084, and J193, carry the exact same *S. meliloti* genomic DNA fragment covering the upstream region of the *ppdK (smc00025)* gene. Two putative promoters, scoring 0.42 and 0.44 respectively, were predicted in this region using the prokaryotic promoter predicting software. The *ppdK* gene encodes the pyruvate phosphate dikinase PpdK. The *S. meliloti* PpdK protein catalyses the formation of phosphoenolpyruvate (PEP) from pyruvate, which constitutes a gluconeogenic pathway independent of PEP carboxykinase (PCK) (160, 161). While the PCK mutants of *S. meliloti* show a reduced level of nitrogen fixation, *ppdK* mutants do not affect symbiotic nitrogen fixation (160, 161).

The nodule isolates, J022 and J024, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of *thuFGK* operon. The fragment covers the second part of *thuE* (*smb20325)* and the intergenic region upstream region of *thuF (smb20326)* and the first part of *thuF* (*smb20326).* Analysis of the fragment with the prokaryotic promoter predicting software (http://www.fruitfly.org) predicted a putative promoter with a score of 0.65 in the intergenic region between *thuE* and *thuF,* which may direct *thuF, thuG* and *thuK* expression. Both ThuF and ThuG are integral membrane proteins, trehalose/maltose transporter permeases. ThuK is the trehalose/maltose transporter ATP-binding protein. Together with ThuE, the trehalose/maltose-binding protein, the four proteins compose a binding protein-dependent disaccharides ABC transporter. The uptake of disaccharides is critical for *S. meliloti* colonization of alfalfa roots but is not important for nodulation and nitrogen fixation (114). ThuF shares 79% identical amino acids with *Agrobacterium tumefaciens* PalF, which is involved in palatinose transport. Palatinose was reported to be important as an osmo-protectant under hyperosmotic conditions in *S. meliloti* (56). These findings raised the possibility that ThuF, ThuG and ThuE may play important roles at the early stages of the symbiosis.

The nodule isolate, J127, carries the *S. meliloti* genomic DNA fragment, which comes from a putative *glxABCD* operon encoding a set of transcription factor and enzymes involved in glutamine metabolism. The genomic fragment in the promoter trap vector covers most part of putative *glxC (smc02611)* open reading frame, which is the immediate upstream of the *glxD (smc02612)* gene. The *glxD* gene could be transcribed

together with the other three *glx* genes in this operon, but it might also be transcribed separately from the promoter(s) overlapping the *glxC* gene. Three putative promoters, with scores of 0.83, 0.81 and 0.75, respectively, were predicted in the trapped region by analysis with the prokaryotic promoter predicting software (http://www.fruitfly.org). A score of 1.0 is an ideal promoter. These findings increase the possibility that the *glxD* gene could be expressed by one of the three putative promoters. This gene encodes an oxidoreductase protein GlxD, which is the glutamate synthase large subunit-like protein. Glutamate synthase has two functions, the synthesis of glutamate and the removal of glutamine. Glutamine is the primary product of ammonia assimilation during nitrogen-limited growth. In *S. meliloti*, the GlxD protein has a homolog, probable glutamate synthase NADPH large chain protein GltB (Smc04028), by sharing 28% identical amino acids (130). The *S. meliloti gltB* mutants were unable to assimilate ammonia and were still able to elicit pink-nodules on alfalfa seedlings (130).

The genomic DNA fragment in the promoter trap vector in nodule isolate J076 comes from the putative *cbbFPTALSX* (Calvin-Benson-Bassham) operon coding enzymes in the Calvin cycle (90). It is not known whether the gene products of this operon play any role in symbiosis. The genomic fragment in the promoter trap vector covers the second half of the putative *cbbS (smb20197)* gene and the first half of the *cbbX (smb20196*) gene, which is the last gene in the operon. Immediately following the *cbbX* gene, which stops at 204392, is the putative *ppe (smb20195)* gene, a pentose phosphate epimerase gene, which starts at 204392 in the same direction as the *cbbFPTALSX* operon. The *ppe* gene could be transcribed together with the *cbb* operon, but it might also be transcribed separately from

the promoter(s) overlapping the *cbbS* and *cbbX* genes (86). Two possible promoters were predicted in the region that was in the promoter trap vector using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putatively upstream promoter scored 0.9, while the putative downstream promoter scored 0.8. A score of 1.0 is an ideal promoter. These findings raised the possibility that the *ppe* gene is expressed by one of these two promoters.

The nodule isolate, J101, carries the *S. meliloti* genomic DNA fragment in the promoter trap vector covering the upstream region of the *sma0340 (wrbA2)* gene. Two possible promoters were predicted in the region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putatively upstream promoter scores 0.94, while the putative downstream promoter scores 0.59. This *sma0340* gene is predicted to encode a trp-repressor binding protein WrbA2. The putative *S. meliloti* WrbA2 is homologous to *E. coli* WrbA with 64% identical amino acids. *E. coli* WrbA functions to block TrpR-specific transcriptional processes, which might be physiologically disadvantageous in the stationary phase of the bacterial life cycle (99, 228). These findings raised the possibility that the *sma0340 (wrbA2)* might encode a functional protein playing roles during symbiosis.

The nodule isolate, J108, carries the *S. meliloti* genomic DNA fragment covering the upstream region of the *smb21662 (exoI2)* gene. Three putative promoters, with scores of 0.72, 0.52 and 0.89, respectively, were predicted in the trapped region by analysis with the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb21662* is

predicted to encode the putative periplasmic protein ExoI2. Its closest homologue in *S. meliloti*, ExoI, is a characteristic signal peptide protein, which is related to symbiosis-essential exopolysaccharide EPS I (succinoglycan) biosynthesis. ExoI mutant synthesized a reduced amount of succinoglycan (13). These findings raised the possibility that the *smb21662* encoding a functional protein playing important roles during symbiosis.

The nodule isolate, J012, carries *S. meliloti* genomic DNA that covers the upstream region of the putative open reading frame *smc00234 (ppiD).* Two promoters were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). Both promoters scored 0.80. The *smc00234* open reading frame is predicted to encode peptidyl-prolyl cis-trans isomerase PpiD. In *E. coli*, the peptidyl-prolyl cis-trans isomerase PpiD is required for folding of outer membrane proteins, and is activated by CpxAR signal transduction system, which was thought to be important in the expression of the stress response network (54). These findings further raised the chance that the putative *smc00234 (ppiD)* may encode a functional protein playing roles in symbiosis.

The nodule isolates, J062, J140 and J160, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the *smc00289* and its immediate downstream *smc00288* open reading frames. One promoter with a score of 0.92 was predicted in this region. *smc00289* is predicted to encode a transcription regulator, cold shock protein CspA5, which is involved in bacterial adaptation to atypical conditions (169). The putative CspA5 shares high-level (82-69%) identical amino acids with putative cold

shock proteins in rhizobial bacteria *A. tumefaciens, Mesorhizobium loti*, and *Bradyrhizobium japonicum*. These findings raised the possibility that *smc00289* might encode a functional protein for symbiosis. The downstream *smc00288* putative gene is predicted to encode hypothetical signal peptide protein, which function is unknown.

The nodule isolate, J189, carries the *S. meliloti* genomic DNA fragment covering the promoter region of putative *smc00643 (purA)* gene. This *purA* gene is predicted to encode adenylosuccinate synthetase IMP-aspartate ligase PurA, which is required for synthesis of purine ribonucleotide (41). The putative open reading frame *smc00644* is downstream of this *smc00643 (purA)* in the same direction. The start code of the open reading frame *smc00644* is only 67bp behind *smc00643 (purA)* stop code. There is a possibility that *smc00644* is also expressed under the trapped promoter as *smc00643 (purA)*. Using the prokaryotic promoter predicting software (http://www.fruitfly.org), one putative promoter with a score of 0.96 was predicted, which overlaps the end of *smc00643* and the intergenic region ahead of smc00644. It is also possible that *smc00644* is expressed under this high-score putative promoter rather than *smc00643*'s promoter(s) trapped in the vector. *smc00644* is predicted to encode a hypothetic protein which has homologs in *A. tumefaciens, M. loti* and *Brucella suis*. Function of Smc00644 is unknown.

The nodule isolate J093 carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc00695-smc00696-smc00697*. One putative promoter scoring 0.95 was predicted in this region using the prokaryotic promoter

predicting software (http://www.fruitfly.org). *smc00695 (aroK)* is predicted to encode Shikimate kinase AroK. AroK (EC 2.7.1.71) catalyzes the conversion of shikimate into shikimate-3-P, the fifth step in the biosynthesis of chorismate necessary for the biosynthesis of aromatic amino acids (100). The *S. meliloti* AroK protein has a homolog in *Rhizobium etli* with 84% identity. The *R. etli aroK* mutation elicited under-developed nodules and the total number of nodules was low (72). The finding further increased the chance that the *S. meliloti* AroK protein may play roles in the symbiosis. *smc00696 (aroB)* is predicted to encode a putative transmembrane 3-dehydroquinate synthase AroB, which is involved in chorismate synthesis. *smc00697* is predicted to encode a hypothetical transmembrane protein, which function is unknown.

The nodule isolate, J171, carries the *S. meliloti* genomic DNA fragment that covers within the putative open reading frame *smc00810*. Immediate downstream of *smc00810* are the putative open reading frames *smc00811 (FtsJ)* and *smc00812*. A putative promoter with a score of 0.45, was predicted in the trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct expression of *smc00811* and *smc00812*. The *smc00811* open reading frame is predicted to encode FtsJ, which is a probable cell division protein. The hypothetical Smc00812 is a conserved hypothetical protein, which shares 38% identical amino acids with *B. japonicum* USDA 110 putative lyase Blr3793.

The nodule isolates, J031, J036, J124 and J126, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative operon *hmuTUV*, which is

involved in hemin transport. One putative promoter with a score of 0.70 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the expression of the putative operon *hmuTUV*. *smc01512 (hmuT)* is predicted to encode hemin binding transmembrane protein HmuT, *smc01511 (hmuU)* is predicted to encode hemin transport system permease transmembrane protein HmuU, and *smc01510 (hmuV)* is predicted to encode hemin transport system ATP-binding protein HmuV. The roles of *S. meliloti hmuTUV* genes during symbiosis were not reported yet. However, *hmu* genes were reported not to be essential for symbiotic nitrogen fixation of *B. japonicum* or *Rhizobium leguminosarum* (154, 225). On the other hand, *hmuT* and *hmuU* genes were also identified as upregulated genes during plant pathogen *Erwinia chrysanthemi* infection to spinach plant by a similar positive GFP-based IVET approach (227).

The nodule isolate, J182, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc02426* and its downstream putative gene *smc02427 (hyuE)*. Three putative promoters with scores of 0.76, 0.71 and 0.64, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc02426* is predicted to encode Smc02426 containing Flavin reductase like domain. Smc02426 shares 40% identical amino acids with proteobacterium *Desulfuromonas acetoxidans* DSM 684 Flavoredoxin, which might be involved in electron transfer chain (95). The *smc02427* is predicted to encode a hydantoin racemase HyuE, which is involved in amino acid biosynthesis.

The nodule isolate, J164, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative gene *smc02653 (lepB)*. One putative promoter with a score of 0.89 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc02653* is predicted to encode a probable signal peptidase I protein LepB, which is related to cell processes or transport of small molecules (41). *S. meliloti* LepB shares 46% identical amino acids with the *B. japonicum* signal peptidase SipS. *B. japonicum* SipS is specifically involved in the interaction with soybean when the bacteria are released from the infection threads (151). Mutations of *sipS* result in instability of the bacteroids within the symbiosome. These findings further increased the chance that *S. meliloti* LepB may play important roles in symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J107, carries *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc02737 (opuC)*. Two promoters were predicted in this trapped region. The putative upstream promoter scored 0.48, while the putative downstream promoter scored 0.47. *smc02737* is predicted to encode an ABC-type glycine betaine transporter protein OpuC. Glycine betaine was reported to be related to osmotic stress tolerance in *S. meliloti* and many other members of the family Rhizobiaceae (22, 32, 210). These findings raised the possibility that OpuC might play important roles in symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J156, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative *glnKamtB* operon, which composes two putative genes

smc03806 (glnK) and *smc03807 (amtB)*. One putative promoter scoring 0.88 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc03806* is predicted to encode probable nitrogen regulatory protein PII GlnK. *smc03807* is predicted to encode ammonium transporter AmtB. Both of GlnK and AmtB are involved in nitrogen metabolism. *glnKamtB* in *R. etli* are activated under nitrogen-limiting, free-living conditions, but are down-regulated just when bacteria are released from the infection threads (211).

The nodule isolate, J058, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative gene *nodN2 (smc03927)* and its immediate downstream putative open reading frame *smc03928*. Two promoters were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putatively upstream promoter scored 0.41, while the putative downstream promoter scored 0.48. NodN2 shares 64% identical amino acids with putative dehydratase NodN in *S. meliloti*, which is involved in biosynthesis of the glucosamine oligosaccharide signal molecules. *S. meliloti nodN* mutants were reported to produce strongly reduced root hair deformation activity and display delayed nodulation of alfalfa (4, 207). The *nodN2* genes was reported to be under a negative control of the NolR repressor (51). My finding that NodN2 might be induced during the early stages of symbiosis establishment is consistent with the previous finding that NolR repressor is expressed at high level both in the free-living bacterium and in the bacteroid but is downregulated by plant signal flavonoids at the early stages of symbiosis establishment (51). However, further induction test showed nodN2 was not upregulated in the presence of flavonoid, apigenin. The immediate

downstream putative *smc03928* gene encodes a conserved hypothetical protein, which function is unknown. Smc03928 shares 28% identical amino acids with proteobacterium *Rhodobacter sphaeroides* putative phosphogluconate dehydratase Rsph03002781.

The nodule isolate, J186, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frames *smc04048* and *smc04047 (Azu2).* One putative promoter with a score of 0.33 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which may direct the expression *smc04047* and *smc04048*. *smc04047* is predicted to encode copper-binding protein Azu2, which shares 44% identical amino acids with *S. meliloti* Azu1 copper-binding pseudoazurin. *smc04048* is predicted to encode putative cytochrome C protein, which shares 26% identical amino acids with putative cytochrome C oxidase chain II protein SMb21368 in *S. meliloti*.

The nodule isolate, J016, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc04234.* Two promoters with scores of 0.40 and 0.35, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc04234* is predicted to encode cold shock-like transcription factor Csp4, which is involved in bacterial adaptation to atypical conditions.

