IDENTIFICATION OF GENES IN *SINORHIZOBIUM MELILOTI* RELEVANT TO SYMBIOSIS AND COMPETITIVENESS USING SIGNATURE TAGGED TRANSPOSONS

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES

DOKTOR DER NATURWISSENSCHAFTEN (DR. RER. NAT)

DER FAKULTÄT FÜR BILOGIE DER UNIVERSITÄT BIELEFELD

VORGELEGT VON **NATALIYA POBIGAYLO** AUS LEMBERG

JUNI 2006

INDEX

A. SUMMARY

Sinorhizobium meliloti is a model organism for studies of plant-microbe interactions. This Gram-negative soil bacterium can enter an endosymbiosis with alfalfa plants through the formation of nitrogen-fixing nodules. *S. meliloti* genome sequence determination has provided a basis for different approaches of functional genomics to this bacterium. One of these approaches is gene disruption with subsequent analysis of mutant phenotypes.

Usually, the selection of mutants that can survive under a certain condition is simple and efficient, and can be performed using a mixture of different mutants. However, the selection of mutants that have an attenuated phenotype in a tested condition is problematic, because all mutants have to be checked separately one by one. A microarray-based signature-tagged mutagenesis (STM) strategy can overcome this problem.

Signature-tagged mutagenesis is based on a collection of mutants split into sets, where each mutant is modified by one or more different signature tags. The tags are short DNA segments that are unique for each mutant in a set and can be amplified using invariant or specific priming sites. Tagged mutants from the same set are pooled prior to an experiment, and each mutant in the mixture can be identified based on the unique tag in its genome. The presence of a particular tag in the mixture can be detected through a hybridization of amplified products to a microarray containing tag-specific probes. In order to integrate the signature tags into the genome, a strategy based on libraries of tag-carrying transposons can be used.

In the transposon-based STM, the number of mutants that can be pooled together in one experiment depends on the quantity of transposons containing different tags. In this study, the construction of a novel set of 412 mTn*5*-based signature-tagged transposons was performed, which is the largest tagged transposon set reported so far, and this set was applied to the mutagenesis of *S. meliloti*. To achieve a high specificity of tag detection, each transposon was bar coded by two signature tags. In order to generate defined, non-redundant sets of signature-tagged *S. meliloti* mutants for subsequent experiments, 12000 mutants were constructed and insertion sites for more than 5000 mutants were determined.

One set consisting of 378 mutants was used in a validation experiment to identify mutants showing altered growth patterns. This and one additional mutant set, which also contained 378 mutants, were used in the experiment for identification of genes in *S. meliloti* relevant to symbiosis and competitiveness. From all the tested mutants, 69 were found to be attenuated. For 25 mutants, attenuated phenotypes were checked by a different method. As a result, the attenuated competitiveness phenotype could be confirmed for 22 from 25 tested mutants. Four mutants were found to have additional symbiotic defects: Fix- phenotype (*ccmC* mutant), low efficiency of nodulation (*asnO* mutant) and delayed nodulation (*cysG* and *metA* mutants).

B. INTRODUCTION

B.1 NITROGEN FIXATION

Nitrogen is one of the most abundant elements on Earth and, next to hydrogen, oxygen and carbon, one of the most important components of the living cells. All of the nutritional nitrogen is obtained by humans and animals directly or indirectly from plants. Plants, in turn, acquire nitrogen from two principal sources: the soil, through commercial fertilizer and/or mineralization of organic material, and the atmosphere, through nitrogen fixation by symbiotic bacteria (Vance, 2001).

The process of biological nitrogen fixation is limited to prokaryotes. The group of prokaryotes that can fix nitrogen due to an evolutionarily conserved nitrogenase protein complex is diverse and contains both eubacteria and archaea (Zehr *et al.*, 2003). Limiting factors in maintaining a high rate of nitrogen fixation are the large quantity of energy needed to break the N – N triple bond, and the high sensitivity of nitrogenase to oxygen. Different mechanisms have evolved to overcome this limitation. For the anaerobic microorganisms (*Clostridium*) that predominate in waterlogged soils, where organic substrates are available but oxygen supply to the micro-environment of the bacteria is severely restricted, this is not problematic (Chen and Johnson, 1993). In cyanobacteria, nitrogen fixation occurs in special cells known as heterocysts which do not photosynthesize but are devoted solely to N_2 fixation (Golden and Yoon, 2003). The obligate aerobe *Azotobacter* protects the nitrogenase in two ways: producing special auxiliary proteins which cause nitrogenase to aggregate when exposed to oxygen and by very high respiratory rate that creates a nearly anoxygenic environment in the cytoplasm of the cells (Poole and Hill, 1997). The most efficient way of nitrogen fixation, however, is the symbiosis between bacteria and plants, where the plant supplies the carbon source for the reduction of nitrogen and at the same time, creates microaerobic environment to protect the nitrogenase. Bacteria, in turn, provide the plant with the nitrogen in form of ammonia. This process is mostly restricted to a limited number of bacterial groups, including the genera *Rhizobium*, *Mesorhizobium*, *Sinorhizobium* (*Ensifer*), and *Bradyrhizobium* (collectively referred to as rhizobia), and *Frankia*. All these genera except *Frankia* belong to the *Rhizobiacae* family in the alpha-proteobacteria, and enter symbiosis with plants from family Leguminosae. *Frankia* is an actinomycete that enters symbiosis with plants from the families Rosaceae, Casuarinaceae, Betulaceae, Myricaceae, Rhamnaceae, Datisticaceae, Eleganaceae and Coriariaceae (Gage, 2004).

The subject of this study is *Sinorhizobium meliloti*, an organism from the genera *Sinorhizobium* that is able to establish nitrogen fixing symbiosis with plants of genera *Medicago*, *Melilotus* and

Trigonella. The *S. meliloti* genome sequence was determined and consists of one chromosome (3.65 Mb) and two megaplasmids, pSymA (1.36 Mb) and pSymB (1.68 Mb) (Galibert *et al.*, 2001).

B.2 SYMBIOSIS BETWEEN RHIZOBIA AND THE LEGUMINOUS PLANTS

B.2.1 Determinate and indeterminate nodules

The symbiosis between Rhizobia and their host plants results in the formation of specialized symbiotic organs called nodules. Nodules are formed by the plant tissue and occupied by the nitrogen-fixing bacteria. Mature nodules are either of determinate or indeterminate type.

Determinate nodules are formed on tropical and subtropical legumes (*Glycine max*, *Phaseolus vulgaris*). They are characterized by meristemic activity that disappears early after nodulation that causes globose shape of nodules. Nodule primordia initiate in the outer cortex (Fig. B1) (Brewin, 1991).

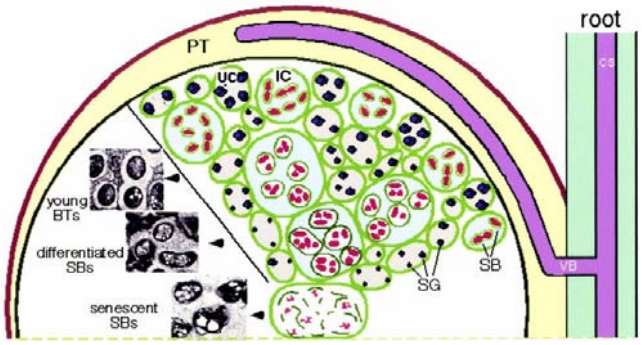


Fig. B1. Scheme of the determinate globose nodule. CS, central stele; VB, vascular bundle; PT, peripheral tissue; BTs, bacteroids; SBs, symbiosomes; SG, starch grain; IC, invaded cell; UC, uninvaded cell (Patriarca *et al.*, 2002).

Indeterminate nodules are usually formed on temperate legumes (e.g., from *Medicago sativa*, *Pisum sativum*, and *Vicia hirsuta*) and are characterized by persistent meristematic activity, that causes elongated shape of nodules. The central tissue of such nodules consists of a number of distinct zones containing invaded cells at different stages of differentiation, in which symbiosomes also show a progressive differentiation (Fig. B2). Nodule primordia initiate in inner cortex (Patriarca *et al.*, 2002; Pawlowski and Bisseling, 1996).

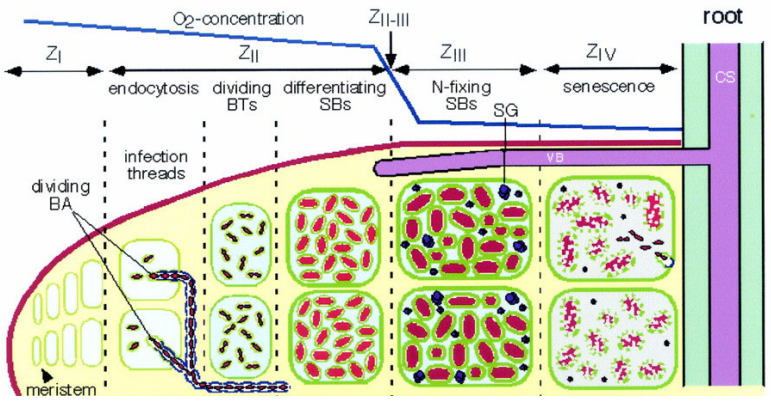


Fig. B2. Scheme of the indeterminate elongated nodule. BA, bacteria; BTs, bacteroids; SBs, symbiosomes; SG, starch grain; CS, central stele; VB, vascular bundle. The nodule zones (Z) are indicated (Patriarca *et al.*, 2002).

B.2.2 Formation of indeterminate nodules

Formation of nodule is a complicated and selective process, characterized by the exchange of molecular signals between the symbiontes. At the first step of the interaction, bacteria sense the specific compounds of plant root exudates. Amino acids, dicarboxylic acids and flavonoids present in the root exudates induce positive chemotaxis of rhizobia (van Rhijn and Vanderleyden, 1995). Furthermore, bacteria respond to the flavonoids by production of the lipochitooligosaccharide signalling molecules known as Nod factors.

At the second step of nodule formation, the rhizobia attach to the plant root surface. For many rhizobia, primary target sites for infection are young growing root hairs but there are no exclusive loci for rhizobial attachment. In some conditions, host plant lectins specifically bind polysaccharides structures present on the bacterial cell surface at the stage of attachment (Brencic and Winans, 2005; van Rhijn and Vanderleyden, 1995). At the same time, the root hair deforms by reinitiating tip growth with a changed growth direction in response to the Nod factors (Fig. B3). On *Vicia sativa* the first root hair deformation can be observed only 1 h after the application of Nod factors, while the maximal deformation is established in 3 h. A period of Nod factor-root contact as short as 10 min is sufficient to induce root hair deformation (Heidstra *et al.*, 1994). The morphological changes of root hair are caused by rearrangements of the actin filaments (Allen *et al.*, 1994) and increased cytoplasmic streaming (Heidstra *et al.*, 1994).

Root hairs curl in the way that the bacterial cells become entrapped in a pocket of host cell wall. After the entrapment, a local lesion of the root hair cell wall is formed by hydrolysis of the cell wall, so that infection can be initiated. The plasma membrane invaginates, and a tube-like structure, called an infection thread, is formed. The infection thread grows down the inside of the

root hair and into the body of the epidermal cell. Rhizobia inside the infection thread grow and divide, thereby keeping the tubule filled with bacteria (Gage, 2004). Even before the infection thread has crossed the epidermis, cells of cortex and pericycle respond in a local manner to the presence of rhizobia. Cells in the inner cortex dedifferentiate by entering the cell cycle; later these cells will form the nodule tissue. The group of dividing cortical cells is called the nodule primordium (Geurts and Bisseling, 2002; Timmers *et al.*, 1999).

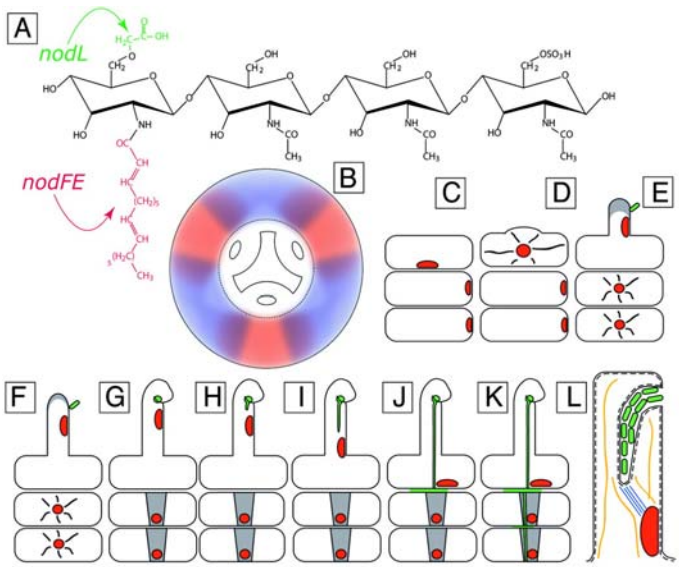


Fig. B3. Overview of the nodulation process in plants that form indeterminate nodules (Gage, 2004). (A) One form of Nod factor synthesized by *S. meliloti*. (B) Diagrammatic cross section of a root, showing gradients of an activating factor at protoxylem poles (blue) and an inhibitor at protophloem poles (red). Nodules are formed next to the protoxylem poles, which are at the ends of the Y-shaped structure depicted in the center the diagram, rather than above the protophloem poles, depicted as ovals. (C)–(K) Steps of the infection thread formation. (L) Enlarged view of the root hair shown in panel I.

At the next step, the infection thread branches in the nodule primordium, thereby increasing the number of sites from which bacteria can exit the thread and enter nodule cells. At this point, cells at the base of primordium establish a radial pattern consisting of a central tissue surrounded by peripheral tissues, and the cells at the apex of primordium form the meristem tissue, thus creating a young nodule (Pawlowski and Bisseling, 1996). Eventually, bacteria exit the infection thread, and the symbiosomes – compartments containing nitrogen-fixing rhizobia surrounded by the plant-derived peribacteroid membrane (Becker *et al.*, 2005) – are formed.

B.2.3 Structure of mature indeterminate nodules

Once inside nodule cells, the bacteria continue to differentiate and synthesize proteins required for nitrogen fixation. Ultrastructural studies of wild type nodules distinguish 5 steps in bacteroid differentiation (types 1 to 5), each of them being restricted to a defined histological region of the nodule (Zones I to IV) (Fig. B2) (Luyten and Vanderleyden, 2000; Vasse *et al.*, 1990).

Zone I contains meristematic tissue, situated at the apex of the nodule. This is a region of

actively dividing plant cells devoid of bacteria.

Zone II is called the infection zone. Here the bacteria enter the root cells via infection threads. Bacteria, released from the infection threads, are called type 1 bacteroids. These bacteroids divide and resemble free-living bacteria by size and cytoplasm content. They have a large periplasmic space, and the peribacteroid membrane (membrane of the plant origin that surrounds invading bacteria) appears irregular in shape due to local fusions with plant cytoplasmic vesicles. In the proximal part of Zone II, type 2 bacteroids are the most abundant. These bacteroids are elongated; their periplasmic and peribacteroid spaces are reduced, and the peribacteroid membrane is more regular in shape. The cell division stops once the type 2 bacteroid stage is reached. DNA replication stops several rounds after the cell division has stopped. Therefore, bacteroids have an increased DNA content comparing to the free-living cells. Nitrogen fixation does not take place in the Zone II.

Interzone II-III is a very restricted zone that contains only 3-4 layers of cells, separating the prefixation zone II and nitrogen-fixing Zone III. The Interzone II-III is easily identified by the abundance of amyloplasts (plastids involved in the synthesis and storage of starch). It contains bacteroids of type 3 which have stopped elongating and are about seven times longer than the free-living bacteria. The membranes surrounding each bacteroid, including the peribacteroid membrane, are smooth, often in contact with each other, with small periplasmic and peribacteroid spaces. Bacteroids do not fix nitrogen in the Interzone II-III, but the expression of the nitrogen fixation genes (*nif* and *fix*) already takes place.

Zone III is filled with the fully differentiated, nitrogen-fixing bacteroids of type 4. In mature nodules, this zone is spread over eight to twelve cell layers. In this zone, the leghaemoglobin is produced that gives a nodule its pink or red color. Leghaemoglobins are essential for symbiotic nitrogen fixation in legume nodules, since they bind oxygen molecules, protecting the nitrogenase.

In the proximal Zone III, the bacteroids stop fixing nitrogen and show a dispersion of the ribosome-enriched areas, thus becoming the bacteroids of type 5.

Zone IV is the senescence zone, located proximal to the point of attachment to the plant root. Here, both symbiotics partners degrade and the number of bacteroids gradually decreases. Ghost membranes of plant and bacteroid origin are the ultimate result of the senescing process.

B.2.4 Bacterial genes involved in competitiveness, efficient nodule formation and nitrogen fixation

B.2.4.1 Genes important for survival and competitiveness in the rhizosphere

Survival of bacteria in the rhizosphere is an important factor of competitiveness, especially in field conditions. Plant root exudates contain a great variety of different compounds, and nearly 5% to 21% of all photosynthetically fixed carbon can be transferred to the rhizosphere through the root exudates (Walker *et al.*, 2003). Therefore, the ability to transport and catabolize the compounds represented in the root exudates is crucial for the survival and competitiveness of bacteria in the plant rhizosphere. This is reflected by the content of rhizobial genomes: rhizobia have approximately 170 ABC transporters, compared to 47 in *E. coli*, so that they can access a far greater range of nutrients present at low concentrations in the soil and in the rhizosphere (Prell and Poole, 2006). Abiotic factors, such as temperature, osmotic pressure, UV light, and pH also play role in the selection and activity of microbes in soils and at the plant surface (Savka *et al.*, 2002).

Bacteria in the rhizosphere compete furthermore via production of and resistance to antibacterial compounds that can originate both from plant root system and from other soil organisms (Savka *et al.*, 2002). An example of such interaction is a mimosine system. The toxin mimosine is produced by the tree legume *Leucaena* and provides a nodulation competition advantage to mimosine-degrading *Rhizobium* strains (Soedarjo *et al.*, 1994). Another example is the production of an antirhizobial peptide trifolitoxin (TFX) by *Rhizobium leguminozarum* bv. trifolii T24. The TFX-producing, TFX-resistant rhizobial strains exhibited at least 20% higher greater nodule occupancy than the non-producing strains in field grown plants 2 years after inoculation (Robleto *et al.*, 1998).

Bacteria can also produce some selective growth-promoting factors, or stimulate the production of such factors by the plant. The best studied system of this type is the *Agrobacterium tumefaciens*-plant interaction. These pathogenic bacteria genetically engineer the host plant to synthesize carbon- and nitrogen-containing molecules (opines) that favour their growth and dissemination of pathogenicity (Savka *et al.*, 2002). Several Rhizobium strains and related legume-associated species have evolved the ability to synthesize carbon compounds called opine-like molecules due to the similarity of their action to the agrobacterial opines. Rhizopine, which is one of such molecules, is synthesized in the bacteroids during symbiosis in the plant nodules (Heinrich *et al.*, 2001). The ability to catabolize rhizopine provides a growth advantage to the rhizopine-degrading rhizobial strains.

B.2.4.2. Genes important for the formation of nodules and infection

At the stage of initiation and growth of infection threads, the genes encoding for the synthesis of signal molecules are especially important. The efficiency of the infection process depends also on the ability of bacteria to adapt to the conditions inside the infection thread and the plant cells. The following genes are crucial for the infection of the host plant by rhizobia.

nod genes

Nod factors initiate many of the developmental changes seen in the host plant early in the nodulation process, including root hair deformation, membrane depolarization, intracellular calcium oscillations, and the initiation of cell division in the root cortex, which establishes a meristem and nodule primordium (Ehrhardt *et al.*, 1992; Gage, 2004). Nod factors consist of an oligomeric backbone of β-1,4-linked N-acetyl-D-glucosaminyl residues, N-acetylated at the nonreducing-terminal residue (Kamst *et al.*, 1998). Qualitative and quantitative aspects of the Nod factor populations are strain-specific. In different strains, 2 to 60 different individual Nod factors can be synthesized (D'Haeze and Holsters, 2002).

Bacterial genes involved in Nod factor synthesis and export are the *nod* (nodulation) genes. They are not expressed in free-living bacteria, with the exception of *nodD*, which is expressed constitutively (Mylona *et al.*, 1995). NodD protein is activated by the binding of the flavonoids and, in turn, activates the transcription of other *nod* genes (Goethals *et al.*, 1992), which encode enzymes involved in the synthesis of Nod factors. *S. meliloti* has three *nodD* genes. NodD1 activates expression of *nod* genes in presence of the plant flavonoids luteolin and methoxychalcone. NodD2 is activated by methoxychalcone and two betaines, trigonelline and stachydrine. NodD3 does not need a plant inducer, but the *nodD3* gene is regulated by NodD1 and another regulator, SyrM (Luyten and Vanderleyden, 2000). The complex regulation of the *nod* gene expression is necessary since the concentration and composition of Nod factors need to be optimal for the efficient nodulation of the host.

Mutations in the *nod* genes result either in the complete abolishment of nodulation, or in a delayed nodulation phenotype. Essential for the nodulation are the *nodABC* genes that encode for the enzymes involved in the synthesis of the chitin oligomer backbone (John *et al.*, 1993; Rohrig *et al.*, 1994), and *nodHPQ*, responsible for the formation of the sulphate group on the reducing sugar (Roche *et al.*, 1991). In addition, *S. meliloti* mutants of *ilvC* gene, which are auxotrophs for isoleucine and valine, have a not-nodulating (Nod^-) phenotype, because *nodABC* genes are not activated by the inducer luteolin in these mutants. This could be caused by the high concentration of the valine and isoleucine precursors that probably inhibit the expression of *nod* genes, or antagonize the inducing effect of luteolin (Aguilar and Grasso, 1991). Other *nod* genes involved in

efficient nodulation, are:

nodEF – products of these genes are involved in the synthesis of C16 polyunsaturated acyl chains (Fig B3). *nodEF* mutants have a delayed nodulation phenotype and show low number of infection threads (Demont *et al.*, 1993)

nodG – encodes, putatively, for a 3-oxoacyl-acyl carrier protein reductase (Lopez-Lara and Geiger, 2001). *nodG* mutants have a minor defect in the infection.

nodL – the product of this gene attaches the O-acetyl decoration at the non-reducing end. *nodL* mutants have a delayed nodulation phenotype, and the overall production of Nod factors is reduced 5- to 10-fold. Depending on the host, the symbiotic defects of *nodL* mutants range from a reduced quantity of infection threads (*M. truncatula* and *M. alba*) to the complete loss of nodulation (*M. lupine* or *M. littoralis*) (Luyten and Vanderleyden, 2000).

Genes of polysaccharides synthesis

Another group of rhizobial genes that play an important role in the infection process and in the survival of rhizobia inside the infection threads are the genes involved in the synthesis of the surface polysaccharides, such as exopolysaccharides (EPS), lipopolysaccharides (LPS), capsular polysaccharides (K antigens) and cyclic β-(1, 2)-glucans.

Exopolysaccharides are extracellular polysaccharides that accumulate on the surface with little or no cell association. *S. meliloti* produces two types of EPS. Succinoglycan (EPS I) is composed of octasaccharide repeating units containing one galactose and seven glucose residues (in molar ratio 1:7), joined by β-1,3, β-1,4 and β-1,6 glycosidic linkages. The single repeating unit is decorated by acetyl, pyruvyl and succinyl groups (Reinhold *et al.*, 1994). Galactoglucan (EPS II) is a polymer of disaccharide repeating unit composed of an acetylated glucose and one pyruvylated galactose coupled by α-1,3 and β-1,3 glycosidic bonds (Skorupska *et al.*, 2006).

Both EPS I and II are secreted in two major fractions of different molecular weight: HMW (High Molecular Weight), consisting of hundreds to thousands of repeating units (polymers of $10^6 – 10^7$ Da) and LMW (Low Molecular Weight) that represent monomers, dimers and trimers in case of EPS I and oligomers (15–20) in case of EPS II (Skorupska *et al.*, 2006).

The gene groups responsible for EPS I synthesis are *exo* and *exs* (Fig. B4). Mutations in genes, responsible for the synthesis of nucleotide sugar precursors, *exoB* and *exoC* not only result in the lack of EPS I production, but also affect the synthesis of EPS II, LPS and β-glucans (Leigh and Lee, 1988). Mutants of the *exoN* gene, which encodes for a UDP-glucose pyrophosphorylase, results in a decreased production of EPS I (Becker *et al.*, 1993b).

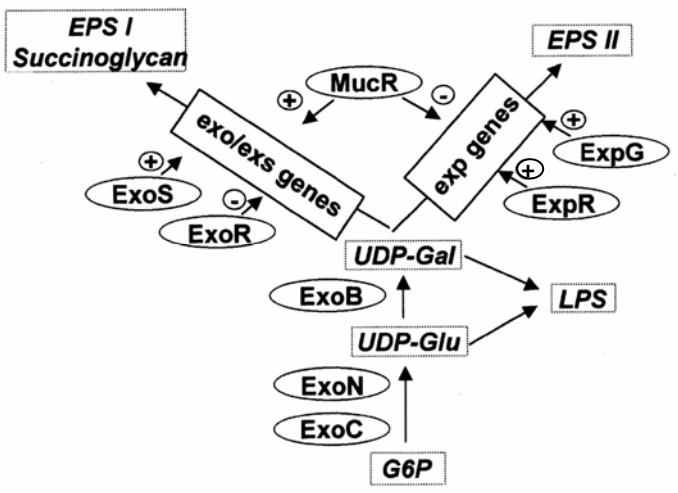


Fig. B4. Biosynthesis of exopolysaccharides. Modified from (Fraysse et al., 2003)

In *S. meliloti* strain Rm1021, mutations in *exoB, exoC* and in other *exo* genes such as *exoA*, *exoF, exoL*, *exoM, exoY, exoO, exoP, exoQ, exoT, exoU, exoV, exoW* or *exoY* that abolish the EPS I production, cause severe defects in symbiosis. *S. meliloti* clones that have a mutation in any of these genes are able to initiate root hair curling and infection thread formation, but the infection threads abort within the peripheral cells of developing nodules, resulting in a formation of empty nodules that do not fix nitrogen (Becker *et al.*, 1993b; Becker *et al.*, 1993c; Fraysse *et al.*, 2003; Reuber and Walker, 1993). *exoH* mutants that produce symbiotically nonfunctional high-molecular-weight EPS I lacking the succinyl modification, also induce formation of ineffective nodules that do not contain intracellular bacteria or bacteroids (Leigh *et al.*, 1987; Skorupska *et al.*, 2006). An *exoZ* mutant, which produces EPS I without the acetyl modification, exhibited a reduced efficiency in the initiation and elongation of infection threads, but produces nitrogen-fixing nodules (Cheng and Walker, 1998). *exoG* and *exoJ* mutants form a mixture of white and pink nodules (Long *et al.*, 1988).

Genes responsible for the synthesis of EPS II form the *exp* gene cluster. EPS II can functionally replace the EPS I in nodule invasion (Zhan *et al.*, 1989). In *S. meliloti* Rm2011 and Rm1021 EPS II is synthesized only under phosphate starvation (Zhan *et al.*, 1991) or when mutation in the regulatory gene *muc*R (Keller *et al.*, 1995) occurs. The restoration of the reading frame of the regulator gene *exp*R by the excision of an insertion sequence also leads to the production of EPS II (Pellock *et al.*, 2002). An additional gene, *expG* (*mucS*) positively regulates the expression of the *exp* genes (Baumgarth *et al.*, 2005; Rüberg *et al.*, 1999).

The role of EPSs in symbiosis is still not well understood. Their possible biological functions include protection of bacteria against environmental stresses, attachment of bacteria to the roots, structural role in the infection thread formation, release of bacteria from infection threads, suppression of plant defence responses, and protection against plant antimicrobial compounds. It

was also assumed that EPSs might serve as signalling molecules that trigger plant developmental response and as determinants of host-specificity (Skorupska *et al.*, 2006).

Rhizobial K polysaccharides (capsular polysaccharides) show structural analogies to group II K-antigens found in *Escherichia coli*. In *S. meliloti*, the synthesis of K-polysaccharides is performed by the products of *rkp* genes. In enteric bacteria, capsular polysaccharides form an adherent cohesive layer that contributes to the protection against abiotic factors (Roberts, 1996). K polysaccharides may have a similar function in rhizobia, protecting the bacteria also against legume defence products during the infection process (Becker *et al.*, 2005; Roberts, 1996). Furthermore, it was shown that *S. meliloti* Rm41 *exoB* mutant, unable to synthesize EPS I and II, successfully infects the host plant due to its symbiotically active K polysaccharide. A mutation in the *rkpZ* gene that affected the size distribution of K-polysaccharide led to the loss of its symbiotic activity. The introduction of *rkpZ* from Rm41 into the *exoB* mutant of strain Rm1021 partially restored the symbiotic efficiency (Becker *et al.*, 2005; Reuhs *et al.*, 1995).

Lipopolysaccharides (LPS) are anchored in the outer membrane of bacteria. They consist of a lipid A anchor, associated with a core polysaccharide or an O-antigen polysaccharide moiety (Becker *et al.*, 2005). The main gene clusters encoding for the lipopolysaccharide synthesis are *lpx*, *lps* and *acp* clusters. Lipopolysaccharides may be involved in symbiosis not only at the infection stage, but also during release of bacteria into plant cells and formation of symbiosomes. Characteristically for rhizobia, their lipid As are acylated with a very long fatty acid chain, 27-hydroxyoctacosanoic acid. Biosynthesis of this substitution requires a special acyl carrier protein, AcpXL, which serves as a donor of C28 (ω-1)-hydroxylated fatty acid (Brozek *et al.*, 1996). The *acpXL* mutant of *S. meliloti* Rm1021 displayed delayed nodulation of *Medicago sativa* and a reduced competitive ability (Sharypova *et al.*, 2003). It was hypothesized that very long-chain fatty acid could stabilize the bacterial membrane, therefore supporting the adaptation of rhizobia to the conditions inside the nodule (Becker *et al.*, 2005). It has been also demonstrated that LPSs of *S. meliloti* suppress the defence response to the yeast elicitor in a *Medicago sativa* cell suspension culture. Addition of either LPS or lipid A, or lipid A lacking the very long acyl chain, inhibited the defence reactions in a concentration-dependent manner (Scheidle *et al.*, 2005).

Cyclic β-(1,2)-glucans are found almost exclusively in rhizobia; their synthesis is directed by the products of *exo* and *ndv* genes. Cyclic β-glucans reach levels from 5% to 20% of cellular dry weight and are predominantly localized within the periplasmic compartment (Luyten and Vanderleyden, 2000). *S. meliloti ndv* mutants, affected in the biosynthesis of cyclic β-glucans, induce small white pseudonodules, devoid of bacteroids and containing no infection threads (Dylan *et al.*, 1986). Accumulation of cyclic β-glucans may allow bacteria to adapt to changing osmotic conditions during plant infection. Their role as signalling molecules has also been

suggested (Dylan *et al.*, 1990; Luyten and Vanderleyden, 2000).

Synthesis of amino acids and nucleotides

The fast-growing cells in the infection thread require supply of nitrogen and carbon from the plant. Indeed, plant provides these bacteria with most of the amino acids at the early stages of symbiosis. *R. etli* strains defective in the biosynthesis of amino acids, such as aromatics, aspartate, cysteine, glutamate, glycine, histidine, isoleucine/valine, leucine, lysine, phenylalanine, and tryptophan, with the exception of arginine and methionine auxotrophs, are able to induce nodules. Analogous observations have been reported in the case of *S. meliloti* (Kerppola and Kahn, 1988; Patriarca *et al.*, 2002). In the case of the isoleucine/valine auxotrophs which have a Nod^- phenotype, the symbiotic defect is not caused by the inaccessibility of the respective amino acids (Chapter B.2.4.2).

In contrast, nucleotide compounds are not provided by the plant in the infection threads. The strains of *B. japonicum*, *R. etli*, *Rhizobium fredii*, and *R. leguminosarum* that are auxotrophic for the synthesis of purines or pirimidines, induce only empty pseudonodules (Patriarca *et al.*, 2002).

phoCDET genes

A *S. meliloti phoCDET* mutant shows a delayed infection, blocked at an early stage, and the bacteria are not released from the infection threads. The *phoCDET* cluster encodes for an ABC-type transport system, responsible for the uptake of phosphate and phosphonates. It was suggested that the symbiotic defect of the *phoCDET* mutant results from the failure to obtain sufficient phosphorus for growth during the infection process (Bardin *et al.*, 1996).

pha genes

A *S. meliloti* mutant carrying a Tn*5* inserion in the *pha* gene cluster induced empty nodules with aborted infection threads. The *pha* gene cluster encodes for a K^+/H^+ antiporter that regulates internal pH and is required for the adaptation to the alkaline pH in the presence of K^+. It was suggested that the propagation of *pha* mutants in the infection thread may be inhibited by the alkaline pH and/or by the presence of K^+ ions (Luyten and Vanderleyden, 2000).

bacA gene

A *S. meliloti bacA* mutant is blocked in the bacteroid development. Mutant bacteria are able to enter the root and form infection threads and are released into the plant cytoplasm. However, in the plant cytoplasm bacteria degrade at the stage of type 1 bacteroids. *bacA* encodes for an inner membrane protein with seven membrane spanning domains. The exact function of the BacA protein is still not identified. *bacA* mutants are sensitive to the detergent deoxycholate and resistant to the antibiotic bleomycin. It was also shown that BacA is involved in the modification of LPS by the very-long-chain fatty acid (Ferguson *et al.*, 2004; Ichige and Walker, 1997).

B.2.4.3. Genes important for the efficient nitrogen fixation

Nitrogen fixation is a highly energy-dependent process. Therefore, it is not surprising that not only the genes encoding for nitrogenase components are crucial for the nitrogen fixation, but also the genes encoding for electron transport chain proteins, for proteins of metabolic pathways and for the dicarboxilic acid transporter protein. Disruption of the genes responsible for the nitrogen fixation might influence the composition of bacteria inside the nodule. It has been proposed that the symbiosome has properties of the lytic compartment continuously being neutralized by ammonia which is exported by the bacteroids (Kannenberg and Brewin, 1989). According to this hypothesis, the lack of bacterial nitrogen fixation would lead to the degradation of bacteroids (Mylona *et al.*, 1995).

The following genes were shown to be important for nitrogen fixation:

nif genes

All the *nif* genes are essential for nitrogen fixation. *nifA* gene encodes a regulatory protein, *nifHDK* genes encode the structural proteins of the nitrogenase enzyme, *nifENB* encode enzymes involved in biosynthesis of the nitrogenase Fe-Mo cofactor, and *nifSWX* genes encode proteins of unknown functions that are required for full nitrogenase activity (Brencic and Winans, 2005). Downstream of the *nifB* gene in *S. meliloti*, a small gene *fdxN*, encoding ferredoxin-like electron transfer protein, is situated. This gene is cotranscribed with *nifENB* and is also absolutely necessary for nitrogen fixation (Klipp *et al.*, 1989).

fix genes

The *fixL*, *fixJ*, and *fixK* genes encode regulatory proteins. In *S. meliloti*, the FixL-FixJ two-component system is the master regulator of all *nif* and *fix* genes. FixL is a membrane-bound histidine kinase, which autoactivates by phosphorylation in response to low levels of oxygen, and then phosphorylates the FixJ protein. Phosphorylated FixJ in turn activates transcription of the regulatory *fixK* and *nifA* genes, whose products regulate transcription of the rest of the nitrogen fixation genes (Brencic and Winans, 2005). Null mutations in *fixL* or *fixJ* genes of *S. meliloti* result in the formation of symbiotically inefficient nodules (Fischer, 1994).

The *fixABCX* genes might code for an electron transport chain to nitrogenase. Mutations in any one of the *fixABCX* genes of *S. meliloti*, *B. japonicum*, and *A. caulinodans* completely abolish nitrogen fixation (Earl *et al.*, 1987; Fischer, 1994).

All four *fixGHIS* gene products are predicted to be transmembrane proteins. FixG is may be involved in a redox process, since its sequence shows similarity to bacterial ferredoxins. FixI is homologous to the catalytic subunit of bacterial and eukaryotic ATPases involved in cation pumping; it is probably a symbiosis-specific cation pump whose function is coupled to the redox

reaction catalyzed by the FixG subunit. *S. meliloti* mutant strains carrying a transposon insertion in *fixG, fixH,* or *fixI* are unable to fix nitrogen (Kahn *et al.*, 1989).

fixNOQP encode the membrane-bound cytochrome oxidase required for respiration of the rhizobia in low-oxygen environments. In *S. meliloti*, *fixNOQP* is a duplicated region linked to the regulatory genes *fixLJ* and *fixK*. *S. meliloti* mutant strains deleted for both *fixNOQP* regions are defective in symbiotic nitrogen fixation (Fischer, 1994).

cycHJKL cluster

The *cycHJKL* gene cluster encodes for a multisubunit cytochrome *c* lyase that is essential for the biogenesis of all cellular *c*-type cytochromes. Particularly, it is involved in the attachment of the heme to *fixO*-encoded cytochrome *c* (Kereszt *et al.*, 1995).

dctA gene

In *S. meliloti*, *dctA* gene encodes for a high affinity transporter (permease) of C_4-dicarboxylic acids malate, fumarate succinate and oxaloacetate, the amino acid aspartate, and the cyclic monocarboxylate orotate. Dicarboxylate transport plays an important role in the operation of an effective, nitrogen-fixing symbiosis. Dicarboxylates were shown to be the major energy and carbon source for the nitrogen-fixing rhizobia. Expression of the *dctA* gene is controlled by a two component kinase regulatory system, DctB/DctD. *dctA* mutants exhibit a clear non-fixing (Fix^-) phenotype (Yurgel and Kahn, 2004).

dme and genes encoding for TCA cycle enzymes

dme gene encodes for NAD^+- dependent malic enzyme which converts malate to pyruvate and simultaneously reduces NAD^+ to NADH. The *dme* gene is expressed constitutively in free-living cells and in the bacteroids (Driscoll and Finan, 1997). DME activity is very important in bacteroids (the *dme* mutants have a Fix^- phenotype), since the nitrogen-fixing bacteria use C_4-dicarboxylic acids as the main energy source for the nitrogen fixation, and feed them directly into the TCA cycle. To use a C_4-dicarboxylic acid such as malate, one molecule is oxidized to oxaloacetate by malate dehydrogenase, and the second one is oxidatively decarboxylated by the malic enzyme to pyruvate. The pyruvate is then oxidized to acetyl-CoA and it condensed with oxaloacetate to citrate, enabling the TCA cycle to proceed (Prell and Poole, 2006). According to this, the *S. meliloti* mutants for other TCA cycle proteins as citrate synthase (Mortimer *et al.*, 1999), isocitrate dehydrogenase (McDermott and Kahn, 1992), succinate dehydrogenase (Gardiol *et al.*, 1982) and α-ketoglutarate dehydrogenase (Duncan and Fraenkel, 1979) also have a Fix^- phenotype. However, α-ketoglutarate dehydrogenase mutant of *Bradyrhizobium japonicum* has a delayed Fix^+ phenotype (Green and Emerich, 1997), and an aconitase mutant of *B. japonicum* can fix nitrogen normally (Thöny-Meyer and Künzler, 1996).

The symbiosis-related genes, listed in this chapter, represent the real picture only partially. Many genes, whose disruption has a more subtle effect on symbiosis, have still not been identified. Moreover, among the 6,204 possible open reading frames of *S. meliloti*, only 59% have been assigned a possible function (Galibert *et al.*, 2001). The methods for identification of unknown gene functions are described in the next chapter.

B.3 SYSTEMATIC APPROACHES TO STUDY THE GENES IMPORTANT FOR THE SYMBIOSIS AND COMPETITIVENESS

The classical approaches to study the function of genes are based on investigation of particular genes and their functions. Gene disruption with subsequent analysis of mutant phenotype, and promoter-reporter fusions are usual tools for such type of studies. However, these methods are laborious and require much time for analysis of every single mutant. The availability of the 6.7 Mb *S. meliloti* genome sequence has enabled (or eased) the application of several large-scale genomic methods to study the function of this organism's genes.

B.3.1 Proteomics

Proteomics is the large-scale study of proteins, particularly their structures and functions. It comprises protein separation and identification, protein sequence analysis, determination of protein structures in three-dimensional space, studies of interactions of proteins on the atomic, molecular and cellular levels, post-translational modifications of proteins, and mapping of the location of proteins and protein-protein interactions in whole cells during key cell events.

The key technologies used in proteomics include (Twyman, 2004; Wilkins *et al.*, 1997):

- gel electrophoresis (one- and two-dimensional), used to identify the relative mass of a protein and its isoelectric point;
- tandem mass spectrometry combined with 2-D electrophoresis, used to identify and quantify proteins;
- affinity chromatography, yeast two hybrid techniques, fluorescence resonance energy transfer (FRET), and Surface Plasmon Resonance (SPR); used to identify protein-protein and protein-DNA binding reactions;
- X-ray crystallography and nuclear magnetic resonance, used to characterize the three-dimensional structure of peptides and proteins;
- X-ray tomography, used to determine the location of labeled proteins or protein

complexes in an intact cell.

A proteomic examination of *S. meliloti* strain Rm1021 was undertaken using a combination of two-dimensional gel electrophoresis and peptide mass fingerprinting by Matrix-assisted laser adsorption ionization time of flight (MALDI-TOF) to analyze the protein expression in the symbiotic state and during carbon or phosphate starvation (Djordjevic *et al.*, 2003). As a result, a total of 27 putative nodule-specific proteins and 35 nutrient-stress-specific proteins were identified. These proteins were related to nitrogen fixation, heme synthesis, protection from oxidative stress, iron uptake, EPS/KPS synthesis and secretion, regulation as well as to other unclassified functions.

Recently, a proteomic approach was also used to analyze identify salt tolerance-related proteins in *Rhizobium etli* and *Sinorhizobium meliloti* (Shamseldin *et al.*, 2006). Using two-dimensional gel electrophoresis and MALDI-TOF, six proteins were found to be induced by 4% NaCl in *R. etli*, but they could not be identified due to the absence of *R. etli* genome sequence. In a second approach used in this work, soluble proteins from salt-induced or non-salt-induced cultures from *R. etli* strain EBRI 26 were separately labeled with different fluorescent cyano-dyes prior to 2D difference in gel electrophoresis. Results revealed that 49 proteins were differentially expressed after the addition of sodium chloride, 14 of them overexpressed and 35 - downregulated. In a similar experiment using *S. meliloti* strain Rm2011, four overexpressed and six downregulated proteins were identified. Particularly, carboxynospermidin decarboxylase, which plays an important role in the biosynthesis of spermidin, was identified as overexpressed, and the enzyme catalase was identified as downregulated.

B.3.2 DNA macro- and microarrays in the analysis of gene expression

DNA macro- and microarrays are used to measure the levels of abundance of defined mRNAs in the experimental conditions. In these methods, the known DNAs (*probe*) corresponding to the mRNAs of the studied organism, are affixed to a support at the pre-defined spots. The mRNAs, isolated from the sample, are reverse transcribed, labeled in a specific way and these molecules, called *targets*, are then hybridized to the array. After washing steps, the spots to which the targets have been hybridized are visualized.

DNA arrays can be classified as macro- or microarrays, depending on the format of the support. In the DNA macroarray method, nylon membranes are used, on which the array of gene-specific DNA fragments is spotted. The targets from the control and test samples are labeled with radioactive phosphate and hybridized to the separate membranes carrying the same array. The hybridization signals are detected by a phosphorimaging instrument. Log-transformed

ratios between signal intensities for each spot reflect the relative quantities of particular mRNAs in the control and experiment samples.

A DNA microarray is a collection of microscopic DNA spots attached to a solid surface, such as glass, plastic or silicon chip forming an array. Due to the small format of the microarray, a much smaller amount of the target has to be used. There are two main types of microarrays. In **spotted microarrays** (or **two-channel microarrays**), the probes are oligonucleotides, cDNA or PCR products, corresponding to mRNAs. Spotted microarray is hybridized with cDNA from two samples that have to be compared (experiment and control) that are labeled with two different fluorofores (in most cases, Cyanine 5 and Cyanine 3). The samples are mixed and hybridized to one single microarray that is then scanned, allowing the visualization of up- and down-regulated genes at once (Cummings and Relman, 2000). In **single-channel microarrays**, the short oligonucleotide probes, synthesized in photolithographic process, are designed to match several parts of single mRNAs. These microarrays give estimations of the absolute value of gene expression and therefore the comparison of two conditions requires the use of two separate microarrays (Lipshutz *et al.*, 1999).

A macroarray (Ampe *et al.*, 2003), a whole-genome spotted microarray (Rüberg *et al.*, 2003) and a whole-genome (including intergenic regions) single-channel microarray (Barnett *et al.*, 2004) have been developed and successfully used to determine the gene expression patterns in different conditions in *S. meliloti.*

B.3.3 Metabolomics

Metabolomics is the study of global metabolite profiles in a system (cell, tissue, or organism) under a given set of conditions. The analysis of the metabolome is particularly challenging due to the diverse chemical nature of metabolites. Metabolites are the result of the interaction of the organism's genome with its environment and are not merely the end product of gene expression but also form part of the regulatory system in an integrated manner (Rochfort, 2005). Recently, metabolite analysis was established for *Sinorhizobium meliloti* (Barsch *et al.*, 2004) The hydrophilic compounds of bacteria grown on minimal medium were analysed via GC-MS, and from about 200 peaks in each chromatogram, 65 compounds could be identified.

B.3.4 Promoter-reporter gene fusion mutant libraries and *in vivo* expression technology

To create a random library of promoter-reporter fusion mutants, a transposon carrying a reporter gene is used to perform the mutagenesis. In the case of the reporter gene having the same direction as the gene interrupted by the transposon, the expression of the interrupted gene can be studied based on the expression of the reporter gene. This method was successfully applied to the analysis of *S. meliloti* genes induced under nutrient deprived conditions and oxygen limitation. The collection of 5000 Rm2011 derivatives carrying Tn*5*-1063, a Tn*5*-based reporter transposon containing the *luxAB* genes of *Vibrio harveyii*, was used to identify genes induced by nitrogen and carbon deprivation (Milcamps *et al.*, 1998) and by oxygen limitation (Trzebiatowski *et al.*, 2001). As a result, 21 genes were found to be induced by nitrogen deprivation, 12 genes – by carbon deprivation, and 11 genes – by oxygen limitation. In a similar approach, transposon Tn*5*-1062, also carrying *luxAB* reporter genes, was used for mutagenesis, and four cold shock induced loci were identified in *S. meliloti* (O'Connell *et al.*, 2000).

For the identification of genes expressed in bacteroids under symbiotic conditions, a method called *in vivo* expression technology (IVET) can be used. This method was originally developed for the identification of genes expressed in parasitic bacteria in the animal host. Here, a library of plasmids is constructed that carry random genomic fragments fused to the promoterless gene, necessary for the survival in the host (reporter gene). This library is transferred into a strain, in which the reporter gene is deleted from the genome. The plasmids integrate into the genome by homologous recombination, and the resulting library of clones is used to infect the host. In the case that genomic fragment in front of reporter gene carries a promoter that functions in the host, the mutant survives and can be isolated from the model (reviewed in Angelichio and Camilli, 2002). IVET was successfully applied to study symbiotically induced genes in *S. meliloti*, using *bacA* (Oke and Long, 1999) and *exoY* (Zhang and Cheng, 2006) as reporter genes.

B.3.5 Signature-tagged mutagenesis

Usually, the selection of mutants that can survive under a certain condition is simple and efficient, and can be performed using a mixture of different mutants. However, the selection of mutants that have an attenuated phenotype in a tested condition is laborious and time consuming, because all mutants have to be checked separately one by one. The signature-tagged mutagenesis (STM) strategy (Hensel *et al.*, 1995) (Fig. B5) can overcome this problem.

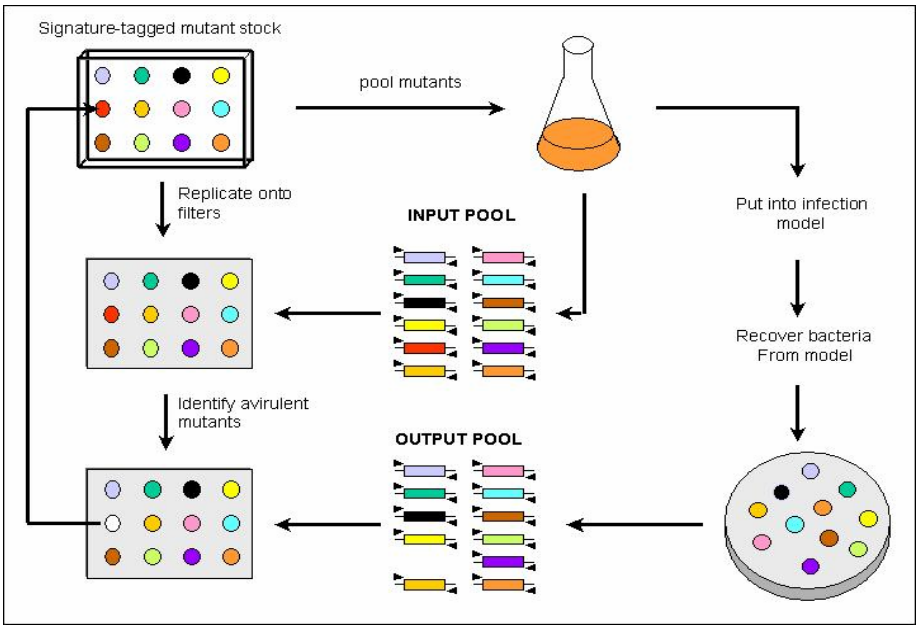


Fig. B5. Scheme of the signature-tagged mutagenesis. Mutants, carrying different signature tags, are mixed together prior to the experiment (in this case, an infection model). The input pool of tags is than compared to the output pool. Mutants, whose tags are not represented in the output pool, are impaired in growth under the experimental condition. Scheme from http://pollux.mpk.med.uni-muenchen.de/alpha1/forschung/haas/dia-stm.jpg.

Signature-tagged mutagenesis, similarly to IVET, was originally created to study the genes in parasitic bacteria relevant to infection and pathogenicity. STM strategy is based on a collection of mutants split in sets, where each mutant is modified by one or more different signature tags. The tags are short DNA segments that are unique for each mutant in a set and can be amplified using invariant (Chiang *et al.*, 1999) or specific (Lehoux *et al.*, 1999) priming sites.

Tagged mutants from the same set are pooled prior to an experiment, and each mutant in the mixture can be identified based on the unique tag in its genome. The presence of a particular tag in the mixture can be detected through hybridization of amplified products to dot blots (Chiang *et al.*, 1999), to a macroarray (Shea *et al.*, 2000) or to a microarray (Groh *et al.*, 2005; Karlyshev *et al.*, 2001; Winzeler *et al.*, 1999) containing tag-specific probes. In the case of variable priming sites being used, individual PCR reactions are performed using individual pairs of primers, and the presence of a tag in the mixture is detected through the presence of the amplification product in the respective PCR reaction (Lehoux *et al.*, 1999).

Two main strategies are used for the integration of the signature tags into the genome: the PCR targeting deletion strategy and the strategy based on libraries of signature tag carrying transposons. The PCR targeting strategy was used in the *Saccharomyces* Genome Deletion Project (Shoemaker *et al.*, 1996; Winzeler *et al.*, 1999) for gene disruption. Application of the targeted disruption protocol resulted in replacement of each gene with a *KanMX* marker cassette, where the cassette for each replacement was marked by an individual pair of 20-bp long tags.

In the transposon-based signature tagged mutagenesis, a library of transposons carrying signature tags is created, and these transposons are then used to perform the mutagenesis. This strategy, unlike the PCR targeted deletion, results in a library of random mutants. The production of transposon mutants is less labor and cost intensive than the production of deletion mutants and can also be performed in the organism whose genome is not sequenced. Moreover, one transposon library can be used for mutagenesis in all the organisms in which the chosen transposon can insert into the genome. The quantity of mutants generated by the transposon-based STM is not limited; however, the size of mutant pools depends on the quantity of differently tagged transposons. The largest tagged transposon library reported so far contains 192 transposons (Karlyshev *et al.*, 2001).

Most STM studies to date have applied the miniTn*5* transposon system (de Lorenzo *et al.*, 1990), which was used in the original STM screen (Hensel *et al.*, 1995). The mTn*5* transposons can transpose in a large spectrum of bacteria. However, some bacteria are nearly refractory to Tn*5*-derived transposons. Several recent studies in *Streptococcus pneumoniae* (Hava and Camilli, 2002), *Neisseria meningitidis* (Geoffroy *et al.*, 2003) and *Campylobacter jejuni* (Hendrixson and DiRita, 2004) have solved this problem by *in vitro* mutagenesis using transposons from the *mariner* family (Saenz and Dehio, 2005).

B.4 OBJECTIVES OF THIS WORK

Signature tagged mutagenesis approach is a powerful and easy-to-use method to identify bacterial genes important for competitiveness and survival in the host. This method can also be used to study symbiotic interactions (Hendrixson and DiRita, 2004), but it has never been applied to study rhizobia-legume symbiosis before. The application of signature-tagged transposon mutagenesis in *S. meliloti* was the main objective of this work. This objective can be split in several smaller tasks. At the first step, a novel, large library of double-tagged transposons had to be constructed. This library had to be used to obtain tagged *S. meliloti* mutants, and the transposon insertion sites had to be determined for a significant part of the mutant library. The tag detection system had to be created and verified in controlled conditions, and finally, the mutant library had to be screened in experiments *in planta* in order to identify genes important for competitiveness and survival in symbiotic conditions. The candidate mutants had also to be checked individually in order to verify the attenuated symbiotic phenotype.

Some crucial steps of this project (design of tags, analysis of transposon insertion sites and of tag - detection data) were performed in cooperation with Faculty of Technology (Technische Fakultät) and Bioinformatics Resource Facility at the Center of Biotechnology (Centrum für Biotechnologie) at Bielefeld University.

B.5 PRESENTATIONS

B.5.1 Publications

Nataliya Pobigaylo, Danijel Wetter, Silke Szymczak, Ulf Schiller, Stefan Kurtz, Folker Meyer, Tim W. Nattkemper, and Anke Becker. Construction of a Large Signature-Tagged Mini-Tn*5* Transposon Library and Its Application to Mutagenesis of *Sinorhizobium meliloti*. Appl. Environ. Microbiol. 2006 72: 4329-4337

Nataliya Pobigaylo, Silke Szymczak, Tim W. Nattkemper, and Anke Becker. Identification of genes in *Sinorhizobium meliloti* relevant to symbiosis and competitiveness using signature tagged transposon mutants. (*in preparation*)

B.5.2 Conference posters

September, 2005. "Prokagen – 2nd European conference on prokaryotic genomes", Göttingen/Germany

July, 2004. "6th European Conference on Nitrogen Fixation", Toulouse/France

May, 2004. "BioPerspectives 2004", Wiesbaden/Germany

C. RESULTS

At the beginning of this PhD project, signature tagged mutagenesis had already been applied to a large spectrum of organisms. The most interesting published works that used this system were the *Saccharomices* Genome Deletion Project (Winzeler *et al.*, 1999) and signature tagged transposon mutagenesis of *Yersinia* (Karlyshev *et al.*, 2001). Both of these projects used two short signature tags to mark each mutant, and a single-channel oligonucleotide microarray to detect the tags in the input and output pools. It was therefore decided to apply a similar approach to *S. meliloti*: double-tagged mutants in a combination with a two-channel microarray for the detection of tags. The design of tags was performed by Ulf Schiller and Stefan Kurtz (Praktische Informatik, Technische Fakultät, Universität Bielefeld). They designed 1498 signature tags, 24 nucleotides in length, with melting temperatures between 69.5°C and 70.5°C using the programs *DNASequenceGenerator* and *vmatch*. Less than 8 contiguous nucleotides of sequence identity were allowed for the tags. Half of the tags were equipped with sticky ends for the *Hin*dIII restriction site, and the other half with sticky ends for the *Kpn*I restriction site (H-tags and K-tags, respectively).

My contribution to the project started with design of the plasmid vector, that would carry the suitable mTn*5* transposon, and continued with cloning of the tags in this transposon.

C.1 CONSTRUCTION OF THE mTn*5*-STM TRANSPOSON AND THE CARRIER PLASMID pG18-STM

Plasmid pG18-STM (Fig. C2) carrying the transposase gene *tnpA*** and a modified mTn*5*-GNm transposon was constructed based on the vector pG18Mob2 (Kirchner and Tauch, 2003). The mini transposon mTn*5*-GNm, used in this study, contains the *nptII* resistance and a promotorless *gusA* reporter gene (Reeve *et al.*, 1999). The *gusA* gene is particularly suited for use as a reporter gene in rhizobia, because these bacteria, as well as their legume hosts, have no β-glucuronidase activity (Reeve *et al.*, 1999; Wilson *et al.*, 1995). The mTn*5*-GNm contains also a terminator sequence, situated after *nptII* gene. The mini Tn*5* contains no transposase gene; therefore, it has to be provided externally (on the plasmid). The plasmid pG18Mob2 already contained the RP4 *mob* region necessary for the transfer into *S. meliloti* from *E. coli* S17-1 cells, which have a chromosomally integrated copy of RP4 able to supply the transfer functions *in trans*. Furthermore, pG18Mob2 replicates in *E. coli*, but not in *S.meliloti* (suicide vector).

Since the tags were equipped with sticky ends for the *Hin*dIII and *Kpn*I restriction sites, these

restriction sites also had to be used for the cloning of tags into the mini transposon. Consequently, the plasmid that carries the mini transposon had to be devoid of *Hin*dIII and *Kpn*I restriction sites, and the mini transposon had to be modified in order to carry the restriction sites for the cloning, and the priming sites for the amplification of the signature tags.

At the first step, the *Hin*dIII restriction site situated 1105 bp downstream of the start codon of the *tnpA** gene from pCRS530 (Reeve *et al.*, 1999) was mutated according to the standard procedure (Carter, 1991). For this single-site mutagenesis, four primers were used: two primers, complementary to the ends of the gene with overhangs containing sites for cloning into pG18Mob2 (Xba_polym_begin and Sph_polym_end), and two primers for the nucleotide replacement in *Hin*dIII restriction site (Hind_mut_1 and Hind_mut_2). The primers for nucleotide replacement are complementary to each other and to the part of the sequence containing the *Hin*dIII restriction site, except for one nucleotide that is exchanged (the same nucleotide in both primers). In the first round of PCR, primers Xba_polym_begin and Hind_mut_2 were used to amplify one part of the gene, whereas primers Sph_polym_end and Hind_mut_1 were used to amplify another part of the gene. In the second round of PCR the two first-round PCR amplification products were used as megaprimers to create the whole gene product containing the replaced nucleotide (Fig. C1). The nucleotide was replaced in the way that the *Hin*dIII restriction site was abolished, but the amino-acid content of the transposase protein was not changed.

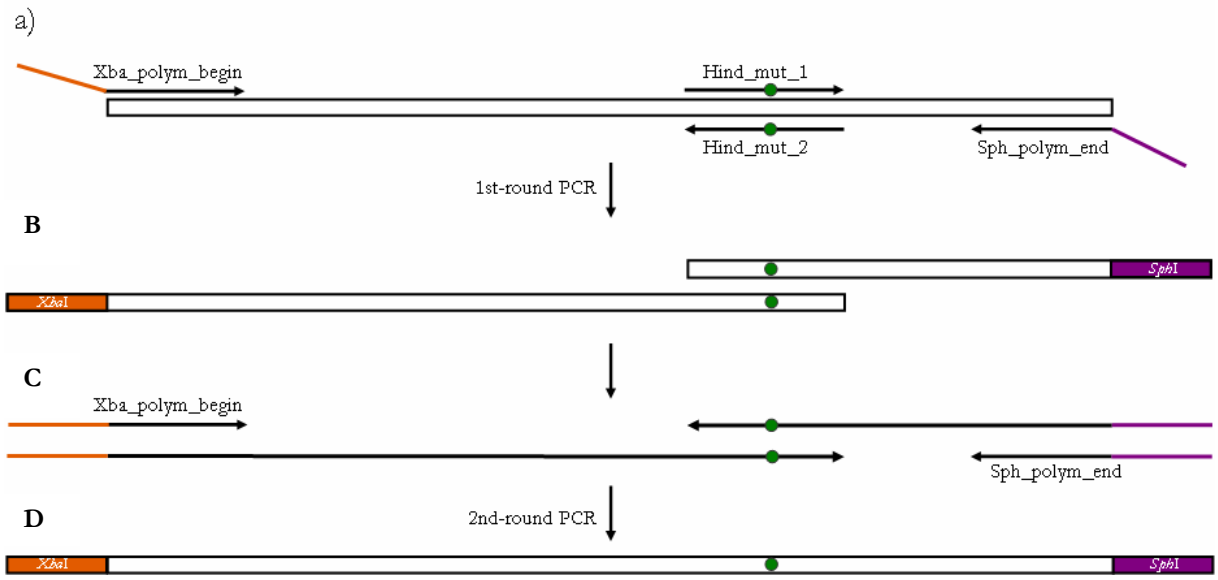


Fig. C1. Single-nucleotide replacement in the *tnpA** gene. (A) The primers used for amplification and nucleotide replacement (the mismatched nucleotide is marked as a green circle). (B) The amplification products from the first-round PCR containing the replaced nucleotide. (C) The products from the first PCR round are used as megaprimers in the second round of PCR together with the primer annealing at the ends of the *tnpA** gene. (D) The product of the second round of PCR containing the replaced nucleotide abolishing the *Hin*dIII restriction site (*tnpA***).

The *tnpA*** gene containing the mutated *Hin*dIII restriction site (GAGCTT) was inserted into the *Sph*I and *Xba*I restriction sites of pG18Mob2. The cloning of the correct insert with mutated *Hin*dIII site was demonstrated by restriction analysis and sequencing.

The mini transposon mTn*5*-GNm was recovered from pCRS530 using the flanking *Xba*I and *Eco*RI restriction sites and inserted into pG18mob2. Subsequently, the *Hin*dIII restriction site in the polylinker of pG18Mob2 was inactivated by treatment of the *Hin*dIII sticky ends with Klenow enzyme and blunt-end ligation.

An artificial linker (AATTCGGCCGCCTAGGCCAAAGGACGTGGTTTACGGGGCACGTAGTTTAAGGAA GTACGGTAA***GGTACC***GGGGGTGGCGGCATTCATATAGCTGCGTGATTTCATTTTA ACTCCCCTCCGCCGC***AAGCTT***AGGTGGACCGTCGTAGAGCTAGTAGGGGCTCAATG CACCAGGACTAGGCCGCCTAGGCCGAATTC) containing four priming sites (underlined) flanking *Kpn*I and *Hin*dIII restriction sites (printed in bold italic) for insertion of variable tag sequences was generated and inserted into the *Sfi*I restriction site of the mTn*5*-GNm transposon resulting in the transposon mTn*5*-STM (Fig. C2 B).

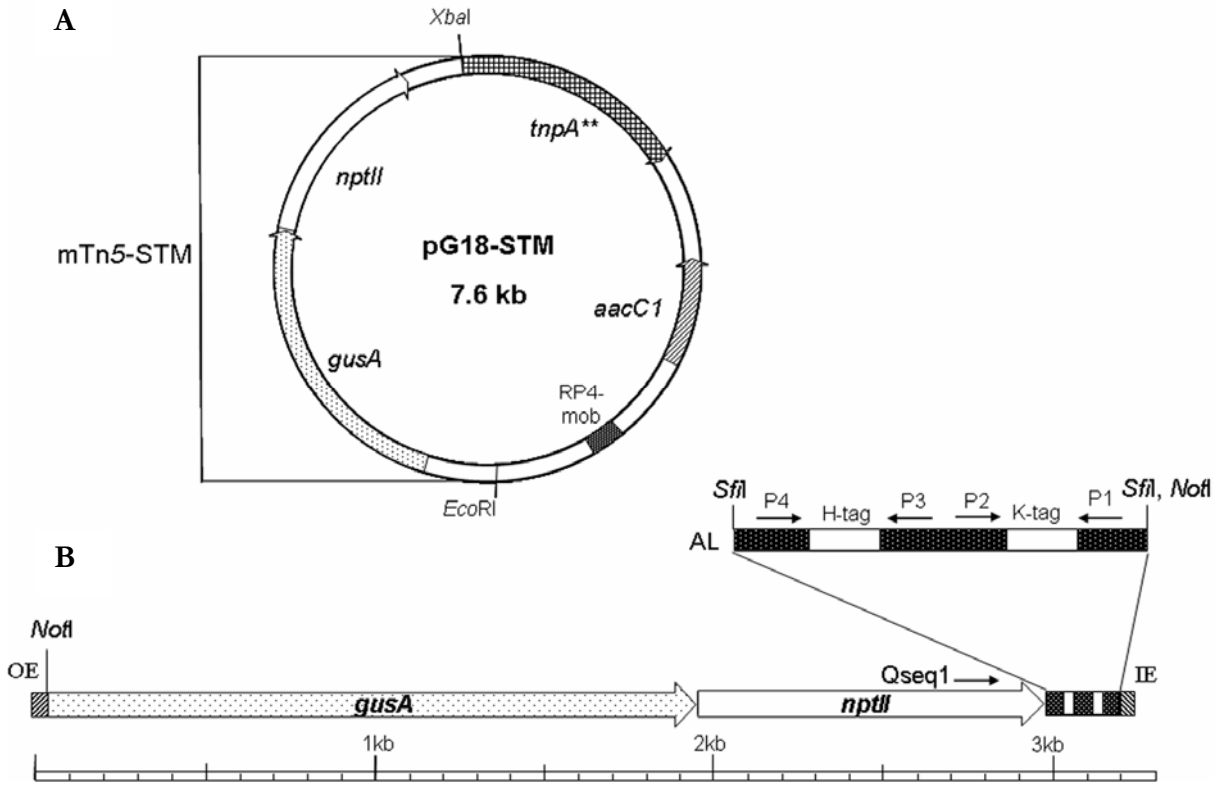


Fig. C2. Vector constructed for sequence-tagged mutagenesis of *S. meliloti*, based on pG18Mob2 containing a modified mTn*5*-GNm transposon and a transposase gene *tnpA*** devoid of the *Hin*dIII restriction site (A) and a detailed representation of the mTn*5*-STM transposon (B). The artificial linker denoted AL was inserted into the *Sfi*I restriction site of mTn*5*-GNm. *Hin*dIII and *Kpn*I restriction sites of the artificial linker were used to clone the signature tags (H-tag and K-tag). P1, P2, P3 and P4 denote the annealing sites for primers P1_Kpn, P2_Kpn, P3_Hind and P4_Hind respectively; Qseq1 denotes the annealing site for the sequencing primer. IE and OE indicate the inside and outside end required for transposition.

The computational design of the artificial linker was performed by Ulf Schiller and Stefan Kurtz (Praktische Informatik, Technische Fakultät, Universität Bielefeld). Candidate sequences for priming sites were designed using the program *DNASequenceGenerator* (Feldkamp *et al.*, 2001). For primer design, 87 candidate sequences were generated that fulfilled the requirements of having less than 7 identical contiguous nucleotides, of length (21bp) and melting temperature (65 to 70°C) and contained no *Kpn*I, *Hind*III, *Eco*RI or *Sfi*I restriction sites. To construct the linker sequence, 4 of these primer sequences were chosen. The primer and linker sequences were checked for similarity to the *S. meliloti* genome using the program *vmatch* (http://www.vmatch.de) (Abouelhoda *et al.*, 2004). Sequences with the lowest degree of similarity to the *S. meliloti* genome were chosen. The linker was synthesized as six separate oligonucleotides that were annealed and ligated using the sticky ends of the restriction sites *Kpn*I and *Hind*III. The plasmid sequence was submitted to GenBank under the accession Nr. DQ408591

C.2 CLONING OF TAGS INTO THE mTn*5*-STM TRANSPOSON

The H- and K-tags were inserted into the linker cassette of transposon mTn*5*-STM. Tags were synthesized as complementary single-stranded oligonucleotides and were annealed prior to insertion into the linker. First, the H-tags were inserted into the *Hind*III restriction site of pG18-STM. Subsequently, K-tags were cloned into the *Kpn*I restriction site of each individual plasmid that contained an H-tag from the first tag cloning step.

The main complication in the cloning of the tags was the inability to screen for the positive clones by blue-white selection or other instant selection methods. Moreover, the ends of tags were not phosphorylated in order to avoid their concatemerization; consequently, the ends of the digested plasmid could not be dephosphorylated and could easily ligate together without the insert. To make the cloning procedure maximally efficient, very low quantity of the digested plasmid was used in the ligation together with high quantity of the insert. Ligation mixtures were then transformed in highly competent *E. coli* cells (Inoue *et al.*, 1990). The tag-containing clones were verified by PCR using primer pairs for amplification of the tags (P1_Kpn, P2_Kpn in the case of K-tags and P3_Hind, P4_Hind in the case of H-tags) and subsequent separation of the amplification products in agarose gels. Since the amplification products from the clones containing tags were only 24 bp longer than the amplification products from the empty clones, the high-percentage agarose gel (2.3 – 2.5%) was used to separate the products.

To check if the tags were cloned correctly, the linker region of each plasmid was sequenced. As a result, a collection of 412 transposons was generated, where each transposon is individually marked by two unique signature tags.

