Research

Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis

Frédéric Ampe, Ernö Kiss, Frédérique Sabourdy and Jacques Batut

Address: Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR 215 Centre National de la Recherche Scientifique - Institut National de la Recherche Agronomique, BP27-31326 Castanet-Tolosan cedex, France.

Correspondence: Jacques Batut. E-mail: jbatut@toulouse.inra.fr

Published: 31 January 2003

Genome ***Biology*** 2003, **4**:R15

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2003/4/2/R15

Received: 19 September 2002
Revised: 14 November 2002
Accepted: 18 December 2002

Abstract

Background: Rhizobia induce the formation on specific legumes of new organs, the root nodules, as a result of an elaborated developmental program involving the two partners. In order to contribute to a more global view of the genetics underlying this plant-microbe symbiosis, we have mined the recently determined *Sinorhizobium meliloti* genome sequence for genes potentially relevant to symbiosis. We describe here the construction and use of dedicated nylon macroarrays to study simultaneously the expression of 200 of these genes in a variety of environmental conditions, pertinent to symbiosis.

Results: The expression of 214 *S. meliloti* genes was monitored under ten environmental conditions, including free-living aerobic and microaerobic conditions, addition of the plant symbiotic elicitor luteolin, and a variety of symbiotic conditions. Five new genes induced by luteolin have been identified as well as nine new genes induced in mature nitrogen-fixing bacteroids. A bacterial and a plant symbiotic mutant affected in nodule development have been found of particular interest to decipher gene expression at the intermediate stage of the symbiotic interaction. *S. meliloti* gene expression in the cultivated legume *Medicago sativa* (alfalfa) and the model plant *M. truncatula* were compared and a small number of differences was found.

Conclusions: In addition to exploring conditions for a genome-wide transcriptome analysis of the model rhizobium *S. meliloti*, the present work has highlighted the differential expression of several classes of genes during symbiosis. These genes are related to invasion, oxidative stress protection, iron mobilization, and signaling, thus emphasizing possible common mechanisms between symbiosis and pathogenesis.

Background

Rhizobia can live either as free-living soil bacteria or as nitrogen-fixing symbionts of plants of the family Leguminosae. The bacteria induce the formation of a specific new organ, the root nodule, as a result of an elaborate developmental program directed by an exchange of signals between the two partners. Plant flavonoids secreted by the roots trigger Nod factor production by the bacteria [1]. In turn, on their specific host plant, Nod factors induce a transduction pathway leading to nodule development [2].

When bacteria growing in the rhizosphere or on the root surface become trapped between two epidermal root-hair cell walls, alteration and invagination of a root-hair cell wall initiates the development of an infection thread, a tubular structure in which the bacteria penetrate the plant and propagate

further towards the inner cortex where the infection threads ramify. Among the few known bacterial genes required for initiation and elongation of infection threads, those directing the production of different types of cell-surface polysaccharides have a major role, possibly by protecting the invading bacteria against plant defense mechanisms [3]. Oxidative stress protection through the production of detoxifying enzymes is also essential for the survival of symbiotic bacteria during infection [4]. Bacterial cells at the tip of the infection threads eventually enter the nodule cells through a process that appears to involve binding to the root-cell cytoplasmic membrane and uptake into the cells by endocytosis [5,6]. Once inside the plant cells, the bacteria differentiate into non-dividing cells called bacteroids. The genetic control of endocytosis and bacteroid differentiation is essentially unknown, with the exception of BacA, an inner membrane protein affecting cell envelope composition [7]. *Sinorhizobium meliloti bacA* mutants are released from the infection threads but senesce rapidly before bacteroid differentiation, possibly because of their increased sensitivity to stress [7,8]. Finally, bacteroids synthesize proteins essential for symbiotic nitrogen fixation such as those responsible for dicarboxylic acid transport (*dct* genes) [9], nitrogenase synthesis (*nif* genes) and microoxic respiration (*fix* genes) [10]. In nodules, *nif* and *fix* gene expression is driven by oxygen limitation [11].

Despite the information outlined above, we lack a global view on how the rhizobium-legume symbiosis develops. In particular, we need to learn more about bacterial genes that control adhesion to plant cells, colonization of infection threads, invasion of plant cells and bacteroid differentiation. Genomics is, of course, of particular interest in this respect. Perret *et al*. pioneered this approach in rhizobia by determining the complete nucleotide sequence of the symbiotic plasmid of *Rhizobium* NGR234 as well as its transcriptome during symbiosis [12].

In this work we have inspected the recently determined *S. meliloti* genome sequence [13] and identified around 200 genes potentially involved in critical steps of the establishment of the rhizobium-legume symbiosis such as attachment, entry into plant cells, development and survival in infection threads and differentiation into nitrogen-fixing bacteroids. Specific macroarrays were designed to monitor expression of these genes during symbiosis with *Medicago truncatula*. Results highlighted genes expressed at different stages of symbiotic interaction, including the stage during which infection takes place.

Results

Array design and bacterial growth conditions

We designed nylon macroarrays featuring 214 genes (3.5% of the total genome content) selected by mining the *S. meliloti* genome sequence. We focused on genes potentially related to adhesion and attachment, oxidative stress protection, iron metabolism, invasion, calcium binding, production of toxins and proteases, cell-surface organization and regulation, on the basis of overall similarity to known genes or the presence of specific motifs. Thirty-four control genes from *S. meliloti* were included, the regulation of which was known under at least some of the conditions studied. The distribution of the target genes between the three replicons of *S. meliloti* was as follows: 126 genes from the chromosome (3.7 megabases (Mb)), 47 from pSymA (1.4 Mb) and 41 from pSymB (1.7 Mb). Finally, five genes from *Corynebacterium* with no homolog in *S. meliloti* were included to assess the quality of the hybridizations. The comprehensive list of the genes included in the custom nylon macroarrays as well as the complete expression results are available as additional data files with the online version of this paper and from [14].

The expression of the 214 *S. meliloti* genes was monitored in cells grown in liquid cultures or in symbiosis with the legume hosts *M. truncatula* (the model legume) and *M. sativa* (alfalfa). Culture conditions included growth in minimal and rich media, oxygen limitation, and growth in the presence of the plant flavonoid luteolin. Growth rate after addition of luteolin was similar to that observed for the control cultures. In oxygen-limited cultures, growth was still observed but was linear rather than exponential, and therefore significantly slower than in the aerated cultures. Several symbiotic situations were assayed. Wild-type *M. truncatula* Jemalong J6 plants were inoculated with wild-type *S. meliloti* strain 1021 and mature nitrogen-fixing nodules were collected 18 days after inoculation (dai). Wild-type *M. sativa* cv. Sitel was similarly inoculated with *S. meliloti* 1021 to enable a comparison between the two host plants.

Medicago indeterminate nodules are characterized by a persisting meristem. As a result, bacterial and plant cells at different developmental stages coexist within a single mature nodule [5]. In order to focus on gene expression at the intermediate stage of the interaction, after nodule formation and prior nitrogen fixation, two mutants were used. First, a *bacA* mutant of *S. meliloti* that is impaired in bacteroid differentiation [8] was inoculated to *M. truncatula* J6 (wild type) to yield *bacA*-J6 nodules lacking a fixation zone. Second, wild-type *S. meliloti* 1021 was inoculated to a *M. truncatula* mutant (TE7) forming nodules enriched in infection threads but lacking a fixation zone [15]. Nodules collected from both combinations (*bacA*-J6 and 1021-TE7 nodules) were expected to be enriched in mRNA synthesized during infection. In addition, young (8 dai) wild-type nodules (1021-J6) were collected as they were also expected to contain a larger proportion of bacteria in the infection zone as compared to mature nodules that mainly contain nitrogen-fixing bacteroids. Total RNA extracted from cultured and symbiotic bacteria was reverse transcribed *in vitro*, yielding complex ^{33}P-labeled cDNA probes. These probes were hybridized to the nylon macroarrays described above and relative mRNA levels of the 214 test genes were determined (see Materials and methods).

Global expression patterns

Global gene expression patterns were analyzed by two complementary methods: principal component analysis and hierarchical clustering (see Materials and methods). The two main axes of principal component analysis accounted for 47% and 17% of the total variability (Figure 1a). Axis 1 (abscissa) clearly separated experiments in liquid cultures (A-E) and symbiotic conditions (F-I). Hence, many of the genes tested were differentially expressed in cultured and symbiotic bacteria. Expression patterns corresponding to the five tested culture conditions grouped tightly together and were therefore highly similar. Interestingly, expression profiles of aerated cultures in rich and minimal media were very much alike, thus suggesting that few of the chosen genes are metabolically regulated. Axis 2 (ordinate) separated the different symbiotic conditions. Data from mature Fix^+ nodules clearly separated from expression results in Fix^- nodules (1021-TE7 and *bacA*-J6). Expression profiles from the two developmental mutants, although close, were not identical. Young wild-type nodules exhibited an intermediate position between mature nodules and the mutants.