The nodule isolates, J083 and J088, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc00880-smc00881-*

smc00882-smc00883. Three putative promoters, scoring 0.79, 0.95 and 0.88 respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). The *smc00880* open reading frame is predicted to encode an oxidoreductase protein, which is involved in small molecule metabolism. The putative Smc00880 shares 35-25% identical amino acids with more than 10 putative oxidoreductases in *S. meliloti*. *smc00881* is predicted to encode 2-dehydro-3-deoxygalactonokinase DgoK1, which is involved in synthesis of carbon compounds. *smc00882* is predicted to encode a putative KHG/KDPG aldolase protein, which is involved in small molecule metabolism. The *smc00883* is open reading frame predicted to encode a conserved hypothetical protein, which function is unknown.

The nodule isolates, J007, J008 and J010, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc01522.* Two promoters scoring 0.83 and 0.51, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01522* is predicted to encode a hypothetical protein, which shares 42% identical amino acids with hypothetical *A. tumefaciens* TetR family transcriptional regulator protein Atu2320 (173). The open reading frame of its downstream putative gene *smc01523 (emrE)* is 51bp away from *smc01522* stop code, and may be expressed under the same promoter. The *smc01523* open reading frame is predicted to encode a transmembrane multi-drug transporter.

The nodule isolate, J148, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc00072*. Two putative promoters with scores of 0.87 and 0.58, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the *smc00072* expression. *smc00072* is predicted to encode a peroxiredoxin protein. Peroxiredoxin of *R. etli* was reported to be strongly expressed under microaerobic conditions and during the symbiosis interaction with plant *Phaseolus vulgaris,* which defenses *R. etli* bacteroids against oxidative stress (65). These findings raised the possibility that the *smc00072* might encode a functional protein for symbiosis.

The nodule isolates, J125, J141 and J208, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the open reading frame *smc00371*. One promoter with a score of 0.22 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc00371* is predicted to encode a conserved hypothetical protein Smc00371, which shares homology with stress response protein YciF in different bacteria (112, 132). The expression of *smc00371* was reported to increase inside bacteroids compared to in free-living cells (9).

The nodule isolate, J184, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc00570*. One putative promoter with a score of 0.34 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc00570* is predicted to encode an oxidoreductase protein. Smc00570 shares 32% identical amino acids with the *S. meliloti* MocA protein, which is

involved in the uptake and subsequent degradation of rhizopine (189). These findings raised the possibility that *S. meliloti smc00570* might encode a functional protein for symbiosis.

The nodule isolates, J136, J137 and J138, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc00722*. Two putative promoters with scores of 0.22 and 0.21, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc00722* is predicted to encode a transmembrane protein Smc00722. Smc00722 shares 25% identical amino acids with proteobacterium *Ralstonia eutropha* putative chemotaxis sensory transducer Reut_A2620. In *S. meliloti*, chemotaxis sensory pathway plays roles to guide the bacteria to the localized sites on the surface of alfalfa root tips during symbiosis (17). These findings raised the possibility that *smc00722* might encode a functional protein playing roles in the early stages of the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J185, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc00856.* Two putative promoters were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), both of which scored 0.45. *smc00856* is predicted to encode a transmembrane protein, which has homologs with 58-54% identical amino acids in rhizobial bacteria *A. tumefaciens*, *M. loti* and *Mesorhizobium sp.* BNC1.

The nodule isolates, J002, J005 and J106, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc00977-smc00976*. Two putative promoters with scores of 0.94 and 0.67, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc00977* is predicted to encode an acyl-coA dehydrogenase, which is involved in fatty acid degradation. *smc00976* is predicted to encode an enoyl-CoA hydratase, which is involved in fatty acid metabolism.

The nodule isolate, J123, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc01214*. Two putative promoters with scores of 0.98 and 0.89, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01214* is predicted to encode a putative zinc-containing alcohol dehydrogenase protein, which is involved in small molecule metabolism. *smc01214* is a homolog of the *S. meliloti tdh* (*smc01564*) gene, which encodes a probable threonine 3-dehydrogenase protein in *S. meliloti*.

The nodule isolate, J177, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc01488*. One putative promoter scoring 0.84 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01488* is predicted to encode a hypothetical transmembrane protein, which shares 41% identical amino acids with pathogen bacterium *B. melitensis* putative zinc-finger protein BMEI1461.

The nodule isolate, J048, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc01571-smc01572-smc01577*. Two putative promoters scoring 0.65 and 0.48, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The *smc01571* and *smc01572* open reading frames are predicted to encode two oxidoreductases. Smc01571 has more than 20 homologs in *S. meliloti*. Smc01572 has eight homologs in *S. meliloti*. *smc01577* is predicted to encode a conserved hypothetical protein, which shares 40-30% identical amino acids with several putative dehydrogenases in different bacteria.

The nodule isolate, J102, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc01731*. One promoter with a score of 0.56 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical Smc01731 shares 76% identical amino acids with the *A. tumefaciens* probable DapE protein, which catalyses the synthesis of L-diaminopimelic acid, one of the last steps in the diaminopimelic acid-lysine pathway, which is related to bacterial cell wall biosynthesis (23, 25).

The nodule isolate, J110, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc01759*. Two promoters scoring 0.97 and 0.86, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01759* is predicted to encode a conserved hypothetical protein, which shares 47% identical amino acids with *B.*

japonicum USDA 110 hypothetical protein Bsl1405. The function of Smc01759 is unknown.

The nodule isolates, J013, J014 and J114, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative gene *smc01817*. Four putative promoters with scores of 0.80, 0.69, 0.84 and 0.96, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The *smc01817* open read frame is predicted to encode a putative LysR-family transcriptional regulator. When *S. meliloti* strain carrying the mutation of *smc01817* were inoculated to alfalfa seedlings, only 77% nodules elicited were pink Fix+ nodules, while wild-type strain elicited 100% pink Fix+ nodules (140). This finding further raised the possibility that the *smc01817* might encode a functional protein for symbiosis.

The nodule isolate, J033, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc01945*. Two putative promoters were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which scored 0.73 and 0.92 respectively. *smc01945* is predicted to encode a transcription regulator, which shares 47% identical amino acids with *Bacillus licheniformis* OhrR. OhrR is a peroxide sensor and acts as a repressor of the organic peroxide resistance gene *ohrA* (80, 81).

The nodule isolates, J115 and J202, carry the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc01957*. One promoter

with a score of 0.55 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01957* is predicted to encode a conserved hypothetical protein, which function is unknown. It shares 28% identical amino acids with the putative *Streptomyces coelicolor* transcriptional regulatory protein Sco2105.

The nodule isolates, J043, J145 and J146, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc01988*. One promoter with a score of 0.61 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc01988* is predicted to encode a hypothetical protein, which shares 28% identical amino acids with proteobacterium *Anaeromyxobacter dehalogenans* hypothetical spermine synthase AdehDRAFT_0174. The function of Smc01988 is unknown.

The nodule isolate, J018, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc02151*. Two putative promoters with scores of 0.73 and 0.57, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical protein Smc02151 is predicted to be virulence associated protein homolog, which shares 50% identical amino acids with animal pathogen bacterium *Dichelobacter nodosus* virulence associated protein VapA (116).

The nodule isolate J072 carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc02171*. Two promoters were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putative upstream promoter scored 0.9, while the putative downstream promoter scored 0.7. *smc02171* is predicted to encode a periplasmic binding protein of an ABC transporter, which is typically involved in the transport of small molecules. Smc02171 shares high levels of homology with the *A. tumefaciens* C58 FrcB with 85% identical amino acids, and the *M. loti* MAFF303099 hypothetical protein Mlr7582 with 80% identical amino acids. The finding of these two close homologs further increased the chance that the *S. meliloti smc02171* open reading frame encodes a functional protein that might play roles in symbiosis.

The nodule isolate, J069, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc02381*. One promoter with a score of 0.97 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc02381* is predicted to be conserved hypothetical protein, which function is unknown. Its closest homolog in *S. meliloti,* hypothetical protein Smb21289, is predicted to interact with ubiquitin. Smc02381 also shares 62% identical amino acids with *B. japonicum* MarR family transcriptional regulatory protein Bll5370.

The nodule isolate, J111, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of a putative operon *smc02619-smc02618-smc02617-smc02616.* One putative promoter scoring 0.99 was predicted in this region using the prokaryotic

promoter predicting software (http://www.fruitfly.org). The *smc02619, smc02618* and *smc02617* open reading frames are predicted to encode three hypothetical transmembrane proteins, which functions are unknown. *smc02616* is predicted to encode a putative transmembrane permease, which may be involved in transport of small molecules.

The nodule isolate, J155, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc02818-smc02817*. Two promoters scoring 0.98 and 0.60, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical Smc02818, a conserved hypothetical protein, shares 47% identical amino acids with pathogen bacterium *B. melitensis* lysophospholipase L2 (EC 3.1.1.5), which is involved in small molecule metabolism. *smc02817* is predicted to encode a transmembrane protein, which shares 40% identical amino acids with *B. melitensis* hypothetical DME (drug/metabolite exporter) family transporter protein BMEII0677 (113).

The nodule isolates, J027 and J153, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smc02988-smc02987*. One promoter with a score of 0.91 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). Smc02988 and Smc02987 share 44% and 46% identical amino acids with the two gene products of putative operon *smc02716-smc02715* in *S. meliloti,* respectively. Smc02987 also shares 28% identical amino acids with a PilT protein of *M. sp*. BNC1, which might be involved in the retraction of bacterial type IV pili (52). The expression of *smc02988-smc02987* was

reported to increase above 2-folds when symbiotic gene *nodD3* is overexpressed in *S. meliloti* (9). These findings raised the possibility that *smc02988* and *smc02987* might encode functional proteins playing roles at the early stages of the symbiosis.

The nodule isolates, J055, J100 and J103, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc03130.* One putative promoter scoring 0.82 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc03130* is predicted to encode a hypothetical protein, which shares 77% identical amino acids with *M. sp.* BNC1 hypothetical protein MesoDRAFT_0565, and 71% identical amino acids with *A. tumefaciens* str. C58 hypothetical protein Atu3565. The function of Smc03130 is unknown.

The nodule isolates, J086 and J187, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc03746*. Two promoters scoring 0.89 and 0.96, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical Smc03746 protein shares 35% identical amino acids with *Pseudomonas fluorescens* glycine cleavage system T protein GcvT. The glycine cleavage (GCV) enzyme system catalyzes the oxidative cleavage of glycine to carbon dioxide, ammonia, and is induced by glycine (89, 205).

The nodule isolate, J183, carries the *S. meliloti* genomic DNA fragment that covers the upstream region and the first part of the putative open reading frame *smc03767*. Two promoters scoring 0.91 and 0.87, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc03767* is predicted to encode a conserved hypothetical protein. Smc03767 shares 25% identical amino acids homology with *E. coli* probable arylsulfatase regulator protein AslB.

The nodule isolate, J191, carries the *S. meliloti* genomic DNA fragment that covers the upstream region and the first part of the putative open reading frame *smc03844*. Two promoters scoring 0.65 and 0.60, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc03844* is predicted to encode a conserved hypothetical protein, Smc03844, which shares 44% identical amino acids with pathogen bacterium *Vibrio vulnificus* CMCP6 hypothetical signal transduction histidine kinase VV12578.

The nodule isolates, J035 and J113, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc03890*. One putative promoter with a score of 0.85 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smc03890* expression. *smc03890* is predicted to encode a transcription regulator. Smc03890 shares high-level homology with DeoR family transcriptional regulators in many different bacteria, which are involved in sugar metabolism (53).

The nodule isolate, J181, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc04054*. Two promoters scoring 0.80 and 0.53, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc04054* is predicted to encode a conserved hypothetical protein Smc04054, which shares 52% identical amino acids with *M. sp*. BNC1 putative glyoxalase MesoDRAFT_4496. The expression of *smc04054* was reported to increase inside bacteroids compared to in free-living cells (9), which raised the possibility that *smc04054* might encode a functional protein playing roles in symbiosis.

The nodule isolate, J168, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc04194.* Three promoters scoring 0.87, 0.57 and 0.53, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc04194* is predicted to encode a putative transmembrane protein. No homolog was found in other bacteria. Upstream gene of *smc04194* is repeat region *sm-4,* while downstream gene of *smc04194* is transposase for insertion sequence element *trm3*. The expression of *smc04194* was reported to increase inside bacteroids compared to in free-living cells (9), which raised the possibility that *smc04054* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J139, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc04311*. One putative promoter

scoring 0.93 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc04311* is predicted to encode a hypothetic transmembrane protein, which function is unknown. It shares 57-52% identical amino acids with putative periplasmic binding transmembrane protein Smc02378, putative periplasmic binding protein Sma1729 and putative amino acid uptake ABC transporter periplasmic solute-binding protein precursor Smb21572 in *S. meliloti*.

The nodule isolates, J073 and J079, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smc04391*. One possible promoter with a score of 0.6 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which may direct *smc04391* expression. *smc04391* is predicted to encode oxidoreductase, which is involved in small molecule metabolism.

The nodule isolates, J019 and J020, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smc04439.* One putative promoter with scores of 0.78 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smc04439* is predicted to encode a glycine betaine transport ATP-binding protein. It shares 46% identical amino acids with putative glycine betaine transport ATP-binding ABC transporter protein OpuA. Glycine betaine was reported to be related to osmotic stress tolerance in *S. meliloti* and many other members of the family Rhizobiaceae (22, 32, 210). These findings raised the possibility

that *smc04439* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolates J004 and J071, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma0126.* One putative promoter with a score of 0.96 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *sma0126* expression. *sma0126* is predicted to encode a putative cold shock protein (CSP) regulator. Sma0126 shares 85-70% identical amino acids with more than 10 probable cold shock proteins in *S. meliloti, M. loti* MAFF303099, *Rhizobium sp.* NGR234 and *A. tumefaciens*. These findings raised the possibility that *sma0126* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolates J119 and J150, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *sma0232.* Two putative promoters with scores of 0.50 and 0.40, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *sma0232* expression. *sma0232* is predicted to encode a hypothetical protein, which shares homology with more than 10 putative transcriptional regulators in different bacteria. The expression of *sma0232* was reported to increase inside bacteroids compared to in free-living cells (9). These findings raised the possibility that *sma0232* may encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolates, J121, J154 and J165, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma0254*. Two putative promoters scoring 0.65 and 0.91, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical Sma0254 protein shares 91% identical amino acids with *A. tumefaciens* str. C58 hypothetical protein AGR_L_2570p, 69% identical amino acids with *B. japonicum* USDA 110 hypothetical protein Bll4163, and 62% identical amino acids with *M. loti* MAFF303099 hypothetical protein Mlr0494. These findings raised the possibility that *sma0254* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J167, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma0447*. One putative promoter scoring 0.91 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The hypothetical protein Sma0447 shares 22% identical amino acids with *S. meliloti* putative permease transmembrane transport protein Smc02160, 32% identical amino acids with *A. tumefaciens* str. C58 putative permease Atu1739, 32% identical amino acids with *B. japonicum* USDA hypothetical protein Blr8064. The function of Sma0447 is unknown.