C.3 TRANSPOSON MUTAGENESIS OF *S. MELILOTI*

The tagged transposon library was used for random mutagenesis of *S. meliloti* Rm2011. *E. coli* S17-1 mediated conjugation was used to transfer the plasmids carrying the tagged transposons into *S. meliloti* cells (Fig. C3). The RP4 mobilizable region of pG18-STM enabled the conjugal transfer of plasmids from the *E. coli* donor strain S17-1 into the *S. meliloti* recipient cells through a procedure of biparental mating. Mutants were selected based on resistance to neomycin due to the *nptII* gene of the transposon. From each conjugation, 24 to 30 clones were picked, resulting in a library of 12000 tagged mutants. The mutant clones were re-arrayed into sets, each containing mutants that differ by their signature tags. Since the mTn*5*-STM transposons used for mutagenesis contain a terminator sequence, the generated mutations can influence the transcription of the whole gene operone (polar mutations).

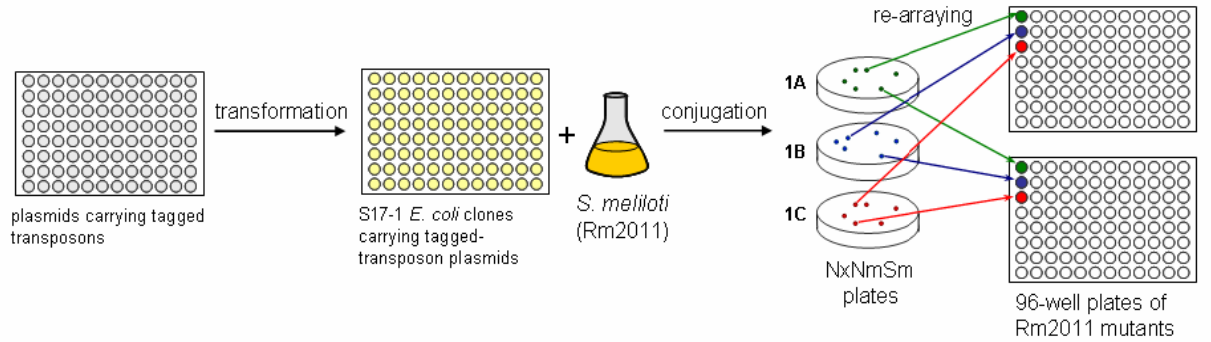


Fig. C3. Construction of the transposon mutant library

C.4 MAPPING OF THE TRANSPOSON INSERTION SITES

The transposon insertion sites of 5089 mutants were determined by sequencing the transposon-genome junction using a primer Qseq1 that binds 64 bp upstream of the linker with the cloned tags (Fig. C2 B). It was therefore possible not only to determine the insertion sites of the transposons, but also to check the tags in the mutants. Sequencing was performed by Qiagen (Hilden, Germany) using genomic DNA as template. No amplification or cloning steps were carried out. The computational analysis of insertion sites was performed by Danijel Wetter (International Graduate School in Bioinformatics and Genome Research, Bielefeld University).

For transposon insertion site mapping, the GenDB annotation system was used (Meyer *et al.*, 2003). A GenDB extension was implemented for automated transposon position identification, where the sequences of transposon-genome junction regions were compared to the *S. meliloti* genome using the BLAST algorithm (Altschul *et al.*, 1997). The transitions from transposon sequence to genome sequence were identified as "jump-in" positions and mapped onto the *S. meliloti* genome. Fig. C4 summarizes the results from transposon mapping. The complete list of

mapped mutations is available on the website of the public *S. meliloti* GenDB genome project (http://www.cebitec.uni-bielefeld.de/groups/nwt/sinogate).

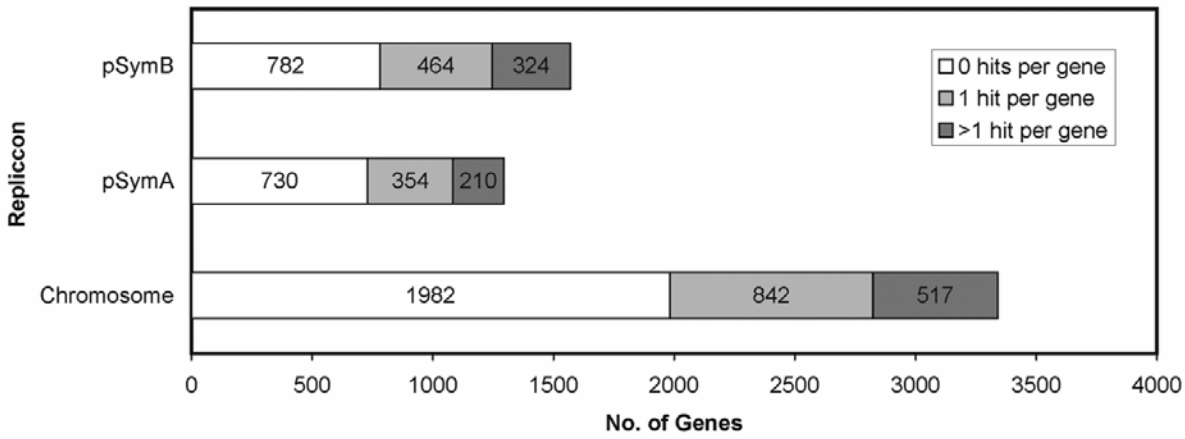


Fig. C4. Distribution of the mTn*5*-STM transposon insertions throughout the *S. meliloti* genome: genes with transposon insertions.

C.5 STATISTICAL ANALYSIS OF THE MUTANT LIBRARY

Statistical verification of the insertion sites was performed in cooperation with Danijel Wetter. Several statistical tests were performed to ensure that the distribution of transposon insertions was random, and that the mTn*5*-STM transposons have no hot spots in the *S. meliloti* genome. An important parameter that shows the randomness of transposon insertions is the quantity of genes that carry a transposon insertion. A low quantity of genes hit by the transposon indicates a bias in the pattern of transposition. In order to determine the theoretical quantity of genes that has to be hit by at least one transposon, the neutral-base pair model (Jacobs *et al.*, 2003) was applied. This model allows the estimation of the number of gene hits based on genome length, number of transposon insertions and gene sizes. Applying this model to the library of 5089 *S. meliloti* transposon mutants, it was predicted that 2890 genes with a standard deviation of ± 378 had to be mutated. The actual number of genes that were hit by at least one transposon was 2711 (43.68 % of the predicted in *S. meliloti* 6207 protein-coding genes), complying with the expectations based on the neutral-base pair model.

Furthermore, a genome wide analysis of all transposon insertion sites in relation to the GC (AT) content was performed. Using a 100 bp window centered at the transposon insertion position, the mean GC (AT) content values were calculated. The difference between the mean GC (AT) content within all of these windows and the mean GC (AT) content of the whole genome is 0.2 % for GC, and 0.1 % for AT. A χ^2 test was also made, to exclude balancing effects

in deviations to the GC (AT) mean. The result was a very low χ^2 score of 12.8 for GC and 14.2 for AT distribution with 2710 degrees of freedom, implying that there was no preference of the mTn*5*-STM transposons to jump into GC or AT rich regions.

In order to test for uniform distribution of all transposon insertions we performed a χ^2 test. Given a p-value of 0.01 and 29 degrees of freedom per replicon, we found a uniform distribution to be highly improbable. This could be explained by the existence of essential genes that are not represented in the mutant library. Moreover, mutations causing slow growth of bacteria under the conditions used for selection of the transconjugants in this study result in underrepresentation of such mutants in the library. The χ^2 test was therefore repeated, assuming that the *S. meliloti* genome contains essential genes. There is not enough information available about quantity and position of essential genes in the *S. meliloti* genome that would allow excluding defined groups of genes from the χ^2 test. To cope with this problem, sets of randomly chosen genes were created, and the χ^2 test was performed many times, leaving out one of the gene sets each time. The number of genes per set ranged from 1 to 100 % of all genes that were localized on a certain replicon and did not have transposon insertions.

Such a modified χ^2 test showed that the distribution of transposon insertions in the genome is likely to be random. The best χ^2 test result for pSymB was a likelihood of 90 % for a random distribution of transposons over this replicon. This result was obtained when a set containing 6 % of genes with no transposon insertion was excluded from the test. For pSymA, an 87 % likelihood of randomness was observed when 10 % of genes with no hit were left out. In contrast, when all genes that do not have a transposon insertion were excluded from the χ^2 test, the probability for the distribution to be random was lower than 11 % for both megaplasmids.

Though the modified χ^2 test worked well for the megaplasmids, it failed in the analysis of the transposon insertion distribution throughout the *S. meliloti* chromosome. A set of genes whose exclusion from the χ^2 test increased the likelihood of randomness to more than 15 % could not be found. The reason for this might be a high proportion of essential genes on the chromosome (Capela *et al.*, 2001) and the great size of the replicon, that leads to the large fraction of genes not hit by a transposon. Therefore, a random search to detect essential genes seems to be unsuitable if the number of transposon insertions is not saturating. Nevertheless, based on the data for pSymA and pSymB and the results from the analysis of transposon insertion sites in relation to the GC (AT) content, a genome wide random distribution of transposon insertion sites can be assumed.

C.6 CREATION OF TWO SETS OF MUTANTS FOR THE SUBSEQUENT EXPERIMENTS.

After the transposon insertion sites were determined for the mutants, this information could be used to create non-redundant sets, containing mutants of interest. These sets could be than used in experiments in free-living and symbiotic conditions. Each set has to contain mutants with different tags; therefore, one set can maximally include 412 mutants.

Two sets of mutants were created: set 1 and set 2, each one containing 378 mutants. The sets were for the biggest part non-redundant, so that just one mutant for each gene was present in the sets. For several genes, two or three mutants were included in the sets in order to test if the results for these mutants will be similar. Particularly, three *fixL*, two *fixN1* and two *ilvC* mutants were included in the sets. Since both of the sets were planned to be screened in symbiotic conditions, I tried to combine the mutants in the way that many potentially attenuated mutants would have been included in the sets. For example, many mutants with a transposon insertion in genes with known symbiotic phenotype were included, as well as in genes that potentially could have a symbiotic phenotype. Also, many mutants with transposon insertion in hypothetical genes were included in the sets.

C.7 PILOT COMPETITION EXPERIMENTS IN FREE-LIVING CONDITIONS

In order to validate the STM approach used here, two pilot experiments were carried out using the mutant set 1. In all cases, the input pool (inoculate) was used as the reference. Each experiment was repeated three times using independent cultures.

As test conditions for the first experiment, growth in rich (TY) medium and minimal medium (VMM) was chosen. The goal of this experiment was not only to compare the differences in the survival of the mutants between rich and minimal medium, but also to study the kinetics of changes in the content of mutants during the growth of culture. Rich and minimal medium were inoculated by the glycerol culture of the mixture of mutants to $o.D._{600}$=0.003.

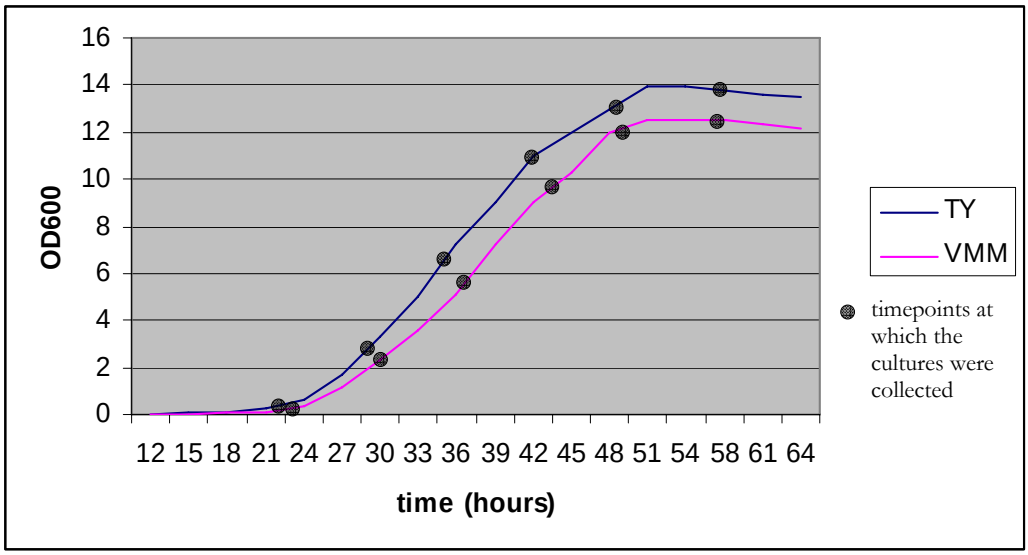


Fig. C5. Growth of S. *meliloti* mutant set in rich (TY) and minimal (VMM) medium

Bacteria were then cultured for 3 days; points at which the cultures were collected are indicated on the Fig. C5. In the second experiment, the same set of mutants was grown in high osmolarity medium and in medium containing SDS as detergent. The cells were collected in the exponential growth phase.

The genome DNA was isolated from the collected cells and the tags were amplified using fluorescently labeled primers (5´-modified by Cy3 in case of the experiment and by Cy5 in case of the input pool) (Fig. C6).

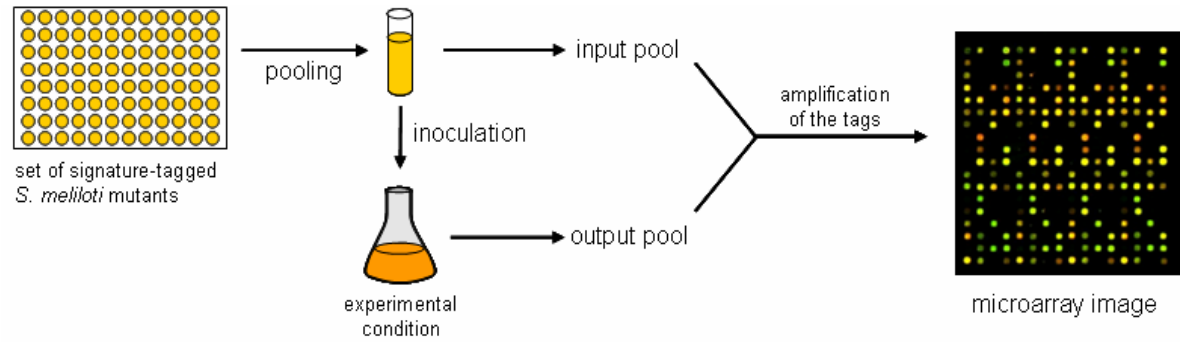


Fig. C6. Schematic representation of a competition experiment using signature-tagged mutants

C.8 TWO-CHANNEL TAG MICROARRAY FOR THE DETECTION OF THE TAGS

In this project, a two-channel microarray was for the first time used in the signature tagged mutagenesis screen. The mTn*5*-STM-1 microarray was created, that contains 24mer oligonucleotides with 5´ C_{12}-amino modifications, directed against the variable sequences of

signature tags in 412 created mTn*5*-STM transposons. Since the signature tags are quite short oligonucleotides, the hybridization and subsequent washing protocols for the similar Sm6kOligo 70-mer microarray (Becker and Rüberg, 2003) were modified in order to optimize the output. The hybridization time was set to only 45 minutes, in comparison to the usual 14 – 18 hours, and the hybridization temperature of 36 °C was used instead of 42 °C. The washing conditions for the tag-microarray were also milder than conditions for the 70-mer oligonucleotide microarray (see chapter E.2.12 for further details).

C.9 ANALYSIS OF TAG-MICROARRAY DATA

A comparison of the STM approach to the standard microarray experiments shows several important differences. First, instead of measuring expression levels of genes, the absence or presence of mutants in a complex pool is analyzed. Second, the PCR fluorescently labeled amplified products are used for hybridization instead of cDNA. Third, the utilization of two tags to barcode each transposon requires modified normalization procedure. Therefore, the specific properties of this kind of data sets have to be taken into account if data from high throughput STM experiments have to be analyzed employing bioinformatics. A framework for appropriate analysis of tag-microarray data was developed based on a carefully designed combination of existing methods for analysis of microarray data as well as selection of suitable parameters. This analysis was performed in cooperation with Silke Szymczak (Angewandte Neuroinformatik, Technische Fakultät, Universität Bielefeld).

C.9.1 Normalization and filtering

The pre-processing of the tag-microarray data was performed similarly to the usual microarray analysis approach. At the next step, LOWESS normalization based on local regression that accounts for intensity and spatial dependence in dye biases was performed for data of the two tags separately. The separate normalization for the two kinds of tags was the most suitable, because the targets for H- and K-tags originate from the separate PCR reactions (Fig. C7).

Then, the c_i distances between H- and K-tag for each mutant were calculated and 10 % of the mutants with the largest c_i values and therefore with the largest discrepancy between H- and K-tag, were removed. The weighted mean of the medians of the m-values was calculated for each mutant combining values of H-tags and K-tags, and used for subsequent analysis.

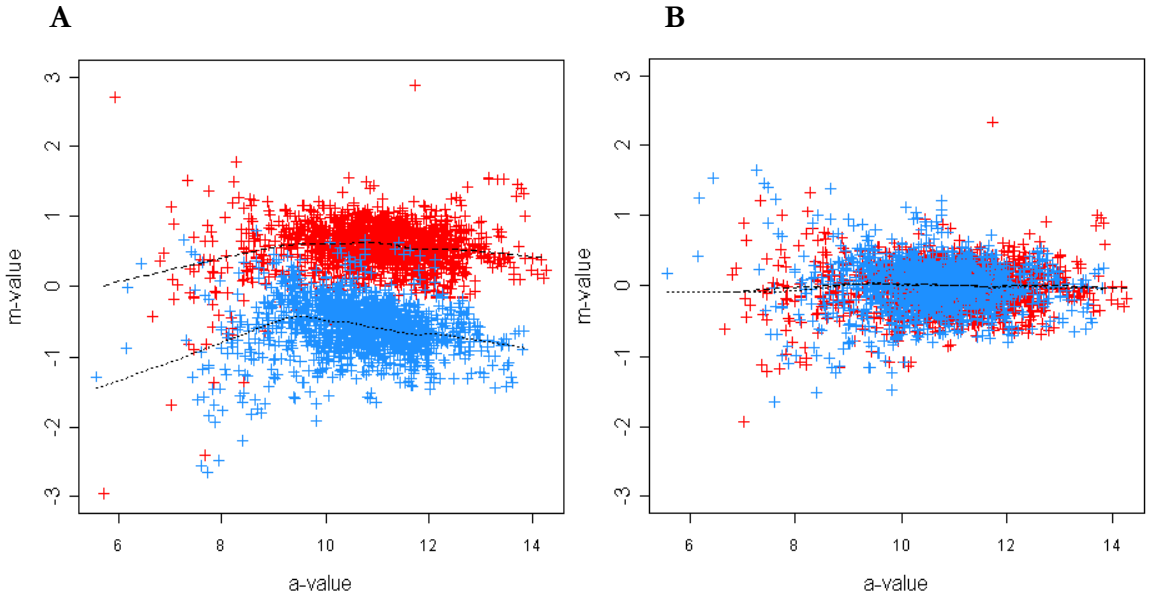


Fig. C7. Reduction of differences in data between the two types of tags by separate normalization. MA plots of data from a single slide before (A) and after (B) normalization (red: H-tag, blue: K-tag, dashed line: LOWESS smoother for H-tag, dotted line: LOWESS smoother for K-tag).

C.9.2 Clustering

The clustering approach was developed using the data from the time course TY-VMM pilot experiment. In order to make the analysis as efficient as possible, different ways of calculating distances and linkage methods for agglomerative clustering were tested. In addition, the results of agglomerative clustering were compared to the results produced by the *k*-means clustering approach. The results were validated based on the similarity of grouped patterns. For a review on clustering methods, see (Heyer *et al*, 1999) and (Zhao and Karypis, 2005).

The first comparison examined the effect of the following distance calculation methods on the result of clustering: Euclidean, centered correlation and uncentered correlation (E.2.13.3), while average linkage method was applied every time. The second comparison was performed to explore the effect of the different linkage methods on clustering (single, complete, average and centroid) while the uncentered correlation distance calculation was used in each of the cases. The results of the different distance calculations are shown in Fig. C8, whereas Fig. C9 presents the dendrograms created using different linkage methods. At all the clustering images, green colour represents attenuated, red color represents fast growing mutants in a respective condition.

Choice of a distance calculation method seems to have a greater influence on the output than the choice of a linkage method. Using centered correlation as distance calculation method, no separation between mutants with increased and attenuated growth can be observed (Fig. C8 A). This is due to the fact that this distance calculation method is based on shape, i. e. global behaviour of each mutant is taken into account. Euclidean distance calculation method makes a

separation of attenuated and increased mutants possible (Fig. C8 B). However, mutants with attenuated growth only in minimal medium are not separated from mutants showing attenuated growth behaviour in both minimal and rich medium. The uncentered correlation coefficient calculation (Fig. C8 C) which combines the advantages of the centered Pearson coefficient (shape relevant) with the advantages of the Euclidean distance (separation based on negative and positive values) allowed the generation of clusters containing mutants with similar behaviour.

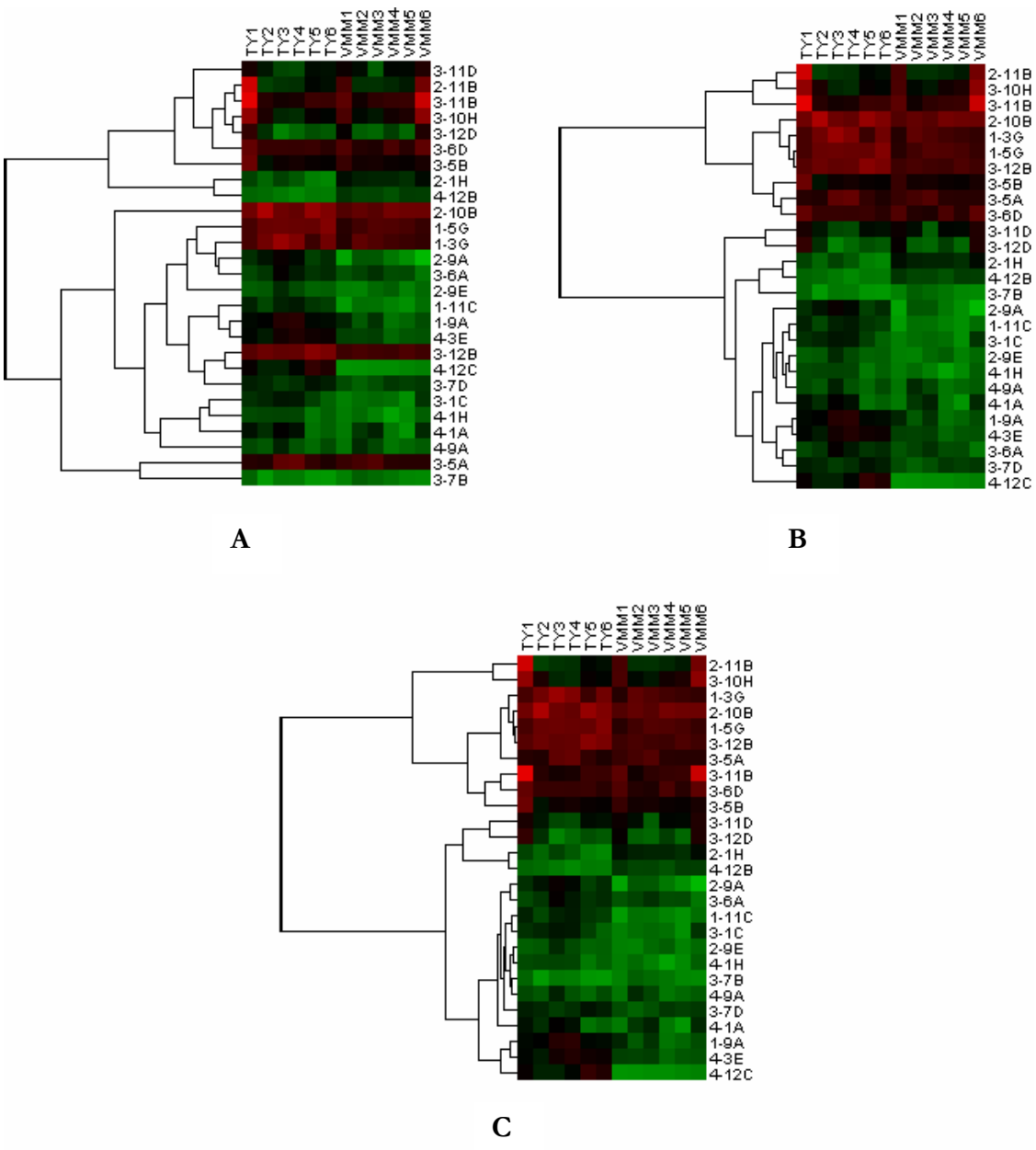


Fig. C8. Effects of different distance calculation methods on clustering. Dendrograms generated with agglomerative clustering (average linkage) on the same data set using centered correlation (A), Euclidean (B) and uncentered correlation (C) distance.

When comparing the effects of various linkage methods on the output of clustering, only small differences in the inner structure of several clusters can be observed. The application of

single linkage, however, generates a dendrogram where the chaining effect of this linkage method is visible. In many joining steps single mutants are merged with other clusters (e. g. mutants 3-5A, 3-6D, 3-11B, 3-10H and 3-5B in Fig. C9 A). Complete, average and centroid linkage show very similar results (see Fig. C9 B, C and D). Complete linkage, however, is sensitive to outliers since only the distance between the two data points that are furthest apart are used for calculation. Centroid linkage sometimes suffers from inversion problems (Morgan and Ray, 1995). Therefore, if the agglomerative clustering is used, the best suitable linkage for this STM system is the average linkage method.

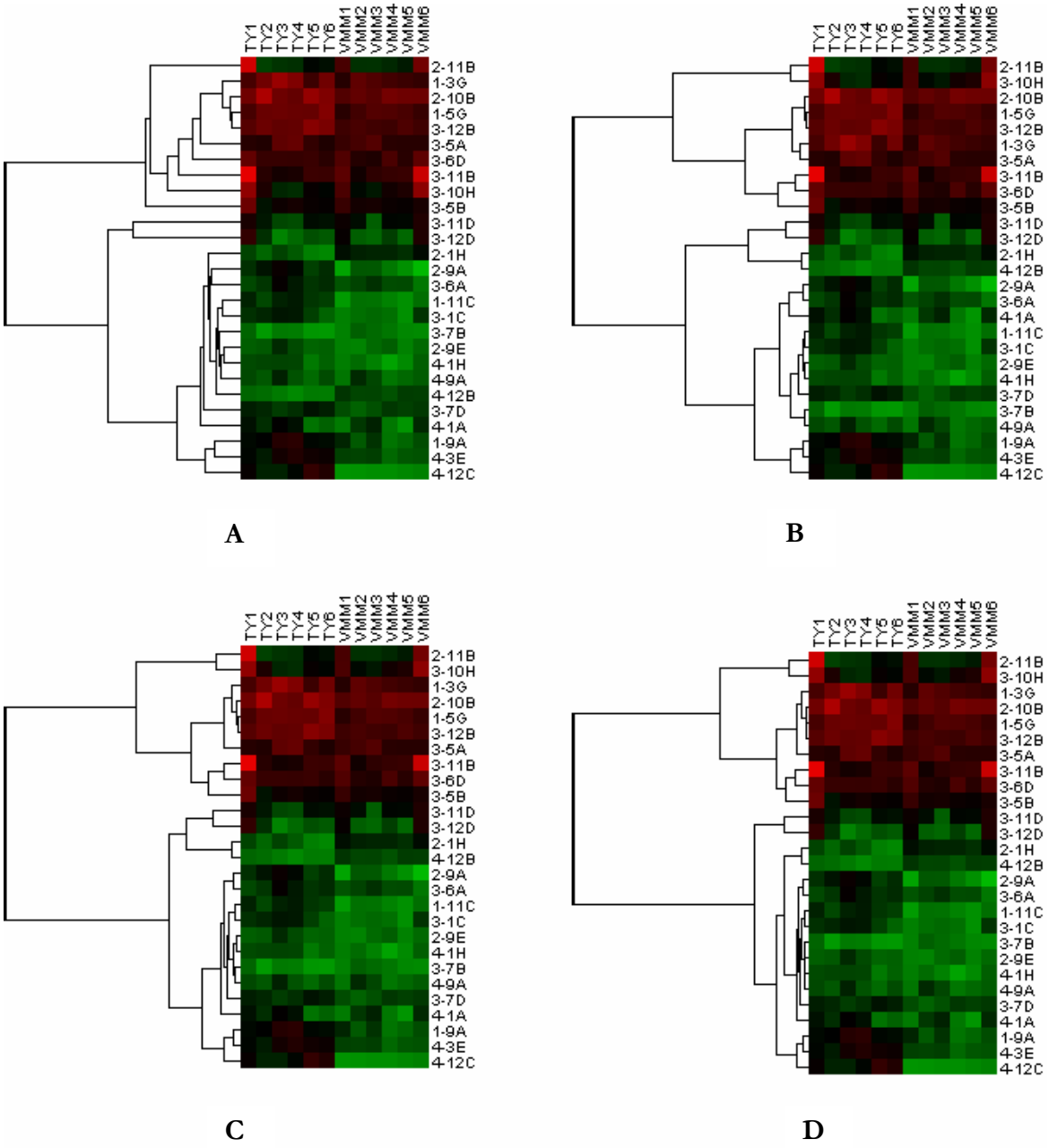


Fig. C9. Effects of different linkage methods on clustering. Dendrograms generated with agglomerative clustering (uncentered correlation distance) on the same data set using single (A), complete (B), average (C) and centroid (D) linkage.

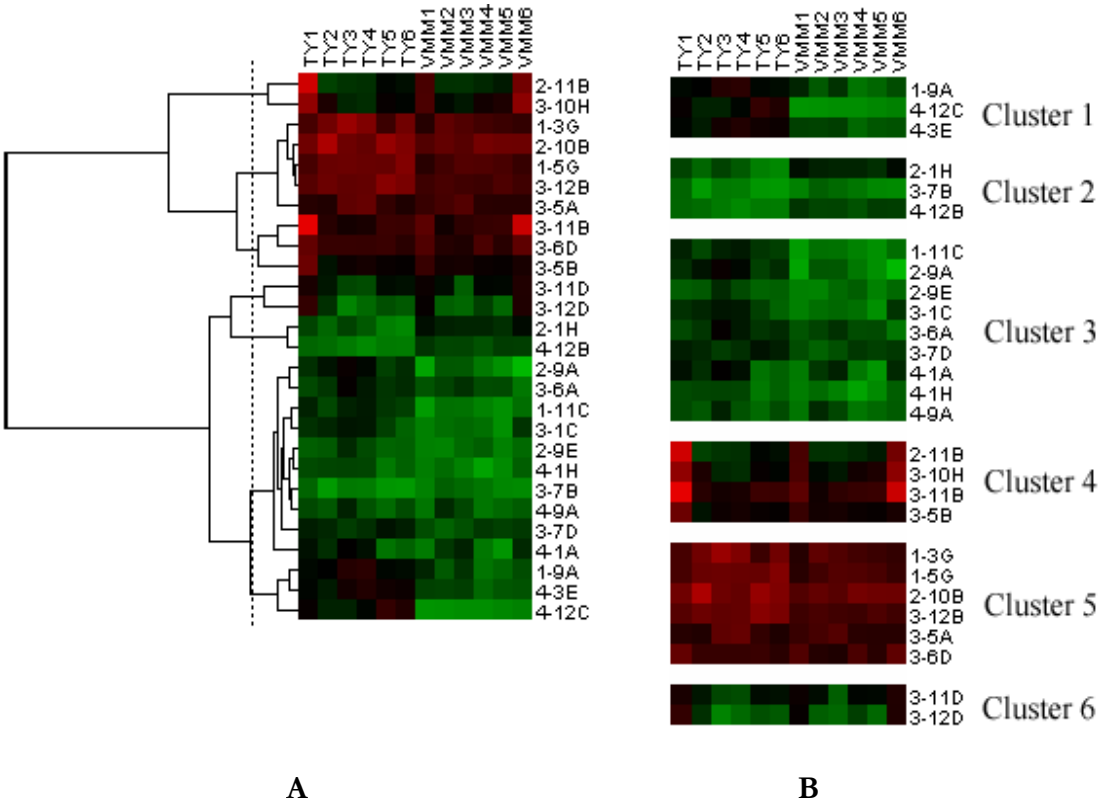
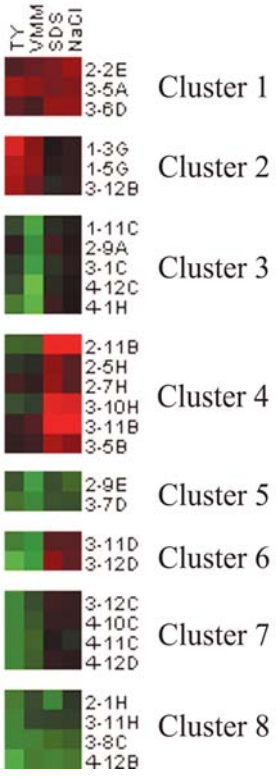


Fig. C10. Comparison of the best results from agglomerative (A) and k-means clustering (B) in the TY-VMM timecourse experiment. Cutting of the dendrogram in (A) at the depth indicated by the dotted line leads to a comparable partition of clusters



After selecting the appropriate distance calculation method, agglomerative and *k*-means (MacQueen, 1967) clustering approaches were compared (Fig. C10). To obtain clusters from the agglomerative clustering approach, the dendrogram was cut at the depth indicated by the dotted line in Fig. C10 A, which results in seven clusters. The number of clusters *k* in *k*-means clustering was set to 6. The resulting separation into clusters is very similar using the two methods, with the difference that in using the *k*-means clustering approach it was possible to unite the mutants, which grew faster early after inoculation and in the stationary phase, in one cluster (cluster 4).

The adjucsted k-means clustering approach was applied to analyze the data from the second pilot experiment, using the number of clusters k=8.

Fig. C11. K-means cluster analysis of tag microarray data from the stress conditions experiment.

The data from the stress conditions were clustered together with the time point 3 (exponential phase) data for VMM and TY cultures (Fig. C11). The clustering pattern of the stress conditions experiment is similar in part to the clustering pattern of the time course experiment. Particularly, the mutants 3-11D and 3-12D are grouped in one cluster in both experiments. Furthermore, all the mutants that belong to cluster 4 in the time course experiment and are characterized by a faster growth in the stationary phase compared to other mutants, grew faster than other mutants also in SDS- and salt induced stress conditions (cluster 4 in the stress conditions experiment).

The characterized mutants in the respective clusters from the time course data and from the stress conditions/VMM/TY comparison are listed in the Tables C1 and C2, respectively. In the Appendix Table G2, the complete list of m-values and p-values is given for the stress conditions experiment and for the time point 3 in VMM and TY cultures.

TABLE C1. Characteristics of mutants that have a changed phenotype in the time course experiment

mutant ID	mutated gene[a]	gene product	mean$_w$-values											
			TY1	**TY2**	**TY3**	**TY4**	**TY5**	**TY6**	**VMM1**	**VMM2**	**VMM3**	**VMM4**	**VMM5**	**VMM6**
Cluster 1	**impaired in growth in VMM**													
1-9A	SMb20616	thiamine biosynthesis oxidoreductase protein ThiO	-0.064	-0.014	0.29	0.336	-0.05	-0.09	-0.283	-0.702	-0.399	-0.918	-0.807	-0.603
4-12C	SMc01053	CysG siroheme synthase protein	0.0595	-0.264	-0.265	-0.04	0.406	0.26	-1.168	-1.154	-1.119	-1.121	-1.07	-1.018
4-3E	SMc02117	MetC cystationine beta-lyase protein	-0.04	-0.207	0.157	0.267	0.138	0.1	-0.51	-0.534	-0.49	-0.826	-0.683	-0.649
Cluster 2	**impaired in growth in TY medium and partially in VMM**													
2-1H	SMc01219	lipopolysaccharide core biosynthesis mannosyltransferase LpsB E3 component of 2-oxoglutarate	-0.534	-0.805	-0.526	-0.7	-1.01	-1.07	-0.113	-0.258	-0.289	-0.285	-0.375	-0.084
3-7B	SMc02487	dehyrogenase complex transmembrane protein LpdA2	-0.793	-1.219	-0.957	-0.98	-1.19	-1.21	-1.023	-0.754	-0.841	-0.933	-1.067	-1.092
4-12B	SMc02144	PstC phosphate transport system permease ABC transporter protein	-0.797	-0.854	-0.964	-1.08	-0.96	-0.98	-0.409	-0.543	-0.542	-0.647	-0.473	-0.491
Cluster 3	**impaired in growth under all conditions, strongly impaired in VMM**													
1-11C	SMc04346	IlvC ketol-acid reductoisomerase protein	-0.267	-0.563	-0.266	-0.2	-0.42	-0.5	-1.221	-0.886	-0.899	-1.027	-1.138	-0.87
2-9A	SMc02899	PheA prephenate dehydratase protein	-0.372	-0.155	0.086	-0.06	-0.51	-0.35	-1.296	-0.692	-0.693	-0.946	-1.134	-1.458
2-9E	SMc03782	hypothetical signal peptide protein	-0.747	-0.695	-0.329	-0.59	-0.79	-0.76	-1.015	-1.034	-0.836	-0.74	-1.023	-0.827
3-1C	SMc01842	putative methyltransferase transcription regulator protein	-0.38	-0.294	-0.167	-0.19	-0.5	-0.78	-1.011	-0.88	-0.825	-0.882	-1.136	-0.404
3-6A	SMc03777	gamma-glutamil phosphate reductase protein ProA	-0.567	-0.422	0.043	-0.2	-0.42	-0.33	-0.663	-0.53	-0.38	-0.6	-0.59	-0.939
3-7D	SMc01174	CysK2 cysteine synthase A protein	-0.229	-0.304	-0.476	-0.36	-0.14	-0.24	-0.605	-0.777	-0.636	-0.439	-0.509	-0.447
4-1A	SMc00917	ATP phosphoribosyltransferase protein HisG	-0.128	-0.356	0.018	-0.12	-0.91	-0.77	-0.961	-0.441	-0.266	-0.94	-1.18	-0.319
4-1H	SMc03776	ProB1 glutamate 5-kinase protein	-0.585	-0.546	-0.535	-0.55	-0.96	-0.77	-0.986	-0.843	-1.021	-1.284	-1.089	-0.747
4-9A	SMc00825	glutamate-cysteine ligase GshA	-0.571	-0.686	-0.263	-0.65	-0.84	-0.6	-0.924	-0.329	-0.557	-0.931	-0.825	-0.717
Cluster 4	**highly competitive at early growth stages and in the stationary phase**													
2-11B	SMc01881	PanB 3-methyl-2-oxobutanoate hydromethyltransferase protein	1.6331	-0.516	-0.409	-0.38	-0.05	-0.11	0.569	-0.378	-0.39	-0.308	-0.202	0.899
3-10H	SMb20360	hypothetical protein. putative protease subunit of ATP-dependent Clp protease	1.1341	0.227	-0.294	-0.35	0.059	-0.03	0.583	-0.07	-0.187	0.147	0.233	1.106

3-11B	pSymB, pos. 1003846	intergenic	1.8135	0.315	0.158	0.203	0.467	0.46	0.718	0.134	0.358	0.425	0.445	1.583
3-5B	SMb20931	sugar uptake ABC transporter periplasmic solute-binding protein precursor	0.8503	-0.145	0.106	0.166	0.072	0.04	0.436	0.157	0.175	0.09	0.058	0.192
Cluster 5	**highly competitive under all conditions**													
1-3G	pSymB, pos. 543971	intergenic	0.5552	0.883	1.211	1.002	0.472	0.89	0.295	0.749	0.636	0.523	0.459	0.4
1-5G	SMa0621	FixI2 E1-E2 type cation ATPase	0.5542	0.824	0.848	0.812	0.797	1.03	0.32	0.518	0.66	0.638	0.567	0.427
2-10B	SMa0564	hypothetical protein	0.88	1.361	0.863	0.786	1.174	1	0.643	0.774	0.679	0.918	0.859	0.853
3-12B	SMb20476	putative ABC transporter periplasmic dipeptide-binding protein	0.6717	0.846	0.762	0.782	1.065	0.95	0.481	0.57	0.507	0.474	0.611	0.499
3-5A	SMb20037	AroE2 shikimate 5-dehydrogenase protein	0.3029	0.254	0.748	0.801	0.397	0.24	0.435	0.549	0.657	0.33	0.315	0.312
3-6D	SMc03032	FlgI flagellar P-ring precursor transmembrane protein	0.777	0.428	0.428	0.418	0.451	0.33	0.635	0.291	0.238	0.589	0.321	0.745
Cluster 6	**clones with low competitiveness under normal conditions and normal competitiveness at early growth stages and in the stationary phase**													
3-11D	SMb20377	putative translation initiation inhibitor protein	0.2047	-0.175	-0.552	-0.6	-0.09	-0.13	0.098	-0.101	-0.744	-0.043	-0.06	0.26
3-12D	SMc00334	Cmk cytidilate kinase protein	0.3993	-0.375	-1.036	-0.83	-0.63	-0.7	0.051	-0.745	-0.817	-0.495	-0.811	0.245

[a]For the mutants that have a transposon insertion in intergenic regions, an exact position of insertion in the respective replicon is given.

TABLE C2. Characteristics of mutants showing an altered phenotype in TY, VMM and/or stress conditions

mutant ID	mutated gene[a]	mean$_w$-values				gene product
		TY	VMM	TY-SDS	TY-NaCl	
Cluster 1	**highly competitive under most of the tested conditions**					
2-2E	SMb21633	0.39	0.55	0.67	0.88	PaaG enoyl-CoA hydratase protein
3-5A	SMb20037	0.90	0.79	0.65	0.71	AroE2 shikimate 5-dehydrogenase protein
3-6D	SMc03032	0.51	0.29	0.83	0.81	FlgI flagellar P-ring precursor transmembrane protein
Cluster 2	**highly competitive in TY and VMM under non-stress conditions**					
1-3G	pSymB, 543971	1.46	0.77	-0.01	0.12	intergenic
1-5G	SMa0621	1.02	0.79	0.07	0.12	FixI2 E1-E2 type cation ATPase
3-12B	SMb20476	0.92	0.61	-0.09	0.14	putative ABC transporter periplasmic dipeptide-binding protein
Cluster 3	**impaired in growth in VMM**					
1-11C	SMc04346	-0.33	-1.08	-0.11	-0.04	IlvC ketol-acid reductoisomerase protein
2-9A	SMc02899	0.10	-0.83	0.26	0.04	PheA prephenate dehydratase protein
3-1C	SMc01842	-0.20	-0.99	-0.15	0.08	putative methyltransferase transcription regulator protein
4-12C	SMc01053	-0.32	-1.35	-0.08	-0.01	CysG siroheme synthase protein
4-1H	SMc03776	-0.65	-1.23	0.21	0.01	ProB1 glutamate 5-kinase protein
Cluster 4	**highly competitive in stress conditions**					
2-11B	SMc01881	-0.50	-0.47	2.09	2.03	PanB 3-methyl-2-oxobutanoate hydromethyltransferase protein
2-5H	SMc03164	-0.18	-0.08	0.77	0.43	XylB xylulose kinase protein. putative
2-7H	pSymB, 96435	0.28	0.11	0.82	0.43	intergenic
3-10H	SMb20360	-0.36	-0.22	1.67	1.50	hypothetical protein. putative protease subunit of ATP-dependent Clp protease
3-11B	pSymB, 1003846	0.19	0.43	1.95	1.85	intergenic
3-5B	SMb20931	0.12	0.21	0.94	0.76	putative sugar uptake ABC transporter periplasmic solute-binding protein precursor
Cluster 5	**impaired in growth under all conditions, strongly impaired in VMM**					
2-9E	SMc03782	-0.40	-1.01	-0.37	-0.52	hypothetical signal peptide protein
3-7D	SMc01174	-0.58	-0.77	-0.40	-0.37	CysK2 cysteine synthase A protein
Cluster 6	**clones with low competitiveness under normal conditions and slightly increased competitiveness under stress conditions**					
3-11D	SMb20377	-0.67	-0.90	0.46	0.46	putative translation initiation inhibitor protein
3-12D	SMc00334	-1.26	-0.98	0.79	0.51	Cmk cytidilate kinase protein
Cluster 7	**clones with low competitiveness under normal conditions but not disadvantaged in stress conditions**					
3-12C	SMc00808	-0.76	-0.35	0.21	0.15	ChrA chromate transport protein
4-10C	SMa0091	-0.75	-0.42	0.13	0.13	conserved hypothetical
4-11C	SMc02597	-0.72	-0.50	-0.03	-0.13	SodC superoxide dismutase Cu-Zn precursor transmembrane protein
4-12D	SMa0070	-0.75	-0.46	0.13	0.08	ABC transporter permease
Cluster 8	**impaired in growth in TY medium and partially under other conditions**					
2-1H	SMc01219	-0.64	-0.35	-0.74	-0.19	lipopolysaccharide core biosynthesis mannosyltransferase LpsB
3-11H	SMc01700	-0.72	-0.31	-0.26	-0.24	peptidyl-prolyl cis-trans isomerase A (PpiA)
3-8C	SMc02050	-0.74	-0.70	-0.54	-0.49	trigger factor protein Tig. probable
4-12B	SMc02144	-1.17	-0.65	-0.70	-0.62	PstC phosphate transport system permease ABC transporter protein

[a] For mutants carrying a transposon insertion in an intergenic region the exact position of insertion in the respective replicon is given.

C.9.3 Confirmation of biological significance of the applied statistical analysis

The applied k-means clustering provides not only a clear visualization of the STM data, but also a basis for the detailed analysis of mutants. To confirm that the analysis of STM data is biologically significant, it is important to support the observed patterns by other studies. Moreover, mutants which have similar defects should belong to the same cluster. Here, the clustering pattern from the stress conditions/VMM/TY clustering (Table C2) is analyzed in detail.

Cluster 1 contains clones that are highly competitive under most of the tested conditions. It includes, particularly, the *flgI* mutant, which is impaired in the synthesis of flagella. The fast-growing phenotype of this mutant supports the observation that the synthesis of flagella is energetically disadvantageous (Macnab, 1996). Two other mutants in this cluster are *paaG* (*SMb21633*) and *aroE2* (*SMb20037*) mutants. *paaG* encodes a putative enoyl-CoA hydratase/isomerase involved in phenylacetate catabolism, and *aroE2* codes for a putative shikimate 5-dehydrogenase protein involved in chorismate metabolism. Both genes have paralogs in the *S. meliloti* genome.

Cluster 2 represents clones that are highly competitive in TY medium and VMM under non-stress conditions. This cluster consists of three mutants bearing a transposon insertion in *fixI2* (encoding an E1-E2 type cation ATPase), in *SMb20476* (coding for a putative ABC transporter periplasmic dipeptide-binding protein) and in the intergenic region between *SMb20518* (encoding a putative endohitinase) and *SMb20519* (encoding a conserved hypothetical protein), probably influencing transcription of *SMb20519*.

Mutants that belong to cluster 3 are strongly impaired in growth in VMM. Characteristically, all mutants in this cluster have a transposon insertion in genes involved in the synthesis of amino acids or cofactors, not present in VMM: isoleucine/valine (*ilvC*), phenylalanine (*pheA*), ubiquinone/menaquinone (*SMc01842*), cysteine (*cysG*), and proline (*proB1*). It was previously shown that *ilvC* mutants of *S. meliloti* are isoleucine/valine auxotrophs (Aguilar and Grasso, 1991) and that *CysG* mutants of *R. etli* are cysteine auxotrophs (Tate *et al.*, 1997).

Cluster 4 contains six mutants that show high competitiveness in stress conditions but not in non-stress conditions. Two of these mutants carry a transposon insertion in the intergenic regions of pSymB: preceding *SMb20088* (encoding a conserved hypothetical protein) and upstream of *SMb21337* (coding for a putative iron-sulfur-binding protein, probably a subunit of an oxidoreductase-like aldehyde oxidase or xanthine dehydrogenase). This cluster also contains a

panB (*SMc01881*) mutant. In *Salmonella enterica*, a *panB* mutation causes auxotrophy for pantothenate (Rubio and Downs, 2002). It can be suggested that in *S. meliloti* the function of PanB can also be performed by another protein, probably by the product of *SMb20821* that at the amino acid level shares 31% identity with *SMc01881* and contains a conserved PanB domain. Three other clones in cluster 4 have mutations in *xylB* (coding for a putative xylulose kinase protein that participates in degradation of D-xylose), *SMb20360* (encoding a putative protease subunit of an ATP-dependent Clp protease) and *SMb20931* (coding for a putative sugar uptake ABC transporter periplasmic solute-binding protein precursor).

Cluster 5 contains two clones, *cysK2* and *SMc03782* mutants, impaired in growth under all conditions tested and more impaired in growth in VMM. *cysK2* encodes a probable cysteine synthase A (O-acetylserine sulfhydrylase A) whereas the gene product of *SMc03782* shows similarities to membrane-bound metallopeptidases involved in cell division and chromosome partitioning.

Cluster 6 and 7 contain mutants which are more impaired in competitiveness under normal conditions than in stress conditions. Such a pattern occurs, probably, due to the fast growth of non-stressed cultures in the exponential phase in comparison to growth of SDS- and salt-stressed cultures. In the fast-growing cultures, mutants that grow and divide slower than others may be less competitive than in stressed slow-growing cultures, if the slow-growing phenotypes are not caused by the stress condition itself. Cluster 6 contains mutants in the *cmk* gene encoding a putative cytidilate kinase and in *SMb20377* encoding a putative translation initiation inhibitor protein. Cluster 7 contains two clones with mutations in transporter genes (*chrA* and *SMa0070*), a *sodC* (coding for a superoxide-dismutase) mutant and a *SMa0091* (encoding a hypothetical protein) mutant.

Cluster 8 contains mutants that are impaired in growth in TY medium, and partially under other conditions. This cluster includes an *lpsB* (encoding a lipopolysaccharide core biosynthesis mannosyltransferase) mutant that was weakened in competitiveness in the fast-growing TY culture and the TY-SDS culture. Since it was previously shown (Clover *et al.*, 1989) that *lpsB* mutants are sensitive to sodium deoxycholate (DOC), this mutant was expected to be attenuated in SDS-containing medium as well. The second mutant in the cluster has a transposon insertion in the *ppiA* gene that encodes a peptidyl-prolyl isomerase (PPIase). This enzyme has a chaperone-like activity and facilitates the cis-trans isomerisation of peptide bonds N-terminal to proline residues within polypeptide chains (Shaw, 2002). Interestingly, another mutant in this cluster carries a transposon insertion in the *tig* gene that encodes a PPIase as well. Trigger factor (TF) encoded by *tig* is a ribosome-bound protein that combines both functions – peptidyl-prolyl isomerization and chaperone-like activities (Genevaux *et al.*, 2004), similarly to the *ppiA* gene

product. Cluster 8 further contains a *pstC* mutant, whose slow-growing phenotype is especially noticeable in the fast-growing TY culture, and is less obvious in VMM and in the stressed cultures. *pstC* encodes a permease protein of a high affinity, high velocity P_i-specific ABC transporter (Yuan *et al.*, 2006a). A comparatively high concentration of inorganic phosphate in VMM might be the reason for the faster growth of the *pstC* mutant in this medium compared to TY medium.

Three mutants that showed an altered growth behavior in cultivation of the mutant pool in VMM compared to TY medium were analyzed individually in competition with the wild type. In these competition experiments the *proB* mutant (cluster 3) was analyzed in VMM, whereas the *chrA* mutant (cluster 7) and the *tig* mutant (cluster 8) were tested in TY. In accordance with the competition experiment analyzing the mutant pool by quantification of the signature tags in microarray hybridizations, the three individually tested mutants showed a reduced competitiveness compared to the wild type.

C.10 STUDY OF SURVIVAL AND COMPETITIVENESS OF MUTANTS *IN PLANTA*

C.10.1 Screening procedure

The most important task in this PhD thesis was the application of the STM approach to identify the rhizobial genes important for the survival and competitiveness in symbiosis. For this, a system had to be used that would allow screening of about 400 mutants (one set of tagged mutants) in one experiment. The quantity of mutants that can be tested simultaneously *in planta* depends on the quantity of nodules that can be obtained in one experiment. In order to produce statistically reasonable results, a quantity of harvested nodules is required that exceeds the quantity of tested mutants in at least 4 – 5 fold and reaches, in this case, 1.5 to 2 thousands nodules. To obtain such a high quantity of nodules in one experiment, the aeroponic system was used. The plant roots grow especially well in this system, so that 150 – 200 nodules can be formed per plant, whereas up to 80 plants are cultivated in one aeroponic tank (Fig. C12). Furthermore, the aeroponics allows spreading of the mixture of mutants equally to all the roots and, since no solid substrate is used, provides an easy access to the roots.

In the experiment *in planta*, both sets of mutants were screened: set 1, used in a validation experiments and set 2. For each set, a separate aeroponic tank with approximately 60 *M. sativa* plants was run. Mature nodules were collected 4 weeks after inoculation. For the first set, 18 ml

of nodules were collected, and for the second set, 16 ml of nodules were collected. Nodules were stored in 9 aliquots at -80 °C (5 aliquots for set 1 and 4 aliquots for set 2).

Fig. C12. Aeroponic tank used in this study

Total plant and bacterial DNA was isolated from the nodules. DNA isolations were performed for each aliquot of nodules (1.5 - 2.5 g) separately and the tags were amplified from this DNA using Cy3-labeled primers. DNA was also isolated from the cultures by which the aeroponic tanks were inoculated. The tags were amplified by Cy5-labeled primers from this DNA and used as reference in the tag-microarray hybridizations.

Four microarray hybridizations were made for each set. Three of these hybridizations were made with Cy3-labeled targets synthesized from the separate DNA isolations from different aliquots of nodules. The fourth hybridization for each set was made using targets synthesized from a mixture of DNA from all three aliquots.

Hybridization, scanning, microarray image processing and data preprocessing were performed in the same way as in the pilot verification experiments. The normalization method used in this experiment was a modification of the global median normalization. Here, 15% of the smallest m-values were excluded from the pool of data used to calculate the median. This was made because a bias toward low m-values was expected. The two mutant sets tested here, contained approximately 5% of mutants that were expected to be attenuated in symbiosis (internal positive controls). Moreover, the chance of finding more attenuated mutants was much higher than the chance of finding mutants that survives especially well in the plant. The calculation of the median excluding 15% of the small m-values resulted in a reasonable number of attenuated and of well-growing mutants. Furthermore, in this way most of internal positive controls could be detected as attenuated.

At the next step of statistical analysis, the mean weighted values were calculated for each

mutant grouping the data from three slides that represent three aliquots of nodules. Based on these $mean_w$ values, 83 mutants were found to be attenuated (37 in the set 1 and 46 in set 2) and 50 mutants were found to be highly represented in the sets under symbiotic conditions (30 in set 1 and 20 in set 2). In order to filter possible false positive clones out, $mean_w$ values from three grouped slides were compared to the data from the fourth slide. The mutants whose absolute m-values in the fourth hybridization were lower than 0.7 were filtered out. Mutants whose respective tags were not detected on the fourth slide were not filtered out from the list of attenuated mutants, since low a-values could originate from very low signals in the Cy3 (experiment) channel (in case of highly represented mutants such cases were not observed). As a result of this filtering, 67 attenuated mutants (30 from the set 1 and 37 from the set 2) and 29 highly represented mutants (22 from set 1 and 7 from set 2) were considered as mutants showing an altered phenotype under symbiotic conditions. The resulting attenuated mutants are listed in the Table C3, and the highly represented mutants are listed in the table C4. All the tested mutants with corresponding m-values and p-values are listed in Appendix (Tables G3 and G4)

TABLE C3. Characteristics of mutants showing an attenuated phenotype under symbiotic conditions

set	mutant ID in the set	mutantID complete	mutated gene	function	m-value slide 4	$mean_w$ value, 3 slides combined
1	1-9C	2011mTn5STM.1.10.C09	SMa0244	hypothetical protein	NA	-0.82
1	2-2D	2011mTn5STM.2.09.D02	SMa0322	hypothetical protein	-0.87	-0.80
2	4-6E	2011mTn5STM.4.09.E06	SMa0414	hypothetical protein	NA	-0.80
1	1-11D	2011mTn5STM.1.05.D11	SMa0814	NifB FeMo cofactor biosynthesis protein	-1.06	-1.09
1	3-1A	2011mTn5STM.3.12.A01	SMa0819	FixB electron transfer flavoprotein alpha chain	NA	-0.87
2	2-6E	2011mTn5STM.2.07.E06	SMa0829	NifK nitrogenase Fe-Mo beta chain	-1.10	-0.96
1	1-9D	2011mTn5STM.1.05.D09	SMa0840	NodD3 transcriptional regulator	NA	-0.93
1	2-5D	2011mTn5STM.2.08.D05	SMa0849	SyrM transcriptional regulator	-0.96	-0.78
2	3-6G	2011mTn5STM.3.06.G06	SMa0863	NodJ membrane transport protein	-1.03	-0.82
2	1-4H	2011mTn5STM.1.10.H04	SMa0873	NifN Nitrogenase Fe-Mo cofactor biosynthesis protein	NA	-0.97
1	1-4A	2011mTn5STM.1.11.A04	SMa1004	hypothetical protein	-0.83	-0.86
1	1-4C	2011mTn5STM.1.01.C04	SMa1229	FixL oxygen regulated histidine kinase	-1.11	-1.11
2	1-8C	2011mTn5STM.1.13.C08	SMa1229	FixL oxygen regulated histidine kinase	-1.08	-0.96
2	3-7H	2011mTn5STM.3.05.H07	SMa1229	FixL oxygen regulated histidine kinase	-1.09	-0.91
2	4-8B	2011mTn5STM.4.10.B08	SMa1742	putative iron uptake protein	-0.72	-0.73
1	1-6G	2011mTn5STM.1.08.G06	SMa1798	Kup2 Potassium uptake protein	-1.05	-0.94
1	4-3F	2011mTn5STM.4.08.F03	SMa1913	putative transport protein	-0.76	-0.86
1	4-6D	2011mTn5STM.4.04.D06	SMb20227	nutrient deprivation-induced protein A (NdiA1)	-0.77	-0.75
1	4-4D	2011mTn5STM.4.04.D04	SMb20481	AsnO asparagine synthase, glutamine-hydrolyzing	-1.05	-1.10

2	3-3H	2011mTn5STM.3.06.H03	SMb20509	putative transcriptional regulator	-0.74	-0.92
2	2-9H	2011mTn5STM.2.06.H09	SMb20612	C4-dicarboxylate transport sensor protein DctB	-0.83	-0.76
2	2-11D	2011mTn5STM.2.05.D11	SMb20615	putative thiamine biosynthesis protein ThiC	-1.06	-0.92
1	1-9A	2011mTn5STM.1.04.A09	SMb20616	ThiO putative thiamine biosynthesis oxidoreductase protein	-0.98	-0.88
1	2-8E	2011mTn5STM.2.08.E08	SMb20757	methylmalonyl-CoA mutase protein BhbA	-0.84	-0.87
1	4-11E	2011mTn5STM.4.04.E11	SMb20942	UDP glucose 4-epimerase protein ExoB	-1.14	-1.07
1	2-6G	2011mTn5STM.2.05.G06	SMb20943	acetyltransferase protein ExoZ	-1.12	-1.05
1	1-8G	2011mTn5STM.1.12.G08	SMb20948	ExoU glucosyltransferase	-1.05	-1.06
1	3-9B	2011mTn5STM.3.10.B09	SMb20957	exopolysaccharide biosynthesis protein ExoA	-1.15	-1.18
1	1-9B	2011mTn5STM.1.01.B09	SMb20958	exopolysaccharide biosynthesis protein ExoM	-1.08	-1.02
1	2-9C	2011mTn5STM.2.09.C09	SMb20959	probable exopolysaccharide biosynthesis protein ExoO	-1.05	-1.08
2	1-9B	2011mTn5STM.1.07.B09	SMb20962	probable phoshomethylpyrimidine kinase protein ThiD	-1.13	-1.05
1	4-12E	2011mTn5STM.4.08.E12	SMb21158	putative transcriptional regulator, probably of sugar phosphate metabolism protein	-1.13	-0.96
2	2-10A	2011mTn5STM.2.12.A10	SMc00129	sensor histidine kinase protein FeuQ	-0.99	-0.96
2	1-1A	2011mTn5STM.1.11.A01	SMc00169	NAD-malic enzyme oxidoreductase protein Dme	-0.94	-0.99
2	2-7C	2011mTn5STM.2.05.C07	SMc00236	probable indole-3-glycerol phosphate synthase TrpC	-0.99	-0.82
1	1-10E	2011mTn5STM.1.12.E10	SMc00349	GTP-binding membrane protein LepA	-0.97	-0.94
2	4-3C	2011mTn5STM.4.02.C03	SMc00644	hypothetical protein	-1.13	-1.08
1	4-6B	2011mTn5STM.4.03.B06	SMc00781	methylmalonate-semialdehyde dehydrogenase MmsA (IolA)	-1.09	-1.03
1	4-3G	2011mTn5STM.4.02.G03	SMc00828	hypothetical protein	-0.87	-0.86
2	3-2B	2011mTn5STM.3.07.B02	SMc00963	biotin ABC transporter, permease component BioN	-1.09	-1.07
1	4-12C	2011mTn5STM.4.04.C12	SMc01053	siroheme synthase protein CysG	-1.06	-1.06
2	4-11D	2011mTn5STM.4.02.D11	SMc01139	RNA polymerase sigma-54 factor	-1.17	-1.03
2	4-2E	2011mTn5STM.4.05.E02	SMc01183	Putative LexA repressor regulator	-0.89	-0.86
1	2-1H	2011mTn5STM.2.12.H01	SMc01219	lipopolysaccharide core biosynthesis mannosyltransferase LpsB	-0.91	-0.84
2	3-5F	2011mTn5STM.3.10.F05	SMc01431	probable acetolactate synthase isozyme III large subunit protein IlvI	-0.99	-0.83
2	1-7G	2011mTn5STM.1.02.G07	SMc01877	probable DNA repair protein RecN	-0.77	-0.73
2	3-3C	2011mTn5STM.3.10.C03	SMc01929	hypothetical protein	-0.70	-0.79
1	2-1D	2011mTn5STM.2.10.D01	SMc01950	high-affinity branched-chain amino acid ABC transporter, permease protein LivM	-0.86	-0.81
2	2-2A	2011mTn5STM.2.09.A02	SMc02109	probable ATP-dependent CLP protease ATP-binding subunit protein ClpA	-0.96	-1.03
2	4-7A	2011mTn5STM.4.10.A07	SMc02123	hypothetical protein	-0.81	-0.77
2	1-6C	2011mTn5STM.1.13.C06	SMc02143	putative phosphate transport system permease ABC transporter protein PstA	-0.95	-0.95
1	4-12B	2011mTn5STM.4.09.B12	SMc02144	phosphate transport system permease ABC transporter protein PstC	-0.76	-0.93
2	2-4D	2011mTn5STM.2.01.D04	SMc02165	probable orotate phosphoribosyltransferase protein PyrE	-0.96	-0.84
1	4-6E	2011mTn5STM.4.02.E06	SMc02226	hypothetical protein	-1.05	-0.97
2	4-1A	2011mTn5STM.4.08.A01	SMc02274	putative capsule polysaccharide export protein precursor RkpU	-0.75	-0.74
2	1-6H	2011mTn5STM.1.10.H06	SMc02361	cytochrome C-type biogenesis transmembrane protein CycH	-0.85	-0.90
2	3-10F	2011mTn5STM.3.06.F10	SMc02562	phosphoenolpyruvate carboxykinase protein PckA	-1.04	-0.85

2	1-12H	2011mTn5STM.1.09.H12	SMc02767	N-5'-phosphoribosylanthranilate isomerase TrpF	-0.95	-0.94
2	2-6A	2011mTn5STM.2.10.A06	SMc03181	pH adaptation potassium efflux system transmembrane protein PhaD1	-1.03	-0.92
2	3-9G	2011mTn5STM.3.08.G09	SMc03277	MFS-type transport protein	-0.93	-0.88
2	4-3F	2011mTn5STM.4.09.F03	SMc03797	homoserine O-succinyltransferase MetA	-1.20	-1.20
2	3-12G	2011mTn5STM.3.08.G12	SMc03849	cytochrome c-type biogenesis protein CcmC	-1.06	-0.98
2	4-5B	2011mTn5STM.4.08.B05	SMc03900	beta-glucan export ATP-binding protein NdvA	-1.06	-1.02
2	3-5A	2011mTn5STM.3.05.A05	SMc04045	probable dihydroxy-acid dehydratase protein IlvD2	-0.99	-0.85
1	1-11C	2011mTn5STM.1.08.C11	SMc04346	IlvC ketol-acid reductoisomerase protein	-0.83	-0.78
2	1-11D	2011mTn5STM.1.03.D11	SMc04346	IlvC ketol-acid reductoisomerase protein	NA	-0.88
2	2-7E	2011mTn5STM.2.02.E07	SMc04405	3-isopropylmalate dehydrogenase LeuB	-1.12	-0.94

TABLE C4. Characteristics of mutants highly represented under symbiotic conditions

set	mutant ID in the set	mutantID complete	mutated gene[a]	function	m-value slide 4	$mean_w$ value, 3 slides combined
1	2-7H	2011mTn5STM.2.09.H07	pSymB, 96435	intergenic	1.20	1.69
1	3-11B	2011mTn5STM.3.12.B11	pSymB, 1003846	intergenic	1.53	1.25
1	4-7F	2011mTn5STM.4.03.F07	S. meliloti chromosome, 2051945	intergenic	0.90	0.86
1	2-12B	2011mTn5STM.2.09.B12	S. meliloti chromosome, 2862221	intergenic	1.11	1.32
1	2-10B	2011mTn5STM.2.04.B10	SMa0565	hypothetical protein	1.20	1.29
1	1-5G	2011mTn5STM.1.12.G05	SMa0621	FixI2 E1-E2 type cation ATPase	0.79	0.86
1	3-5A	2011mTn5STM.3.11.A05	SMb20037	shikimate 5-dehydrogenase protein AroE2	1.20	1.12
1	3-10H	2011mTn5STM.3.08.H10	SMb20360	hypothetical protein	1.21	1.21
2	3-9D	2011mTn5STM.3.05.D09	SMb20426	hypothetical transcriptional regulator protein	1.06	1.38
1	3-9D	2011mTn5STM.3.10.D09	SMb20582	putative transcriptional regulator, LysR family protein	1.06	1.10
2	3-11B	2011mTn5STM.3.10.B11	SMb20607	hypothetical protein	0.72	0.88
1	2-10D	2011mTn5STM.2.03.D10	SMb20748	glycosyltransferase protein PssF	0.87	0.75
2	1-10E	2011mTn5STM.1.03.E10	SMb21630	hypothetical protein	1.58	1.70
1	2-2E	2011mTn5STM.2.04.E02	SMb21633	PaaG enoyl-CoA hydratase protein	1.00	0.74
1	3-12D	2011mTn5STM.3.07.D12	SMc00334	Cmk cytidilate kinase protein	1.09	0.80
1	1-1C	2011mTn5STM.1.09.C01	SMc00513	putative amino-acid binding periplasmic protein	0.91	1.10
2	1-3A	2011mTn5STM.1.06.A03	SMc00790	probable dipeptide transport ATP-binding ABC transporter protein DppF	0.78	0.71
2	1-1E	2011mTn5STM.1.02.E01	SMc00898	probable glutathione-regulated potassium-efflux system transmembrane protein KefB	0.91	1.23
2	3-9C	2011mTn5STM.3.07.C09	SMc01127	OlsB protein required for ornithine-containing lipid biosynthesis	1.17	0.75
1	3-5D	2011mTn5STM.3.07.D05	SMc01406	putative transcription regulator protein	0.80	0.76

1	2-11B	2011mTn5STM.2.06.B11	SMc01881	PanB 3-methyl-2-oxobutanoate hydromethyltransferase protein	2.84	2.01
1	3-8H	2011mTn5STM.3.09.H08	SMc01965	putative spermidine/putrescine transport ATP-binding ABC transporter protein	0.94	0.84
1	1-9F	2011mTn5STM.1.02.F09	SMc02612	glutamate synthase family protein GlxD	0.98	0.82
1	3-6D	2011mTn5STM.3.11.D06	SMc03032	flagellar P-ring precursor transmembrane protein	1.23	1.08
2	4-12A	2011mTn5STM.4.05.A12	SMc03112	methionine synthase MetH	0.85	0.84
1	2-5H	2011mTn5STM.2.05.H05	SMc03164	xylulose kinase protein XylB	0.78	1.10
1	4-11D	2011mTn5STM.4.09.D11	SMc03873	RNA polymerase sigma factor protein SigC	1.23	1.15
1	3-1B	2011mTn5STM.3.08.B01	SMc04217	hypothetical protein	0.71	0.79
1	2-11D	2011mTn5STM.2.01.D11	SMc04452	NADH dehydrogenase transmembrane protein Ndh	1.17	1.02

[a] For mutants carrying a transposon insertion in an intergenic region the exact position of insertion in the respective replicon is given.

C.10.2 Determination of quantity of mutants in single nodules

An additional experiment was performed in order to determine how many mutants enter one nodule in aeroponics. To answer this question, 9 nodules from each aeroponic tank were tested. The DNA from each of these nodules was used to perform separate target preparations and microarray hybridizations. The presence of a specific mutant in a nodule was detected based on high average spot intensity values for both tags. An example of the results from such hybridization is presented on Fig. C13. As a result of this experiment, it could be determined that most of nodules (15 from 18 nodules) were occupied by a single mutant clone. Two nodules contained two different mutant clones, and one nodule contained three different mutant clones. Two of the nodules (set 1, nodules 7 and 9) contained the same mutant clone; all other nodules contained different mutant clones (Table C5). It can be therefore concluded that most of the nodules in aeroponic conditions were occupied by a single mutant clone.