Hierarchical clustering first confirmed the conclusions of principal component analysis (Figure 1b): patterns from

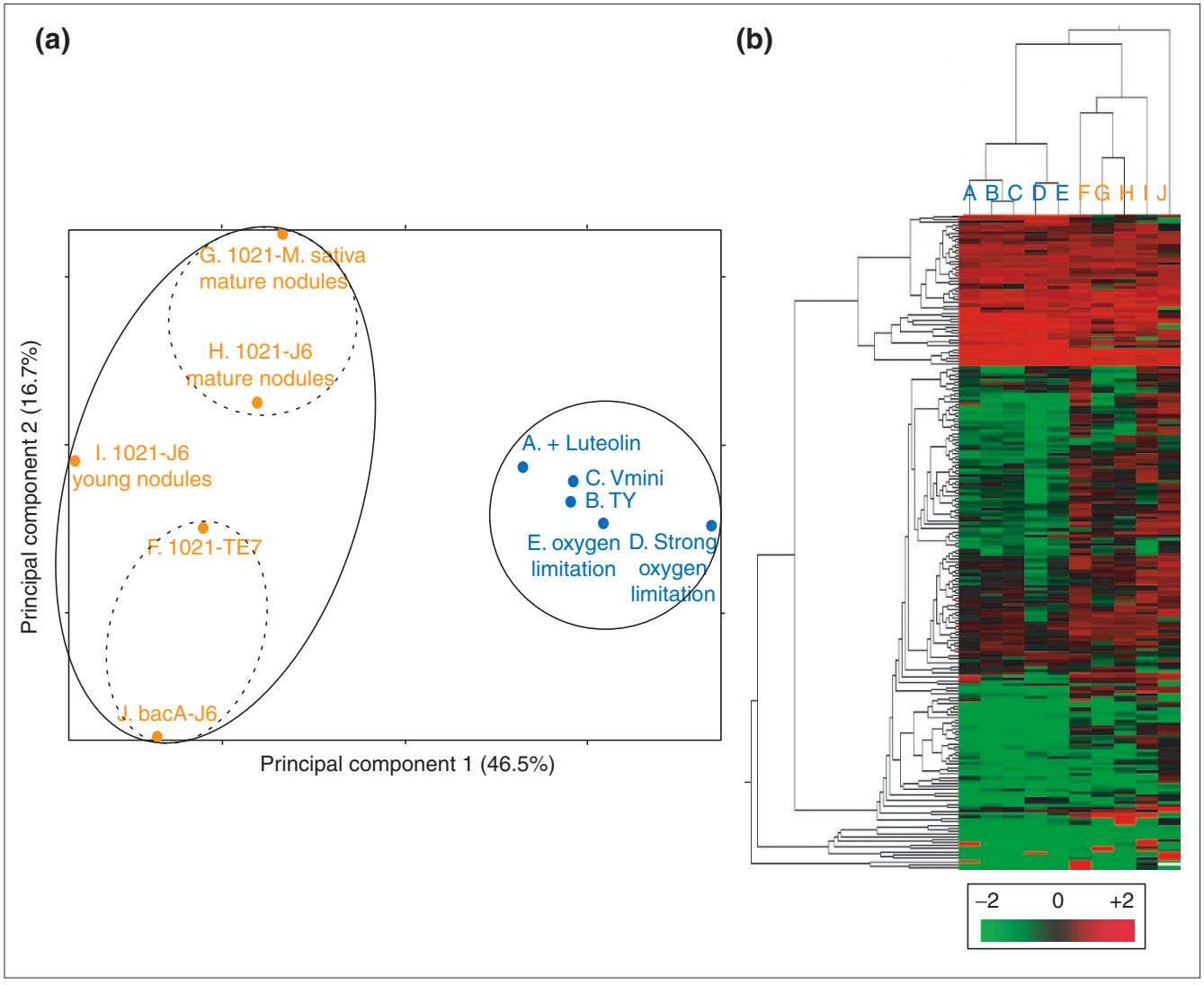

Figure 1

Global analysis of gene expression profiles in liquid cultures and symbiotic conditions. **(a)** Principal component analysis. Projection on the two main axes of principal gene analysis. A-E (blue) are profiles derived from cultured bacteria. A, minimal medium (Vmini) + 10 μM luteolin; B, rich medium; C, aerated Vmini; D, microoxic Vmini (strong oxygen limitation); E, microoxic Vmini. F-J (orange) are profiles derived from symbiotic bacteria. F, 1021-TE7; G, 1021-*M. sativa*; H, 1021-J6 mature nodules; I, 1021-J6 young nodules; J, *bacA*-J6. **(b)** Hierarchical clustering of gene-expression data. Each row represents a separate gene target (PCR product) on the macroarrays and each column (A-J) a specific bacterial growth condition as indicated above. The results presented represent the normalized expression level and are depicted according to the color scale shown at the bottom.

liquid cultures (A-E) formed a group distinct from those derived from symbiotic bacteria (F-I). It also allowed the identification of several clusters of genes sharing similar expression profiles. One cluster contained genes, which were highly and constantly expressed in all conditions tested (for example, *SMc00715*, *SMc00979*). Genes induced or repressed *in planta*, genes induced during infection, and genes induced by luteolin also formed individual clusters. A careful analysis of expression profiles and numerical data enabled us to sort out genes whose might be particularly relevant to the plant-microbe interaction. The following sections present the most significant results.

Control genes

To assess the overall validity of our experimental procedure, we carefully analyzed the expression of a set of 34 control genes with partly known regulation, including up- and down-regulated genes as well as genes expected to be constitutive under the conditions tested. The Nod factor biosynthesis genes were the controls for luteolin induction. As expected [1], *nodC* and *nodA* were induced (5-7-fold) by this plant flavonoid (Table 1). Genes whose expression is induced in *S. meliloti* bacteroids from *Medicago* nodules were also used as positive controls on our macroarrays. These included genes coding for nitrogenase (*nifH*, *nifN* and the *nifA* regulator), a microoxic respiratory chain (*fixN*, *fixP*), regulators (*fixL*, *fixJ*, *fixK* and *fixT*), the pyruvate dehydrogenase (*pdhA*) the nodule-expressed gene *nex18* [10,16,17] and the oxygen-regulated gene *tspO* [18]. As previously documented, the level of expression of *nif* and *fix* genes was close to, or below, the detection limit in liquid aerated cultures whereas they were all detected within nodules (Table 2), with the exception of *fixJ*. Very high *nifH* expression levels were observed in bacteroids, particularly when isolated from *M. sativa*. Our results also showed significantly higher expression levels of *nex18 in planta* than in aerated cultures as already described (*nex18* was previously identified as being specifically expressed in nodules) [17]. In addition, we measured a 2.5-fold induction of *nex18* in oxygen-limited cultures, suggesting that this gene, just like the gene next to it, *tspO*, is under oxygen control [18]. We show here that *tspO* is also significantly induced in nodules. The *pdhA* gene, coding for pyruvate dehydrogenase, was induced in nodules as previously reported [16]. The *dctA* gene was induced in minimal medium containing succinate as expected.

Contrary to the situation in mature bacteroids, little is known about gene expression during the infection process. However, gene fusion experiments have shown that *nifH*, which is highly expressed in bacteroids, was not expressed in bacteria from the infection zone of mature nodules [11]. Consistently, our data confirmed high expression of *nifH* in mature bacteroids from both *M. truncatula* and *M. sativa* and very low or no expression in symbiotic bacteria from the two developmental mutants (Table 2). The same applied to the transcriptional regulator *nifA*. Conversely, *nod* genes

Table 1

Genes induced by luteolin

Gene ID	Gene name	Description	Cultured bacteria					Symbiotic bacteria				
								M. truncatula				*M. sativa*
			V + O_2	TY	VμO_2 weak limitation	VμO_2 strong limitation	Luteolin	J6-1021 18 dai	J6-1021 8 dai	TE7-1021	J6-bacA	Sitel-1021
SMa0866	*nodC*	Nod factor biosynthesis protein	1	0.8	0.7	0.6	4.7 (+)	1.4	1.4	3.1 (+)	2.9 (+)	0.6
SMa0869	*nodA*	Nod factor biosynthesis protein	1	1.1	0.7	1.0	6.7 (+)	2.4	4.3 (+)	7.1 (+)	ND	1.0
SMa2339	*sma2339*	Probable siderophore biosynthesis protein	1	0.8	0.7	0.6	3.6 (+)	1.4	1.5	1.4	2.4 (+)	1.4
SMc01516	*smc01516*	Unknown function	1	1.0	1.5	1.2	5.7 (+)	3.5 (+)	4.8 (+)	3.5 (+)	1.5	3.0 (+)
SMc02509	*sitA*	Iron transporter	1	1.5	1.1	0.7	12.2 (+)	3.6 (+)	4.3 (+)	3.8 (+)	4.8 (+)	4.3 (+)
SMb21087	*traA2**	Probable conjugal transfer protein	1	0.6	0.7	0.9	2.8 (+)	1.0	1.1	0.9	0.4 (-)	1.6 (+)
SMc03785	*ialA*	Probable invasion protein	1	1.0	4.3 (+)	2.8 (+)	9.2 (+)	2.9 (+)	2.8 (+)	1.6	ND	3.8 (+)

Numerical values give induction folds (ratios) compared to growth in aerated minimal medium. +, -, Genes significantly upregulated or downregulated, respectively, compared to growth in aerated minimal medium (V + O_2). TY, growth in rich medium; VμO_2, microoxic growth; luteolin, induction by luteolin. ND, Not determined. *In the DNA region amplified for the macroarray studies, *traA1* and *traA2* share 95% identity. The signal is therefore a combination of the expression of the two genes.