The nodule isolates, J026 and J056, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *sma0537-sma0538*. One

putative promoter scoring 0.76 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma0537* is predicted to encode hypothetical protein Sma0537, which shares 34% identical amino acids with *Symbiobacterium thermophilum* IAM 14863 putative peptidase STH913. *sma0538* is predicted to encode a hypothetical protein. Sma0538 shares 27% identical amino acids with *M. loti* MAFF303099 hypothetical protein Mlr0200. The functions of Sma0537 and Sma0538 are unknown.

The nodule isolate, J067, carries the *S. meliloti* genomic DNA fragment covering the upstream region of putative open reading frame *sma0945*. Three putative promoters with scores of 0.75, 0.72 and 0.42, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the putative *sma0945* gene expression. The putative *sma0945* gene is predicted to encode a hypothetical protein, which shares 27% identical amino acids with proteobacterium *R. eutropha* hypothetical permease Reut_A1129, which belongs to the drug/metabolite transporter (DMT) superfamily (113). The function of Sma0945 is unknown.

The nodule isolate, J021, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma0969*. Two putative promoters were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), both of which scored 0.98. *sma0969* is predicted to encode a response regulator of two-component regulatory system. Sma0969 shares 28% identical amino acids with *S. meliloti* phosphate regulon transcriptional regulatory protein PhoB. The *S. meliloti* PhoB and PhoR proteins constitute a two-component regulatory system,

regulating phosphate assimilation system *phoCDET* (7, 8, 45, 159). The phosphate assimilation system *phoCDET* is required for the early stages of the infection process during symbiosis. The mutations on the phosphate assimilation system block the nodule development in the early infection process before the release of bacteria from the infection threads(45, 67). The *S. meliloti phoB* gene was also required for the activation of *exp* gene expression under phosphate limitation, so PhoB also controls the exopolysaccharide EPSII galactoglucan production in *S. meliloti* (152, 190). These findings raised the possibility that *sma0969* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The promoter trap vector in nodule isolate, J038, carries the *S. meliloti* genomic DNA fragment covering the upstream region of the putative open reading frame *sma1124*. Two putative promoters with scores of 0.63 and 0.45, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the putative gene *sma1124* expression. The putative *sma1124* gene is predicted to encode a hypothetical signal peptide protein, which function is unknown. It shares 43% identical amino acids with *M. sp.* BNC1 hypothetical HlyD-family secretion protein MesoDRAFT_2732.

The genomic DNA fragment in the promoter trap vector carried by nodule isolate J053 comes from the putative operon *sma1746-sma1745-sma1742-sma1741* encoding enzymes of a putative iron uptake system. The genomic fragment in the promoter trap vector covers the second half of *sma1741,* which is the last putative gene in the operon.

Immediately following *sma1741*, is the putative open reading frame *sma1740*, which in the same direction as the putative operon *sma1746-sma1745-sma1742-sma1741*. The *sma1740* could be transcribed together with this putative operon, but it might also be transcribed separately from the promoter(s) overlapping *sma1741*. Two putative promoters with scores of 0.43 and 0.37, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the expression of the putative open reading frame *sma1740*. *sma1740* is predicted to encode a conserved hypothetical protein, which shares 36% identical amino acids with *M. sp*. BNC1 hypothetical siderophore-interacting protein MesoDRAFT_0445 and shares 34% *A. tumefaciens* str. C58 hypothetical iron-chelator utilization protein Atu2454.

The nodule isolates, J116, J161, J162 and J198, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *sma1760-sma1761*. One putative promoter scoring 0.27 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma1760* is predicted to encode a protein carrying Serine/Threonine protein kinase catalytic domain. In *S. meliloti,* Sma1760 shares 32% identical amino acids with putative enzyme with aminotransferase class-III domain protein Smb20973, 25% identical amino acids with putative aminotransferase protein Smc00677 and 28% putative aminotransferase protein Smc00675, which are involved in small molecule metabolism. The putative *sma1761* gene is predicted to encode a putative aminotransferase, which is also involved in small molecule metabolism. These findings raised the possibility that the *sma1760* and

sma1761 might encode functional proteins playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J011, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma1898.* Two putative promoters scoring 0.34 and 0.31, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putative gene product Sma1898 shares 54-34% identical amino acids with more than 10 putative OsmC-like proteins in different bacteria. The *osmC* gene and its homologues are osmotically inducible (2, 98, 102, 199). These findings raised the possibility that *sma1898* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J135, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma1921.* Three putative promoters scoring 0.86, 0.82 and 0.63, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma1921* is predicted to encode a putative protein, which shares 29% identical amino acids homology with the human pathogen bacterium *Corynebacterium jeikeium* peptide methionine sulfoxide reductase MsrA. MsrA catalyzes the reduction of oxidized methionine residues of proteins, a modification leading to functional inactivation of numerous proteins (149). The expression of *sma1921* was reported to increase inside bacteroids compared to in

free-living cells (9). These findings further raised the chance that *sma1921* might encode a protein that plays important roles in the symbiosis.

The nodule isolate, J099, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative gene *sma1954*. Three putative promoters with scores of 0.78, 0.76 and 0.70, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma1954* encodes a putative LysR-family transcriptional regulator. When *S. meliloti* bacteria carrying the mutation of the *sma1954* gene were inoculated to alfalfa seedlings, only 97% nodules elicited were pink Fix+ nodules, while the wild-type strain always elicited 100% pink Fix+ nodules (140). This finding further raised the possibility that *sma1954* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J151, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma2049.* One putative promoter with a score of 0.89 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma2049* open reading frame is predicted to encode a putative LacI-family transcription regulator. Bacterial LacI-family transcription regulators respond to the presence of small effector molecules (212). In *S. meliloti*, the *sma2049* shares 36-23% identical amino acids with 6 other putative open reading frames. These findings raised the possibility that the *sma2049* might encode a functional protein for symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J059, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma2199*. Two putative promoters with scores of 0.68 and 0.51, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma2199* is predicted to encode a putative periplasmic solute-binding ABC transporter protein, which is involved in transport of small molecules. Sma2199 shares 31% identical amino acids with putative amino acid-binding periplasmic ABC transporter protein Smc03864, 27% identical amino acids with probable amino acids, amines or peptides uptake ABC transporter periplasmic solute-binding protein precursor SMb20976 and 25% identical amino acids with SMb20706 in *S. meliloti*. It also shares 30% identical amino acids with *M. loti* MAFF303099 putative ABC transporter, periplasmic amino acid-binding protein Mlr5654. These findings raised the possibility that the *S. meliloti sma2199* might encode a functional protein playing roles in the symbiosis establishment between *S. meliloti* and alfalfa.

The nodule isolate, J179, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *sma2235*. One putative promoter scoring 0.54 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma2235* encodes a hypothetical protein, which function is unknown. It shares 22% identical amino acids with *Bacillus anthracis* str. A2012 putative Beta-lactamase class A protein Bant_01002356.

The nodule isolates, J032, J158 and J159, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *sma2239-sma2241*. Two

putative promoters scoring 0.74 and 0.56, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *sma2239* is predicted to encode a hypothetical protein Sma2239, which shares 62% identical amino acids with *M. loti* MAFF303099 putative ABC transporter Mlr7981. *sma2241* is predicted to encode a hypothetical protein Sma2241. Sma2241 shares 26% identical amino acids with *Pseudomonas aeruginosa* aromatic-amino-acid aminotransferase TyrB, which is involved in aromatic biosynthesis (101).

The nodule isolates, J074, J075, J087 and J120, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smb20156.* Two putative promoters with scores of 0.78 and 0.58, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the putative *smb20156* gene expression. *smb20156* is predicted to encode an ABC transporter ATP-binding protein, which may be involved in transporting small molecules.

The genomic DNA fragment in the promoter trap vector in nodule isolate, J034, covers the promoter region of the *smb20162*, a putative *luxR* family transcription factor (165). Two promoters were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putative upstream promoter scored 0.97, while the putative downstream promoter scored 0.69. The hypothetical gene product Smb20162 shares homology with the soil bacterium *Paracoccus denitrificans* FlhR protein with 41% identical amino acids, which is the regulator of the two-component regulatory system

FlhRS. The FlhRS two-component regulatory system is involved in the regulation of formaldehyde, methanol, and methylamine oxidation (104). Smb20162 also shares homology with *R. sp*. NGR234 putative transcription regulator RngR00478 with 39% identical amino acids and *B. japonicum* USDA 110 putative two-component response regulator Bll0331 with 36% identical amino acids. As a putative LuxR family quorum-sensing transcription factor, Smb20162 may play roles in the early stages of the symbiosis establishment between *S. meliloti* and alfalfa.

The promoter trap vector in nodule isolate, J172, carries the *S. meliloti* genomic DNA fragment covering the second part of the putative open reading frames *smb20178* and the first part of the putative open reading frames *smb20179*. Two putative promoters with scores 0.63 and 0.60, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20179* and its downstream putative gene *smb20180* expression. The *smb20179* and *smb20180* open reading frames are predicted to encode hypothetical signal peptide proteins. Smb20179 has many high-identity homologs in different bacteria including rhizobial bacteria and pathogen bacteria, but its functions are unknown. Smb20180 shares 26% identical amino acids with Cyanobacterium *Synechococcus elongatus* PCC6301 blue-copper-protein-like protein Syc1570_d.

The nodule isolates, J082 and J090, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smb20243.* One possible promoter with a score of 0.80 was predicted in this region using the

prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20243* expression. *smb20243* is predicted to encode the putative glycosyltransferase protein Smb20243, which is involved in surface polysaccharides or antigens synthesis. Smb20243 shares 32% identical amino acids with *Pseudomonas aeruginosa* PAO1 glycosyltransferase WbpY, which is required for *P. aeruginosa* A-band lipopolysaccharide (LPS) biosynthesis (34, 57). *P. aeruginosa* A-band lipopolysaccharide (LPS) is involved in the first steps of bacterial adherence leading to host colonization (57). These findings raised the possibility that *smb20243* might encode a functional protein playing roles in the early stages of the symbiosis.

The nodule isolate, J112, carries the *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smb20278*. Two putative promoters with scores of 0.96 and 0.63, respectively, were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20278* expression. *smb20278* is predicted encode a hypothetical protein, which shares 43-25% identical amino acids with peroxidase-related alkylhydroperoxidase AhpD in different bacterial genera.

The nodule isolates, J188 and J195, carries the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smb20399.* Two promoters scoring 0.90 and 0.85, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20399* expression. *smb20399* is predicted to encode a putative protein,

which has homologs in Rhizobial bacteria *M. loti, A. tumefaciens*. However, the function of Smb20399 and its homologs are unknown.

The nodule isolates, J025, J050 and J068, carry the exact same *S. meliloti* genomic DNA fragment locates in the second part of the open reading frame *smb20449* and the first part of its immediately downstream open reading frame *smb20450*. Four putative promoters with scores of 0.91, 0.82, 0.67 and 0.61, respectively, were predicted in this trapped region ahead of *smb20450* using the prokaryotic promoter predicting software (http://www.fruitfly.org). Some of these putative promoters might direct the expression of *smb20450* and its immediately downstream putative open reading frames, *smb20451* and *smb20452*. The putative *smb20450* gene is predicted to encode a regulatory protein, while putative *smb20451* and *smb20452* genes are predicted to encode two conserved hypothetical proteins, which functions are unknown. The putative Smb20450 shares 86% identical amino acids with *R. sp.* NGR234 sigma factor SigB regulation protein RsbU and 26% identical amino acids with *Bacillus subtilis* sigma factor SigB regulation protein RsbU (1, 60, 206). The RsbU activates SigB transcription factor, which responds to signals of environmental and metabolic stress by inducing over 40 general stress genes in *B. subtilis* (1, 60). These findings raised the possibility that *smb20450* might encode a functional protein activated for the symbiosis establishment between *S. meliloti* and alfalfa.

The genomic DNA fragment in the promoter trap vector carried by nodule isolate J169 locates within a putative operon *smb20458-sm20b459-smb20460-smb20461-smb20462-*

smb20463. The *smb20458, sm20b459, smb20460* and *smb20462* open reading frames are predicted to encode putative dTDP-glucose 4,6-dehydratase protein, putative UDP-glucose 4-epimerase protein, putative cellulose synthase protein, and putative cellulase H precursor protein, respectively. These proteins are involved in surface polysaccharides synthesis and macromolecule metabolism. *smb20461* and *smb20463* are predicted to encode hypothetical proteins with signal sequences. The genomic fragment in the promoter trap vector covers the second half of the putative *smb20462* open reading frame and the first part of *smb20463*. Five promoters were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). Three of the putative promoters locating ahead of *smb20463,* all with a score of 0.34, might direct *smb20463* expression. The putative gene product Smb20463 (169Aa) shares 42% identical amino acids with hypothetical protein Smc02017 in *S. meliloti*. It also shares 76-40% identical amino acids with 7 other hypothetic proteins in *M. loti, A. tumefaciens, R. sp. NGR234* and *Brucella abortus.* Two of the putative promoters with scores of 0.98 and 0.46, respectively, locate within the *smb20463* open reading frame. Two open reading frames *smb20463S1* and *smb20463S2* downstream of the two putative promoters might be expressed under the two putative promoters. The *smb20463S1* uses the same reading frame as *smb20463*. Its gene product Smb20463S1 (107Aa) is the same as Smb20463 but 62Aa shorter in N-terminal. The *smb20463S2* gene product is an 89Aa hypothetical protein, sharing low homology with hypothetical sucrose phosphorylase in cyanobacterium *Synechococcus sp.* WH8102.

The nodule isolate, J017, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20476.* Two promoters were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putatively upstream promoter scored 0.72, while the putative downstream promoter scored 0.64. *smb20476* is predicted to encode periplasmic dipeptide-binding ABC transporter protein Smb20476, which may play roles in uptaking peptide from environment. Smb20476 shares 33% and 32% identical amino acids with *Bacillus firmus* dipeptide transporter protein DppA and *E. coli* DppA, respectively. DppA functions as dipeptide permease in bacteria. In *E. coli,* bacterial cell transcripts *dppA* and alters its dipeptide transport capabilities in response to an environmental signal (158). For some pathogen bacteria of *Streptococci*, the oligopeptide uptake systems are involved in host infection (193, 220). These findings raised the possibility that *smb20476* might encode a functional protein playing roles in the early stages of the symbiosis.