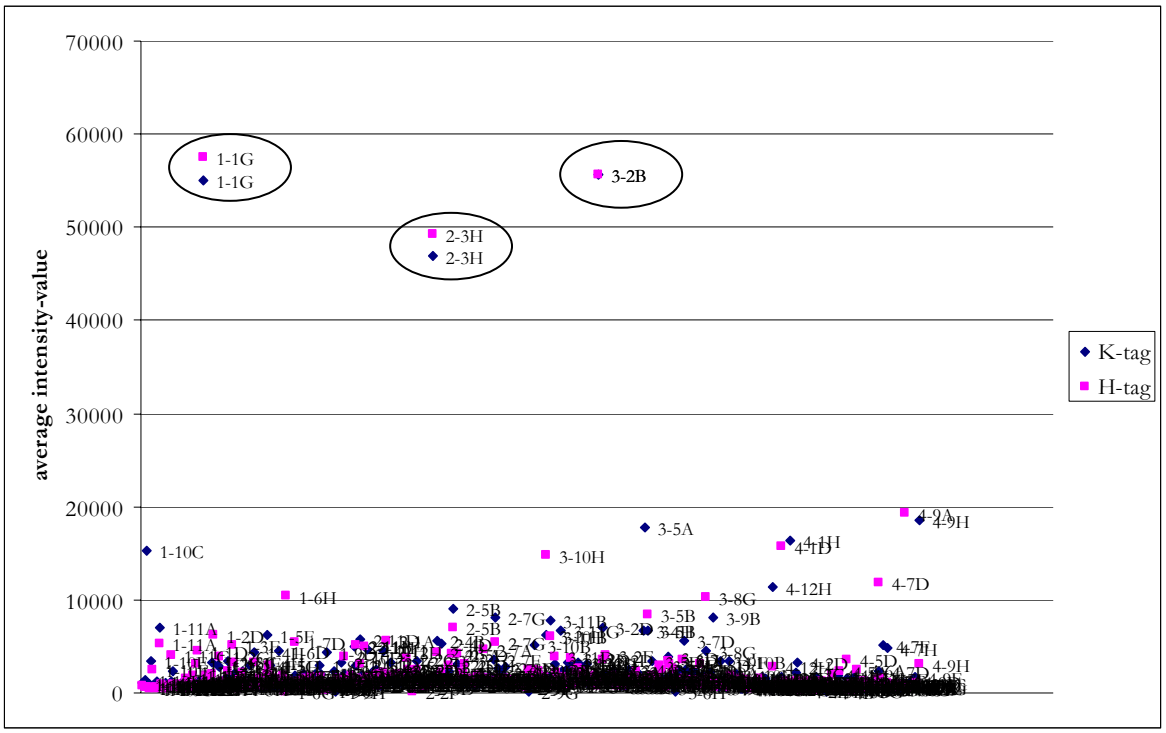


Fig. C13. Determination of quantity of different mutant clones in a single nodule from the aeroponic tank. Results of hybridization for the nodule 2 from set 1. Here, mutants with IDs 1-1G, 2-3H and 3-2B were detected. The average H-tag and K-tag intensity values for the clone 3-2B lay so close that they overlap on the diagram.

TABLE C5. Quantity of different mutant clones in single nodules from the aeroponic tanks

set	nodule Nr	quantity of mutants in the nodule	Short mutant IDs
1	1	2	1-5G, 2-7G
1	2	3	1-1G, 2-3H, 3-2B
1	3	1	3-9D
1	4	1	3-12A
1	5	1	1-7H
1	6	1	2-2B
1	7	1	2-10B
1	8	1	2-11F
1	9	1	2-10B
2	1	1	1-4F
2	2	1	2-6D
2	3	1	2-3A
2	4	2	1-8A, 4-7D
2	5	1	3-8F
2	6	1	2-11C
2	7	1	1-10A
2	8	1	1-9F
2	9	1	2-2B

C.10.3 Phenotype confirmation for attenuated mutants

Since the signature tagged mutagenesis is a screening strategy, some of the attenuated mutants were tested individually for their competitiveness phenotype in order to confirm or to reject the phenotype observed in the STM experiment. From 67 attenuated mutants, 25 mutants with no previously known fixation or nodulation defect were chosen for the testing procedure. *M. sativa* plants, growing on agar plates, were inoculated by a 1 : 1 mixture of wild type Rm2011 and a checked mutant, and after 4-5 weeks the nodules were harvested. Since the mTn*5* transposon carries a neomycin resistance gene, the presence of a mTn*5* mutant in a nodule can be revealed by plating the nodule content on an agar plate with neomycin-containing medium. In order to identify the wild type/mutant ratio of dividing bacteria in the nodules, the equal quantity of diluted nodule content is plated on neomycin plates and on plates without neomycin, and the number of clones grown on these plates is compared. For each mutant, 25 to 30 nodules from 15-18 plants were analyzed, and the average value from the percentage of mutant in each nodule was calculated. The attenuated competitiveness phenotype could be confirmed for 22 mutants from 25 (Table C6).

TABLE C6. Competition phenotypes of 25 clones in individual tests

mutantID complete	mutated gene	function	$mean_w$ value, 3 slides combined	number of analyzed nodules	average percentage of mutant in nodules
2011mTn5STM.4.09.F03	SMc03797	homoserine O-succinyltransferase MetA	-1.20	26	0.04
2011mTn5STM.4.04.C12	SMc01053	siroheme synthase protein CysG	-1.06	30	0.45
2011mTn5STM.1.04.A09	SMb20616	ThiO putative thiamine biosynthesis oxidoreductase protein	-0.88	30	1.89
2011mTn5STM.1.12.E10	SMc00349	GTP-binding membrane protein LepA	-0.94	30	3.97
2011mTn5STM.4.09.B12	SMc02144	phosphate transport system permease ABC transporter protein PstC	-0.93	29	4.06
2011mTn5STM.2.12.A10	SMc00129	sensor histidine kinase protein FeuQ	-0.96	29	4.54
2011mTn5STM.1.07.B09	SMb20962	probable phoshomethylpyrimidine kinase protein ThiD	-1.05	25	4.96
2011mTn5STM.3.08.G12	SMc03849	cytochrome c-type biogenesis protein CcmC	-0.98	27	7.93
2011mTn5STM.4.02.E06	SMc02226	hypothetical protein	-0.97	33	1.24
2011mTn5STM.1.13.C06	SMc02143	putative phosphate transport system permease ABC transporter protein PstA	-0.95	30	11.45
2011mTn5STM.2.09.A02	SMc02109	probable ATP-dependent CLP protease ATP-binding subunit protein clpA	-1.03	28	13.31
2011mTn5STM.2.10.D01	SMc01950	high-affinity branched-chain amino acid ABC transporter, permease protein LivM	-0.81	31	15.31
2011mTn5STM.4.05.E02	SMc01183	Putative LexA repressor regulator	-0.86	30	16.67
2011mTn5STM.4.04.D04	SMb20481	AsnO asparagine synthase, glutamine-hydrolyzing	-1.10	25	18.04

2011mTn5STM.4.02.C03	SMc00644	hypothetical protein	-1.08	30	18.98
2011mTn5STM.1.02.G07	SMc01877	probable DNA repair protein RecN	-0.73	30	21.70
2011mTn5STM.4.04.D06	SMb20227	nutrient deprivation-induced protein A (NdiA1)	-0.75	27	25.46
2011mTn5STM.4.10.A07	SMc02123	hypothetical protein	-0.77	25	27.16
2011mTn5STM.3.06.H03	SMb20509	putative transcriptional regulator	-0.92	25	28.25
2011mTn5STM.4.02.G03	SMc00828	hypothetical protein	-0.86	31	31.61
2011mTn5STM.4.09.E06	SMa0414	hypothetical protein	-0.80	29	32.03
2011mTn5STM.3.10.C03	SMc01929	hypothetical protein	-0.79	28	34.24
2011mTn5STM.3.08.G09	SMc03277	MFS-type transport protein	-0.88	28	49.43
2011mTn5STM.2.09.D02	SMa0322	hypothetical protein	-0.80	28	53.15
2011mTn5STM.4.08.F03	SMa1913	putative transport protein	-0.86	27	77.73

C.10.4 Symbiotic phenotypes other than reduced symbiotis competitiveness

Each of 25 mutants whose competition phenotypes were tested individually was also tested to see if any defect in symbiosis can be observed if the plants are inoculated only by the mutant culture. Only severe symbiotic defects, like Fix^- phenotype or strong delay in nodulation were considered. Mutants in four genes were found to have such symbiotic defects.

Clone 2011mTn5STM.3.08.G1 that has a transposon insertion in the *ccmC* (SMc03849) gene showed a Fix^- phenotype: the induced nodules were white (Fig. C14 B) and the plant shoots started to turn yellow 2.5 – 3 weeks after inoculation.

Clone 2011mTn5STM.4.04.D04 that has a mutation in *asnO* (SMb20481) gene could induce only very few nodules. On one plant, the nodules formed by this mutant had a strongly abnormal morphology (Fig. C14 C), but all nodules formed on other plants looked normal and had a Fix^+ phenotype (Fig. C14 D). Most of nodules, induced by *asnO* mutant, were bifurcated.

Clones 2011mTn5STM.4.04.C12 and 2011mTn5STM.4.09.F03 carrying mTn*5* insertion in *cysG* (SMc01053) gene and in *metA* (SMc03797) respectively, showed a delayed nodulation phenotype, and the nodules formed by these mutants turned pink 6-8 days later than the wild type induced nodules. While the *metA* mutant formed pink effective nodules three weeks after inoculation (Fig C14 E) on all plants, *cysG* mutant induced different kinds of nodules: on some plants, small white nodules were formed, and on other plants nodules were pink and large. Fig C14 F shows three plants, inoculated by the *cysG* mutant and growing in the same plate, that formed nodules of different morphology.

To analyze these phenotypes more closely, the kinetics of nodule formation was observed. Additionally, acetylene reduction assay and histochemical GUS staining of nodules were performed.

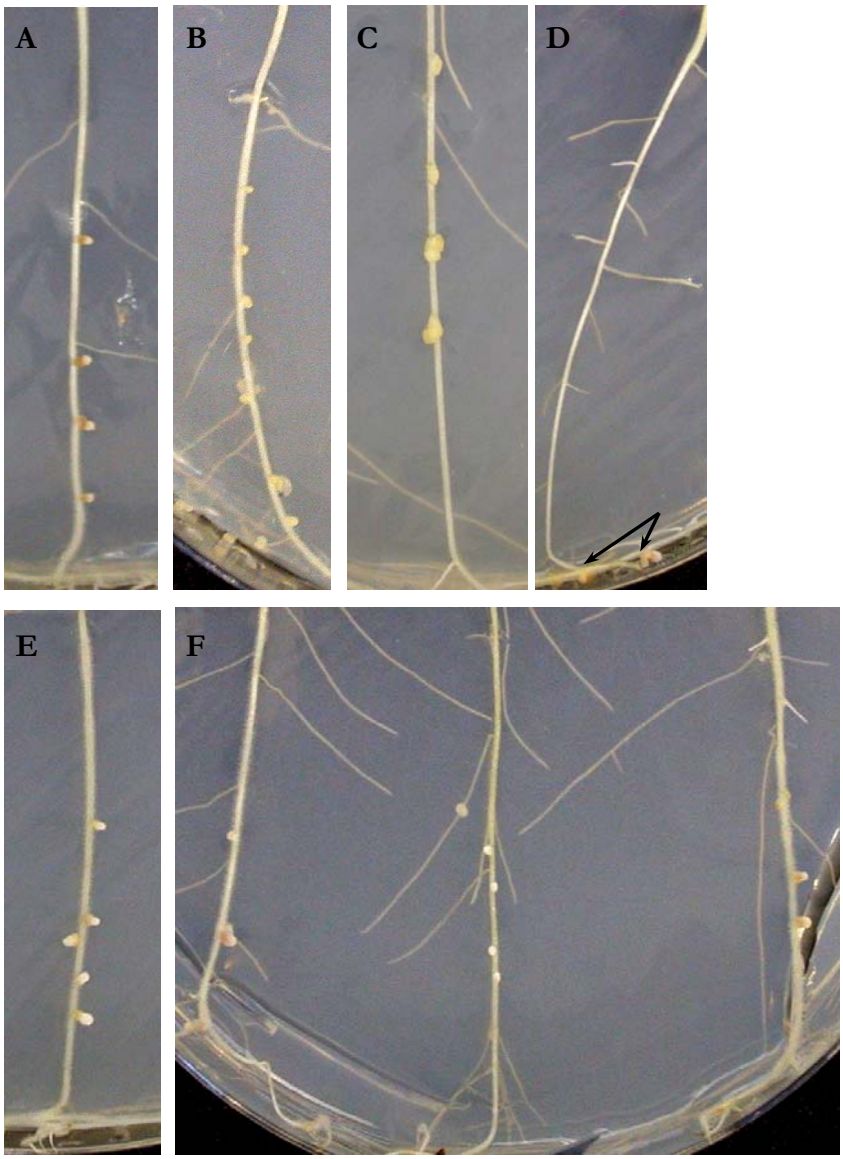


Fig. C14. Phenotype of nodules at 21 dpi induced by: (A) wild type *S. meliloti*, (B) *ccmC* mutant, (C) and (D) *asnO* mutant, (E) *metA* mutant, and (F) *cysG* mutant. In (D), the arrows indicate the nodules.

C.10.4.1 Kinetics of nodule formation

In case of *asnO*, *cysG* and *metA* mutants and wild type Rm2011, formation of nodules on 110 plants were analyzed. In case of *ccmC*, 27 plants were observed. Plants were inoculated by a respective bacterial culture ($o.D._{600}$=0.02) and nodules were counted at 9, 12, 17 and 21 days post inoculation (dpi). *M. sativa* plants were grown on agar plates, with 3 plants per plate. The results of this experiment are depicted on Fig. C15.

As expected, the *ccmC* mutant that showed a Fix^- phenotype, induced more nodules than the wild type (Paau *et al.*, 1985). Plants inoculated by *cysG* mutant and plants inoculated by the *metA* mutant started to form nodules later than plants inoculated by the wild type *S. meliloti*, but formed

the same, or even higher, quantity of nodules at 17-21 dpi. The *asnO* mutant induced strikingly few nodules: even at 21 dpi only 21 nodules were formed on 110 plants in generally.

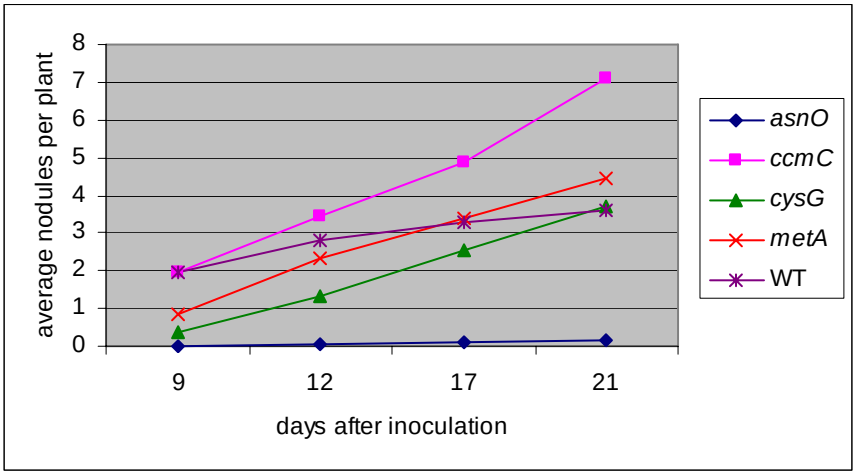


Fig. C15. Kinetics of nodule formation by the *asnO*, *cysG*, *ccmC* and *metA* mutants comparing to the wild type Rm2011

C.10.4.2 Acetylene reduction assay

In order to investigate the nitrogen fixation rate in nodules induced by *cysG*, *ccmC* and *metA* mutants, acetylene reduction assay was performed. The *asnO* mutant was not included in this experiment because the quantity of nodules produced by this mutant was not sufficient.

The whole plants were incubated with acetylene in order to reduce the stress factor in the measurements. Three plants were placed in each tube. In case of plants inoculated by wild type *S. meliloti* and *cysG*, and *metA* mutants, 10 tubes were analyzed. In case of *ccmC* mutant, 5 tubes were analyzed. The results of the assay (Fig. C16) correlate with the observed phenotype of mutants. Particularly, the nodules induced by *ccmC* mutant did not show any acetylene assimilation higher than the background. Nodules induced by the *metA* mutant reduced the same quantity of acetylene as nodules induced by the wild type *S. meliloti* (Fig. C16 B). However, since the quantity of nodules formed on plants inoculated by *metA* mutant was higher than on the plants inoculated by the wild type Rm2011 (Fig. C15), the *metA*-inoculated plants generally reduced more acetylene (Fig. C16 A).

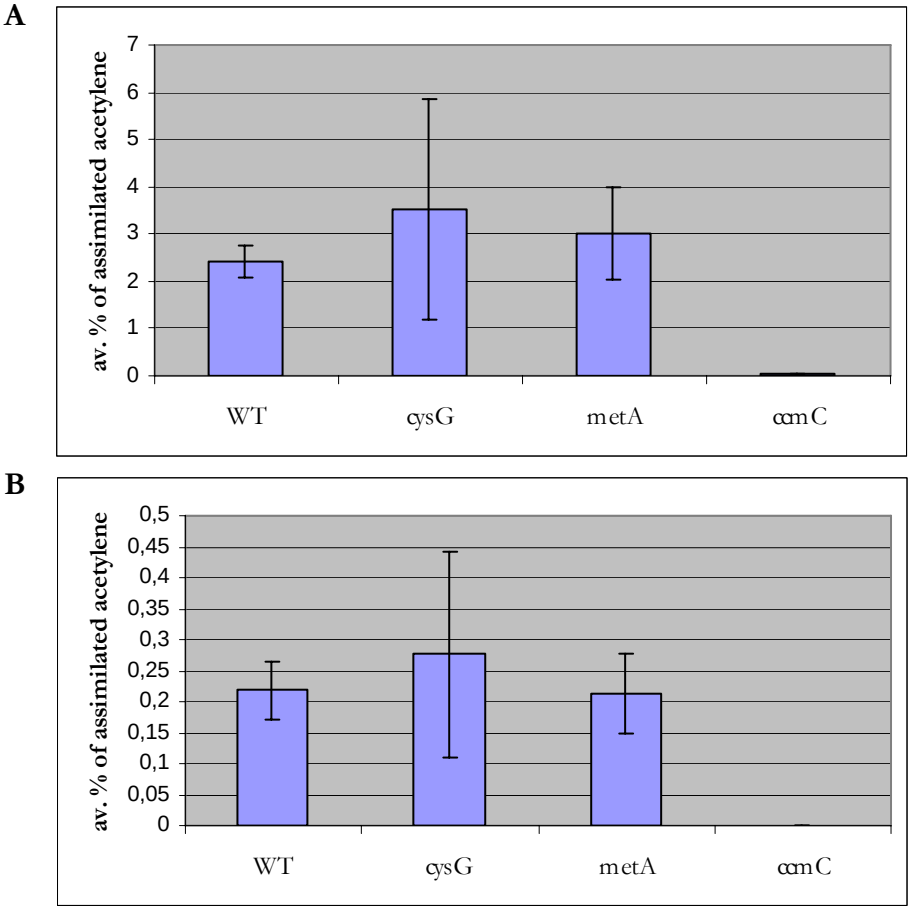


Fig. C16. Percentage of assimilated acetylene per tube (3 plants) (A) and per nodule (B).

The results of acetylene reduction measures for the *cysG* mutant reflect the heterogeneity in the phenotypes of nodules induced by this mutant (fig. C14 F). While in some tubes low levels of acetylene were reduced (1.2-1.4 %), in other tubes the percentage of reduced acetylene was much higher than expected (4.5% or even 9.5% reduced acetylene per tube in comparison to the average of 2.4% in the case of the wild type inoculated plants). Generally, nodules induced by the *cysG* mutant reduced more acetylene than the wild type induced nodules.

C.10.4.3 Histochemical GUS staining

In the tested *ccmC*, *asnO*, *metA* and *cysG* mutants, the mTn*5* transposon was inserted in a way that the promotorless *gusA* gene had the same direction as the interrupted gene. It was therefore possible to visualize the promoter activity of the mutated gene in the nodules based on the expression of the *gusA* gene. Nodules were collected at 22 dpi, sectioned, stained for GUS activity and observed by microscopy. Additionally, nodule sections were stained by 0.1 M potassium iodide in order to visualize the starch granules.

ccmC mutant

Nodules induced by the *ccmC* mutant contain bacteria only in the apical infection zone, where expression of the reporter *gusA* gene is especially strong (Fig. C17, C18). This corresponds with the non-fixing phenotype observed in the acetylene reduction assay and indicates the possible role of the *ccmC* gene in the infection process.

Another interesting feature of the nodules induced by the *ccmC* mutant is the high concentration of starch in the apex part and in the periphery of the nodule (Fig. C18 B)

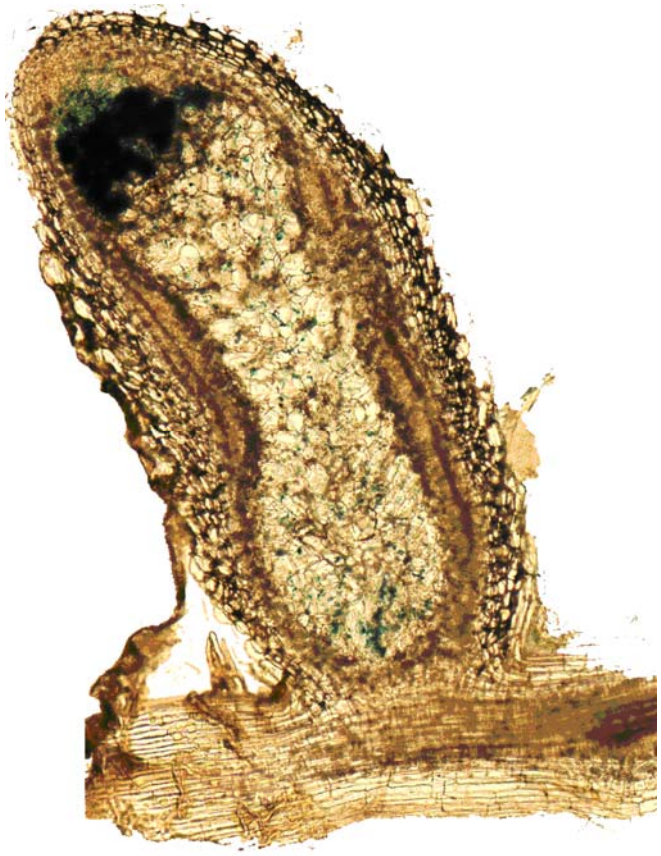


Fig. C17. GUS staining (blue) of a longitudinal section of a nodule induced by the *ccmC* mutant.

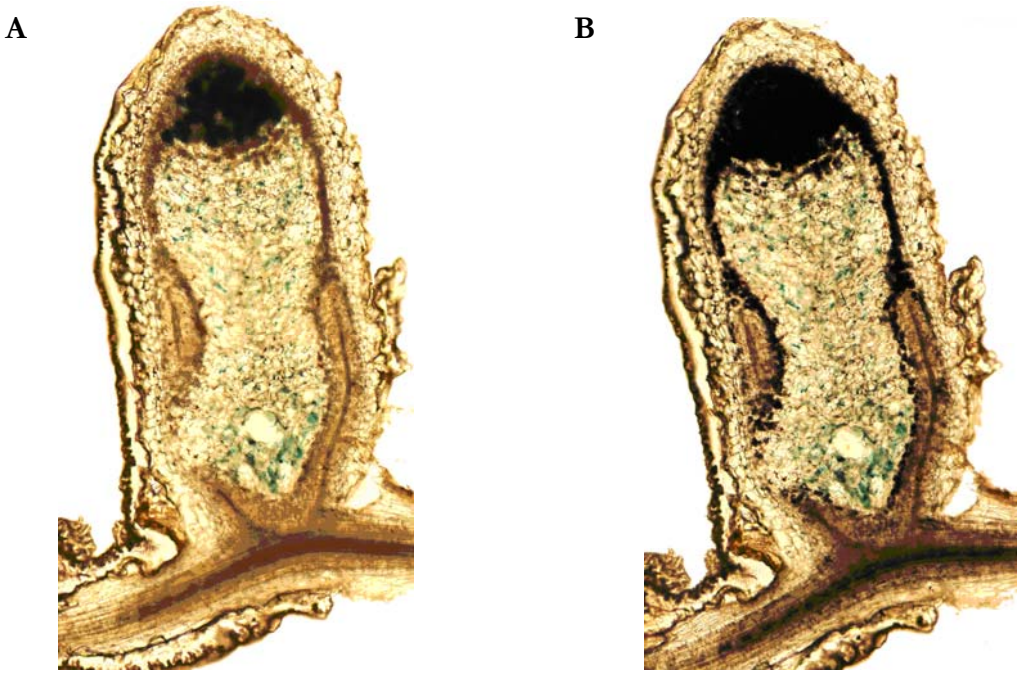


Fig. C18. Longitudinal sections of a nodule induced by the *ccmC* mutant. (A) GUS staining; (B) GUS and K-iodide staining. Starch is stained dark brown.

asnO mutant

One of the pink nodules, induced by this mutant, was analyzed. The nodule is filled with bacteria and the expression of the reporter gene take place in bacteroids of all developmental stages (Fig. C19). Starch granules are observed in the interzone, like in the wild type induced nodules. Nevertheless, the structure of the nodule is quite unusual: it does not have the typical prolonged form with a clear definition of zones throughout all longitudinal sections. In Fig. C19 A C and D, three longitudinal sections of the same nodule are shown. Whereas in section A the meristematic zone I and infection zone II occupy a large part of the nodule, in sections C and D these zones are almost not observed; most of the plant cells on these sections are swollen and filled with bacteroids. Furthermore, the proximal part of the nodule is unusually broad, which probably originates from difficulties in infection.

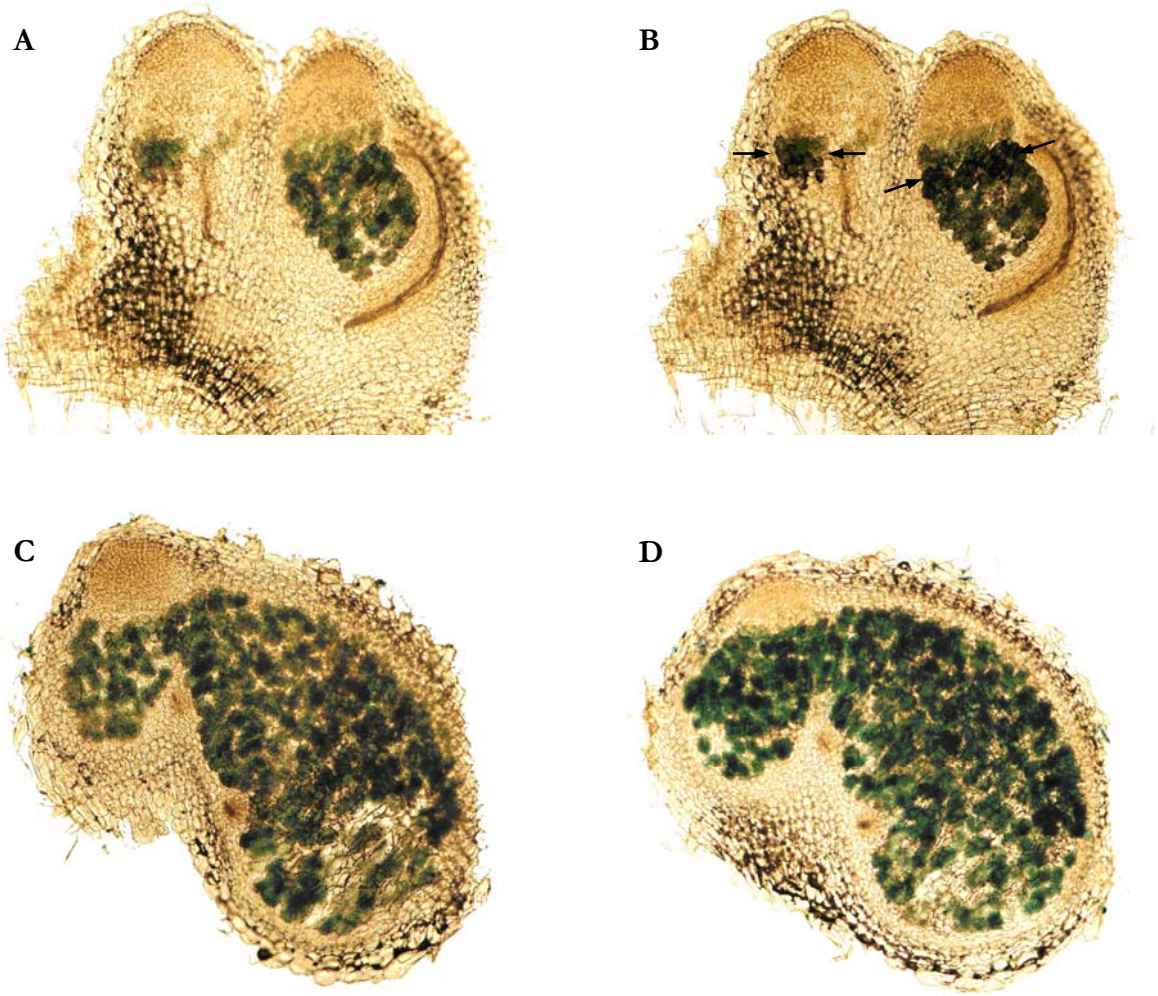


Fig. C19. Longitudinal sections of a nodule induced by the *asnO* mutant. (A), (C), (D) GUS staining; (B) GUS and K-iodide staining. Starch accumulation in the interzone is indicated by arrows.

metA mutant

Longitudinal sections of a nodule induced by the *metA* mutant are shown on Fig. C20. Expression of the reporter gene was detected only in the middle and in the proximal part of the nodule. In addition, some plant cells, filled with starch, could be detected in the middle part of the nodule. Since the nodules induced by *metA* can fix nitrogen at the same rate as the wild type induced nodules, it can be suggested that the *metA* gene is expressed only at a low level in the Zone III nitrogen-fixing bacteroids, and at a higher level in the Zone IV bacteroids.

cysG mutant

Longitudinal sections of a *cysG* mutant-induced pink nodule are shown on the Fig. C21. The general morphology of the nodule appears normal, except for a high quantity of non-infected Zone III cells, located mostly in the center of the nodule. Expression of the reporter gene takes place in the Zone III and Zone IV of the nodule.

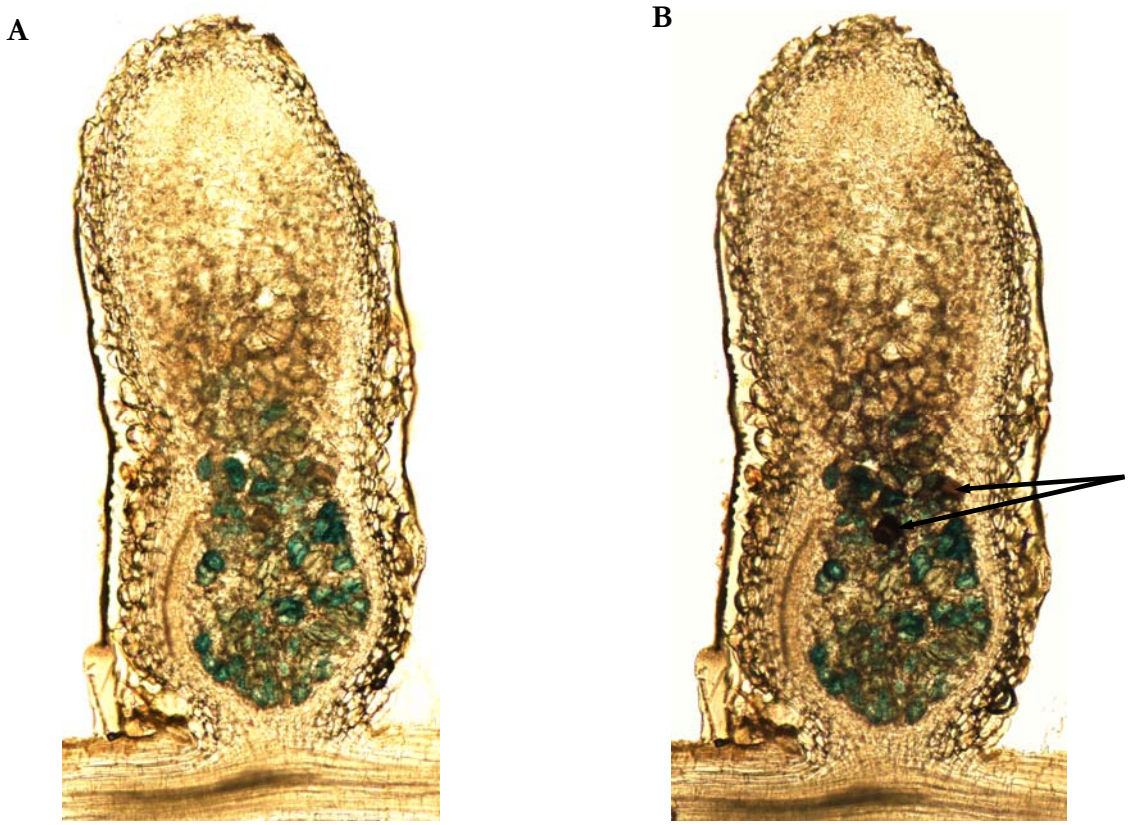


Fig. C20. Longitudinal sections of a nodule induced by the *metA* mutant. (A), GUS staining; (B), GUS and K-iodide staining. Arrows indicate the cells in the middle part of the nodule, filled with starch.

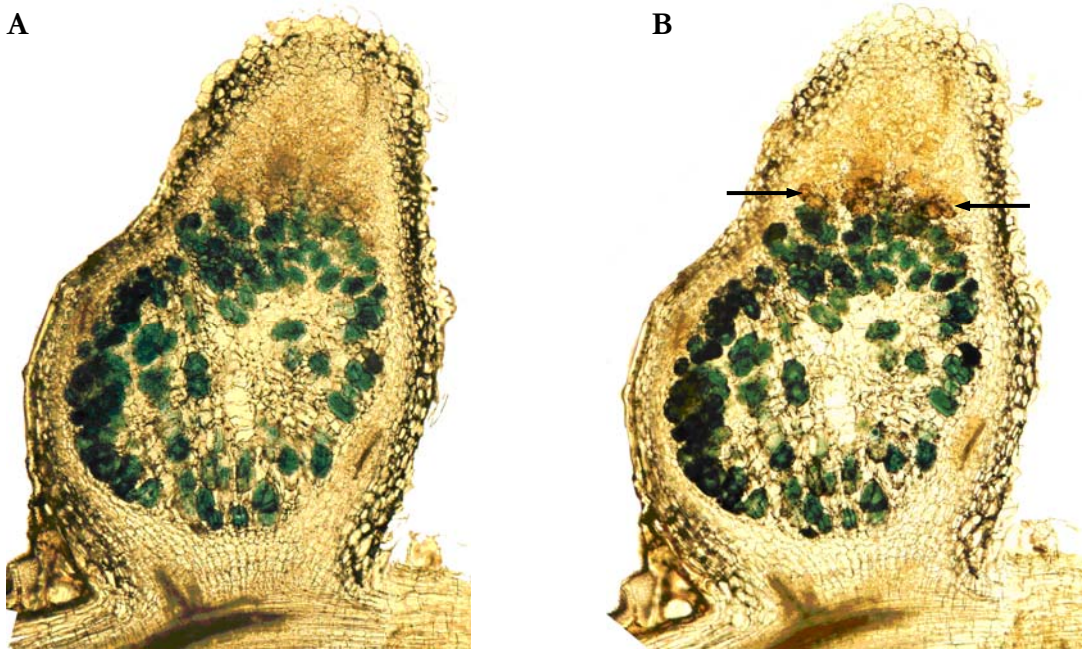


Fig. C21. Longitudinal sections of a nodule induced by the *cysG* mutant. (A), GUS staining; (B), GUS and K-iodide staining. Starch accumulation in the interzone is indicated by arrows.

D. DISCUSSION

D.1 SIGNATURE TAGGED TRANSPOSON MUTAGENESIS IS AN EFFICIENT APPROACH TO STUDY GENES INVOLVED IN SURVIVAL OF RHIZOBIA IN SYMBIOTIC AND FREE-LIVING CONDITIONS

Symbiosis between rhizobia and their legume hosts is a complicated and fragile process that requires participation of many gene products from both symbionts. At present, most of rhizobial genes and many legume genes that are crucial for the establishment of successful symbiosis and/or nitrogen fixation have been discovered. However, many of the genes that influence symbiosis in a more subtle way are still not identified. In order to discover genes which are specifically expressed in symbiosis, novel high-throughput methods such as transcriptomics (Ampe *et al.*, 2003; Becker *et al.*, 2004; Capela *et al.*, 2006; Long *et al.*, 1988), proteomics (Djordjevic *et al.*, 2003), and *in vivo* expression technology (Oke and Long, 1999; Zhang and Cheng, 2006) were successfully applied to *Sinorhizobium meliloti*. Data from these experiments, however, provides no information about the involvement of the identified genes in survival and competitiveness of rhizobia in the host. Usually, to study the function of genes in symbiosis and competitiveness, gene disruption or deletion with subsequent analysis of symbiotic phenotypes is applied. This method is efficient when one or several few genes have to be tested. However, it is very laborious and time consuming when used for the screening of hundreds or thousands of mutants. A pre-selection of mutants in defined free-living conditions is often performed in order to reduce the number of mutants that have to be tested *in planta* (Milcamps *et al.*, 1998; Summers *et al.*, 1998; Trzebiatowski *et al.*, 2001; Uhde *et al.*, 1997). At the same time, such a pre-selection reduces the chances to find unknown genes involved in symbiosis, if this genes are not expressed in the tested free-living conditions. The method of signature tagged mutagenesis, in contrast, enables the screening of hundreds of mutants in one passage through the host, so that the whole library of mutants can be screened quite easily.

Originally developed to identify genes involved in the pathogen-host interactions, signature-tagged transposon mutagenesis was used in this project for the first time to study the rhizobia-legume symbiosis. Furthermore, the efficiency of the STM approach was improved here through the creation of a novel, largest to date, library of tagged transposons, and through the combination of the STM approach with two-channel microarrays. The utilization of two signature tags per transposon together with an appropriate data normalization and filtering technique allowed high specificity of the tag-microarray data. The determination of the

transposon insertion sites for more than 5000 mutants provided the possibility to influence the composition of the tested mutant sets and to immediately evaluate the results of the STM experiments. The pilot validation experiments performed using one set of transposon tagged mutants showed biologically reasonable results.

D.2 *S. MELILOTI* MUTANTS, KNOWN TO BE IMPAIRED IN SYMBIOSIS, WERE DETECTED AS ATTENUATED IN STM EXPERIMENT

The two sets of mutants, tested *in planta*, were generated in a way that a number of mutants in genes with known role in symbiosis were included in order to ensure the reliability of the obtained results. For some of these genes, two or three mutants were included in the sets. In fact, 27 from 67 mutants that were detected as attenuated were expected to be impaired in competitiveness in *S. meliloti*. These include mutants containing a transposon insertion in the following genes:

***exo* (exopolysaccharide synthesis) genes**

6 from generally 9 *exo* mutants included in the test sets were detected as attenuated. The attenuated mutants included those with a transposon insertion in *exoA*, *exoB*, *exoM*, *exoO*, *exoU* and *exoZ* genes. The $mean_w$ values of *exoK, exoN2* and *exoP2* mutants were -0.33, -0.55 and -0.15 respectively by a threshold of -0.7 and were therefore filtered out. *exoA*, *exoB, exoO exoM* and *exoU* mutants do not produce succinoglycan and were shown to form empty Fix^- nodules (Becker *et al.*, 1993b; Becker *et al.*, 1993c; Skorupska *et al.*, 2006). *exoZ* mutant was shown to have a reduced efficiency in the initiation and elongation of infection threads (Cheng and Walker, 1998), which can lead to the impaired competitiveness. In contrast, *exoK* mutant could induce pink nodules normally (Becker *et al.*, 1993a). Symbiotic phenotypes of *exoP2* and *exoN2 S. meliloti* mutants have not been reported till now.

***nod* (nod factor synthesis and transport) genes**

Two *nod* mutants were included in the test sets. Both *nodD3* and *nodJ* mutants were attenuated in the STM experiment. NodD3 protein is involved in the regulation of the nod factor production, whereas NodJ is responsible for the transport of nod factors. Both *S. meliloti nodD3* (Barran *et al.*, 2002; Honma and Ausubel, 1987) and *nodJ* (Barran *et al.*, 2002) mutants exhibit delayed nodulation on alfalfa.

***fix* (nitrogen fixation) genes**

From 13 *fix* mutants in the sets, 4 could be identified as attenuated: all three *fixL* mutants present in the sets (one *fixL* mutant in the set 1 and 2 *fixL* mutants in the set 2) and a *fixB*

mutant. Attenuated phenotype of these mutants was expected since *fixL* (Batut *et al.*, 1985; Vasse *et al.*, 1990) and *fixB* (Hirsch and Smith, 1987) *S. meliloti* mutants have a Fix$^-$ phenotype, that can lead to the degradation of the symbiosomes containing these mutants (Mylona *et al.*, 1995).

Two *fixN1*mutants and mutants with a transposon insertion in *fixM*, *fixN2*, *fixO3*, *fixP1*, *fixP2* and in *fixT3* genes were not attenuated. A *fixI2* mutant was detected as highly represented under symbiotic conditions. For *fixN1*, *fixN2*, *fixO3*, *fixP1*, *fixP2* mutants, the non-attenuated phenotype originates from the duplication of the *fixNOQP* region in the *S. meliloti* genome, so that a single mutation in one of these regions does not influence symbiotic efficiency. The disruption of *fixM* was reported not to affect nodulation or nitrogen fixation efficiency (Cosseau *et al.*, 2002). Symbiotic phenotypes of *fixT3* and *fixI2 S. meliloti* mutants have not been reported till now.

***nif* (nitrogen fixation) genes**

All three *nif* mutants (with transposon insertion in *nifB*, *nifK* and *nifN* genes), included in the sets, were attenuated. *nifN* mutants were shown to induce nodule formation that was similar to that induced by the wild type strains both in time of appearance of nodules and in distribution of nodules. However, no acetylene reduction was detected 3, 4, and 5 weeks after inoculation. After 4 weeks plants showed symptoms of nitrogen starvation (Aguilar *et al.*, 1985; Aguilar *et al.*, 1987). *nifB* (Buikema *et al.*, 1987) and *nifK* (Hirsch *et al.*, 1983; Zimmerman *et al.*, 1983) mutants are unable to symbiotically fix nitrogen.

syrM encodes a transcriptional regulator that activates expression of *nod* genes through NodD3. The *syrM* mutation was shown to cause a delayed nodulation phenotype (Dusha *et al.*, 1999).

dctB encodes a C4-dicarboxylate transport sensor protein. The *dctB* mutant was shown to induce nodules on alfalfa with about half the nitrogen fixation activity of the wild type (Engelke *et al.*, 1987).

dme encodes NAD-malic enzyme oxidoreductase protein. The *dme* mutants induce non-fixing nodules (Driscoll and Finan, 1997).

bioN encodes the permease component of biotin ABC transporter. Mutation in *bioN* eliminated growth on alfalfa roots and reduced bacterial capacity to maintain normal intracellular levels of biotin (Entcheva *et al.*, 2002). The deficiency in root colonization leads to the weak competitiveness of the mutant strains (Guillen-Navarro *et al.*, 2005).

rpoN codes for RNA polymerase sigma-54 factor. The *rpoN* mutant was shown to elicit Fix$^-$ root nodules on alfalfa (Ronson *et al.*, 1987). The product of ***lpsB*** gene - lipopolysaccharide core biosynthesis mannosyltransferase – is involved in the synthesis of LPS. *lpsB* mutants were shown to be slightly compromised for infection thread formation and fixed nitrogen at reduced levels

(Campbell *et al.*, 2002; Campbell *et al.*, 2003).

cycH encodes cytochrome C-type biogenesis transmembrane protein. The *cycH* mutant in *S. meliloti* strain AK631 was defective in symbiotic nitrogen fixation, "respiratory" nitrate reduction, and cytochrome *c* oxidase activity (Kereszt *et al.*, 1995).

The product of ***pckA*** gene – phosphoenolpyruvate carboxykinase – catalyzes the the decarboxylation and phosphorylation of oxaloacetate to phosphoenolpyruvate. This reaction is the first step in the gluconeogenic pathway. The *pckA* mutant was shown to elicit nodules with reduced nitrogen-fixing activity. It had acetylene-reducing activities of approximately 60% of plants inoculated with the wild type bacteria (Finan *et al.*, 1991; Osteras *et al.*, 1995).

phaD1 encodes a pH adaptation potassium efflux system transmembrane protein. It was shown that the *S. meliloti* mutants defective for the Pha K^+ efflux system are affected in the invasion of alfalfa root nodules, K^+ sensitive and unable to adapt to alkaline pH in the presence of K^+ (Putnoky *et al.*, 1998).

ndvA codes for beta-glucan export ATP-binding protein involved in cyclic β-(1,2)-glucan production. On alfalfa, the *ndvA S. meliloti* mutant formed a small number of early infection threads, roughly 3% to 15% of the level seen with the wild type. Infection threads aborted and were not observed in mature ndv-induced pseudonodules. The mutant delayed formation of numerous small white nodules that are distributed throughout the root system. These nodules were not invaded by the mutant bacteria and consequently did not fix nitrogen (Dickstein *et al.*, 1988; Stanfield *et al.*, 1988).

ilvC codes for ketol-acid reductoisomerase protein, involved in isoleucine / valine synthesis. Both *ilvC* mutants in two different sets were detected as attenuated. The *ilvC* mutant was shown to be strongly impaired in nodulation. It did not induce nodules or induced small ineffective nodules in about 10% of inoculated *M. sativa* plants. The mutant also induced abnormal root hair deformations: alfalfa plants showed abundant hypertrophied root hairs with numerous branchings and swellings and, only exceptionally, curled root hairs (Aguilar and Grasso, 1991).

The fact, that many mutants that were expected to be attenuated, could be detected in the STM experiment, supports the reliability of the two-tag microarray detection system. Importantly, the mutants that had a defect on late stages of symbiosis, like some *fix* or *nif* mutants, or mutants with only a slight symbiotic deficiency, like *lpsB* or *dctB* mutants, could also be detected by this method as weakly competitive.

D.3 INDIVIDUAL COMPETITIVENESS TEST CONFIRMED THE STM EXPERIMENT RESULTS FOR THE MOST OF CHECKED MUTANTS.

The main interest of the STM experiment *in planta* was the identification of the genes whose attenuated competitiveness phenotype was not observed before. It was important to prove the attenuated phenotypes of such mutants by another method. Therefore, individual competitiveness test was performed for 25 signature tagged mutants with no previously known fixation or nodulation defect. For 22 mutants, the attenuated phenotype could be confirmed (Table C6). The mutants that had a defect in competitiveness, but no other obvious nodulation or nitrogen fixation defect, had a transposon insertion the following genes:

thiD gene encodes probable phoshomethylpyrimidine kinase, involved in thiamine synthesis and ***thiO*** encodes putative thiamine biosynthesis oxidoreductase protein, probably involved in thiamine synthesis. Both genes are situated on pSymB, but belong to different clusters. *S. meliloti* is able to synthesize thiamine, but still depends on external supply of this vitamin. It was shown that colonization of roots can be limited by availability of thiamine (Streit *et al.*, 1996). On the other hand, it was shown that in *R. etli* mutant strain, constitutively expressing *thiCOGE* gene cluster, the *fixNOQP* operon was constitutively expressed (Miranda-Rios *et al.*, 1997). This mutant strain also showed an increased capacity to fix nitrogen during symbiosis (Soberon *et al.*, 1990). It was suggested that a metabolite related to the purine-thiamine biosynthetic pathway, probably AICAR (5-amino-4-imidazolecarboxamide), is a negative effector for the *fixNOQP* expression. The constitutive expression of *thiC* could lower the concentration of several metabolites of the purine biosynthetic pathway, particularly AICAR, and, as a result, block the negative regulation of *fixNOQP* (Miranda-Rios *et al.*, 1997). Mutations in thiamine-synthesis genes could therefore have an effect on symbiotic competitiveness through a slight deficit in the thiamine, and/or through the down-regulation of nitrogen fixation genes.

ndiA1 was shown to be induced by carbon and nitrogen deprivation, osmotic stress, oxygen limitation and during entry into stationary phase. Its expression is also positively regulated by the products of *tspO* and *fixL* genes (Davey and de Bruijn, 2000). Mutation in *ndiB* gene, situated downstream of *ndiA1* and *ndiA2*, did not cause any defect in nodulation efficiency and nitrogen fixation (symbiotic competitiveness was not tested) (Davey and de Bruijn, 2000). Here, *ndiA1* mutant was weakly impaired in competitiveness in STM experiment as well as in the individual test (25.46% nodule occupation when co-inoculated with wild type). Since the function of *ndiA1*, and of the downstream genes *ndiA2* and *ndiB* is still unknown, it is difficult to speculate on the nature of attenuated competitiveness phenotype of the *ndiA1* mutant. The putative explanation

for this phenotype could be the impaired ability of the *ndiA1* mutant to adapt to the stress conditions faced by rhizobial cells during infection and nodule formation.

feuQ gene codes for sensor histidine kinase. In *R. leguminosarum*, a mutation in *feuQ* led to the loss of the high affinity iron uptake system, although siderophores were still produced (Yeoman *et al.*, 1997). Iron is particularly important in symbiosis, since many enzymes and proteins involved in nitrogen fixation, including nitrogenase, ferredoxin, cytochromes and leghaemoglobin are iron-containing. Interestingly, *feuPQ* operon is situated closely upstream of the cytochrome maturation operon *cycHJKL* in both *R. leguminosarum* (Yeoman *et al.*, 1997) and *S. meliloti*. However, a mutation in *feuQ* did not affect cytochrome *c* function or expression of the *cycHJKL* operon in *R. leguminosarum* (Yeoman *et al.*, 1997). Still, the expression of the *feuPQ* and *cycHJKL* clusters could be in some way connected, since the *cycK* mutant, similarly to *feuQ* mutant, had a defect in high affinity iron acquisition (Nadler *et al.*, 1990; Yeoman *et al.*, 1997). The defect of *feuQ* transposon mutant in symbiotic competitiveness, observed here, could be probably explained by the internal deficit of this mutant in iron that could lead to slower growth and reduced rates of nitrogen fixation.

pstA and ***pstC*** encode for the parts of the high-affinity phosphate transporter system PstSCAB (Yuan *et al.*, 2006a). *pstA* and *pstC* belong to the same gene cluster, where *pstC* is situated downstream of *pstS* gene and upstream of *pstA*. Downstream of *pstA* gene, genes *pstB*, *phoU* and *phoB* are situated. It was recently shown that *pstSCAB* genes and the *phoUB* genes are transcribed from a single promoter (Yuan *et al.*, 2006a). Since the signature tagged mutants generated by the mTn*5*-STM transposon contain a transcription terminator, the transcription of all the downstream genes in the operon is disrupted. It is especially important in the case of *phoB* that encodes a transcriptional regulator. PhoB protein, which is activated by the sensor kinase PhoR in conditions of low external phosphate concentration, regulates the transcription of a large spectrum of genes (Pho regulon). In *S. meliloti*, PhoB was shown to induce the expression of *pstSCAB* and *phoCDET* operons which encode ABC-type high affinity phosphate transport systems (Bardin *et al.*, 1996; Bardin and Finan, 1998), *exp* operons which encode enzymes for EPS II biosynthesis (Rüberg *et al.*, 1999), operon *phnGHIJK* that codes for phophonate degradation enzymes (Parker *et al.*, 1999) and *pta-ackA* operon involved in acetyl phosphate metabolism (Summers *et al.*, 1999), and to repress the espression of *orfA-pit* operon that encodes a low affinity phosphate transport system (Bardin *et al.*, 1998). In recent studies (Krol, 2004; Yuan *et al.*, 2006b), many more putative Pho regulon members were identified, that led to a suggestion that the Pho regulon overlaps and interacts with several other control circuits, such as the oxidative stress response and iron homeostasis (Yuan *et al.*, 2006b).

Furthermore, PstSCAB transporter system can have an influence on the PhoB activity

through the inactivation of PhoR in presence of high external phosphate. In *S. meliloti* Rm1021 (but not in Rm2011) a single nucleotide deletion led to a frameshift in the *pstC* gene, and as a result, to the truncation of the PctC protein. This resulted in the deregulation of the phosphate starvation response in Rm1021 (Krol, 2004).

The attenuated symbiotic competitiveness phenotype of the *pstA* and *pstC* transposon mutants originates, probably, from a combinational effect of PstSCAB$^-$ and PhoB$^-$ phenotypes. In order to identify the impact of single mutations on the symbiotic competitiveness, a comparison between non-polar *pstA, pstC* and *phoB* mutants needs to be performed.

livM encodes a permease protein of high-affinity branched-chain amino acid ABC transporter that belongs to hydrophobic amino acid transporter (HAAT) family. *S. meliloti livM* gene shows a 69% identity to *braE* from *Rhizobium leguminosarum* bv. *trifolii*. In *R. leguminosarum*, amino acid uptake is dominated by two ABC transporters, the general amino acid permease (Aap) and the branched-chain amino acid transporter permease (Bra_{Rl}) (Hosie *et al.*, 2002). The double *aap-bra* mutant was shown to retain a nitrogenase activity at rates per bacteroid that were equal or exceeded the wild type rates, but the plants, nodulated by this mutant, were severely nitrogen starved (Lodwig *et al.*, 2003). To explain these results, it was suggested that in *R. leguminosarum* an amino acid cycle operates where an amino acid such as glutamate, or a derivative of it, is supplied by the plant to the bacteroid. The bacteroid uses the amino acid to transaminate oxaloacetate or pyruvate to produce aspartate or alanine, respectively, and either or both of these amino acids are secreted (Lodwig *et al.*, 2003; Prell and Poole, 2006). The mutation on both of the amino-acid transporters would disrupt this cycle and result in nitrogen starvation of the nodulated plant.

S. meliloti genome, in contrast to *R. leguminosarum* genome, contains several branched-chain amino acid transporters. Particularly, gene *SMa0677*, encoding a periplasmic solute-binding protein of such a transporter, showed induced expression in bacteroids (Becker *et al.*, 2004). The expression of another branched-chain amino acid transporter operon, that includes genes *SMb20602* to *SMb20605* and *SMb21707*, was induced during nitrogen deprivation (Milcamps *et al.*, 1998). The *livM* gene, in contrary, was repressed in bacteroids in comparison to growth in free-living conditions (Becker *et al.*, 2004). The possible explanation for the attenuated competitiveness phenotype of the *livM* mutant could be that this gene is active during root colonization. The rhizobial ABC transporters are very important for the survival in the rhizosphere, since the plant exudates contain many different compounds that can be catabolised by rhizobia. Inability to transport some of these compounds can negatively influence competitiveness of *S. meliloti* cells carrying the *livM* mutation.

lexA codes for a repressor transcription regulator. LexA negatively regulates the SOS genes under normal growth conditions by binding to a 20-bp consensus sequence (the SOS box) in the

operator region of these genes. Activation of the SOS genes occurs after DNA damage by the accumulation of single-stranded DNA regions. The RecA protein, stimulated by single-stranded DNA, interacts with LexA causing an autocatalytic cleavage which disrupts the DNA-binding part of LexA, leading to derepression of the SOS regulon and eventually DNA repair (Brent and Ptashne, 1980). In rhizobia, the effect of *lexA* mutation has not been studied yet.

recN codes for a DNA repair protein. RecN participates in SOS response and is involved in recombination and double strand repair (Picksley *et al.*, 1984) *recN* is a part of SOS regulon (Rostas *et al.*, 1987) and is controlled by multiple SOS boxes (Schnarr *et al.*, 1991). Since the transposon insertion is situated only 131 nucleotides before the end of the gene, it could be possible that the truncated RecN is partly functional.

Downstream of *recN*, *ligA* gene is situated that encodes a DNA ligase protein. In view of the fact that the *ligA* mutation is lethal for the most of bacteria (Lavesa-Curto *et al.*, 2004), its transcription is probably not disrupted by the transposon insertion in *recN*.

It was shown that the *recN* mutation, similarly to *lexA* mutation, causes constitutive activation of the SOS regulon (Chua *et al.*, 1993; Dunman *et al.*, 2000). Since the *lexA* and *recN* mutants have similar symbiotic phenotypes, it could be possible that the constitutively active SOS regulon causes a disadvantage in the establishment of symbiosis. It is known that the SOS-response is error-prone and that in the cells expressing SOS genes, the rate of mutagenesis is enhanced (Smith and Walker, 1998). These high mutagenesis rates, particularly, could be detrimental for rhizobia in symbiosis.

lepA encodes GTP-binding membrane protein that belongs to the GTP-binding elongation factor family. This family includes three factors involved in the translation process (EF-G, EF-Tu and IF2) two GTPases of the FtsY family (FtsY and Ffh), Era, ThdF/TrmE, EngA, YchF, Obg and LepA (Caldon *et al.*, 2001). Similarity of LepA to protein synthesis elongation-factor GTPases indicates its involvement in transcription. It was suggested that LepA can regulate ribosome function (Caldon *et al.*, 2001; March, 1992). With a help of two-hybrid screens in *Bacillus subtilis*, an interaction between HolA (delta subunit of DNA polymerase III holoenzyme) and LepA protein was detected, suggesting that LepA might act to coordinate DNA replication and protein synthesis (Noirot-Gros *et al.*, 2002). The defect of the *lepA* mutant in symbiotic competitiveness indicates that an unknown symbiosis-relevant intracellular signaling mechanism might exist that involves the LepA GTPase protein.

clpA encodes a the regulatory subunit of ATP-dependent Clp protease. ClpA is a member of the Hsp100/Clp family of molecular chaperones. It associates with the ClpP serine-type protease protein, forming the active holoenzyme. Usually, Clp proteases are present at low levels during normal growth, but are inducible during many types of stress. Their functions include

degradation of damaged proteins that occur in stress conditions and control of diverse cell activities. In *Bacillus subtilis*, ClpP is involved in sporulation, cell competence, stress tolerance and regulation of gene expression (Porankiewicz *et al.*, 1999).

Interestingly, in *Bradyrhizobium japonicum* USDA 110, a mutation in the clpA gene caused changes in EPS synthesis. The mutant strain produced approximately twofold higher levels of low-molecular-weight EPS when compared to the wild type USDA 110 strain (Louch and Miller, 2001). An effect of *clpA* mutation on EPS synthesis might occur in *S. meliloti* as well. In this case, changes in EPS could influence the symbiotic competitiveness of the *clpA* mutant.

SMb20509 codes for a putative transcriptional regulator from a helix-turn-helix (HTH) -type GntR family. This family was named after the *Bacillus subtilis* repressor of the gluconate operon (Haydon and Guest, 1991). The members of this regulator family contain a well-conserved DNA-binding HTH domain at the N terminus, and a variable effector-binding or oligomerisation domain at the C terminus. (Rigali *et al.*, 2002).

Genes ***SMa0414***, ***SMc00644***, ***SMc01929***, ***SMc02123***, ***SMc00828*** and ***SMc02226*** encode hypothetical proteins. *SMa0414* has an ATP dependent DNA ligase domain and shows similarity to the thermostable DNA ligases. *SMc01929* shows similarity to metallo-beta-lactamases, and *SMc00828* has a similarity to the diadenosine and diphosphoinositol polyphosphate phosphohydrolases. Genes *SMc00644* and *SMc02226* do not show significant homologies to known genes. In the gene *SMc02123*, which also shows no homologies to the known genes, the transposon insertion is situated nearly at the end of the gene. The downstream gene *fpr*, which is probably co-transcribed with *SMc02123*, codes for a probable ferredoxin--NADP reductase.

D.4 FOUR MUTANTS, IMPAIRED IN SYMBIOTIC COMPETITIVENESS, EXHIBIT ALSO OTHER SYMBIOTIC PHENOTYPES

While the main goal of the signature tagged mutagenesis experiment *in planta* was to identify genes important for symbiotic competitiveness, it was also interesting to test if the attenuated mutants exhibit other symbiotic phenotypes. Four mutants were identified that showed a strong defect in symbiosis. These mutants were studied in more detail using GUS-staining of nodules, monitoring of nodulation dynamics, and acetylene reduction assay. In this chapter, the phenotypes of these four mutants will be discussed.

ccmC mutant was the only mutant in the sets which showed a Fix$^-$ phenotype. *ccmC* (*cycZ*) gene codes for cytochrome c-type biogenesis protein. This gene belongs to the cluster *ccmABCDG*, where *ccmA*, *ccmB*, *ccmC* and *ccmD* encode parts of the heme transporter, and *ccmG*

encodes periplasmic thiol:disulfide oxidoreductase. Cytochrome biogenesis is very important in symbiosis since the respiratory pathways that involve cytochromes are required for normal nitrogen fixation. Moreover, the *ccmC* mutant, studied here, has even more severe symbiotic defect than the *fix* mutants (Hirsch and Smith, 1987), since it induces empty nodules devoid of bacteroids. In a recent work (Capela *et al.*, 2006) *ccmA* gene was detected as induced during early stages of the *S. meliloti*-alfalfa symbiosis. The *ccmA* mutant, constructed by Capela *et al.*, showed a Fix$^-$ phenotype.

The symbiotic phenotypes of *ccm* mutants are similar to those of *cyc* mutants. It was shown that the *cycHJKL* gene cluster is necessary for the biogenesis of all cellular *c*-type cytochromes in *S. meliloti* and respiratory nitrate reduction *ex planta* (Kereszt *et al.*, 1995). In *Rhizobium phaseoli*, cytochrome *c* – deficient mutant induced empty nodules on *Phaseolus vulgaris* (Soberon *et al.*, 1993). A polar *cycH* mutant of *Rhizobium etli*, in which expression of downstream *cycJKL* genes was disrupted, also induced empty nodules on *Phaseolus vulgaris* (Tabche *et al.*, 1998).

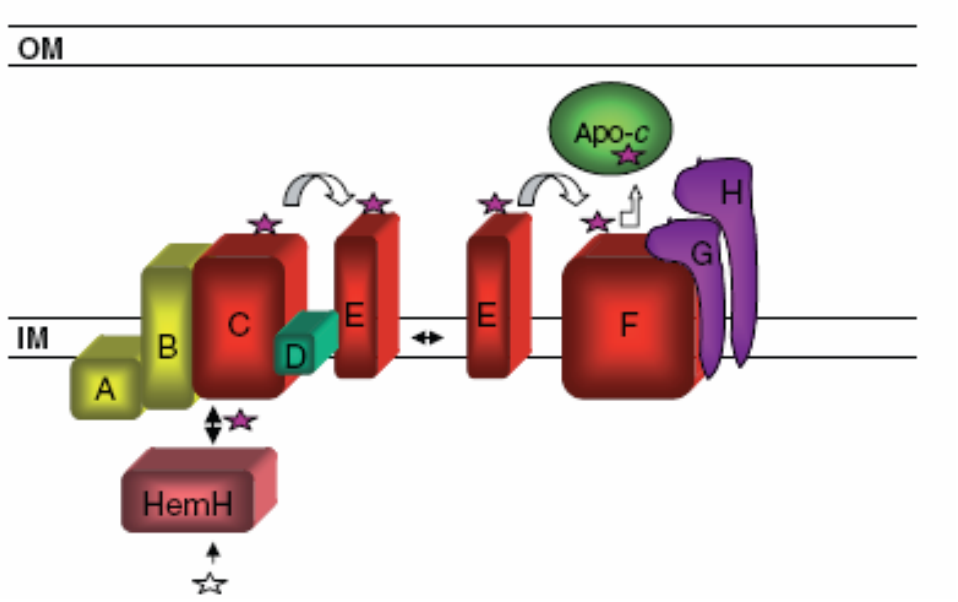


Fig. D1. Representation of the different proteins involved in the type I system for the maturation of *c*-type cytochromes (Cianciotto *et al.*, 2005). CcmA (A) and CcmB (B) form an ABC transporter for the transport of yet-to-be-discovered molecule. In the cytoplasm, but associated with membrane, the enzyme ferrochelatase (HemH) performs the last step of haem biosynthesis by converting protoporphyrin IX (white star) to haem (red star). Once haem reaches the periplasm, perhaps via CcmC (C), it is transferred to the haem chaperone CcmE (E). The protein CcmD (D) has been shown to facilitate the interaction between CcmC and CcmE. Haem is passed from CcmE to CcmF (F), which, in association with CcmG (G) and CcmH (H), forms a haem-lyase complex. CcmG and CcmH maintain the apocytochrome (apo-*c*) in a reduced state.

TABLE D1. Homology between *ccm* and *cyc* genes.

	genes[a]								
ccm	*ccmA*	*ccmB*	*ccmC*	*ccmD*	*ccmE*	*ccmF*	*ccmG*	*ccmH*	*ccmI*
cyc	*cycV*	*cycW*	*cycZ*	*cycX*	*cycJ*	*cycK*	*cycY*	*cycL*	*cycH*

[a]Genes, annotated in *S. meliloti*, are written in red font.

The explanation for the similarity in phenotypes of *ccm* and *cyc* mutants could be that both gene clusters encode different parts of one cytochrome maturation system, but were annotated differently (Table D1). Fig. D1 illustrates the known structure of the Ccm cytochrome *c* maturation system that could be also true for *S. meliloti*.

The severe symbiostic defect of *ccm* and *cyc* mutants, together with the fact that the expression of *ccmA* gene was induced at the early stage of symbiosis, allows suggestion that *ccm* and *cys* genes are important not only for nitrogen fixation, but also for the successful infection.

asnO mutant displayed a severe defect in nodulation efficiency. Nodules, formed by this mutant, could fix nitrogen, but their morphology differed from those of wild type induced nodules. The broad proximal part of the nodules and bifurcated morphology indicate the difficulties in the infection. When the plants were inoculated with *asnO* mutant only, it could form nodules at less than 5% of the wild type nodulation rate. Interestingly, when the plants were inoculated with the mixture of wild type Rm2011 and *asnO* mutant, the average percentage of *asnO* mutant in nodules reached 18%. This indicates that the presence of the wild type Rm2011 might facilitate infection of plants by the *asnO* mutant.

asnO gene codes for a protein homologous to glutamine-dependent asparagine synthetases, but it is not absolutely required for the asparagine synthesis in *S. meliloti* GMI211 background (Berges *et al.*, 2001). Here, the signature tagged *asnO* mutant did not show an attenuated growth in Vincent minimal medium; it can be therefore suggested that *asnO* mutation does not cause auxotrophy in Rm2011 background as well.

In the *S. meliloti*, constitutive expression of *fixT* gene inhibits expression of FixLJ dependent genes. Berges *et al.* (2001) have shown that a Tn*5* insertion in the *asnO* gene impairs this inhibition. Moreover, they demonstrated that *asnO* gene expression was induced in microoxic conditions. This data indicates that *asnO* gene has a regulatory function in symbiosis. However, it is not clear how the phenotype observed by Berges *et al.* is connected to the impaired infection phenotype of the signature tagged *asnO* mutant, since the activation of the *fix* genes does not take place at the infection step, but later in symbiosis.

CysG mutant showed a delayed nodulation phenotype and low occupation of plant cells by bacteroids, but high rates of nitrogenase activity. This mutant induced pink and white nodules on different plants. Moreover, *cysG* mutant was impaired in growth in Vincent minimal medium,

probably, due to cysteine auxotrophy.

cysG gene encodes for siroheme synthase (uroporphyrinogen III methylase), which is necessary for siroheme synthesis. Siroheme is the prosthetic group of many nitrite and sulfite reductases that function in the conversion of the highly oxidized forms of nitrogen and sulfur to the fully reduced forms (NH_4^+ and S^{2-}). Unlike other *cys* genes, *cysG* is transcribed together with nitrite reductase gene, and not with genes for cysteine synthesis. A *cysG* mutant was obtained in *R. etli* by Tn*5* mutagenesis (Tate *et al.*, 1997). This mutant was identified as a cysteine auxotroph based on its inability to utilize sulfate as the sole sulfur source. Also, *cysG* mutant could not grow with nitrate as the sole nitrogen source, because *cysG* gene product is required for the activity of the nitrite reductase complex. Furthermore, this mutant had a Nod^+ Fix^+ phenotype, but was impaired in competitiveness 5-fold. Poor competitiveness phenotype could be restored by addition of an organic sulfur source. Acetylene reduction activity of the *R. etli cysG* mutant-induced nodules was the same as of the wild type-induced nodules (Tate *et al.*, 1997), whereas the *S. meliloti cysG* mutant studied here showed higher acetylene reduction rates than the wild type. It should be noted, however, that the difference between the *S. meliloti* and *R. etli* acetylene reduction data could originate from the utilization of different approaches in acetylene reduction measurements: in *R. etli*, the quantity of reduced acetylene per gram of nodules was estimated, whereas here, the quantity of reduced acetylene per plant/nodule was estimated.