Table 2

Genes induced in symbiosis

Gene ID	Gene name	Description	Cultured bacteria					Symbiotic bacteria				
								M. truncatula				*M. sativa*
			V + O_2	TY	VμO_2 weak limitation	VμO_2 strong limitation	Luteolin	J6-1021 18 dai	J6-1021 8 dai	TE7-1021	J6-bacA	Sitel-1021
Controls												
SMa0815	*nifA*	Transcriptional activator of nitrogenase genes	1	1.5	2.8 (+)	1.4	2.1	20 (+)	24 (+)	6.0 (+)	1.9	27 (+)
SMa0825	*nifH*	Nitrogenase Fe protein	1	0.9	0.6	0.7	1.7	13 (+)	13 (+)	2.7 (+)	0.6	58 (+)
SMa0873	*nifN*	Nitrogenase Fe-Mo cofactor biosynthesis protein	1	0.9	ND	0.7	0.9	2.4 (+)	4.2 (+)	3.5 (+)	4.8 (+)	4.4 (+)
SMa1225	*fixK1*	Transcriptional activator of fix genes	1	0.8	2.1 (+)	2.4 (+)	2.3	6.1 (+)	9.6 (+)	6.6 (+)	8.3 (+)	4.2 (+)
SMa1220	*fixN1*	Heme b/copper cytochrome c oxidase subunit	1	2.0	15 (+)	17 (+)	3.7	13 (+)	31 (+)	5.3 (+)	19 (+)	8.8 (+)
SMa1213	*fixP1*	Di-heme cytochrome c	1	1.9	8.7 (+)	11 (+)	3.7	11 (+)	8.9 (+)	6.7 (+)	4.6 (+)	4.4 (+)
SMa1226	*fixT1*	Inhibitor of FixL autophosphorylation	1	1.9	3.4	2.3	2.9	3.7 (+)	7.9 (+)	5.3 (+)	ND	1.2
SMa1229	*fixL*	Oxygen-regulated histidine kinase	1	1.5	10 (+)	ND	1.6	18 (+)	17.5 (+)	17.5 (+)	11.5 (+)	0.5
SMc00819	*katA*	Catalase	1	0.6	1.1	1.0	1.3	3.9 (+)	3.8 (+)	1.9	2.0	1.9 (+)
SMa1077	*nex18*	Hypothetical protein	1	1.3	2.4 (+)	2.6 (+)	1.2	2.3 (+)	1.5	3.5 (+)	2.3 (+)	1.4
SMa1079	*tspO*	Tryptophan rich sensory protein homolog	1	1.1	2.4 (+)	1.6	1.4	1.9 (+)	1.8 (+)	2.7 (+)	1.7 (+)	2.0 (+)
SMc01030	*pdhAa*	Pyruvate dehydrogenase	1	1.3	1.3	1.9	1.5	2.0 (+)	1.8	2.5	ND	2.1 (+)
New genes induced in bacteroids												
SMc02285	*cyaE*	Adenylate cyclase	1	1.0	1.0	0.7	0.9	1.4 (+)	2.5 (+)	1.1	2.5 (+)	1.5 (+)
SMc03099	*cyaF1*	Adenylate cyclase	1	1.0	0.8	0.8	0.8	1.9 (+)	2.1 (+)	1.3	1.4 (+)	1.1
SMb20539	*cyaF6*	Adenylate cyclase	1	1.0	1.3	0.9	0.8	3.3 (+)	3.0 (+)	2.1 (+)	2.1 (+)	2.4 (+)
SMa0841	*sma0841*	Transposase (hypothetical protein)	1	1.1	0.7	0.5	0.9	3.7 (+)	8.8 (+)	3.5 (+)	3.8 (+)	2.7 (+)
SMb20518	*smb20518*	Unknown function (possible lytic enzyme)	1	0.9	0.8	0.5 (-)	1.0	2.9 (+)	3.5 (+)	1.6 (+)	2.4 (+)	1.6 (+)
SMc02394	*smc02394*	Virulence associated protein homolog	1	1.2	1.2	1.1	1.0	1.8 (+)	1.0	1.2	1.8 (+)	1.7 (+)
Genes induced in symbiosis and by luteolin												
SMc03785	*ialA*	Probable invasion protein	1	1.0	4.3 (+)	2.8 (+)	9.2 (+)	2.9 (+)	2.8 (+)	1.6	ND	3.8 (+)
SMc02509	*sitA*	Iron transporter	1	1.5	1.1	0.7	12.2 (+)	3.6 (+)	4.3 (+)	3.8 (+)	4.8 (+)	4.3 (+)
SMc01516	*smc01516*	Unknown function	1	1.0	1.5	1.2	5.7 (+)	3.5 (+)	4.8 (+)	3.5 (+)	1.5	3.0 (+)
Other genes induced in *M. truncatula* but not in *M. sativa*												
SMb20838	*smb20838**	Hemolysin-type calcium-binding protein	1	1.4	1.1	2.6 (+)	0.9	2.1 (+)	2.2 (+)	2.4 (+)	3.0 (+)	0.8
SMa0627	*aqpZ2*	Aquaporin	1	1.0	1.2	0.8	1.1	2.2 (+)	1.8 (+)	1.1	1.4	0.7
SMc03784	*smc03784*	Unknown function	1	1.3	1.2	0.8	1.0	2.7 (+)	2.1 (+)	0.8	1.2	0.9

Legend as for Table 1. *In the DNA region amplified, SMb8038 and SMc00287 share > 90% identity. The signal is therefore a combination of the expression of the two genes.

were expressed in 8 dai nodules as well as in 1021-TE7 and *bacA*-J6 nodules, whereas expression levels were lower in 18 dai wild-type bacteroids from both *M. sativa* and *M. truncatula*. This result is consistent with the observation that the *nodABC* genes from *R. leguminosarum* are transcribed in infection threads and that transcription is switched off at the onset of bacteroid differentiation [19] and with the detection of Nod factors in the infection zone of alfalfa nodules, which contains the infection threads [20].

Twelve control genes showed no significant differential expression in the conditions tested. This behavior was expected for *dnaJ*, *dnaK*, *exoX*, *exoY*, *ftsK*, *glnD*, *nodD1*, *ntrB*, *ntrC* and *trpE*. Instead, the induced expression of *ndiA* and *ndiB* by oxygen limitation [18] was not confirmed in this study. This discrepancy could be due to differences in oxygen-limitation conditions. Conversely, *hemA*, which was thought to be constitutive [16], was actually slightly induced under oxygen limitation, although expression levels in aerobic conditions were always high. This finding was confirmed by RT-PCR experiments (data not shown). Six other control genes (*ftsZ1*, *fusA*, *glnB*, *groEL1*, *groES1*, *rpsA*) were repressed in symbiosis as anticipated (see below and Table 3).

Genes induced by luteolin

Plant flavonoids were recently shown to influence the expression of several genes besides *nod* genes in both transcriptome [12] and proteome analyses [21], although the identity of the differentially expressed proteins could not be clearly determined from the latter study. In our work, five *S. meliloti* genes were significantly induced by luteolin in addition to the control *nod* genes (Table 1). The two genes coding for the conjugal transfer protein TraA of pSymA and pSymB were induced 2.8-fold by luteolin. Although the overall biological function of these genes is presently unknown, this result suggests that their associated transfer functions could be stimulated by the presence of luteolin. Three other genes forming a tight cluster in Figure 1b were induced by luteolin, and also in nodules (Table 1). Most interestingly, these three genes are related to iron metabolism. *SMa2339* codes for a probable siderophore, *sitA* for an iron transporter and *SMc01516* is homologous with a gene of the *tonB-hmu* cluster of *R. leguminosarum* whose precise role is not known [22]. The fifth gene induced by the plant flavonoid was *ialA*, a gene coding for an NTPase required for the infection phenotype of *Bartonella* [23]. This gene showed a complex expression pattern: it was induced by luteolin, by oxygen limitation and in symbiotic bacteria from mature Fix$^+$ nodules but not in cultured oxygenated bacteria. This suggests that IalA may indeed play a part in different stages of the symbiotic interaction.