The nodule isolate, J170, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20498.* One putative promoter with a score of 0.37 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb20498* is predicted to encode a hypothetical protein, which shares homology with *Silicibacter sp.* M1040 hypothetical protein deoxyribose-phosphate aldolase RoseDRAFT_3695 with 74% identical amino acids and with the *A. tumefaciens* C58 hypothetical aldolase Atu3875 with 74% identical amino acids.

The nodule isolates, J080 and J081, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20504*. Two putative promoters with score 0.81 and 0.94, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20504* expression. *smb20504* is predicted to encode a putative ABC transporter periplasmic sugar-binding protein Smb20504. Smb20504 shares 28-23% identical amino acids with 11 other hypothetic proteins in *S. meliloti*.

The nodule isolate, J178, carries the *S. meliloti* genomic DNA fragment locating within the putative gene *smb20592 (rpo),* which is predicted to encode a putative RNA polymerase sigma factor *rpo*. The putative promoters carried in the trap promoter are in opposite direction to the open reading frame *smb20592.* Two open reading frames, *smb20591* and *smb20590*, are immediately located upstream of the *smb20592* in opposite direction, which might are under the direction by the trapped putative promoters. One putative promoter with a score of 0.92 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20591* and *smb20590* expression. The *smb20591* and *smb20590* open reading frames are predicted to encode hypothetical proteins Smb20591 and Smb20590. Smb20591 *s*hares 40% identical amino acids with proteobacterium *Magnetospirillum magnetotacticum* MS-1 hypothetical FAD/FMN-containing dehydrogenases Magn03005788. Smb20590 shares 58% identical amino acids with putative *M. loti* MAFF303099 acylphosphatase Mll7735. These findings raised the possibility that the

smb20591 and *smb20590* encode functional proteins which play roles during the symbiosis.

The nodule isolate, J006, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20600*. One putative promoter scoring 0.91 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb20600* is predicted to encode a hypothetical protein Smb20600, which shares 29% identical amino acids with pathogen bacterium *Bartonella quintana* putative Electron transfer flavoprotein alpha-subunit BruAb2_0445.

The nodule isolate, J045, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative operon *smb20724-smb20725*. Two putative promoters with scores of 0.54 and 0.99, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb20724* is predicted to encode a conserved hypothetical exported protein, which shares 33-27% identical amino acids with conserved hypothetical proteins Sma1927, Smc00499 and Smb20771 in *S meliloti*. The putative *smb20725* gene is predicted to encode a hypothetical membrane protein, which shares 29% identical amino acids with conserved hypothetical membrane protein Smb20770.

The nodule isolates, J104, J105 and J117, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20861*. One putative promoter with a score of 0.79 was predicted using the prokaryotic promoter

predicting software (http://www.fruitfly.org). *smb20861* is predicted to encode a putative oxidoreductase FAD flavoprotein Smb20861, which is involved in small molecules' metabolism. Smb20861 shares 45-25% identical amino acids with more than 10 hypothetic proteins in *S. meliloti, B. japonicum, R. sp.* NGR234*, M. sp.* BNC1*,* and *R. etli*.

The genomic DNA fragment in the promoter trap vector in nodule isolate J174 comes from the putative operon *smb20886-smb20887-smb20888* coding hypothetical proteins, which functions are unknown. The genomic DNA fragment covers the second half of the putative *smb20887* open reading frame. One putative promoter with a score of 0.55 was predicted in this trapped region ahead of the *smb20888* open reading frame using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb20888* expression. *smb20888* is predicted to encode a conserved hypothetical protein, which shares homology with more than 10 putative transglutaminase family proteins in different bacteria.

The nodule isolate, J030, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb20910* and its downstream open reading frame *smb20909.* Three putative promoters with scores of 0.83, 0.91 and 0.64, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The two putative genes encode two hypothetical proteins Smb20910 and Smb20909. Smb20910 shares 78% identical amino

acids with *S meliloti* hypothetical protein Smc02316. No homolog of Smb20909 has been found in *S. meliloti* or other bacteria.

The nodule isolate, J197, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21157*. Two putative promoters were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which scored 0.96 and 0.83 respectively. The putative *smb21157* gene is predicted to encode a putative translation initiation inhibitor Smb21157. Smb21157 shares 54-27% identical amino acids with more than 10 hypothetic proteins in *A. tumefaciens, M. loti*, *R. sp.* NGR234, *B. japonicum*. These findings raised the possibility that *smb21157* might encode a functional protein for symbiosis. Its downstream open reading frame *smb21158,* only 24bp away from *smb21157,* may also be expressed under the same promoter. *smb21158* is predicted to encode a transcriptional regulator, probably involved in sugar phosphate metabolism. However, only the expression of *smb21158,* not *smb21157,* was reported to be increased in bacteroids compared to in free-living cells (9).

The nodule isolate, J206, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21202*. One putative promoter with a score of 0.47 was predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb21202* is predicted to encode a conserved hypothetical protein Smb21202, which shares 77% identical amino acids with *A. tumefaciens* hypothetical aromatic compounds dioxygenase Atu2453*,* 75% identical

amino acids with *M. loti* MAFF303099 hypothetical glyoxalase and bleomycin resistance protein MesoDRAFT_4275. The expression of *smb21202* was reported to increase inside bacteroids compared to in free-living cells (9). These findings raised the possibility that *smb21202* might encode a functional protein playing roles during the symbiosis.

The nodule isolates, J049, J077 and J122, carry the exact same *S. meliloti* genomic DNA fragment that covers the upstream region of the putative open reading frame *smb21259.* One possible promoter with a score of 0.81 was predicted in the trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct *smb21259* expression. *smb21259* is predicted to encode hypothetical exported protein precursor. The function of Smb21259 is unknown.

The nodule isolates, J051, J078, J098, J152 and J196, carry the exact same *S. meliloti* genomic DNA fragment that covers the first part of the putative open reading frame *smb21333* and its upstream region including second part of *smb21332* and the 56bp intergenic region between *smb21332* and *smb21333*. One promoter with a score of 0.58 was predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). It is possible for *smb21333* to be expressed under this putative promoter. No homolog of the hypothetical protein Smb21333 was found in other bacteria.

The nodule isolate, J094, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21397*. Two putative promoters

with scores of 0.66 and 0.44, respectively, were predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb21397* is predicted to encode a hypothetical calcium-binding protein, which function is unknown. It shares 51% identical amino acids with putative secreted calcium-binding protein Smb20838 in *S. meliloti*. It also shares 48% identical amino acids with *P. fluorescens* protease PrtA. Protease was reported to play important role in the interaction between pathogen bacteria and their hosts (26).

The nodule isolates, J054 and J163, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21400* and *smb21401*. Two promoters were predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). The putatively upstream promoter scored 0.86 while the putative downstream promoter scored 0.52. The hypothetical protein Smb21400 shares 68-50% identical amino acids with four other hypothetical proteins in *S. meliloti,* Smc00284, Sma1797, Sma0450 and Smc01186. The functions of Smb21400 and its homologs are unknown. The expression of *smb21400* was reported to increase inside bacteroids compared to free-living cells (9). The open reading frame of its downstream putative open reading frame *smb21401* is 52bp away from *smb21400* stop code, and may be expressed under the same promoter. *smb21401* is predicted to encode a hypothetical protein, which has no homolog to be found in other bacteria so far.

The nodule isolate, J015, carries the *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21402*. One putative promoter

with a score of 0.97 was predicted using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb21402* is predicted to encode calcium-binding protein Smb21402. Smb21402 shares 35% identical amino acids with putative secreted calcium-binding protein *S. meliloti* ExpE1, which is involved in synthesis of exopolysaccharide EPS II (133, 165). Smb21402 also shares 70% identical amino acids with *S. meliloti* putative calcium binding protein Smb21229, and 29% identical amino acids with *S. meliloti* putative hemolysin-adenlyate cyclase protein Smb20079. These findings raised the possibility that *smb21402* might encode a functional protein playing roles during the symbiosis.

The nodule isolates, J134 and J147, carry the exact same *S. meliloti* genomic DNA fragment that covers the promoter region of the putative open reading frame *smb21468*. One putative promoter scoring 0.80 was predicted in this region using the prokaryotic promoter predicting software (http://www.fruitfly.org). *smb21468* is predicted to encode a conserved hypothetical membrane protein, which shares 21% identical amino acids with proteobacterium *Caulobacter crescentus* hypothetical sulfate ABC transporter permease CC1597.

The nodule isolates, J129, J192 and J194, carry the exact same *S. meliloti* genomic DNA fragment that covers upstream region of putative open reading frame *smb21595.* Two putative promoters with scores of 0.80 and 0.70, respectively, were predicted in this trapped region using the prokaryotic promoter predicting software (http://www.fruitfly.org), which might direct the expression of *smb21595*. *smb21595* is

predicted to encode a sugar-uptake ABC transporter periplasmic solute-binding protein precursor.

The promoter trap vectors in nodule isolate, J144, carries the *S. meliloti* genomic DNA fragment, locating within putative open reading frame *smc01203.* The direction of fragment showing promoter activities is opposite to *smc01203*. One promoter, score of 0.38, is predicted in the direction of this fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames *smc01203R1* and *smc01203R2* in the direction of the putative promoter. The putative Smc01203R1 protein (157Aa) shares 31% identical amino acids with proteobacterium *Burkholderia cenocepacia* putative short-chain dehydrogenase Bcen2424DRAFT_5053. The putative Smc01203R2 protein (88Aa) shares 35% identical amino acids with *B. pseudomallei* putative non-ribosomal peptide synthase BPSS1197, 32% identical amino acids with *B. japonicum* hypothetical protein blr4533. These findings raised the possibility that the Smc01203R1 and Smc01203R2 might encode a functional protein playing roles in the symbiosis of *S. meliloti* and alfalfa.

The promoter trap vector in nodule isolate, J143, carries the *S. meliloti* genomic DNA fragment locating in the second part of open reading frame *smc02230.* The direction of fragment showing promoter activities is opposite to *smc02230*. *smc02230* is predicted to encode an unknown-function hypothetical transmembrane protein. Two promoters with scores of 0.41 and 0.37, respectively, are predicted in the direction of this fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are five

open reading frames, the sizes of which are more than 500bp, in the direction of the putative promoters. All of them have some homologues in other bacteria. These findings raised the possibility that one of these putative reading frames might encode a functional protein for symbiosis.

The promoter trap vectors in three nodule isolates, J132, J180, and J199, carry the exact same *S. meliloti* genomic DNA fragment, covering most of the *smc02772* open reading frame (86). But the direction of fragment showing promoter activities is opposite to that of the *smc02772* and *smc02773* open reading frames, which have been predicted to code ABC transporter for sugar (86). Two putative promoters were identified in this region opposite to that of the *smc02772* and *smc02773* open reading frames using the prokaryotic promoter predicting software (http://www.fruitfly.org). Both of the two promoter scored 0.97. There are three possible open reading frames, *smc02773R1, smc02773R2,* and *smc02773R3,* in the direction of the putative promoters. The putative Smc02773R1 *(436Aa)* protein shares 28% identical amino acids with the *Kineococcus radiotolerans* TrkA-C protein SRS30216, 26% identical amino acids with the *Chlorobium phaeobacteroides* heat shock protein HslU, and 26% to 32% identical amino acids with more than 10 hypothetical proteins from different bacteria including *M. loti*, *B. cenocepacia*, *Azotobacter vinelandii, K. radiotolerans, C. phaeobacteroides.* These findings raised the possibility that *smc02773R1* might encode a functional protein for symbiosis.

The promoter trap vectors in nodule isolates, J131 and J203, carry the exact same *S. meliloti* genomic DNA fragment locating in the middle of open reading frame *smc03181*. The putative *smc03181* gene is predicted to encode a pH adaptation potassium efflux system protein D, PhaD. The direction of fragment showing the promoter activities is opposite to *smc03181*. Three promoters with scores of 0.40, 0.38 and 0.37, respectively, were predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames *smc03181R1* and *smc03181R2* in the direction of the putative promoters. The hypothetical Smc03181R1 protein (106Aa) shares 36% identical amino acids with proteobacterium *Xanthomonas axonopodis* hypothetical protein XAC3375. The putative Smc03181R2 protein (90Aa) shares 33% identical amino acids with *Actinobacteria Frankia sp.* EAN1pec hypothetical protein Franean1DRAFT_7093. These findings raised the possibility that *smc03181R1* or *smc03181R2* might encode a functional protein playing roles in the symbiosis of *S. meliloti* and alfalfa.

The nodule isolate J175 carries the *S. meliloti* genomic DNA fragment covering second half the *purU1* (*smc03205*) gene and the intergenic region between the *purU1* and the *ipdA3* (*smc03204*) (86). The PurU1 protein is putative formyltetrahydrofolate deformylase, which is involved in small molecule metabolism and nucleotide biosynthesis (50). The direction of the trapped promoter, however, is opposite to that of the *purU1* and the *ipdA3 (smc03204)* genes. Three promoters with scores of 0.90, 0.64 and 0.62, respectively, were predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are three

possible open reading frames in the *purU1* gene in the putative promoters' directions. The putative Smc03205R1 (220Aa) shares 43 to 38% identities with the four hypothetical proteins of *K. radiotolerans* SRS30216. This raises the possibility that the Smc03205R1 might encode a functional protein. The hypothetical Smc03205R2 (241Aa) shares 38 to 42% identical amino acids with four hypothetical proteins from *K. radiotolerans* SRS30216, and 29% identical amino acids with a putative *Archaeoglobus fulgidus* DSM 4304 acetylglutamate kinase. The hypothetical Smc03205R3 (267Aa) shares 25% identical amino acids with a archaebacterium *Haloarcula japonica* TR-1 cell division protein FtsZ1, 38 to 42% identical amino acids with four hypothetical proteins of *K. radiotolerans* SRS30216, and 29% identical amino acids with a putative *A. fulgidus* acetylglutamate kinase DSM4304.

The promoter trap vectors in nodule isolates, J201 and J207, carry the exact same *S. meliloti* genomic DNA fragment locating in the second part of putative open reading frame *smc04135* and the first part of putative open reading frame *smc04136.* The putative *smc04135* and *smc04136* genes are predicted to encode ABC transporter proteins. The direction of fragment showing promoter activities is opposite to both of them. Two promoters with scores of 0.60 and 0.53, respectively, were predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There is one possible open reading frame *smc04135R* in the direction of the putative promoters. The hypothetic Smc04135R protein (285Aa) shares 30% identical amino acids with proteobacterium *Pseudomonas syringae pv. syringae* B ATP-dependent DNA helicase RecQ and 30-25% identical amino acids with several

other putative proteins. These findings raised the possibility that the *smc04135R* might encode a functional protein for symbiosis.