In *S. meliloti*, the symbiotic defect of the *cysG* mutant seems to be generally more acute than in case of *R. etli cysG* mutant. Particularly, *S. meliloti cysG* mutant is stronger impaired in the symbiotic competitiveness. Here, in 30 analyzed nodules from the plants inoculated by the 1:1 mixture of wild type Rm2011 and *cysG* mutant, only trace amounts of the mutant was detected, indicating that all of the tested nodules were occupied by the wild type. Furthermore, shape and histological organization of nodules induced by *R. etli cysG* mutant were the same as of wild type-induced nodules. In case of *S. meliloti cysG* mutant, both of these characteristics were changed in comparison to the wild type Rm2011. The symbiotic defect of the *cysG* mutant is probably caused by the pleyotropic character of the *cysG* mutation, which influences several important biochemical pathways.

metA mutant had a delay in nodulation, but formed pink nodules that could fix nitrogen at the same rate as the wild type. Besides, *metA* mutant was very strongly impaired in symbiotic competitiveness: when co-inoculated with the wild type Rm2011, the mutant practically did not enter the nodules at all. GUS-staining of the nodules induced by the *metA* mutants has shown that *metA* gene is mainly expressed in the proximal part of the nodules.

metA gene codes for homoserine O-succinyltransferase which catalyses the first step in methionine biosynthesis in gram-negative bacteria. *E. coli met*A mutant was shown to be a

methionine auxotroph (Michaeli *et al.*, 1981). Interestingly, *S. meliloti metA* gene contains a riboswitch. Riboswitches are structured elements typically found in the 5´ untranslated regions of mRNAs. Riboswitches can directly bind a ligand, and in this bound state they induce or repress the gene they are encoded in. Thus, an mRNA that contains a riboswitch is directly involved in regulating its own activity, depending on the presence or absence of its target molecule. Riboswitches regulate several metabolic pathways including the biosynthesis of vitamins (e.g. riboflavin, thiamin and cobalamin) and the metabolism of methionine, lysine and purines (Tucker and Breaker, 2005; Vitreschak *et al.*, 2004).

Mutation that cause methionine auxotrophy were shown to result in a Fix$^-$ phenotype in *S. meliloti* 104A14 (Kerppola and Kahn, 1988) and in *Rhizobium fredii* HH303 (Kim *et al.*, 1988). Unfortunately, in these studies the genes, which were disrupted in the auxotrophic mutants, were not deduced. Since the *metA* mutant studied here, does not have a Fix$^-$ phenotype, it is probably not strictly auxotrophic for methionine. On the other side, the impaired competitiveness phenotype indicates that the *metA* mutant might suffer from the methionine unavailability during growth in the infection threads.

It is surprising that the expression of the *gusA* reporter gene in the *metA* mutant was detected mainly in mature and senescing bacteroids, since it was shown that the protein synthesis declines in the mature nodules (Bisseling *et al.*, 1979). This could indicate an unknown function of the *metA* gene in the senescing bacteroids.

D.5 CONCLUSIONS AND OUTLOOK

In this study, signature-tagged mutagenesis (STM) strategy is presented which for the first time was applied to study rhizobium-legume symbiosis. The STM approach was modified in order to enhance the efficiency of screening. A library of 412 different double-tagged transposons was created using a novel set of signature tags. Also, a new two-channel microarray hybridization approach was designed and used to identify and quantify individual mutants in the pool.

A number of studies demonstrated a broad host range for transposition of the mTn*5* transposon (Saenz and Dehio, 2005) including, e.g. *S. typhimurium*, *E. coli*, *K. pneumoniae*, *V. cholerae*, *P. mirabilis*, *B. melitensis*, *Y. pestis*, and *C. rodentium*. This broad application spectrum in combination with the large number of signature tags and the tag-specific microarray makes the mTn*5*-STM transposon set a powerful and easy-to-use tool that can be applied to a large spectrum of bacteria.

An extensive library of transposon mutants containing more than 12000 clones was created applying the set of tagged transposons. The transposon insertion sites were determined for 42 %

of the mutants in this library. As a result, 44 % coverage of all predicted protein-coding genes by mapped transposon insertions was achieved. Analysis of the transposon library suggests that insertion sites of the mTn*5*-STM transposons are random and show no hotspots.

Pilot experiments were performed to verify the modified STM approach. The statistical processing of the tag microarray data comprising normalization and clustering allowed identification of clusters of mutants that showed a similar growth pattern throughout different growth conditions. As expected, clones, carrying a similar kind of mutation, were grouped into the same cluster.

Two sets of mutants, each one containing 378 mutants, were screened *in planta* in order to identify mutants impaired in symbiotic competitiveness. 67 mutants were detected as attenuated in symbiotic conditions. Many of these mutants were already known to have a Nod^-, Fix^-, or low competitiveness phenotype. For 25 mutants, whose symbiotic defect was not observed in *S. meliloti* before, the symbiotic competitiveness was checked by another method, and for 22 mutants the attenuated competitiveness phenotype could be confirmed. Furthermore, four of the checked mutants showed additional symbiotic defects: Fix^- phenotype (*ccmC* mutant), low efficiency of nodulation (*asnO* mutant) and delayed nodulation (*cysG* and *metA* mutants). These phenotypes were verified by histochemical GUS-staining of nodules, acetylene reduction assay and by monitoring the dynamics of nodule formation.

The signature tagged mutagenesis system, applied in this work, offers many possibilities for the following research. In this study, only a small part of the library of signature tagged mutants has been screened. The screening of the rest of the library can reveal many new mutants with reduced competitiveness. Furthermore, the genes, identified in this study as important for symbiotic competitiveness, need a closer inspection. This is especially important in case of genes annotated as hypothetical proteins. At last, the mutants, which were identified as highly competitive in the STM experiment, also have to be studied in detail.

E. METHODS AND MATERIALS

E.1 MATERIALS

E.1.1 Sources for chemicals and kits

Amersham	dNTPs
DUCHEFA	X-Gluc (5-Bromo-4-chloro-3-indolyl-ß-D-glucuronide)
Fermentas	Restriction enzymes, Shrimp alkaline phosphatase
Fluka	NaFe-EDTA
Invitrogen	Select Agar
Macherey-Nagel	NucleoSpin Plasmid kit, NucleoSpin Tissue kit, NucleoSpin Extract II PCR clean-up kit
Merck	$CaCl_2$, $CaCl_2 \times 2\ H_2O$, $CoCl_2 \times 6\ H_2O$, $CuSO_4 \times 5\ H_2O$, EDTA, $FeCl_3 \times 6\ H_2O$, Fe-Citrate $\times 3\ H_2O$, H_3BO_3, HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), K-Acetate, K_2HPO_4, KH_2PO_4, KI, KOH, K_2SO_4, Mg-Acetate, $MgCl_2 \times 6\ H_2O$, $MgSO_4 \times 7\ H_2O$, $MnSO_4 \times 4\ H_2O$, Na-Acetate, Na_2-citrate $\times 2\ H_2O$, Na_3-Citrate, NaCl, NaOH, Na_2-succinate, NH_4Cl, NH_4NO_3, $(NH_4)_2SO_4$, $NaMoO_4 \times 2\ H_2O$, $ZnSO_4 \times 7\ H_2O$
Oxoid	Yeast extract, Agar Nr. 1, Tryptone
Peqlab	Taq-polymerase, agarose
Qiagen	DNeasy Plant Mini kit
QMT	4 × Blocking solution
Roche	T4 DNA ligase, exo- Klenow fragment, DNA marker X, DNA marker IX, DIG EasyHyb solution
Roth	Ethanol, acetic acid, sulfuric acid, orthophosphoric acid, hydrochloric acid
SERVA	All antibiotics, bromphenol blue, ethidium bromide, Triton-X-100, polyethylene glycol (PEG), RNase A, glucose, glycerol, sodium dodecyl sulfate (SDS)
Sigma-Aldrich	Biotin, Bovine serum albumin (BSA), $K_3Fe(CN)_6$, $K_4Fe(CN)_6 \times 3H_2O$, Tris-HCl, Dithiothreitol (DTT)

E.1.2 Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics[a]	Source or reference
S. meliloti		
Rm2011	wild type, Sm^r	J. Dénarié, France
E. coli		
DH5α	F$^-$ *sup*E44 Δ*lac*U169 (φ80*lacZ*ΔM15) *hsdR*17 *rec*A1 *end*A17 *gyr*A96 *thi*-1 *rel*A1	(Sambrook *et al.*, 1989)
S17-1	MM294, RP4-2-Tc::Mu-Km::Tn7 chromosomally integrated	(Simon *et al.*, 1983)
Plasmids		
pBC KS (-)	Cloning vector; Cam^r; 3.4kb	Stratagene
pG18Mob2	pK18Mob derivative, Gm^r, 2.8kb	(Kirchner and Tauch, 2003)
pCRS530	Contains mTn*5*GNm, Km^r, Ap^r	(Reeve *et al.*, 1999)
pG18-STM	pG18Mob2 derivative containing modified mTn*5*GNm, Km^r, Gm^r	This study

[a] Cam^r, chloramphenicol resistance, Sm^r, streptomycin resistance, Km^r, kanamycin resistance, Gm^r, gentamicin resistance, Ap^r, ampicillin resistance

E.1.3 Plant material

In this study, *Medicago sativa* cv *Europe* was used (KWS, Einbeck, Germany).

E.1.4 Primers and other oligonucleotides

All the oligonucleotides listed here were synthesized by Qiagen Operon. Primers P1_Kpn, P2_Kpn, P3_Hind and P4_Hind were also synthezised with a Cy3 or Cy5 5´ modification. The list of cloned tags is given in the Appendix (Table G1)

Name	**Sequence**	**Comment**
Primers		
P1_Kpn	AAAGGACGTGGTTTACGGGGC	Amplification of K-tag
P2_Kpn	TATATGAATGCCGCCACCCCC	Amplification of K-tag
P3_Hind	ATTTTAACTCCCCTCCGCCGC	Amplification of H-tag
P4_Hind	TAGTCCTGGTGCATTGAGCCC	Amplification of H-tag
Qseq1	ATCTAGCCCGCCTAATGAGC	Sequencing of the transposon-genome junction site

M13 universal forward	TGTAAAACGACGGCCAGT	Amplification and sequencing at some steps of the construction of pG18-STM
M13 universal reverse	CAGGAAACAGCTATGACC	Amplification and sequencing at some steps of the construction of pG18-STM
Hind_mut_1	GCTCAGAGAGAGCTTCACGCT	Mutagenesis of *Hin*dIII site in *tnpA** gene
Hind_mut_2	AGCGTGAAGCTCTCTCTGAGC	Mutagenesis of *Hin*dIII site in *tnpA** gene
Sph_polym_end	ACATGCATGCGATCAGATCTTGATCCC	Mutagenesis of *Hin*dIII site in *tnpA** gene
Xba_polym_begin	ATAGTCTAGACACAAGTAGCGTCCTGAACG	Mutagenesis of *Hin*dIII site in *tnpA** gene
4-30_long	GATGCAGCCAAGCTAGCTTGCGG	Sequencing of the tags in transposon
Other oligonucleotides		
L1 up	AATTCGGCCGCCTAGGCCAAAGGACGTGGTT TACGGGGCACGTAGTTTAAGGAAGTACGGTA AGGTAC	Used for the construction of L123
L1 down	CTTACCGTACTTCCTTAAACTACGTGCCCCGTA AACCACGTCCTTTGGCCTAGGCGGCCG	Used for the construction of L123
L2 up	CGGGGGTGGCGGCATTCATATAGCTGCGTG ATTTCATTTTAACTCCCCTCCGCCGCA	Used for the construction of L123
L2 down	AGCTTGCGGCGGAGGGGAGTTAAAATGAAAT CACGCAGCTATATGAATGCCGCCACCCCCGGT AC	Used for the construction of L123
L3 up	AGCTTAGGTGGACCGTCGTAGAGCTAGTAGGG CTCAATGCACCAGGACTAGGCCGCCTAGGCCG	Used for the construction of L123
L3 down	AATTCGGCCTAGGCGGCCTAGTCCTGGTGCAT TGAGCCCTACTAGCTCTACGACGGTCCACCTA	Used for the construction of L123

E.1.5 Growth media and supplements

E.1.5.1 Media

All the media, if not indicated other, were dissolved by distilled water and autoclaved.

LB-medium (Luria-Bertrani Broth, (Sambrook *et al.*, 1989))

10 g/l	Tryptone
5 g/l	Yeast Extract
5 g/l	NaCl

LBG-medium (Luria-Bertrani Broth containing glucose)

10 g/l	Tryptone
5 g/l	Yeast extract

	5 g/l	NaCl
	1 g/l	Glucose

TY-medium (Beringer, 1974)

	5 g/l	Tryptone
	3 g/l	Yeast extract
	0.4 g/l	$CaCl_2$

2 × TY-medium

	10 g/l	Tryptone
	6 g/l	Yeast extract
	0.4 g/l	$CaCl_2$

Modified Vincent minimal medium (Becker *et al.*, 2004; Vincent, 1970)

Stock 1:	2.56 g/l	K_2HPO_4
	1.56 g/l	KH_2PO_4
	0.246 g/l	$MgSO_4 \times 7\ H_2O$
	1 g/l	NH_4Cl
	1.62 g/l	Na_2-succinate

Ingredients were dissolved in 800 ml Millipore and pH value was set to 7.0 by KOH or H_3PO_4. Than, Millipore water was added to fill 1 liter and the stock was autoclaved.

Stock 2:	67 g/l	$CaCl_2$

Autoclaved

Stock 3:	1 mg/ml	Biotine

Dissolved by 0.1 N NaOH and sterile filtered

Stock 4:	10 g/l	$FeCl_3 \times 6\ H_2O$

Sterile filtered

Stock 5:	3 g/l	H_3BO_3
	2.23 g/l	$MnSO_4 \times 4\ H_2O$
	0.287 g/l	$ZnSO_4 \times 7\ H_2O$

	0.125 g/l	$CuSO_4 \times 5\ H_2O$
	0.065 g/l	$CoCl_2 \times 6\ H_2O$
	0.12 g/l	$NaMoO_4 \times 2\ H_2O$
	Sterile filtered	

For full strength medium, 1 ml of each solution 2, 3, 4 and 5 were added one by one to 1 l of solution 1.

SOB Medium (Sambrook *et al.*, 1989), used for preparation of the competent *E. coli* cells

Stock A:	20 g	Tryptone
	5 g	Yeast extract
	0.5 g	NaCl
	dissolved in 980 ml H_2O	
Stock B:	18.65 g/l	KCl
Stock C:	406.6 g/l	$MgCl_2 \times 6\ H_2O$
Stock D:	493 g/l	$MgSO_4 \times 7\ H_2O$

10ml of the stock solution B were added to the stock solution A prior to autoclaving. Solutions C and D were autoclaved separately, and 5 ml of each were added to the autoclaved and cooled mixture of A and B solutions.

SOC medium

SOB medium supplemented with 20 ml 180 g/l solution of glucose

10 × Hogness Freezing Medium (Werner *et al.*, 1997)

Stock 1:	520 ml	Glycerol (87%)
	4.99 g	Na_2-citrate $\times 2\ H_2O$
	9 g	$(NH_4)_2SO_4$
	0.99 g	$MgSO_4 \times 7\ H_2O$
	Filled-up with water till 800ml and autoclaved	
Stock 2:	6.2 g	K_2HPO_4
	1.796 g	KH_2PO_4
	Filled-up with water till 200ml and autoclaved	

After autoclaving solutions were cooled down and mixed together.

Plant agar medium (Broughton and Dilworth, 1971)

Stock A:	294 g/l	$CaCl_2 \times 2\ H_2O$
Stock B	136.1 g/l	KH_2PO_4
Stock C:	6.7 g/l	Fe-Citrate × 3 H_2O
	123 g/l	$MgSO_4 \times 7\ H_2O$
	87 g/l	K_2SO_4
	0.338 g/l	$MnSO_4 \times H_2O$
Stock D:	0.247 g/l	H_3BO_3
	0.288 g/l	$ZnSO_4 \times 7\ H_2O$
	0.1 g/l	$CuSO_4 \times 5\ H_2O$
	0.056 g/l	$CoSO_4 \times 7\ H_2O$
	0.048 g/l	$Na_2MoO_4 \times 2\ H_2O$

The stock solutions A, B, C and D were dissolved and autoclaved. 15g of agar (Agar Nr. 1, Oxoid) were autoclaved in 1 l of Millipore water. For the final medium, 0.5 ml of each stock solution A, B, C and D were added to the liquid agar. The plates were poured using an inclined support.

Plant medium for aeroponics (Journet *et al.*, 2001)

Stock 1 (0.5 ml/l):	294 g/l	$CaCl_2$ x 2 H_2O
Stock 2 (0.5 ml/l):	123 g/l	$MgSO_4$ x 7 H_2O
Stock 3 (1.5 ml/l):	60.9 g/l	K_2SO_4
Stock 4 (2.5 ml/l):	7.35 g/l	NaFe-EDTA
Stock 5 (5 ml/l):	131.9 g/l	K_2HPO_4
	46.7 g/l	KH_2PO_4
Stock 6 (1 ml/l):	2 g/l	H_3BO_3
	1.8 g/l	$MnSO_4$ x H_2O
	0.2 g/l	$ZnSO_4$ x 7 H_2O
	0.08 g/l	$CuSO_4$ x 5 H_2O
	0.25 g/l	Na_2MoO_4 x 2 H_2O
	0.02 g/l	$CoCl_2$ x 6 H_2O
Stock 7 (5 ml/l):	80 g/l	NH_4NO_3

The stock solutions were added to the sterile Millipore water one by one, stirring (solutions 1 and 4 were added at the end). Stock solution 7 was left out in the nitrogen-free variant of the medium (E.2.2.3). After 15 min of stirring pH value was measured and set to pH 6.5 – 7.0 using 20 % H_2SO_4 or 10 % KOH. 17 – 21 liter of this medium was prepared for one aeroponic tank.

E.1.5.2 Supplements

Supplements for solid media

Select Agar (Invitrogen)

For the solid bacterial media, 16 g/l were added.

Agar Nr. 1 (Oxoid)

Nitrogen-depleted agar. For the solid plant medium, 15 g/l were added.

Antibiotics

Chloramphenicol (Cam)

For the selection of Chloramphenicol -resistant *E. coli* clones, 12.5 mg/l of antibiotic were added to solid media.

Gentamycin (Gm)

For the selection of Gentamycin-resistant *E. coli* clones, 8 – 10 mg/l of antibiotic were added to solid media.

Kanamycin (Km)

For the selection of Kanamycin-resistant *E. coli* clones 50 mg/l of antibiotic were added to solid media; 25 mg/l were of antibiotic added to liquid media to prevent the loss of plasmids.

Nalidixic acid (Nx)

For the selection of *S. meliloti* strains Rm2011 and Rm1021, 8 - 10 mg/l of antibiotic were added to solid media; 2.5 – 5 mg/l were added to liquid media.

Neomicin (Nm)

For the selection of *S. meliloti* carrying mTn*5* transposon, 120 mg/l of antibiotic were added to solid media.

Streptomycin (Sm)

For the selection of *S. meliloti* strains Rm2011 and Rm1021, 600 mg/l of antibiotic were added to solid media.

Other

X-Gluc

For the detection of *gusA* expression in *S. meliloti*, 20 mg/l of X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid) were added to solid media.

E.1.6 Buffers and solutions

E.1.6.1 DNA- and enzyme buffers

TE-buffer	10 mM	Tris-HCl
	1 mM	EDTA
	pH 7.5	
10 × TA- restriction buffer	660 mM	K-Acetate
	330 mM	Tris-HCl
	100 mM	Mg-Acetate
	5 mM	Dithiothreitol (DTT)
	1 mg/ml	Bovine serum albumin (BSA)
	pH 7.5	(adjust with acetic acid)
10 mM dNTP-mix	10 mM	dATP
	10 mM	dCTP
	10 mM	dGTP
	10 mM	dTTP
RNase A	20 mg/ml	

E.1.6.2 DNA electrophoresis buffers

1 × TA-buffer	40 mM	Tris-HCl
	10 mM	Na-Acetate
	1 mM	EDTA
	pH 7.8	(adjust with acetic acid)
BPB loading buffer	80 ml	Glycerol
	10 ml	TA-buffer
	2.5 g	Bromphenol blue

E.1.6.3 Solutions for microarray-hybridizations

20 × SSC	3 M	NaCl
	0.3 M	Na_3-Citrate
	pH 7.4	
Rinsing solution 1	250 ml	H_2O
	250 µl	Triton X100

Dissolved at 80 °C for 5 min; cooled down to the room temperature

Rinsing solution 2	500 ml	H_2O
	50 µl	32 % HCl
Rinsing solution 3	225 ml	H_2O
	25 ml	1 M KCl
Blocking solution	150 ml	H_2O
	40 µl	32 % HCl
	50 ml	4 × QMT Blocking solution

MilliQ/HCL mix is pre-warmed to 50°C. The 4 × QMT Blocking solution was added 5 min before use and the complete Blocking solution was pre-warmed to 50°C for at least 5 min.

Washing buffer 1	2 × SSC
	0.2 % SDS
	30 °C
Washing buffer 2	0.5 × SSC
	20 °C

E.1.6.4 Buffer for the preparation of competent *E. coli* cells

TB buffer	10 mM	HEPES
	15 mM	CaCl2 × 2 H_2O
	55 mM	$MnCl_2$ × 2 H_2O
	250 mM	KCl

All components, except $MnCl_2$, were mixed and the pH was adjusted to 6.7 with KOH. Then the $MnCl_2$ was added and the mixture was filter sterilized.

E.1.6.5 GUS staining buffer

X-Gluc stock solution:

25 mg X-Gluc (5-Bromo-4-chloro-3-indolyl-ß-D-glucuronide)
dissolved in 500 µl N,N-Dimethylformamid (N,N-DMF)
Use immediately

Potassium-Ferri/Ferrocyanide stock solution

100 mM Potassium Ferricyanide $K_3Fe(CN)_6$
100 mM Potassium Ferrocyanide $K_4Fe(CN)_6 \times 3H_2O$
Store at -20° C

Tris-HCl/NaCl-Buffer

100 mM Tris-HCl
50 mM NaCl
pH 7.0

For 1 ml of GUS-Buffer, add 20 µl of Potassium-Ferri/Ferrocyanide stock solution and 20 µl of X-Gluc stock solution to 960 µl of Tris-HCl/NaCl-buffer. The resulting buffer can be stored at 4° C for maximally 1 – 2 days.

E.1.7 Software and databases

Program/ resource	URL	Reference
BioEdit 7.00	http://www.mbio.ncsu.edu/BioEdit/bioedit.html	(Hall, 1999)
Primer3	http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi	(Rozen and Skaletsky, 2000)
GenDB	http://www.cebitec.uni-bielefeld.de/groups/brf/software/gendb_info/index.html	(Meyer *et al.*, 2003)
Java TreeView	http://genetics.stanford.edu/_alok/TreeView/	(Saldanha, 2004)
Cluster 3.0	2http://bonsai.ims.u-tokyo.ac.jp/_mdehoon/software/cluster/software.htm#ctv	(Yang *et al.*, 2002)
S. meliloti genome project	http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/	(Galibert *et al.*, 2001)

E.1.8 Contents and layout of the tag microarray

The mTn*5*-STM-1 microarray contains 23mer oligonucleotides carrying 5´ C_{12}-amino modifications. The probes are directed against the variable sequences of signature tags of the 412 mTn*5*-STM transposons. A microarray slide contains two arrays each with 4608 spots in 16 grids of 18 rows and 16 columns. The 16 grids were arrayed in a 4x4 pattern. Each oligonucleotide is present in at least four replicates per array. Besides, 3 23-mer genomic control sequences were printed in 192 replicates. Spotting was performed as described by Rüberg *et al.* (2003). Tag sequences and the layout of the mTn*5*-STM-1 microarray are described at http://www.cebitec.uni-bielefeld.de/groups/nwt/transcriptomics_facility/services_and_printed_arrays/.

E.2 METHODS

E.2.1. Cultivation of bacteria

E.2.1.1 Growth conditions

Bacteria were grown on using solid media, as well as liquid media. In special cases, the media were supplemented with additives listed above (E.1.5.2). *E. coli* cells were cultivated at 37 °C in LB, LBG, SOB or SOC medium. For the preparation of competent *E. coli* cells, they were cultivated at the room temperature (18 – 22 °C). *S. meliloti* cells were cultivated at 28 – 30 °C in TY, 2 × TY, or in Vincent minimal medium.

E.2.1.2 Determination of bacterial cell number

Baterial cell numbers were determined using photometrical measures of cell density at 600 nm. Optical density of 0.1 corresponds to 2×10^7 cells/ml if *E. coli* culture is measured, and to 10^8 cells/ml if *S. meliloti* culture is measured. The number of living cells was determined by the quantity of colonies grown on solid medium from the diluted culture.

E.2.1.3 Storage of bacterial strains

Mutant library was stored in standard flat well 96-well plates (Greiner). Other strains were stored in eppendorf tubes.

Storage in glycerol solution

Bacteria were grown in the liquid medium overnight (*E. coli*) or for 2 – 3 days (*S. meliloti*); 100 µl of bacterial culture were mixed with 120 µl of 87 % (v/v) glycerol and stored at -20 °C.

HFM culture

Bacteria were grown as above. 180 µl of bacterial culture were mixed with 20 µl of Hogness Freezing Medium (E.1.5.1) and stored at -80 °C.

E.2.2 Cultivation of plants and harvesting of nodules

E.2.2.1 Sterilization of *M. sativa* seeds

The seeds of *M. sativa* are treated by the concentrated sulfuric acid in order decontaminate seeds from bacteria and fungi, and to stimulate the germination. This method also helps to increase the fraction of germinating seeds to over 90 %.

- Place 20 seeds in an eppendorf tube (in case large quantities of seeds have to be sterilized, use blue caps and multiply all volumes by 40).
- Add 0. 5 ml of concentrated (95 – 97 %) sulfuric acid. Incubate for 12 min agitating.
- Pipet the acid away. It is important to remove as much acid as possible.
- Add 1 ml of sterile Millipore water, invert 2 – 3 times, and remove the water.
- Add 1 ml of fresh Millipore water. Incubate for 2 – 3 min, agitating.
- Repeat the washing 2 more times.
- Let the seeds covered with water (150 – 200 µl) stand in light for 30 – 40 min.
- Place the seeds on the plant agar plates using sterile glass pipette, seal the plates with Parafilm and wrap them in aluminum foil.
- Incubate for 24 hours at 4 °C and subsequently for 24 hours at 28 °C.

E.2.2.2 Cultivation of plants on Petri dishes and inoculation by *S. meliloti*

- Culture the *S. meliloti* strains overnight on TY agar plates. Resuspend the bacteria in 0.85 % NaCl solution, and dilute the cell suspensions so that they contain 2×10^7 cells/ml ($o.D._{600}$=0.02).
- Using tweezers, put the prepared germ buds on the upper 1/3 of plant agar plates in the way that the roots are directed to the thick side of the agar. Three germ buds are put on one agar plate.
- Use 150 µl of bacterial cell suspension to inoculate each plate. Apply the bacterial suspension on the lower part of the plate and along the roots.
- Let the plates dry and place them in the plant incubator roots down. Incubate for 3 weeks or more so that the nodules can form.

E.2.2.3 Cultivation of plants in aeroponic system

(Journet et al., 2001)

The aeroponic system is used when big quantity of plant root material is needed. In this work, this system was used in order to obtain several thousands of nodules from one experiment. In the aeroponics, plants are cultivated without any solid substrate, and the roots are kept moist by the plant medium aerosol.

Seeds of *M. sativa* cv *Europe* were sterilized and cultivated for 6 – 8 days on the plates with plant medium. During this time, the aeroponic tank was sterilized using 0.1 % H_2O_2 for 2 days and subsequently dried for 3 days. The tank was partly filled with 17 – 21 l of aeroponics plant medium supplemented with 5 mM ammonium nitrate. The defensor was switched on for 1 day before the planting to reach the correct humidity inside the tank. The plants were picked from the agar plates and their roots were carefully inserted in the holes of the aeroponic tank lid. After 10 days of cultivation in the presence of nitrogen, the medium containing nitrogen was changed to the nitrogen-free medium. After two more days of cultivation plants were inoculated with 17 ml of *S. meliloti* culture ($o.D._{600}$=1.0, washed twice in plant medium) to promote nodulation. The nodules were harvested 4 weeks after inoculation.

E.2.2.4 Harvesting and sterilization of nodules from the agar plates

The surface of the nodules must be sterilized and washed to avoid the contamination by bacteria that stick to the surface of the nodules.

- Harvest nodules using tweezers and put each one into an eppendorf tube.
- Add 1ml of sterile Millipore water and vortex. Remove the water.
- Add 500 µl of 70 % ethanol and vortex. Incubate for 1 min. Remove the ethanol.
- Wash the nodules three times with 1 ml Millipore, vortex each time.

E.2.2.5 Harvesting and sterilization of nodules from the aeroponic tank

The roots of plants that were grown in the aeroponics are more than 1 m long and carry about 150 - 200 nodules. Nodules were harvested using tweezers and immediately frozen in liquid nitrogen in aliquots of 1.5 – 2.5 g wet weight (3 – 4 ml) in 50 ml blue caps. Before the DNA isolation, nodules had to be washed in order to remove bacteria that stick to their surface.

Protocol:

- Thaw the nodules
- Add 45 ml of sterile Millipore water and invert several times. Remove the water with a glass pipette.
- Add 10 ml of 70% ethanol and invert several times. Incubate for 1 min. Remove the ethanol.
- Wash the nodules three times with 45 ml sterile Millipore.
- Remove as much water as possible. Transfer the nodules into a mortar and immediately cover with liquid nitrogen. Continue with the protocol for DNA isolation (E.2.4.3).

E.2.2.6 Exogenous climatic factors for plant growth

	Plant incubator (agar plates)	Climatic chamber (aeroponics)
Temperature	19 °C	22 °C
Light hours	17 h	18 h

E.2.3 DNA transfer

E.2.3.1. Transformation of *E. coli*

Bacterial transformation is the process by which bacterial cells take up naked DNA molecules. In the transformation procedure, the cell walls of bacteria are treated in a way that the DNA can pass into the cytoplasm. Such treated *E. coli* are called competent cells. In this work, a highly efficient chemical method of competent cells preparation was used (Inoue et al., 1990), resulting in 1 x 10^7 to 1 x 10^8 transformants per microgram of plasmid DNA.

Preparation of competent cells

- Pick 7 – 10 large *E.coli* colonies from the LB-agar plate and resuspend them in 3 ml of SOB. Use approximately 1 ml of the cell suspension to inoculate 150 ml of SOB. Incubate the flask in a shaker (150 rpm) overnight at the room temperature (18 – 22 °C). Harvest the culture at the $o.D._{600}$ of 0.5 – 0.6. Pre-cool the TB buffer and prepare liquid nitrogen.

After this point, all the procedures are made on ice.

- Transfer the culture to blue caps and cool them 10 min on ice.
- Centrifuged for 10 min, at 3000 rpm, 4 °C, discard the supernatant.
- 50 ml ice-cold TB buffer (E.1.6.4) and incubate for 10 min on ice.
- Centrifugation step as above.

- Gently resuspend the pellets in 10 ml of ice-cold TB buffer. Add 0.7 ml of DMSO and incubate on ice for 10 min more.
- Aliquot the cells using in pre-cooled eppendorf tubes and plastic tips. Glass pipettes should not be used for aliquoting.
- Freeze the aliquots in liquid nitrogen and store at -80 °C.

Transformation of competent cells by DNA

- Thaw 150 µl of competent cells at room temperature. Add DNA in the volume of 10 µl or less. Incubate the cells together with DNA on ice for 25 – 40 min.
- Heat-shock the cells by placing in 43 °C water bath for 30 sec and subsequently cooling them on ice for 5 min.
- Add 750 µl of SOC to the cells and incubate at 36 °C for 45 min.
- Plate the transformed cells (10 – 100 µl) on LB plates containing the antibiotic for selection.

E.2.3.2 Conjugation

(Simon et al., 1983)

The *E.coli* S17-1 cells were grown overnight in liquid LB medium supplemented with kanamycin to $o.D._{600}$ = 0.5 - 0.6. *S. meliloti* cells were grown in liquid TY to $o.D._{600}$=0.8 - 1.0. The two cultures were mixed together so that quantity of *S. meliloti* cells was equal to the quantity of *E.coli* cells, e.g 700 µl of rhizobial culture of $o.D._{600}$=0.8 would have been mixed with 190 µl of *E.coli* culture that of $o.D._{600}$ = 0.6.

The mixture was centrifuged and the supernatant was discarded (poured away, not pipetted out). The sells were resuspended in the rest of medium (about 40 µl) and applied on the cellulose acetate filter (Sartorius) placed on TY plate without antibiotic. After growth overnight at 30 °C the filter was taken from the plate and cells were washed from it using 700 µl of 0.85 % NaCl. 100 µl of the suspension were then plated on plates with antibiotic.

E.2.4. Isolation and purification of DNA

E.2.4.1. Isolation of plasmid DNA from bacteria (*E. coli*)

Plasmid DNA was isolated using NucleoSpin® Plasmid kit from Macherey-Nagel. In this method, the plasmid DNA is liberated from the *E. coli* host cells by SDS/alkaline lysis and purified by binding to the NucleoSpin®Plasmid column with subsequent washing end elution steps.

Protocol (Columns, and all the components marked by the sign ▫ are a part of the kit):

- Plate plasmid containing cells on the medium containing antibiotic. Incubate overnight.
- Using a glass pipette, scrape some cells from the plate and resuspend them in 250 µl buffer A1▫ by vigorous vortexing.
- Add 250 µl of buffer A2▫ and mix gently with the cell suspension by inverting the tube 6 – 8 times, than incubate at room temperature for 5 min.
- Add 300 µl of buffer A3▫ and mix gently by inverting the tube 6 – 8 times.
- Centrifuge for 5 - 10 min at 11,000 × g at room temperature.
- Place a NucleoSpin® Plasmid column in a 2 ml collecting tube and load the supernatant from step 3 onto the column. Centrifuge for 1 min at 11,000 × g. Discard flow-through.
- Place the NucleoSpin® Plasmid column back into the 2 ml collecting tube and add 600 µl of buffer A4▫ mixed with ethanol as indicated on the bottle. Centrifuge for 1 min at 11,000 × g. Discard flow-through.
- To dry the silica membrane completely, reinsert the NucleoSpin® Plasmid column into the 2 ml collecting tube. Centrifuge for 2 min at 11,000 x g.
- Place the NucleoSpin® Plasmid column in a 1.5 ml microcentrifuge tube and add 50 µl buffer AE▫. Incubate 1 min at room temperature. Centrifuge for 1 min at 11,000 × g.
- Store the DNA at -20 °C

E.2.4.2 Isolation of genomic DNA from bacteria (*S. meliloti*)

The whole genomic DNA was isolated from *S. meliloti* cells using NucleoSpin® Tissue kit from Macherey-Nagel. With the NucleoSpin Tissue method, lysis is achieved by incubation of the samples in a solution containing SDS and proteinase K at 56 °C. Appropriate conditions for binding of DNA to the silica membrane of the NucleoSpin® Tissue columns are created by addition of large amounts of chaotropic ions and ethanol to the lysate. Contaminations are removed by efficient washing with buffer. Pure genomic DNA is finally eluted under low ionic strength conditions in a slightly alkaline elution buffer.

Protocol (Columns, and all the components marked by the sign ▫ are a part of the kit):

- Centrifuge up to 1 ml culture with $o.D._{600}$ = 0.8 for 5 min at 8,000 × g. Remove supernatant (for more dense cultures, less culture was used; for less dense cultures, several centrifugation steps were performed adding more culture to the same tube).

- Resuspend the pellet in 180 µl buffer T1▫ by pipeting up and down. Add 25 µl proteinase K▫. Vortex vigorously and incubate at 56 °C until complete lysis is obtained (at least 1 – 3 h). Vortex occasionally during incubation or use a shaking incubator.
- Add 20 µl RNase A solution and incubate for an additional 5 min at room temperature.
- Vortex the samples. Add 200 µl of buffer B3▫, vortex vigorously and incubate at 70 °C for 10 min. Vortex briefly.
- Add 210 µl of ethanol (96 – 100 %) to the sample and vortex vigorously.
- For each sample, place one NucleoSpin® Tissue column into a 2 ml collecting tube. Apply the sample to the column. Centrifuge for 1 min at 11,000 × g. Discard the flow-through and place the column back into the collecting tube.
- Add 500 µl buffer BW▫. Centrifuge for 1 min at 11,000 × g. Discard flow-through and place the column back into the collecting tube.
- Add 600 µl buffer B5▫ to the column and centrifuge for 1 min at 11,000 × g. Discard flow-through and place the column back into the collecting tube.
- Centrifuge the column for 1 min at 11,000 × g.
- Place the NucleoSpin® Tissue column into a 1.5 ml microcentrifuge tube and add 100 µl prewarmed elution buffer BE▫ (70 °C). Incubate at room temperature for 1 min. Centrifuge 1 min at 11,000 × g

E.2.4.3. Isolation of pure DNA from *M. sativa* nodules

The mixed bacterial and plant DNA was isolated from *M. sativa* nodules using DNeasy® Plant Mini kit from Qiagen. In the DNeasy Plant procedure, plant material is first mechanically disrupted and then lysed by addition of lysis buffer and incubation at 65 °C. RNase A in the lysis buffer digests the RNA in the sample. After lysis, proteins and polysaccharides are salt precipitated. Cell debris and precipitates are removed in a single step by a spin through the QIAshredder Column that serves as filtration and homogenization unit. The cleared lysate is transferred to a new tube and binding buffer and ethanol are added to promote binding of the DNA to the DNeasy membrane. Contaminants such as proteins and polysaccharides are removed by two wash steps. Pure DNA is eluted in a low-salt buffer.

Protocol (Columns, and all the components marked by the sign ▫ are a part of the kit):

- Grind the washed nodule tissue (E.2.2.5) under liquid nitrogen to a fine powder using a mortar and pestle. Transfer tissue powder to eppendorf tubes (approximately 50 mg of tissue powder per tube).

- Add 400 µl of Buffer AP1▫ and 4 µl of RNase A▫ stock solution (100 mg/ml) to each tube and vortex vigorously (from here all the steps are described for one tube).
- Incubate the mixture for 10 min at 65 °C. Mix 2 – 3 times during incubation by inverting tube.
- Add 130 µl of Buffer AP2▫ to the lysate, mix, and incubate for 5 min on ice.
- Centrifuge the lysate for 5 min at 20,000 × g.
- Apply the lysate to the QIAshredder Mini Spin Column (lilac) placed in a 2 ml collection tube and centrifuge for 2 min at 20,000 × g.
- Transfer flow-through fraction from step 4 to a new tube without disturbing the cell-debris pellet.
- Add 1.5 volumes of Buffer AP3/E▫ to the cleared lysate and mix by pipetting.
- Apply 650 µl of the mixture from step 6, including any precipitate which may have formed, to the DNeasy Mini Spin Column sitting in a 2 ml collection tube. Centrifuge for 1 min at ≥6000 × g and discard flow-through.
- Repeat the previous step with remaining sample. Discard flow-through and collection tube.
- Place DNeasy Mini Spin Column in a new 2 ml collection tube, add 500 µl Buffer AW▫ to the DNeasy Mini Spin Column and centrifuge for 1 min at ≥6000 × g. Discard flow-through and reuse the collection tube in the next step.
- Add 500 µl Buffer AW▫ to the DNeasy Mini Spin Column and centrifuge for 2 min at 20,000 × g to dry the membrane.
- Transfer the DNeasy Mini Spin Column to a 1.5 ml or 2 ml microcentrifuge tube and pipet 100 µl of Buffer AE▫ directly onto the DNeasy membrane. Incubate for 5 min at room temperature (15 – 25°C) and then centrifuge for 1 min at ≥6000 × g to elute.

E.2.4.4. Isolation of DNA from *M. sativa* nodules by boiling

This method was used to isolate DNA from single nodules.

Protocol:

- Place the washed nodule (E.2.2.4) in an eppendorf tube together with 100 µl of sterile Millipore water.
- Grind the nodule using a sterile micropestle.
- Place the tube in a boiling water bath for 5 min.
- Centrifuge 1 min at 6000 × g
- Use 5 µl of supernatant for the PCR.

E.2.4.5. Purification of PCR products

Purification of PCR products was performed using a NucleoSpin® Extract II PCR clean-up kit from from Macherey-Nagel. With this method, DNA binds to a silica membrane in the presence of chaotropic salt added by binding buffer NT. The binding mixture is loaded directly onto NucleoSpin® Extract II columns. Contaminations like salts and soluble macromolecular components are removed by a washing step with ethanolic buffer NT3. Pure DNA is finally eluted under low ionic strength conditions with slightly alkaline buffer NE (5 mM Tris/HCl, pH 8.5).

Protocol (Columns, and all the components marked by the sign ▫ are a part of the kit):

- Mix 1 volume of sample with 2 volumes of buffer NT▫ (e.g. mix 100 µl PCR reaction and 200 µl NT).
- Place a NucleoSpin® Extract II column into a 2 ml collecting tube and load the sample.
- Add 600 µl buffer NT3▫. Centrifuge for 1 min at 11,000 x g. Discard flow-through and place the NucleoSpin® Extract II column back into the collecting tube.
- Centrifuge for 2 min at 11,000 x g to remove buffer NT3▫ quantitatively. Make sure the spin column doesn't come in contact with the flow-through while removing it from the centrifuge and the collecting tube.
- Place the NucleoSpin® Extract II column into a clean 1.5 ml microcentrifuge tube. Add 15 – 50 µl elution buffer NE▫ and incubate at room temperature for 1 min to increase the yield of eluted DNA. Centrifuge for 1 min at 11,000 × g.

E.2.5 Visualization and quantification of DNA

E.2.5.1 Agarose gel electrophoresis

This method is used for visualization, as well as quantification of double-stranded DNA. The agarose gels used in this study were prepared of 1 % or 2.3 % (m/v) strength, depending on the size of DNA fragments that had to be visualized.

Protocol:

- Mix the agarose with TA buffer and boil till diluted.
- Pour the agarose-TA, cooled to 65 °C, in a horizontal chamber with a comb.
- After the gel becomes solid, cover it with TA buffer and remove the comb.
- Mix the DNA probes mixed with BPB loading buffer 2 : 1 Vol. Pipet them into the pockets of the gel.

- Run the gel using the voltage of 40 – 80 V dependent on the gel strength (the higher the strength of a gel, the lower the voltage).
- Dye the gel using a weak Ethidium Bromide solution for 5 min and wash the gel for 7 min with water.
- Make a photograph on the transilluminator under UV light.

E.2.5.2 Checking fluorescently labeled targets on agarose gels prior to microarray hybridizations

(Becker and Rüberg, 2003)

- combine 2 µl of the labeled target with appr. 4 µl of 80 % (v/v) glycerol (NO BPB or other dyes!)
- run a 0.8 % (w/v) agarose gel in TA buffer for 20 min at 80 V
- do not touch gel with gloves or hands at places where targets are loaded
- load BPB-marker at least 2 slots away from the target as a positive control
- switch on the Typhoon Imager at least 30 min before use
- place the agarose gel in -90° orientation on the bottom left corner
- start the "Typhoon Scanner Control" software
- load the Cy5/Cy3_medium template file (click on the menu: Template, Load, Cy5/Cy3_medium)
- the template file specifies the following:
    - aquisition: fluorescence
    - orientation: -90°
    - pixel size: 200 µm
    - focal plane: 3 mm
    - setup is as follows (due to bleaching, Cy5 should be scanned first, then Cy3)

Use	Emission filter	PMT	Laser	Sensitivity
√	Cy5 670 BP30	800	red 633	medium
√	Cy3 555 BP20	800	green 532	medium

- scan the gel at 633 and 532 nm to detect Cy5-red and Cy3-green labeled target cDNA. Note that BPB markers are only detected together with the Cy5 targets at 633 nm.
- save and store the *.ds file obtained on a network drive
- load the *.ds file into ImageQuant. Selecting laser buttons 1 and/or 2 allows to individually evaluate the Cy5 and Cy3 labelings for separately labeled targets or the joint

detection of Cy5/3 images for combined targets. Copy/paste selected regions into Powerpoint, print and archive the bitmap images together with the slide hybridization performed using these targets.

E.2.5.3 Quantification of DNA

The concentration and purity of DNA was determined using ND-1000 Spectrophotometer (NanoDrop). Analysis of dsDNA was performed using a programmed method as recommended by the manufacturer.

E.2.6 Polymerase chain reaction (PCR)

Polymerase chain reaction is a technique for enzymatic replication of DNA *in vitro*. PCR uses two primers that anneal to the sence and anti-sence strands at the ends of the DNA fragment which has to be amplified. The cycles of DNA denaturing - primer annealing - second strand DNA synthesis lead to the amplification of the chosen DNA fragment. Termostable DNA polymerase *Taq* was used in all PCR reactions in this work. The following termal cyclers were used: PTC-100™ (Programmable Thermal Controller, MJ Research), PTC-200™ (Programmable Thermal Cycler, MJ Research) and DYAD™ (Programmable Thermal Cycler, MJ Research)

PCR mixture:

5 µl	DNA template (50-100 ng/µl)
2.5 µl	primer mix (25 pmol/µl)
1 µl	Taq-Polymerase
2.5 µl	dNTPs (10 mM)
10 µl	PCR buffer containing MgSO4
79 µl	H_2O

PCR program:

1) Initial denaturation	95 °C	1min
2) Denaturation	95 °C	30 sec
3) Annealing	50-65 °C	25 sec
4) Elongation	72 °C	1-2 min
5) Final elongation	72 °C	5 min
6) Storage	4 °C	forever

Steps 2 to 4 are cycled 25-35 times.

E.2.7 Cloning

E.2.7.1 DNA digestion using restriction enzymes

Restriction endonucleases cleave double-stranded DNA molecules at the specific recognition sequences. This results in forming of DNA fragments with "blunt" or "sticky" ends, depending on the restriction enzyme. The restriction digestion was performed as recommended by the

manufacturer of the restriction endonucleases and using a suitable restriction buffer. The inactivation of the enzymes was performed by incubation at 65 °C for 20 min. In some cases, a subsequent purification was performed using a NucleoSpin® Extract II PCR clean-up kit from from Macherey-Nagel (E.2.4.5).

E.2.7.2 DNA annealing

DNA annealing is performed in case two single-stranded complementary DNA molesules have to be joined in a single double-stranded molecule. For this, two molecules have to be mixed in proportion 1 : 1, heated in order to denaturate the DNA molecules and slowly cooled down.

An annealing mixture could be prepared as follows:

7.2 µl	DNA 1, 100 mM
7.2 µl	DNA 2, 100 mM
1.6 µl	NaCl, 0.5 M

The annealing mixture is heated to 95 °C and slowly colled to 4 °C (e.g. in a water bath).

E.2.7.3 DNA ligation

The digested DNA fragments with "blunt" ends or compatible "sticky" ends can be joined together using the enzyme T4-Ligase. DNA ligation involves creating a phosphodiester bond between the 3´ hydroxyl of one nucleotide and the 5´ phosphate of another. The concentrations and proportions of ligated molecules may vary depending on application. In case an insert have to be cloned in the plasmid DNA, the concentration of the insert DNA normally should exceed the concentration of the plasmid DNA 3 – 4 fold.

A ligation mixture could be prepared as follows:

3 µl	plasmid DNA
6 µl	insert DNA
2 µl	10 × ligation buffer
1 µl	T4-Ligase
8 µl	H_2O

The ligation mixture is incubated overnight at 4 °C or at a gradient from 17 °C to 4 °C.

E.2.7.4 Blunting of overhangs with Klenow fragment

Klenow fragment of DNA polymerase I has 5´ -> 3´ DNA polymerase and 3´ -> 5´ exonuclease activity. Exo- Klenow fragment, used here, lacks 3´ -> 5´ exonuclease activity. In the

"fill-in" blunting reaction, the 5´ overhangs, created by cleavage with many restriction enzymes, are blunted by the synthesis of the second strand. Klenow fragment and dNTPs can be directly added to the restriction mixture after the deactivation of restriction enzyme (1 µl Klenow and 1 µl of 10 mM dNTPs / 20 µl restriction mixture). Reaction is stopped by heating the mixture for 10 minutes at 75 °C.

E.2.7.5 Dephosphorylation of DNA 5´-termini

Dephosphorylation of 5´ termini is often performed for the linearized plasmid vector in order to prevent self-ligation of the vector. Here, the shrimp alkaline phosphatase (SAP) was used, which is easier to inactivate by heating than Calf Intestinal Alkaline Phosphatase. SAP can be added directly to the restriction mixture so that restriction and dephosphorylation steps can be combined. 1 unit of SAP is added per 1 picomole of DNA 5´-termini and the mixture is incubated at 37 °C for 30 min. The reaction is stopped by heating at 65 °C for 15 min (one heating step is enough to deactivate restriction enzyme and SAP).

E.2.8 Large scale methods (96-well plate format)

The procedures, described here, were used for cloning of the H-tags and K-tags into plasmids and for the transfer of these plasmids into *S. meliloti*. All the methods were optimized to enhance their efficiency and adjusted to 96-well format. Most of the pipetting was performed using 8-channel automatic pipettes.

E.2.8.1 Annealing of tags

Since the tags were synthesized as single-stranded nucleotides, they had to be annealed before cloning into pG18-STM. Annealing was performed in low profile microtiter plates (ABgene) as follows.

An annealing mixture:

7.2 µl	Up-tag, 100 mM
7.2 µl	Down-tag, 100 mM
1.6 µl	NaCl, 0.5 M

The plate was placed in thermal cycler for 3 min with heating at 95 °C. The block was let to cool down till the room temperature, and then the plate was taken out and cooled at +4 °C for at least 15 min. The mix was stored at +4 °C

E.2.8.2 DNA digestion with *Kpn*I restriction enzyme

In the second round of tag cloning, plasmids carrying H-tags, were digested with *Kpn*I restriction enzyme in order to clone K-tags. Prior to digestion, the concentration of H-tag carrying plasmids was measured. The restriction was then performed in a way that the final restriction mixture contained 200 - 250 ng of the plasmid. Restriction enzyme, restriction buffer an water were pre-mixed and pipetted into the wells. The corresponding volume of each plasmid was added, and the missing reaction volumes were filled-up with water in the respective wells. Restriction was performed in low profile microtiter plates (ABgene)

The restriction mixture was prepared as follows:

	1 reaction	105 reactions
Plasmid DNA	0.5 - 3 µl	-
10 × restriction buffer	2.5 µl	262.5 µl
*Kpn*I	1 µl	105 µl
H_2O	16.5 µl	1732.5 µl

The plate was incubated at 37 °C overnight in a termal cycler and stored at -20 °C.

E.2.8.3 Ligation

Ligation mixture was prepared as follows:

	1 reaction	105 reactions
Digested plasmid	0.5 µl	52.5 µl for H-tag cloning
Annealing mix	7.3 µl	766.5 µl
T4 ligase buffer	1 µl	105 µl
T4 DNA ligase	0.7 µl	73.5 µl
PEG	0.5 µl	52.5 µl

Ligase, PEG and ligase buffer were pre-mixed and dispensed in low profile microtiter plates (ABgene) (2.2 µl of mix per well). Then annealing mix and the digested plasmid were added. The plate was placed in thermal cycler and slowly cooled from 16 °C to 8 °C.

E.2.8.4 Transformation of *E. coli*

Transformation of competent DH5α cells by ligation mix

150 µl of competent cells were added to each well of the low profile microtiter plate (ABgene) containing 10 µl of the ligation mix. The mixture was placed on ice for 25 minutes, then in the 42 °C water bath for 30 sec, and subsequently transferred to the ice bath.

The transformation mix was transferred from low profile microtiter plate to the deep-well plate (ABgene) and 750 µl of SOC were added to each well. The plate was shaken vigorously for 1 hour at 37 °C. After centrifugation at 3000 rpm for 6 min, 600 µl of supernatant were removed and discarded, and 120 µl of cells were plated on LBG plates supplemented with antibiotic.

Tranformation of competent S17-1 *E. coli* cells by purified plasmid

In the case of transformation by a purified plasmid DNA, 50 µl of competent cells were used, and 250 µl of SOC medium were added to each well instead of 750 µl. 50 µl of each transformation mix were then plated on Omnitray plates (Nunc) filled with LBG plates supplemented with antibiotic. Plating was performed using an eight-channel pipette with 4 tips instead of 8. In this way, 4 transformations were be plated on one Omnitray plate.

E.2.8.5 Conjugation

The transformed S17-1 *E. coli* cells were picked from LB-kanamycin plates grown overnight in 700 µl of LB supplemented with kanamycin in deep-well plate at 37 °C with shaking. *S. meliloti* cells were scraped from a TY plate, resuspended in 100 ml of TY and grown overnight at 30 °C in a 500 ml flask. 150 µl of each *E. coli* overnight culture were added to 600 µl of fresh LB - kanamycin and grown till $o.D._{600}$ = 0.5 - 0.6. 50 ml of overnight culture of *S. meliloti* 1021 were added to 150 ml of fresh TY and grown till $o.D._{600}$ = 0.8 - 1.0. Then, 150 µl of *E. coli* culture from each well were mixed with 700 µl of *S. meliloti* culture in a new deep-well plate. The plate was centrifuged 10 minutes at 2500 rpm, the supernatant was discarded and the cells resuspended in the rests of medium (20 – 30 µl).

For better conjugation efficiency, cellulose membrane filters from Sartorius (product code 11107--47------N) were used. The filters were placed into wells of 24-well plates (Nunc), filled with 2 × TY agar medium. The resuspended cells were dropped at the filters, dried and incubated at 30°C overnight.

The filters with bacteria were then transferred to 700 µl of 0.85% NaCl in deep-well plate and vortexed. 120 µl of suspension were plated on 2 × TY agar plates supplemented with neomycin, streptomycin and nalidixic acid, and incubated at 30 °C for 3 - 4 days. 25 to 30 colonies were picked to the fresh 2 × TY plates with the same antibiotics and incubated for 2 days at 30 °C.

E.2.9 Growth conditions for bacteria in competition experiments in free-living conditions

378 mutants containing different tags were mixed in roughly equal quantities and the mixture was stored as a glycerol stock. A part of the mutant mixture was stored separately and used as a reference (input pool). In all experiments, 3 biological replicates were used. Rich (TY) and minimal (VMM) media were inoculated by the glycerol culture stock of the mixture of mutants to $o.D._{600}$=0.003. Cultures reached the stationary growth phase at $o.D._{600}$=14 (in TY medium), $o.D._{600}$=12.5 (in VMM) and $o.D._{600}$=8.5 (in stress conditions). The timepoints, at which the cells were collected, are denoted in the Table E1.

For salt- and detergent-induced stress experiments, *S. melioti* cells were cultured in TY medium for 6 h after inoculation and then NaCl or sodium dodecyl sulfate (SDS) were added to a concentration of 400 mM or 0.87 mM, respectively.

TABLE E1. Absolute time after inoculation and optical density $o.D._{600}$ of the time points at which the samples were collected

medium	time	$o.D._{600}$
TY	20 h	0.28
TY	28 h	2.9
TY[a]	36 h	7.2
TY	43 h	11
TY	49 h	13.5
TY	60 h	12.5
VMM	24 h	0.28
VMM	30 h	2.3
VMM[a]	37 h	5.9
VMM	43 h	10
VMM	49 h	12.5
VMM	60 h	11.5
TY-NaCl	24h	4.5
TY-SDS	24h	5

[a]These cultures were used in both experiments

E.2.10 Acetylene reduction assay

In the acetylene reduction assay, nodules or the whole plants are incubated with acetylene, and then the quantity of ethylene that has formed over time is measured. Since nitrogenase can reduce acetylene to ethylene, the measured ethylene/acetylene ratio reflects nitrogenase activity in the nodules.

Plants were grown on the plant agar plates and harvested 22 days post inoculation. Plants were picked from the plates with the whole root system and placed in 12 ml screw cap glass tubes,

covered with rubber septa and closed with perforated screw caps. In each tube, 3 plants were placed. 1 ml of acetylene was injected in each tube using a 1 ml Hamilton GASTIGHT syringe. The tubes with plants were placed in the plant incubator. After 20 hours, the ethylene / acetylene ratio was determined using Hewlett Packard 5710A gas chromatograph.

E.2.11 Histochemical GUS staining

GUS staining was used to detect expression of the promoters in the studied genes in the nodules. Fresh nodules were sectioned using VT1000S vibratom (Leica) to obtain 80 µm thick sections. The sections were covered by GUS buffer (500 µl for 5-8 sections) and incubated at 37 °C until blue staining could be observed (1 to 8 hours, depending on the strength of the promotor). The stained sections were rinsed two times in the Tris-HCl / NaCl Buffer. Additionally, the sections were stained by 0.1 M potassium iodide by dropping about 20 µl of the solution at the edge of cover slip immediately before microscoping. Photographs were taken using a Nicon ECLIPSE 80i microscope with 10-fold magnification and Nicon Digital Sight DS-2Mv camera.

E.2.12 Microarray hybridization

Sm6kOligo microarrays were spotted on Quantifoil QMT epoxy slides. These had to be processed prior to hybridizations to block free epoxy groups.

QMT slide processing (Becker and Rüberg, 2003)

- place the slides in a plastic rack and carry out the processing by transferring the racks from one container to the other, occasionally lift the rack up and down during washing
- wash slides for 5 min at room temperature in 250 ml of rinsing solution 1
- wash slides for 2 min at room temperature in 250 ml of rinsing solution 2, repeat this step
- wash slides for 10 min at room temperature in 250 ml of rinsing solution 3
- wash slides for 1 min at room temperature in 250 ml of MilliQ H_2O
- incubate slides for 15 min at 50 °C in 200 ml prewarmed blocking solution in a glass container, shake at least every 5 min or apply constant shaking. Use a flat bottom glass container to process 1-2 slides (20 ml of blocking solution) and a multiple glass container to process multiple slides (200 ml blocking solution)
- wash slides for 1 min at room temperature in 250 ml of MilliQ H_2O
- place rack on an 12 × 8 cm plastic microplate cover (Genomics Solutions) containing 2 Kim-wipes and immediately spin in the microplate centrifuge at 1200 rpm for 3 min. Use a stack of 3 used glass slides at every side of the plastic dish to lift up the rack with the

slides, this avoids precipitation artetacts at the side of the slide. Be sure to counter-balance using an appropriate balance.

Fluorescently labeled purified PCR products were were lyophilized and resuspended in 110 µl of DIG Easy Hyb solution. Hybridization was carried out at 36 °C for 1 h in the HS4800 Hybridization Station (Tecan). Before applying the hybridization sample to the microarray, it was denatured for 3 minutes at 95 °C. Following hybridization, the arrays were washed twice in 2 × SSC, 0.2 % SDS for 5 minutes at 30 °C and subsequently twice in 0.5 × SSC for 2 minutes at 20 °C.

E.2.13 Tag-microarray data analysis

E.2.13.1 Pre-processing

Image processing was performed with ImaGene (version 6.0.1). For each spot the background corrected spot intensities were calculated using the means of all chosen pixels for background and signal. Negative spots or spots that were flagged as empty or having bad quality were removed. Mutants for which less than half of the replicated spots (for each tag separately) were detected, were sorted out prior to further analysis.

The mean intensity (a-value) was calculated for each spot using the standard formula $a_i=\log_2(R_iG_i)^{0.5}$. $R_i=I_{ch1i}-Bg_{ch1i}$ and $G_i= I_{ch2i}-Bg_{ch2i,}$ where I_{ch1i} or I_{ch2i} is the intensity of a spot in channel 1 or channel 2 and Bg_{ch1i} or Bg_{ch2i} is the background intensity of a spot in channel 1 or channel 2, respectively. The logarithm to the base 2 of the ratio of intensities (the m-value) was calculated for each spot using the formula $m_i=\log_2(R_i/G_i)$ (Becker *et al.*, 2004).

E.2.13.2 Normalization and filtering

Normalization accounts for systematic differences between the two fluorescent dyes caused by different labelling and detection efficiencies (Quackenbush, 2001). Due to these differences, the data distribution is not centered around zero although often most of the genes or mutants should have similar values in experiment and reference.

In a normalization procedure, a normalization factor c is calculated and used to move the center of the distribution to zero:

$$\text{m-value}_{norm}(i) = \text{m-value}(i) - c \text{ for spot } i$$

Several approaches for computation of the normalization factor exist. In this study, global median and LOWESS normalization were used. Data used for the calculation is assumed to have an m-value of zero.

The two types of tags used in this STM experiment introduce additional differences. They may be caused by variable efficiencies of dye incorporation depending on the primers used for amplification of the different tags. Normalization was therefore performed for each type of tag separately.

The data from the pilot verification experiments was normalized using **locally weighted regression (LOWESS)** (Cleveland, 1979). This method accounts for differences between the two dyes that depend on intensity (a-values) (Yang *et al.*, 2002).

The data from the experiment *in planta* was analysed using a modified **global median normalization**. This normalization method assumes that the differences between the two dyes can be described by a constant factor calculated as the median of the m-values (Yang *et al.*, 2002):

c = median (m-value(i)), where i=1,...,nr of spots.

15% of the smallest m-values were excluded from the pool of data used to calculate the median.

For the filtering, the distances between H- and K-tag for each mutant were calculated, using the formula

$$c_t = \left(\left|mean_H - mean_K\right|\right) \Big/ \left(\sqrt{\frac{std_H^2}{n_H} + \frac{std_K^2}{n_K}}\right),$$

where $mean_{tag}$, std_{tag} and n_{tag} denote mean, standard deviation and number of m-values for a single mutant, respectively. The 90%-quantiles of the distribution of each criterion for all mutants were determined and the mutants exceeding this threshold were removed from the data set.

In case several technical replicates were used, the weighted mean of the medians of the m-values derived from the biological replicates was calculated for each mutant combining values of H-tags and K-tags:

$$mean_w = \frac{1}{n}\sum_{i=1}^{n_{rep}} n_i * median_i,$$

Where n_{rep}, n_i and mediani denote the number of replicates, the number of spots of replicate i and the median of the m-values of replicate i respectively, and $n = \sum_{i=1}^{rep} n_i$ is the overall number of spots. Mutants, whose $mean_w$-values were higher than 0.7 or lower than -0.7, were considered as having an induced or attenuated phenotype, respectively.

E.2.13.3 Cluster analysis

Cluster analysis was performed using Cluster 3.0 (Yang *et al.*, 2002), a successor of Eisen's Cluster software (Eisen *et al.*, 1998).

The following distances between two data points x and y were applied in the clustering algorithms:

1. Euclidean distance: $d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$

2. Centered correlation: $d(x,y) = 1 - r_c$, where $r_c = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$;

3. Uncentered correlation: $d(x,y) = 1 - r_u$, where $r_u = \frac{\sum_{i=1}^{n}(x_i)(y_i)}{\sqrt{\sum_{i=1}^{n}(x_i)^2}\sqrt{\sum_{i=1}^{n}(y_i)^2}}$

Where $\bar{x}$ and $\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ denote the mean and standard deviation of x, respectively.

k-means clustering was performed in the following way. The algorithm was repeated 105 times. Each times a different set of k rows (= mutants) was selected as prototype vectors in the initialization step. The solution with the smallest sum of distances within clusters (i. e. distances between data points in one cluster) was chosen as the final result.

Java Treeview (Saldanha, 2004) was used for visualization where the contrast value for all plots was set to 2.0.

In the TY-VMM time course experiment, the data matrix for clustering analysis after filtering of absolute m-values consisted of 27 mutants in rows and 12 time points (6 per medium) in columns. In the *k*-means clustering, number of clusters *k* was set to 6. In the stress conditions experiment, the data from the stress conditions were clustered together with the timepoint 3 data for VMM and TY cultures. As a result, the data matrix for the clustering consisted of 29 mutants in rows and 4 conditions in columns. Number of clusters *k* was set to 8. In both experiments, the distance was calculated based on uncentered Pearson correlation.

F. REFERENCES

Abouelhoda, M.I., Kurtz, S., and Ohlebusch, E. (2004) Replacing Suffix trees with enhanced Suffix arrays. *J. Discrete Algorithms* **2**: 53-86.

Aguilar, O.M., Kapp, D., and Pühler, A. (1985) Characterization of a *Rhizobium meliloti* fixation gene (*fixF*) located near the common nodulation region *J. Bacteriol.* **164**: 245-254.

Aguilar, O.M., Reilander, H., Arnold, W., and Pühler, A. (1987) *Rhizobium meliloti nifN* (*fixF*) gene is part of an operon regulated by a *nifA*-dependent promoter and codes for a polypeptide homologous to the *nifK* gene product *J. Bacteriol.* **169**: 5393-5400.

Aguilar, O.M., and Grasso, D.H. (1991) The product of the *Rhizobium meliloti ilvC* gene is required for isoleucine and valine synthesis and nodulation of alfalfa. *J. Bacteriol.* **173**: 7756-7764.

Allen, N.S., Bennett, M.N., Cox, D.N., Shipley, A., Ehrhardt, D.W., and Long, S.R. (1994) Effects of Nod factors on alfalfa root hair Ca2+ and H+ currents and on cytoskeleton behavior. In *In Advances in Molecular Genetics of Plant-Microbe Interactions*. Vol. 3. M.G.Daniels, Downie, J.A. and Osbourne, A.E. (eds): Dordrecht: Kluwer Academic Publishers, pp. 107-114.

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

Ampe, F., Kiss, E., Sabourdy, F., and Batut, J. (2003) Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. *Genome Biology* **4**: R15.

Angelichio, M.J., and Camilli, A. (2002) *In Vivo* Expression Technology. *Infect. Immun.* **70**: 6518-6523.

Bardin, S., Dan, S., Osteras, M., and Finan, T.M. (1996) A phosphate transport system is required for symbiotic nitrogen fixation by *Rhizobium meliloti*. *J. Bacteriol.* **178**: 4540-4547.

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

Bardin, S.D., Voegele, R.T., and Finan, T.M. (1998) Phosphate Assimilation in *Rhizobium (Sinorhizobium) meliloti*: Identification of a *pit*-Like Gene. *J. Bacteriol.* **180**: 4219-4226.

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

Barran, L.R., Bromfield, E.S., and Brown, D.C. (2002) Identification and cloning of the bacterial nodulation specificity gene in the *Sinorhizobium meliloti--Medicago laciniata* symbiosis. *Can. J. Microbiol.* **48**: 765-767.

Barsch, A., Patschkowski, T., and Niehaus, K. (2004) Comprehensive metabolite profiling of *Sinorhizobium meliloti* using gas chromatography-mass spectrometry. *Funct. Integr. Genomics.* **4**: 219-230.

Batut, J., Terzaghi, B., Gherardi, M., Huguet, M., Terzaghi, E., Garnerone, A.M., Bistard, P., and Huguet, T. (1985) Localization of a symbiotic *fix* region on *Rhizobium meliloti* pSym megaplasmid more than 200 kilobases from the *nod-nif* region *Mol. Gen. Genet.* **199**: 232-239.

Baumgarth, B., Bartels, F.W., Anselmetti, D., Becker, A., and Ros, R. (2005) Detailed studies of the binding mechanism of the *Sinorhizobium meliloti* transcriptional activator ExpG to DNA. *Microbiology* **151**: 259-268.

Becker, A., Kleickmann, A., Arnold, W., and Puhler, A. (1993a) Analysis of the *Rhizobium meliloti exoH/exoK/exoL* fragment: ExoK shows homology to excreted endo-beta-1,3-1,4-glucanases and ExoH resembles membrane proteins. *Mol. Gen. Genet.* **238**: 145-154.

Becker, A., Kleickmann, A., Keller, M., Arnold, W., and Pühler, A. (1993b) 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.

Becker, A., Kleickmann, A., Küster, H., Keller, M., Arnold, W., and Pühler, A. (1993c) 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.

Becker, A., and Rüberg, S. (2003) *Sinorhizobium meliloti* SM6KOligo microarray manual. Department of Genetics, Institute for Genome Research, University of Bielefeld.

Becker, A., Bergès, H., Krol, E., Bruand, C., Rüberg, S., Capela, D., Lauber, E., Meilhoc, E., Ampe, F., Bruijn, F.J.d., Fourment, J., Francez-Charlot, A., D. Kahn, H.K., Liebe, C., Pühler, A., Weidner, S., and Batut, J. (2004) Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. *Mol. Plant-Microbe Interact.* **17**: 292-303.

Becker, A., Fraysse, N., and Sharypova, L. (2005) Recent advances in studies on structure and symbiosis-related function of rhizobial K-antigens and lipopolysaccharides. *Mol. Plant-Microbe Interact.* **18**: 899–905.

Berges, H., Checroun, C., Guiral, S., Garnerone, A.-M., Boistard, P., and Batut, J. (2001) A glutamine-amidotransferase-like protein modulates FixT anti-kinase activity in *Sinorhizobium meliloti*. *BMC Microbiology* **1**: 6.

Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* **84**: 188-198.

Bisseling, T., van den Bos, R.C., Weststrate, M.W., Hakkaart, M.J., and van Kammen, A. (1979) Development of the nitrogen-fixing and protein-synthesizing apparatus of bacteroids in pea root nodules. *Biochim. Biophys. Acta* **562**: 515-526.

Brencic, A., and Winans, S.C. (2005) Detection of and Response to Signals Involved in Host-Microbe Interactions by Plant-Associated Bacteria. *Microbiol. Mol. Biol. Rev.* **69**: 155-194.

Brent, R., and Ptashne, M. (1980) The *lexA* Gene Product Represses Its Own Promoter. *PNAS* **77**: 1932-1936.

Brewin, N.J. (1991) Development of the Legume Root Nodule. *Annual Review of Cell Biology* **7**: 191-226.

Broughton, W.J., and Dilworth, M.J. (1971) Control of leghaemoglobin synthesis in snake beans. *Biochemical Journal* **125**: 1075-1080.