Genes induced in mature N_2-fixing nodules

Six genes were specifically induced in symbiotic bacteria from wild-type Fix$^+$ nodules. Interestingly, none of them was significantly stimulated by oxygen limitation *ex planta* (Table 2). Three of them code for nucleotide cyclases (*cyaE*, *cyaF1* and *cyaF6*), two of which (*cyaF1* and *cyaF6*) belong to a family so far unique to *S. meliloti* [24]. *SMa0841*, a gene coding for the transposase lying in the *syrM-nodD3* intergenic region, was also induced in symbiotic bacteria. Induction of transposases *in planta* was also reported for the symbiotic plasmid of *Rhizobium* NGR234 [12]. Two other genes induced in nodules were *SMb20518* and *SMc02394*. The function of SMb20518 is unknown but this protein harbors a domain found in lytic enzymes of *Mycobacterium*, *Haemophilus* or *Pseudomonas* phages. SMc02394 is homologous to virulence-associated proteins of the VapA/VapI family. The best documented homolog is VapI from the footrot disease agent *Dichelobacter nodosus* [25], but homologs are also present in both enterohemorrhagic and nonpathogenic strains of *Escherichia coli*, in *Vibrio cholerae* and in *Mesorhizobium loti*. Finally, our results show an induction of *katA* in all nodules tested with higher levels of transcripts in bacteroids than during infection. Consistently, a KatA-catalase activity was found in 5-week-old nodules of *M. sativa* [26]. We show here that this is the result of higher transcription of *katA* in bacteroids. In addition to these six genes, three genes (*ialA*, *sitA* and *SMc01516*) were induced in nodules, as well as by luteolin as previously discussed.

Genes potentially induced during infection

Among the genes induced *in planta*, a number were found to be preferentially induced in the young nodules as well as in the two developmental mutants, as compared to mature nodules. These genes are thus good candidates for being specific of the infection stage. Fifteen of the 214 tested genes were significantly induced in *bacA*-J6 and 1021-TE7 nodules and in young wild-type nodules (8 dai) compared to mature nodules (Table 4). The highest level of induction was found for a calcium-binding protease (SMa0034) from pSymA. Most interestingly, three of the 26 nucleotide cyclases tested appeared to be characteristic of the infection stage (*cyaD1*, *cyaF4*, and *SMa1591*). Also, several genes from pSymA coding for a *virB-like (avhB)* type-IV secretion system were induced in *bacA*-J6 and 1021-TE7 nodules. Two genes coding for calcium-binding proteins (*SMa2111* and *SMc04171*) and a third one coding for a toxin secretion protein (*SMc02661*) were also induced during infection. *SMc02156* was also induced by oxygen limitation. The known homologs of *SMc02156* are involved in pathogenesis in enterohemorrhagic *E. coli*, *V. cholerae* (*acfC*) and *Campylobacter jejuni*. However, no homolog was found in *M. loti*. Oxygen limitation is probably one of the signals triggering expression of *SMc02156 in planta*. Of the 14 outer membrane proteins tested, one (RopB2) was specifically expressed during infection. Two genes related to oxidative stress were also induced during infection: a newly described superoxide dismutase (SodC) and a transmembrane efflux protein (SMb20338) homologous to PqrA (77% identity) from *Ochrobactrum anthropi*, which confers resistance to paraquat. Finally, the gene

Table 3

Genes repressed in symbiosis

Gene ID	Gene name	Description	Cultured bacteria					Symbiotic bacteria				
								M. truncatula				*M. sativa*
			V + O_2	TY	VμO_2 weak limitation	VμO_2 strong limitation	Luteolin	J6-1021 18 dai	J6-1021 8 dai	TE7-1021	J6-bacA	Sitel-1021
Cell structure, energy metabolism and protein synthesis												
SMc02400	*smc02400*	Outer membrane protein	1	1.3	2 (+)	2 (+)	1.2	0.2 (-)	0.1 (-)	0.3 (-)	0.2 (-)	0.1 (-)
SMc02396	*smc02396*	Outer membrane protein	1		2 (+)	2 (+)	2.1	0.8	0.3 (-)	0.5 (-)	0.6 (-)	0.6 (-)
SMc01312	*fusA1*	Elongation factor G	1	1.3	0.4 (-)	1.3	0.9	0.3 (-)	0.1 (-)	0.4 (-)	0.5 (-)	0.1 (-)
SMc01874	*ftsZ1*	Cell division protein	1	1.1	0.9		0.8	0.6 (-)	0.6 (-)	0.5 (-)	0.8	0.6 (-)
SMc00335	*rpsA*	30 S ribosomal protein S1	1	1.1	0.3 (-)	0.6	0.4	0.3 (-)	0.2 (-)	0.3 (-)	0.3 (-)	0.3 (-)
SMc03239*	*ppa*	Inorganic pyrophosphatase	1	1	0.4 (-)	0.8	0.8	0.2 (-)	0.2 (-)	0.3 (-)	0.4 (-)	0.1 (-)
SMc00870	*atpE*	ATP synthase	1	0.7	0.6	1.1	2	0.4 (-)	0.4 (-)	0.3 (-)	0.4 (-)	0.2 (-)
SMc00287	*smc00287*	(NADH ubiquinone) oxidoreductase	1	0.9	0.7	1.1	0.7	0.5 (-)	0.7	0.4 (-)	0.6	0.5 (-)
SMc00187	*fbcF*	Ubiquinol cytochrome c reductase	1	0.8	2 (+)	2.5 (+)	0.7	0.3 (-)	0.4 (-)	0.3 (-)	0.2 (-)	0.6 (-)
Transcriptional regulators												
SMc01110	*phrR*	DNA-binding protein - Transcriptional regulator	1		0.8		0.9	0.9	0.4 (-)	0.6	0.5 (-)	0.3 (-)
SMb21117	*smb21117*	Transcriptional regulator	1		0.8	0.6	0.8	0.7 (-)	0.4 (-)	0.6 (-)	0.9	0.6 (-)
SMc00170	*smc00170*	LuxR transcriptional regulator	1	1.1	0.6 (-)		0.9	0.6 (-)	0.3 (-)	0.3 (-)	0.2 (-)	0.3 (-)
SMc00829	*smc00829*	Transcriptional regulator (probable pdhR)	1	0.6	0.8	0.8	0.5	0.3 (-)	0.2 (-)	0.3 (-)	0 (-)	0.2 (-)
Others												
SMc00912	*groES1*	Chaperonin	1	0.8	0.3 (-)	0.5 (-)	0.4	0.3 (-)	0.3 (-)	0.3 (-)	0.1 (-)	0.3 (-)
SMc00913	*groEL1*	Chaperonin	1	1.3	0.3 (-)	0.5 (-)	0.6	0.5 (-)	0.6 (-)	0.5 (-)	0.1 (-)	0.4 (-)
SMc03786	*bfr*	Bacterioferritin	1	0.7	1.7	1.3	0.8	0.4 (-)	0.3 (-)	0.6	0.5	0.2 (-)
SMc04458	*secA*	Preprotein translocase	1		0.9	1.1	0.7	0.4 (-)	0.4 (-)	0.4 (-)	0.1 (-)	0.3 (-)
SMc00947	*glnB*	Nitrogen regulatory protein PII	1	0.9	2 (+)	2 (+)	0.7	0.6 (-)	0.6 (-)	0.4 (-)	0.5 (-)	0.7 (-)

*Two different PCR products for *ppa* were spotted and signals were similar. Legend as for Table 1.

coding for the transcriptional regulator ChvI was slightly overexpressed during the infection of *M. truncatula* by *S. meliloti*. ChvI is part of the ChvI-ChvG(ExoS) two-component system regulating the synthesis of succinoglycan, a polysaccharide essential for infection [27]. The expression of *chvG*, close to the detection limit, appeared to be slightly induced in *bacA* and young nodules, but not in TE7 nodules.