The promoter trap vectors in nodule isolate, J064, carries the *S. meliloti* genomic DNA fragment locating within the second part of the putative open reading frame *sma1252*. The direction of fragment showing promoter activities is opposite to *sma1252*. Two promoters with scores of 0.87 and 0.77, respectively, were predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames *sma1252R1* and *sma1252R2* in the direction of the putative promoters. The hypothetical Sma1252R1 protein (60Aa) shares 31% identical amino acids with *Pseudomonas oleovorans* PhaF protein, which controls the medium-chain-length poly(R)-3-hydroxyalkanoate (PHA) synthesis in *P. oleovorans* under conditions of limiting nutrients in the presence of an excess of carbon source (171). The hypothetical Sma1252R2 protein (150Aa) shares 28% identical amino acids with *S. pyogenes* collagen-like surface protein SclB, which is involved in the adhesion of *S. pyogenes* bacteria to human cells. These findings raised the possibility that *sma1252R1* and *sma1252R2* might encode a functional protein playing roles in the symbiosis.

The promoter trap vectors in two nodule isolates, J149 and J173, carry the exact same genomic fragment covering the middle part of *sma1523* open reading frames. The direction of fragment showing promoter activity is same to that of *sma1523*, but is opposite to the downstream putative open reading frame *sma1521*. One promoter with

score of 0.80 was predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames, *sma1523S1,* and *sma1523S2,* in the direction of the putative promoter. The hypothetical Sma1523S1 (240Aa) protein shares 39% identical amino acids with the *K. radiotolerans* SRS30216's putative short-chain dehydrogenase protein, KradDRAFT_0730. The hypothetical Sma1523S2 (109Aa) protein shares 36% identical amino acids with the proteobacterium *Rhodoferax ferrireducens* putative extracellular solute-binding protein, RferDRAFT_0763. These findings raised the possibility that the *sma1523S1* and *sma1523S2* might encode a functional protein playing roles in the symbiosis.

The nodule isolate, J200, carries the *S. meliloti* genomic DNA fragment locating with the second part of putative open reading frame *smb20276* and the first part of its downstream putative *smb20277.* The direction of fragment showing promoter activities is opposite to that of *smb20276*, which is predicted to encode a transcriptional regulator. Two promoters with scores of 0.91 and 0.74, respectively, were predicted in the direction of the trapped fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There is one possible open reading frame *smb20276R* in the direction of the putative promoters. The hypothetical Smb20276R (187Aa) protein shares 34% identical amino acids with a *K. radiotolerans* hypothetical regulator DeoR protein, *KradDRAFT_1350*, 39% identical amino acids with *A. dehalogenans* hypothetical protein AdehDRAFT_1169. These findings raised the possibility that *smb20276R* might encode a functional protein playing roles in the symbiosis.

The nodule isolate, J029, carries the a *S. meliloti* genomic DNA fragment covering second half the open reading frame *smb20714* and the intergenic region between the *smb20714* and the open reading frame *smb20715.* The open reading frame *smb20714* is predicted to encode a sugar uptake ABC transporter permease protein. The direction of the trapped promoter, however, is opposite to that of *smb20714*. Three promoters with scores of 0.94, 0.87 and 0.82, respectively, were predicted in the direction of this fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are four possible open reading frames overlapping the putative *smb20714* open reading frame in the direction of these putative promoters. One of the four open reading frame *smb20714R1* is predicted to encode a 119 Aa long hypothetical protein Smb20714R1, which shares 32% identity with the hypothetical glycosyl transferase KradDRAFT_2460 of *K. radiotolerans* SRS30216 protein and shares 30% identity with the alpha-acetolactate decarboxylase PpenA01001500 of *Pediococcus pentosaceus* ATCC 25745. This raises the possibility that *smb20714R1* might encode a functional protein. The hypothetical proteins encoded by the other three open reading frames have no homologs in other bacteria.

The promoter trap vectors in nodule isolates, J039 and J176, carry the exact same genomic fragment, covering the middle of the putative open reading frame *smb21065*. The direction of fragment showing promoter activities is opposite to *smb21065* and its upstream putative open reading frame *smb21066*. Five putative promoters with scores of 0.99, 0.98, 0.97, 0.80 and 0.77, respectively, were predicted in the direction of the

fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames *smb21066R1* and *smb21066R2* overlapping *smb21066* under the direction of the five putative promoters. The hypothetical protein Smb21066R1 (253Aa) shares 31% identical amino acids with *B. subtilis* hypothetical tripeptidase homolog YqjE protein. No homolog of Smb21066R2 (230Aa) was found in bacteria.

The promoter trap vectors in nodule isolate, J128, carries the genomic fragment, covering the intergenic region between two putative open reading frames *smb21573* and *smb21574*, and the first part of *smb21574.* The direction of fragment showing promoter activities is opposite to either of them. One promoter with score of 0.32 was predicted in the direction of the fragment using the prokaryotic promoter predicting software (http://www.fruitfly.org). There are two possible open reading frames *smb21573R1* and *smb21573R2* in the direction of the putative promoter. The hypothetical Smb21573R1 protein (137Aa) shares 36% identical amino acids with *Mycobacterium leprae* mycocerosic acid synthase MasA. These findings raised the possibility that *smb21573R1* might encode a functional protein for symbiosis. No homolog of Smb21573R2 (170Aa) was found in other bacteria.

REFERENCES

1. **Akbar, S., C. M. Kang, T. A. Gaidenko, and C. W. Price.** 1997. Modulator protein RsbR regulates environmental signalling in the general stress pathway of *Bacillus subtilis*. *Mol Microbiol* **24:**567-578.

2. **Atichartpongkul, S., S. Loprasert, P. Vattanaviboon, W. Whangsuk, J. D. Helmann, and S. Mongkolsuk.** 2001. Bacterial Ohr and OsmC paralogues define two protein families with distinct functions and patterns of expression. *Microbiology* **147:**1775-1782.

3. **Badet, B., P. Vermoote, P. Y. Haumont, F. Lederer, and F. LeGoffic.** 1987. Glucosamine synthetase from *Escherichia coli*: purification, properties, and glutamine-utilizing site location. *Biochemistry* **26:**1940-1948.

4. **Baev, N., M. Schultze, I. Barlier, D. C. Ha, H. Virelizier, E. Kondorosi, and A. Kondorosi.** 1992. Rhizobium *nodM* and *nodN* genes are common *nod* genes: *nodM* encodes functions for efficiency of *nod* signal production and bacteroid maturation. *J Bacteriol* **174:**7555-7565.

5. **Balzergue, S., B. Dubreucq, S. Chauvin, I. Le-Clainche, F. Le Boulaire, R. de Rose, F. Samson, V. Biaudet, A. Lecharny, C. Cruaud, J. Weissenbach, M. Caboche, and L. Lepiniec.** 2001. Improved PCR-walking for large-scale isolation of plant T-DNA borders. *Biotechniques* **30:**496-8, 502, 504.

6. **Banfalvi, Z., V. Sakanyan, C. Koncz, A. Kiss, I. Dusha, and A. Kondorosi.** 1981. Location of nodulation and nitrogen fixation genes on a high molecular weight plasmid of *R. meliloti*. *Mol Gen Genet* **184:**318-325.

7. **Bardin, S., S. Dan, M. Osteras, and T. M. Finan.** 1996. A phosphate transport system is required for symbiotic nitrogen fixation by *Rhizobium meliloti*. *J Bacteriol* **178:**4540-4507.

8. **Bardin, S. D., and T. M. Finan.** 1998. Regulation of phosphate assimilation in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* **148:**1689-1700.

9. **Barnett, M. J., C. J. Toman, R. F. Fisher, and S. R. Long.** 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.

10. **Battisti, L., J. C. Lara, and J. A. Leigh.** 1992. Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. *Proc Natl Acad Sci U S A* **89:**5625-5629.

11. **Becker, A., H. Berges, E. Krol, C. Bruand, S. Ruberg, D. Capela, E. Lauber, E. Meilhoc, F. Ampe, F. J. de Bruijn, J. Fourment, A. Francez-Charlot, D. Kahn, H. Kuster, C. Liebe, A. Puhler, S. Weidner, and J. Batut.** 2004. Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. *Mol Plant Microbe Interact* **17:**292-303.

12. **Becker, A., A. Kleickmann, M. Keller, W. Arnold, and A. Puhler.** 1993. Identification and analysis of the *Rhizobium meliloti exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. *Mol Gen Genet* **241:**367-379.

13. **Becker, A., A. Kleickmann, H. Kuster, M. Keller, W. Arnold, and A. Puhler.** 1993. Analysis of the *Rhizobium meliloti* genes *exoU, exoV, exoW, exoT*, and *exoI* involved in exopolysaccharide biosynthesis and nodule invasion: *exoU* and *exoW* probably encode glucosyltransferases. *Mol Plant Microbe Interact* **6:**735-744.

14. **Becker, A., H. Kuster, K. Niehaus, and A. Puhler.** 1995. Extension of the *Rhizobium meliloti* succinoglycan biosynthesis gene cluster: identification of the *exsA* gene encoding an ABC transporter protein, and the *exsB* gene which probably codes for a regulator of succinoglycan biosynthesis. *Mol Gen Genet* **249:**487-497.

15. **Becker, A., S. Ruberg, B. Baumgarth, P. A. Bertram-Drogatz, I. Quester, and A. Puhler.** 2002. Regulation of succinoglycan and galactoglucan biosynthesis in *Sinorhizobium meliloti*. *J Mol Microbiol Biotechnol* **4:**187-190.

16. **Begum, A. A., S. Leibovitch, P. Migner, and F. Zhang.** 2001. Specific flavonoids induced nod gene expression and pre-activated nod genes of *Rhizobium leguminosarum* increased pea (*Pisum sativum L.*) and lentil (*Lens culinaris L.*) nodulation in controlled growth chamber environments. *J Exp Bot* **52:**1537-1543.

17. **Bergman, K., M. Gulash-Hoffee, R. E. Hovestadt, R. C. Larosiliere, P. G. Ronco, 2nd, and L. Su.** 1988. Physiology of behavioral mutants of *Rhizobium meliloti*: evidence for a dual chemotaxis pathway. *J Bacteriol* **170:**3249-3254.

18. **Bertram-Drogatz, P. A., I. Quester, A. Becker, and A. Puhler.** 1998. The *Sinorhizobium meliloti* MucR protein, which is essential for the production of high-molecular-weight succinoglycan exopolysaccharide, binds to short DNA regions upstream of *exoH* and *exoY*. *Mol Gen Genet* **257:**433-441.

19. **Beynon, J. L., M. K. Williams, and F. C. Cannon.** 1988. Expression and functional analysis of the *Rhizobium meliloti nifA* gene. *Embo J* **7:**7-14.

20. **Bigot, A., H. Pagniez, E. Botton, C. Frehel, I. Dubail, C. Jacquet, A. Charbit, and C. Raynaud.** 2005. Role of FliF and FliI of Listeria monocytogenes in flagellar assembly and pathogenicity. *Infect Immun* **73:**5530-5539.

21. **Bolanos, L., M. Redondo-Nieto, R. Rivilla, N. J. Brewin, and I. Bonilla.** 2004. Cell surface interactions of *Rhizobium* bacteroids and other bacterial strains with symbiosomal and peribacteroid membrane components from pea nodules. *Mol Plant Microbe Interact* **17:**216-223.

22. **Boncompagni, E., M. Osteras, M. C. Poggi, and D. le Rudulier.** 1999. Occurrence of choline and glycine betaine uptake and metabolism in the family rhizobiaceae and their roles in osmoprotection. *Appl Environ Microbiol* **65:**2072-2077.

23. **Born, T. L., and J. S. Blanchard.** 1999. Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis. *Curr Opin Chem Biol* **3:**607-613.

24. **Bourdon, M. A., T. Krusius, S. Campbell, N. B. Schwartz, and E. Ruoslahti.** 1987. Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proc Natl Acad Sci U S A* **84:**3194-3198.

25. **Bouvier, J., C. Richaud, W. Higgins, O. Bogler, and P. Stragier.** 1992. Cloning, characterization, and expression of the *dapE* gene of *Escherichia coli*. *J Bacteriol* **174:**5265-5271.

26. **Bowen, D. J., T. A. Rocheleau, C. K. Grutzmacher, L. Meslet, M. Valens, D. Marble, A. Dowling, R. Ffrench-Constant, and M. A. Blight.** 2003. Genetic and biochemical characterization of PrtA, an RTX-like metalloprotease from Photorhabdus. *Microbiology* **149:**1581-1591.

27. **Breedveld, M. W., and K. J. Miller.** 1994. Cyclic beta-glucans of members of the family Rhizobiaceae. *Microbiol Rev* **58:**145-161.

28. **Breedveld, M. W., and K. J. Miller.** 1995. Synthesis of glycerophosphorylated cyclic (1,2)-beta-glucans in *Rhizobium meliloti* strain 1021 after osmotic shock. *Microbiology* **141 (Pt 3):**583-588.

29. **Brencic, A., and S. C. Winans.** 2005. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol Mol Biol Rev* **69:**155-194.

30. **Brewin, N. J.** 1991. Development of the legume root nodule. *Annu Rev Cell Biol* **7:**191-226.

31. **Brewin, N. J.** 1993. The *Rhizobium*-legume symbiosis: plant morphogenesis in a nodule. *Semin Cell Biol* **4:**149-156.

32. **Brhada, F., M. C. Poggi, and D. Le Rudulier.** 1997. Choline and glycine betaine uptake in various strains of Rhizobia isolated from nodules of *Vicia faba var. major* and *Cicer arietinum l.*: modulation by salt, choline, and glycine betaine. *Curr Microbiol* **34:**167-172.

33. **Buendia, A. M., B. Enenkel, R. Koplin, K. Niehaus, W. Arnold, and A. Puhler.** 1991. The *Rhizobium meliloti exoZl exoB* fragment of megaplasmid 2: ExoB functions as a UDP-glucose 4-epimerase and ExoZ shows homology to NodX of *Rhizobium leguminosarum biovar viciae* strain TOM. *Mol Microbiol* **5:**1519-1530.

34. **Burrows, L. L., D. Chow, and J. S. Lam.** 1997. *Pseudomonas aeruginosa* B-band O-antigen chain length is modulated by Wzz (Ro1). *J Bacteriol* **179:**1482-1489.

35. **Caetano-Anolles, G., D. K. Crist-Estes, and W. D. Bauer.** 1988. Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin requires functional nodulation genes. *J Bacteriol* **170:**3164-3169.

36. **Caetano-Anolles, G., A. Lagares, and W. D. Bauer.** 1990. *Rhizobium meliloti* exopolysaccharide mutants elicit feedback regulation of nodule formation in alfalfa. *Plant Physiol.* **91:**368-374.

37. **Camilli, A., and J. J. Mekalanos.** 1995. Use of recombinase gene fusions to identify *Vibrio cholerae* genes induced during infection. *Mol Microbiol* **18:**671-683.