Brozek, K.A., Carlson, R.W., and Raetz, C.R.H. (1996) A special acyl carrier protein for transferring long hydroxylated fatty acids to lipid A in *Rhizobium*. *J. Biol. Chem.* **271**: 32126-32136.

Buikema, W.J., Klingensmith, J.A., Gibbons, S.L., and Ausubel, F.M. (1987) Conservation of structure and location of *Rhizobium meliloti* and *Klebsiella pneumoniae nifB* genes. *J. Bacteriol.* **169**: 1120-1126.

Caldon, C.E., Yoong, P., and March, P.E. (2001) Evolution of a molecular switch: universal bacterial GTPases regulate ribosome function. *Molecular Microbiology* **41**: 289-297.

Campbell, G.R.O., Reuhs, B.L., and Walker, G.C. (2002) Chronic intracellular infection of alfalfa nodules by *Sinorhizobium meliloti* requires correct lipopolysaccharide core. *PNAS* **99**: 3938-3943.

Campbell, G.R.O., Sharypova, L.A., Scheidle, H., Jones, K.M., Niehaus, K., Becker, A., and Walker, G.C. (2003) Striking Complexity of Lipopolysaccharide Defects in a Collection of *Sinorhizobium meliloti* Mutants. *J. Bacteriol.* **185**: 3853-3862.

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

Capela, D., Filipe, C., Bobik, C., Batut, J., and Bruand, C. (2006) *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: a transcriptomic dissection. *Mol. Plant-Microbe Interact.* **19**: 363-372.

Carter, P. (1991) Mutagenesis facilitated by the removal or introduction of unique restriction sites. In *Directed Mutagenesis: A Practical Approach*. McPherson, M.J. (ed). NY: Oxford University Press, pp. 1-25.

Chen, J.S., and Johnson, J.L. (1993) Molecular biology of nitrogen fixation in the clostridia. *Biotechnology* **25**: 371-392.

Cheng, H.-P., and Walker, G.C. (1998) Succinoglycan Is Required for Initiation and Elongation of Infection Threads during Nodulation of Alfalfa by *Rhizobium meliloti*. *J. Bacteriol.* **180**: 5183–5191.

Chiang, S.L., Mekalanos, J.J., and Holden, D.W. (1999) *In vivo* genetic analysis of bacterial virulence. *Annu. Rev. Microbiol.* **53**: 129-154.

Chua, K.L., Mak, Y.K., and Oliver, P. (1993) Expression of the *recA* gene in recombination-deficient (rec-) strains of *Escherichia coli*. *Biochimie* **75**: 775-783.

Cianciotto, N.P., Cornelis, P., and Baysse, C. (2005) Impact of the bacterial type I cytochrome *c* maturation system on different biological processes. *Molecular Microbiology* **56**: 1408-1415.

Cleveland, W.S. (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**: 829-836.

Clover, R.H., Kieber, J., and Signer, E.R. (1989) Lipopolysaccharide mutants of *Rhizobium meliloti* are not defective in symbiosis. *J. Bacteriol.* **171**: 3961-3967.

Cosseau, C., Garnerone, A.M., and Batut, J. (2002) The FixM flavoprotein modulates inhibition by AICAR or 5'AMP of respiratory and nitrogen fixation gene expression in *Sinorhizobium meliloti*. *Mol. Plant-Microbe Interact.* **15**: 598-607.

Cummings, C.A., and Relman, D.A. (2000) Using DNA microarrays to study host-microbe interactions. *Genomics* **6**: 513-525.

D'Haeze, W., and Holsters, M. (2002) Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* **12**: 79R-105.

Davey, M.E., and de Bruijn, F.J. (2000) A homologue of the tryptophan-rich sensory protein TspO and FixL regulate a novel nutrient deprivation-induced *Sinorhizobium meliloti* locus. *Appl. Environ. Microbiol.* **66**: 5353-5359.

de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K.N. (1990) Mini-Tn*5* transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.* **172**: 6568-6572.

Demont, N., Debelle, F., Aurelle, H., Denarie, J., and Prome, J.C. (1993) Role of the *Rhizobium meliloti nodF* and *nodE* genes in the biosynthesis of lipo-oligosaccharidic nodulation factors. *J. Biol. Chem.* **268**: 20134-20142.

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

Djordjevic, M.A., Chen, H.C., Natera, S., Van Noorden, G., Menzel, C., Taylor, S., Renard, C., Geiger, O., and Weiller, G.F. (2003) A global analysis of protein expression profiles in *Sinorhizobium meliloti*: discovery of new genes for nodule occupancy and stress adaptation. *Mol. Plant-Microbe Interact.* **16**: 508-524.

Driscoll, B.T., and Finan, T.M. (1997) Properties of NAD(+)- and NADP(+)-dependent malic enzymes of *Rhizobium* (*Sinorhizobium*) *meliloti* and differential expression of their genes in nitrogen-fixing bacteroids. *Microbiology* **143**: 489-498.

Duncan, M.J., and Fraenkel, D.G. (1979) α-Ketoglutarate dehydrogenase mutant of *Rhizobium meliloti*. *J. Bacteriol.* **37**: 415-419.

Dunman, P.M., Ren, L., Rahman, M.S., Palejwala, V.A., Murphy, H.S., Volkert, M.R., and Humayun, M.Z. (2000) *Escherichia coli* cells defective for the *recN* gene display constitutive elevation of mutagenesis at 3,N4-ethenocytosine via an SOS-induced mechanism. *Molecular Microbiology* **37**: 680-686.

Dusha, I., Olah, B., Szegletes, Z., Erdei, L., and Kondorosi, A. (1999) *syrM* is involved in the determination of the amount and ratio of the two forms of the acidic Exopolysaccharide EPSI in *Rhizobium meliloti*. *Mol. Plant-Microbe Interact.* **12**: 755-765.

Dylan, T., Ielpi, L., Stanfield, S., Kashyap, L., Douglas, C., Yanofsky, M., Nester, E., Helinski, D.R., and Ditta, G. (1986) *Rhizobium meliloti* genes required for nodule development are related to chromosomal virulence genes in *Agrobacterium tumefaciens*. *PNAS* **83**: 4403-4407.

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

Earl, C.D., Ronson, C.W., and Ausubel, F.M. (1987) Genetic and structural analysis of the *Rhizobium meliloti fixA*, *fixB*, *fixC*, and *fixX* genes. *J. Bacteriol.* **169**: 1127-1136.

Ehrhardt, D.W., Atkinson, E.M., and Long, S.R. (1992) Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* **256**: 998-1000.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *PNAS* **95**: 14863-14868.

Engelke, T.H., Jagadish, M.N., and Pühler, A. (1987) Biochemical and genetical analysis of *Rhizobium meliloti* mutants defective in C4-dicarboxylate transport. *J. Gen. Microbiol.* **133**: 3019-3029.

Entcheva, P., Phillips, D.A., and Streit, W.R. (2002) Functional Analysis of *Sinorhizobium meliloti* Genes Involved in Biotin Synthesis and Transport. *Appl. Environ. Microbiol.* **68**: 2843-2848.

Feldkamp, U., Saghafi, S., Banzhaf, W., and Rauhe, H. (2001) DNA sequence generator: A program for the construction of DNA sequences. In *Proceedings of the Seventh International Workshop on DNA Based Computers (DNA 7)*. Jonoska, N. and Seeman, N.C. (eds.) University of South Florida,Tampa, FL, pp. 23-32.

Ferguson, G.P., Datta, A., Baumgartner, J., Roop, R.M., 2nd, Carlson, R.W., and Walker, G.C. (2004) Similarity to peroxisomal-membrane protein family reveals that *Sinorhizobium* and *Brucella BacA* affect lipid-A fatty acids. *PNAS* **101**: 5012-5017.

Finan, T.M., McWhinnie, E., Driscoll, B., and Watson, R.J. (1991) Complex symbiotic phenotypes result from glucogeogenic mutations in *Rhizobium meliloti*. *Mol. Plant-Microbe Interact.* **4**: 386-392.

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

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

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.

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

Gardiol, A., Arias, A., Cervenansky, C., and Martinez-Drets, G. (1982) Succinate dehydrogenase mutant of *Rhizobium meliloti*. *J. Bacteriol.* **151**: 1621-1623.

Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004) *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **5**: 195-200.

Geoffroy, M.-C., Floquet, S., Metais, A., Nassif, X., and Pelicic, V. (2003) Large-Scale Analysis of the Meningococcus Genome by Gene Disruption: Resistance to Complement-Mediated Lysis. *Genome Res.* **13**: 391-398.

Geurts, R., and Bisseling, T. (2002) *Rhizobium* Nod Factor Perception and Signalling. *Plant Cell* **14**: S239-249.

Goethals, K., Montagu, M.V., and Holsters, M. (1992) Conserved Motifs in a Divergent nod Box of *Azorhizobium caulinodans* ORS571 Reveal a Common Structure in Promoters Regulated by LysR-Type Proteins. *PNAS* **89**: 1646-1650.

Golden, J.W., and Yoon, H.S. (2003) Heterocyst development in *Anabaena*. *Curr. Opin. Microbiol.* **6**: 557-563.

Green, L.S., and Emerich, D.W. (1997) The Formation of Nitrogen-Fixing Bacteroids Is Delayed but Not Abolished in Soybean Infected by an [alpha]-Ketoglutarate Dehydrogenase-Deficient Mutant of *Bradyrhizobium japonicum*. *Plant Physiol.* **114**: 1359-1368.

Groh, J.L., Luo, Q., Ballard, J.D., and Krumholz., L.R. (2005) A method adapting microarray technology for signature-tagged mutagenesis of *Desulfovibrio desulfuricans* G20 and *Shewanella oneidensis* MR-1 in anaerobic sediment survival experiments. . *Appl. Environ. Microbiol.* **71**: 7064-7074.

Guillen-Navarro, K., Araiza, G., Garcia-de los Santos, A., Mora, Y., and Dunn, M.F. (2005) The *Rhizobium etli bioMNY* operon is involved in biotin transport. *FEMS Microbiol. Lett.* **250**: 209-219.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**: 95-98.

Hava, D., and Camilli, A. (2002) Large-scale identificationof serotype 4 *Streptococcus pneumoniae* virulence factors. *Molecular Microbiology* **45**: 1389-1406.

Haydon, D.J., and Guest, J.R. (1991) A new family of bacterial regulatory proteins. *FEMS Microbiol. Lett.* **63**: 291-295.

Heidstra, R., Geurts, R., Franssen, H., Spaink, H.P., van Kammen, A., and Bisseling, T. (1994) Root Hair Deformation Activity of Nodulation Factors and Their Fate on *Vicia sativa*. *Plant Physiol.* **105**: 787-797.

Heinrich, K., Ryder, M.H., and Murphy, P.J. (2001) Early production of rhizopine in nodules induced by *Sinorhizobium meliloti* strain L5-30. *Can. J. Microbiol.* **47**: 165-171.

Hendrixson, D.R., and DiRita, V.J. (2004) Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Molecular Microbiology* **52**: 471-484.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**: 400-403.#

Heyer, L.J., Kruglyak, S. and Yooseph, S. (1999) Exploring Expression Data: Identification and Analysis of Coexpressed Genes, *Genome Research* 9:1106-1115.

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

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

Honma, M.A., and Ausubel, F.M. (1987) *Rhizobium meliloti* Has Three Functional Copies of the *nodD* Symbiotic Regulatory Gene. *PNAS* **84**: 8558-8562.

Hosie, A.H.F., Allaway, D., Galloway, C.S., Dunsby, H.A., and Poole, P.S. (2002) *Rhizobium leguminosarum* has a second general amino acid permease with unusually broad substrate specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family. *J. Bacteriol.* **184**: 4071-4080.

Ichige, A., and Walker, G.C. (1997) Genetic analysis of the *Rhizobium meliloti bacA* gene: functional interchangeability with the *Escherichia coli sbmA* gene and phenotypes of mutants. *J. Bacteriol.* **179**: 209-216.

Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**: 23-28.

Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M.V., and Manoil, C. (2003) Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *PNAS* **100**: 14339-14344.

John, M., Rohrig, H., Schmidt, J., Wieneke, U., and Schell, J. (1993) Rhizobium NodB protein involved in nodulation signal synthesis is a chitooligosaccharide deacetylase. *PNAS* **90**: 625-629.

Journet, E.P., El-Gachtouli, N., Vernoud, V., de Billy, F., Pichon, M., Dedieu, A., Arnould, C., Morandi, D., Barker, D.G., and Gianinazzi-Pearson, V. (2001) *Medicago truncatula* ENo.D.11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. *Mol. Plant-Microbe Interact.* **14**: 737-748.

Kahn, D., David, M., Domergue, O., Daveran, M.L., Ghai, J., Hirsch, P.R., and Batut, J. (1989) *Rhizobium meliloti fixGHI* sequence predicts involvement of a specific cation pump in symbiotic nitrogen fixation. *J. Bacteriol.* **171**: 929-939.

Kamst, E., Spaink, H.P., and Kafetzopoulos, D. (1998) Biosynthesis and secretion of rhizobial lipochitin-oligosaccharide signal molecules. *Subcell. Biochem.* **29**: 29-71.

Kannenberg, E.L., and Brewin, N.J. (1989) Expression of a cell surface antigen from *Rhizobium leguminosarum* 3841 is regulated by oxygen and pH. *J. Bacteriol.* **171**: 4543-4548.

Karlyshev, A.V., Oyston, P.C., Williams, K., Clark, G.C., Titball, R.W., Winzeler, E.A., and Wren, B.W. (2001) Application of high-density array-based signature-tagged mutagenesis to discover novel *Yersinia* virulence-associated genes. *Infect. Immun.* **69**: 7810-7819.

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

Kereszt, A., Slaska-Kiss, K., Putnoky, P., Banfalvi, Z., and Kondorosi, A. (1995) The *cycHJKL* genes of *Rhizobium meliloti* involved in cytochrome c biogenesis are required for "respiratory" nitrate reduction ex planta and for nitrogen fixation during symbiosis. *Mol. Gen. Genet.* **247**: 39-47.

Kerppola, T.K., and Kahn, M.L. (1988) Symbiotic phenotypes of auxotrophic mutants of *Rhizobium meliloti* 104A14. *J. Gen. Microbiol.* **134**: 913-919.

Kim, C.H., Kuykendall, L.D., Shah, K.S., and Keister, D.L. (1988) Induction of Symbiotically Defective Auxotrophic Mutants of *Rhizobium fredii* HH303 by Transposon Mutagenesis **54**: 423-427.

Kirchner, O., and Tauch, A. (2003) Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. *J. Biotechnol.* **104**: 287-299.

Klipp, W., Reilander, H., Schlüter, A., Krey, R., and Pühler, A. (1989) The *Rhizobium meliloti fdxN* gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with *nifA* and *nifB*. *Mol. Gen. Genet.* **216**: 293-302.

Krol, E. (2004) Microarray-based screening of the *Sinorhizobium meliloti* genome for genes specifically regulated under defined environmental conditions: logarithmic growth, stress response and symbiosis. In *Fakultät für biologie*. Vol. PhD Bielefeld: Universität Bielefeld, pp. 56.

Lavesa-Curto, M., Sayer, H., Bullard, D., MacDonald, A., Wilkinson, A., Smith, A., Bowater, L., Hemmings, A., and Bowater, R.P. (2004) Characterization of a temperature-sensitive DNA ligase from *Escherichia coli*. *Microbiology* **150**: 4171-4180.

Lehoux, D.E., Sanschagrin, F., and Levesque, R.C. (1999) Defined oligonucleotide tag pools and PCR screening in signature-tagged mutagenesis of essential genes from bacteria. *Biotechniques* **26**: 473-480.

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

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

Lipshutz, R.J., Fodor, S.P.A., Gingeras, T.R., and Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nature Genetics* **21**: 20 - 24.

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

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

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

Louch, H.A., and Miller, K.J. (2001) Synthesis of a low-molecular-eeight form of exopolysaccharide by *Bradyrhizobium japonicum* USDA 110. *Appl. Environ. Microbiol.* **67** 1011-1014.

Luyten, E., and Vanderleyden, J. (2000) Survey of genes identified in *Sinorhizobium meliloti* spp., necessary for the development of an efficient symbiosis. *Eur. J. Soil Biol.* **36**: 1-26.

Macnab, R.M. (1996) Flagella and motility. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. Neidhardt, F.C., Curtiss, R.I., Ingraham, J.L., Lin, E.C.C., Low, K.B., Masaganik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umbarger, H.E. (eds). Washington D. C.: American Society for Microbiology, pp. 123-145.

MacQueen, J.B. (1967) Some Methods for classification and Analysis of Multivariate Observations. In *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. Le Cam, L.M. and Neyman, J. (eds.) Berkeley, Califonia: University of California Press, pp. 281-297.

March, P.E. (1992) Membrane-associated GTPases in bacteria. *Mol. Microbiol.* **6**: 1253-1257.

McDermott, T.R., and Kahn, M.L. (1992) Cloning and mutagenesis of the *Rhizobium meliloti* isocitrate dehydrogenase gene. *J. Bacteriol.* **174**: 4790-4797.

Meade, H.M., and Signer, E.R. (1977) Genetic Mapping of *Rhizobium Meliloti*. *PNAS* **74**: 2076-2078.

Meyer, F., Goesmann, A., McHardy, A.C., Bartels, D., Bekel, T., Clausen, J., Kalinowski, J., Linke, B., Rupp, O., Giegerich, R., and Pühler, A. (2003) GenDB--an open source genome annotation system for prokaryote genomes. *Nucl. Acids Res.* **31**: 2187-2195.

Michaeli, S., Ron, E.Z., and Cohen, G. (1981) Construction and physical mapping of plasmids containing the *metA* gene of *Escherichia coli* K-1. *Mol. Gen. Genet.* **182**: 349-354.

Milcamps, A., Ragatz, D.M., Lim, P., Berger, K.A., and de Bruijn, F.J. (1998) Isolation of carbon- and nitrogen-deprivation-induced loci of *Sinorhizobium meliloti* 1021 by Tn*5-luxAB* mutagenesis. *Microbiology* **144**: 3205-3218.

Miranda-Rios, J., Morera, C., Taboada, H., Davalos, A., Encarnacion, S., Mora, J., and Soberon, M. (1997) Expression of thiamin biosynthetic genes (*thiCOGE*) and production of symbiotic terminal oxidase cbb3 in *Rhizobium etli*. *J. Bacteriol.* **179**: 6887-6893.

Morgan, B.J.T., and Ray, A.P.G. (1995) Non-uniqueness and inversions in cluster analysis. *Applied Statistics* **44**: 117-134.

Mortimer, M.W., McDermott, T.R., York, G.M., Walker, G.C., and Kahn, M.L. (1999) Citrate synthase mutants of *Sinorhizobium meliloti* are ineffective and have altered cell surface polysaccharides. *J. Bacteriol.* **181**: 7608-7613.

Mylona, P., Pawlowski, K., and Bisseling, T. (1995) Symbiotic Nitrogen Fixation. *Plant Cell* **7**: 869-885.

Nadler, K.D., Johnston, A.W.B., Chen, J.W., and John, T.R. (1990) A *Rhizobium leguminosarum* mutant defective in symbiotic iron acquisition. *J. Bacteriol.* **172**: 670-677.

Noirot-Gros, M.-F., Dervyn, E., Wu, L.J., Mervelet, P., Errington, J., Ehrlich, S.D., and Noirot, P. (2002) An expanded view of bacterial DNA replication. *Proc Natl Acad Sci USA* **99**: 8342–8347.

O'Connell, K.P., Gustafson, A.M., Lehmann, M.D., and Thomashow, M.F. (2000) Identification of cold shock gene loci in *Sinorhizobium meliloti* by using a *luxAB* reporter transposon. *Appl. Environ. Microbiol.* **66**: 401-405.

Oke, V., and Long, S.R. (1999) Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis. *Molecular Microbiology* **32**: 837-849.

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

Paau, A.S., Leps, W.T., and Brill, W.J. (1985) Regulation of nodulation by *Rhizobium meliloti* 102F15 on its mutant which forms an unusually high number of nodules on alfalfa. *Appl. Environ. Microbiol.* **50**: 1118–1122.

Parker, G.F., Higgins, T.P., Hawkes, T., and Robson, R.L. (1999) *Rhizobium* (*Sinorhizobium*) meliloti *phn* genes: characterization and identification of their protein products. *J. Bacteriol.* **181**: 389-395.

Patriarca, E.J., Tate, R., and Iaccarino, M. (2002) Key role of bacterial NH4+ metabolism in *Rhizobium*-plant symbiosis. *Microbiol. Mol. Biol. Rev.* **66**: 203-222.

Pawlowski, K., and Bisseling, T. (1996) Rhizobial and actinorhizal symbioses: what are the shared features? *Plant Cell* **8**: 1899-1913.

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

Picksley, S.M., Attfield, P.V., and Lloyd, R.G. (1984) Repair of DNA double-strand breaks in *Escherichia coli* K12 requires a functional *recN* product. *Mol. Gen. Genet.* **195**: 267-274.

Poole, R.K., and Hill, S. (1997) Respiratory protection of nitrogenase activity in *Azotobacter vinelandii*--roles of the terminal oxidases. *Biosci. Rep.* **17**: 303-317.

Porankiewicz, J., Wang, J., and Clarke, A.K. (1999) New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. *Molecular Microbiology* **32**: 449-458.

Prell, J., and Poole, P. (2006) Metabolic changes of rhizobia in legume nodules. *Trends Microbiol.* **14**: 161-168.

Putnoky, P., Kereszt, A., Nakamura, T., Endre, G., Grosskopf, E., Kiss, P., and Kondorosi, Á. (1998) The *pha* gene cluster of *Rhizobium meliloti* involved in pH adaptation and symbiosis encodes a novel type of K+ efflux system. *Mol. Microbiol.* **28**: 1091-1101.

Quackenbush, J. (2001) Computional analysis of microarray data. *Nat. Rev. Genet.* **2**: 418-427.

Reeve, W.G., Tiwari, R.P., Worsley, P.S., Dilworth, M.J., Glenn, A.R., and Howieson, J.G. (1999) Constructs for insertional mutagenesis, transcriptional signal localization and gene regulation studies in root nodule and other bacteria. *Microbiology* **145**: 1307-1316.

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

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

Reuhs, B.L., Williams, M.N., Kim, J.S., Carlson, R.W., and Cote, F. (1995) Suppression of the Fix- phenotype of *Rhizobium meliloti exoB* mutants by *lpsZ* is correlated to a modified expression of the K polysaccharide. *J. Bacteriol.* **177**: 4289-4296.

Rigali, S., Derouaux, A., Giannotta, F., and Dusart, J. (2002) Subdivision of the Helix-Turn-Helix GntR Family of Bacterial Regulators in the FadR, HutC, MocR, and YtrA Subfamilies. *J. Biol. Chem.* **277**: 12507-12515.

Roberts, I.S. (1996) The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu. Rev. Microbiol.* **50**: 285-315.

Robleto, E.A., Kmiecik, K., Oplinger, E.S., Nienhuis, J., and Triplett, E.W. (1998) Trifolitoxin production increases nodulation competitiveness of *Rhizobium etli* CE3 under agricultural conditions. *Appl. Environ. Microbiol.* **64**: 2630-2633.

Roche, P., Debelle, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarie, J., and Prome, J.C. (1991) Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* **67**: 1131-1143.

Rochfort, S. (2005) Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. *J. Nat. Prod.* **68**: 1813-1820.

Rohrig, H., Schmidt, J., Wieneke, U., Kondorosi, E., Barlier, I., Schell, J., and John, M. (1994) Biosynthesis of lipooligosaccharide nodulation factors: *Rhizobium* NodA protein is involved in N-acylation of the chitooligosaccharide backbone. *PNAS* **91**: 3122-3126.

Ronson, C.W., Nixon, B.T., Albright, L.M., and Ausubel, F.M. (1987) *Rhizobium meliloti ntrA* (*rpo*N) gene is required for diverse metabolic functions *J. Bacteriol.* **169**: 2424-2431.

Rostas, K., S., Morton, J., Picksley, S.M., and Lloyd, R.G. (1987) Nucleotide sequence and LexA regulation of the *Escherichia coli* recN gene. *Nucleic Acids Res.* **15**: 5041-5050.

Rozen, S., and Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Krawetz, S., Misener, S. (ed). Totowa, NJ: Humana Press, pp. 365-386.

Rüberg, S., Pühler, A., and Becker, A. (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**: 603-611.

Rüberg, S., Tian, Z.X., Krol, E., Linke, B., Meyer, F., Wang, Y., Pühler, A., Weidner, S., and Becker, A. (2003) Construction and validation of a *Sinorhizobium meliloti* whole genome DNA microarray: genome-wide profiling of osmoadaptive gene expression. *J. Biotechnol.* **106**: 255-268.

Rubio, A., and Downs, D.M. (2002) Elevated levels of ketopantoate hydroxymethyltransferase (PanB) lead to a physiologically significant coenzyme A elevation in *Salmonella enterica* serovar *Typhimurium*. *J. Bacteriol.* **184**: 2827-2832.

Saenz, H.L., and Dehio, C. (2005) Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. *Curr. Opin. Microbiol.* **8**: 612-619.

Saldanha, A.J. (2004) Java Treeview - extensible visualization of microarray data. *Bioinformatics* **20**: 3246-3248.

Sambrook, J., Fritsch, E.F., and Maniatis, T., (eds) (1989) *Molecular cloning: a laboratory manual*: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Savka, M.A., Dessaux, Y., Oger, P., and Rossbach, S. (2002) Engineering bacterial competitiveness and persistence in the phytosphere. *Mol. Plant-Microbe Interact.* **15**: 866-874.

Scheidle, H., Gross, A., and Niehaus, K. (2005) The Lipid A substructure of the *Sinorhizobium meliloti* lipopolysaccharides is sufficient to suppress the oxidative burst in host plants. *New Phytologist* **165**: 559-566.

Schnarr, M., Oertel-Buchheit, P., Kazmaier, M., and Granger-Schnarr, M. (1991) DNA binding properties of the LexA repressor. *Biochimie* **73**: 423-431.

Shamseldin, A., Nyalwidhe, J., and Werner, D. (2006) A proteomic approach towards the analysis of salt tolerance in *Rhizobium etli* and *Sinorhizobium meliloti* Strains. *Curr. Microbiol.* **52**: 333-339.

Sharypova, L.A., Niehaus, K., Scheidle, H., Holst, O., and Becker, A. (2003) *Sinorhizobium meliloti acpXL* mutant lacks the C28 hydroxylated fatty acid moiety of lipid A and does not express a slow migrating form of lipopolysaccharide. *J. Biol. Chem.* **278**: 12946–12954.

Shaw, P.E. (2002) Peptidyl-prolyl isomerases: a new twist to transcription. *EMBO Rep.* **3**: 521-526.

Shea, J.E., Santangelo, J.D., and Feldman, R.G. (2000) Signature-tagged mutagenesis in the identification in virulence genes in pathogens. *Curr. Opin. Microbiol.* **3**: 451-458.

Shoemaker, D.D., Lashkari, D.A., Morris, D., Mittman, M., and Davis, R.W. (1996) Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat. Genet.* **14**: 450-456.

Simon, R., Priefer, U., and Pühler, A. (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Bio/Technology* **1**: 784-791.

Skorupska, A., Janczarek, M., Marczak, M., Mazur, A., and Król, J. (2006) Rhizobial exopolysaccharides: genetic control and symbiotic functions. *Microb. Cell. Fact.* **5**.

Smith, B.T., and Walker, G.C. (1998) Mutagenesis and More: *umuDC* and the *Escherichia coli* SOS Response. *Genetics* **148**: 1599-1610.

Soberon, M., Membrillo-Hernandez, J., Aguilar, G.R., and Sanchez, F. (1990) Isolation of *Rhizobium phaseoli* Tn*5*-induced mutants with altered expression of cytochrome terminal oxidases *o* and *aa3*. *J. Bacteriol.* **172**: 1676-1680.

Soberon, M., Aguilar, G.R., and Sanchez, F. (1993) *Rhizobium phaseoli* cytochrome *c*-deficient mutant induces empty nodules on *Phaseolus vulgaris* L **8**: 159-166.

Soedarjo, M., Hemscheidt, T.K., and Borthakur, D. (1994) Mimosine, a toxin present in leguminous trees (*Leucaena* spp.), induces a mimosine-degrading enzyme activity in some *Rhizobium* strains. *Appl. Environ. Microbiol.* **60**: 4268-4272.

Stanfield, S.W., Ielpi, L., O'Brochta, D., Helinski, D.R., and Ditta, G.S. (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.

Streit, W.R., Joseph, C.M., and Phillips, D.A. (1996) Biotin and other water-soluble vitamins are key growth factors for alfalfa root colonization by *Rhizobium meliloti* 1021. *Mol. Plant-Microbe Interact.* **9**: 330-338.

Summers, M.L., Elkins, J.G., Elliott, B.A., and McDermott, T.R. (1998) Expression and regulation of phosphate stress inducible genes in *Sinorhizobium meliloti*. *Mol. Plant-Microbe Interact.* **11**: 1094-1101.

Summers, M.L., Denton, M.C., and McDermott, T.R. (1999) Genes coding for phosphotransacetylase and acetate kinase in *Sinorhizobium meliloti* are in an operon that is inducible by phosphate stress and controlled by PhoB. *J. Bacteriol.* **181**: 2217-2224.

Tabche, M.L., Garcia, E.G., Miranda, J., Escamilla, J.E., and Soberon, M. (1998) Rhizobium etli *cycHJKL* gene locus involved in *c*-type cytochrome biogenesis: sequence analysis and characterization of two *cycH* mutants. *Gene* **208**: 215-219.

Tate, R., Riccio, A., Iaccarino, M., and Patriarca, E.J. (1997) A *cysG* mutant strain of *Rhizobium etli* pleiotropically defective in sulfate and nitrate assimilation. *J. Bacteriol.* **179**: 7343-7350.

Thöny-Meyer, L., and Künzler, P. (1996) The *Bradyrhizobium japonicum* aconitase gene (*acnA*) is important for free-living growth but not for an effective root nodule symbiosis *J. Bacteriol.* **178**: 6166-6172.

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

Trzebiatowski, J.R., Ragatz, D.M., and de Bruijn, F.J. (2001) Isolation and Regulation of *Sinorhizobium meliloti* 1021 Loci Induced by Oxygen Limitation. *Appl. Environ. Microbiol.* **67**: 3728-3731.

Tucker, B.J., and Breaker, R.R. (2005) Riboswitches as versatile gene control elements. *Curr. Opin. Struct. Biol.* **15**: 342-348.

Twyman, R.M. (2004) *Principles of proteomics*. New York: BIOS Scientific Publishers.

Uhde, C., Schmidt, R., Jording, D., Selbitschka, W., and Pühler, A. (1997) Stationary-phase mutants of *Sinorhizobium meliloti* are impaired in stationary-phase survival or in recovery to logarithmic growth. *J. Bacteriol.* **179**: 6432-6440.

van Rhijn, P., and Vanderleyden, J. (1995) The *Rhizobium*-plant symbiosis. *Microbiol. Rev.* **59**: 124-142.

Vance, C.P. (2001) Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world of declining renewable resources. *Plant Physiol.* **127**: 390-397.

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

Vincent, J.M. (1970) A manual for the practical study of root nodule bacteria. In *IBP Handbook No. 15* Blackwell Scientific Publication, Oxford.

Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2004) Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet.* **20**: 44-50.

Walker, T.S., Bais, H.P., Grotewold, E., and Vivanco, J.M. (2003) Root exudation and rhizosphere biology. *Plant Physiol.* **132**: 44-51.

Werner, E., Holder, A.A., and Hoheisel, J.D. (1997) Growth and storage of YAC clones in Hogness Freezing Medium. *Nucl. Acids Res.* **25**: 1467–1468.

Wilkins, M.R., Williams, K.L., Appel, R.D., and Hochstrasser, D.F. (1997) *Proteome research: new frontiers in functional genomics.* Berlin Heidelberg: Springer Verlag.

Wilson, K.J., Sessitsch, A., Corbo, J.C., Giller, K.E., Akkermans, A.D., and Jefferson, R.A. (1995) beta-Glucuronidase (GUS) transposons for ecological and genetic studies of rhizobia and other gram-negative bacteria. *Microbiology* **141**: 1691-1705.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., Chu, A.M., Connelly, C., Davis, K., Dietrich, F., Dow, S.W., Bakkoury, M.E., Foury, F., Friend, S.H., Gentalen, E., Giaever, G., Hegemann, J.H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D.J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J.L., Riles, L., Roberts, C.J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R.K., Veronneau, S., Voet, M., Volckaert, G., Ward, T.R., Wysocki, R., Yen, G.S., Yu, K., Zimmermann, K., Philippsen, P., Johnston, M., and Davis, R.W. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**: 901-906.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**: e15.

Yeoman, K.H., Delgado, M.J., Wexler, M., Downie, J.A., and Johnston, A.W. (1997) High affinity iron acquisition in *Rhizobium leguminosarum* requires the *cycHJKL* operon and the *feuPQ*

gene products, which belong to the family of two-component transcriptional regulators. *Microbiology* **143**: 127-134.

Yuan, Z.-C., Zaheer, R., and Finan, T.M. (2006a) Regulation and properties of PstSCAB, a high-affinity, high-velocity phosphate transport system of *Sinorhizobium meliloti*. *J. Bacteriol.* **188**: 1089-1102.

Yuan, Z.-C., Zaheer, R., Morton, R., and Finan, T.M. (2006b) Genome prediction of PhoB regulated promoters in *Sinorhizobium meliloti* and twelve proteobacteria. *Nucl. Acids Res.* **34**: 2686-2697.

Yurgel, S.N., and Kahn, M.L. (2004) Dicarboxylate transport by rhizobia. *FEMS Microbiol. Rev.* **28**: 489-501.

Zehr, J.P., Jenkins, B.D., Short, S.M., and Steward, G.F. (2003) Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ. Microbiol.* **5**: 539-554.

Zhan, H.J., Levery, S.B., Lee, C.C., and Leigh, J.A. (1989) A second exopolysaccharide of *Rhizobium meliloti* strain SU47 that can function in root nodule invasion. *PNAS* **86**: 3055-3059.

Zhan, H.J., Lee, C.C., and Leigh, J.A. (1991) Induction of the second exopolysaccharide (EPSb) in *Rhizobium meliloti* SU47 by low phosphate concentrations. *J. Bacteriol.* **173**: 7391–7394.

Zhang, X.-S., and Cheng, H.-P. (2006) Identification of *Sinorhizobium meliloti* early symbiotic genes by use of a positive functional screen. *Appl. Environ. Microbiol.* **72**: 2738-2748.

Zhao Y, Karypis G. (2005) Data clustering in life sciences. *Mol Biotechnol.* **31**:55-80

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

G. APPENDIX

TABLE G.1. CLONED SIGNATURE TAGS

Tag ID	H-tag sequence	K-tag sequence
1-10A	GCGCGATCTCTGTGAACAATGAGA	GGGAGTTCTAAACGCTGGCTGACA
1-10B	TGAAAGGCCAGCGGACTGAATAGA	TAAGTAATCACCAGCCAACCGGCA
1-10C	TCCGAATGCTAGAAATTCGGGTCA	CTTTAGATATAGCCTGGCCGGGCC
1-10D	TCCGGTCGAGCGGTTTTTCTTACT	CTCAATTCACGAGCCACAGGAACA
1-10E	GCGGGTACACTGGCTAAGAAGCCT	CGGGCAGCGCTCAATAACTTTTTT
1-10F	GGCGGCTTACATGTTGGTTCAGAC	CCGGGAAACCACCTACTACGAAGG
1-10G	ATGATGCAAAGCTTAAGTTGCGCG	CCCCCCCCTGGAGTATACCAGTCT
1-10H	GAAGTAGGGTCCCCCGGACATTCT	CGCAGGCCCTTTGCTTTTTGTACT
1-11A	ATTGACCTACTCCACCTCTGCGCA	TCATTCCCCCGTCTTCTACCCAAG
1-11B	ACGCAATCTAAGCGTGCATTACGG	CCCGCCGGTGCATACATAATACAA
1-11C	CCTCTCCCAATCCCACACGTTCTT	GGCCAGTCGCCTCGCTGAATATAT
1-11D	AGCGCCGAACTTCGGTAACATAGG	AACTGACTTGTTAATGCCGACGCG
1-11E	GAGGAGTGAGTTGAGCCGATTTCG	GTACGAGATGCTGCCGAAAGGCTA
1-11F	CGCCTTACACACGCACGGGTATAT	CGCATGTCTTCATTTAGTCCCCCA
1-11G	TGCGGGCGGGGAGTAACTAGTAAG	TCGGTCTATGTAAATGTTGGCGG
1-11H	ATATTTTATATTGTCGCCGCCGCG	ACAAGTTGGTGTGGGTCGGAAATG
1-12A	ACGGTGGAATATGACACCCACTGC	ATCTTTGTTGGAAACTGAGGGCCG
1-12B	TCGCGTTCTTGGATGTCCAAATTC	CCTACCAACTCGACGACGAAGTGC
1-12C	TCTACAGCGGTCACTCTGACAGCG	GGCGCACCAATATCTATCGGCTTT
1-12D	CCCCCGCTCGGCGTCTTAATATAT	TAGGACAGTCCAAGCATAGGCCGA
1-12E	CGCAAACTGTCTAACGGTATGGGC	TCGCCCTGGGTTTCTACCGATTAG
1-12F	ATTATGACCCACGTGAACACCCCA	TCTCTCTGCCAGCATGAACCGAAT
1-12G	GTCATTGTCCCATCTTCGGCCATA	CGTCACTGGAGATTCTAAAGCGCG
1-12H	CGTGTCAGCTCCATGATTGTCCTC	TATCGCGAGGAGATGGAAGTGGAG
1-1A	GGGCGTAGCCTCTAAGCAACCTTG	CGCATCCTTTCACCAGGCCTTAGT
1-1B	CGCGCGGGGGTTTATAAGCATATA	GTATTGGTCACGCATTCCTTCCGA
1-1C	GGAAAATATGGGTCCACGATCGCT	TCCGGGTATTTATCACCCGTTGTG
1-1D	CGCCTATCTGTCGATCTAGCACGC	TAACGCCAATAAGCAATTCGTGGG
1-1E	CACCAGATGAGTGCAATCGCTCAT	TGGGAGCATCGTTCTCTTTTCTGG
1-1F	TTCCGATGGGTCACTATCGGTGAC	TGGGTCCCTGTTTAGACGAAGCCT
1-1G	TTCCCGGCCTACCTTCCTACATGT	CGCAGCGGCCTCCCATAGTAGTAT
1-1H	ACGCCCGCCACTATTCAATCATTT	TCCAAATGGATTCACCGGAGAATG
1-12A	TCGGCGTACTTATCAGGGCCTTCT	CATAAGAGGCGGGATCGGAATCTC
1-2B	AGACATCGAGATATTCCGGCAGGG	CCACATTTTGTCCTTTAGCGTCGC
1-2C	CGTAGCGTTGATTGGAATAACGCG	TGCCCAAGCGAAGTTAGTCACTCC
1-2D	TACTAAACCTGCGGCTGCACGGTA	CCGCATGGCTTACCCATACGACTA
1-2E	ACTTTCGTCGGCTGTGAGAATCCA	TCAGGTGAACATGGGAATGGACAC
1-2F	AACGATATTCGCGTTAAGGGTCCG	GGGGTTCCAGCCGATCCTATAGTG
1-2G	GTGGATGGCAGGCTAAATGGGTCT	GCCCGTGTGGTTGTTCCCTTTACT
1-2H	AGTACCGTCCCATTACTCCGAGCG	TAATTCAATTCCGTGGTCAACCGG
1-3A	GACGCGTTAGATGCACAGGAGGAC	CGATCTGGCCTCTCTTCGGTTAGC
1-3B	CCGAATCGTCGGAATTACTTTGGC	GGCCATTGCTGTTCTACTGGATCG
1-3C	GTCTGGGGGCAGTAAACGCGATAT	TGGCATAGACAGCAGAGCGAAGGT

1-3D	TCAGGGAGTGCATTCGATTCACTG	ACGCTGTGGAGGTTTTGCGGTAAT
1-3E	CGCGGGCCGAGATAGTATTATTGC	TGGGGCCGAATAGTTTGGAGAAAG
1-3F	AGCATTAAGAGGGGTTGATCCCGA	CCGCTCCAGAGGTTAAGACCTGGT
1-3G	TGGCAGTACATGGACGAAACAACG	GTCCTGGCTCAGGGCGTTCTATTC
1-3H	GCGCCTAAAGCTCAAGGTGATCAA	TGGGAAGAGACCCTGGTGTACTGG
1-4A	TATGTTCCCGCAGTAGGCAAGCTG	ATCTCATGCATTCTAAGGCACGCG
1-4B	GGGGTTGTTGGTGCCGACCTATAC	GCGCTTACCGATGACACAAATTGG
1-4C	CAACTGGACCATACCAAATGCCGT	AGCTGCTGGGGTCTTGACTCCATT
1-4D	ATGTTAAGCTGATACCTCCCGGCG	CAGTTTGGCTGCATGGTTAGTGCA
1-4E	GCGCCATGAGGGTGAGAACGTATA	CGGCGCTGGGCAAATAATAAAATC
1-4F	TATATCCGTCACCCCACCAGTGGT	CTGCCCTGTCTGCGTCTGCTATTT
1-4G	GCGGCGCTTAACCAGTGAAGACTA	TGCCACATTCACAATGTGTAGCCC
1-4H	AGTGAGCCGGCCGCTTCGTATATA	CGGATAAGCAGTGACTCGAGCGTC
1-5A	CGCGCATGCTGTTACTCGTACAAC	TTAACGGATACCCTCACATGGCCA
1-5B	CGCACTCACCAGTAGCTGTGCAGT	CGTCCCGAAAAAAGCGTCAGAATT
1-5C	GAGTCACCGTGGCAAATCGGTTAA	GCTGCCTGTCTAGTCGGTCAAACG
1-5D	CGCGGCTATCTTCTGCACTTTCAT	GCGCCAACAGTCCCTAATGTATCG
1-5E	ACGCAACTCGTAGCGGAAAGAGGT	GAATCGGGGTCCCAGTAGGGATTT
1-5F	GTTATCCGATTATTGGAGGGGGGC	GCGAGGACGTCACATCGTGTAAT
1-5G	GGTCGCCTTATACATGTCCACCCA	CTTGCCGGTGGTCTGTGGTACCTA
1-5H	CTGCGGGGAAATGGAAATACTTC	TGGCGGAAGTTAAACACGAAGCTG
1-6A	TGCCCTCGCATATAGTCATCACCA	CGCGGCAAAAGAAGTCTCGTAACA
1-6B	GATTGGCAGTGCATCCACTGATGT	AAGAGACATGGCAATTGGGCTGAG
1-6C	CTTTGTATCGTAACCGTGCCAGCG	CGGTTGCGGCTTGTTACTTGGTAG
1-6D	GGTGATGTGTATCGCACGTGACCT	GGGGGGGCCGTTCTTAATTAGATG
1-6E	TAATTATCCCCCGAGGATCGTCGA	CTGGTCTCCCCTTCGAGCGATAAA
1-6F	ACGGAGTACCACCGCTGGATCTCT	GATGCCGTACGCCCTCGACATATA
1-6G	AAGGCGCCGTAATACCTCGTGAAG	TTGCTAGTACACAGGTTGCCGCCT
1-6H	CAAGAGCCAGTTGCCTGATCACAA	CGGGAACTATATTCTTGGCTCGCC
1-7A	TGGGATCTTTTAGATGGTGCGAT	CACCCCCCACGGACCATGTAATAT
1-7B	TCCGCAATTGAGAATAAGCGTTGG	GGATAAAGTGGCCAGCCACATCCT
1-7C	AAAGAATCTGGACATTGAGGCGGC	TTTAAAGAAATCAGCAGCAGCGCG
1-7D	GGGAGCAAACACGGTTCCACTATG	GTGGCATATGGAATACGAGTCGCG
1-7E	CCGCCGCTACTAGACCTCGTTTGT	TGTTATAAGGGCTGTACGACGCGC
1-7F	TTGACATACAGATCCGCTCCGTGA	TCTAGTACCATCGTGCCACTCGCA
1-7G	TCTCCCCGATGCAGAGTCTCAACT	TAACCCATCCGTAGACGGAGTCGA
1-7H	ATCCTGCCAGAAATAGTCGCTGCC	CGCGCGTCGGTGTATAGAAATAA
1-8A	CGTCATGATCGTGTGTCAATGTGC	GCACCGCTAGGTTACGTCTTGGTG
1-8B	TCCGCGTCCTAATATGGAGGTGAA	GCCTGCACCTCAACATAGTCACCA
1-8C	AGAAGACCCAAGGGGACAAAGGTG	TTAAAGCATTCAACAGACCCCCGA
1-8D	CTGGACCGCTTTACTTTTTGGCGT	ACGGCGAGCAGTAGTGGATTCGTA
1-8E	GGGCGCGTCCATTATAAGTTTCGT	GAGGATGGGAAACTATCCATGGGGG
1-8F	CCGCTCCTCTTGTACTGACGCAAT	CCACAGTGCAAAGGCAAGGTAAGC
1-8G	GATAGCCTATCAGCAATCGGGCGT	AGCGTGACAGAAAGGTTGAGGCAC
1-8H	TCTGCATGACCTTGGTCTGCCTTT	CACGGCGGCTAGCGGATTAGTATC
1-9A	TAGACAAGTCTCAGGCCGAGGCAT	ATTACAATGCCATGGACCCCATTG
1-9B	ATATTGCGAGGTGTATTGCCGGGT	CGCGTGGGTTTGGTAGAACCTGTA
1-9C	CGTTTTATTCCGTCTCTGGTGCCA	CTTGATAGACGGGCGTTAGCCATG
1-9D	TCCCAAGCCTCTGTGCTGATGAAT	CGAGAAAATGCTGGTGGTGGCTTA
1-9E	TCCTTGTAAGTGTCAAAAGCCGCG	TCAGATAACGTACGCGGTTGGACA

1-9F	GCCCGTTGGCACTTACTGTTGCTA	ACACAGCTGCGCTCATAGACGATG
1-9G	AAGCCATCTGCCCGTACACAAGAA	AGACAGAGGGTGCAGTACGCTTGC
1-9H	AAATCCACCGCGGCCAACATATAT	AATTTTGGGAGTGGCGTGGTGAAT
2-10A	CCCCCGTGAGGCGCTGTTTATATA	CAACAGTCGGGCAGTTACCCTCT
2-10B	ATACTCTCGCAGGCACTTTGCGAA	GCCGTTCCTCCTCGGTTTACAATC
2-10C	GCGTATTGCTCTGTTGGGGTGGTA	GGATAACCACGGGGCTCGGTATAA
2-10D	AATACCTGGGGCTGACCTAGCCCT	AGCGGTCTCCGAGTAATCTTTCGG
2-10E	CTGGCTTCAAGGACCACGAAAGAA	AGGAGTGTTCTTCTTTGGACCCGG
2-10F	TTCTGGAAGTTTACTTCACGCGCG	TGAGTAAGGCAATTACCCCGGAGC
2-10G	GCCCCTATCAATCTCTGGACGCTT	GATCTGAACTCTATGGCTCCCCGG
2-10H	CGGCCACACCGTGCGAGTATATAG	GCCTTTTTTAGACAAAGCCGTCGC
2-11A	GTGTTGGTACGGAATCATCTCGGG	CTGTCTCGGCGAACAACGTTTCTT
2-11B	CGGATTCGATCCACTCGTCTGATT	TACAGGAGTGCTCCATCCGATGTG
2-11C	TCATAATCCTCTTCAGTGGGCGGA	TACTTAACGATACCGCCGAGTGCG
2-11D	CCAAAGGTCCCGTGTTGAAGTGTC	CGGCTAAAGAACGAACGGGGTAGA
2-11E	TAATCGGCAAACCGATGCTTTCAC	CGCAGTGAAAGAAAATCTGCTCGG
2-11F	TGCGTGAGGTGTTGTGATTTACCG	TTCCGCGACTGTTCACTAGGTGCT
2-11G	CGCGGGCATATTGAGTACTACCGA	ACCCGGTAGGGGCAATGGGTATAG
2-11H	AATGGATGTTGGAGAGGATCACGC	CCCCCCGACCGAATATAAGTGCTT
2-12A	ATGCCGCAGTGTAGAAGTGGCATT	TTCCTGCGGTAGGTGACGTTGACT
2-12B	TGTGAGACTGAAATAAGGGGGCCA	GCGTGATGTCCGTAACGTTTGTCA
2-12C	TTGCCAATGCGGTCGTAGGGTATA	GAGGTAGTCTCTCTTGGTCGGCGG
2-12D	TCGGGGTTACGCAAATTCCCTAAC	TGGGACATTATTGTCTTGGGCGTC
2-12E	CTACTTCTATTTGATGCGCCGGCA	GCCGCAGCCTCAGAGATGGTAATA
2-12F	CGGCAGAGGAACTCCGTAAAGAGC	AGCGCTGGAAATCCGTACAAGGTC
2-12G	AATTGTACTAAAGATCGCCGGGCG	GGTACTCAGACAACGCCTGCTTGG
2-12H	GGGTTCGGTCATTATGAACTCGGG	AGGGGGGTCATGGGCATACAATAG
2-1A	CCGTCGTTAAACGACGCTTCCTTT	ATTCTCGTATGTCCTCAACGCCCA
2-1B	CAACTTCAGATATCTGCGGCCC	GCCTTATGCAATACCGAACGCTCA
2-1C	TGCCTTCTCAGGTAGATGCGGAAC	TCGGATCCTCATAGCACCGCATAC
2-1D	CAATCGGTCCATCACCCTAATCGA	TGTAAATCGACTTCATGGCAGCCA
2-1E	GGACATCTTCATGACAAGTGGCGA	CGGCGTAGGCTACCAATCGATAGC
2-1F	ACACATGGAGATCAGGGACGCTGT	TCCCAGCTGATTGTTCTGAGAGCA
2-1G	AGGATAGTCAGTATCGACCCGGCG	GAGCTGAATTGGCCTGAAGGATGA
2-1H	ACGAACTGCACTCTTAACAGCCGG	TTCTGCAGTTTCACAACGACGACC
2-2A	CTCTCTCAATACTCGGTCGTGGCG	GGAGCTATGTTTCGCCCATTCCAT
2-2B	CCCTCGGTCAGAGACTTAAGGCGT	CGCGGACCCAAATCTATGCTAGGT
2-2C	GCAGAATAAATAAGAGCGCCCGGA	ATTAGCCCCCTCTAAATTGGGGGG
2-2D	CGTGCGCCCTTACATCTAGTTCCA	CCTCATATCTTCCAGTTCACGCCG
2-2E	TCGGGCCCTATAATACCCCGACTT	TTAATCCTACGGTGTCTGCAGGCG
2-2F	TACGGAGAGTGGGGTTATGCACGT	GCCGCGTCAGCCTCTTTAGGTTTA
2-2G	CTAAAGTCCACGTTGCCGACAGCT	GGTAGAATCGCAGATCACTCCCCA
2-2H	CCGGGTTAAGGACTCCTAACGGGT	ATTAAATTTTTAGGGGCGTCGCCG
2-3A	GCGACAAGCCCGCGTCTTTAATAC	AGCCCCTAGAGAAACCTTCATGCG
2-3B	TCCCTTAGTTTAGGCCCCCAAGGA	CGACGCATTTTGATTTCCGTTTGT
2-3C	TAGCTTAAACCCGGTTTGGCGATC	CCTCGACCCAATCATAACCGTCTG
2-3D	TCGGATAGGTAAAGATGCCGGTCC	AGACCACTTCACAGGACACTCGCC
2-3E	CGCCGCGCTATAACCCTAGTATGG	GACATCTACATGAAAGCGAGCCGC
2-3F	CTGGCCCCACTGGCACTATCAGGTA	GTGCGCTTCTCCTAGTCGCACATT
2-3G	ATATATGGCGAGTTACCGCCCCCT	TATGGTTATTCACGGAAAGCGCCC

2-3H	AGGTATATTCCTGTACGGCGGGCA	AAGGCTACAGACACACAATCGCCC
2-4A	ATTTTCAGGTGCAATTCAAGCCCC	GGACGACGCAAGAAAGAGTGGTTG
2-4B	TTAGATTTTGACGGGTTTCGGCGT	CAAACGTAGGACGACCGTATTCGC
2-4C	GTTTCGGTGGTGAGCTTAGACCCC	ACAAATATCCACCATGGGTTGCCA
2-4D	ATCACAATAACGTTCAGCGGCGAG	TCACTATTTTTCACAGCACGCCCA
2-4E	CGATTACGAGTGTGAGGTTCGCCT	CCCGACCTGTAATGTAAGTCCGCA
2-4F	CGCGCGGACTCTAGAAGTTGTCAT	CCGGGTGAGCCTGATAGAAAACGT
2-4G	TACAAGGCGCAAAGATCATACCGG	TGAATTCCGCGTAGTGCTCTTTGG
2-4H	TGGGATATCTAGAAAAGGCCGGCA	TCGGCCCACATAATTTGGGATAGG
2-5A	CGGCGGCAGGTACATTTAAGTCAG	CGAAATTCCTCCCACGCTCACATA
2-5B	GGCGGGGCGAAGTCTTAGTACTTG	GCGGGGGAGGCTATGGATTTTAGT
2-5C	GGCCGTTTACGTGTAGGTGTCACC	GACGTCAAGAATGAAGGGCGTCTG
2-5D	CGGGGCATGATCTACAGTAACCGA	CCTGTACACTTGTAGGTTGGGGCG
2-5E	CCTTTTGTAGGGTTTTATGCGGCG	CGGGCTCCCATTTCATGGACTACT
2-5F	CGGGCCACAAAAAAACATGTCATT	CCGGCTATGATTCTTGCTCCAGGT
2-5G	CGTTGGCCCTTAGCAGGTCATATG	AGTCGGTAATACGCTAGGCTGCCG
2-5H	CGGGCTACGATGACTCAGATCCAT	GACAAGCAAATATTGGGGCCTCGT
2-6A	CGTAGATTCCCAAGGTTCGTCGCT	CCGAGGCGTATCAGAGTTAGCGAA
2-6B	GAGGCTTCACCGCAATCCTGCTAT	ACCAAAGGGGCTCTTCAAACATGC
2-6C	CGCACCTTGTCGTATTAAGGGCAG	TACGTACTGTCTGACGGGGTGCAA
2-6D	CTGGTGTGAGATCGTAGTCAGCCG	GCCCGGCTGAAACCAGGTACTATG
2-6E	CGATGGGCGCGAATCCTTAATTTA	TTAGCCGCTGTCTTGCTGCAAGAT
2-6F	CAAAGGAGCAACCCCGCTTCTAAA	CGAGCCCGACACAGAGTGCTTACT
2-6G	CTTCACGAGGTATTACGGCGCCTT	CGGGCGAGAGAAAAGCCTATGAGT
2-6H	GCGCGGATCAGTATGGTTCAAAAC	CGTTATATGTGCGTAAGCAGGGCG
2-7A	ATACAGGGCGGCGCAATAACCTAA	AGCGCGCAGGGCATAGGAGTATTA
2-7B	AATACACGACTCAACTGCGTGCGA	ACGCCCGGTCATAGGGTTCAATAG
2-7C	GGCAAGTCAATAATTGTGAGGCCG	TCTTGATCGCCTGGCGAGAAATTA
2-7D	GCGCGCTACATCTCCAATCTGGTA	AAGTTCTTTTCGGAGGGTACGCGA
2-7E	TGTCGTGTGGAACGGGCTGATATA	CCCCCGGCGTTTCACTACCTTAAT
2-7F	TCCCACTCGGAAGCTGTTATCCAA	CCGGCACGCACCACTAAATTCTAA
2-7G	CCAAAGATATCAACACCCTGGCCA	CAGATCGCTGGACCTCCTCTGTTC
2-7H	AGTAGGAGATTACGTTCGCGCGCT	AAGAAGGAAGGGGGAAACGGTTTG
2-8A	TCCGATCTAAAACGGGAGCAGGAC	ATCGAACCACATGTACCCCCTGTG
2-8B	CGCGCGTACAGTGTTTCGAATTTT	GTAAAGCTTGTGCACGGGCAGCT
2-8C	AACGTAACGCCTCCGCATGTGATA	CTATGTCATGCTTGCATTGGAGCG
2-8D	TAGCAGCTCATCCACATGCAAGGA	GGCGAGGATTCAAGTGGTTTGTGT
2-8E	GATTTGAACGGTGAGAGTACCGCG	CGCGAGGGTGTGATACTAGATCCG
2-8F	AATCGCGGCCCGGTTAATTCTTAC	TCCAATGGATCCTGGTCCCTTTTT
2-8G	CCCGGCGGCCTAGTAATTGATTTT	GCCAGCCCGATGAACTGAGATCTA
2-8H	GCGAAACGTCAATATCCGCCTTTC	CCCCCCATTCTATTGGCATGAAAC
2-9A	CTATGACGTGGCGCCTACAACTCC	GGCGACCGGGATTCGGTACTAATA
2-9B	CAGACCAGAGGATACATCGGCACC	GCGCATTTATTGATACCAGCGCAC
2-9C	TACTTAGCACGAATCGAGGTGCGC	CGGTCTGATGGCTGAGAGCATTG
2-9D	ATAGGACGGCTGCGAATCATGCTA	TCGCACAGTATTACTCACGCGGAC
2-9E	CCGCGACCGATCTTGGTAACACTA	TCGGGTTCGAATCGGAAGAACTCT
2-9F	ATATATCGGACGGATGCCCGTCAT	CTGATAGCATGCAACCGGATCGAT
2-9G	CTCAGTACGAGCACTGTCTTCGCG	CTATACAGGCACGAAGGATCGCGA
2-9H	CGCCTGTTTTAACCCAGTGATTGC	GGCCTCATTCATTCACGACATCCT
3-10A	GAGTTTGACGTATAGACCCGGGGC	GATGGCGCCCACCCGATTACTATA

3-10B	TTCGAACATTCTCATACGGTGCGA	CGCCGGAAGCCTGTAGACCTTAAA
3-10C	CCCGGGAGTTTAAATCCGACACTC	AGTATAAGCGGATGGCTCGAACCG
3-10D	CGCGCCTGGGGGTCGTACTATATA	CGCTTCACAAGGGTTTGCGATTAA
3-10E	TCTTATCGTCACAGCGAAGAGGCC	TTCAACGGTGTACGTGACTGCTCC
3-10F	GACCAGACTGAGTGGAATTGCCGT	CCGTATTAGTCAATGCCCTTCGCC
3-10G	ACCTACAGTGAGGCAAACGTCGGA	TCCAACGTACCGGATTGGTGGTTA
3-10H	CGCGCGATCGATTGAAAGTAAGTC	CCAATGAGGACGGAGCCCAGATAT
3-11A	CGCGCAATGGTAGCGGGATATAAT	GCAGAGGGGAAGAATGGATAGGCA
3-11B	CTTTAACTAGAATCGCGGCGCGC	GACCTCGCGTAGGGGTCAGCTAAT
3-11C	ATGGACCTTCCACCTGTTATGCC	CTATCTCAAGCTGGCCCAATGGCT
3-11D	TGTCGTCGGGCGCATTCACTATAT	GGCAGCCGAGTTAAGGTGGAACTT
3-11E	ACGGGATAAGTTGTTAACCCCCCG	TCTGGATACGCTGCATAGCGTCAA
3-11F	CGTGGAAAGATAGTGTCGCATCGA	CGCGTGCTTTTCGATTATCTAGCG
3-11G	TTGCGCTGTACTCGAACGTTGATC	GGGGAGGGTGGCTAAAATACAGCC
3-11H	GGCAGAGTGGAGTTGAACAGGCAT	CGTGATGAAGGCAGTGATACAGCG
3-12A	GCGAGACCCGCGAATGTGATTATA	CGGGCCGTGACGGGATTATATTAG
3-12B	CTAAAACTTGATCCTTGACGCGCG	TGCCGACTTGCGGTTGCTACTAAG
3-12C	GTTCATTTCAGGCAGTCAGGTCGC	ATTGAGGAGCTCGGGTTTGTCGTT
3-12D	TCAGACTCTACTCTCTGTCGCGCG	GCGCGAGTTGACGGTATCTACACA
3-12E	TCCCTCCCGAAGCCGACATAGTAT	AAGAGACTCGAAGCGAGGGGTGTC
3-12F	CGGGACAACAACTCAATCGAGAGG	GGCTCATCTCTTATCATTGCGGCA
3-12G	CAGGTTCCTCGCTACTCCTGGAGG	TTAGTCGCGCTCCTTCTCCACTGT
3-12H	GAGTACGGCTTCATTACACCGCCA	CACTCAGGACTGGCATTTACGGCT
3-1A	TGCCACGCAACATTACACTGACC	GTGCGAGCACAGCAACGACTACAT
3-1B	CCCACAGAACAGGAAGTTGCTTGG	TATCCACAATAGGCTCCACGTCGG
3-1C	AATAGATTAGCGAGCGAATCGCCG	ACGGCTTAGCTGGAGGTATCCCTC
3-1D	GGCTCTATACCTGCTCGTGTGCGT	GACTTCAATTTTAAGCACCCGCCG
3-1E	CGAAGTCGTCGCGAGTCTGTACAA	CCGCCCCAGATGCCATTGTATATT
3-1F	CGCGGCGTCATGTGTACATTCTTA	ACAGGAACGTAGTCCGGTCCAAT
3-1G	CGTCTTCCCTTGTTATCTCGTGCG	AATTAGTGACGCCAGCTCGCTGC
3-1H	CTATCCGCGGTTTGATTAAACGGG	GCTCCAAATAGCGCTGTCGTTCAT
3-2A	GCGAGTGATATGCATGTTGTCGGA	AATCGGGAACCATCATGCAAATCA
3-2B	CGAGCGAGTCGTTACGCGAAATAG	GATCAGCCACCGAGGTTGAAACAC
3-2C	TGGAACATTGCACAACAGTGACCA	ACGGCCATCCGGTCAAGAGTAAG
3-2D	CGGTTAACGGCTGTAGGCTGAGTG	ACTAAAGCAGGAGTCGGATGTGCG
3-2E	TGGGCGGGTAGATTTCAGAACTGA	TAACCCGTTCTACATTCGAGGGGC
3-2F	AGTCTAGCAAATTAACGGGGCGCA	GGCGCTGAGACGAACGAAAGGTAT
3-2G	CACCCCTTCAGCAACATCTATGGG	TCCATGCCGTGCATCGTATAACTG
3-2H	CTAATAACGGCAAAGCAACGCTGC	CAACGCAATGTGAGTGTATGCGTG
3-3A	GCGTCTCCAGGCATGCGATCTATA	CTGAACAAACATCTCTCGGCTCGG
3-3B	ACCTGGCGTTGCATCTCAGGATAG	AGTCATTAACCTGGCACAGCGGAA
3-3C	TGCAAGTCCCTGCACTGACATGAT	CAACCCAGCAAACCTGAATGACC
3-3D	GGGCCCCGCACTTAATTCACTCTA	ATATAACATGAACGCCGTTCGCCA
3-3E	TAAGGCAGCAACTGCAAGTGGGAT	TGAACCCAATGCTGACAGGTTTCA
3-3F	TCCGGTCTACGAACGTGCCTAGAA	CGGATCATGTTCCTATGGGGGATC
3-3G	GAAGCCCCTCAACTACAAACGCCT	CGCTCGGGATTTCACGTATATGGA
3-3H	TCGGCCACGTGCACATAGGATAAT	CCCTCTGATTTTCAAACGACTGCG
3-4A	GCACTGTTCCGTAAGCTGGAAGGA	TCTTTGGTTGGAACGAATTACGCG
3-4B	TCGGTATCACCTGGGCCGTCTAAT	TATTCCGAATTCCTAGCGATCCGC
3-4C	TCATGCGACGATGGACGTCTTTTA	GATCGGACCTTACTGCATCAACCG

3-4D	CGCGCAAAATGAGGTTAGAAGTCG	CCGGGTGCGTTTTTTGGTTTCTTA
3-4E	TACACCGGGGTTTTTTCAAGCTCC	ATTCGCCAGTGCTGTCACCAGAAT
3-4F	CGGGCGTGTAAACACATTTATCGG	CGATGCAAGAGGTCTCCATTTTGG
3-4G	GATGATGCGAGAACGCAGAACATG	CTGCCGATGTTTTCGACTTTGGAA
3-4H	CAGCAAATGGGCCTCCTCCAGTAT	CGGTCGGAGAAGATGACTTGGAAA
3-5A	GCCCGCGGCTCATTATTTAGACTC	TACCTTTGCGGCCGAAGGGATTA
3-5B	TGCGGATTCAGGAGATCTGTCTCC	CTGTGTTACACGGAGTGCGAGGCTA
3-5C	TCGGGAAGATTCACGTCGAAGATC	CACGCTGCTCCTAGGCAAATTGTC
3-5D	TTACTATGCGCGAAGCGTAATCCG	GGGCCACCAACGAAATTAATCTGG
3-5E	TATACGCTCCCAGCGTGGTCGTAT	GCTTGGAGGAAACTCACGACCTCA
3-5F	TCTCGATCGTTTGGTTAGCGCAGT	GCCGGACACAGTTACATACGGGTC
3-5G	CGCAAGATGAGGTCAAATCGAGCT	CCTAGCCAATGATGGAACTGGCAA
3-5H	GCCACTAACGACCTTCTCGGTGAA	CTACTGCAATGGAGGGAATCGTGG
3-6A	ACCGCAACAGAAGTGCGAAGCTAG	TCCCCACATTGGATGACTGAACT
3-6B	GGGTGCTGTAGTACAGGCGCTACG	GGCAAGCGACTCACAGAAACATCA
3-6C	GCAGTTAGTTAACGCAACGCAGGC	GGTGGCATCCATCCTGTCACACTA
3-6D	CGCGCGTTGTACCCTAGAAGAGGT	GTACTAACCCTCTTTGCCAGCCGG
3-6E	GCCATGGGGTGCGGTTACACTATA	TGGGGGAAGCAGTAATGAGACGTC
3-6F	CCATGGCCGGTTCTGCTTGCTATA	GTCGACGCTGACCAAAGTTCATGA
3-6G	GTACATCGTTTCCACAACCAGGGG	GTAGGCGGTACGCACTGTATTGGC
3-6H	AGTCGCCGAATTATGGGTGCACTA	ATCTATTCGATAGGACCGCGTGCA
3-7A	CGCGCTGATCCATGGGCTACTATA	CGGCACATTTGCAACATAAGGACA
3-7B	AATAAGGCTAAGGCGCGCTTGCTT	GGCCGCATTTCAAGTCCTTTTACG
3-7C	GGGCGGCACAGAATTGTGTAAAAG	CAGAAGAAGCGTGAATTCGCCTTG
3-7D	GTACCCTTCCGCTTTTGTGCAGGT	TGCCTGCCGTGGTATGAAACAGTA
3-7E	ACTACCGCCATAAATTTGGTGGGC	TGCTTAACGTCTGAGGGATGACCC
3-7F	TCCCGCAAGAGTCAGTCCTACACC	TTATAGTGCAAGCGGGCTCTCTCG
3-7G	AAACCGGACAGTTATGGGACCTGC	CAAGCACGTAACTTCTGAAGGCGC
3-7H	TGGACAAGACGATCGGTGGGTACT	GGATCCCCTCGTCGAACTTGAAGT
3-8A	AGCCCATCTCAAATGGCCCACTAG	CCGCACTGATTGACGATCTTAGCA
3-8B	TTCCTGTCTCATGGTGAGACGGGT	CCTACCGTAGCTAGGTGGGGGTTG
3-8C	CGCTGCAACTAAAGGTGCTCCTGT	ACTGTCGGAAGTCGCAGTCATTCC
3-8D	GAGGGCATAAACCAGCGGCACTAC	GAGTGCCGTCATTGCCGAGTAGAC
3-8E	CGGTTCCGGTAAACTGATTCGACA	AATTTGCAGGGCCGATGAGTTTTC
3-8F	GCCCAGTTGTAATGGCAGACGACT	GGCCCCATGCGTGTTTCCTTTATA
3-8G	GCTAGGACATCGCTATGTGTGCGA	GGATGCAGGCAAAATACTGCGATC
3-8H	TCCGGCCAAGTACAGAAATGCAAT	CACAACTATCGTGGCCTGCAACCT
3-9A	GAAAACGACCGGTTGTTTATGGGC	GCAGGTGACAAGGATTTACGCCAG
3-9B	TGGAGCGTTAATCGAGTCGGCTAA	CAGAGGGACGGGTGTCTATCGTTG
3-9C	CGCTATCATATCGCTTATGGGGCA	CATTAATACAGACGTGCTCCGGCC
3-9D	CGCGCACATGATATTTCCCTTACG	GAAAGGGTTCGCGTGATACGATTG
3-9E	CGCCGAAGCTTGGTTATACCAAGG	AGCTCCCTTAAAGGCCTACGTCGC
3-9F	GCGCGGGTTAGAGGAAGACTCATC	TGGCGGCTCCGCTGATAACTTAAT
3-9G	ATATATAGGCGCGGCCTGGCTACC	TGGCCCGTGACACTCATTTAAAGG
3-9H	CCGGGGCTTACGATGTTACGTAGG	CTCCCTAGGTATTCCGCAAGGACG
4-10A	TATCAGTTGTTTGGTGAACGCGGA	CGGTTTTACGTTACATTTGGCCCC
4-10B	CTCAGGCGCCCTAAGAGTCTTGTG	AGGTCACTTAGCGATACCCCACGC
4-10C	TGCGGCACCTCTCATGTCTAATCA	CGAGATCCTTAGTCGTCAGTCGGC
4-10D	GAGCATCTTACCTCGCTCAACCCA	TCCCCGCGGACATAACCTTTAATG
4-10E	TTATGTCGTACCGTTGGGTCAGGC	ACGTCCTAGACGACACGCAGGTGT