Genes repressed in symbiosis

Eighteen of the genes tested exhibited a significantly reduced expression in nodules (Table 3). Nine of them are

Table 4

Genes specifically induced during infection

Gene ID	Gene name	Description	Cultured bacteria					Symbiotic bacteria				
								M. truncatula				*M. sativa*
			V + O_2	TY	VμO_2 weak limitation	VμO_2 strong limitation	Luteolin	J6-1021 18 dai	J6-1021 8 dai	TE7-1021	J6-bacA	Sitel-1021
SMa0034	*sma0034*	Protease	1	1.3	1.2	1.2	1.5	1.7	4.6 (+)	5.5 (+)	4.6 (+)	1.0
SMa2111	*sma2111*	Hypothetical protein hemolysin-type calcium-binding protein	1	1.1	0.9	0.7	1.1	1.0	1.8 (+)	2.1 (+)	1.7	1.4
SMc04171	*smc04171*	Hemolysin-type calcium-binding	1	0.9	1.1	0.7	1.0	0.9	2.2 (+)	2.7 (+)	3.4 (+)	0.9
SMc02661	*smc02661*	Toxin secretion ATP-binding protein	1	1.0	0.9	0.7	0.9	1.4	1.5	2.1 (+)	2.9 (+)	1.1
SMa1302	*avhB11*	Type IV secretion protein	1	1.1	0.8	0.6	1.0	1.0	2.8 (+)	2.0 (+)	3.1 (+)	1.7
SMa1306	*avhB9*	Type IV secretion protein	1	1.0	0.8	0.5	0.8	1.1	1.4 (+)	1.8 (+)	1.8 (+)	1.3
SMa1310	*avhB7*	Type IV secretion protein	1	0.9	0.6	ND	0.5	1.2	1.2	2.4 (+)	2.9 (+)	1.2
SMa1591	*sma1591*	Adenylate cyclase	1	1.1	1.1	0.7	1.0	1.5	2.6 (+)	2.6 (+)	2.2 (+)	1.4
SMa0570	*cyaF4*	Adenylate cyclase	1	0.8	1.1	0.7	0.9	1.5	2.8 (+)	2.6 (+)	3.4 (+)	2.5 (+)
SMc02176	*cyaD1*	Adenylate cyclase	1	1.1	0.9	0.8	0.9	1.3	3.1 (+)	3.7 (+)	4.0 (+)	1.3
SMc02560	*chvI*	Transcriptional regulator	1	0.6	0.9	0.6	0.9	1.1	2.1 (+)	1.4 (+)	1.5	0.7
SMc04446	*chvG exoS*	Histidine kinase sensory transmembrane protein	1	1.2	0.7	1.4	1.2	2.0	2.8 (+)	1.4	3.2 (+)	0.8
SMc00257	*ropB2*	Outer membrane protein	1	1.0	1.1	1.0	0.9	1.5	1.3	2.4 (+)	2.4 (+)	1.2
SMc02597	*sodC*	Superoxide dismutase	1	1.1	1.1	0.8	1.0	0.5	2.2 (+)	2.1 (+)	2.7 (+)	1.1
SMb20338	*smb20338*	Transmembrane efflux protein (paraquat resistance)	1	0.9	0.7	0.5	1.0	1.4	2.6 (+)	2.3 (+)	1.4	1.5
SMc02156	*smc02156*	Hypothetical protein	1	1.1	2.9 (+)	1.7 (+)	1.0	1.0	0.7	2.1 (+)	1.6 (+)	0.7

Legend as for Table 1.

related to the general cell machinery, in particular to cell structure, energy metabolism and protein synthesis. The decreased expression of *ftsZ1* is in agreement with the down-shift of FtsZ protein level observed in *Melilotus alba* nodules [28]. Cell-structure genes were generally less expressed in mature nodules (bacteroids) than in infecting bacteria, probably reflecting the drastic changes inferred by bacteroid differentiation. The repression of genes linked to energy metabolism (*ppa*, *atpE*), as well as to protein synthesis (*rpsA*, *fusA1*) may correlate with the slow growth within nodules as compared to liquid culture conditions. In infection threads, generation time was estimated at around 11 hours [29], whereas mature bacteroids do not divide. Correlation with growth rate is reinforced by the fact that these genes were also repressed under oxygen limitation where growth is slower (generation time greater than 5 hours) than in aerated cultures (generation time around 1.5 to 2 hours). However, this was not confirmed for cultures under severe oxygen limitation and complementary studies are thus needed to determine the origin of this variation in cultured cells. Two uncharacterized *S. meliloti* genes coding for proteins potentially involved in aerobic respiration (*SMc00287* and *fbcF*) were also repressed in bacteroids. As the ubiquinol cytochrome *c* reductase FbcF was induced by oxygen limitation in liquid cultures, it is most probably involved in a microoxic respiratory chain different from the known symbiotic *fixNOQP*-encoded respiratory chain. The expression of two more genes, *hemA* (heme biosynthesis) and *omp10* (outer membrane protein), was induced by oxygen limitation (4 and 2.5 times respectively) but not in symbiotic bacteria (data not shown), demonstrating that genes induced by oxygen limitation are not necessarily induced in nodules. Additional examples are *SMc02400* and *SMc02396*, which are induced by oxygen limitation, but strongly repressed in nodules.

Four genes encoding *S. meliloti* transcriptional regulators were repressed in nodules: those for the pH regulator *phrR* and its paralog *SMb21117*, *SMc00170* (*sinR*), homolog of the quorum-sensing regulator *luxR*, and *SMc00829*, homologous to the gene for PdhR, the repressor of *pdhA* in *E. coli* [30].

The other genes repressed in nodules code for the bacterioferritin Bfr, the preprotein translocase SecA, and the nitrogen-regulatory protein PII (GlnB). Interestingly, expression of the chaperonin-encoding *groES* and *groEL* seemed to decrease both in symbiosis and in oxygen-limited cultures. The *groES* and *groEL* signals varied in a similar manner, a result consistent with these genes probably belonging to the same operon [31]. The signal measured for *groES* actually corresponds to the expression of *groES1* and *groES2* as they share 99% identity. The same applies for *groEL*. Hence, because of the cross-hybridization of the two copies in our experiments, the actual regulation of each individual copy remains to be established. The widespread Sec system enables the translocation of proteins across the inner membrane. We found *secA* to be repressed during all stages of symbiosis. Hence symbiotic bacteria probably use a SecA-independent system for protein secretion. Finally, the expression of *glnB* (PII) was downregulated in nodules, as recently shown in *R. leguminosarum* [32] and in *S. meliloti*, by proteome analysis of *M. alba* nodules [28].

Gene expression in *M. sativa* nodules

The data presented above were obtained with the model plant *M. truncatula*. We have also monitored the expression of the 214 genes in wild-type (Fix^+) nodules from alfalfa (*M. sativa*). Most of the genes tested had similar expression levels in both legumes. This is nicely illustrated by the expression patterns derived from Fix^+ nodules from both plants grouping together (Figure 1a,b). Nevertheless, a few genes were differentially expressed in the two plant hosts (Table 2). The most striking difference was observed with *nifH*, whose expression was four times higher in *M. sativa* than in *M. truncatula*. This is in good agreement with the higher nitrogen-fixation efficiencies observed with alfalfa in our laboratory conditions (data not shown). Conversely, some genes were induced in *M. truncatula* but not in *M. sativa*. These code for an aquaporin (AqpZ2), a calcium-binding protein (SMb20838 and/or the duplicated SMc00286) and a protein of unknown function from the *ialA* cluster (SMc03784).

Discussion

A transcriptome approach to symbiosis

Symbiotic life with legumes implies a profound adaptation of the rhizobial life style, involving changes in expression of a large number of genes. Transcriptome technology is thus a promising tool for deciphering the rhizobium-legume symbiosis. However, apart from the pioneer work of Perret *et al.* who have followed by reverse northern blots the transcription of DNA segments from the symbiotic plasmid of *Rhizobium* NGR234 in mature nodules of two plant species [12], transcriptome analysis of rhizobia is at its beginnings. In this work, we have used DNA macroarrays to monitor expression, in a gene-specific and quantitative approach, of around 200 *S. meliloti* genes selected for their potential relevance for symbiosis.

Despite its obvious appropriateness, a transcriptome approach to symbiosis has particular difficulties. One major limitation is obviously the difficulty of comparing cultured cells to symbiotic bacteria, the physiology of which is not well described. In the present work, principal component and hierarchical cluster analyses indeed showed that expression patterns of the genes tested differed significantly between cultured and symbiotic bacteria. This might not be specific of the set of genes tested. Indeed, it was similarly reported by others that several hundred proteins of *S. meliloti* were up- or down-regulated in *M. alba* and *M. truncatula* nodules [28,33]. This problem is reinforced by the paucity of genuinely constitutive bacterial genes that could be used for calibration purposes. We addressed these limitations in two ways: first, 34 genes of known regulation were included as internal controls; and second, a statistical evaluation of the data was conducted (see Materials and methods) and only the most solid data ($p < 0.01$) have been considered here. In addition, in several instances, variations in expression were substantiated by circumstantial evidence, for example, similar expression patterns for different genes belonging to the same operon or a similar pattern of expression for a gene in closely related free-living or symbiotic conditions.