38. **Campbell, G. R., L. A. Sharypova, H. Scheidle, K. M. Jones, K. Niehaus, A. Becker, and G. C. Walker.** 2003. Striking complexity of lipopolysaccharide defects in a collection of *Sinorhizobium meliloti* mutants. *J Bacteriol* **185:**3853-3862.

39. **Cangelosi, G. A., L. Hung, V. Puvanesarajah, G. Stacey, D. A. Ozga, J. A. Leigh, and E. W. Nester.** 1987. Common loci for *Agrobacterium tumefaciens* and *Rhizobium meliloti* exopolysaccharide synthesis and their roles in plant interactions. *J Bacteriol* **169:**2086-2091.

40. **Canter Cremers, H. C., M. Batley, J. W. Redmond, L. Eydems, M. W. Breedveld, L. P. Zevehuizen, E. Pees, C. A. Wijffelman, and B. J. Lugtenberg.** 1990. *Rhizobium leguminosarum exoB* mutants are deficient in the synthesis of UDP-glucose 4'-epimerase. *J Biol Chem* **265:**21122-21127.

41. **Capela, D., F. Barloy-Hubler, J. Gouzy, G. Bothe, F. Ampe, J. Batut, P. Boistard, A. Becker, M. Boutry, E. Cadieu, S. Dreano, S. Gloux, T. Godrie, A. Goffeau, D. Kahn, E. Kiss, V. Lelaure, D. Masuy, T. Pohl, D. Portetelle, A. Puhler, B. Purnelle, U. Ramsperger, C. Renard, P. Thebault, M. Vandenbol, S. Weidner, and F. Galibert.** 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc Natl Acad Sci U S A* **98:**9877-9882.

42. **Cardenas, L., J. Dominguez, O. Santana, and C. Quinto.** 1996. The role of the nodI and nodJ genes in the transport of Nod metabolites in *Rhizobium etli*. Gene **173:**183-187.

43. **Cebolla, A., and A. J. Palomares.** 1994. Genetic regulation of nitrogen fixation in *Rhizobium meliloti*. *Microbiologia* **10:**371-384.

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

45. **Charles, T. C., W. Newcomb, and T. M. Finan.** 1991. *ndvF*, a novel locus located on megaplasmid pRmeSU47b (pEXO) of *Rhizobium meliloti*, is required for normal nodule development. *J Bacteriol* **173:**3981-3992.

46. **Cheng, H. P., and G. C. Walker.** 1998. Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J Bacteriol* **180:**5183-5191.

47. **Cheng, H. P., and G. C. Walker.** 1998. Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *J Bacteriol* **180:**20-26.

48. **Cheng, H. P., and S. Y. Yao.** 2004. The key *Sinorhizobium meliloti* succinoglycan biosynthesis gene exoY is expressed from two promoters. *FEMS Microbiol Lett* **231:**131-136.

49. **Chesnokova, O., J. B. Coutinho, I. H. Khan, M. S. Mikhail, and C. I. Kado.** 1997. Characterization of flagella genes of *Agrobacterium tumefaciens*, and the effect of a bald strain on virulence. *Mol Microbiol* **23:**579-590.

50. **Chiribau, C. B., C. Sandu, G. L. Igloi, and R. Brandsch.** 2005. Characterization of PmfR, the transcriptional activator of the pAO1-borne *purU-mabO-folD* operon of *Arthrobacter nicotinovorans*. *J Bacteriol* **187:**3062-3070.

51. **Cren, M., A. Kondorosi, and E. Kondorosi.** 1995. NolR controls expression of the *Rhizobium meliloti* nodulation genes involved in the core Nod factor synthesis. *Mol Microbiol* **15:**733-747.

52. **Crowther, L. J., R. P. Anantha, and M. S. Donnenberg.** 2004. The inner membrane subassembly of the enteropathogenic *Escherichia coli* bundle-forming pilus machine. *Mol Microbiol* **52:**67-79.

53. **Dandanell, G., K. Norris, and K. Hammer.** 1991. Long-distance *deoR* regulation of gene expression in *Escherichia coli*. *Ann N Y Acad Sci* **646:**19-30.

54. **Dartigalongue, C., and S. Raina.** 1998. A new heat-shock gene, *ppiD*, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in *Escherichia coli*. *Embo J* **17:**3968-3980.

55. **David, M., M. L. Daveran, J. Batut, A. Dedieu, O. Domergue, J. Ghai, C. Hertig, P. Boistard, and D. Kahn.** 1988. Cascade regulation of nif gene expression in *Rhizobium meliloti*. Cell **54:**671-683.

56. **De Costa, D. M., K. Suzuki, and K. Yoshida.** 2003. Structural and functional analysis of a putative gene cluster for palatinose transport on the linear chromosome of *Agrobacterium tumefaciens* MAFF301001. *J Bacteriol* **185:**2369-2373.

57. **de Lima Pimenta, A., P. Di Martino, E. Le Bouder, C. Hulen, and M. A. Blight.** 2003. In vitro identification of two adherence factors required for in vivo virulence of *Pseudomonas fluorescens*. *Microbes Infect* **5:**1177-1187.

58. **Dean, D. R., J. T. Bolin, and L. Zheng.** 1993. Nitrogenase metalloclusters: structures, organization, and synthesis. *J Bacteriol* **175:**6737-6744.

59. **DeAngelis, P. L.** 2002. Microbial glycosaminoglycan glycosyltransferases. *Glycobiology* **12:**9R-16R.

60. **Delumeau, O., S. Dutta, M. Brigulla, G. Kuhnke, S. W. Hardwick, U. Volker, M. D. Yudkin, and R. J. Lewis.** 2004. Functional and structural characterization of RsbU, a stress signaling protein phosphatase 2C. *J Biol Chem* **279:**40927-40937.

61. **Demont, N., M. Ardourel, F. Maillet, D. Prome, M. Ferro, J. C. Prome, and J. Denarie.** 1994. The *Rhizobium meliloti* regulatory *nodD3* and *syrM* genes control the synthesis of a particular class of nodulation factors N-acylated by (omega-1)-hydroxylated fatty acids. *Embo J* **13:**2139-2149.

62. **Dickstein, R., T. Bisseling, V. N. Reinhold, and F. M. Ausubel.** 1988. Expression of nodule-specific genes in alfalfa root nodules blocked at an early stage of development. *Genes Dev* **2:**677-687.

63. **Djordjevic, M. A.** 2004. *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. *Proteomics* **4:**1859-1872.

64. **Doherty, D., J. A. Leigh, J. Glazebrook, and G. C. Walker.** 1988. *Rhizobium meliloti* mutants that overproduce the *Rhizobium meliloti* acidic calcofluor-binding exopolysaccharide. *J Bacteriol* **170:**4249-4256.

65. **Dombrecht, B., C. Heusdens, S. Beullens, C. Verreth, E. Mulkers, P. Proost, J. Vanderleyden, and J. Michiels.** 2005. Defence of *Rhizobium etli* bacteroids against oxidative stress involves a complexly regulated atypical 2-Cys peroxiredoxin. *Mol Microbiol* **55:**1207-1221.

66. **Downie, J. A., and S. A. Walker.** 1999. Plant responses to nodulation factors. *Curr Opin Plant Biol* **2:**483-489.

67. **Dylan, T., P. Nagpal, D. R. Helinski, and G. S. Ditta.** 1990. Symbiotic pseudorevertants of *Rhizobium meliloti ndv* mutants. *J Bacteriol* **172:**1409-1417.

68. **Elsen, S., O. Duche, and A. Colbeau.** 2003. Interaction between the H2 sensor HupUV and the histidine kinase HupT controls HupSL hydrogenase synthesis in *Rhodobacter capsulatus*. *J Bacteriol* **185:**7111-7119.

69. **Elzer, P. H., R. W. Phillips, M. E. Kovach, K. M. Peterson, and R. M. Roop, 2nd.** 1994. Characterization and genetic complementation of a *Brucella abortus* high-temperature-requirement A (*htrA*) deletion mutant. *Infect Immun* **62:**4135-4139.

70. **Engler, G., M. Holsters, M. Van Montagu, J. Schell, J. P. Hernalsteens, and Schilperoort.** 1975. Agrocin 84 sensitivity: a plasmid determined property in *Agrobacterium tumefaciens*. *Mol Gen Genet* **138:**345-349.

71. **Epple, G., K. M. van der Drift, J. E. Thomas-Oates, and O. Geiger.** 1998. Characterization of a novel acyl carrier protein, RkpF, encoded by an operon involved in capsular polysaccharide biosynthesis in *Sinorhizobium meliloti*. *J Bacteriol* **180:**4950-4954.

72. **Ferraioli, S., R. Tate, M. Cermola, R. Favre, M. Iaccarino, and E. J. Patriarca.** 2002. Auxotrophic mutant strains of *Rhizobium etli* reveal new nodule development phenotypes. *Mol Plant Microbe Interact* **15:**501-510.

73. **Finan, T. M., E. Hartweig, K. LeMieux, K. Bergman, G. C. Walker, and E. R. Signer.** 1984. General transduction in *Rhizobium meliloti*. *J Bacteriol* **159:**120-124.

74. **Finan, T. M., B. Kunkel, G. F. De Vos, and E. R. Signer.** 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J Bacteriol* **167:**66-72.

75. **Fischer, H. M.** 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol Rev* **58:**352-386.

76. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 1988. Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. *Genes Dev* **2:**282-293.

77. **Fisher, R. F., and S. R. Long.** 1989. DNA footprint analysis of the transcriptional activator proteins NodD1 and NodD3 on inducible *nod* gene promoters. *J Bacteriol* **171:**5492-5502.

78. **Fox, J. E., M. Starcevic, P. E. Jones, M. E. Burow, and J. A. McLachlan.** 2004. Phytoestrogen signaling and symbiotic gene activation are disrupted by endocrine-disrupting chemicals. *Environ Health Perspect* **112:**672-677.

79. **Fraysse, N., F. Couderc, and V. Poinsot.** 2003. Surface polysaccharide involvement in establishing the rhizobium-legume symbiosis. *Eur J Biochem* **270:**1365-1380.

80. **Fuangthong, M., S. Atichartpongkul, S. Mongkolsuk, and J. D. Helmann.** 2001. OhrR is a repressor of *ohrA*, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J Bacteriol* **183:**4134-4141.

81. **Fuangthong, M., and J. D. Helmann.** 2002. The OhrR repressor senses organic hydroperoxides by reversible formation of a cysteine-sulfenic acid derivative. *Proc Natl Acad Sci U S A* **99:**6690-6695.

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

83. **Gage, D. J., T. Bobo, and S. R. Long.** 1996. Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). *J Bacteriol* **178:**7159-7166.

84. **Gage, D. J., and W. Margolin.** 2000. Hanging by a thread: invasion of legume plants by rhizobia. *Curr Opin Microbiol* **3:**613-617.

85. **Gal, M., G. M. Preston, R. C. Massey, A. J. Spiers, and P. B. Rainey.** 2003. Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. *Mol Ecol* **12:**3109-3121.

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

87. **Geiger, O., and I. M. Lopez-Lara.** 2002. Rhizobial acyl carrier proteins and their roles in the formation of bacterial cell-surface components that are required for the development of nitrogen-fixing root nodules on legume hosts. *FEMS Microbiol Lett* **208:**153-162.

88. **Georgiadis, M. M., H. Komiya, P. Chakrabarti, D. Woo, J. J. Kornuc, and D. C. Rees.** 1992. Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* **257:**1653-1659.

89. **Ghrist, A. C., and G. V. Stauffer.** 1995. The *Escherichia coli* glycine transport system and its role in the regulation of the glycine cleavage enzyme system. *Microbiology* **141 (Pt 1):**133-140.

90. **Gibson, J. L., and F. R. Tabita.** 1997. Analysis of the *cbbXYZ* operon in *Rhodobacter sphaeroides*. *J Bacteriol* **179:**663-669.

91. **Glazebrook, J., A. Ichige, and G. C. Walker.** 1996. Genetic analysis of Rhizobium meliloti *bacA-phoA* fusion results in identification of *degP*: two loci required for symbiosis are closely linked to *degP*. *J Bacteriol* **178:**745-752.

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

93. **Glucksmann, M. A., T. L. Reuber, and G. C. Walker.** 1993. Family of glycosyl transferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. *J Bacteriol* **175:**7033-7044.

94. **Glucksmann, M. A., T. L. Reuber, and G. C. Walker.** 1993. Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. *J Bacteriol* **175:**7045-7055.

95. **Gomes, C. M., J. B. Vicente, A. Wasserfallen, and M. Teixeira.** 2000. Spectroscopic studies and characterization of a novel electron-transfer chain from *Escherichia coli* involving a flavorubredoxin and its flavoprotein reductase partner. *Biochemistry* **39:**16230-16237.

96. **Gonzalez, J. E., B. L. Reuhs, and G. C. Walker.** 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.

97. **Gonzalez, J. E., C. E. Semino, L. X. Wang, L. E. Castellano-Torres, and G. C. Walker.** 1998. Biosynthetic control of molecular weight in the polymerization of the octasaccharide subunits of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Proc Natl Acad Sci U S A* **95:**13477-13482.

98. **Gordia, S., and C. Gutierrez.** 1996. Growth-phase-dependent expression of the osmotically inducible gene osmC of *Escherichia coli* K-12. *Mol Microbiol* **19:**729-736.

99. **Grandori, R., P. Khalifah, J. A. Boice, R. Fairman, K. Giovanielli, and J. Carey.** 1998. Biochemical characterization of WrbA, founding member of a new family of multimeric flavodoxin-like proteins. *J Biol Chem* **273:**20960-20966.

100. **Griffin, H. G., and M. J. Gasson.** 1995. The gene (*aroK*) encoding shikimate kinase I from *Escherichia coli*. *DNA Seq* **5:**195-197.

101. **Gu, W., J. Song, C. A. Bonner, G. Xie, and R. A. Jensen.** 1998. PhhC is an essential aminotransferase for aromatic amino acid catabolism in *Pseudomonas aeruginosa*. *Microbiology* **144 (Pt 11):**3127-3134.

102. **Gutierrez, C., and J. C. Devedjian.** 1991. Osmotic induction of gene *osmC* expression in *Escherichia coli* K12. *J Mol Biol* **220:**959-973.

103. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166:**557-580.

104. **Harms, N., W. N. Reijnders, S. Koning, and R. J. van Spanning.** 2001. Two-component system that regulates methanol and formaldehyde oxidation in *Paracoccus denitrificans*. *J Bacteriol* **183:**664-670.