4-10F	CGCGCCCGCTAACGTAGAGTAATT	AGGCTGATGTTAACGAACATGCGC
4-10G	AGATTGCCTCGACTCATGCGTCTT	ATGCGATAGGTTGACTGGTGTCCG
4-10H	TGGGGAATACTGGTCGTCTTCGTC	GTGCCTGGGTAGCGCTTGTTGTTA
4-11A	TGGCCGCTAAATCTCAACCAGCTA	CCGGAAACTTCTTATCCACGGAGG
4-11B	CCTCGGGACCGTAGATGACCAACT	CTCTTAGCGGACGCTCAAGTGTCG
4-11C	TTGTCGTCCGAGCCAGCTACCTAC	CGCTCTCCTTTACAACAAATCGCG
4-11D	TCTACTGAGACAATGTTGCGCCCA	TGAATCAAGCTAGCGCACGACTGA
4-11E	ACTCTAAGGAGCCCTGACGAGCGT	TGGGCCCTCCTTATCCTGTGTCTT
4-11F	GGACAGCTTTAACGGTTAGGCGGA	AGAGTCAAATTCACCCGAAAGCGG
4-11G	TCGCGGGGCTATACCATAAGGTGT	TTTGTTTTATCCTTGCCCGTGCAA
4-11H	GCAAACAGTTGCATAGGGGTGGAG	ATGGGAACGCTTCAGCCTTTATGG
4-12A	TCGACTAGCTTACTGGGGCCCTTC	CGGCGGGACTATCTGATTAACACG
4-12B	GTTGGTCCCCACCCCGAAGTAATA	TGACTGAGGCCCGGAGATAGGAAT
4-12C	CATAAACTCGAGAGACGACGGGGA	GAGCCGAAGTGATTTTTCCTGGGA
4-12D	ACCACTGACTGGGAGAAGTCAGCG	CACCAAAGCGCTGACGGAGAAATA
4-12E	TGGCTCTGTGGCAGGTGGATCTAT	ATTGAGTGGTAAGGCTGCTGTGGG
4-12F	ACTCATGTTGACCGCCCTCCCTAT	TGCGAATTTACAAATGTCTGCCGA
4-12G	CTCGCGCACCGGATGGTTAATAAT	ATGGTTGCCCGAACAAGCAGTTTT
4-12H	ATACAAGCCTGGTATGCCAAGGGC	CCGGGGTGTGCCCAATAAACTTAC
4-1A	CCGCGGCAGTAGAGATGCCTTAAT	ACGTCCAGGATACTGTACCGACGC
4-1B	ACCGACTGGTTTTTGGGGTCAACT	CTGCTCTGAGCCGTAGTCGACCAT
4-1C	GAACACCAGCGTAGCTCAGGTTGG	TATGTGTAACGGCGGTGCACAGAC
4-1D	GTTAACACCTAGAACCCGTGGGCC	GTTTTGGGCTTAGGGTTGACGCAT
4-1E	CTCAGTTGCGATGCGCGGTTATAT	CCTCTAGTTTGATGATGGGCACCG
4-1F	ACACGATATAAAGCGTGGCGGATG	TCTGCTGAACCCCATCCCATACAT
4-1G	CGCGGCATGAGCCTACTTGAGTAA	AGGCTCGCAGAGATAGATGGGTGA
4-1H	CGGTACTTCGATGAGCAGCCACTT	CGGCCGAACGACTTACGAACAATA
4-2A	GCCAGCAGGCCGTTGACAATACTA	GGAGCGGTCGCTGGGTAATAGCTA
4-2B	TGACACGTCCGCCCCTGAGTATAG	GTTTAAGACAGGTGGCTCTCCCGG
4-2C	CGCCGCGAAAGTCTAGTGCCTATT	AAGGGATGGCCAAAGCAGACAATT
4-2D	CGCTAGACAATCTCGTCCAGTCCG	TCTGAGGTCTGCGGTCCGTCTTAC
4-2E	AGATTGGGCGACATGGTAGAGGGT	GAGTAAGCCTGCTCTCCATCAGCG
4-2F	TATCTCGCAAGCAATGCAGCTGAA	TTGCCCCAAGAAGAGAAATGAGCA
4-2G	ACAAGGAGGAGAGTTTTGGAGGCG	GTAATGACAAAAGTCCCGGGTGGC
4-2H	ACGTCTGGTTACAGTACTCGGGCG	ATCCTCGACGGATTGAGACATTGG
4-3A	CATAAAGCTACCGACAGGACGCCA	CGCGAGCGCGTAGAGGAGTAAGTA
4-3B	AGAACTCGTTGGAGTACGTCGGA	AGACTGTGAGCTAGAACACCGCGC
4-3C	CACCGCGTCGCTTAGTCCTATTGA	ACCGGCGTCCGTATCTCTCCTAAG
4-3D	TGCAAGTTACTAATTTTCGCCGCG	CGATTGATTAGCTTGCAGGACCCA
4-3E	CGCCCCGGAATTTTGTGATAGATT	TTAATTGTCGGGAGCTGCCTCCTT
4-3F	AACACCACATATGGGGAAGGACGA	TAGGATTCTAGCCGGGACGGACAC
4-3G	TAATAGAGGCGTGTCATTTGCGCA	GTTAAATTATCTCCCTCGGCGCGA
4-3H	ACCAATTGTAGTTTCGGGGGCCC	GCACAGTGAACGAAGAAAGGCACA
4-4A	GACGGACTATTCTCGAACTGGGGG	GCCGTGCCGATCAAATCATACATC
4-4B	GATTTACTCAGCGGGGAATTGGGA	TATGTGAAGAGATCTTGTGCGCCG
4-4C	TAAGCCCTTCTTCCAAGCTCTGGC	ATTAGAGGTACTGGCTGGGACGCG
4-4D	TAAAGCCCACCTTTCATTGTGGCA	CCACCCATGCACCTTCAGTTGAAT
4-4E	TCGGGGCCATGCAATGAGTTATAG	TTTACCCAACTTGCCCCGAGAAA
4-4F	CGCGACGGTAACGAGAAGTAACCA	TACGGAAGGTGTCGTAACTGGCGT
4-4G	CTCCGCACTCTCGGGCACTCTATA	AGGCTCCCTCATCATGTACAAGCG

4-4H	AACTTTACCGCGGGAGGACTTTCC	TTCGTGAACCCGGAGTCCAGACTA
4-5A	CGCACTAGTCTCGAGGCGGTTAGA	CCGCCGTCACTCAACAAGGTATCT
4-5B	TCCCTGGTAGCTATACTGCTGCGC	CGGCCTTGTCTTTTGGGTTTTCAT
4-5C	GCCAGCAGGCCGTTGACAATACTA	GGAGCGGTCGCTGGGTAATAGCTA
4-5D	GAGGGAGGTCGATGTTCCGGATAG	CGTCGGGGATGTTTAGTTTTCCGT
4-5E	AATCTGCGTAATGGAGCCGGAACT	TGTAAAGTCGGTGCTTGAGCACGA
4-5F	TCGGCGGGGGGAGAATATTAAAAC	GGGCGAATGAGTCCTCACAGGTAA
4-5G	ACGGAGCAGAAAGATGGAGTGTGG	TTGACAACCTGCATGCCGATACAG
4-5H	TAGCGTAGGTCGTAAGATCCGGCA	ACTAGCAGGGAGGCAACGCGTAAG
4-6A	GTCGTTGGCTACGTACCTCGGCTT	CCGGCAACAAAAATCCTATTCGCT
4-6B	TGAACGATAAAAGTGGGCTCGACG	GTCAACTTAGTGAGGAAGGCCGCA
4-6C	GTTTACCTTATCTGCAAGGGCCGC	AGAAACGGCAGCTGGATGTGTGTT
4-6D	CGAGGTGGTCACAGGCCTAGACAT	GGAACGTCGACTTAACACAGGCGT
4-6E	CGGGGGTAAAAAGTCTCCGTTTCA	TGTGCCGAGAATCACACTGAAAGG
4-6F	TAATAGACGTCGCGCAGTGGAAGC	GTTCTCAAGATCACAGGGTTGGCG
4-6G	CGAGAAACTAATTGCGTATCGCGC	GGCGGGAGCCGTTGCATAGATAC
4-6H	TACTACCTGGAAGCCGGCATTGGT	CAGGAGGCTCCGGAAATTCTTCAG
4-7A	CGCGCGAGAAGCATACTCTAGCAG	ACGCGGATGCTGTCCGAAATAAAT
4-7B	GCCGGCGAGATCAATGTAGGTCTT	GGAAGTATCTGAGGAAATTGCGAA
4-7C	ATAATGGCCGAGAAGCCGCTCAT	GGTGACACCGAACAGTTTTGTGTT
4-7D	AGAAATCGTTACAGGCCAGGTGCC	TCGGGCGACGGAGGTAACTTTAAC
4-7E	ACGCTCTCACATCTGGCTAGACCG	GATTAGCAGATAGCGTTTGGGCCG
4-7F	CAACTAGGCGTCAAAGATTTGCGG	TGGGTGGCATATTCATGTGACAA
4-7G	CTCTCCGGACAACTGCCATAACCA	AAACTTTGACTGCAGCACACACGG
4-7H	CTGTGCTATCCAATCGACTCGGT	TTGGGTGTGGTGATAGGTGTGACG
4-8A	ATATGACGGTGCAGACTATCGCGG	CGCAAAAATGTTTTAGGTGCACGC
4-8B	GTTGCACCATAGGTACAAAGCGGC	ACTGTCAATACGGCTGAGGCGAAG
4-8C	GATATGAACACACGTAGCACGGCG	CACACGCGATGAGAAGCTGAAAGA
4-8D	CAGTATATTGATCCAGAGCGCGGG	GACGGTGGTTCTGACCCGCTTAAT
4-8E	GCGGGCGAACTCAAAATTTGATTC	AGGCTTGCTCACTGACGTCGTTCT
4-8F	TCGAATCTGAGATTGGCTCCAACC	GGAGCGGGTAAACGTTCTGTTCGT
4-8G	AGCGGTTGAGTATCGCCCTAGGAC	GGCTTGCAATCTCCCGAGACCTAT
4-8H	TAGTAGTTAATACCGGCCCCGGCC	ACTGACGGCTCGAGATGTGGTAGC
4-9A	GACAGCCCGGGGATTTAGGCTAAT	CTGGTACACGAACACATGCCAGCT
4-9B	CGGATTATTCGGACTAACGGACGG	CATCACGAGAGCTACCCCCATCAC
4-9C	CATGGTCGTTGATAGCCCTCTTGG	ACGTCTCGGAATGATTTTGCTCGA
4-9D	CTTAGGCTGGCTATTCACACCGCA	TGTGGGTAGTCCCTCGTGTCGACT
4-9E	ATTTACTAGAAGCCCGGTGACGCG	CGGCGAGGCACCTTAGATACTTCC
4-9F	CCTGCGGAAACAGAAAACCTAGGC	CACATTACGTACGGGAGTCCCCAA
4-9G	CTTTTGTTCGAGGAAGTCCGGGTC	GCGGCCCCTGGGAAATTTTACTTA
4-9H	CGCGCCGTTTTTTATAGCTTCTTGC	GCAATGAACAGACATAGCCCAGGG
5-1A	ATAAGAACCAAGTGCCGGGCAAAA	TCTATGATGGCCTGGGAGGTGCTA
5-1B	GGGGAACTGCCTCGTTAGAGACGT	ATACCCACGACTTTTACTCGCGGG
5-1C	GCGCGCATTACCCTGAGGTTTAAC	CAGAGGACCAGTTGGATGCGCTAT
5-1D	CACGTAGGTGCGGGACCATTAGTC	TTGTGTCCTAAGGCTCTAGCGGCA
5-1E	CGCGAAACCTATTTGTAGCGTGGA	AACGCAGCGTCTGTGTTGAGAAGA
5-1F	TCAGGAAGAAACTCTGTTTGGCCG	TCGAAGCCATTCAAATACCAACGC
5-1G	CGGGAGAACGGAACTGTGAATCAG	CCCGTCGCTCCGAGGAATAAGTAG
5-1H	ACGGGCTCATATTTGGTCGCTTCT	CGGTGTGTTAGAGTGGCTCCGACT
5-2A	TTGCGTTCCAGATGGATCGTATGA	CCCCTGAACTTAACCGCTAATGCG

5-2B	GGGGATTCCGCTCAGTAATTAGCG	GATGCCTGAGTCCGTAGGTTTGGA
5-2C	CTATATGGTCACCTTCGTCCGGGC	ATGTCAATCGTTCGATTTTCCCGA
5-2D	CAGGAACCGGGAAGCGCTCTAGTA	GCTAGCACCCAAGAGTTGGAAGGG
5-2E	GAGCTGCGGGTCCCGGATACTATA	CGGCCCATGCTACATAGCAATCAT
5-2F	TCGGCTAGGGCTTTCTAGGAACGA	CTTCGATACGTACAAACTGGCCGC
5-2G	TAGTCGGCGCCAGGATGAGACTAA	GGCCGCCTTGAAACTAGGACTTCA
5-2H	TATAGACTCCCATGCCCTCACCCA	AAGATTGGTCGAAGGCCACAGCTT
5-3A	CAGCAAAACTTAAAGTCACCCGCG	TACGGTCTGGCATCAGGTTCAAGG
5-3B	ACATTAGCATACGTGGATTGCCCG	TTCGACATGTGAATTGATGCCCAC
5-3C	AATACATCAGGCCCAAGTTGCCGT	GTGCAGTGCTAGGGGGACATATGC
5-3D	TTCCCAGACTCAATGATCGAGGGA	CGTGCAAACAACGCACTTGATGC
5-3E	CTCGCGCTAGGGTCTTTGGCTTAA	TTCGCTTTAAACAAGAGCTGGGGC
5-3F	TATAGAGGGCAGGCGGAACAAAGC	GCGCGTGGATCACATCCCTAGATA
5-3G	TCGCAGGGTTTCACGGGAATAATT	TTGTTGAACTATTGATTCGGGCGG
5-3H	ACAAGCTGAGCAATTGTGCTCACG	ACAAGTGCAACCCTCGAGCTCAAA
5-4A	CGAGTAGCTCGAAAGAGACGCCAA	AGGTTGGATTCCCGCCATTCTCTA
5-4B	CCCCCGTAGTGTGGCTACTTAGGC	CAGAACGGTCGCAACCTAGTTCGT
5-4C	CCAGCATCATACTTACCGGGCATG	AGGTCGAATGACAGCCTTCCCTCT
5-4D	TGGGGGGCACAGTCAGAAAGTATG	CGACCAGATCATCCGGCTAACTGT
5-4E	CAGCATACCCGAACTATGCTGGGA	TGCCCACTCCAGCACCAGTTATTT
5-4F	CGATCATATTATGCTACCGCGCCA	GATCGCTTGACGGCCGAGTCTTAT
5-4G	GTCGTACGTGGTCTCATATGCGCA	AGGCCGATCGATATGGCAAGAACT
5-4H	TGCTGCTCTTCTGGTACTTGCGTG	GCCCCGACATCTGTAACTCTCGAA

TABLE G.2. CHARACTERISTICS OF ALL MUTANTS IN THE EXPERIMENTAL SET 1 IN FREE-LIVING CONDITIONS

complete mutantID	mutated gene[a]	predicted function	$mean_w$ -value (TY)	p-value (t) (TY)	$mean_w$-value (VMM)	p-value (t) (VMM)	$mean_w$ -value (SDS)	p-value (t) (SDS)	$mean_w$-value (NaCl)	p-value (t) (NaCl)
2011mTn5STM.1.10.A01	SMb21639	phenylacetic acid degradation protein PaaB	0.01	7.49E-01	0.20	1.26E-02	0.40	1.98E-20	0.48	1.95E-22
2011mTn5STM.1.09.B01	SMa0675	cation (Ca) exchange protein, possible	-0.34	4.04E-06	-0.25	8.06E-10	-0.25	1.25E-14	-0.38	6.64E-08
2011mTn5STM.1.09.C01	SMc00513	putative amino-acid binding periplasmic protein	-0.14	3.11E-02	0.16	4.10E-03	0.26	1.28E-09	0.19	2.65E-07
2011mTn5STM.1.02.D01	SMa0259	hypothetical protein	NA	NA	0.26	1.15E-02	0.39	1.79E-05	0.58	1.76E-06
2011mTn5STM.1.01.E01	SMb20834	lipopolysaccharide processing protein RkpZ1	0.14	5.50E-02	0.30	2.34E-03	-0.48	2.36E-22	-0.39	4.51E-25
2011mTn5STM.1.13.F01	SMa0113	Sensory histidine kinase	0.14	2.24E-01	0.37	7.18E-06	NA	NA	0.29	1.08E-03
2011mTn5STM.1.02.G01	SMa1497	putative Gst12 glutathione-S-transferase	-0.08	2.00E-01	0.00	3.54E-01	0.26	8.18E-02	NA	NA
2011mTn5STM.1.13.H01	SMb20167	hypothetical protein	-0.01	7.69E-01	0.10	8.29E-02	0.27	3.63E-18	0.42	1.85E-18
2011mTn5STM.1.13.B02	SMc02327	ribonuclease HI protein RnhA2	-0.11	1.55E-01	-0.11	2.21E-02	0.09	6.32E-05	0.01	7.04E-01
2011mTn5STM.1.09.C02	SMa1903	putative protease	NA	NA	NA	NA	-0.02	5.39E-01	-0.16	3.93E-02
2011mTn5STM.1.03.D02	SMc02466	SdhB succinate dehydrogenase iron-sulfur protein	-0.44	7.20E-06	-0.29	1.11E-01	0.19	1.78E-10	0.20	2.14E-08
2011mTn5STM.1.11.E02	SMc02738	ABC cholin transporter, permease component (OpuB)	-0.12	1.25E-01	0.00	9.16E-01	-0.07	2.14E-03	-0.14	1.23E-12
2011mTn5STM.1.08.F02	SMa0525	putative ABC-type iron transport system protein	-0.44	1.51E-06	-0.20	2.59E-03	-0.08	4.29E-04	-0.11	7.65E-07
2011mTn5STM.1.07.G02	S. meliloti chromosome, 2021711	intergenic	-0.23	1.22E-03	-0.09	3.25E-01	-0.23	8.34E-17	-0.11	5.96E-07
2011mTn5STM.1.06.H02	SMc01090	ATP-dependent RNA-helicase protein DeaD	-0.45	2.19E-06	-0.34	6.30E-05	0.17	7.04E-11	0.31	1.03E-29
2011mTn5STM.1.10.A03	SMc01046	Trk system potassium uptake protein	-0.23	1.03E-02	-0.31	3.20E-07	-0.11	1.20E-04	-0.21	9.08E-05
2011mTn5STM.1.01.B03	SMc01432	hypothetical protein	-0.04	4.27E-01	0.01	8.56E-01	0.20	8.01E-10	0.23	3.22E-13
2011mTn5STM.1.05.C03	SMb20071	putative efflux protein	0.38	7.42E-06	0.30	4.50E-06	0.30	1.84E-14	0.28	1.11E-17
2011mTn5STM.1.03.D03	SMb20349	putative ABC transporter	0.19	7.35E-03	0.17	2.67E-08	0.04	2.47E-03	0.15	1.01E-05

2011mTn5STM.1.11.E03	SMa0300	ABC transporter, permease	0.31	1.61E-03	0.23	4.99E-07	-0.16	2.78E-08	-0.15	3.49E-04
2011mTn5STM.1.06.F03	SMc00300	putative oxidoreductase protein	0.04	2.92E-01	0.11	5.20E-04	0.05	3.81E-01	0.11	7.78E-04
2011mTn5STM.1.08.G03	pSymB, 543971	intergenic	1.17	5.59E-04	0.74	7.58E-21	0.02	8.70E-01	0.15	3.01E-07
2011mTn5STM.1.08.H03	SMa1236	NapA periplasmic nitrate reductase	0.29	2.27E-06	0.26	9.04E-06	0.09	1.52E-06	0.12	3.92E-04
2011mTn5STM.1.11.A04	SMa1004	hypothetical protein	0.20	8.53E-04	0.06	3.65E-01	-0.22	6.89E-09	-0.27	3.57E-07
2011mTn5STM.1.06.B04	SMb21707	putative short-chain amide or branched-chain amino acid uptake ABC transporter ATP-binding protein	0.17	9.31E-03	0.01	4.67E-01	0.30	3.38E-14	0.32	6.43E-19
2011mTn5STM.1.01.C04	SMa1229	FixL Oxygen regulated histidine kinase	0.04	4.95E-01	0.17	1.94E-03	-0.28	2.67E-19	-0.26	9.49E-17
2011mTn5STM.1.03.D04	SMa1531	NuoC2 NADH I chain C	0.21	5.84E-04	0.05	2.64E-01	0.08	1.10E-04	0.11	4.77E-06
2011mTn5STM.1.13.E04	SMa1503	hypothetical protein	0.47	6.19E-06	0.36	2.12E-05	0.19	2.02E-08	0.22	1.56E-12
2011mTn5STM.1.02.F04	SMa1292	hypothetical protein	0.30	8.57E-04	0.17	2.67E-02	0.00	4.64E-01	0.00	7.14E-01
2011mTn5STM.1.01.G04	SMc01497	sorbitol/mannitol transporter SmoF	0.16	1.98E-02	0.29	1.57E-06	0.44	7.61E-18	0.38	2.46E-16
2011mTn5STM.1.07.H04	SMb21422	putative sugar uptake ABC transporter ATP-binding protein	0.17	2.60E-05	0.07	4.75E-01	0.12	2.18E-08	0.15	4.14E-11
2011mTn5STM.1.02.A05	SMb21645	ABC transporter, permease component	0.10	9.46E-02	0.20	4.10E-04	0.30	2.86E-13	0.40	6.33E-10
2011mTn5STM.1.07.B05	SMb20142	putative oligopeptide ABC transporter permease protein	-0.09	2.52E-01	-0.07	5.97E-01	0.19	1.00E-05	0.30	5.97E-10
2011mTn5STM.1.03.C05	SMc02121	general L-amino acid transport ATP-binding ABC transporter protein AapP	0.05	4.75E-01	0.21	5.30E-03	0.44	1.93E-16	0.48	1.69E-18
2011mTn5STM.1.08.D05	SMb20219	putative response regulator protein	0.13	1.37E-01	0.23	3.54E-02	0.13	1.84E-13	0.37	1.45E-22
2011mTn5STM.1.09.E05	SMb21527	putative taurin uptake ABC transporter ATP-binding protein TauB	0.05	4.06E-01	0.20	4.02E-05	0.05	1.77E-03	0.17	7.04E-07
2011mTn5STM.1.12.F05	SMa0803	putative ABC-transporter ATP-binding protein	-0.16	3.01E-03	-0.09	2.89E-02	-0.10	2.74E-04	-0.26	3.79E-13
2011mTn5STM.1.12.G05	SMa0621	FixI2 E1-E2 type cation ATPase	0.88	1.68E-03	0.77	3.69E-09	0.11	5.74E-05	0.16	5.19E-10
2011mTn5STM.1.07.H05	SMa0875	NolG efflux transporter	-0.12	9.63E-02	0.12	1.05E-01	0.04	8.71E-03	-0.16	2.91E-16
2011mTn5STM.1.09.A06	SMa1507	hypothetical protein	-0.20	1.64E-03	0.03	8.27E-01	0.04	1.68E-02	0.13	4.52E-10
2011mTn5STM.1.04.B06	SMb20930	putative sugar uptake ABC transporter ATP-binding protein	-0.12	3.15E-01	-0.21	7.09E-02	-0.32	9.84E-13	-0.29	1.43E-12
2011mTn5STM.1.11.C06	SMc02507	manganese ABC transporter permease SitC	0.03	3.36E-01	0.22	1.40E-03	0.61	1.00E-12	0.72	1.63E-06
2011mTn5STM.1.03.D06	SMa1760	hypothetical protein	NA	NA	NA	NA	0.12	1.12E-03	-0.10	2.26E-06

2011mTn5STM.1.06.E06	pSymA, 305100	intergenic	-0.09	1.15E-01	0.10	6.25E-02	0.27	8.46E-14	0.19	7.43E-11
2011mTn5STM.1.12.F06	SMb20390	hypothetical protein	-0.11	1.18E-01	-0.07	4.77E-01	0.25	9.19E-21	0.28	1.05E-22
2011mTn5STM.1.08.G06	SMa1798	Kup2 Potassium uptake protein	-0.25	3.02E-04	-0.26	1.31E-08	-0.23	3.45E-08	NA	NA
2011mTn5STM.1.11.H06	SMa0214	putative KduI DKI isomerase	-0.10	8.27E-02	-0.19	5.15E-07	0.20	9.86E-18	0.13	8.55E-08
2011mTn5STM.1.12.B07	S. Meliloti chromosome, 1067681	intergenic	-0.06	3.68E-02	-0.02	3.69E-01	-0.06	9.26E-02	0.01	4.86E-01
2011mTn5STM.1.12.C07	SMa0803	putative ABC-transporter ATP-binding protein	0.23	5.54E-07	0.08	3.87E-02	-0.04	7.33E-01	-0.21	8.66E-05
2011mTn5STM.1.11.D07	SMb20819	ferredoxin, Rieske [2Fe-2S] domain MocE	0.10	6.04E-03	0.02	2.79E-01	-0.22	8.23E-16	-0.32	1.83E-30
2011mTn5STM.1.08.E07	SMa1467	Probable inner-membrane permease	0.07	1.08E-01	0.17	4.38E-05	-0.01	1.58E-01	-0.12	1.47E-01
2011mTn5STM.1.11.F07	SMa1115	Putative manganese transport protein	-0.08	1.72E-02	-0.18	1.98E-05	-0.13	2.55E-04	-0.18	1.77E-11
2011mTn5STM.1.04.G07	S. meliloti chromosome, 2181489	intergenic	-0.08	1.59E-01	-0.15	2.89E-01	-0.14	3.90E-05	-0.24	9.33E-16
2011mTn5STM.1.05.H07	S. meliloti chromosome, 1621189	intergenic	-0.02	7.03E-01	0.14	1.20E-04	0.25	5.97E-16	0.23	1.02E-16
2011mTn5STM.1.02.A08	SMc00185	putative ABC-transporter ATP-binding protein	0.07	1.32E-01	0.15	5.70E-04	0.09	6.56E-03	0.05	2.54E-02
2011mTn5STM.1.03.B08	SMb20513	hypothetical protein	-0.13	2.42E-01	-0.21	4.85E-07	-0.09	4.33E-04	-0.11	1.44E-05
2011mTn5STM.1.05.C08	SMb21477	putative reverse transcriptase maturase protein	0.12	2.28E-02	-0.06	1.30E-01	-0.09	1.13E-05	-0.14	2.56E-05
2011mTn5STM.1.11.D08	SMa1638	hypothetical protein	-0.40	1.02E-06	-0.52	2.68E-13	-0.48	3.33E-25	-0.41	3.96E-17
2011mTn5STM.1.03.E08	SMc02236	hypothetical protein	-0.27	4.70E-05	-0.04	6.68E-01	-0.17	6.76E-08	0.13	1.69E-04
2011mTn5STM.1.09.F08	SMa1146	hypothetical protein	-0.34	1.19E-06	-0.36	8.67E-13	-0.14	6.63E-05	-0.22	6.07E-11
2011mTn5STM.1.12.G08	SMb20948	ExoU glucosyltransferase	-0.09	1.05E-01	-1.02	6.81E-15	-0.43	3.41E-19	-0.57	2.65E-29
2011mTn5STM.1.09.H08	SMa0809	hypothetical protein	-0.27	2.64E-03	-0.58	2.97E-12	0.22	2.06E-02	0.35	5.16E-02
2011mTn5STM.1.04.A09	SMb20616	ThiO thiamine biosynthesis oxidoreductase protein	0.41	5.44E-06	-0.49	4.56E-07	0.31	4.60E-12	0.37	1.16E-14
2011mTn5STM.1.01.B09	SMb20958	exopolysaccharide biosynthesis protein ExoM	0.25	5.29E-02	0.10	1.08E-01	-0.06	1.16E-03	-0.03	6.88E-02
2011mTn5STM.1.10.C09	SMa0244	hypothetical protein	0.14	8.36E-03	0.19	3.76E-02	-0.05	6.07E-02	0.06	1.28E-01
2011mTn5STM.1.05.D09	SMa0840	NodD3 transcriptional regulator	0.09	1.75E-01	-0.11	5.26E-03	-0.15	3.38E-01	-0.26	7.82E-03

2011mTn5STM.1.05.E09	SMc02773	sugar transport ATP-binding protein	0.28	8.85E-06	0.23	4.68E-06	0.11	2.05E-05	0.24	1.31E-08
2011mTn5STM.1.02.F09	SMc02612	glutamate synthase family protein GlxD	0.14	4.67E-02	0.00	4.30E-01	0.30	1.19E-06	0.15	3.20E-05
2011mTn5STM.1.03.G09	SMa1128	DegP4 protease like protein	0.33	1.89E-06	0.27	4.90E-04	-0.19	4.87E-12	-0.20	7.39E-08
2011mTn5STM.1.12.H09	SMb21087	conjugal transfer protein TraA2	-0.12	9.37E-02	0.02	7.41E-01	-0.23	3.34E-02	-0.15	2.62E-01
2011mTn5STM.1.13.A10	SMa0364	hypothetical protein	0.62	7.02E-06	0.34	2.56E-12	0.32	2.36E-22	0.19	5.57E-17
2011mTn5STM.1.08.B10	SMb21578	copper-transporting P-type ATPase protein AtcU2	-0.02	6.69E-01	-0.29	5.40E-04	-0.41	2.69E-15	-0.31	1.00E-14
2011mTn5STM.1.11.C10	SMb20070	putative sulfate permease protein	0.21	4.68E-02	0.09	3.03E-01	0.07	4.51E-03	0.18	1.40E-06
2011mTn5STM.1.12.E10	SMc00349	GTP-binding membrane protein LepA	NA	NA	-0.83	3.59E-03	-0.38	1.80E-22	-0.29	1.08E-18
2011mTn5STM.1.08.F10	SMc02406	hypothetical protein	-0.34	6.92E-04	0.01	2.65E-01	0.53	3.39E-16	0.09	2.74E-02
2011mTn5STM.1.13.G10	SMa0104	ABC transporter, solute-binding protein	0.14	1.99E-02	0.03	5.65E-01	0.06	1.25E-02	0.12	9.63E-02
2011mTn5STM.1.04.H10	SMa0252	hypothetical protein	0.14	3.55E-02	-0.07	1.91E-01	0.06	5.48E-01	0.06	3.47E-01
2011mTn5STM.1.02.A11	SMc00413	hypothetical protein	-0.09	4.84E-01	0.30	1.58E-02	NA	NA	0.75	1.79E-06
2011mTn5STM.1.07.B11	SMb20967	putative transcriptional regulator, lacI family protein	0.19	1.11E-01	0.02	5.73E-01	-0.15	2.42E-07	-0.09	9.60E-03
2011mTn5STM.1.08.C11	SMc04346	IlvC ketol-acid reductoisomerase protein	-0.20	1.95E-03	-1.09	2.20E-13	-0.08	9.33E-03	-0.02	7.34E-01
2011mTn5STM.1.05.D11	SMa0814	NifB FeMo cofactor biosynthesis protein	NA	NA	-0.29	1.85E-02	-0.43	7.52E-09	-0.48	6.61E-16
2011mTn5STM.1.01.E11	SMa0765	FixN2 cytochrome c oxidase polypeptide I	0.44	1.84E-05	0.28	1.82E-01	-0.10	8.93E-05	-0.30	6.95E-17
2011mTn5STM.1.06.F11	SMa0922	hypothetical protein	-0.03	4.88E-01	0.08	4.25E-01	0.04	6.62E-02	-0.01	4.36E-01
2011mTn5STM.1.12.G11	SMa1195	hypothetical protein	0.45	1.42E-03	0.58	5.02E-14	0.29	2.10E-17	0.25	4.08E-16
2011mTn5STM.1.05.H11	SMa1493	putative LysR-type transcription factor	0.03	2.91E-01	0.07	1.41E-02	-0.16	1.05E-04	-0.23	2.26E-07
2011mTn5STM.1.11.A12	SMc01664	putative transcription regulator protein	-0.06	3.78E-01	-0.05	4.16E-01	-0.13	3.64E-04	-0.15	2.29E-12
2011mTn5STM.1.11.B12	SMc00322	hypothetical protein	-0.28	3.51E-04	-0.30	1.90E-04	0.03	3.89E-01	-0.06	1.68E-01
2011mTn5STM.1.11.C12	SMa0489	putative ABC transporter, ATP-binding protein	0.21	3.84E-03	0.25	8.07E-05	0.03	1.71E-02	-0.13	7.55E-05
2011mTn5STM.1.02.D12	SMa1288	hypothetical protein	0.02	9.07E-01	-0.25	2.20E-04	-0.20	2.62E-08	-0.21	9.20E-07
2011mTn5STM.1.11.E12	SMc04337	hypothetical protein	-0.14	3.03E-04	0.09	6.59E-02	0.10	3.51E-02	0.05	5.06E-01
2011mTn5STM.1.04.F12	SMc04363	hypothetical protein	-0.32	5.72E-04	-0.02	7.54E-01	0.32	1.13E-13	0.41	2.57E-12

2011mTn5STM.1.04.H12	SMb21641	regulator of phenylacetic acid degradation PaaX	0.04	8.09E-01	0.04	7.34E-01	0.25	1.87E-10	0.33	2.12E-22
2011mTn5STM.2.05.A01	S. meliloti chromosome, 2044300	intergenic	NA	NA	NA	NA	NA	NA	1.13	1.56E-05
2011mTn5STM.2.07.B01	SMa1182	NosZ N2O reductase	-0.03	8.50E-01	0.14	5.60E-04	-0.01	9.42E-01	0.05	5.13E-02
2011mTn5STM.2.11.C01	SMa1677	hypothetical protein	0.32	6.58E-03	0.00	7.61E-01	0.23	1.42E-04	0.08	1.07E-01
2011mTn5STM.2.10.D01	SMc01950	high-affinity branched-chain amino acid ABC transporter, permease protein LivM	-0.23	1.25E-03	-0.14	1.20E-04	-0.04	8.52E-02	-0.11	5.85E-04
2011mTn5STM.2.12.E01	SMc03249	hypothetical protein	0.49	8.91E-10	0.39	5.31E-14	0.14	3.96E-06	0.02	4.60E-02
2011mTn5STM.2.07.F01	SMa0607	hypothetical protein	0.10	8.05E-02	0.10	8.89E-02	0.01	6.23E-01	0.06	7.59E-03
2011mTn5STM.2.01.G01	pSymA, 854083	intergenic	0.22	2.91E-04	0.20	1.54E-08	0.03	2.40E-02	0.09	5.74E-03
2011mTn5STM.2.12.H01	SMc01219	lipopolysaccharide core biosynthesis mannosyltransferase LpsB	-0.39	4.39E-04	-0.36	3.09E-08	-0.75	2.70E-17	-0.18	1.99E-06
2011mTn5STM.2.07.A02	SMc01607	putative permease protein	0.28	3.44E-05	0.32	1.06E-12	-0.11	6.66E-06	0.02	3.90E-02
2011mTn5STM.2.05.B02	SMb21144	putative choline uptake ABC transporter periplasmic solute-binding protein precursor	0.11	2.21E-02	0.28	1.95E-07	0.06	2.82E-01	0.15	3.29E-04
2011mTn5STM.2.05.C02	SMa0876	NolF secretion protein	0.21	1.03E-04	0.24	1.56E-05	-0.02	5.64E-01	-0.23	1.25E-01
2011mTn5STM.2.09.D02	SMa0322	hypothetical protein	-0.03	7.03E-01	-0.17	3.04E-07	-0.27	2.36E-16	-0.36	2.54E-17
2011mTn5STM.2.04.E02	SMb21633	PaaG enoyl-CoA hydratase protein	0.05	9.52E-01	0.53	5.89E-05	0.75	1.38E-08	0.96	5.23E-09
2011mTn5STM.2.05.F02	S. meliloti chromosome, 1627869	intergenic	0.23	1.15E-03	0.54	4.13E-11	0.34	2.81E-03	0.25	7.77E-06
2011mTn5STM.2.02.G02	SMb20273	hypothetical protein	0.03	1.28E-01	0.11	7.90E-04	0.20	1.43E-08	0.21	4.97E-07
2011mTn5STM.2.12.H02	SMb20851	putative transcriptional regulator, sorC family protein	0.17	3.18E-02	-0.10	1.02E-01	-0.14	3.92E-09	-0.15	1.13E-06
2011mTn5STM.2.12.A03	SMb20209	hypothetical protein	0.65	2.10E-08	0.43	2.97E-17	0.27	1.02E-09	0.35	1.30E-09
2011mTn5STM.2.10.B03	S. meliloti chromosome, 3367064	intergenic	-0.04	3.75E-01	-0.04	1.89E-01	-0.37	1.09E-19	-0.42	1.21E-09
2011mTn5STM.2.02.C03	SMc00556	DNA repair protein RadA	0.00	7.22E-01	0.12	6.16E-02	-0.05	5.32E-01	-0.10	1.04E-01
2011mTn5STM.2.12.D03	SMa0485	probable ThrC2 threonine synthase	-0.01	4.46E-01	-0.13	1.47E-03	-0.22	8.19E-13	-0.32	4.84E-15
2011mTn5STM.2.01.E03	SMb20712	rhizopine uptake ABC transporter periplasmic solute-binding protein precursor MocB	0.39	1.33E-05	0.41	1.71E-12	0.07	1.50E-02	0.25	8.56E-08

2011mTn5STM.2.07.F03	SMb20168	hypothetical protein	0.29	3.42E-06	0.37	1.25E-06	NA	NA	0.07	2.16E-02
2011mTn5STM.2.06.G03	SMc03044	chemotaxis protein MotD	0.13	3.51E-02	0.19	1.22E-05	0.21	3.34E-09	0.33	9.40E-06
2011mTn5STM.2.09.H03	SMa1833	hypothetical protein	0.07	4.91E-02	-0.12	1.02E-05	NA	NA	NA	NA
2011mTn5STM.2.11.A04	SMc02270	capsular polysaccharide biosynthesis transmembrane protein RkpI	0.10	1.68E-02	0.38	3.58E-06	-0.28	2.81E-13	0.04	4.29E-04
2011mTn5STM.2.09.B04	SMb20112	hypothetical protein	0.07	6.97E-02	0.21	8.04E-09	-0.04	1.10E-01	0.13	4.27E-08
2011mTn5STM.2.12.C04	SMc02659	GTP pyrophosphokinase protein RelA	NA	NA	-0.34	1.13E-01	NA	NA	NA	NA
2011mTn5STM.2.09.D04	SMb21301	putative membrane-anchored aldehyde dehydrogenase protein	-0.12	1.24E-03	0.01	9.07E-01	NA	NA	0.37	1.03E-04
2011mTn5STM.2.01.E04	SMa0306	putative histidine ammonia-lyase	0.32	5.94E-05	0.12	3.11E-05	-0.11	9.40E-07	-0.22	2.57E-11
2011mTn5STM.2.01.F04	SMa0431	hypothetical protein	0.21	8.33E-05	0.30	1.69E-10	-0.08	4.15E-03	-0.08	3.77E-03
2011mTn5STM.2.11.G04	SMb20472	probable carbamoyltransferase	0.11	2.77E-02	0.14	5.60E-05	-0.11	1.87E-05	-0.19	1.49E-07
2011mTn5STM.2.10.H04	SMa0110	ABC transporter, ATP-binding protein	-0.15	2.27E-03	-0.12	1.06E-02	-0.11	3.92E-04	-0.12	3.49E-03
2011mTn5STM.2.05.A05	SMb21199	oligopeptide uptake ABC transporter ATP-binding protein OppD	-0.05	9.64E-01	0.20	3.44E-05	0.03	1.50E-01	0.30	2.52E-07
2011mTn5STM.2.10.B05	SMb21367	cytochrome c class I protein CycA	0.23	4.01E-02	0.53	4.20E-16	0.67	1.54E-16	0.49	7.96E-18
2011mTn5STM.2.11.C05	SMa0250	hypothetical protein	0.09	2.34E-02	0.13	3.17E-02	0.11	2.65E-02	-0.03	1.69E-01
2011mTn5STM.2.08.D05	SMa0849	SyrM transcriptional regulator	0.25	1.91E-04	0.15	4.52E-05	-0.10	1.42E-04	-0.34	1.13E-15
2011mTn5STM.2.06.E05	SMb21177	phosphate uptake ABC transporter ATP-binding protein	0.28	1.90E-03	0.32	1.40E-04	0.09	9.38E-07	0.18	4.48E-08
2011mTn5STM.2.06.F05	pSymB, 1225168	intergenic	0.03	3.97E-01	0.26	2.21E-06	0.52	4.71E-18	0.57	1.07E-25
2011mTn5STM.2.06.G05	SMa1766	hypothetical protein	0.08	5.39E-02	-0.19	1.32E-07	-0.18	1.66E-04	-0.34	6.20E-18
2011mTn5STM.2.05.H05	SMc03164	xylulose kinase protein XylB	-0.32	4.18E-04	-0.10	8.61E-01	0.86	2.17E-11	0.48	1.36E-11
2011mTn5STM.2.07.A06	SMb20866	hypothetical protein	0.37	3.60E-08	0.38	4.19E-13	-0.09	1.25E-05	-0.15	2.40E-07
2011mTn5STM.2.11.B06	SMb20111	putative oligopeptide ABC transporter ATP-binding protein	0.23	2.59E-02	0.22	1.50E-04	-0.06	4.92E-02	0.02	9.08E-01
2011mTn5STM.2.01.C06	SMa2163	hypothetical protein	0.36	4.56E-06	0.09	8.96E-03	-0.13	2.21E-05	-0.26	7.48E-15
2011mTn5STM.2.06.D06	SMc00864	hypothetical protein	-0.08	4.26E-01	0.05	3.60E-01	0.09	5.67E-02	0.08	1.40E-02
2011mTn5STM.2.03.E06	SMb21174	phosphate uptake ABC transporter permease protein	0.27	1.19E-03	0.33	2.56E-07	0.11	9.75E-02	0.19	6.16E-06
2011mTn5STM.2.10.F06	SMb21000	putative transport protein	0.30	5.20E-05	0.44	1.03E-08	0.38	6.90E-09	0.30	3.99E-05

2011mTn5STM.2.05.G06	SMb20943	acetyltransferase protein ExoZ	-0.07	4.23E-02	0.07	1.51E-01	-0.22	5.40E-12	-0.14	2.07E-04
2011mTn5STM.2.11.H06	SMb20114	hypothetical protein	0.01	2.41E-01	0.00	9.97E-01	0.11	2.48E-07	0.22	5.18E-16
2011mTn5STM.2.07.A07	SMc04204	iron transport regulator transmembrane protein FecR	0.10	6.35E-02	-0.03	1.92E-01	0.03	8.94E-01	-0.10	8.52E-04
2011mTn5STM.2.07.B07	SMb21493	putative transcriptional regulator, asnC family protein	0.11	2.15E-02	0.03	7.91E-02	0.01	5.19E-01	0.04	8.58E-03
2011mTn5STM.2.02.C07	SMc01528	dipeptide transport ATP-binding ABC transporter protein	0.01	6.74E-01	0.18	4.23E-03	NA	NA	0.26	3.71E-04
2011mTn5STM.2.01.D07	SMa1957	hypothetical protein	-0.11	1.62E-02	-0.35	2.49E-10	-0.22	9.54E-12	-0.33	1.14E-19
2011mTn5STM.2.12.E07	SMc00383	glutatione S-transferase protein	0.44	4.37E-09	0.38	1.55E-06	0.62	2.36E-15	0.43	6.22E-13
2011mTn5STM.2.06.F07	SMc00774	acetoacetyl-coenzyme A synthetase, additional	0.25	3.24E-04	0.33	1.84E-09	0.23	2.52E-12	0.25	9.05E-16
2011mTn5STM.2.04.G07	SMb20123	putative transcriptional regulator protein	0.08	1.63E-02	0.00	9.84E-01	NA	NA	0.80	5.08E-04
2011mTn5STM.2.09.H07	pSymB, 96435	intergenic	-0.07	4.52E-01	0.09	7.05E-01	0.91	8.13E-09	0.48	2.14E-04
2011mTn5STM.2.04.A08	SMa0039	putative LysR-family transcriptional regulator	0.19	3.99E-05	0.24	1.04E-06	0.12	2.98E-03	0.12	7.10E-02
2011mTn5STM.2.04.B08	S. meliloti chromosome, 3364101	intergenic	0.32	5.35E-05	0.22	2.56E-08	0.04	5.23E-03	-0.13	2.15E-07
2011mTn5STM.2.05.C08	SMa1956	putative LysR-family transcriptional regulator	0.03	2.92E-01	-0.15	2.77E-03	-0.13	4.34E-03	-0.30	4.53E-14
2011mTn5STM.2.04.D08	SMc03149	hypothetical protein	0.12	2.52E-03	0.07	3.31E-02	0.06	1.04E-04	0.06	9.04E-03
2011mTn5STM.2.08.E08	SMb20757	methylmalonyl-CoA mutase protein BhbA	0.24	3.75E-03	0.15	9.30E-04	0.24	2.44E-13	0.26	4.60E-13
2011mTn5STM.2.12.F08	SMa1513	putative ABC transporter permease	0.23	1.39E-06	0.33	3.87E-09	-0.05	4.11E-01	0.03	6.92E-01
2011mTn5STM.2.11.G08	SMc00527	hypothetical protein	-0.26	6.00E-05	0.00	6.69E-01	0.00	8.19E-01	0.03	3.33E-01
2011mTn5STM.2.10.H08	SMa0394	putative ABC transporter, permease	-0.19	5.54E-02	-0.27	2.26E-08	-0.19	3.83E-11	-0.23	9.89E-13
2011mTn5STM.2.11.A09	SMc02899	PheA prephenate dehydratase protein	0.12	8.98E-01	-0.85	6.20E-04	0.31	3.97E-06	0.06	4.22E-02
2011mTn5STM.2.09.B09	SMc00148	hypothetical protein	0.06	9.21E-01	0.04	5.35E-01	-0.27	2.13E-07	-0.11	8.09E-07
2011mTn5STM.2.09.C09	SMb20959	probable exopolysaccharide biosynthesis protein ExoO	-0.11	8.75E-02	-0.12	9.20E-04	-0.39	2.99E-19	-0.21	7.94E-12
2011mTn5STM.2.09.D09	SMc04456	chaperon protein CsaA	-0.25	1.07E-03	-0.24	1.13E-07	-0.27	3.70E-13	-0.18	3.27E-14
2011mTn5STM.2.12.E09	SMc03782	hypothetical protein	-0.24	7.97E-03	-1.02	4.77E-14	-0.36	1.14E-06	-0.54	2.49E-11
2011mTn5STM.2.07.F09	SMa0260	GabD3 succinate-semialdehyde dehdyrogenase	0.17	8.40E-02	0.47	1.91E-06	0.06	6.03E-01	0.08	7.01E-02

2011mTn5STM.2.10.G09	SMa1746	putative iron uptake protein	-0.30	4.14E-07	-0.39	2.82E-12	-0.44	6.52E-23	-0.59	6.23E-37
2011mTn5STM.2.03.H09	SMc02045	putative oxidoreductase protein	-0.28	2.70E-06	-0.16	1.10E-03	0.22	1.66E-09	0.27	1.82E-19
2011mTn5STM.2.10.A10	S. meliloti chromosome, 2106334	intergenic	0.21	3.62E-04	0.10	5.85E-03	0.08	7.77E-04	0.06	5.09E-02
2011mTn5STM.2.04.B10	SMa0565	hypothetical protein	0.82	1.69E-08	0.79	1.21E-13	NA	NA	0.54	6.22E-16
2011mTn5STM.2.04.C10	SMc00086	diheme cytochrome C-type signal peptide protein CycG	-0.09	5.57E-02	-0.13	4.28E-03	-0.08	1.64E-03	0.05	4.20E-04
2011mTn5STM.2.03.D10	SMb20748	glycosyltransferase protein PssF	0.03	5.64E-01	-0.16	2.10E-05	0.66	4.72E-16	0.51	6.21E-14
2011mTn5STM.2.11.E10	SMb21273	spermidineputrescine ABC transporter periplasmic solute-binding protein precursor	0.23	2.49E-03	0.26	3.20E-04	0.16	3.96E-08	0.15	5.24E-08
2011mTn5STM.2.09.F10	SMc01818	adenylate cyclase transmembrane protein CyaC	-0.11	7.42E-02	-0.16	4.50E-04	-0.06	6.50E-01	NA	NA
2011mTn5STM.2.05.G10	SMb20522	hypothetical protein	-0.20	1.30E-03	-0.21	1.16E-05	-0.25	4.27E-09	-0.41	2.22E-18
2011mTn5STM.2.01.H10	SMa1017	hypothetical protein	-0.35	9.02E-04	-0.57	6.02E-18	-0.47	3.74E-07	-0.26	4.25E-05
2011mTn5STM.2.06.A11	SMc01042	nitrogen regulation protein NtrB	-0.09	1.69E-02	-0.08	1.14E-02	NA	NA	-0.06	3.12E-01
2011mTn5STM.2.06.B11	SMc01881	PanB 3-methyl-2-oxobutanoate hydromethyltransferase protein	-0.60	3.38E-04	-0.48	1.22E-05	2.28	1.18E-21	2.20	1.38E-16
2011mTn5STM.2.04.C11	SMb20515	putative CheR methyltransferase, SAM binding domain	-0.15	6.19E-04	-0.23	8.61E-06	-0.17	7.66E-05	-0.25	1.28E-07
2011mTn5STM.2.01.D11	SMc04452	NADH dehydrogenase transmembrane protein Ndh	-0.24	7.05E-03	-0.02	5.63E-01	0.42	1.42E-04	0.68	2.40E-27
2011mTn5STM.2.05.E11	SMa1122	hypothetical protein	0.16	9.60E-03	0.09	2.93E-02	-0.15	1.08E-01	0.06	9.72E-01
2011mTn5STM.2.04.F11	pSymA, 510651	intergenic	-0.05	6.96E-02	0.11	2.96E-03	NA	NA	NA	NA
2011mTn5STM.2.11.G11	SMb21345	putative sugar ABC transporter	-0.11	7.17E-02	-0.05	1.73E-01	0.09	3.67E-05	0.16	1.18E-11
2011mTn5STM.2.10.H11	SMb20549	hypothetical protein	-0.40	7.99E-08	-0.52	1.37E-13	-0.20	1.07E-09	-0.41	1.71E-26
2011mTn5STM.2.03.A12	SMc00739	hypothetical protein	-0.01	6.68E-01	-0.03	9.64E-01	-0.11	8.41E-02	-0.18	7.23E-08
2011mTn5STM.2.09.B12	S. meliloti chromosome, 2862221	intergenic	-0.32	1.03E-03	-0.18	7.44E-05	NA	NA	NA	NA
2011mTn5STM.2.04.C12	SMc01814	probable glutamate synthase small chain protein	-0.53	5.40E-08	-0.28	2.64E-09	0.17	7.85E-05	0.11	1.40E-02
2011mTn5STM.2.01.D12	pSymB, 610814	intergenic	-0.34	2.79E-06	-0.42	1.90E-06	-0.07	1.83E-02	-0.28	4.22E-21
2011mTn5STM.2.03.E12	SMb20548	hypothetical protein	0.17	1.52E-02	-0.19	4.30E-06	-0.37	1.70E-08	-0.41	3.48E-08

2011mTn5STM.2.09.F12	SMc03110	hypothetical protein	-0.06	4.26E-01	0.10	3.54E-03	0.32	3.00E-10	-0.15	1.76E-04
2011mTn5STM.2.06.G12	SMc00094	betaine aldehyde dehydrogenase protein BetB	-0.15	2.11E-02	0.05	4.49E-01	0.31	6.52E-05	0.26	6.60E-08
2011mTn5STM.2.06.H12	SMc01711	hypothetical protein	-0.34	6.57E-06	-0.11	1.18E-01	0.20	6.23E-15	0.24	1.08E-21
2011mTn5STM.3.12.A01	SMa0819	FixB electron transfer flavoprotein alpha chain	0.41	3.77E-08	0.15	3.68E-05	-0.16	1.60E-02	-0.22	5.44E-02
2011mTn5STM.3.08.B01	SMc04217	hypothetical protein	-0.04	8.65E-01	0.48	1.25E-05	0.64	2.64E-06	0.64	2.65E-10
2011mTn5STM.3.10.C01	SMc01842	putative methyltransferase transcription regulator protein	-0.12	8.01E-01	-1.00	1.62E-07	-0.13	2.89E-01	0.11	1.01E-02
2011mTn5STM.3.10.D01	SMc00228	ATP-dependent DNA helicase protein RecG	0.08	7.07E-02	-0.15	4.08E-02	-0.14	1.65E-04	-0.13	1.01E-02
2011mTn5STM.3.10.E01	SMc04386	aspartate aminotransferase B protein	0.01	9.02E-01	0.11	1.17E-01	0.06	7.66E-02	0.13	9.18E-05
2011mTn5STM.3.08.F01	SMa2333	KdpA potassium-transporting ATPase A chain	0.27	9.25E-04	0.31	1.81E-06	0.03	5.84E-01	0.01	9.47E-03
2011mTn5STM.3.05.G01	SMb20178	hypothetical protein	0.13	6.34E-03	0.07	1.89E-02	0.03	3.18E-02	0.10	1.34E-03
2011mTn5STM.3.11.A02	pSymA, 1215937	intergenic	0.27	5.97E-06	0.11	8.98E-05	-0.19	2.67E-09	-0.31	9.97E-07
2011mTn5STM.3.08.B02	SMc02691	putative membrane transport protein	0.28	3.47E-04	0.17	1.46E-05	0.13	1.39E-06	0.14	7.13E-05
2011mTn5STM.3.08.C02	SMb20121	hypothetical protein	0.08	4.16E-03	0.02	5.63E-01	0.02	8.65E-02	0.02	6.68E-02
2011mTn5STM.3.07.D02	SMc02376	putative heat shock protein	-0.13	3.26E-02	-0.03	8.64E-02	0.01	2.45E-01	0.03	7.06E-01
2011mTn5STM.3.09.E02	SMa2215	putative GntR-family transcriptional regulator	0.16	2.67E-03	0.17	5.68E-05	-0.22	2.05E-17	-0.38	2.84E-11
2011mTn5STM.3.10.F02	SMc00420	hypothetical protein	0.00	9.34E-01	0.32	4.10E-05	0.52	3.79E-10	0.43	3.80E-09
2011mTn5STM.3.08.G02	SMa2075	probable extracellular solute-binding protein	0.22	7.17E-05	0.01	3.10E-01	-0.11	1.02E-02	-0.15	9.94E-02
2011mTn5STM.3.07.H02	SMc02363	cytochrome C-type biogenesis transmembrane protein CycK	0.16	3.51E-01	0.19	2.87E-01	0.01	6.31E-01	0.09	7.73E-01
2011mTn5STM.3.09.A03	S. meliloti chromosome, 2823999	intergenic	0.19	4.04E-03	0.09	3.54E-03	-0.16	1.35E-02	0.01	6.76E-01
2011mTn5STM.3.11.B03	SMb20029	hypothetical protein	0.38	3.05E-05	0.29	2.49E-08	0.19	7.33E-09	0.34	5.32E-09
2011mTn5STM.3.11.C03	SMb20067	hypothetical protein	0.11	9.58E-03	0.09	2.96E-05	0.14	6.48E-09	0.21	4.66E-06
2011mTn5STM.3.09.D03	SMa0929	TraG conjugal transfer protein	-0.04	6.01E-01	0.02	3.89E-01	0.12	1.46E-02	0.07	3.46E-02
2011mTn5STM.3.05.E03	SMc01552	hypothetical protein	0.20	2.49E-03	0.31	2.11E-03	-0.04	3.58E-01	-0.08	6.74E-01

2011mTn5STM.3.08.F03	S. meliloti chromosome, 2178672	intergenic	0.10	1.39E-01	0.34	4.93E-09	-0.20	2.84E-11	-0.19	1.06E-01
2011mTn5STM.3.11.G03	SMa0981	NtrR2 transcription regulator	0.01	9.43E-01	0.08	3.56E-01	-0.27	3.16E-14	-0.36	1.25E-11
2011mTn5STM.3.08.H03	SMb20436	probable nitrate transporter	-0.08	6.70E-02	-0.23	3.23E-07	-0.22	1.32E-18	-0.18	8.00E-09
2011mTn5STM.3.07.A04	SMc01131	hypothetical protein	-0.26	2.30E-02	-0.19	9.80E-04	0.46	1.57E-09	0.42	8.16E-17
2011mTn5STM.3.11.B04	SMb21431	hypothetical protein	-0.01	6.43E-01	0.04	2.28E-01	-0.12	1.11E-04	-0.11	4.89E-02
2011mTn5STM.3.06.C04	SMc01668	hypothetical protein	0.30	2.06E-01	0.26	4.91E-03	0.16	3.19E-03	0.08	1.55E-01
2011mTn5STM.3.10.D04	SMb20672	putative sugar ABC transporter	0.02	3.76E-01	-0.05	2.34E-02	0.12	1.45E-06	0.22	5.64E-08
2011mTn5STM.3.06.E04	SMa1179	NosR Regulatory protein for N2O reductase	0.05	2.91E-01	0.26	4.47E-07	-0.11	3.43E-03	-0.07	2.08E-01
2011mTn5STM.3.12.F04	SMb20009	putative transcriptional regulator	0.11	4.44E-02	0.24	5.03E-08	0.19	9.48E-08	0.25	9.06E-17
2011mTn5STM.3.06.G04	SMc00609	anaerobic dimethyl sulfoxide reductase chain A protein DmsA	-0.02	6.46E-01	0.05	1.04E-01	-0.14	3.25E-03	0.01	3.32E-01
2011mTn5STM.3.05.H04	SMb20152	hypothetical protein	0.08	2.29E-01	0.17	2.55E-05	NA	NA	NA	NA
2011mTn5STM.3.11.A05	SMb20037	shikimate 5-dehydrogenase protein AroE2	0.63	1.39E-04	0.76	4.76E-06	0.73	5.12E-05	0.78	3.72E-08
2011mTn5STM.3.09.B05	SMb20931	putative sugar uptake ABC transporter periplasmic solute-binding protein precursor	0.05	2.54E-01	0.19	3.80E-04	1.04	6.40E-14	0.83	7.13E-09
2011mTn5STM.3.08.C05	SMa2293	probable beta lactamase transcriptional activator	-0.39	2.41E-07	-0.08	3.52E-02	NA	NA	NA	NA
2011mTn5STM.3.07.D05	SMc01406	putative transcription regulator protein	-0.38	6.05E-02	-0.02	7.60E-01	0.51	2.09E-04	0.32	2.69E-02
2011mTn5STM.3.07.E05	SMa0890	hypothetical protein	0.35	2.79E-02	0.40	3.68E-08	-0.04	1.32E-01	-0.15	1.62E-01
2011mTn5STM.3.08.F05	SMb21527	putative taurin uptake ABC transporter ATP-binding protein	0.08	2.54E-01	0.10	1.40E-04	-0.11	1.13E-04	-0.04	2.98E-01
2011mTn5STM.3.10.G05	SMc01717	mercuric reductase protein MerA1	-0.07	6.54E-01	0.00	3.84E-01	0.08	5.73E-01	0.14	8.70E-03
2011mTn5STM.3.07.H05	SMa2000	Putative ABC transporter, periplasmic solute-binding protein	0.01	9.99E-01	-0.05	3.88E-02	NA	NA	-0.11	6.05E-04
2011mTn5STM.3.10.A06	SMc03777	gamma-glutamyl phopsphate reductase protein ProA	0.25	2.09E-02	-0.47	5.18E-02	0.26	1.62E-05	0.29	3.12E-04
2011mTn5STM.3.10.B06	SMc03899	hypothetical protein	NA	NA	NA	NA	NA	NA	NA	NA
2011mTn5STM.3.10.C06	SMc02070	hypothetical protein	-0.51	7.21E-03	-0.23	3.90E-03	0.06	8.45E-01	-0.29	4.52E-06
2011mTn5STM.3.11.D06	SMc03032	flagellar P-ring precursor transmembrane protein	0.34	1.08E-05	0.27	3.25E-09	0.92	1.61E-30	0.89	1.78E-34
2011mTn5STM.3.08.E06	SMa0953	AttA1-like ABC transporter, ATP binding protein	NA	NA	0.14	5.10E-02	NA	NA	NA	NA

2011mTn5STM.3.11.F06	SMa2245	hypothetical protein	0.35	1.69E-07	0.36	9.77E-12	0.00	1.83E-01	-0.24	3.41E-10
2011mTn5STM.3.10.G06	SMc02677	pyroline-5-carboxylate reductase protein ProC	-0.05	7.69E-02	-0.37	4.67E-08	-0.13	2.10E-09	-0.13	1.48E-09
2011mTn5STM.3.07.H06	SMa1052	hypothetical protein	0.01	9.62E-01	-0.04	1.76E-01	-0.20	7.25E-10	-0.28	4.29E-18
2011mTn5STM.3.09.A07	SMb21079	putative cAMP binding protein	0.31	6.48E-06	0.11	1.45E-03	-0.11	5.02E-05	-0.22	9.73E-16
2011mTn5STM.3.06.B07	SMc02487	dihydrolipoamide dehydrogenase LpdA2	-0.76	1.27E-02	-1.02	4.27E-05	NA	NA	-0.31	6.21E-02
2011mTn5STM.3.05.C07	SMb20678	hydroxypyruvate reductase protein TtuD2	0.15	6.20E-04	0.12	7.52E-03	0.40	1.37E-14	0.41	4.48E-22
2011mTn5STM.3.08.D07	SMc01174	cysteine synthase A protein CysK2	-0.53	3.04E-08	-0.78	2.67E-18	-0.39	5.37E-20	-0.37	4.97E-25
2011mTn5STM.3.05.E07	SMc01949	high-affinity branched-chain amino acid ABC transporter, ATP-binding protein LivG	0.02	7.59E-01	0.06	6.85E-01	-0.06	6.17E-03	-0.06	1.16E-01
2011mTn5STM.3.06.F07	SMc03977	hypothetical protein	-0.12	6.11E-02	-0.25	6.91E-08	-0.20	7.20E-08	-0.24	1.18E-16
2011mTn5STM.3.07.G07	SMc01419	RNA polymerase sigma factor protein SigE1	-0.38	1.12E-06	-0.23	1.41E-07	-0.26	1.68E-22	-0.22	3.80E-18
2011mTn5STM.3.08.H07	SMc04258	putative transport system permease ABC transporter protein	-0.41	8.87E-07	-0.41	1.92E-12	-0.34	1.94E-02	-0.20	1.82E-04
2011mTn5STM.3.06.A08	SMc01198	hypothetical protein	-0.04	4.75E-01	-0.12	1.20E-03	-0.40	5.75E-20	-0.38	2.76E-24
2011mTn5STM.3.08.B08	SMb20888	hypothetical protein	1.06	NA	0.14	3.99E-01	NA	NA	NA	NA
2011mTn5STM.3.10.C08	SMc02050	trigger factor protein Tig	-0.67	3.66E-07	-0.71	7.20E-08	-0.55	5.97E-06	-0.50	5.52E-08
2011mTn5STM.3.05.D08	SMc01554	hypothetical protein	NA	NA	-1.21	6.74E-01	NA	NA	NA	NA
2011mTn5STM.3.10.E08	SMa2037	Putative oxidoreductase	0.04	2.85E-01	-0.12	3.43E-03	-0.22	5.65E-12	-0.26	3.68E-15
2011mTn5STM.3.05.F08	SMc02857	heat shock protein 70 (HSP70) chaperone	NA	NA	NA	NA	NA	NA	NA	NA
2011mTn5STM.3.08.G08	SMa0583	NrtB, Nitrate transport permease protein	0.26	1.79E-03	0.27	6.32E-07	0.01	8.64E-01	0.05	7.38E-02
2011mTn5STM.3.09.H08	SMc01965	putative spermidine/putrescine transport ATP-binding ABC transporter protein	-0.24	1.67E-02	-0.13	2.66E-05	0.27	3.10E-11	0.26	1.18E-14
2011mTn5STM.3.11.A09	SMa0796	hypothetical protein	0.12	6.30E-03	-0.01	4.53E-01	-0.37	7.51E-21	-0.51	2.12E-32
2011mTn5STM.3.10.B09	SMb20957	exopolysaccharide biosynthesis protein ExoA	-0.06	9.67E-02	-0.29	8.26E-09	-0.19	2.64E-12	-0.20	3.67E-13
2011mTn5STM.3.06.C09	SMb21596	hypothetical protein	-0.14	3.45E-03	-0.29	1.32E-05	-0.40	7.59E-21	-0.37	1.56E-10
2011mTn5STM.3.10.D09	SMb20582	putative transcriptional regulator, LysR family protein	-0.29	3.81E-04	-0.26	4.39E-10	0.00	9.33E-01	0.03	1.44E-01
2011mTn5STM.3.08.E09	SMb20933	putative two-component sensor histidine kinase protein	0.44	2.44E-03	-0.02	8.67E-01	NA	NA	-0.06	8.03E-01
2011mTn5STM.3.12.F09	SMa1087	putative cation transport P-type ATPase	-0.03	2.62E-01	0.04	5.33E-01	NA	NA	NA	NA

2011mTn5STM.3.05.G09	SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit	-0.05	1.90E-01	0.03	1.29E-01	-0.16	3.73E-07	-0.17	8.28E-08
2011mTn5STM.3.07.H09	pSymB, 1322754	intergenic	-0.21	3.69E-05	-0.52	3.34E-12	-0.24	1.37E-10	-0.35	1.56E-18
2011mTn5STM.3.08.A10	SMb20860	probable non-heme chloroperoxidase	0.09	5.18E-02	-0.05	2.31E-01	-0.30	2.49E-16	-0.38	2.10E-08
2011mTn5STM.3.07.B10	SMa0316	hypothetical protein	0.10	3.95E-02	0.00	7.22E-01	0.33	4.33E-07	0.06	3.89E-02
2011mTn5STM.3.06.C10	SMa0615	FixO3 cytochrome-c oxidase subunit	NA	NA	-0.58	8.19E-10	-0.16	1.68E-08	-0.30	7.22E-15
2011mTn5STM.3.05.D10	SMa1916	hypothetical protein	-0.06	1.60E-01	-0.50	3.85E-13	-0.16	2.74E-07	-0.35	1.82E-15
2011mTn5STM.3.09.E10	SMc01000	hypothetical protein	-0.30	1.48E-04	-0.37	8.24E-14	-0.53	1.56E-20	-0.51	6.80E-21
2011mTn5STM.3.09.F10	S. meliloti chromosome, 110795	intergenic	-0.07	5.10E-02	-0.08	5.63E-03	0.10	1.15E-02	-0.05	2.05E-01
2011mTn5STM.3.11.G10	SMc03025	flagellum-specific ATP synthase protein FliI	0.15	1.31E-03	0.11	1.90E-04	0.40	4.49E-16	0.56	1.70E-21
2011mTn5STM.3.08.H10	SMb20360	hypothetical protein	-0.22	1.42E-04	-0.24	1.02E-06	1.82	7.28E-20	1.63	2.20E-17
2011mTn5STM.3.09.A11	SMc01968	putative transcription regulator protein	0.09	1.22E-01	0.12	3.09E-03	0.41	7.09E-18	0.37	3.65E-13
2011mTn5STM.3.12.B11	pSymB, 1003846	intergenic	0.66	1.51E-02	0.41	2.53E-03	2.12	1.37E-16	2.01	1.43E-22
2011mTn5STM.3.07.C11	pSymA, 291094	intergenic	-0.34	3.99E-06	-0.36	8.24E-11	-0.03	5.56E-01	-0.19	6.91E-05
2011mTn5STM.3.08.D11	SMb20377	putative translation initiation inhibitor protein	-0.84	3.46E-05	-0.91	4.25E-07	0.53	3.35E-20	0.52	2.89E-09
2011mTn5STM.3.09.E11	SMc04208	putative toxin secretion transmembrane protein	-0.13	1.83E-01	-0.24	9.45E-01	NA	NA	NA	NA
2011mTn5STM.3.12.F11	SMc01946	leucine-specific binding protein precursor	-0.24	1.35E-04	-0.16	7.97E-05	0.26	1.41E-03	0.21	3.22E-04
2011mTn5STM.3.11.G11	S. meliloti chromosome, 2065862	intergenic	-0.65	3.46E-03	0.02	9.98E-01	NA	NA	NA	NA
2011mTn5STM.3.06.H11	SMc01700	peptidyl-prolyl cis-trans isomerase A	-0.64	5.75E-06	-0.33	1.66E-13	-0.24	9.17E-11	-0.23	1.41E-13
2011mTn5STM.3.08.A12	SMc00779	probable FAD-linked oxidoreductase	0.14	1.14E-01	0.09	8.90E-04	0.33	2.01E-10	0.34	8.50E-16
2011mTn5STM.3.12.B12	SMb20476	putative ABC transporter periplasmic dipeptide-binding protein	0.77	4.42E-10	0.59	8.79E-13	-0.06	5.70E-02	0.18	2.55E-07
2011mTn5STM.3.09.C12	SMc00808	ChrA chromate transport protein	-0.63	5.48E-10	-0.37	9.55E-11	0.27	6.84E-05	0.18	2.94E-06
2011mTn5STM.3.07.D12	SMc00334	Cmk cytidilate kinase protein	-1.10	2.01E-06	-0.99	4.31E-11	0.88	4.27E-09	0.57	5.81E-05