Medicago species produce nodules that display an indeterminate developmental program with the consequence that all stages of bacterial development coexist within a single mature nodule. We postulated that bacterial and plant developmental mutants could be used to dissect the interaction further and focus on putative infection genes. Cytological studies have indeed shown that *bacA* and TE7 mutant nodules contain bacteria at the infection stage but lack mature bacteroids [8,15] The pattern of expression of a few specific marker genes (*nifH*, *nodA*, *nodC* and *katA*) was consistent with this observation. As infection in mutant backgrounds may differ from infection in a wild-type context - in particular as it was recently shown that a *bacA* mutant is pleiotropic [7] - we have only considered here as putative infection genes those that displayed enhanced expression under the three conditions tested (*bacA* and TE7 mutant nodules and young nodules). Genes whose expression was enhanced under only one or two of the conditions tested were not considered further. Altogether, these elements strongly suggest that the genes induced in 1021-TE7 and *bacA*-J6 nodules as well as in young nodules (Table 4) correspond to genes expressed during the infection stage, thus providing insight into this yet poorly understood stage of symbiosis.

Common molecular mechanisms of symbiosis and pathogenesis

Increasing evidence suggests that there are mechanisms in common in the development of rhizobium-legume symbiosis and the pathogenic interactions between proteobacteria and eukaryotic cells [34]. Analogy between the endocytic process of infecting mammalian pathogens and rhizobia was nicely illustrated by the finding of a *bacA* homolog in *Brucella* that is required for macrophage invasion [35]. In addition to *bacA*, genes homologous to those of an invasion associated locus of *Bartonella bacilliformis* [23] and *Brucella suis* were identified in the *S. meliloti* genome. Erythrocyte invasion by *Bartonella* requires three virulence determinants: the deformin protein - not present in *S. meliloti* - which induces invagination of the red blood cells; the flagella; and the *ialAB* locus. IalA is an NTPase activated within the erythrocyte and probably inducing the production of energy from Mg-ATP and Ca^{2+} efflux. The finding that *ialA* expression is induced by luteolin and also in symbiotic *S. meliloti* reinforces its potential role as an invasion protein. Significantly, all the genes from the *ialA* cluster (*SMc03782* through *SMc03785*) were induced in *M. truncatula* nodules.

Intracellular pathogens and symbiotic bacteria often use specialized type-III or type-IV secretion systems to deliver proteins into host cells. In contrast to several other rhizobia, *S. meliloti* does not have any type-III secretion system. pSymA sequencing revealed a potential *virB*-encoded type-IV secretion system [36]. We found the corresponding genes to be preferentially synthesized in infecting bacteria. However, in the light of recent work [37], we favor the view that the *virB*-like genes of *S. meliloti* are actually orthologs of the *Agrobacterium avhB* genes involved in plasmid conjugal transfer rather than genuine virulence or infection genes. Accordingly, *S. meliloti* strains containing deletions of the *virB*/*avhB* operon appeared normal for nodulation and nitrogen fixation (D. Wells, unpublished work cited in [36]; E.K. and J.B., unpublished results).

Oxidative stress

To combat invading bacteria, eukaryotic hosts go into an iron-withholding mode [38], change environmental conditions such as pH and generate reactive oxygen species (ROS). Alfalfa responds to infection by *S. meliloti* by the production of superoxide ions and hydrogen peroxide [39]. To overcome these plant defenses, *S. meliloti* possesses three catalases (KatA, KatB and KatC) and two superoxide dismutases (SodA and SodC). Although a *katA* mutant is not impaired in either nodulation or nitrogen fixation [40], the KatA protein is induced in 5-week old nodules of *M. sativa* [26] as a result of an increased transcription (the present work). Both catalases contribute to the degradation of hydrogen peroxide *in planta* as only a *katA-katC* double mutant shows decreased nitrogen fixation [26]. In addition to catalases, superoxide dismutases (SOD) were shown to be important for symbiosis, as a *sodA* mutant nodulated poorly and displayed abnormal infection [4]. We observed that the other *S. meliloti* SOD (SodC) is specifically induced during infection. Although widely found in eukaryotes, this CuZnSOD is rare in bacteria. SodC from *S. meliloti* has a signal peptide and is probably secreted in the periplasm as is *E. coli* SodC [41]. Because superoxide anions cannot cross membranes, the SodC enzyme probably protects the bacterial cell wall against superoxide ions of external origin. Accordingly, alfalfa was shown to produce O_2^- in infection threads and infected cells in up to 9-day-old nodules [39]. Therefore, we speculate that SodC might detoxify plant ROS, whereas the cytoplasmic SodA would be dedicated to the detoxification of ROS synthesized by the bacteria. Finally, SMb20338, a transmembrane efflux protein sharing 77% identity with PqrA from *O. anthropi* conferring resistance to paraquat [42] is expressed during infection and could also participate in oxidative stress responses.

Iron metabolism

In plants, free iron is scarce, and bacteria should acquire it from their host, as do many pathogenic bacteria. Rhizobia possess several systems for iron acquisition. In *R. leguminosarum* and *Bradyrhizobium japonicum*, *tonB*-like and heme-uptake *hmu* mutants are not affected in nodulation and nitrogen fixation [22,43]. Similarly, *S. meliloti* mutants unable to synthesize the siderophore rhizobactin induce effective nitrogen-fixing nodules [44]. However, it is still not known how symbiotic bacteria acquire their iron [45]. The complete sequencing of the *S. meliloti* genome has revealed the presence of a new system allowing the utilization of chelated iron, the Sit system. Our results have shown that *sitA* is induced by luteolin and in symbiotic bacteria. SitA is homologous to proteins from *Yersinia pestis* and *Salmonella typhimurium* that contribute to enteropathogenicity [46,47]. In *S. meliloti*, the gene coding for the general regulator Fur is sitting next to *sitA*. The Fur protein is a common regulator for most genes regulated in response to iron shortage in pathogenic bacteria [38]. Conversely, under high-iron conditions specific genes are activated, including genes for the biosynthesis of bacterioferritin. Our results showing that the Sit system is induced during symbiosis while bacterioferritin (*bfr*) synthesis is repressed, suit very well the model described for pathogens and suggest that the SitABC-transporter is an important system for iron acquisition *in planta*. In addition, the probable siderophore encoded by *SMa2339* - a homolog of *iucB* coding for the siderophore aerobactin in *E. coli* and *Shigella flexneri* - was also induced by the plant flavonoid luteolin. Hence, several systems may cooperate for iron mobilization *in planta*.

Signaling in the rhizobium-legume symbiosis

Calcium-binding proteins

The transcriptome analysis carried out here identified a restricted list of proteins with calcium-binding motifs specifically induced during the establishment of symbiosis. The biochemical role of these proteins (SMa0034, SMa2111,

SMb20838/SMc00286 and SMc04171) cannot be determined on the basis of sequence homology, although SMa0034 is likely to be a protease. Interestingly, SMa2111 and SMc04171 appear to be paralogs. Calcium has a role in a large number of cellular processes, and may be of particular importance in the establishment of symbiosis as calcium spiking is one of the earliest responses of plant root hairs to Nod factors [48]. All together, these data point to a possible role of calcium in controlling infection by rhizobia.

Nucleotide cyclases: key regulators of symbiosis?

The cyclic nucleotides cAMP and cGMP are key universal second messengers that participate in a large number of signal transduction pathways [49]. So far, *S. meliloti* encodes the most nucleotide cyclases (26) of any bacterial genome [13]. These nucleotide cyclases belong to two distinct classes [24]: the widespread class III and a new class - the *cyaF* class - so far unique to *S. meliloti* and *M. loti* (it was not found in the genome of the closely related plant pathogen *Agrobacterium tumefaciens*). So far, only three of the 26 *S. meliloti* cyclases have been studied. Mutants in the class III cyclases *cyaA* and *cyaB* (formerly *cya1* and *cya2*) did not show any alteration in their symbiotic phenotypes [50]. Conversely, a mutation in *cyaF5* (formerly *cya3*) significantly increased alfalfa shoot dry weight, thus enhancing symbiotic abilities [51]. In this work, three other cyclases belonging to the *cyaF* class were significantly upregulated during symbiosis. The role of these cyclases remains to be elucidated. Cyclic nucleotides in prokaryotes, particularly cAMP, are mostly known as regulators of catabolic functions. In other cases they are involved in communication with host cells. The most fascinating examples may be found in the interaction between the human pathogens *B. pertussis* and *Bacillus anthracis* with their target cells. In both bacteria, secreted adenylate cyclases are translocated into the target cells where they have a key role in pathogenic conversion [52,53]. The accumulating data on bacterial genome sequences show that adenylate cyclases with similarity to eukaryotic enzymes are not exceptional, as found for *Mycobacterium tuberculosis* [49], but their role in signal transduction in such complex interactions as those involving rhizobia and leguminous plants remains to be investigated.