105. **Hartwig, U. A., C. A. Maxwell, C. M. Joseph, and D. A. Phillips.** 1990. Effects of alfalfa nod gene-inducing flavonoids on *nodABC* transcription in *Rhizobium meliloti* strains containing different *nodD* genes. *J Bacteriol* **172:**2769-2773.

106. **Her, G. R., J. Glazebrook, G. C. Walker, and V. N. Reinhold.** 1990. Structural studies of a novel exopolysaccharide produced by a mutant of *Rhizobium meliloti* strain Rm1021. *Carbohydr Res* **198:**305-312.

107. **Hirsch, A. M.** 1992. Developmental biology of legume nodulation. *New. Phytol* **122:**211-237.

108. **Hirsch, A. M., M. Bang, and F. M. Ausubel.** 1983. Ultrastructural analysis of ineffective alfalfa nodules formed by *nif*::Tn*5* mutants of *Rhizobium meliloti*. *J Bacteriol* **155:**367-380.

109. **Hirsch, A. M., and C. A. Smith.** 1987. Effects of *Rhizobium meliloti nif* and *fix* mutants on alfalfa root nodule development. *J Bacteriol* **169:**1137-1146.

110. **Hoheisel, J. D.** 1994. Creation of a rare cutter multiple cloning site. *Biotechniques* **17:**456, 458, 460.

111. **Honma, M. A., M. Asomaning, and F. M. Ausubel.** 1990. *Rhizobium meliloti nodD* genes mediate host-specific activation of nodABC. *J Bacteriol* **172:**901-911.

112. **Ibanez-Ruiz, M., V. Robbe-Saule, D. Hermant, S. Labrude, and F. Norel.** 2000. Identification of RpoS (sigma(S))-regulated genes in *Salmonella enterica serovar typhimurium*. *J Bacteriol* **182:**5749-5756.

113. **Jack, D. L., N. M. Yang, and M. H. Saier, Jr.** 2001. The drug/metabolite transporter superfamily. *Eur J Biochem* **268:**3620-3639.

114. **Jensen, J. B., N. K. Peters, and T. V. Bhuvaneswari.** 2002. Redundancy in periplasmic binding protein-dependent transport systems for trehalose, sucrose, and maltose in *Sinorhizobium meliloti*. *J Bacteriol* **184:**2978-2986.

115. **Karlin, S., M. J. Barnett, A. M. Campbell, R. F. Fisher, and J. Mrazek.** 2003. Predicting gene expression levels from codon biases in alpha-proteobacterial genomes. *Proc Natl Acad Sci U S A* **100:**7313-7318.

116. **Katz, M. E., R. A. Strugnell, and J. I. Rood.** 1992. Molecular characterization of a genomic region associated with virulence in *Dichelobacter nodosus*. *Infect Immun* **60:**4586-4592.

117. **Keller, M., A. Roxlau, W. M. Weng, M. Schmidt, J. Quandt, K. Niehaus, D. Jording, W. Arnold, and A. Puhler.** 1995. Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. *Mol Plant Microbe Interact* **8:**267-277.

118. **Kijne, J. W.** 1992. The *Rhizobium* Infection Process, p. 349-398. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological Nitrogen Fixation. Chapman & Hall, New York.

119. **Kim, Y. K., and L. L. McCarter.** 2000. Analysis of the polar flagellar gene system of *Vibrio parahaemolyticus*. *J Bacteriol* **182:**3693-3704.

120. **Kjellen, L., and U. Lindahl.** 1991. Proteoglycans: structures and interactions. *Annu Rev Biochem* **60:**443-475.

121. **Kobayashi, H., Y. Naciri-Graven, W. J. Broughton, and X. Perret.** 2004. Flavonoids induce temporal shifts in gene-expression of *nod*-box controlled loci in *Rhizobium sp.* NGR234. *Mol Microbiol* **51:**335-347.

122. **Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop, 2nd, and K. M. Peterson.** 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166:**175-176.

123. **Le, T. T., S. Harlepp, C. C. Guet, K. Dittmar, T. Emonet, T. Pan, and P. Cluzel.** 2005. Real-time RNA profiling within a single bacterium. *Proc Natl Acad Sci U S A.* **102:**9160–9164.

124. **Leigh, J. A., and C. C. Lee.** 1988. Characterization of polysaccharides of *Rhizobium meliloti exo* mutants that form ineffective nodules. *J Bacteriol* **170:**3327-3332.

125. **Leigh, J. A., J. W. Reed, J. F. Hanks, A. M. Hirsch, and G. C. Walker.** 1987. *Rhizobium meliloti* mutants that fail to succinylate their calcofluor-binding exopolysaccharide are defective in nodule invasion. *Cell* **51:**579-587.

126. **Leigh, J. A., E. R. Signer, and G. C. Walker.** 1985. Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. *Proc Natl Acad Sci U S A* **82:**6231-6235.

127. **Leigh, J. A., and G. C. Walker.** 1994. Exopolysaccharides of *Rhizobium*: synthesis, regulation and symbiotic function. *Trends Genet* **10:**63-67.

128. **Lestrate, P., A. Dricot, R. M. Delrue, C. Lambert, V. Martinelli, X. De Bolle, J. J. Letesson, and A. Tibor.** 2003. Attenuated signature-tagged mutagenesis mutants of *Brucella melitensis* identified during the acute phase of infection in mice. *Infect Immun* **71:**7053-7060.

129. **Levery, S. B., H. Zhan, C. C. Lee, J. A. Leigh, and S. Hakomori.** 1991. Structural analysis of a second acidic exopolysaccharide of *Rhizobium meliloti* that can function in alfalfa root nodule invasion. *Carbohydr Res* **210:**339-347.

130. **Lewis, T. A., R. Gonzalez, and J. L. Botsford.** 1990. *Rhizobium meliloti* glutamate synthase: cloning and initial characterization of the glt locus. *J Bacteriol* **172:**2413-2420.

131. **Li, L., Y. Jia, Q. Hou, T. C. Charles, E. W. Nester, and S. Q. Pan.** 2002. A global pH sensor: Agrobacterium sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. *Proc Natl Acad Sci U S A* **99:**12369-12374.

132. **Liu, D., Y. Zhao, X. Fan, Y. Sun, and R. O. Fox.** 2004. *Escherichia coli* stress protein YciF: expression, crystallization and preliminary crystallographic analysis. *Acta Crystallogr D Biol Crystallogr* **60:**2389-2390.

133. **Lloret, J., B. B. Wulff, J. M. Rubio, J. A. Downie, I. Bonilla, and R. Rivilla.** 1998. Exopolysaccharide II production is regulated by salt in the halotolerant strain *Rhizobium meliloti* EFB1. *Appl Environ Microbiol* **64:**1024-1028.

134. **Loewus, M. W., and D. R. Briggs.** 1952. The number of catalytically active sites present on the chymotrypsin molecule. *J Biol Chem* **199:**857-864.

135. **Long, S., J. W. Reed, J. Himawan, and G. C. Walker.** 1988. Genetic analysis of a cluster of genes required for synthesis of the calcofluor-binding exopolysaccharide of *Rhizobium meliloti*. *J Bacteriol* **170:**4239-4248.

136. **Long, S. R.** 2001. Genes and signals in the rhizobium-legume symbiosis. *Plant Physiol* **125:**69-72.

137. **Long, S. R.** 1996. *Rhizobium* symbiosis: *nod* factors in perspective. *Plant Cell* **8:**1885-1898.

138. **Long, S. R.** 1989. *Rhizobium*-legume nodulation: life together in the underground. *Cell* **56:**203-214.

139. **Lopez-Lara, I. M., and O. Geiger.** 2001. The nodulation protein NodG shows the enzymatic activity of an 3-oxoacyl-acyl carrier protein reductase. *Mol Plant Microbe Interact* **14:**349-357.

140. **Luo, L., S. Y. Yao, A. Becker, S. Ruberg, G. Q. Yu, J. B. Zhu, and H. P. Cheng.** 2005. Two new *Sinorhizobium meliloti* LysR-type transcriptional regulators required for nodulation. *J Bacteriol* **187:**4562-4572.

141. **Lyon, W. R., and M. G. Caparon.** 2004. Role for serine protease HtrA (DegP) of *Streptococcus pyogenes* in the biogenesis of virulence factors SpeB and the hemolysin streptolysin S. *Infect Immun* **72:**1618-1625.

142. **Mahan, M. J., J. M. Slauch, and J. J. Mekalanos.** 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* **259:**686-688.

143. **Maillet, F., F. Debelle, and J. Denarie.** 1990. Role of the nodD and syrM genes in the activation of the regulatory gene *nodD3*, and of the common and host-specific nod genes of *Rhizobium meliloti*. *Mol Microbiol* **4:**1975-1984.

144. **Mandon, K., P. A. Kaminski, and C. Elmerich.** 1994. Functional analysis of the *fixNOQP* region of *Azorhizobium caulinodans. J Bacteriol* **176:**2560-2568.

145. **Mandon, K., P. A. Kaminski, C. Mougel, N. Desnoues, B. Dreyfus, and C. Elmerich.** 1993. Role of the *fixGHI* region of *Azorhizobium caulinodans* in free-living and symbiotic nitrogen fixation. *FEMS Microbiol Lett* **114:**185-189.

146. **Marschner, H.** 1995. Mineral nutrition of higher plants, 2nd ed. Academic Press, Ltd., London.

147. **Meade, H. M., S. R. Long, G. B. Ruvkun, S. E. Brown, and F. M. Ausubel.** 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. *J Bacteriol* **149:**114-122.

148. **Mendrygal, K. E., and J. E. Gonzalez.** 2000. Environmental regulation of exopolysaccharide production in *Sinorhizobium meliloti*. *J Bacteriol* **182:**599-606.

149. **Merkamm, M., and A. Guyonvarch.** 2001. Cloning of the *sodA* gene from *Corynebacterium melassecola* and role of superoxide dismutase in cellular viability. *J Bacteriol* **183:**1284-1295.

150. **Miller, K. J., M. W. McKinstry, W. P. Hunt, and B. T. Nixon.** 1992. Identification of the diacylglycerol kinase structural gene of *Rhizobium meliloti* 1021. *Mol Plant Microbe Interact* **5:**363-271.

151. **Muller, P., K. Ahrens, T. Keller, and A. Klaucke.** 1995. A TnphoA insertion within the *Bradyrhizobium japonicum sipS* gene, homologous to prokaryotic

signal peptidases, results in extensive changes in the expression of PBM-specific nodulins of infected soybean (*Glycine max*) cells. *Mol Microbiol* **18:**831-840.

152. **Muller, P., M. Keller, W. M. Weng, J. Quandt, W. Arnold, and A. Puhler.** 1993. Genetic analysis of the *Rhizobium meliloti exoYFQ* operon: ExoY is homologous to sugar transferases and ExoQ represents a transmembrane protein. *Mol Plant Microbe Interact* **6:**55-65.

153. **Mulligan, J. T., and S. R. Long.** 1985. Induction of *Rhizobium meliloti nodC* expression by plant exudate requires *nodD*. *Proc Natl Acad Sci U S A* **82:**6609-6613.

154. **Nienaber, A., H. Hennecke, and H. M. Fischer.** 2001. Discovery of a haem uptake system in the soil bacterium *Bradyrhizobium japonicum*. *Mol Microbiol* **41:**787-800.

155. **Ogawa, J., and S. R. Long.** 1995. The *Rhizobium meliloti groELc* locus is required for regulation of early nod genes by the transcription activator NodD. *Genes Dev* **9:**714-729.

156. **Oke, V., and S. R. Long.** 1999. Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. *Mol Microbiol* **32:**837-849.

157. **Oke, V., and S. R. Long.** 1999. Bacteroid formation in the *Rhizobium*-legume symbiosis. *Curr Opin Microbiol* **2:**641-646.

158. **Olson, E. R., D. S. Dunyak, L. M. Jurss, and R. A. Poorman.** 1991. Identification and characterization of *dppA*, an Escherichia coli gene encoding a periplasmic dipeptide transport protein. *J Bacteriol* **173:**234-244.

159. **Oresnik, I. J., T. C. Charles, and T. M. Finan.** 1994. Second site mutations specifically suppress the Fix- phenotype of *Rhizobium meliloti ndvF* mutations on alfalfa: identification of a conditional *ndvF*-dependent mucoid colony phenotype. Genetics **136:**1233-1243.

160. **Osteras, M., B. T. Driscoll, and T. M. Finan.** 1995. Molecular and expression analysis of the *Rhizobium meliloti* phosphoenolpyruvate carboxykinase (*pckA*) gene. *J Bacteriol* **177:**1452-1460.

161. **Osteras, M., S. A. O'Brien, and T. M. Finan.** 1997. Genetic analysis of mutations affecting pckA regulation in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* **147:**1521-1531.

162. **Osteras, M., J. Stanley, and T. M. Finan.** 1995. Identification of *Rhizobium*-specific intergenic mosaic elements within an essential two-component regulatory system of *Rhizobium* species. *J Bacteriol* **177:**5485-5494.

163. **Paulsson, J.** 2004. Summing up the noise in gene networks. *Nature* **427:**415-418.

164. **Pellock, B. J., H. P. Cheng, and G. C. Walker.** 2000. Alfalfa root nodule invasion efficiency is dependent on *Sinorhizobium meliloti* polysaccharides. J Bacteriol **182:**4310-4318.

165. **Pellock, B. J., M. Teplitski, R. P. Boinay, W. D. Bauer, and G. C. Walker.** 2002. A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. *J Bacteriol* **184:**5067-5076.

166. **Perret, X., C. Staehelin, and W. J. Broughton.** 2000. Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev* **64:**180-201.

167. **Peters, N., and S. Long.** 1988. Alfalfa root exudates and compounds which promote or inhibit induction of *Rhizobium meliloti* nodulation genes. *Plant Physiol* **88:**396-400.

168. **Peters, N. K., J. W. Frost, and S. R. Long.** 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* **233:**977-980.

169. **Phadtare, S.** 2004. Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* **6:**125-136.

170. **Piehler, J.** 2005. New methodologies for measuring protein interactions in vivo and in vitro. *Curr Opin Struct Biol* **15:**4-14.

171. **Prieto, M. A., B. Buhler, K. Jung, B. Witholt, and B. Kessler.** 1999. PhaF, a polyhydroxyalkanoate-granule-associated protein of *Pseudomonas oleovorans* GPo1 involved in the regulatory expression system for *pha* genes. *J Bacteriol* **181:**858-868.

172. **Rainey, P. B.** 1999. Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. *Environ Microbiol* **1:**243-257.

173. **Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan, and R. Tobes.** 2005. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69:**326-356.

174. **Rediers, H., P. B. Rainey, J. Vanderleyden, and R. De Mot.** 2005. Unraveling the secret lives of bacteria: use of in vivo expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression. *Microbiol Mol Biol Rev* **69:**217-261.