2011mTn5STM.3.09.E12	S. meliloti chromosome, 2288070	intergenic	-0.31	1.62E-02	0.01	3.38E-01	0.51	1.15E-03	NA	NA
2011mTn5STM.3.08.F12	SMa0470	putative ABC transporter, ATP-binding protein	-0.16	6.42E-04	-0.03	4.78E-01	0.18	2.42E-04	-0.06	2.22E-04
2011mTn5STM.3.06.G12	SMc01340	putative aminotransferase protein	-0.38	4.43E-06	-0.25	5.82E-07	0.03	4.86E-01	-0.08	6.41E-04
2011mTn5STM.3.05.H12	SMc03198	molybdenum transport ATP-binding ABC transporter protein ModC	-0.27	4.08E-05	-0.20	2.04E-06	0.33	5.01E-12	0.42	2.64E-21
2011mTn5STM.4.01.A01	SMc00917	ATP phosphoribosyltransferase protein HisG	0.21	4.11E-01	-0.34	2.83E-02	-0.02	6.14E-01	0.48	2.93E-04
2011mTn5STM.4.10.B01	SMa1141	putative fnr/crp family transcriptional regulator	0.12	1.59E-03	0.16	8.60E-04	-0.10	5.84E-04	-0.26	7.33E-01
2011mTn5STM.4.02.C01	SMc02893	probable transcriptional regulator	-0.09	1.66E-01	-0.08	1.69E-01	0.28	1.77E-09	0.22	6.42E-07
2011mTn5STM.4.03.D01	SMc01549	hypothetical protein	0.06	6.21E-02	0.12	2.48E-05	NA	NA	-0.13	9.70E-01
2011mTn5STM.4.10.E01	SMa1077	Nex18 Symbiotically induced conserved protein	0.35	2.19E-05	0.26	1.49E-07	-0.19	1.31E-07	-0.22	1.77E-04
2011mTn5STM.4.07.F01	SMa0166	hypothetical protein	-0.14	3.11E-02	-0.04	1.04E-01	NA	NA	0.08	7.86E-01
2011mTn5STM.4.06.G01	SMb21313	bifunctional glycosyltransferase, forming alpha-glycosyl and beta-glycosyl linkages protein ExpE2	0.02	9.89E-01	0.16	1.10E-04	0.19	4.94E-06	0.28	7.45E-06
2011mTn5STM.4.01.H01	SMc03776	glutamate-5-kinase protein ProB1	-0.58	3.68E-05	-1.24	1.07E-16	0.26	2.13E-07	0.04	2.16E-01
2011mTn5STM.4.04.B02	SMa1131	hypothetical protein	0.05	4.39E-01	0.14	6.66E-05	0.10	2.44E-01	0.07	2.75E-01
2011mTn5STM.4.01.C02	SMa0229	hypothetical protein	0.17	7.80E-03	0.13	1.90E-04	0.07	1.39E-02	0.06	4.48E-03
2011mTn5STM.4.04.D02	SMb21133	putative sulfate uptake ABC transporter periplasmic solute-binding protein	0.01	6.75E-01	-0.09	1.27E-02	-0.27	1.87E-16	-0.39	1.58E-19
2011mTn5STM.4.08.E02	SMc00080	hypothetical protein	-0.04	2.28E-01	0.12	1.98E-03	-0.20	8.20E-09	0.04	2.25E-02
2011mTn5STM.4.07.F02	SMa0695	ArcB catabolic ornithine carbamoyl transferase	-0.07	5.18E-01	-0.02	3.19E-01	-0.13	7.32E-04	-0.26	9.64E-01
2011mTn5STM.4.02.G02	SMb20868	putative two-component sensor histidine kinase protein	0.14	5.43E-04	0.14	1.42E-05	-0.11	1.41E-06	-0.15	6.23E-07
2011mTn5STM.4.08.H02	pSymB, 1408087	intergenic	0.10	8.41E-02	-0.09	5.20E-03	-0.16	3.32E-11	-0.11	3.55E-07
2011mTn5STM.4.08.A03	SMb21536	hypothetical protein	0.24	3.66E-02	0.41	1.46E-10	0.09	1.08E-02	0.14	9.00E-05
2011mTn5STM.4.07.B03	S. meliloti chromosome, 2005840	intergenic	0.19	7.88E-04	0.19	5.43E-09	-0.25	3.99E-14	-0.02	9.39E-01
2011mTn5STM.4.09.C03	SMc00606	putative ribonuclease protein	-0.05	3.19E-01	-0.01	9.10E-01	-0.05	6.75E-03	-0.01	2.36E-02

2011mTn5STM.4.01.D03	SMc00886	hypothetical protein	-0.07	5.96E-02	0.05	7.60E-01	0.10	3.21E-01	0.14	9.15E-04
2011mTn5STM.4.09.E03	SMc02117	cystathionine beta-lyase protein MetC	0.15	5.64E-02	-0.60	1.64E-06	-0.02	7.40E-01	0.20	3.94E-04
2011mTn5STM.4.08.F03	SMa1913	putative transport protein	0.15	1.73E-03	0.11	5.37E-01	-0.27	4.69E-09	-0.43	4.38E-04
2011mTn5STM.4.02.G03	SMc00828	hypothetical protein	-0.08	5.52E-02	-0.43	1.84E-08	-0.15	3.47E-09	-0.11	2.58E-01
2011mTn5STM.4.01.H03	SMb20151	hypothetical protein	0.18	1.47E-02	0.23	2.43E-10	0.26	1.33E-16	0.31	1.10E-20
2011mTn5STM.4.08.A04	SMc02139	hypothetical protein	-0.16	7.68E-02	0.20	1.70E-04	0.30	3.97E-09	0.21	6.10E-08
2011mTn5STM.4.05.B04	SMa1371	probable ABC transporter ATP-binding protein	0.17	1.92E-02	0.41	2.50E-04	0.05	8.50E-02	-0.02	1.87E-01
2011mTn5STM.4.12.C04	SMa1465	putative ABC transporter permease	-0.08	7.12E-02	-0.04	1.06E-01	-0.12	4.00E-05	0.01	9.02E-01
2011mTn5STM.4.04.D04	SMb20481	AsnO asparagine synthase, glutamine-hydrolyzing	-0.06	1.85E-01	-0.13	7.32E-03	-0.11	1.28E-07	-0.13	3.29E-11
2011mTn5STM.4.02.E04	pSymA, 1165092	intergenic	0.42	2.37E-05	0.21	2.15E-07	-0.06	2.52E-04	-0.17	2.50E-07
2011mTn5STM.4.09.F04	SMa2115	Gst13 glutathione S-transferase	0.08	2.45E-01	0.02	2.57E-01	-0.38	5.63E-10	-0.46	1.12E-20
2011mTn5STM.4.03.G04	SMa0752	FAD/NAD(P)-binding oxidoreductase protein	0.01	6.15E-01	-0.07	2.87E-02	-0.27	3.71E-14	-0.39	4.87E-26
2011mTn5STM.4.10.H04	SMc00428	hypothetical protein	-0.23	6.96E-05	-0.12	2.90E-04	0.33	1.68E-10	0.45	1.11E-23
2011mTn5STM.4.01.A05	pSymB, 376527	intergenic	0.22	2.40E-02	0.39	7.25E-09	0.16	4.34E-04	0.25	1.17E-09
2011mTn5STM.4.06.B05	SMc00127	choline sulfatase protein BetC	0.16	1.02E-01	0.37	8.64E-07	0.16	8.56E-05	0.29	1.00E-06
2011mTn5STM.4.03.C05	SMc00937	TRK system potassium uptake transmembrane protein TrkH	-0.05	7.04E-01	-0.22	4.88E-07	-0.30	4.06E-14	-0.46	7.14E-29
2011mTn5STM.4.05.D05	SMa0222	putative GntR-family transcriptional regulator	-0.08	2.54E-01	0.07	2.97E-01	0.45	1.43E-09	0.25	1.76E-15
2011mTn5STM.4.11.E05	SMb20534	hypothetical protein	0.21	1.03E-04	0.03	2.32E-01	NA	NA	-0.26	1.42E-03
2011mTn5STM.4.11.F05	SMc02149	hypothetical protein	0.97	3.49E-05	0.77	5.33E-08	NA	NA	0.42	1.25E-03
2011mTn5STM.4.05.G05	S. meliloti chromosome, 532543	intergenic	-0.06	2.94E-01	-0.02	5.27E-01	-0.03	1.82E-01	0.01	8.95E-01
2011mTn5STM.4.02.H05	SMc00523	hypothetical protein	0.09	2.98E-01	0.16	1.70E-04	0.09	1.38E-04	0.14	3.04E-09
2011mTn5STM.4.09.A06	SMb20149	hypothetical protein	0.29	1.74E-05	0.19	2.34E-06	0.14	3.22E-08	0.10	5.02E-07
2011mTn5STM.4.03.B06	SMc00781	methylmalonate-semialdehyde dehydrogenase MmsA (IolA)	-0.10	1.46E-02	0.13	5.40E-04	-0.24	2.88E-11	-0.24	8.68E-16
2011mTn5STM.4.03.C06	SMc01267	hypothetical protein	-0.14	2.66E-02	-0.08	6.00E-04	-0.31	5.94E-08	-0.28	3.13E-20

2011mTn5STM.4.04.D06	SMb20227	nutrient deprivation-induced protein A (NdiA1)	-0.13	7.63E-04	-0.23	2.86E-08	-0.14	4.25E-08	-0.05	2.18E-02
2011mTn5STM.4.02.E06	SMc02226	hypothetical protein	0.09	2.44E-02	0.17	2.04E-03	0.25	2.69E-05	0.39	2.70E-19
2011mTn5STM.4.05.F06	SMb20820	putative ferredoxin reductase MocF	0.10	1.10E-02	0.10	1.92E-05	-0.24	8.48E-07	-0.46	8.46E-23
2011mTn5STM.4.07.G06	SMc00955	hypothetical protein	0.13	8.36E-02	0.26	3.72E-05	-0.11	4.06E-03	-0.08	1.58E-04
2011mTn5STM.4.05.H06	SMa1480	probable LysR-type activator	-0.15	9.57E-03	-0.13	6.50E-04	-0.19	1.67E-11	-0.07	1.95E-02
2011mTn5STM.4.11.A07	SMa1734	hypothetical protein	0.18	9.06E-04	-0.07	5.97E-02	-0.18	3.44E-09	-0.34	3.87E-29
2011mTn5STM.4.03.B07	SMa0417	hypothetical protein	0.19	2.37E-04	0.10	1.32E-02	-0.10	1.07E-04	-0.13	3.30E-07
2011mTn5STM.4.07.C07	SMb20941	msbA-like saccharide exporting ABC transporter protein	-0.21	4.39E-04	-0.15	1.20E-04	-0.20	3.58E-11	-0.14	1.45E-05
2011mTn5STM.4.07.D07	SMb21592	putative sugar uptake ABC transporter ATP-binding protein	-0.11	1.30E-01	-0.18	4.48E-06	NA	NA	NA	NA
2011mTn5STM.4.09.E07	SMc00893	hypothetical protein	0.37	3.00E-03	0.22	7.23E-06	0.05	3.54E-01	0.05	2.32E-03
2011mTn5STM.4.03.F07	S. meliloti chromosome, 2051945	intergenic	-0.16	4.09E-03	0.04	1.49E-01	NA	NA	0.22	1.72E-02
2011mTn5STM.4.09.G07	SMa2279	hypothetical protein	-0.10	9.55E-02	-0.12	7.43E-03	0.15	6.28E-08	-0.03	4.77E-01
2011mTn5STM.4.04.H07	S. meliloti chromosome, 81040	intergenic	-0.20	6.41E-02	-0.09	1.35E-03	0.20	5.80E-10	-0.11	5.04E-04
2011mTn5STM.4.08.A08	SMc02372	putative transport transmembrane protein	0.17	6.13E-04	0.18	1.70E-04	0.21	8.08E-03	0.15	2.33E-04
2011mTn5STM.4.12.B08	SMa1355	hypothetical protein	0.18	1.78E-03	0.12	2.60E-04	-0.31	3.92E-18	-0.31	3.23E-17
2011mTn5STM.4.06.C08	SMa0021	hypothetical protein	-0.33	7.70E-03	-0.18	3.60E-04	0.10	1.57E-01	-0.04	5.09E-02
2011mTn5STM.4.04.D08	SMa1782	putative LysR-family transcriptional regulator	0.31	1.78E-04	0.03	5.20E-01	-0.14	3.10E-07	-0.21	4.29E-15
2011mTn5STM.4.01.E08	SMb20925	hypothetical protein	0.16	6.40E-03	0.06	1.80E-01	-0.23	3.00E-03	-0.16	2.68E-05
2011mTn5STM.4.01.F08	SMc01611	ferrichrome-iron receptor precursor protein	-0.03	6.72E-01	0.01	3.79E-01	-0.14	5.46E-08	-0.09	2.67E-06
2011mTn5STM.4.10.G08	S. meliloti chromosome, 2816977	intergenic	-0.03	7.61E-01	0.03	4.74E-01	0.13	1.19E-06	0.07	2.78E-01
2011mTn5STM.4.08.H08	SMb20007	KatC catalase protein	-0.08	1.70E-01	-0.13	4.90E-04	0.36	6.53E-10	0.43	4.30E-14
2011mTn5STM.4.03.A09	SMc00825	glutamate-cysteine ligase GshA	-0.36	2.16E-02	-0.68	6.90E-04	NA	NA	NA	NA
2011mTn5STM.4.05.B09	SMa0574	hypothetical protein	-0.04	2.62E-01	0.09	9.31E-03	-0.06	5.17E-02	-0.10	1.71E-04

2011mTn5STM.4.11.C09	S. meliloti chromosome, 3123223	intergenic	-0.24	7.62E-04	-0.10	4.18E-02	0.33	6.35E-10	0.34	2.81E-15
2011mTn5STM.4.08.D09	SMa1084	hypothetical protein	-0.32	4.37E-06	-0.43	2.50E-15	-0.34	9.78E-14	-0.40	3.50E-21
2011mTn5STM.4.08.E09	SMa1163	putative cation transport P-type ATPase	0.08	2.10E-02	-0.06	7.97E-02	-0.15	2.96E-04	-0.16	2.09E-06
2011mTn5STM.4.01.F09	SMa0734	hypothetical protein	0.07	1.46E-02	0.06	2.93E-01	-0.16	3.20E-02	-0.33	1.90E-11
2011mTn5STM.4.07.G09	SMa1582	hypothetical protein	-0.21	1.10E-04	-0.25	4.64E-08	-0.27	3.33E-15	-0.34	9.84E-27
2011mTn5STM.4.02.H09	SMb20049	putative elongation factor G protein	-0.06	1.31E-01	-0.01	2.48E-01	0.15	3.57E-05	0.14	1.74E-03
2011mTn5STM.4.04.A10	pSymA, 248361	intergenic	-0.16	5.18E-03	-0.16	4.85E-05	-0.01	7.95E-01	-0.08	6.32E-03
2011mTn5STM.4.11.B10	SMb21300	putative deoxyribose-phosphate aldolase protein	-0.10	6.51E-02	-0.01	3.01E-01	0.02	7.67E-01	-0.01	4.60E-02
2011mTn5STM.4.11.C10	SMa0091	hypothetical protein	-0.57	6.69E-08	-0.44	1.45E-13	0.18	1.86E-04	0.16	3.10E-04
2011mTn5STM.4.02.D10	SMa1362	putative inner-membrane permease	-0.36	6.49E-06	-0.37	1.10E-12	-0.18	2.48E-10	-0.25	3.89E-15
2011mTn5STM.4.06.E10	SMb20909	hypothetical protein	0.11	4.04E-02	-0.12	2.97E-03	-0.28	1.97E-02	NA	NA
2011mTn5STM.4.12.F10	SMc02419	putative peptide transport system permease protein	-0.01	7.53E-01	-0.08	1.81E-02	-0.10	7.34E-05	-0.01	3.88E-01
2011mTn5STM.4.06.G10	SMa0769	FixP2 cytochrome c oxidase	-0.21	9.62E-05	-0.36	2.80E-11	-0.32	1.05E-08	-0.56	4.65E-26
2011mTn5STM.4.04.H10	SMc02273	RkpA fatty acid synthase protein	-0.28	2.10E-06	-0.23	8.62E-06	-0.35	1.18E-23	0.05	5.25E-04
2011mTn5STM.4.01.A11	SMa1397	hypothetical protein	-0.08	1.91E-01	0.07	1.48E-01	0.01	8.11E-01	0.07	7.07E-02
2011mTn5STM.4.05.B11	SMa0583	NrtB, Nitrate transport permease protein	-0.30	2.49E-05	-0.08	7.91E-01	NA	NA	NA	NA
2011mTn5STM.4.01.C11	SMc02597	superoxide dismutase Cu-Zn precursor protein SodC	-0.56	1.26E-09	-0.52	1.32E-10	0.00	4.02E-01	-0.11	1.69E-03
2011mTn5STM.4.09.D11	SMc03873	RNA polymerase sigma factor protein SigC	-0.57	4.38E-07	-0.49	6.43E-09	0.36	2.21E-11	0.22	4.15E-06
2011mTn5STM.4.04.E11	SMb20942	UDP glucose 4-epimerase protein ExoB	-0.08	4.51E-02	-0.17	1.90E-04	-0.34	1.01E-14	-0.36	8.65E-15
2011mTn5STM.4.09.F11	SMa0592	hypothetical protein	-0.27	1.71E-04	-0.08	1.11E-01	0.11	8.42E-02	-0.04	2.70E-03
2011mTn5STM.4.01.G11	SMb20355	hypothetical protein	-0.42	4.25E-07	-0.48	2.73E-13	-0.24	5.84E-04	-0.12	1.70E-03
2011mTn5STM.4.08.H11	SMc00339	adenylate cyclase 1 protein CyaA	-0.28	1.44E-04	-0.28	1.50E-08	0.37	1.06E-18	0.40	1.44E-26
2011mTn5STM.4.04.A12	SMb21165	putative histidine ammonia-lyase histidase protein HutH1	-0.11	5.37E-01	-0.20	1.62E-06	-0.02	4.66E-01	0.03	6.30E-01
2011mTn5STM.4.09.B12	SMc02144	phosphate transport system permease ABC transporter protein	-0.91	1.46E-08	-0.66	1.07E-11	-0.71	4.16E-26	-0.64	4.17E-23

2011mTn5STM.4.04.C12	SMc01053	siroheme synthase protein CysG	-0.29	1.74E-04	-1.35	1.40E-21	-0.05	2.99E-02	0.01	5.69E-01
2011mTn5STM.4.07.D12	SMa0070	ABC transporter permease	-0.51	5.96E-05	-0.48	1.80E-13	0.17	1.51E-06	0.11	1.12E-02
2011mTn5STM.4.08.E12	SMb21158	putative transcriptional regulator, probably of sugar phosphate metabolism protein	-0.27	1.28E-04	-0.22	1.62E-07	-0.29	5.40E-24	-0.20	9.56E-18
2011mTn5STM.4.08.F12	SMb20018	probable sugar transport ATP-binding protein	-0.05	2.87E-01	0.07	1.73E-01	0.37	3.46E-12	0.44	2.15E-18
2011mTn5STM.4.03.G12	SMc00775	FbpB ABC iron transporter, permease component	-0.31	5.84E-04	-0.26	3.20E-08	0.03	3.43E-01	0.09	2.68E-03
2011mTn5STM.4.07.H12	SMa0398	probable HisD2 histidinol dehydrogenase	-0.32	3.32E-06	-0.45	4.08E-16	-0.14	1.71E-04	-0.37	5.59E-07

[a] For mutants carrying a transposon insertion in an intergenic region the exact position of insertion in the respective replicon is given.

TABLE G.3. RESULTS OF *IN PLANTA* STM EXPERIMENT FOR SET 1

mutantID complete	aliases	function	Slide 4, combined DNA			slides 1-3		
			n	m-value	p-value(t)	n	m-value	p-value(t)
2011mTn5STM.1.13.A10	SMa0364	hypothetical protein	16	-0.51	1.8E-12	40	-0.48	2.7E-15
2011mTn5STM.1.08.B10	SMb21578	copper-transporting P-type ATPase protein AtcU2	16	-0.59	1.1E-14	39	-0.59	1.7E-17
2011mTn5STM.1.11.C10	SMb20070	putative sulfate permease protein	16	0.35	1.1E-05	38	0.52	9.0E-13
2011mTn5STM.1.12.E10	SMc00349	GTP-binding membrane protein LepA	16	-0.97	8.9E-16	40	-0.94	4.6E-32
2011mTn5STM.1.08.F10	SMc02406	hypothetical protein	0	NA	NA	40	0.49	4.2E-11
2011mTn5STM.1.13.G10	SMa0104	ABC transporter. solute-binding protein	0	NA	NA	40	0.22	7.3E-02
2011mTn5STM.1.04.H10	SMa0252	hypothetical protein	16	0.60	1.2E-06	28	0.59	1.6E-07
2011mTn5STM.1.02.A11	SMc00413	hypothetical protein	0	NA	NA	24	0.32	4.7E-03
2011mTn5STM.1.07.B11	SMb20967	putative transcriptional regulator, lacI family protein	16	-0.08	2.6E-03	40	-0.22	3.6E-03
2011mTn5STM.1.08.C11	SMc04346	IlvC ketol-acid reductoisomerase protein	11	-0.83	2.3E-05	19	-0.78	4.1E-09
2011mTn5STM.1.05.D11	SMa0814	NifB FeMo cofactor biosynthesis protein	16	-1.06	1.2E-16	40	-1.09	3.9E-36
2011mTn5STM.1.01.E11	SMa0765	FixN2 cytochrome c oxidase polypeptide I	16	-0.61	1.2E-12	24	-0.27	3.6E-03
2011mTn5STM.1.06.F11	SMa0922	hypothetical protein	16	0.50	8.1E-10	40	0.12	3.1E-01
2011mTn5STM.1.12.G11	SMa1195	hypothetical protein	16	0.32	5.2E-05	40	0.30	7.6E-10
2011mTn5STM.1.05.H11	SMa1493	putative LysR-type transcription factor	12	-0.76	1.0E-09	12	-0.37	1.3E-03
2011mTn5STM.1.11.A12	SMc01664	putative transcription regulator protein	16	-0.36	2.9E-09	40	-0.57	1.1E-09
2011mTn5STM.1.11.B12	SMc00322	hypothetical protein	16	-0.42	2.4E-06	39	-0.37	3.8E-08
2011mTn5STM.1.11.C12	SMa0489	putative ABC transporter. ATP-binding protein	16	-0.19	3.7E-05	40	-0.12	5.5E-03
2011mTn5STM.1.02.D12	SMa1288	hypothetical protein	16	0.06	1.6E-01	39	-0.18	1.0E-01
2011mTn5STM.1.11.E12	SMc04337	hypothetical protein	16	0.56	1.4E-05	40	0.36	1.3E-05
2011mTn5STM.1.04.F12	SMc04363	hypothetical protein	16	0.53	3.1E-07	40	0.35	6.4E-06
2011mTn5STM.1.04.H12	SMb21641	regulator of phenylacetic acid degradation PaaX	16	1.13	4.6E-14	40	0.67	1.1E-08
2011mTn5STM.1.10.A01	SMb21639	phenylacetic acid degradation protein PaaB	16	0.37	2.9E-07	40	0.24	1.9E-02
2011mTn5STM.1.09.B01	SMa0675	cation (Ca) exchange protein, possible	16	-0.70	1.5E-13	40	-0.71	1.5E-17
2011mTn5STM.1.09.C01	SMc00513	putative amino-acid binding periplasmic protein	16	0.91	9.8E-12	16	1.10	2.8E-09
2011mTn5STM.1.02.D01	SMa0259	hypothetical protein	16	0.17	4.4E-02	24	-0.42	2.7E-02
2011mTn5STM.1.01.E01	SMb20834	lipopolysaccharide processing protein RkpZ1	16	-0.49	1.2E-08	40	-0.51	2.4E-15
2011mTn5STM.1.13.F01	SMa0113	Sensory histidine kinase	13	0.45	2.3E-04	34	0.40	1.1E-06
2011mTn5STM.1.02.G01	SMa1497	putative Gst12 glutathione-S-transferase	0	NA	NA	24	0.22	3.7E-05
2011mTn5STM.1.13.H01	SMb20167	hypothetical protein	16	-0.04	4.7E-02	40	-0.09	1.9E-01
2011mTn5STM.1.13.B02	SMc02327	ribonuclease HI protein RnhA2	16	0.36	4.8E-09	40	0.30	5.0E-03
2011mTn5STM.1.09.C02	SMa1903	putative protease	16	0.08	2.1E-02	28	0.27	2.4E-07
2011mTn5STM.1.03.D02	SMc02466	SdhB succinate dehydrogenase iron-sulfur protein	12	0.63	2.6E-11	28	0.31	2.3E-03
2011mTn5STM.1.11.E02	SMc02738	ABC cholin transporter, permease component (OpuB)	16	-0.21	1.1E-05	40	-0.41	2.6E-03
2011mTn5STM.1.08.F02	SMa0525	putative ABC-type iron transport system protein	16	-0.48	7.0E-12	40	-0.43	5.8E-10
2011mTn5STM.1.07.G02	S. meliloti chromosome, 2021711	intergenic	16	-0.33	6.1E-08	40	-0.52	5.3E-09
2011mTn5STM.1.06.H02	SMc01090	ATP-dependent RNA-helicase protein DeaD	0	NA	NA	40	-0.09	5.9E-01

2011mTn5STM.1.10.A03	SMc01046	Trk system potassium uptake protein	16	0.39	9.4E-09	40	0.35	6.3E-15
2011mTn5STM.1.01.B03	SMc01432	hypothetical protein	16	-0.03	6.2E-02	40	-0.22	6.7E-02
2011mTn5STM.1.05.C03	SMb20071	putative efflux protein	16	-0.67	3.1E-11	23	-0.66	2.9E-09
2011mTn5STM.1.03.D03	SMb20349	putative ABC transporter	16	-0.40	6.3E-08	40	-0.50	1.4E-07
2011mTn5STM.1.11.E03	SMa0300	ABC transporter. permease	16	-0.47	2.7E-10	16	-0.18	7.7E-06
2011mTn5STM.1.06.F03	SMc00300	putative oxidoreductase protein	16	0.50	4.0E-09	40	0.52	1.1E-13
2011mTn5STM.1.08.G03	pSymB, 543971	intergenic	16	0.15	1.3E-04	24	0.11	1.5E-01
2011mTn5STM.1.08.H03	SMa1236	NapA periplasmic nitrate reductase	16	-0.36	5.2E-09	24	-0.13	7.3E-03
2011mTn5STM.1.11.A04	SMa1004	hypothetical protein	16	-0.83	2.0E-13	40	-0.86	3.1E-23
2011mTn5STM.1.06.B04	SMb21707	putative ureashort-chain amide or branched-chain amino acid uptake ABC transporter, ATP-binding protein	16	0.17	5.2E-04	40	0.11	2.2E-01
2011mTn5STM.1.01.C04	SMa1229	FixL Oxygen regulated histidine kinase	16	-1.11	1.7E-15	39	-1.11	8.3E-29
2011mTn5STM.1.03.D04	SMa1531	NuoC2 NADH I chain C	16	-0.10	1.5E-02	40	-0.33	4.1E-05
2011mTn5STM.1.13.E04	SMa1503	hypothetical protein	16	0.03	1.7E-01	40	-0.12	1.3E-02
2011mTn5STM.1.02.F04	SMa1292	hypothetical protein	0	NA	NA	40	-0.30	1.5E-06
2011mTn5STM.1.01.G04	SMc01497	sorbitol/mannitol transporter SmoF	16	0.81	3.5E-09	34	0.48	2.3E-02
2011mTn5STM.1.07.H04	SMb21422	putative sugar uptake ABC transporter ATP-binding protein	16	-0.24	1.4E-05	40	-0.11	2.5E-02
2011mTn5STM.1.02.A05	SMb21645	ABC transporter. permease component	16	-0.32	8.3E-07	40	-0.35	3.0E-11
2011mTn5STM.1.07.B05	SMb20142	putative oligopeptide ABC transporter permease protein	16	0.38	2.3E-06	40	0.43	1.3E-12
2011mTn5STM.1.03.C05	SMc02121	general L-amino acid transport ATP-binding ABC transporter protein AapP	16	0.51	4.3E-08	40	0.51	6.8E-10
2011mTn5STM.1.08.D05	SMb20219	putative response regulator protein	16	-0.07	2.9E-03	40	-0.05	1.5E-01
2011mTn5STM.1.09.E05	SMb21527	putative taurin uptake ABC transporter ATP-binding protein TauB	16	0.20	1.2E-02	40	-0.04	7.2E-01
2011mTn5STM.1.12.F05	SMa0803	putative ABC-transporter ATP-binding protein	16	-0.26	7.3E-08	40	-0.38	1.9E-07
2011mTn5STM.1.12.G05	SMa0621	FixI2 E1-E2 type cation ATPase	16	0.79	5.4E-13	40	0.86	4.7E-19
2011mTn5STM.1.07.H05	SMa0875	NolG efflux transporter	0	NA	NA	32	0.11	7.6E-04
2011mTn5STM.1.09.A06	SMa1507	hypothetical protein	16	0.02	5.1E-01	40	0.09	7.9E-02
2011mTn5STM.1.04.B06	SMb20930	putative sugar uptake ABC transporter ATP-binding protein	16	0.20	1.7E-03	40	0.27	2.3E-03
2011mTn5STM.1.11.C06	SMc02507	Manganese ABC transporter permease SitC	16	-0.10	4.4E-02	40	-0.02	3.7E-01
2011mTn5STM.1.03.D06	SMa1760	Miscellaneous; Unknown	16	-0.44	5.2E-12	40	-0.38	9.8E-05
2011mTn5STM.1.06.E06	pSymA, 305100	intergenic	16	-0.16	5.8E-06	40	-0.30	8.5E-04
2011mTn5STM.1.12.F06	SMb20390	hypothetical protein	16	-0.34	5.0E-07	40	-0.27	5.0E-09
2011mTn5STM.1.08.G06	SMa1798	Kup2 Potassium uptake protein	16	-1.05	8.1E-16	24	-0.94	7.0E-19
2011mTn5STM.1.11.H06	SMa0214	putative KduI DKI isomerase	0	NA	NA	40	-0.17	1.8E-07
2011mTn5STM.1.12.B07	S. meliloti chromosome, 1067681	intergenic	12	0.19	2.4E-03	22	0.41	4.7E-02
2011mTn5STM.1.12.C07	SMa0803	putative ABC-transporter ATP-binding protein	0	NA	NA	0	NA	NA
2011mTn5STM.1.11.D07	SMb20819	Ferredoxin, Rieske [2Fe-2S] domain MocE	0	NA	NA	16	-0.34	4.0E-08
2011mTn5STM.1.08.E07	SMa1467	Probable inner-membrane permease	16	-0.31	1.2E-10	32	-0.84	4.4E-21
2011mTn5STM.1.11.F07	SMa1115	Putative manganese transport protein	16	0.00	6.2E-01	40	-0.01	4.3E-01
2011mTn5STM.1.04.G07	S. meliloti chromosome, 2181489	intergenic	16	0.22	5.6E-06	40	0.24	1.9E-02
2011mTn5STM.1.05.H07	S. meliloti chromosome, 1621189	intergenic	16	0.54	2.5E-15	40	0.48	1.8E-06

2011mTn5STM.1.02.A08	SMc00185	putative ABC-transporter ATP-binding protein	16	0.64	7.7E-10	40	0.49	3.2E-14
2011mTn5STM.1.03.B08	SMb20513	hypothetical protein	16	-0.21	4.5E-08	40	-0.31	2.4E-02
2011mTn5STM.1.05.C08	SMb21477	putative reverse transcriptase maturase protein	0	NA	NA	16	0.36	6.1E-04
2011mTn5STM.1.11.D08	SMa1638	hypothetical protein	16	-0.08	2.5E-02	40	-0.03	4.4E-01
2011mTn5STM.1.03.E08	SMc02236	hypothetical protein	0	NA	NA	40	0.20	1.1E-03
2011mTn5STM.1.09.F08	SMa1146	hypothetical protein	16	-0.38	9.0E-10	40	-0.33	2.8E-07
2011mTn5STM.1.12.G08	SMb20948	ExoU glucosyltransferase	15	-1.05	3.8E-14	40	-1.06	5.1E-27
2011mTn5STM.1.09.H08	SMa0809	hypothetical protein	16	0.08	2.4E-01	38	-0.12	1.8E-01
2011mTn5STM.1.04.A09	SMb20616	ThiO thiamine biosynthesis oxidoreductase protein	14	-0.98	1.7E-09	37	-0.88	6.2E-20
2011mTn5STM.1.01.B09	SMb20958	exopolysaccharide biosynthesis protein ExoM	13	-1.08	6.8E-14	34	-1.02	3.2E-20
2011mTn5STM.1.10.C09	SMa0244	hypothetical protein	0	NA	NA	11	-0.82	1.4E-05
2011mTn5STM.1.05.D09	SMa0840	NodD3 transcriptional regulator	0	NA	NA	16	-0.93	3.0E-11
2011mTn5STM.1.05.E09	SMc02773	sugar transport ATP-binding protein	15	0.27	9.7E-04	32	0.17	7.6E-02
2011mTn5STM.1.02.F09	SMc02612	glutamate synthase family protein GlxD	16	0.98	6.2E-12	40	0.82	1.2E-22
2011mTn5STM.1.03.G09	SMa1128	DegP4 protease like protein	15	0.00	6.9E-01	40	0.01	8.8E-01
2011mTn5STM.1.12.H09	SMb21087	conjugal transfer protein TraA2	0	NA	NA	0	NA	NA
2011mTn5STM.2.10.A10	S. meliloti chromosome, 2106334	intergenic	16	0.30	1.1E-06	40	0.30	9.9E-03
2011mTn5STM.2.04.B10	SMa0565	hypothetical protein	16	1.20	1.6E-08	40	1.29	3.3E-22
2011mTn5STM.2.04.C10	SMc00086	diheme cytochrome C-type signal peptide protein CycG	16	0.11	1.1E-02	24	0.13	8.6E-02
2011mTn5STM.2.03.D10	SMb20748	glycosyltransferase protein PssF	16	0.87	3.8E-07	37	0.75	8.3E-13
2011mTn5STM.2.11.E10	SMb21273	spermidineputrescine ABC transporter periplasmic solute-binding protein precursor	16	0.19	7.3E-03	40	0.43	4.1E-07
2011mTn5STM.2.09.F10	SMc01818	adenylate cyclase transmembrane protein CyaC	14	-0.02	2.4E-01	23	0.09	4.2E-02
2011mTn5STM.2.05.G10	SMb20522	hypothetical protein	16	-0.17	5.7E-04	40	0.01	6.6E-01
2011mTn5STM.2.01.H10	SMa1017	hypothetical protein	0	NA	NA	13	0.29	3.4E-04
2011mTn5STM.2.06.A11	SMc01042	nitrogen regulation protein NtrB	0	NA	NA	0	NA	NA
2011mTn5STM.2.06.B11	SMc01881	PanB 3-methyl-2-oxobutanoate hydromethyltransferase protein	10	2.84	1.6E-04	19	2.01	1.4E-08
2011mTn5STM.2.04.C11	SMb20515	putative CheR methyltransferase, SAM binding domain	13	-0.17	8.5E-02	35	-0.01	9.1E-01
2011mTn5STM.2.01.D11	SMc04452	NADH dehydrogenase transmembrane protein Ndh	16	1.17	1.8E-21	32	1.02	2.3E-19
2011mTn5STM.2.05.E11	SMa1122	hypothetical protein	16	-0.12	3.6E-01	34	0.04	9.1E-01
2011mTn5STM.2.04.F11	pSymA, 510651	intergenic	0	NA	NA	16	0.59	3.9E-04
2011mTn5STM.2.11.G11	SMb21345	putative sugar ABC transporter	16	0.23	2.4E-09	40	0.30	5.8E-13
2011mTn5STM.2.10.H11	SMb20549	hypothetical protein	16	0.11	3.9E-04	40	0.04	1.1E-02
2011mTn5STM.2.03.A12	SMc00739	hypothetical protein	16	0.74	2.6E-09	40	0.63	2.2E-07
2011mTn5STM.2.09.B12	S. meliloti chromosome, 2862221	intergenic	16	1.11	1.4E-07	24	1.32	1.7E-13
2011mTn5STM.2.04.C12	SMc01814	probable glutamate synthase small chain protein	16	0.49	2.2E-07	40	0.20	4.3E-02
2011mTn5STM.2.01.D12	pSymB. 610814	intergenic	16	-0.02	7.4E-01	24	-0.18	6.8E-06
2011mTn5STM.2.03.E12	SMb20548	hypothetical protein	16	0.24	1.0E-03	40	0.22	8.4E-06
2011mTn5STM.2.09.F12	SMc03110	hypothetical protein	0	NA	NA	16	-0.23	1.9E-02
2011mTn5STM.2.06.G12	SMc00094	betaine aldehyde dehydrogenase protein BetB	0	NA	NA	22	0.01	3.3E-01
2011mTn5STM.2.06.H12	SMc01711	hypothetical protein	16	0.22	9.8E-07	40	0.37	1.9E-08

2011mTn5STM.2.05.A01	S. meliloti chromosome, 2044300	intergenic	10	1.42	2.2E-04	11	0.38	2.6E-01
2011mTn5STM.2.07.B01	SMa1182	NosZ N2O reductase	16	0.27	2.5E-06	39	0.24	1.2E-03
2011mTn5STM.2.11.C01	SMa1677	hypothetical protein	16	0.02	9.8E-01	40	-0.11	1.8E-01
2011mTn5STM.2.10.D01	SMc01950	high-affinity branched-chain amino acid ABC transporter. permease protein LivM	15	-0.86	4.2E-08	37	-0.81	1.8E-14
2011mTn5STM.2.12.E01	SMc03249	hypothetical protein	16	-0.28	1.0E-06	40	-0.20	1.9E-08
2011mTn5STM.2.07.F01	SMa0607	hypothetical protein	13	-1.09	6.4E-03	20	-0.70	1.4E-04
2011mTn5STM.2.01.G01	pSymA, 854083	intergenic	16	-0.15	2.9E-03	40	-0.02	3.5E-01
2011mTn5STM.2.12.H01	SMc01219	lipopolysaccharide core biosynthesis mannosyltransferase LpsB	13	-0.91	6.1E-05	25	-0.84	4.8E-11
2011mTn5STM.2.07.A02	SMc01607	putative permease protein	16	-0.12	7.4E-04	40	-0.03	5.6E-01
2011mTn5STM.2.05.B02	SMb21144	putative choline uptake ABC transporter periplasmic solute-binding protein precursor	16	0.04	3.6E-01	40	0.01	2.1E-01
2011mTn5STM.2.05.C02	SMa0876	NolF secretion protein	14	-0.19	1.4E-03	34	-0.23	5.8E-03
2011mTn5STM.2.09.D02	SMa0322	hypothetical protein	16	-0.87	2.3E-13	40	-0.80	1.2E-23
2011mTn5STM.2.04.E02	SMb21633	PaaG enoyl-CoA hydratase protein	11	1.00	2.2E-05	36	0.74	1.7E-05
2011mTn5STM.2.05.F02	S. meliloti chromosome, 1627869	intergenic	14	0.82	6.6E-06	33	0.65	4.0E-08
2011mTn5STM.2.02.G02	SMb20273	hypothetical protein	16	-0.47	2.2E-09	40	-0.59	3.3E-12
2011mTn5STM.2.12.H02	SMb20851	putative transcriptional regulator, SorC family protein	16	-0.29	3.3E-04	37	-0.26	1.3E-04
2011mTn5STM.2.12.A03	SMb20209	hypothetical protein	16	-0.16	2.8E-04	37	-0.35	8.8E-06
2011mTn5STM.2.10.B03	S. meliloti chromosome, 3367064	intergenic	16	0.49	5.7E-06	24	0.45	1.4E-05
2011mTn5STM.2.02.C03	SMc00556	DNA repair protein RadA	13	-0.52	1.3E-04	36	-0.73	3.9E-07
2011mTn5STM.2.12.D03	SMa0485	probable ThrC2 threonine synthase	16	-0.30	1.2E-07	24	-0.26	3.7E-06
2011mTn5STM.2.01.E03	SMb20712	rhizopine uptake ABC transporter periplasmic solute-binding protein precursor MocB	0	NA	NA	40	0.10	3.1E-01
2011mTn5STM.2.07.F03	SMb20168	hypothetical protein	12	-0.03	3.5E-01	26	0.00	9.9E-01
2011mTn5STM.2.06.G03	SMc03044	chemotaxis protein MotD	15	0.32	2.5E-02	39	0.02	9.7E-01
2011mTn5STM.2.09.H03	SMa1833	hypothetical protein	0	NA	NA	16	-0.13	1.0E-02
2011mTn5STM.2.11.A04	SMc02270	capsular polysaccharide biosynthesis transmembrane protein RkpI	16	0.05	1.3E-02	40	-0.15	2.5E-01
2011mTn5STM.2.09.B04	SMb20112	hypothetical protein	16	-0.30	1.3E-06	24	-0.37	6.1E-07
2011mTn5STM.2.12.C04	SMc02659	GTP pyrophosphokinase protein RelA	0	NA	NA	0	NA	NA
2011mTn5STM.2.09.D04	SMb21301	putative membrane-anchored aldehyde dehydrogenase protein	0	NA	NA	16	-0.01	5.5E-01
2011mTn5STM.2.01.E04	SMa0306	putative histidine ammonia-lyase	16	-0.62	5.8E-12	22	-0.72	4.3E-08
2011mTn5STM.2.01.F04	SMa0431	hypothetical protein	16	0.00	9.3E-02	40	0.11	2.1E-02
2011mTn5STM.2.11.G04	SMb20472	probable carbamoyltransferase	16	-0.24	1.6E-06	38	-0.01	2.2E-01
2011mTn5STM.2.10.H04	SMa0110	ABC transporter. ATP-binding protein	16	-0.12	6.1E-02	40	-0.08	3.6E-01
2011mTn5STM.2.05.A05	SMb21199	oligopeptide uptake ABC transporter ATP-binding protein OppD	16	-0.18	2.8E-05	40	-0.04	1.9E-01
2011mTn5STM.2.10.B05	SMb21367	cytochrome c class I protein CycA	16	0.17	2.7E-04	16	0.85	3.1E-14
2011mTn5STM.2.11.C05	SMa0250	hypothetical protein	16	0.40	4.7E-04	40	0.16	4.2E-04
2011mTn5STM.2.08.D05	SMa0849	SyrM transcriptional regulator	16	-0.96	1.4E-18	24	-0.78	1.8E-13
2011mTn5STM.2.06.E05	SMb21177	phosphate uptake ABC transporter ATP-binding protein	16	-0.40	1.4E-09	40	-0.54	2.2E-13
2011mTn5STM.2.06.F05	pSymB, 1225168	intergenic	16	-0.12	6.7E-01	19	-0.11	6.5E-01
2011mTn5STM.2.06.G05	SMa1766	hypothetical protein	15	-0.61	1.4E-06	38	-0.45	1.7E-10

2011mTn5STM.2.05.H05	SMc03164	xylulose kinase protein XylB	16	0.78	5.2E-08	24	1.10	4.4E-11
2011mTn5STM.2.07.A06	SMb20866	hypothetical protein	16	0.01	6.8E-01	40	0.13	5.3E-02
2011mTn5STM.2.11.B06	SMb20111	putative oligopeptide ABC transporter ATP-binding protein	16	-0.29	8.1E-07	40	-0.19	2.1E-02
2011mTn5STM.2.01.C06	SMa2163	hypothetical protein	16	0.05	9.0E-02	40	0.18	9.2E-05
2011mTn5STM.2.06.D06	SMc00864	hypothetical protein	16	0.68	1.0E-09	24	0.83	3.4E-09
2011mTn5STM.2.03.E06	SMb21174	phosphate uptake ABC transporter permease protein	16	0.07	2.8E-02	23	0.33	8.6E-07
2011mTn5STM.2.10.F06	SMb21000	putative transport protein	14	0.27	2.7E-03	34	0.36	1.0E-09
2011mTn5STM.2.05.G06	SMb20943	acetyltransferase protein ExoZ	16	-1.12	2.0E-14	31	-1.05	3.1E-22
2011mTn5STM.2.11.H06	SMb20114	hypothetical protein	16	0.05	4.3E-01	40	0.06	7.9E-01
2011mTn5STM.2.07.A07	SMc04204	iron transport regulator transmembrane protein FecR	16	-0.03	1.2E-01	32	0.17	5.9E-06
2011mTn5STM.2.07.B07	SMb21493	putative transcriptional regulator. asnC family protein	16	-0.17	1.6E-03	40	-0.07	9.8E-03
2011mTn5STM.2.02.C07	SMc01528	dipeptide transport ATP-binding ABC transporter protein	10	0.21	4.7E-02	16	0.66	6.2E-10
2011mTn5STM.2.01.D07	SMa1957	hypothetical protein	16	-0.43	6.0E-11	40	-0.40	8.5E-06
2011mTn5STM.2.12.E07	SMc00383	glutatione S-transferase protein	12	0.19	1.9E-02	16	0.10	9.9E-03
2011mTn5STM.2.06.F07	SMc00774	acetoacetyl-coenzyme A synthetase, additional	16	-0.36	1.4E-07	36	-0.45	2.8E-07
2011mTn5STM.2.04.G07	SMb20123	putative transcriptional regulator protein	0	NA	NA	0	NA	NA
2011mTn5STM.2.09.H07	pSymB. 96435	intergenic	16	1.20	3.0E-11	24	1.69	1.8E-15
2011mTn5STM.2.04.A08	SMa0039	putative LysR-family transcriptional regulator	16	-0.18	1.3E-05	20	-0.05	3.9E-01
2011mTn5STM.2.04.B08	S. meliloti chromosome, 3364101	intergenic	16	-0.20	6.4E-04	40	-0.08	1.9E-01
2011mTn5STM.2.05.C08	SMa1956	putative LysR-family transcriptional regulator	16	-0.15	2.7E-04	40	0.00	8.8E-01
2011mTn5STM.2.04.D08	SMc03149	hypothetical protein	0	NA	NA	24	0.37	8.3E-10
2011mTn5STM.2.08.E08	SMb20757	methylmalonyl-CoA mutase protein BhbA	16	-0.84	3.9E-11	32	-0.87	1.4E-26
2011mTn5STM.2.12.F08	SMa1513	putative ABC transporter permease	16	-0.02	4.8E-01	38	0.08	3.2E-01
2011mTn5STM.2.11.G08	SMc00527	hypothetical protein	16	0.65	6.5E-11	36	0.55	5.5E-16
2011mTn5STM.2.10.H08	SMa0394	putative ABC transporter. permease	16	-0.07	1.0E-02	40	-0.03	2.9E-01
2011mTn5STM.2.11.A09	SMc02899	PheA prephenate dehydratase protein	0	NA	NA	0	NA	NA
2011mTn5STM.2.09.B09	SMc00148	hypothetical protein	16	-0.07	9.1E-03	32	-0.12	2.8E-02
2011mTn5STM.2.09.C09	SMb20959	probable exopolysaccharide biosynthesis protein ExoO	16	-1.05	9.8E-17	40	-1.08	1.8E-31
2011mTn5STM.2.09.D09	SMc04456	chaperon protein CsaA	16	0.28	3.0E-06	24	0.29	5.3E-06
2011mTn5STM.2.12.E09	SMc03782	hypothetical protein	16	0.10	2.1E-02	34	-0.31	8.6E-02
2011mTn5STM.2.07.F09	SMa0260	GabD3 succinate-semialdehyde dehdyrogenase	0	NA	NA	34	-0.02	8.5E-01
2011mTn5STM.2.10.G09	SMa1746	putative iron uptake protein	16	-0.75	1.1E-15	40	-0.65	3.0E-09
2011mTn5STM.2.03.H09	SMc02045	putative oxidoreductase protein	16	-0.08	2.7E-01	40	-0.26	1.0E-02
2011mTn5STM.3.08.A10	SMb20860	probable non-heme chloroperoxidase	16	-0.62	1.2E-09	24	-0.90	1.4E-13
2011mTn5STM.3.07.B10	SMa0316	hypothetical protein	16	-0.10	1.5E-02	40	-0.30	5.5E-04
2011mTn5STM.3.06.C10	SMa0615	FixO3 cytochrome-c oxidase subunit	16	-0.55	1.3E-11	24	-0.69	2.8E-07
2011mTn5STM.3.05.D10	SMa1916	hypothetical protein	16	-0.10	5.4E-02	38	-0.35	5.5E-05
2011mTn5STM.3.09.E10	SMc01000	hypothetical protein	16	-0.64	3.0E-10	38	-0.83	4.3E-19
2011mTn5STM.3.09.F10	S. meliloti chromosome, 110795	intergenic	13	0.62	8.2E-06	24	0.69	6.2E-11
2011mTn5STM.3.11.G10	SMc03025	flagellum-specific ATP synthase protein FliI	14	-0.31	6.4E-01	34	-0.47	1.5E-03

2011mTn5STM.3.08.H10	SMb20360	hypothetical protein	16	1.21	2.0E-06	24	1.21	3.0E-09
2011mTn5STM.3.09.A11	SMc01968	putative transcription regulator protein	16	0.55	1.7E-13	23	-0.16	4.3E-01
2011mTn5STM.3.12.B11	pSymB. 1003846	intergenic	16	1.53	2.2E-06	40	1.25	6.8E-11
2011mTn5STM.3.07.C11	pSymA. 291094	intergenic	16	0.32	1.7E-02	35	0.18	5.5E-02
2011mTn5STM.3.08.D11	SMb20377	putative translation initiation inhibitor protein	16	-0.05	2.0E-01	40	-0.30	4.0E-06
2011mTn5STM.3.09.E11	SMc04208	putative toxin secretion transmembrane protein	0	NA	NA	0	NA	NA
2011mTn5STM.3.12.F11	SMc01946	leucine-specific binding protein precursor	11	0.40	2.2E-05	24	0.39	9.2E-04
2011mTn5STM.3.11.G11	S. meliloti chromosome, 2065862	intergenic	0	NA	NA	28	0.84	2.2E-09
2011mTn5STM.3.06.H11	SMc01700	peptidyl-prolyl cis-trans isomerase A	16	-0.67	9.4E-16	24	-0.34	7.0E-07
2011mTn5STM.3.08.A12	SMc00779	probable FAD-linked oxidoreductase	16	0.08	6.7E-02	40	0.07	1.5E-01
2011mTn5STM.3.12.B12	SMb20476	putative ABC transporter periplasmic dipeptide-binding protein	16	0.83	3.4E-08	34	0.60	3.8E-06
2011mTn5STM.3.09.C12	SMc00808	ChrA chromate transport protein	16	0.27	7.4E-05	40	0.09	8.6E-02
2011mTn5STM.3.07.D12	SMc00334	Cmk cytidilate kinase protein	15	1.09	1.7E-04	22	0.80	1.4E-03
2011mTn5STM.3.09.E12	S. meliloti chromosome, 2288070	intergenic	10	0.38	7.2E-05	16	0.36	5.3E-04
2011mTn5STM.3.08.F12	SMa0470	putative ABC transporter. ATP-binding protein	16	-0.40	1.3E-13	24	-0.16	9.1E-05
2011mTn5STM.3.06.G12	SMc01340	putative aminotransferase protein	16	0.67	1.5E-09	35	0.45	1.3E-08
2011mTn5STM.3.05.H12	SMc03198	molybdenum transport ATP-binding ABC transporter protein ModC	16	0.65	7.8E-12	39	0.59	1.1E-05
2011mTn5STM.3.12.A01	SMa0819	FixB electron transfer flavoprotein alpha chain	0	NA	NA	40	-0.87	5.0E-20
2011mTn5STM.3.08.B01	SMc04217	hypothetical protein	16	0.71	6.0E-06	40	0.79	9.0E-15
2011mTn5STM.3.10.C01	SMc01842	putative methyltransferase transcription regulator protein	14	-0.51	3.0E-03	37	-0.43	1.0E-04
2011mTn5STM.3.10.D01	SMc00228	ATP-dependent DNA helicase protein RecG	14	0.29	3.4E-01	16	0.00	3.6E-01
2011mTn5STM.3.10.E01	SMc04386	aspartate aminotransferase B protein	16	0.27	2.1E-02	40	0.14	1.2E-01
2011mTn5STM.3.08.F01	SMa2333	KdpA potassium-transporting ATPase A chain	16	-0.39	1.0E-02	36	-0.56	9.7E-10
2011mTn5STM.3.05.G01	SMb20178	hypothetical protein	15	-0.20	2.0E-03	39	-0.37	4.0E-07
2011mTn5STM.3.11.A02	pSymA, 1215937	intergenic	16	-0.64	2.5E-13	32	-0.54	1.4E-11
2011mTn5STM.3.08.B02	SMc02691	putative membrane transport protein	16	0.38	1.1E-05	40	0.45	5.8E-06
2011mTn5STM.3.08.C02	SMb20121	hypothetical protein	14	0.01	7.8E-01	40	0.11	1.0E-02
2011mTn5STM.3.07.D02	SMc02376	putative heat shock protein	16	-0.62	1.6E-11	40	-0.56	3.6E-22
2011mTn5STM.3.09.E02	SMa2215	putative GntR-family transcriptional regulator	16	-0.47	8.0E-11	24	-0.69	7.0E-13
2011mTn5STM.3.10.F02	SMc00420	hypothetical protein	16	0.64	2.0E-04	39	0.47	1.5E-08
2011mTn5STM.3.08.G02	SMa2075	probable extracellular solute-binding protein	15	-0.06	3.6E-01	39	0.07	1.2E-01
2011mTn5STM.3.07.H02	SMc02363	cytochrome C-type biogenesis transmembrane protein CycK	0	NA	NA	14	-0.19	4.4E-01
2011mTn5STM.3.09.A03	S. meliloti chromosome, 2823999	intergenic	16	0.00	3.3E-01	40	0.16	1.2E-03
2011mTn5STM.3.11.B03	SMb20029	hypothetical protein	16	-0.05	5.9E-01	40	-0.07	1.7E-01
2011mTn5STM.3.11.C03	SMb20067	hypothetical protein	16	-0.11	6.0E-03	40	-0.23	2.0E-04
2011mTn5STM.3.09.D03	SMa0929	TraG conjugal transfer protein	16	0.65	6.7E-09	39	0.80	8.4E-16
2011mTn5STM.3.05.E03	SMc01552	hypothetical protein	16	0.02	9.5E-01	37	-0.19	3.6E-02
2011mTn5STM.3.08.F03	S. meliloti chromosome, 2178672	intergenic	16	0.01	8.9E-01	40	0.06	9.9E-01
2011mTn5STM.3.11.G03	SMa0981	NtrR2 transcription regulator	16	-0.45	5.0E-11	40	-0.37	3.8E-09

2011mTn5STM.3.08.H03	SMb20436	probable nitrate transporter	16	-0.12	3.3E-03	40	-0.14	2.2E-01
2011mTn5STM.3.07.A04	SMc01131	hypothetical protein	16	0.32	8.1E-03	30	-0.19	2.9E-01
2011mTn5STM.3.11.B04	SMb21431	hypothetical protein	16	-0.02	7.8E-01	38	0.08	2.1E-01
2011mTn5STM.3.06.C04	SMc01668	hypothetical protein	0	NA	NA	16	-0.01	2.0E-01
2011mTn5STM.3.10.D04	SMb20672	putative sugar ABC transporter	16	0.07	5.8E-02	24	-0.27	5.3E-06
2011mTn5STM.3.06.E04	SMa1179	NosR Regulatory protein for N2O reductase	16	-0.04	5.9E-01	34	-0.28	2.2E-02
2011mTn5STM.3.12.F04	SMb20009	putative transcriptional regulator	16	0.10	1.8E-03	39	0.20	1.2E-03
2011mTn5STM.3.06.G04	SMc00609	anaerobic dimethyl sulfoxide reductase chain A protein DmsA	16	-0.14	2.7E-02	38	-0.27	2.2E-02
2011mTn5STM.3.05.H04	SMb20152	hypothetical protein	0	NA	NA	24	0.06	4.8E-01
2011mTn5STM.3.11.A05	SMb20037	shikimate 5-dehydrogenase protein AroE2	12	1.20	2.1E-05	31	1.12	5.1E-12
2011mTn5STM.3.09.B05	SMb20931	putative sugar uptake ABC transporter periplasmic solute-binding protein precursor	16	0.42	3.3E-03	32	0.43	1.4E-09
2011mTn5STM.3.08.C05	SMa2293	probable beta lactamase transcriptional activator	0	NA	NA	16	0.80	9.7E-08
2011mTn5STM.3.07.D05	SMc01406	putative transcription regulator protein	15	0.80	1.1E-05	29	0.76	9.8E-11
2011mTn5STM.3.07.E05	SMa0890	hypothetical protein	16	0.14	7.8E-02	38	0.09	1.8E-02
2011mTn5STM.3.08.F05	SMb21527	putative taurin uptake ABC transporter ATP-binding protein	16	0.25	3.5E-06	40	0.37	2.1E-11
2011mTn5STM.3.10.G05	SMc01717	mercuric reductase protein MerA1	16	0.59	1.4E-07	24	0.57	6.7E-09
2011mTn5STM.3.07.H05	SMa2000	Putative ABC transporter periplasmic solute-binding protein	0	NA	NA	40	-0.33	9.2E-06
2011mTn5STM.3.10.A06	SMc03777	gamma-glutamyl phopsphate reductase protein ProA	0	NA	NA	12	-0.02	8.2E-01
2011mTn5STM.3.10.B06	SMc03899	hypothetical protein	0	NA	NA	0	NA	NA
2011mTn5STM.3.10.C06	SMc02070	hypothetical protein	16	-0.20	9.7E-03	30	-0.13	4.1E-01
2011mTn5STM.3.11.D06	SMc03032	flagellar P-ring precursor transmembrane protein	16	1.23	1.8E-14	40	1.08	3.6E-25
2011mTn5STM.3.08.E06	SMa0953	AttA1-like ABC transporter. ATP binding protein	0	NA	NA	8	0.28	3.1E-03
2011mTn5STM.3.11.F06	SMa2245	hypothetical protein	16	-0.02	4.1E-01	40	-0.06	1.8E-01
2011mTn5STM.3.10.G06	SMc02677	pyroline-5-carboxylate reductase protein ProC	16	0.21	2.9E-08	39	0.33	6.9E-05
2011mTn5STM.3.07.H06	SMa1052	hypothetical protein	16	0.02	8.6E-01	40	-0.15	1.9E-01
2011mTn5STM.3.09.A07	SMb21079	putative cAMP binding protein	16	-0.17	2.7E-04	40	-0.36	2.0E-04
2011mTn5STM.3.06.B07	SMc02487	dihydrolipoamide dehydrogenase LpdA2	0	NA	NA	10	-0.64	4.6E-05
2011mTn5STM.3.05.C07	SMb20678	hydroxypyruvate reductase protein TtuD2	15	-0.47	1.2E-09	37	-0.34	5.8E-06
2011mTn5STM.3.08.D07	SMc01174	cysteine synthase A protein CysK2	16	-0.31	7.7E-07	40	-0.54	2.0E-09
2011mTn5STM.3.05.E07	SMc01949	high-affinity branched-chain amino acid ABC transporter, ATP-binding ABC transporter protein LivG	16	0.14	5.0E-03	29	0.39	2.8E-09
2011mTn5STM.3.06.F07	SMc03977	hypothetical protein	16	-0.45	2.2E-07	40	-0.61	2.9E-10
2011mTn5STM.3.07.G07	SMc01419	RNA polymerase sigma factor protein SigE1	16	0.44	4.6E-13	24	0.46	3.6E-02
2011mTn5STM.3.08.H07	SMc04258	putative transport system permease ABC transporter protein	0	NA	NA	28	-0.01	7.3E-01
2011mTn5STM.3.06.A08	SMc01198	hypothetical protein	16	0.20	4.1E-05	32	0.07	4.1E-01
2011mTn5STM.3.08.B08	SMb20888	hypothetical protein	0	NA	NA	0	NA	NA
2011mTn5STM.3.10.C08	SMc02050	trigger factor protein Tig	16	-0.06	4.3E-01	24	0.23	1.8E-02
2011mTn5STM.3.05.D08	SMc01554	hypothetical protein	0	NA	NA	0	NA	NA
2011mTn5STM.3.10.E08	SMa2037	Putative oxidoreductase	16	-0.42	1.4E-11	39	-0.50	6.2E-06
2011mTn5STM.3.05.F08	SMc02857	heat shock protein 70 (HSP70) chaperone	0	NA	NA	0	NA	NA
2011mTn5STM.3.08.G08	SMa0583	NrtB. Nitrate transport permease protein	15	0.37	6.8E-05	32	0.79	3.6E-10
2011mTn5STM.3.09.H08	SMc01965	putative spermidine/putrescine transport ATP-binding ABC transporter protein	16	0.94	9.0E-13	39	0.84	6.9E-16

2011mTn5STM.3.11.A09	SMa0796	hypothetical protein	16	0.39	2.1E-13	40	0.40	4.3E-13
2011mTn5STM.3.10.B09	SMb20957	exopolysaccharide biosynthesis protein ExoA	16	-1.15	2.7E-14	40	-1.18	8.5E-33
2011mTn5STM.3.06.C09	SMb21596	hypothetical protein	16	-0.12	1.0E-02	8	-0.23	3.7E-03
2011mTn5STM.3.10.D09	SMb20582	putative transcriptional regulator, LysR family protein	16	1.06	5.1E-16	40	1.10	4.7E-26
2011mTn5STM.3.08.E09	SMb20933	putative two-component sensor histidine kinase protein	0	NA	NA	13	-0.34	1.9E-02
2011mTn5STM.3.12.F09	SMa1087	putative cation transport P-type ATPase	0	NA	NA	16	-0.54	4.2E-16
2011mTn5STM.3.05.G09	SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit	16	-0.50	2.1E-07	40	-0.35	5.6E-11
2011mTn5STM.3.07.H09	pSymB, 1322754	intergenic	16	0.13	1.6E-02	34	0.19	1.0E-03
2011mTn5STM.4.04.A10	pSymA, 248361	intergenic	16	-0.17	5.5E-05	40	-0.28	1.1E-07
2011mTn5STM.4.11.B10	SMb21300	putative deoxyribose-phosphate aldolase protein	16	0.06	2.0E-02	40	-0.29	9.0E-03
2011mTn5STM.4.11.C10	SMa0091	hypothetical protein	16	0.13	2.0E-02	39	-0.22	3.9E-03
2011mTn5STM.4.02.D10	SMa1362	putative inner-membrane permease	16	-0.47	4.6E-10	40	-0.46	1.2E-11
2011mTn5STM.4.06.E10	SMb20909	hypothetical protein	0	NA	NA	10	-0.46	3.1E-03
2011mTn5STM.4.12.F10	SMc02419	putative peptide transport system permease protein	16	-0.31	6.8E-08	40	-0.22	5.3E-03
2011mTn5STM.4.06.G10	SMa0769	FixP2 cytochrome c oxidase	16	-0.84	1.9E-12	32	-0.67	7.0E-09
2011mTn5STM.4.04.H10	SMc02273	RkpA fatty acid synthase protein	16	0.10	7.8E-02	16	-0.32	6.3E-09
2011mTn5STM.4.01.A11	SMa1397	hypothetical protein	16	0.09	1.5E-02	36	0.01	6.5E-01
2011mTn5STM.4.05.B11	SMa0583	NrtB. Nitrate transport permease protein	0	NA	NA	10	-0.10	1.5E-01
2011mTn5STM.4.01.C11	SMc02597	superoxide dismutase Cu-Zn precursor protein SodC	12	0.16	1.0E-01	24	0.37	1.9E-06
2011mTn5STM.4.09.D11	SMc03873	RNA polymerase sigma factor protein SigC	16	1.23	2.1E-10	24	1.15	4.6E-14
2011mTn5STM.4.04.E11	SMb20942	UDP glucose 4-epimerase protein ExoB	15	-1.14	3.9E-13	38	-1.07	3.8E-22
2011mTn5STM.4.09.F11	SMa0592	hypothetical protein	16	-0.37	7.2E-11	40	-0.24	2.3E-03
2011mTn5STM.4.01.G11	SMb20355	hypothetical protein	0	NA	NA	28	-0.52	9.3E-06
2011mTn5STM.4.08.H11	SMc00339	adenylate cyclase 1 protein CyaA	16	0.04	5.3E-02	40	0.22	2.7E-04
2011mTn5STM.4.04.A12	SMb21165	putative histidine ammonia-lyase histidase protein HutH1	16	-0.64	2.0E-07	39	-0.83	6.0E-17
2011mTn5STM.4.09.B12	SMc02144	phosphate transport system permease ABC transporter protein	13	-0.76	3.0E-09	32	-0.93	1.1E-11
2011mTn5STM.4.04.C12	SMc01053	siroheme synthase protein CysG	16	-1.06	9.2E-18	40	-1.06	2.3E-32
2011mTn5STM.4.07.D12	SMa0070	ABC transporter permease	16	0.75	6.4E-09	40	0.60	8.4E-13
2011mTn5STM.4.08.E12	SMb21158	putative transcriptional regulator, probably of sugar phosphate metabolism protein	16	-1.13	3.1E-16	16	-0.96	4.5E-17
2011mTn5STM.4.08.F12	SMb20018	probable sugar transport ATP-binding protein	14	0.34	2.2E-05	34	-0.16	6.6E-01
2011mTn5STM.4.03.G12	SMc00775	FbpB ABC iron transporter, permease component	8	-0.18	1.0E-01	24	-0.27	5.0E-02
2011mTn5STM.4.07.H12	SMa0398	probable HisD2 histidinol dehydrogenase	0	NA	NA	40	-0.20	4.2E-11
2011mTn5STM.4.01.A01	SMc00917	ATP phosphoribosyltransferase protein HisG	9	-0.21	5.9E-02	13	-0.23	2.1E-01
2011mTn5STM.4.10.B01	SMa1141	putative fnr/crp family transcriptional regulator	16	-0.01	8.2E-01	39	0.06	9.9E-02
2011mTn5STM.4.02.C01	SMc02893	probable transcriptional regulator	15	0.67	2.6E-07	40	0.18	9.2E-03
2011mTn5STM.4.03.D01	SMc01549	hypothetical protein	0	NA	NA	8	1.03	6.1E-06
2011mTn5STM.4.10.E01	SMa1077	Nex18 Symbiotically induced conserved protein	16	-0.24	1.1E-05	40	-0.04	9.0E-01
2011mTn5STM.4.07.F01	SMa0166	hypothetical protein	16	-0.14	1.5E-04	40	-0.26	6.4E-04
2011mTn5STM.4.06.G01	SMb21313	bifunctional glycosyltransferase. forming alpha-glycosyl and beta-glycosyl linkages protein ExpE2	15	-0.47	7.8E-08	40	-0.48	9.9E-12
2011mTn5STM.4.01.H01	SMc03776	glutamate-5-kinase protein ProB1	15	-0.27	7.4E-03	38	-0.66	1.1E-05
2011mTn5STM.4.04.B02	SMa1131	hypothetical protein	16	0.18	8.1E-07	40	-0.05	7.3E-01

2011mTn5STM.4.01.C02	SMa0229	hypothetical protein	15	-0.48	1.1E-05	38	-0.50	8.3E-06
2011mTn5STM.4.04.D02	SMb21133	putative sulfate uptake ABC transporter periplasmic solute-binding protein	16	-0.65	1.7E-11	38	-0.55	1.1E-10
2011mTn5STM.4.08.E02	SMc00080	hypothetical protein	16	-0.19	3.1E-04	40	0.02	8.7E-01
2011mTn5STM.4.07.F02	SMa0695	ArcB catabolic ornithine carbamoyl transferase	16	-0.30	6.4E-04	40	-0.32	6.3E-10
2011mTn5STM.4.02.G02	SMb20868	putative two-component sensor histidine kinase protein	16	-0.49	1.6E-10	40	-0.54	3.6E-08
2011mTn5STM.4.08.H02	pSymB. 1408087	intergenic	16	-0.15	7.7E-04	40	0.02	9.6E-01
2011mTn5STM.4.08.A03	SMb21536	hypothetical protein	16	-0.50	5.2E-08	37	-0.47	3.7E-10
2011mTn5STM.4.07.B03	S. meliloti chromosome, 2005840	intergenic	16	0.14	4.5E-03	40	0.21	1.9E-06
2011mTn5STM.4.09.C03	SMc00606	putative ribonuclease protein	15	0.02	3.4E-01	38	-0.11	2.9E-01
2011mTn5STM.4.01.D03	SMc00886	hypothetical protein	15	-0.24	6.9E-03	20	-0.20	7.7E-02
2011mTn5STM.4.09.E03	SMc02117	cystathionine beta-lyase protein MetC	14	-0.25	1.9E-02	36	-0.03	1.5E-01
2011mTn5STM.4.08.F03	SMa1913	putative transport protein	15	-0.76	1.1E-09	40	-0.86	7.7E-18
2011mTn5STM.4.02.G03	SMc00828	hypothetical protein	14	-0.87	5.8E-09	38	-0.86	1.2E-15
2011mTn5STM.4.01.H03	SMb20151	hypothetical protein	16	0.10	7.5E-04	40	0.24	1.0E-07
2011mTn5STM.4.08.A04	SMc02139	hypothetical protein	16	-0.51	1.4E-09	23	-0.43	9.4E-09
2011mTn5STM.4.05.B04	SMa1371	probable ABC transporter ATP-binding protein	16	-0.25	2.1E-05	32	0.05	5.4E-01
2011mTn5STM.4.12.C04	SMa1465	putative ABC transporter permease	16	-0.14	3.2E-06	40	0.02	5.9E-01
2011mTn5STM.4.04.D04	SMb20481	AsnO asparagine synthase. glutamine-hydrolyzing	16	-1.05	2.1E-18	38	-1.10	4.8E-30
2011mTn5STM.4.02.E04	pSymA, 1165092	intergenic	0	NA	NA	32	-0.02	8.8E-02
2011mTn5STM.4.09.F04	SMa2115	Gst13 glutathione S-transferase	0	NA	NA	0	NA	NA
2011mTn5STM.4.03.G04	SMa0752	FAD/NAD(P)-binding oxidoreductase protein	16	-0.50	2.2E-09	40	-0.43	1.5E-10
2011mTn5STM.4.10.H04	SMc00428	hypothetical protein	0	NA	NA	24	-0.17	7.3E-02
2011mTn5STM.4.01.A05	pSymB. 376527	intergenic	16	0.70	1.7E-12	35	0.62	1.3E-08
2011mTn5STM.4.06.B05	SMc00127	choline sulfatase protein BetC	16	-0.06	1.7E-01	32	-0.22	9.8E-02
2011mTn5STM.4.03.C05	SMc00937	TRK system potassium uptake transmembrane protein TrkH	16	-0.47	2.4E-10	40	-0.33	1.5E-04
2011mTn5STM.4.05.D05	SMa0222	putative GntR-family transcriptional regulator	16	0.49	1.4E-10	40	0.05	7.9E-01
2011mTn5STM.4.11.E05	SMb20534	hypothetical protein	0	NA	NA	16	-0.16	1.8E-04
2011mTn5STM.4.11.F05	SMc02149	hypothetical protein	0	NA	NA	0	NA	NA
2011mTn5STM.4.05.G05	S. meliloti chromosome, 532543	intergenic	16	-0.22	3.3E-05	32	-0.35	8.0E-13
2011mTn5STM.4.02.H05	SMc00523	hypothetical protein	16	0.41	9.6E-09	40	-0.06	6.1E-01
2011mTn5STM.4.09.A06	SMb20149	hypothetical protein	16	0.17	4.6E-05	40	0.20	2.5E-03
2011mTn5STM.4.03.B06	SMc00781	methylmalonate-semialdehyde dehydrogenase MmsA (IolA)	16	-1.09	8.9E-16	23	-1.03	2.0E-16
2011mTn5STM.4.03.C06	SMc01267	hypothetical protein	16	-0.49	6.9E-12	24	-0.44	1.6E-06
2011mTn5STM.4.04.D06	SMb20227	nutrient deprivation-induced protein A (NdiA1)	16	-0.77	7.0E-15	40	-0.75	1.7E-25
2011mTn5STM.4.02.E06	SMc02226	hypothetical protein	15	-1.05	2.4E-16	40	-0.97	8.4E-28
2011mTn5STM.4.05.F06	SMb20820	putative ferredoxin reductase MocF	16	-0.40	1.3E-11	24	0.06	5.6E-01
2011mTn5STM.4.07.G06	SMc00955	hypothetical protein	16	0.49	9.6E-06	39	0.38	1.4E-07
2011mTn5STM.4.05.H06	SMa1480	probable LysR-type activator	16	-0.28	1.9E-09	40	-0.35	5.0E-11
2011mTn5STM.4.11.A07	SMa1734	hypothetical protein	16	-0.50	7.0E-11	40	-0.42	2.0E-07
2011mTn5STM.4.03.B07	SMa0417	hypothetical protein	16	-0.41	1.0E-08	40	-0.26	2.4E-03

2011mTn5STM.4.07.C07	SMb20941	msbA-like saccharide exporting ABC transporter protein	16	-0.52	1.2E-10	39	-0.67	3.9E-12
2011mTn5STM.4.07.D07	SMb21592	putative sugar uptake ABC transporter ATP-binding protein	0	NA	NA	0	NA	NA
2011mTn5STM.4.09.E07	SMc00893	hypothetical protein	16	0.51	5.9E-08	40	0.36	2.8E-07
2011mTn5STM.4.03.F07	S. meliloti chromosome, 2051945	intergenic	16	0.90	1.2E-11	24	0.86	4.4E-10
2011mTn5STM.4.09.G07	SMa2279	hypothetical protein	16	-0.25	3.9E-02	24	0.09	2.0E-01
2011mTn5STM.4.04.H07	S. meliloti chromosome, 81040	intergenic	16	0.62	1.4E-13	24	0.71	3.6E-10
2011mTn5STM.4.08.A08	SMc02372	putative transport transmembrane protein	14	0.40	2.7E-04	36	0.48	2.6E-07
2011mTn5STM.4.12.B08	SMa1355	hypothetical protein	16	-0.05	1.9E-01	40	-0.17	5.6E-02
2011mTn5STM.4.06.C08	SMa0021	hypothetical protein	14	-0.16	1.2E-02	35	-0.27	7.5E-05
2011mTn5STM.4.04.D08	SMa1782	putative LysR-family transcriptional regulator	16	0.47	1.3E-16	40	0.66	4.0E-07
2011mTn5STM.4.01.E08	SMb20925	hypothetical protein	0	NA	NA	27	-0.33	7.7E-03
2011mTn5STM.4.01.F08	SMc01611	ferrichrome-iron receptor precursor protein	16	0.14	6.5E-04	40	0.31	4.7E-05
2011mTn5STM.4.10.G08	S. meliloti chromosome, 2816977	intergenic	16	1.10	1.9E-14	40	0.67	2.3E-08
2011mTn5STM.4.08.H08	SMb20007	KatC catalase protein	16	0.00	6.2E-01	35	-0.30	1.7E-03
2011mTn5STM.4.03.A09	SMc00825	glutamate-cysteine ligase GshA	0	NA	NA	0	NA	NA
2011mTn5STM.4.05.B09	SMa0574	hypothetical protein	16	-0.83	2.4E-12	40	-0.68	1.9E-19
2011mTn5STM.4.11.C09	S. meliloti chromosome, 3123223	intergenic	14	0.15	1.2E-01	31	0.27	3.8E-07
2011mTn5STM.4.08.D09	SMa1084	hypothetical protein	16	-0.42	1.3E-09	40	-0.69	1.8E-12
2011mTn5STM.4.08.E09	SMa1163	putative cation transport P-type ATPase	16	-0.42	2.4E-07	39	-0.35	1.9E-08
2011mTn5STM.4.01.F09	SMa0734	hypothetical protein	14	-0.40	1.2E-04	24	-0.18	5.9E-04
2011mTn5STM.4.07.G09	SMa1582	hypothetical protein	0	NA	NA	40	-0.14	1.5E-04
2011mTn5STM.4.02.H09	SMb20049	putative elongation factor G protein	16	-0.10	5.8E-02	38	-0.14	2.4E-01

[a] For mutants carrying a transposon insertion in an intergenic region the exact position of insertion in the respective replicon is given.