Materials and methods

Microbiological techniques

The two *S. meliloti* strains used in this study are 1021, the reference strain whose genome sequence is available [54], and VO2119 [17], a *bacAΔ::aadA* mutant of 1021 (Sm^R Sp^R) kindly provided by S. Long and V. Oke. Cultured bacteria were grown either in TY rich medium or in Vincent minimal medium (Vmini) [55] in 1 liter bioreactors (Setric Toulouse, France) at 30°C and pH 6.8. Oxic cultures were grown with an oxygen pressure (pO_2) always greater than 60% of saturation by air. In microoxic cultures, pO_2 was initially kept high (> 60% as in the aerated cultures) until an OD of 0.2. The reactors were then rapidly flushed with N_2 and the pO_2 was maintained below the detection limit of the electrodes (apparent pO_2 = 0, that is less than 1 μM dissolved oxygen) although a constant input of oxygen was maintained to ensure bacterial growth and energy metabolism. This was achieved through a tight control of agitation - reduced to around 200 rpm - and aeration. Two airflows were tested: 3 volume air/volume liquid/minute (vvm) and 1 vvm denominated as 'oxygen limitation' and 'strong limitation', respectively, in the text. At an OD of 0.5, 25 ml of culture was sampled and bacterial cells were rapidly separated from supernatant by filtration (0.2 μm) and immediately frozen in liquid nitrogen until further RNA extraction. This complete operation did not exceed 15 sec. Induction by luteolin (12 h) was carried out by the addition of 10 μM luteolin to bacteria growing in Vmini medium [56]. Three independent repetitions of each liquid culture were performed.

Plant methods

Seeds of *M. truncatula* Jemalong J6 (wild type) and TE7 (Nod^+, Inf^+, Fix^-) [15], as well as *M. sativa cv* Sitel were sterilized and grown in aeroponic tanks as previously described [57]. Plants were first grown in a medium supplemented with 5 mM ammonium nitrate. After 10 days of growth in the presence of nitrogen, the medium was changed to nitrogen-free medium. Plants were inoculated after two more days with *S. meliloti* to promote nodulation. Wild-type nodules were harvested 8 days (young nodules) and 18 days (mature nodules) after inoculation, whereas Fix^- nodules were collected 11 days after inoculation as a longer period would lead to degeneracy. One hundred mature wild-type plants yielded approximately 10 g of nodules, whereas this yield was around 2-5 g for young and Fix^- plants. Nodules were frozen in liquid nitrogen until further RNA extraction. Three independent repetitions of each symbiotic condition were performed.

RNA isolation

Total RNA from cultured bacteria was isolated as previously described [16]. For symbiotic bacteria, approximately 1 g of frozen nodules were crushed, and RNA was directly extracted from the powder by a protocol similar to that used for liquid cultures. In order to limit artifacts due to differences in RNA extractions of a given sample, four independent RNA preparations for each of the three repetitions of each biological condition were pooled [58]. RNA preparations from symbiotic bacteria are likely to be contaminated by plant RNA. To evaluate their influence on hybridization results, plant mRNA was extracted from mature *M. truncatula* nodules: total RNA was first extracted as described by Jackson and Larkins [59] and mRNAs were further purified using Dyna beads from Dynal Biotech (Compiègne, France) as described by the manufacturer. When membranes were hybridized with control 'plant' probes generated with 200 ng of purified poly(A)$^+$ RNA using the protocol described below, no cross-hybridization between plant RNA and the spotted

PCR products could be detected, and the influence of plant RNA on hybridization signals with symbiotic bacterial RNAs was then considered to be negligible.

Probe synthesis

Labeled cDNA was synthesized by random priming in a protocol adapted from that described by Tao *et al*. [60] except that random primers were preferred to specific primers for each gene as it was shown that results yielded by the former are less biased [58]. Ten micrograms of total RNA, 250 ng of random hexamer and nonamer primers (New England Biolabs) and 1 ng of five different *in vitro* transcribed mRNAs from *Corynebacterium striatum* not present in *S. meliloti* were heated for 10 min at 70°C and quickly cooled on ice. cDNA synthesis was performed at 42°C for 3 h in a 50 μl reaction mixture containing the RNA and primer mixture, reverse transcriptase buffer (Gibco BRL), 0.5 mM each dATP, dGTP and dTTP, 100 μCi [α^{33}P]dCTP (NEN), 60 units of ribonuclease inhibitor (Promega), 10 mM dithiothreitol, and 200 units Superscript II (GibcoBRL). The template RNA was then degraded with 50 μl (100 mM EDTA 100 mM, 0.8% SDS, 0.72 M NaOH) and incubated for 30 min at 68°C. After neutralization with 50 μl 1 M TrisCl pH 7.4/0.6 M HCl, the labeled cDNAs were purified with MicroSpin S-200 columns (Amersham-Pharmacia Biotech, Little Chalfont, UK). The quality of several cDNA probes was checked on polyacrylamide gels and the large majority of labeled cDNAs ranged from 300 base-pairs (bp) to 1.4 kilobases (kb) (data not shown).

PCR amplification and DNA array construction

Specific 20-mer primers were designed for each of the 214 genes using Primer Select (DNAStar, Madison, WI). Melting temperatures of the primers and of the amplified products were set constant (58-62°C and 80-85°C, respectively). PCR products of constant size (280-350 bp) were generated for each gene. The cycling conditions (30 cycles) were as follows: 30 sec of denaturing at 94°C, 1 min of annealing at 55°C and 1 min of elongation at 72°C. The quality of the PCRs was checked by electrophoresis on agarose (2%) gels. Sequencing of 40 PCR fragments provided another quality control. With one exception, all the sequences generated corresponded to the expected gene fragment. All information relating to problematic reactions was used to redesign oligonucleotides. Of the primers, 89% gave satisfying results after the first run. The design of new primers and the optimization of PCR conditions increased the percentage of successful PCRs to 95%. The primary PCRs were eluted from agarose gels to get rid of genomic DNA. One microliter of a 1/100 dilution was then used as template for the secondary PCR and 300 μl of secondary PCR were precipitated with isopropanol and resuspended in 50% DMSO to a final concentration of around 400 ng/μl. These were spotted on Immobilon-Ny+ membranes (Millipore, Bedford, MA) using the Eurogridder robot (Eurogentec, Seraing, Belgium) available at the Toulouse Génopôle [61]. Each spot consisted of 100 nl of concentrated PCR, and each PCR product was spotted in duplicate. To assess the reproducibility of the spotting process, test membranes with 48 identical spots were first generated. When hybridized with a specific probe, the coefficient of variation of spot intensities was 3.6%. A similar result was obtained when identical PCR products were spotted on independent membranes in the same batch. All the membranes with the approximately 200 *S. meliloti* genes used in this study were then produced during a single batch and variability between membranes was considered to be limited. Five external standards corresponding to PCR products of *C. striatum* genes that have no homolog in *S. meliloti* were included in the membranes and used to check the quality of all hybridizations. Plasmids containing these genes were kindly provided by A. Becker.

DNA array hybridization

The nylon filters were prehybridized in 10 ml hybridization solution (5x SSC, 5x Denhardt's solution, 100 μg/ml sheared salmon sperm DNA, 0.5% SDS) for 2 hours at 65°C. The cDNA probe was heated for 5 min at 95°C, rapidly cooled on ice, and added to 5 ml hybridization solution. The prehybridization solution was removed and replaced with the hybridization solution. Hybridization was carried out for 17 hours at 65°C. Following hybridization, each filter was rinsed with 2x SSC containing 0.5% SDS for 5 min at 65°C, followed by two washes with 0.2x SSC containing 0.5% SDS at 65°C for 20 min each. The filters were partially air-dried, wrapped in Saran, and exposed to a phosphor screen for 72 hours.