175. **Reed, J. W., M. Capage, and G. C. Walker.** 1991. *Rhizobium meliloti exoG* and *exoJ* mutations affect the *exoX-exoY* system for modulation of exopolysaccharide production. *J Bacteriol* **173:**3776-88.

176. **Reed, J. W., J. Glazebrook, and G. C. Walker.** 1991. The *exoR* gene of *Rhizobium meliloti* affects RNA levels of other *exo* genes but lacks homology to known transcriptional regulators. *J Bacteriol* **173:**3789-3794.

177. **Reed, J. W., and G. C. Walker.** 1991. The *exoD* gene of *Rhizobium meliloti* encodes a novel function needed for alfalfa nodule invasion. *J Bacteriol* **173:**664-677.

178. **Rees, D. C., and J. B. Howard.** 1999. Structural bioenergetics and energy transduction mechanisms. *J Mol Biol* **293:**343-350.

179. **Reeve, W. G., M. J. Dilworth, R. P. Tiwari, and A. R. Glenn.** 1997. Regulation of exopolysaccharide production in *Rhizobium leguminosarum biovar viciae* WSM710 involves *exoR*. *Microbiology* **143 (Pt 6):**1951-1958.

180. **Reinhold, B. B., S. Y. Chan, T. L. Reuber, A. Marra, G. C. Walker, and V. N. Reinhold.** 1994. Detailed structural characterization of succinoglycan, the major exopolysaccharide of *Rhizobium meliloti* Rm1021. *J Bacteriol* **176:**1997-2002.

181. **Relic, B., X. Perret, M. T. Estrada-Garcia, J. Kopcinska, W. Golinowski, H. B. Krishnan, S. G. Pueppke, and W. J. Broughton.** 1994. Nod factors of *Rhizobium* are a key to the legume door. *Mol Microbiol* **13:**171-178.

182. **Reuber, T. L., S. Long, and G. C. Walker.** 1991. Regulation of R*hizobium meliloti exo* genes in free-living cells and in planta examined by using TnphoA fusions. *J Bacteriol* **173:**426-434.

183. **Reuber, T. L., and G. C. Walker.** 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* **74:**269-280.

184. **Reuhs, B. L., R. W. Carlson, and J. S. Kim.** 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.

185. **Roche, P., P. Lerouge, C. Ponthus, and J. C. Prome.** 1991. Structural determination of bacterial nodulation factors involved in the *Rhizobium meliloti*-alfalfa symbiosis. *J Biol Chem* **266:**10933-10940.

186. **Roden, L., T. Koerner, C. Olson, and N. B. Schwartz.** 1985. Mechanisms of chain initiation in the biosynthesis of connective tissue polysaccharides. *Fed Proc* **44:**373-380.

187. **Roden, L., and R. Smith.** 1966. Structure of the neutral trisaccharide of the chondroitin 4-sulfate-protein linkage region. *J Biol Chem* **241:**5949-5954.

188. **Rosinha, G. M., D. A. Freitas, A. Miyoshi, V. Azevedo, E. Campos, S. L. Cravero, O. Rossetti, G. Splitter, and S. C. Oliveira.** 2002. Identification and characterization of a *Brucella abortus* ATP-binding cassette transporter homolog to *Rhizobium meliloti* ExsA and its role in virulence and protection in mice. *Infect Immun* **70:**5036-5044.

189. **Rossbach, S., D. A. Kulpa, U. Rossbach, and F. J. de Bruijn.** 1994. Molecular and genetic characterization of the rhizopine catabolism (*mocABRC*) genes of *Rhizobium meliloti* L5-30. *Mol Gen Genet* **245:**11-24.

190. **Ruberg, S., A. Puhler, and A. Becker.** 1999. Biosynthesis of the exopolysaccharide galactoglucan in *Sinorhizobium meliloti* is subject to a complex control by the phosphate-dependent regulator PhoB and the proteins ExpG and MucR. *Microbiology* **145 (Pt 3):**603-611.

191. **Ruvkun, G. B., V. Sundaresan, and F. M. Ausubel.** 1982. Directed transposon Tn*5* mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. *Cell* **29:**551-559.

192. **Sambrook, J., E. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press.

193. **Samen, U., B. Gottschalk, B. J. Eikmanns, and D. J. Reinscheid.** 2004. Relevance of peptide uptake systems to the physiology and virulence of *Streptococcus agalactiae*. *J Bacteriol* **186:**1398-1408.

194. **Scharf, B., H. Schuster-Wolff-Buhring, R. Rachel, and R. Schmitt.** 2001. Mutational analysis of the *Rhizobium lupini* H13-3 and *Sinorhizobium meliloti* flagellin genes: importance of flagellin A for flagellar filament structure and transcriptional regulation. J *Bacteriol* **183:**5334-5342.

195. **Schlaman, H. R., B. Horvath, E. Vijgenboom, R. J. Okker, and B. J. Lugtenberg.** 1991. Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum biovar viciae*. *J Bacteriol* **173:**4277-4287.

196. **Schultze, M., and A. Kondorosi.** 1995. The role of Nod signal structures in the determination of host specificity in the *Rhizobium*-legume symbiosis. *World J Microbiol Biotechnol* **12:**137-149.

197. **Shah, V. K., P. Rangaraj, R. Chatterjee, R. M. Allen, J. T. Roll, G. P. Roberts, and P. W. Ludden.** 1999. Requirement of NifX and other *nif* proteins for in vitro biosynthesis of the iron-molybdenum cofactor of nitrogenase. *J Bacteriol* **181:**2797-2801.

198. **Sharma, S. B., and E. R. Signer.** 1990. Temporal and spatial regulation of the symbiotic genes of *Rhizobium meliloti* in planta revealed by transposon Tn*5*-*gusA*. *Genes Dev* **4:**344-356.

199. **Shin, D. H., I. G. Choi, D. Busso, J. Jancarik, H. Yokota, R. Kim, and S. H. Kim.** 2004. Structure of OsmC from *Escherichia coli*: a salt-shock-induced protein. *Acta Crystallogr D Biol Crystallogr* **60:**903-911.

200. **Silby, M. W., P. B. Rainey, and S. B. Levy.** 2004. IVET experiments in *Pseudomonas fluorescens* reveal cryptic promoters at loci associated with recognizable overlapping genes. *Microbiology* **150:**518-520.

201. **Smith, H. E., H. Buijs, R. R. de Vries, H. J. Wisselink, N. Stockhofe-Zurwieden, and M. A. Smits.** 2001. Environmentally regulated genes of *Streptococcus suis*: identification by the use of iron-restricted conditions in vitro and by experimental infection of piglets. *Microbiology* **147:**271-280.

202. **Soberon, M., C. Morera, A. Kondorosi, O. Lopez, and J. Miranda.** 2001. A purine-related metabolite negatively regulates *fixNOQP* expression in *Sinorhizobium meliloti* by modulation of *fixK* expression. *Mol Plant Microbe Interact* **14:**572-576.

203. **Spaink, H. P.** 2000. Root nodulation and infection factors produced by rhizobial bacteria. *Annu Rev Microbiol* **54:**257-288.

204. **Stanfield, S. W., L. Ielpi, D. O'Brochta, D. R. Helinski, and G. S. Ditta.** 1988. The *ndvA* gene product of *Rhizobium meliloti* is required for beta-(1----2)glucan production and has homology to the ATP-binding export protein HlyB. *J Bacteriol* **170:**3523-3530.

205. **Stauffer, L. T., A. Ghrist, and G. V. Stauffer.** 1993. The *Escherichia coli gcvT* gene encoding the T-protein of the glycine cleavage enzyme system. *DNA Seq* **3:**339-346.

206. **Streit, W. R., R. A. Schmitz, X. Perret, C. Staehelin, W. J. Deakin, C. Raasch, H. Liesegang, and W. J. Broughton.** 2004. An evolutionary hot spot: the pNGR234b replicon of *Rhizobium sp.* strain NGR234. *J Bacteriol* **186:**535-542.

207. **Surin, B. P., and J. A. Downie.** 1988. Characterization of the Rhizobium leguminosarum genes *nodLMN* involved in efficient host-specific nodulation. *Mol Microbiol* **2:**173-183.

208. **Swanson, J. A., J. T. Mulligan, and S. R. Long.** 1993. Regulation of *syrM* and *nodD3* in *Rhizobium meliloti*. *Genetics* **134:**435-444.

209. **Szeto, W. W., B. T. Nixon, C. W. Ronson, and F. M. Ausubel.** 1987. Identification and characterization of the *Rhizobium meliloti ntrC* gene: *R. meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. *J Bacteriol* **169:**1423-1432.

210. **Talibart, R., M. Jebbar, G. Gouesbet, S. Himdi-Kabbab, H. Wroblewski, C. Blanco, and T. Bernard.** 1994. Osmoadaptation in rhizobia: ectoine-induced salt tolerance. *J Bacteriol* **176:**5210-5217.

211. **Tate, R., A. Riccio, M. Merrick, and E. J. Patriarca.** 1998. The *Rhizobium etli amtB* gene coding for an NH4+ transporter is down-regulated early during bacteroid differentiation. *Mol Plant Microbe Interact* **11:**188-198.

212. **Thomsen, L. E., M. Pedersen, M. Norregaard-Madsen, P. Valentin-Hansen, and B. H. Kallipolitis.** 1999. Protein-ligand interaction: grafting of the uridine-specific determinants from the CytR regulator of *Salmonella typhimurium* to *Escherichia coli* CytR. *J Mol Biol* **288:**165-175.

213. **Timmers, A. C., M. C. Auriac, and G. Truchet.** 1999. Refined analysis of early symbiotic steps of the *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. *Development* **126:**3617-3628.

214. **Triccas, J. A., W. J. Britton, and B. Gicquel.** 2001. Isolation of strong expression signals of *Mycobacterium tuberculosis*. *Microbiology* **147:**1253-1258.

215. **Uttaro, A. D., G. A. Cangelosi, R. A. Geremia, E. W. Nester, and R. A. Ugalde.** 1990. Biochemical characterization of avirulent *exoC* mutants of *Agrobacterium tumefaciens*. *J Bacteriol* **172:**1640-1646.

216. **van Brussel, A. A. N., R. Bakhuizen, P. C. van Spronsen, H. P. Spaink, T. Tak, B. J. J. Lugtenberg, and J. W. Kijne.** 1992. Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharides of *Rhizobium. Science* **257:**70-72.

217. **van Spronsen, P. C., R. Bakhuizen, A. A. van Brussel, and J. W. Kijne.** 1994. Cell wall degradation during infection thread formation by the root nodule bacterium *Rhizobium leguminosarum* is a two-step process. *Eur J Cell Biol* **64:**88-94.

218. **Vasse, J., F. de Billy, S. Camut, and G. Truchet.** 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol* **172:**4295-4306.

219. **Walker, S. A., and J. A. Downie.** 2000. Entry of *Rhizobium leguminosarum bv.viciae* into root hairs requires minimal Nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. *Mol Plant Microbe Interact* **13:**754-762.

220. **Wang, C. H., C. Y. Lin, Y. H. Luo, P. J. Tsai, Y. S. Lin, M. T. Lin, W. J. Chuang, C. C. Liu, and J. J. Wu.** 2005. Effects of oligopeptide permease in group a streptococcal infection. *Infect Immun* **73:**2881-2890.

221. **Wang, J., A. Mushegian, S. Lory, and S. Jin.** 1996. Large-scale isolation of candidate virulence genes of *Pseudomonas aeruginosa* by in vivo selection. *Proc Natl Acad Sci U S A* **93:**10434-10439.

222. **Wang, L. X., Y. Wang, B. Pellock, and G. C. Walker.** 1999. Structural characterization of the symbiotically important low-molecular-weight succinoglycan of Sinorhizobium meliloti. *J Bacteriol* **181:**6788-6796.

223. **Wei, X., and W. D. Bauer.** 1999. Tn*5*-induced and spontaneous switching of *Sinorhizobium meliloti* to faster-swarming behavior. *Appl Environ Microbiol* **65:**1228-35.

224. **Weinstein, M., R. C. Roberts, and D. R. Helinski.** 1992. A region of the broad-host-range plasmid RK2 causes stable in planta inheritance of plasmids in *Rhizobium meliloti* cells isolated from alfalfa root nodules. *J Bacteriol* **174:**7486-7489.

225. **Wexler, M., K. H. Yeoman, J. B. Stevens, N. G. de Luca, G. Sawers, and A. W. Johnston.** 2001. The *Rhizobium leguminosarum tonB* gene is required for the uptake of siderophore and haem as sources of iron. *Mol Microbiol* **41:**801-816.

226. **Williams, K., P. C. Oyston, N. Dorrell, S. Li, R. W. Titball, and B. W. Wren.** 2000. Investigation into the role of the serine protease HtrA in *Yersinia pestis* pathogenesis. *FEMS Microbiol Lett* **186:**281-286.

227. **Yang, S., N. T. Perna, D. A. Cooksey, Y. Okinaka, S. E. Lindow, A. M. Ibekwe, N. T. Keen, and C. H. Yang.** 2004. Genome-wide identification of plant-upregulated genes of *Erwinia chrysanthemi* 3937 using a GFP-based IVET leaf array. *Mol Plant Microbe Interact* **17:**999-1008.

228. **Yang, W., L. Ni, and R. L. Somerville.** 1993. A stationary-phase protein of *Escherichia coli* that affects the mode of association between the trp repressor protein and operator-bearing DNA. *Proc Natl Acad Sci U S A* **90:**5796-5800.

229. **Yao, S. Y., L. Luo, K. J. Har, A. Becker, S. Ruberg, G. Q. Yu, J. B. Zhu, and H. P. Cheng.** 2004. *Sinorhizobium meliloti* ExoR and ExoS proteins regulate both succinoglycan and flagellum production. *J Bacteriol* **186:**6042-6049.

230. **Yeh, K. C., M. C. Peck, and S. R. Long.** 2002. Luteolin and GroESL modulate in vitro activity of NodD. *J Bacteriol* **184:**525-530.

231. **York, G. M., and G. C. Walker.** 1997. The *Rhizobium meliloti exoK* gene and *prsD/prsE/exsH* genes are components of independent degradative pathways which contribute to production of low-molecular-weight succinoglycan. *Mol Microbiol* **25:**117-134.

232. **Zhan, H. J., and J. A. Leigh.** 1990. Two genes that regulate exopolysaccharide production in *Rhizobium meliloti*. *J Bacteriol* **172:**5254-5259.

233. **Zimmerman, J. L., W. W. Szeto, and F. M. Ausubel.** 1983. Molecular characterization of Tn*5*-induced symbiotic (Fix-) mutants of *Rhizobium meliloti*. *J Bacteriol* **156:**1025-1034.