TABLE G.4. RESULTS OF *IN PLANTA* STM EXPERIMENT FOR SET 2

mutantID complete	aliases	function	Slide 4, combined DNA			Slides 1-3		
			n	m-value	p-value(t)	n	m-value	p-value(t)
2011mTn5STM.1.11.A10	SMb20030	putative ECF-sigma factor protein	8	0.25	2.7E-03	32	-0.02	8.8E-01
2011mTn5STM.1.05.B10	SMc03248	putative transposase protein	8	0.22	7.3E-03	32	0.24	4.5E-12
2011mTn5STM.1.02.C10	SMb20097	hypothetical oxidoreductase protein	8	-0.09	2.6E-01	24	-0.68	2.3E-15
2011mTn5STM.1.01.D10	SMc02642	hypothetical protein	8	0.79	2.6E-07	24	0.40	1.1E-08
2011mTn5STM.1.03.E10	SMb21630	hypothetical protein	8	1.58	1.9E-09	24	1.70	2.2E-28
2011mTn5STM.1.06.F10	SMb20443	putative TRAP dicarboxylate transporter	8	-0.41	2.8E-05	32	-0.40	1.7E-14
2011mTn5STM.1.02.G10	SMc00490	transcription regulator GntR family	8	0.54	9.4E-09	32	0.61	9.6E-13
2011mTn5STM.1.06.H10	SMc00708	putative hydroxyacylglutathione hydrolase protein Glx II	8	-0.12	8.6E-02	32	-0.21	2.2E-05
2011mTn5STM.1.12.A11	SMa2089	hypothetical protein	8	-0.37	4.3E-04	32	-0.74	4.6E-09
2011mTn5STM.1.08.B11	SMb20935	putative transcriptional regulator protein ExsI	0	NA	NA	24	0.26	5.1E-05
2011mTn5STM.1.11.C11	SMc02269	capsule polysaccharide biosynthesis protein RkpJ	8	0.30	4.9E-04	32	-0.40	5.1E-03
2011mTn5STM.1.03.D11	SMc04346	ketol-acid reductoisomerase protein IlvC	0	NA	NA	32	-0.88	1.7E-26
2011mTn5STM.1.12.E11	SMc02613	glutamine synthetase III protein GlnT	8	0.70	3.3E-06	32	0.53	7.7E-08
2011mTn5STM.1.09.F11	SMc00033	diguanylate cyclase/phosphodiesterase protein	8	0.28	1.8E-03	32	0.36	1.2E-05
2011mTn5STM.1.11.G11	SMa1252	hypothetical protein	8	0.11	4.0E-02	32	0.07	7.1E-01
2011mTn5STM.1.02.H11	SMc03125	putative ABC-transporter ATP-binding protein	8	-0.07	3.4E-02	32	-0.28	4.3E-03
2011mTn5STM.1.07.A12	SMc00594	beta-etherase protein LigE	8	0.32	1.9E-06	32	0.22	2.0E-04
2011mTn5STM.1.13.B12	SMb21114	nitrate transport protein NrtA	8	0.15	4.4E-03	32	0.20	9.1E-05
2011mTn5STM.1.07.C12	SMa1646	putative ABC-transporter ATP-binding protein	8	-0.13	8.1E-02	16	0.10	8.8E-04
2011mTn5STM.1.07.D12	SMa1660	putative acetyltransferase	8	-0.30	3.4E-05	32	-0.21	2.0E-06
2011mTn5STM.1.13.E12	SMb20153	hypothetical protein	8	0.12	4.8E-03	32	0.18	6.4E-03
2011mTn5STM.1.06.F12	SMa1213	FixP1 Di-heme cytochrome c protein	8	0.48	1.9E-06	32	0.46	4.6E-13
2011mTn5STM.1.12.G12	SMc01909	hypothetical protein	8	0.69	1.9E-08	32	0.60	9.1E-11
2011mTn5STM.1.09.H12	SMc02767	N-5'-phosphoribosylanthranilate isomerase TrpF	8	-0.95	2.5E-07	32	-0.94	2.5E-27
2011mTn5STM.1.11.A01	SMc00169	NAD-malic enzyme oxidoreductase protein Dme	8	-0.94	7.0E-07	32	-0.99	5.0E-29
2011mTn5STM.1.12.B01	SMa1256	hypothetical protein	0	NA	NA	32	-0.33	5.6E-06
2011mTn5STM.1.07.C01	SMb20007	catalase protein CatC	0	NA	NA	24	0.23	3.2E-04
2011mTn5STM.1.06.D01	SMa1084	hypothetical protein	8	-0.50	2.2E-07	8	-0.24	3.0E-04
2011mTn5STM.1.02.E01	SMc00898	probable glutathione-regulated potassium-efflux system transmembrane protein KefB	8	0.91	2.7E-07	24	1.23	9.9E-11
2011mTn5STM.1.02.F01	SMc01443	putative glutathione S-transferase protein	8	0.62	7.6E-05	32	0.50	1.6E-13
2011mTn5STM.1.03.G01	SMb20346	putative efflux transmembrane protein	8	-0.37	3.3E-05	32	-0.52	4.0E-07
2011mTn5STM.1.10.H01	SMb20551	hypothetical protein	8	-0.14	2.7E-03	32	-0.22	1.8E-04
2011mTn5STM.1.11.A02	SMc01165	putative sugar kinase protein IolC	0	NA	NA	10	0.82	6.6E-07
2011mTn5STM.1.10.B02	S. meliloti chromosome, 3368649	intergenic	8	-0.36	3.0E-06	32	-0.26	9.5E-08
2011mTn5STM.1.11.C02	SMc02851	putative transcriptional regulator	0	NA	NA	32	0.79	1.2E-19
2011mTn5STM.1.11.D02	SMa1272	NorQ protein required for nitric oxide reductase activity	8	0.53	8.0E-05	32	0.48	2.5E-20

2011mTn5STM.1.01.E02	SMc04023	probable UTP-glucose-1-phosphate uridylyltransferase protein ExoN2	8	-0.42	2.0E-06	16	-0.55	8.4E-08
2011mTn5STM.1.06.F02	SMb20216	putative epoxide hydrolase protein	8	-0.09	1.5E-02	16	0.02	1.2E-01
2011mTn5STM.1.05.H02	SMb21236	hypothetical protein	8	0.24	5.8E-03	32	0.11	7.9E-03
2011mTn5STM.1.06.A03	SMc00790	probable dipeptide transport ATP-binding ABC transporter protein DppF	8	0.78	5.5E-04	32	0.71	1.5E-08
2011mTn5STM.1.02.B03	SMb20893	probable sugar uptake ABC transporter permease protein	8	-0.15	1.8E-02	32	-0.23	4.8E-03
2011mTn5STM.1.12.C03	SMc00830	hypothetical transmembrane protein	8	0.15	6.4E-04	32	0.42	8.6E-05
2011mTn5STM.1.04.D03	pSymB, 551929	intergenic	8	-0.35	2.3E-04	32	-0.30	1.5E-07
2011mTn5STM.1.13.E03	SMc04051	putative RNA polymerase sigma-E factor protein	8	1.06	4.8E-05	8	-0.38	2.8E-04
2011mTn5STM.1.09.F03	SMa2251	hypothetical protein	8	0.36	3.8E-06	32	0.19	1.2E-04
2011mTn5STM.1.05.G03	SMa0689	hypothetical protein	8	0.53	4.4E-06	32	0.37	9.2E-05
2011mTn5STM.1.11.H03	SMb20274	hypothetical protein	8	0.32	7.1E-05	16	0.20	5.7E-05
2011mTn5STM.1.04.A04	SMa0894	hypothetical protein	0	NA	NA	16	0.09	1.1E-01
2011mTn5STM.1.05.B04	pSymA, 1142848	intergenic	8	-0.26	4.8E-03	32	-0.21	7.9E-05
2011mTn5STM.1.06.C04	SMa1787	hypothetical protein	8	-0.13	1.6E-02	32	-0.08	2.9E-01
2011mTn5STM.1.06.D04	SMa2119	hypothetical protein	8	-0.35	7.3E-05	32	-0.68	2.0E-10
2011mTn5STM.1.06.E04	S. meliloti chromosome, 3366182	intergenic	8	0.06	9.2E-01	24	-0.81	9.3E-11
2011mTn5STM.1.07.F04	SMa1434	probable ABC transporter ATP-binding protein	8	-0.45	4.5E-08	32	-0.43	4.1E-11
2011mTn5STM.1.08.G04	SMc00913	GroEL1 60 kD chaperonin A protein	8	0.12	1.2E-01	32	0.22	1.5E-06
2011mTn5STM.1.10.H04	SMa0873	NifN Nitrogenase Fe-Mo cofactor biosynthesis protein	0	NA	NA	32	-0.97	1.4E-26
2011mTn5STM.1.05.A05	SMb20969	putative sugar ABC transporter permease	8	0.17	2.3E-04	24	-0.01	4.4E-01
2011mTn5STM.1.08.B05	SMc00289	putative cold shock transcription regulator protein CspA5	8	-0.32	4.5E-06	32	-0.51	1.8E-06
2011mTn5STM.1.06.C05	SMc01143	probable heat-inducible transcription repressor protein HrcA	8	0.71	5.6E-08	32	0.69	3.5E-23
2011mTn5STM.1.09.D05	SMb21176	PhoD phosphate uptake ABC transporter periplasmic solute-binding protein precursor	0	NA	NA	24	-0.21	1.8E-01
2011mTn5STM.1.06.E05	SMb20179	hypothetical protein	8	0.42	3.3E-06	32	0.34	1.9E-09
2011mTn5STM.1.10.F05	SMb21185	putative succinate-semialdehyde dehydrogenase (NAD(P)+) protein GabD	0	NA	NA	16	-0.21	1.9E-03
2011mTn5STM.1.05.G05	SMb20024	ribosomal large subunit pseudouridine synthase B	8	0.39	3.7E-06	32	0.33	6.2E-05
2011mTn5STM.1.11.H05	SMa1696	hypothetical protein	0	NA	NA	8	0.00	6.6E-01
2011mTn5STM.1.10.A06	SMb20552	putative peptidase	8	-0.46	9.5E-05	32	-0.47	2.2E-16
2011mTn5STM.1.03.B06	SMb21202	probable dioxygenase	8	-0.28	4.1E-04	32	-0.41	3.9E-05
2011mTn5STM.1.13.C06	SMc02143	putative phosphate transport system permease ABC transporter protein PstA	8	-0.95	5.0E-08	32	-0.95	4.9E-27
2011mTn5STM.1.02.D06	SMc00270	putative transketolase alpha subunit protein	0	NA	NA	0	NA	NA
2011mTn5STM.1.04.E06	SMa0909	hypothetical protein	8	-0.04	3.7E-01	32	-0.01	7.5E-01
2011mTn5STM.1.06.F06	SMb21281	putative integral membrane transporter protein. xanthineuracil permeases family	0	NA	NA	32	-0.14	1.6E-04
2011mTn5STM.1.09.G06	SMc02780	probable succinate-semialdehyde dehydrogenase (NADP+) protein GabD	0	NA	NA	8	0.34	2.3E-02
2011mTn5STM.1.10.H06	SMc02361	cytochrome C-type biogenesis transmembrane protein CycH	8	-0.85	2.7E-09	32	-0.90	8.9E-17
2011mTn5STM.1.10.A07	SMa1643	hypothetical protein	8	-0.25	1.4E-05	32	-0.17	3.7E-02
2011mTn5STM.1.07.B07	SMc00616	putative permease protein PerM	8	0.15	2.3E-03	32	0.14	2.8E-01
2011mTn5STM.1.06.C07	SMc00416	probable tetrapyrrole methylase	8	0.32	6.4E-04	24	0.09	1.5E-01
2011mTn5STM.1.02.D07	SMc01373	putative DNA-damage-inducible protein P (DinP)	0	NA	NA	24	0.58	2.7E-07
2011mTn5STM.1.09.E07	SMc00729	putative electron transfer flavoprotein beta-subunit EtfB	8	0.59	6.2E-10	32	0.64	5.4E-23

2011mTn5STM.1.08.F07	SMa1442	putative GntR-family transcription factor	8	-0.56	2.3E-06	32	-0.47	5.7E-19
2011mTn5STM.1.02.G07	SMc01877	probable DNA repair protein RecN	8	-0.77	6.2E-07	32	-0.73	4.0E-14
2011mTn5STM.1.10.H07	SMb21070	putative tyrosine-protein kinase ExoP2	8	-0.27	2.5E-04	32	-0.15	3.2E-02
2011mTn5STM.1.09.A08	SMb21076	putative glycosyltransferase	8	0.60	2.3E-07	32	0.74	2.8E-16
2011mTn5STM.1.09.B08	SMa1153	hypothetical protein	8	0.15	1.7E-02	32	0.14	7.5E-06
2011mTn5STM.1.13.C08	SMa1229	FixL oxygen regulated histidine kinase	8	-1.08	1.0E-07	24	-0.96	8.2E-19
2011mTn5STM.1.06.D08	SMa1840	hypothetical protein	8	-0.24	7.5E-04	32	-0.19	7.0E-04
2011mTn5STM.1.04.E08	SMc03885	putative histidinol-phosphate aminotransferase protein HisC	8	-0.36	2.2E-04	32	-0.19	1.5E-02
2011mTn5STM.1.10.F08	SMa0626	putative membrane protein	8	-0.50	6.0E-06	32	-0.31	2.3E-03
2011mTn5STM.1.13.G08	SMb20416	probable ABC transporter periplasmic glycerol-3-phosphate-binding protein precursor	8	0.70	1.5E-06	16	-0.14	5.0E-01
2011mTn5STM.1.11.H08	SMb20457	putative transcriptional regulator	8	-0.30	1.9E-04	32	-0.27	1.5E-05
2011mTn5STM.1.08.A09	SMa1301	putative transmembrane transport protein	8	-0.57	3.6E-06	32	-0.57	1.9E-07
2011mTn5STM.1.07.B09	SMb20962	probable phoshomethylpyrimidine kinase protein ThiD	8	-1.13	3.6E-08	32	-1.05	2.0E-31
2011mTn5STM.1.02.C09	pSymA, 597300	intergenic	8	0.02	8.6E-01	32	-0.15	1.0E-01
2011mTn5STM.1.10.D09	pSymB, 1120637	intergenic	8	-0.20	6.1E-04	32	-0.08	8.8E-02
2011mTn5STM.1.13.E09	SMb20110	putative oligopeptide ABC transporter permease protein	8	-0.26	2.8E-04	32	-0.44	2.1E-06
2011mTn5STM.1.04.F09	SMb21196	putative oligopeptide uptake ABC transporter periplasmic solute-binding protein precursor OppA	8	-0.05	2.5E-01	32	-0.13	2.0E-02
2011mTn5STM.1.12.G09	SMc02285	putative adenylate cyclase transmembrane protein CyaE	8	0.24	7.2E-02	32	0.23	4.6E-03
2011mTn5STM.1.08.H09	SMc01184	hypothetical transmembrane protein	8	-0.54	1.3E-04	16	0.13	3.6E-02
2011mTn5STM.2.12.A10	SMc00129	sensor histidine kinase protein FeuQ	8	-0.99	1.6E-07	16	-0.96	5.5E-12
2011mTn5STM.2.10.B10	SMa2277	hypothetical protein	8	0.65	5.1E-06	32	0.65	7.7E-16
2011mTn5STM.2.06.C10	SMb21319	putative membrane protein ExpA1	8	0.30	1.4E-06	32	-0.05	7.4E-01
2011mTn5STM.2.12.D10	SMc02705	hypothetical transmembrane protein	8	-0.79	6.9E-07	32	-0.61	7.3E-20
2011mTn5STM.2.06.E10	SMb20900	putative diguanylate cyclase/phosphodiesterase protein	8	0.20	6.7E-04	32	0.09	3.4E-01
2011mTn5STM.2.10.F10	SMa0280	conserved hypothetical protein	8	-0.33	4.4E-05	16	-0.49	5.7E-05
2011mTn5STM.2.09.G10	SMa0298	putative ABC transporter ATP-binding protein	8	-0.37	2.4E-06	32	-0.23	9.2E-07
2011mTn5STM.2.07.H10	SMc03133	putative amino-acid transport system permease ABC transporter protein	8	0.43	6.5E-06	16	0.63	2.2E-09
2011mTn5STM.2.09.A11	SMb20317	putative sugar ABC transporter ATP-binding protein	8	0.64	5.0E-08	32	0.79	1.1E-24
2011mTn5STM.2.09.B11	SMc00296	poly-beta-hydroxybutyrate polymerase PhbC	8	-0.22	3.3E-03	32	-0.05	9.1E-01
2011mTn5STM.2.10.C11	SMc01957	hypothetical protein	8	0.55	1.5E-06	32	0.65	1.2E-22
2011mTn5STM.2.05.D11	SMb20615	putative thiamine biosynthesis protein ThiC	8	-1.06	3.9E-09	8	-0.92	6.2E-08
2011mTn5STM.2.09.E11	SMa0453	putative plasmid stability protein y4JK	8	0.27	3.3E-03	32	0.25	2.0E-09
2011mTn5STM.2.07.F11	SMc00012	putative cytochrome C oxidase assembly transmembrane protein CtaG	8	0.07	8.2E-01	32	0.26	1.4E-06
2011mTn5STM.2.06.G11	SMa1819	hypothetical protein	8	-0.59	1.2E-06	32	-0.37	1.1E-09
2011mTn5STM.2.04.H11	SMa0160	transcriptional regulator, GntR family	8	-0.25	6.7E-05	32	-0.28	4.1E-06
2011mTn5STM.2.04.A12	SMc04204	putative iron transport regulator transmembrane protein FecR	8	0.38	9.2E-07	32	0.30	1.6E-04
2011mTn5STM.2.06.B12	SMb20232	putative suagr ABC transporter permease protein	8	0.36	7.7E-05	8	0.59	8.3E-06
2011mTn5STM.2.10.C12	SMb20772	probable pyridoxal phosphate biosynthetic protein PdxA	8	0.50	1.6E-04	32	0.52	9.9E-04
2011mTn5STM.2.12.D12	SMa0013	hypothetical protein	8	-0.31	2.9E-04	8	-0.72	1.0E-06
2011mTn5STM.2.04.E12	SMa0794	hypothetical protein	8	-0.16	2.6E-02	32	0.14	4.9E-02

2011mTn5STM.2.05.F12	SMc03114	probable transposase for insertion sequence element ISRM3-like	8	0.12	3.1E-03	24	0.05	6.3E-02
2011mTn5STM.2.12.G12	SMb21183	probable chaperonine protein HtpG	8	-0.47	1.7E-07	32	-0.52	6.0E-07
2011mTn5STM.2.07.H12	SMb20498	probable aldolase	0	NA	NA	32	-0.12	3.1E-02
2011mTn5STM.2.07.A01	SMa1292	hypothetical protein	8	0.14	4.0E-01	31	-0.10	9.6E-02
2011mTn5STM.2.01.B01	SMa1667	probable ArcD1 arginine/ornithine antiporter	8	-0.52	1.6E-04	32	-0.58	1.7E-11
2011mTn5STM.2.03.C01	SMc00361	hypothetical protein	8	-0.06	4.4E-02	32	-0.26	5.7E-05
2011mTn5STM.2.02.D01	SMc02268	KpsF3 involved in KPS assembly. putative arabinose 5-phosphate isomerase	8	-0.33	1.3E-06	32	-0.57	4.1E-14
2011mTn5STM.2.01.E01	SMb21000	putative transport protein	8	0.03	6.2E-01	32	0.10	1.1E-02
2011mTn5STM.2.09.F01	pSymA, 280377	intergenic	8	0.00	3.3E-01	32	-0.19	4.8E-02
2011mTn5STM.2.07.G01	SMc00372	putative oxidoreductase protein	0	NA	NA	32	0.34	6.0E-06
2011mTn5STM.2.02.H01	SMa0224	putative permease of the major facilitator superfamily	8	0.28	2.1E-02	32	0.23	2.1E-05
2011mTn5STM.2.09.A02	SMc02109	probable ATP-dependent CLP protease ATP-binding subunit protein	8	-0.96	4.0E-07	32	-1.03	1.9E-26
2011mTn5STM.2.02.B02	SMa1966	putative LysR-family transcriptional regulator	8	-0.02	2.1E-01	32	-0.32	2.9E-03
2011mTn5STM.2.01.C02	SMa0493	putative ABC transporter permease protein	8	0.10	2.9E-01	32	0.05	8.5E-01
2011mTn5STM.2.06.D02	SMb21351	putative C4-dicarboxylate large membrane transport protein DctM	8	-0.43	5.8E-05	32	-0.58	4.7E-11
2011mTn5STM.2.07.E02	SMc01166	putative malonic semialdehyde oxidative decarboxylase protein IolD	8	-0.36	9.7E-04	32	-0.48	1.1E-05
2011mTn5STM.2.03.F02	SMa0352	hypothetical protein	0	NA	NA	20	-0.14	5.5E-04
2011mTn5STM.2.05.G02	SMc02768	hypothetical protein	8	-0.03	5.7E-01	32	0.05	8.7E-01
2011mTn5STM.2.10.H02	SMa0664	hypothetical protein	8	0.01	7.0E-01	32	-0.35	8.9E-05
2011mTn5STM.2.01.A03	SMb20881	hypothetical exported protein	8	0.33	6.4E-06	31	0.54	2.5E-09
2011mTn5STM.2.08.B03	SMb21262	putative mureinpeptideoligopeptide ABC transporter permease protein	8	-0.40	2.3E-04	32	-0.46	6.0E-06
2011mTn5STM.2.07.C03	SMb21446	probable glycogen operon protein GlgX	8	-0.29	2.4E-05	32	-0.27	3.3E-08
2011mTn5STM.2.08.D03	SMb20261	putative malate dehydrogenase protein	8	0.29	8.6E-02	32	0.26	4.8E-05
2011mTn5STM.2.09.E03	SMc01182	hypothetical protein	8	0.26	1.7E-02	24	0.41	6.8E-10
2011mTn5STM.2.05.F03	SMb20775	hypothetical exported protein. TonB-dependent receptor protein	8	0.45	2.3E-03	32	0.50	2.9E-12
2011mTn5STM.2.04.G03	SMa2137	probable glycerate	8	-0.05	1.8E-01	32	-0.05	3.4E-01
2011mTn5STM.2.04.H03	S. meliloti chromosome, 894903	intergenic	8	0.38	9.0E-04	32	0.27	2.1E-04
2011mTn5STM.2.01.A04	SMa1250	putative NirK Cu-nitrite reductase	8	0.00	9.6E-01	24	-0.36	6.0E-09
2011mTn5STM.2.02.B04	SMc02599	probable L-aspartate oxidase protein NadB	0	NA	NA	24	0.24	7.8E-05
2011mTn5STM.2.09.C04	SMc03184	pH adaptation potassium efflux system G transmembrane protein PhaG1	0	NA	NA	0	NA	NA
2011mTn5STM.2.01.D04	SMc02165	probable orotate phosphoribosyltransferase protein PyrE	8	-0.96	1.7E-09	32	-0.84	4.9E-24
2011mTn5STM.2.09.E04	SMa1882	putative transcriptional activator	8	-0.28	4.1E-06	32	-0.29	8.6E-08
2011mTn5STM.2.07.F04	SMc00196	putative MFS permease	8	0.37	1.0E-06	32	0.38	4.6E-04
2011mTn5STM.2.09.G04	SMb20879	putative protein required for attachment to host cells	8	0.54	6.2E-06	32	0.62	2.8E-22
2011mTn5STM.2.05.H04	SMc00380	probable PecS HTH-type transcriptional regulator. MarR family	8	0.41	3.9E-05	32	0.59	1.3E-10
2011mTn5STM.2.09.A05	SMa1233	NapB periplasmic nitrate reductase	8	-0.67	6.6E-06	8	-0.90	4.6E-08
2011mTn5STM.2.03.B05	SMb20092	hypothetical protein	8	0.34	4.8E-07	32	0.12	7.1E-02
2011mTn5STM.2.06.C05	SMc01281	hypothetical protein	8	0.18	2.2E-03	32	0.35	5.5E-05
2011mTn5STM.2.10.D05	SMa0302	putative ABC transporter. periplasmic solute-binding protein	8	0.24	5.5E-03	32	0.09	1.3E-01
2011mTn5STM.2.04.E05	SMc02176	putative adenylate cyclase transmembrane protein CyaD1	0	NA	NA	16	1.03	9.8E-07

2011mTn5STM.2.09.F05	SMc01496	probable sorbitol-binding periplasmic protein SmoE	8	0.34	1.0E-04	24	-0.65	1.7E-17
2011mTn5STM.2.02.G05	SMb20955	endo-beta-1.3-1.4-glycanase protein ExoK	8	-0.47	1.1E-04	32	-0.33	1.4E-10
2011mTn5STM.2.04.H05	SMa0257	NADH:flavin oxidoreductase	8	-0.49	8.3E-06	32	-0.47	2.1E-06
2011mTn5STM.2.10.A06	SMc03181	pH adaptation potassium efflux system transmembrane protein PhaD1	8	-1.03	1.4E-10	32	-0.92	6.4E-27
2011mTn5STM.2.07.B06	SMa0224	putative permease of the major facilitator superfamily	0	NA	NA	32	0.04	5.4E-02
2011mTn5STM.2.04.C06	SMa0059	dehydrogenase related to short chain alcohol dehydrogenases	8	-0.18	8.1E-04	32	-0.25	2.4E-06
2011mTn5STM.2.10.D06	SMa0557	probable LysR family transcriptional regulator	8	-0.21	1.3E-04	32	-0.29	5.1E-04
2011mTn5STM.2.07.E06	SMa0829	NifK nitrogenase Fe-Mo beta chain	8	-1.10	7.6E-11	32	-0.96	3.6E-31
2011mTn5STM.2.04.F06	SMa2404	rhizobactin siderophore biosynthesis protein	8	0.05	8.5E-02	32	-0.02	7.3E-01
2011mTn5STM.2.10.G06	SMc01228	streptomycin resistance protein StrB	8	0.73	2.4E-08	32	0.55	1.8E-16
2011mTn5STM.2.12.H06	SMa0101	amidohydrolase	8	0.66	5.6E-08	32	0.72	9.5E-09
2011mTn5STM.2.09.A07	pSymB, 1668997	intergenic	8	0.16	5.4E-04	24	0.10	5.1E-01
2011mTn5STM.2.05.B07	SMb20213	hypothetical protein	8	0.39	2.4E-06	32	0.36	1.6E-07
2011mTn5STM.2.05.C07	SMc00236	probable indole-3-glycerol phosphate synthase TrpC	8	-0.99	1.7E-08	32	-0.82	5.9E-26
2011mTn5STM.2.06.D07	SMa1223	flavoprotein oxidoreductase FixM	8	-0.60	1.0E-07	24	-0.37	2.0E-03
2011mTn5STM.2.02.E07	SMc04405	3-isopropylmalate dehydrogenase LeuB	8	-1.12	9.2E-09	32	-0.94	5.5E-31
2011mTn5STM.2.07.F07	SMb20680	putative hydroxypyruvate isomerase protein	8	0.31	4.4E-06	24	0.18	7.5E-04
2011mTn5STM.2.08.G07	SMb21263	putative mureinpeptideoligopeptide ABC transporter permease protein	8	-0.12	1.3E-02	16	-0.71	6.1E-10
2011mTn5STM.2.03.H07	SMc01547	hypothetical protein	8	0.21	1.4E-02	24	-0.18	1.7E-02
2011mTn5STM.2.08.A08	SMa0179	putative transcriptional activator	8	0.02	1.1E-01	32	0.07	5.8E-01
2011mTn5STM.2.07.B08	SMa0506	putative ABC transporter. periplasmic solute-binding protein	8	0.49	1.4E-05	32	0.46	3.7E-11
2011mTn5STM.2.02.C08	SMc02365	probable serine protease DegP1	8	0.73	1.7E-07	32	0.56	1.2E-13
2011mTn5STM.2.03.D08	SMc00399	magnesium and cobalt transport protein corA	0	NA	NA	24	0.85	5.3E-11
2011mTn5STM.2.10.E08	SMa2367	putative ABC transporter. permease	0	NA	NA	32	0.38	2.8E-06
2011mTn5STM.2.03.F08	SMa0554	hypothetical protein	0	NA	NA	4	0.31	3.6E-01
2011mTn5STM.2.07.G08	SMc03766	hypothetical protein	8	0.19	4.9E-04	32	0.27	1.7E-05
2011mTn5STM.2.05.H08	SMa0036	ABC transporter. ATP-binding protein	8	0.26	7.4E-04	32	-0.13	2.3E-01
2011mTn5STM.2.10.A09	SMc00819	KatA catalase	8	-0.08	5.2E-02	32	-0.03	7.7E-01
2011mTn5STM.2.12.B09	SMc01162	putative glutamine amidotransferase	8	0.88	2.0E-06	0	NA	NA
2011mTn5STM.2.06.C09	SMb21343	putative sugar uptake ABC transporter permease protein	8	-0.09	1.9E-02	32	-0.03	3.6E-01
2011mTn5STM.2.11.D09	pSymA, 1265903	intergenic	8	0.20	7.4E-04	32	0.39	4.6E-04
2011mTn5STM.2.01.E09	SMc00371	hypothetical protein	0	NA	NA	24	0.78	4.8E-18
2011mTn5STM.2.12.F09	SMb21073	putative OMA family outer membrane saccharide export protein. similar to ExoF	8	-0.18	1.6E-03	32	-0.21	1.4E-03
2011mTn5STM.2.07.G09	SMb20454	hypothetical protein	0	NA	NA	20	-0.11	3.6E-01
2011mTn5STM.2.06.H09	SMb20612	C4-dicarboxylate transport sensor protein DctB	8	-0.83	9.5E-07	32	-0.76	1.6E-17
2011mTn5STM.3.09.A10	SMb20264	putative amino acid ABC transporter permease protein	8	0.06	8.2E-02	32	0.22	8.4E-03
2011mTn5STM.3.09.B10	SMc03106	hypothetical transmembrane protein	8	0.64	2.4E-09	32	0.56	7.3E-14
2011mTn5STM.3.08.C10	SMb20813	lipid A export ATP-binding/permease protein MsbA	8	-0.49	2.6E-06	32	-0.40	1.9E-06
2011mTn5STM.3.06.D10	SMa1536	NuoM2 NADH-Ubiquinone/plastoquinone (complex I) oxidoreductase	8	-0.34	3.1E-05	32	-0.21	2.3E-08
2011mTn5STM.3.06.E10	SMb20707	putative adenylate cyclase protein CyaG2	8	0.20	4.3E-04	32	0.42	1.2E-08
2011mTn5STM.3.06.F10	SMc02562	phosphoenolpyruvate carboxykinase protein PckA	8	-1.04	3.3E-08	24	-0.85	1.1E-21

2011mTn5STM.3.11.H10	SMb20806	hypothetical protein	8	-0.16	5.7E-03	32	-0.16	2.8E-03
2011mTn5STM.3.05.A11	SMc01551	membrane bound O-acyl transferase	0	NA	NA	24	0.16	2.5E-02
2011mTn5STM.3.10.B11	SMb20607	hypothetical protein	8	0.72	4.6E-06	16	0.88	7.1E-07
2011mTn5STM.3.10.C11	SMb20976	putative amino acid uptake ABC transporter periplasmic solute-binding protein precursor	8	0.04	6.1E-01	32	0.05	6.4E-01
2011mTn5STM.3.09.D11	SMb20885	hypothetical membrane protein	8	0.15	7.8E-03	32	0.15	1.3E-04
2011mTn5STM.3.10.E11	SMc04091	putative protease transmembrane protein HtpX	8	0.05	2.5E-01	8	0.22	1.9E-02
2011mTn5STM.3.05.F11	SMb20508	probable sugar ABC transporter	8	0.00	8.3E-01	32	-0.06	1.9E-03
2011mTn5STM.3.09.G11	SMc02337	ABC sugar transporter. fused ATP-binding domains	0	NA	NA	8	1.24	7.2E-06
2011mTn5STM.3.08.H11	SMb20623	hypothetical protein	8	0.64	1.3E-05	32	0.67	1.9E-20
2011mTn5STM.3.08.B12	SMb20231	putative ABC transporter sugar-binding protein	8	-0.03	3.3E-01	32	-0.08	2.3E-01
2011mTn5STM.3.12.C12	SMb21530	putative acetolactate synthase isozyme I large subunit protein IlvB2	8	0.11	5.4E-02	32	-0.02	5.7E-01
2011mTn5STM.3.08.D12	SMa0166	hypothetical protein	0	NA	NA	24	-0.43	3.7E-06
2011mTn5STM.3.07.E12	SMb21039	putative oligopeptidemurein peptide ABC transporter permease protein	8	-0.36	6.9E-05	32	-0.43	1.4E-06
2011mTn5STM.3.07.F12	pSymA, 465898	intergenic	8	-0.21	4.1E-03	32	-0.31	8.6E-04
2011mTn5STM.3.08.G12	SMc03849	cytochrome c-type biogenesis protein CcmC	8	-1.06	2.4E-07	16	-0.98	1.1E-14
2011mTn5STM.3.07.H12	SMc04011	probable nitrogen assimilation regulatory protein	8	0.68	1.5E-07	32	0.70	2.7E-23
2011mTn5STM.3.07.A01	SMa0447	hypothetical protein	8	-0.37	8.9E-05	32	-0.24	3.2E-02
2011mTn5STM.3.10.B01	SMc04244	probable high-affinity zinc uptake system ATP-binding protein znuC	8	0.34	1.2E-03	32	0.26	8.1E-09
2011mTn5STM.3.11.C01	SMb20701	hypothetical protein	8	-0.17	1.2E-03	32	-0.06	1.7E-01
2011mTn5STM.3.08.D01	SMb21512	hypothetical protein	8	-0.22	1.2E-02	32	-0.43	6.1E-09
2011mTn5STM.3.11.E01	SMc01101	probable polypeptide deformylase protein Def1	8	0.35	1.2E-04	32	0.09	5.2E-01
2011mTn5STM.3.09.F01	SMb20937	putative ABC transporter protein. fused ATP-binding and permease components	8	-0.50	6.4E-05	32	-0.72	6.0E-13
2011mTn5STM.3.08.G01	SMc01816	alcohol dehydrogenase	8	0.14	1.2E-02	32	-0.43	2.6E-03
2011mTn5STM.3.10.A02	SMb20211	putative transcriptional regulator protein	8	-0.12	3.5E-01	32	-0.38	4.1E-05
2011mTn5STM.3.07.B02	SMc00963	biotin ABC transporter. permease component BioN	8	-1.09	2.5E-08	32	-1.07	2.5E-29
2011mTn5STM.3.09.C02	SMc03099	putative adenylate cyclase protein CyaF1	8	-0.32	1.1E-03	32	-0.45	7.5E-07
2011mTn5STM.3.11.D02	SMc04169	transcription regulator. AraC family	8	-0.25	4.4E-04	32	-0.32	3.1E-04
2011mTn5STM.3.08.E02	SMa1306	VirB9 type IV secretion protein	8	-0.11	5.2E-01	32	-0.29	3.0E-06
2011mTn5STM.3.08.F02	SMc03006	probable CheY chemotaxis regulator protein	8	0.40	9.0E-05	32	0.44	2.9E-07
2011mTn5STM.3.05.G02	SMc01550	hypothetical protein	8	-0.15	1.2E-01	32	-0.18	4.1E-03
2011mTn5STM.3.08.H02	SMa1418	putative ABC transporter permease protein	8	-0.06	1.8E-01	32	-0.10	1.1E-01
2011mTn5STM.3.08.A03	SMc01657	putative ferrichrome-iron receptor precursor protein FhuA2	8	0.25	6.2E-04	32	0.26	1.5E-07
2011mTn5STM.3.10.B03	SMc01817	putative transcription regulator protein	8	0.69	1.6E-06	24	-0.11	3.2E-01
2011mTn5STM.3.10.C03	SMc01929	hypothetical protein	8	-0.70	1.1E-07	32	-0.79	1.4E-16
2011mTn5STM.3.10.D03	SMb21647	periplasmic alpha-galactoside binding protein AgpA	8	0.29	1.9E-03	32	0.37	3.3E-04
2011mTn5STM.3.07.E03	SMc00109	putative transcription regulator protein	8	-0.32	2.1E-03	32	-0.21	2.6E-02
2011mTn5STM.3.07.F03	SMb21251	hypothetical membrane protein	8	-0.22	3.3E-03	32	-0.33	1.0E-05
2011mTn5STM.3.07.G03	pSymA, 127834	intergenic	8	0.09	1.5E-02	32	-0.19	3.8E-03
2011mTn5STM.3.06.H03	SMb20509	putative transcriptional regulator	8	-0.74	1.8E-04	32	-0.92	2.3E-16
2011mTn5STM.3.09.A04	SMa1220	FixN1 Heme b / copper cytochrome c oxidase subunit	8	-0.06	7.5E-02	32	-0.10	1.5E-04
2011mTn5STM.3.12.B04	SMc01579	hypothetical protein	8	0.58	9.6E-07	32	0.44	3.5E-14

2011mTn5STM.3.11.C04	SMb20771	hypothetical protein	8	0.15	1.4E-03	32	0.19	7.1E-04
2011mTn5STM.3.06.D04	SMc02360	high-affinity branched-chain amino acid transport system permease protein livH	8	0.55	6.7E-07	16	0.46	7.0E-04
2011mTn5STM.3.08.E04	SMc01986	hypothetical transmembrane protein	8	1.00	7.3E-07	8	0.62	2.2E-06
2011mTn5STM.3.09.F04	SMb20467	putative sensor kinase protein	8	-0.05	9.7E-01	32	-0.05	5.1E-01
2011mTn5STM.3.08.G04	SMa0911	hypothetical protein	8	-0.04	2.7E-01	32	-0.07	5.4E-01
2011mTn5STM.3.07.H04	SMc00926	probable glycolate oxidase protein, iron-sulfur subunit	0	NA	NA	32	0.13	2.4E-01
2011mTn5STM.3.05.A05	SMc04045	probable dihydroxy-acid dehydratase protein IlvD2	8	-0.99	2.8E-10	32	-0.85	8.9E-23
2011mTn5STM.3.11.B05	SMc01203	hypothetical transmembrane protein	8	0.35	2.1E-06	16	0.21	1.4E-02
2011mTn5STM.3.05.C05	SMa0059	dehydrogenase related to short chain alcohol dehydrogenases	8	0.17	2.6E-04	32	0.08	8.4E-02
2011mTn5STM.3.05.D05	SMa0900	possible anti-restriction protein	8	-0.33	1.9E-03	32	-0.27	8.8E-04
2011mTn5STM.3.11.E05	SMb20478	putative dipeptide ABC transporter permease and ATP-binding protein	8	-0.22	6.5E-04	32	-0.12	1.9E-02
2011mTn5STM.3.10.F05	SMc01431	probable acetolactate synthase isozyme III large subunit protein IlvI	8	-0.99	2.4E-08	32	-0.83	7.9E-23
2011mTn5STM.3.08.G05	SMb20012	hypothetical protein	8	-0.05	1.3E-01	32	-0.17	7.6E-05
2011mTn5STM.3.06.H05	SMa0247	hypothetical protein	0	NA	NA	32	-0.37	8.7E-07
2011mTn5STM.3.06.A06	SMb21298	hypothetical protein, possibly exported	0	NA	NA	32	0.16	5.1E-03
2011mTn5STM.3.11.B06	S. meliloti chromosome, 1712504	intergenic	0	NA	NA	0	NA	NA
2011mTn5STM.3.12.C06	SMa2339	putative siderophore biosynthesis protein	8	-0.48	2.1E-06	16	-0.29	1.9E-04
2011mTn5STM.3.07.D06	SMa1731	putative BetB2 betaine aldehyde dehydrogenase	8	-0.72	1.3E-07	32	-0.59	7.6E-15
2011mTn5STM.3.06.E06	SMc03254	putative antikinase protein FixT3	8	-0.16	1.3E-01	16	-0.24	6.7E-06
2011mTn5STM.3.06.G06	SMa0863	NodJ membrane transport protein	8	-1.03	1.5E-07	32	-0.82	1.0E-22
2011mTn5STM.3.12.H06	SMc02784	hypothetical protein	8	-0.52	9.2E-07	32	-0.65	2.1E-08
2011mTn5STM.3.12.A07	SMb20105	putative transcriptional regulator protein	8	-0.18	5.0E-03	32	-0.02	8.8E-01
2011mTn5STM.3.05.B07	SMa1198	possible Copper export protein	0	NA	NA	32	1.24	9.5E-32
2011mTn5STM.3.06.C07	SMb20340	hypothetical protein	8	0.15	1.1E-03	24	0.14	4.3E-03
2011mTn5STM.3.05.D07	SMb20866	hypothetical protein	8	-0.28	1.6E-05	32	-0.10	5.5E-02
2011mTn5STM.3.06.E07	pSymB, 1126745	intergenic	8	-0.34	2.0E-06	32	-0.12	5.2E-02
2011mTn5STM.3.05.F07	SMb20019	putative dihydrolipoamide succinyltransferase protein SucB	8	-0.30	3.0E-07	32	-0.37	8.1E-08
2011mTn5STM.3.08.G07	SMa1745	putative iron uptake protein	8	-0.19	3.3E-04	32	-0.15	2.3E-07
2011mTn5STM.3.05.H07	SMa1229	FixL Oxygen regulated histidine kinase	8	-1.09	5.7E-08	32	-0.91	3.4E-22
2011mTn5STM.3.07.A08	SMc00470	hypothetical transmembrane protein	8	1.32	1.1E-06	16	0.44	9.3E-02
2011mTn5STM.3.11.B08	SMa2329	probable KdpC potassium-transporting ATPase C chain	8	0.02	1.4E-01	32	0.03	3.6E-01
2011mTn5STM.3.07.C08	SMb20325	sugar ABC transporter. extracellular solute-binding protein ThuE	8	-0.15	2.7E-02	32	-0.13	7.4E-04
2011mTn5STM.3.12.D08	SMc02041	oxidoreductase. short-chain dehydrogenase/reductase family	0	NA	NA	32	0.25	4.0E-02
2011mTn5STM.3.06.E08	SMc02019	sugar transport system permease protein	8	0.76	3.1E-07	32	0.60	3.5E-09
2011mTn5STM.3.06.F08	SMc01112	probale hemolysin TlyC	8	-0.19	2.7E-03	16	0.05	5.7E-02
2011mTn5STM.3.06.G08	pSymA, 600702	intergenic	8	-0.39	9.2E-06	32	-0.25	4.0E-07
2011mTn5STM.3.10.H08	SMa2414	probable tonB-dependent receptor yncD precursor	8	-0.44	1.6E-06	32	-0.58	2.7E-06
2011mTn5STM.3.06.A09	SMc01230	putative transglycosylase	8	0.52	6.7E-06	24	0.18	2.7E-02
2011mTn5STM.3.08.B09	SMa1978	probable haloacid dehalogenase-like hydrolase	8	0.04	7.0E-02	32	0.02	3.0E-01
2011mTn5STM.3.07.C09	SMc01127	OlsB protein required for ornithine-containing lipid biosynthesis	8	1.17	2.6E-06	16	0.75	1.7E-11

2011mTn5STM.3.05.D09	SMb20426	hypothetical transcriptional regulator protein	8	1.06	2.4E-07	16	1.38	1.2E-14
2011mTn5STM.3.09.E09	SMa1057	hypothetical protein	8	-0.13	4.2E-03	32	-0.07	1.7E-02
2011mTn5STM.3.07.F09	SMc03762	hypothetical transmembrane protein	0	NA	NA	0	NA	NA
2011mTn5STM.3.08.G09	SMc03277	MFS-type transport protein	8	-0.93	1.9E-07	32	-0.88	2.6E-22
2011mTn5STM.3.06.H09	SMc00378	putative exodeoxyribonuclease protein	8	0.73	2.0E-08	24	0.34	6.3E-04
2011mTn5STM.4.11.A10	S. meliloti chromosome, 3408566	intergenic	8	-0.16	6.4E-04	24	0.04	6.1E-01
2011mTn5STM.4.06.B10	S. meliloti chromosome, 2576040	intergenic	8	0.13	2.4E-04	32	0.30	1.1E-10
2011mTn5STM.4.03.C10	SMa0561	hypothetical protein	8	-0.27	7.0E-04	32	-0.12	1.7E-02
2011mTn5STM.4.09.D10	SMa0118	hypothetical protein	8	-0.25	4.7E-05	32	-0.18	7.8E-10
2011mTn5STM.4.12.E10	SMb20492	putative short chain oxidoreductase protein	8	-0.12	8.2E-02	32	0.13	2.0E-02
2011mTn5STM.4.08.F10	SMc02514	probable amino-acid ABC transporter periplasmic binding protein precursor.	8	0.76	3.8E-08	32	0.69	4.4E-20
2011mTn5STM.4.12.G10	SMb21175	phosphate uptake ABC transporter permease protein PhoE	8	0.27	1.0E-05	32	0.34	2.4E-07
2011mTn5STM.4.08.H10	SMc04382	NdvB glycosyltransferase involved in the production of beta-(1.2)-glucan	0	NA	NA	0	NA	NA
2011mTn5STM.4.02.A11	SMb20358	putative two-component hybrid sensor and regulator, adenylate and guanylate cyclase catalytic domain	8	0.09	1.4E-02	32	0.01	8.8E-01
2011mTn5STM.4.02.B11	SMb20619	probable transcriptional regulator	8	0.23	4.7E-03	16	0.39	1.6E-03
2011mTn5STM.4.03.C11	SMa1732	hypothetical protein	8	-0.28	3.2E-07	32	-0.33	1.6E-13
2011mTn5STM.4.02.D11	SMc01139	RNA polymerase sigma-54 factor	8	-1.17	3.0E-08	24	-1.03	2.2E-25
2011mTn5STM.4.10.E11	SMa0969	putative response regulator of two-component system	8	0.06	1.7E-01	32	-0.04	6.8E-01
2011mTn5STM.4.06.F11	SMb20725	hypothetical membrane protein	8	-0.25	6.2E-06	32	-0.45	2.6E-06
2011mTn5STM.4.09.G11	SMa1273	NorB nitric oxide reductase	8	-0.20	4.0E-04	32	-0.07	3.2E-01
2011mTn5STM.4.02.H11	SMc04054	hypothetical protein	8	0.24	2.8E-04	32	0.13	4.0E-02
2011mTn5STM.4.05.A12	SMc03112	methionine synthase MetH	8	0.85	9.8E-06	31	0.84	4.5E-19
2011mTn5STM.4.10.B12	SMa0136	hypothetical protein	8	-0.13	8.3E-02	32	-0.25	2.0E-02
2011mTn5STM.4.05.C12	SMa1564	hypothetical protein	8	-0.50	5.4E-06	32	-0.58	1.2E-15
2011mTn5STM.4.10.D12	SMc03195	probable uracil-DNA glycosylase protein	8	-0.01	4.5E-01	32	-0.34	1.7E-02
2011mTn5STM.4.12.E12	SMa2085	probable ABC transporter. permease protein	8	-0.57	6.3E-07	32	-0.61	7.6E-17
2011mTn5STM.4.02.F12	SMb21588	probable sugar ABC transporter	8	-0.41	6.2E-07	24	-0.27	2.6E-02
2011mTn5STM.4.05.G12	SMa1142	FixL-related histidine kinase	8	-0.31	3.7E-04	32	-0.17	2.3E-02
2011mTn5STM.4.01.H12	SMb20609	putative two-component sensor histidine kinase protein	8	0.13	1.3E-02	24	-0.06	6.5E-02
2011mTn5STM.4.08.A01	SMc02274	putative capsule polysaccharide export protein precursor RkpU	8	-0.75	6.0E-06	32	-0.74	4.6E-10
2011mTn5STM.4.12.B01	SMc04207	rhizobiocin secretion protein RspD	8	-0.16	2.3E-02	32	-0.18	5.1E-06
2011mTn5STM.4.01.C01	SMa1649	putative ABC transporter. permease protein	8	-0.41	5.9E-05	32	-0.56	8.9E-12
2011mTn5STM.4.09.D01	SMb21240	putative MPA1 family auxiliary surface saccharide export protein	8	-0.55	1.9E-03	24	-0.66	6.2E-11
2011mTn5STM.4.04.E01	SMb20055	putative transcriptional regulator protein	8	-0.07	3.2E-01	24	-0.27	4.6E-03
2011mTn5STM.4.03.F01	SMa0172	hypothetical protein	8	-0.57	1.5E-04	32	-0.71	2.0E-09
2011mTn5STM.4.11.G01	SMb21010	D-beta-hydroxybutyrate dehydrogenase protein BdhA	8	0.24	1.2E-05	32	0.21	3.0E-05
2011mTn5STM.4.05.H01	SMc01837	urease alpha subunit protein UreC	8	0.61	2.8E-06	32	0.75	2.2E-14
2011mTn5STM.4.03.B02	SMb21488	putative cytochrome o ubiquinol oxidase chain I protein	8	-0.30	2.5E-04	32	-0.33	1.9E-10
2011mTn5STM.4.09.C02	SMc00831	probable transcription regulator	8	-0.01	6.7E-01	32	-0.01	7.7E-01

2011mTn5STM.4.09.D02	SMc00234	Putative peptidyl-prolyl cis-trans isomerase D	8	-0.40	3.9E-07	32	-0.35	9.1E-13
2011mTn5STM.4.05.E02	SMc01183	Putative LexA repressor regulator	8	-0.89	1.7E-06	32	-0.86	7.4E-19
2011mTn5STM.4.08.F02	SMc04403	probable peptidyl-dipeptidase protein	8	-0.13	5.7E-02	32	-0.31	1.4E-05
2011mTn5STM.4.11.G02	pSymB, 179255	intergenic	8	-0.38	2.0E-04	32	-0.83	9.6E-11
2011mTn5STM.4.04.H02	SMa2147	hypothetical protein	8	-0.53	8.3E-06	32	-0.72	5.1E-16
2011mTn5STM.4.02.A03	SMc02172	FrcR transcriptional regulator. ROK family	8	0.51	4.6E-03	32	0.48	7.8E-11
2011mTn5STM.4.12.B03	SMa1750	hypothetical protein	8	-0.54	6.0E-06	32	-0.55	6.7E-13
2011mTn5STM.4.02.C03	SMc00644	hypothetical protein	8	-1.13	1.0E-08	32	-1.08	3.5E-29
2011mTn5STM.4.11.D03	SMa0937	hypothetical protein	0	NA	NA	32	0.52	9.8E-09
2011mTn5STM.4.10.E03	SMa0601	hypothetical protein	8	0.33	1.4E-03	32	0.17	5.5E-03
2011mTn5STM.4.09.F03	SMc03797	homoserine O-succinyltransferase MetA	8	-1.20	6.7E-08	32	-1.20	1.9E-28
2011mTn5STM.4.01.G03	SMb20610	putative two-component response regulator protein	8	0.65	3.5E-05	32	0.42	2.0E-10
2011mTn5STM.4.10.H03	SMc01802	putative agmatinase SpeB	8	0.20	3.6E-03	32	0.15	6.8E-02
2011mTn5STM.4.06.A04	SMb21021	putative transcriptional regulator protein	8	-0.30	9.0E-04	24	-0.61	3.7E-05
2011mTn5STM.4.07.B04	SMc04317	solute-binding periplasmic protein precursor AfuA	0	NA	NA	20	0.74	6.9E-16
2011mTn5STM.4.09.C04	SMb20986	putative nitrate reductase. large subunit protein NarB	8	0.04	9.8E-01	24	-0.07	3.7E-01
2011mTn5STM.4.06.D04	SMc04244	probable high-affinity zinc uptake system ATP-binding ABC transporter protein ZnuC	8	-0.46	3.0E-07	32	-0.32	1.8E-03
2011mTn5STM.4.12.E04	SMc02271	probable ribitol type dehydrogenase RkpH	8	0.03	7.6E-01	32	-0.41	2.6E-04
2011mTn5STM.4.03.F04	SMa1455	hypothetical protein	8	-0.19	1.2E-02	24	-0.26	9.7E-11
2011mTn5STM.4.11.G04	SMb21211	putative membrane protein	8	-0.61	3.0E-08	32	-0.58	1.3E-10
2011mTn5STM.4.04.H04	SMc01265	hypothetical protein	8	-0.27	9.3E-04	32	-0.10	1.4E-01
2011mTn5STM.4.11.A05	SMc01128	putative sensor histidine kinase transmembrane protein	0	NA	NA	16	0.31	1.7E-01
2011mTn5STM.4.08.B05	SMc03900	beta-glucan export ATP-binding protein NdvA	8	-1.06	1.1E-06	16	-1.02	2.8E-14
2011mTn5STM.4.04.C05	SMa1247	NirV periplasmic nitrate reductase	8	-0.47	9.4E-07	32	-0.40	2.7E-09
2011mTn5STM.4.01.D05	pSymA, 12217	intergenic	0	NA	NA	0	NA	NA
2011mTn5STM.4.08.E05	SMa2271	hypothetical protein	8	0.53	2.6E-07	16	0.19	3.3E-01
2011mTn5STM.4.12.F05	SMc02224	putative calcium/proton antiporter transmembrane protein ChaA	0	NA	NA	8	0.26	3.7E-03
2011mTn5STM.4.12.G05	SMa2402	L-2.4-diaminobutyrate decarboxylase	8	-0.38	3.1E-06	32	-0.45	1.1E-10
2011mTn5STM.4.09.H05	SMa2229	rhizobactin/siderophore synthesis protein RhbB hypothetical protein	8	-0.33	9.3E-06	32	-0.37	1.6E-04
2011mTn5STM.4.04.A06	SMb20627	hypothetical protein	8	-0.06	1.8E-01	32	-0.32	3.2E-04
2011mTn5STM.4.11.B06	SMa0647	hypothetical protein	8	-0.13	6.0E-04	32	-0.31	1.7E-05
2011mTn5STM.4.10.C06	SMc00539	hypothetical transmembrane protein	8	0.02	4.8E-01	16	-0.45	3.3E-10
2011mTn5STM.4.08.D06	SMc02244	putative transmembrane protein	8	0.32	7.4E-05	32	-0.14	2.8E-01
2011mTn5STM.4.09.E06	SMa0414	hypothetical protein	0	NA	NA	32	-0.80	4.6E-15
2011mTn5STM.4.07.F06	SMa1471	hypothetical protein	0	NA	NA	32	-0.49	1.4E-07
2011mTn5STM.4.09.G06	SMa0941	hypothetical protein	8	-0.10	3.4E-02	32	-0.45	1.5E-05
2011mTn5STM.4.09.H06	SMa1896	putative methionine sulfoxide reductase	8	-0.41	3.7E-06	32	-0.36	5.9E-10
2011mTn5STM.4.10.A07	SMc02123	hypothetical protein	8	-0.81	1.4E-08	24	-0.77	8.4E-18
2011mTn5STM.4.06.B07	SMa0196	putative gluconolactonase precursor	8	-0.29	9.9E-05	32	-0.13	8.1E-02
2011mTn5STM.4.08.C07	SMc00114	probable protease II	8	0.01	9.1E-01	32	0.11	6.5E-04
2011mTn5STM.4.01.D07	SMc03146	type I secretion membrane fusion protein. HlyD family	8	0.23	2.9E-04	32	0.21	8.7E-05

2011mTn5STM.4.03.E07	SMa2029	hypothetical protein	0	NA	NA	32	-0.52	2.2E-17
2011mTn5STM.4.05.F07	SMc03015	VisN transcriptional regulator of motility genes LuxR family	0	NA	NA	24	0.31	1.0E-02
2011mTn5STM.4.03.G07	SMb21492	hypothetical protein	8	-0.44	2.8E-06	32	-0.39	1.2E-17
2011mTn5STM.4.08.H07	SMb20054	putative non-heme chloroperoxidase protein	8	-0.07	7.5E-03	32	0.02	3.4E-01
2011mTn5STM.4.07.A08	SMa0244	hypothetical protein	8	-0.48	3.9E-06	24	-0.19	5.3E-04
2011mTn5STM.4.10.B08	SMa1742	putative iron uptake protein	8	-0.72	2.7E-08	20	-0.73	1.0E-14
2011mTn5STM.4.01.C08	SMb21155	hypothetical protein	8	-0.34	9.0E-05	32	-0.31	1.4E-09
2011mTn5STM.4.02.D08	SMa1377	Putative amidase	8	-0.52	1.7E-06	32	-0.50	8.4E-09
2011mTn5STM.4.12.E08	SMb20486	probable sugar ABC transporter	8	0.31	1.0E-06	32	0.36	6.1E-13
2011mTn5STM.4.09.F08	SMb20362	putative inositol monophosphatase protein	8	0.12	1.9E-02	32	0.27	1.2E-05
2011mTn5STM.4.11.G08	SMb20420	putative glutathione S-transferase protein	8	-0.17	4.0E-05	32	-0.09	4.6E-01
2011mTn5STM.4.08.H07	SMc03784	hypothetical transmembrane protein	8	0.75	2.4E-05	32	0.09	5.6E-01
2011mTn5STM.4.11.A09	SMb21579	putative transcriptional regulator. merR family protein	8	-0.26	8.9E-05	32	-0.55	2.1E-07
2011mTn5STM.4.11.B09	SMc04344	hypothetical signal peptide protein	8	-0.22	2.9E-05	16	0.11	4.3E-05
2011mTn5STM.4.12.C09	SMc00726	putative thiol:disulfide interchange redox-active center transmembrane protein TlpA	8	-0.17	1.3E-04	32	-0.40	1.4E-03
2011mTn5STM.4.06.D09	SMb20614	hypothetical protein	8	-0.38	4.2E-07	32	-0.33	2.7E-07
2011mTn5STM.4.05.E09	SMc02147	probable phosphate regulon sensor histidine kinase transmembrane protein PhoR	8	0.32	4.7E-05	32	0.13	1.1E-01
2011mTn5STM.4.06.F09	SMb21222	putative transcriptional regulator	8	-0.78	2.5E-07	32	-0.66	5.5E-19
2011mTn5STM.4.02.G09	SMc00925	hypothetical protein	8	0.53	6.2E-06	32	0.45	4.7E-14
2011mTn5STM.4.08.H09	SMc03004	putative methyl-accepting chemotaxis protein McpE	8	0.01	9.4E-01	20	0.32	7.4E-03

[a] For mutants carrying a transposon insertion in an intergenic region the exact position of insertion in the respective replicon is given.

G.5 LIST OF ABBREVIATIONS

% (v/v)	Volume percent	K-iodide	Potassium iodide
% (v/v)	Weight percent	Km	Kanamycin
°C	Degrees Celsius	LMW	Low molecular weight
AT	Sum of all the bases A and T	mg	Milligram
BPB	Bromphenol blue	min	minute
BSA	Bovine serum albumin	ml	Milliliter
Cam	Chloramphenicol	mM	Millimolar
cv	Cultivar	mRNA	Messenger RNA
Cy3	Cyanine 3	µ	Micro
Cy5	Cyanine 5	µg	Microgram
dATP	2'-deoxyadenosine 5'-triphosphate	µl	Microliter
dCTP	2'-deoxycytidine 5'-triphosphate	µm	Micrometer
dGTP	2'-deoxyguanosine 5'-triphosphate	µM	Micromolar
dTTP	2'-deoxythymidine 5'-triphosphate	Nod^-	Non-nodulating phenotype
DTT	Dithiothreitol	Nod^+	Nodulating phenotype
DMSO	Dimethyl Sulfoxide	Nm	Neomicin
DNA	Deoxyribonucleic acid	Nx	Nalidixic acid
dNTPs	Deoxynucleotide triphosphates	o.D.	Optical density
dpi	Days post inoculation	PEG	Polyethylene glycol
Dr.	Doctor	pmol	Pikomole
dsDNA	Double-stranded DNA	PCR	Polymerase chain reaction
EDTA	Ethylenediaminetetraacetic acid	pH	the logarithm of the reciprocal of hydrogen-ion concentration in gram atoms per liter
et al.	Et alii (lat.)/and others	RNA	Ribonucleic acid
EtOH	Ethanol	RNase	Ribonuklease
ed/eds	Editor/editors	rpm	Rotations per minute
Fig.	Figure(s)	SAP	Shrimp alkaline phosphatase
Fix^-	Non-nitrogen fixing phenotype	sec	Seconds
Fix^+	Nitrogen fixing phenotype	SDS	Sodium dodecyl sulfate
g	Gramm	Sm	Streptomycin
× g	Times gravity	STM	Signature tagged mutagenesis
GC	Sum of all the bases G and C	U	Unit
Gm	Gentamycin	UV	Ultra-Violet
h	Hour(s)	V	Volt
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid	Vol	Volume
HMW	High molecular weight	X-Gluc	5-Bromo-4-chloro-3-indolyl-ß-D-glucuronide

G.6 LIST OF FIGURES

G.7 LIST OF TABLES

G.8 ACKNOWLEDGEMENTS

At the first place, I would like to thank all the members of the Chair of Genetics, the Faculty of Technology and the NRW Graduate School of Bioinformatics and Genome Research at Bielefeld University, who directly or indirectly contributed to this work.

I appreciate the possibility to accomplish my PhD work at the Chair of Genetics led by Prof. Dr. Alfred Pühler. I am grateful also to the NRW Graduate School of Bioinformatics and Genome Research for financing my scholarship and conference travelling.

I am especially grateful to my tutor, HD Dr. Anke Becker, for the formulating the idea and providing resources for this project, for the great mentoring and advice and for all the help in the writing of the papers and the thesis. Her patience and encouragement, and her bright ideas have been a permanent source of motivation for me.

Furthermore, I thank Juniorprof. Dr. - Ing. Tim W. Nattkemper for his advice in the statistics and bioinformatics, Danijel Wetter and Dr. Alexander Goesmann for the statistical analysis of the mutant library, Ulf Schiller and Stefan Kurtz for the design of tags and the artificial linker, and Silke Szymczak, for the STM microarray data analysis, which was essential for this project.

I would like to thank also Chip and Exopol group members for all the support and advice. Special thanks go to Dr. Lisa Krol and Dr. Larissa Sharypova, for their professional help and for their friendship. I thank Eva Schulte-Berndt and Manuela Mayer for the assistance in microarray hybridizations and maintenance of the mutant library. Thanks also go to Dr. Birgit Baumgarth, Christelle Bahlawane and Javier Serrania for the everyday help in the lab, and for the help in the harvesting of nodules from the aeroponics. I would like to thank also Dr. Matthew McIntosh for the proof-reading of the thesis.

I am appreciated to Dr. Natalija Hohnjec and to Markus Baier, whose help with aeroponics and acetylene reduction assay was of great value for this project.

I thank also my family and my fiancée Héctor López Romero, for understanding and emotional support, and for always staying by my side.