Data analysis

Hybridization signals were quantified and analyzed with ImaGene and Genesight (BioDiscovery, Marina del Rey, CA). The experimental design used in this study consisted of three independent ^{33}P-labeled cDNA preparations derived from the three independent repetitions of each biological condition separately hybridized to filters containing PCRs spotted in duplicate. For each spot, a background-subtracted estimate of expression level (median) was obtained and scaled to total counts on the membrane. This normalization procedure was validated *a posteriori* by the fact that a large proportion (> 75%) of the genes tested did not show significant differential expression, although they were selected for their potential implication in symbiosis. Another line of control was provided by the internal standards derived from *C. striatum* genes. Relative counts of *S. meliloti* material with respect to these controls were always comparable from one membrane to another. We then used Student's *t*-test to compare means of standardized expressions between growth conditions as described by Arfin *et al*. [58]. Differences were considered to be significant for a p-value < 0.01, and genes were then designated as 'induced' or 'repressed' as compared to the reference condition. Hierarchical clustering [62] was applied on log-transformed data using Euclidean distances and the average cluster linkage algorithm, and principal

component analysis was performed as principal gene analysis as implemented in the Genesight software.

Duplicated genes and cross-hybridization

In the washing conditions used, cross-hybridization is likely to occur only in the case of duplicated genes, that is if the percent identity at the DNA level is greater than 80%. Therefore, a few of the PCR products spotted here may hybridize with cDNAs from several duplicated genes because of sequence similarities, and the signals measured for these particular genes are the combination of the expression of two or sometimes three duplicated genes. In our study, the expression of the following genes cannot be separated: *nodD1*, *nodD2* and *nodD3*; *groEL1* and *groEL2*; *groES1* and *groES2*; *traA1* and *traA2*; *fixK1* and *fixK2*; *fixN1* and *fixN2*; *fixT1*, *fixT2* and *fixT3*; *SMc00286* and *SMb20838*.

Additional data files

A list of the genes tested and the comprehensive expression results are available as an additional data file with the online version of this article and from [14].

Acknowledgements

During this project, F.A. was supported by the Institut de Recherche pour le Développement (IRD, Paris, France) and E.K. by an INRA and a Marie-Curie postdoctoral fellowship. We thank A-M. Garnerone for help with transcriptome experiments, P. Gamas's group for plant mRNA preparations, A. Becker for providing the *C. striatum* control genes, D. Capela for her contribution to *S. meliloti* genome data mining, V. Guillou and H. Feau for technical assistance with the bioreactors, the staff of the 'Plate-forme Génomique' at the Toulouse Génopôle for assistance in macroarray construction, T. Huguet for the TE7 seeds, and the 'JB' team for helpful discussions and critical reading of the manuscript. This work was supported by the Toulouse Génopôle, the department Santé des Plantes et Environnement of INRA (Grant 0441-04) and by the CNRS program Puces à ADN.

References

1. Denarié J, Debellé F, Promé JC: ***Rhizobium*** **lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis.** *Annu Rev Biochem* 1996, **65:**503-535.
2. Catoira R, Galera C, de Billy F, Penmetsa RV, Journet EP, Maillet F, Rosenberg C, Cook D, Gough C, Denarie J: **Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway.** *Plant Cell* 2000, **12:**1647-1666.
3. Pellock BJ, Cheng HP, Walker GC: **Alfalfa root nodule invasion efficiency is dependent on *Sinorhizobium meliloti* polysaccharides.** *J Bacteriol* 2000, **182:**4310-4318.
4. Santos R, Herouart D, Puppo A, Touati D: **Critical protective role of bacterial superoxide dismutase in *Rhizobium*-legume symbiosis.** *Mol Microbiol* 2000, **38:**750-759.
5. Vasse J, de Billy F, Camut S, Truchet G: **Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules.** *J Bacteriol* 1990, **172:**4295-4306.
6. Brewin NJ: **Tissue and cell invasion by *Rhizobium*: the structure and development of infection threads and symbiosomes.** In *The Rhizobiaceae*. Edited by Spaink HP, Kondorosi A, Hooykaas PJJ. Dordrecht: Kluwer Academic Publishers; 1998: 417-429.
7. Ferguson GP, Roop II RM, Walker GC: **Deficiency of a *Sinorhizobium meliloti bacA* mutant in alfalfa symbiosis correlates with alteration of the cell envelope.** *J Bacteriol* 2002, **184:**5625-5632.
8. Glazebrook J, Ichige A, Walker GC: **A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development.** *Genes Dev* 1993, **7:**1485-1497.
9. Watson RJ, Chan YK, Wheatcroft R, Yang AF, Han SH: ***Rhizobium meliloti*** **genes required for C4-dicarboxylate transport and symbiotic nitrogen fixation are located on a megaplasmid.** *J Bacteriol* 1988, **170:**927-934.
10. Kaminski PA, Batut J, Boistard P: **A survey of nitrogen fixation by rhizobia.** In *The Rhizobiaceae*. Edited by Spaink HP, Kondorosi A, Hooykaas PJJ. Dordrecht: Kluwer Academic Publishers; 1998: 431-460.
11. Soupène E, Foussard M, Boistard P, Truchet G, Batut J: **Oxygen as a key developmental regulator of *Rhizobium meliloti* N2-fixation gene expression within the alfalfa root nodule.** *Proc Natl Acad Sci USA* 1995, **92:**3759-3763.
12. Perret X, Freiberg C, Rosenthal A, Broughton WJ, Fellay R: **High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234.** *Mol Microbiol* 1999, **32:**415-425.
13. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, *et al.*: **The composite genome of the legume symbiont *Sinorhizobium meliloti*.** *Science* 2001, **293:**668-672.
14. **Infection macroarray** [http://sequence.toulouse.inra.fr/rhime/Expression/Infection/Infection_macro-array_2002.htm]
15. Bénaben V, Duc G, Lefebvre V, Huguet T: **TE7, an inefficient mutant of *Medicago truncatula* Gaertn. cv Jemalong.** *Plant Physiol* 1995, **107:**53-62.
16. Cabanes D, Boistard P, Batut J: **Symbiotic induction of pyruvate dehydrogenase genes from *Sinorhizobium meliloti*.** *Mol Plant-Microbe Interact* 2000, **13:**483-493.
17. Oke V, Long SR: **Bacterial genes induced within the nodule during the *Rhizobium*-legume symbiosis.** *Mol Microbiol* 1999, **32:**837-849.
18. Davey ME, de Bruijn FJ: **A homologue of the tryptophan-rich sensory protein TspO and FixL regulate a novel nutrient deprivation-induced *Sinorhizobium meliloti* locus.** *Appl Environ Microbiol* 2000, **66:**5353-5359.
19. Schlaman HR, Horvath B, Vijgenboom E, Okker RJ, Lugtenberg BJ: **Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum* biovar *viciae*.** *J Bacteriol* 1991, **173:**4277-4287.
20. Timmers AC, Auriac MC, de Billy F, Truchet G: **Nod factor internalization and microtubular cytoskeleton changes occur concomitantly during nodule differentiation in alfalfa.** *Development* 1998, **125:**339-349.
21. Chen H, Higgins J, Oresnik IJ, Hynes MF, Natera S, Djordjevic MA, Weinman JJ, Rolfe BG: **Proteome analysis demonstrates complex replicon and luteolin interactions in pSymA-cured derivatives of *Sinorhizobium meliloti* strain 2011.** *Electrophoresis* 2000, **21:**3833-3842.
22. Wexler M, Yeoman KH, Stevens JB, de Luca NG, Sawers G, Johnston AW: **The *Rhizobium leguminosarum tonB* gene is required for the uptake of siderophore and haem as sources of iron.** *Mol Microbiol* 2001, **41:**801-816.
23. Minnick MF, Mitchell SJ, McAllister SJ: **Cell entry and the pathogenesis of *Bartonella* infections.** *Trends Microbiol* 1996, **4:**343-347.
24. Capela D, Barloy-Hubler F, Gouzy J, Bothe G, Ampe F, Batut J, Boistard P, Becker A, Boutry M, Cadieu E, *et al.*: **Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti*.** *Proc Natl Acad Sci USA* 2001, **98:**9877-9882.
25. Katz ME, Wright CL, Gartside TS, Cheetham BF, Doidge CV, Moses EK, Rood JI: **Genetic organization of the duplicated *vap* region of the *Dichelobacter nodosus* genome.** *J Bacteriol* 1994, **176:**2663-2669.
26. Sigaud S, Becquet V, Frendo P, Puppo A, Herouart D: **Differential regulation of two divergent *Sinorhizobium meliloti* genes for HPII-like catalases during free-living growth and protective role of both catalases during symbiosis.** *J Bacteriol* 1999, **181:**2634-2639.
27. Cheng HP, Walker GC: **Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system.** *J Bacteriol* 1998, **180:**20-26.
28. Natera SH, Guerreiro N, Djordjevic MA: **Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis.** *Mol Plant Microbe Interact* 2000, **13:**995-1009.

29. Gage DJ, Bobo T, Long SR: **Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*).** *J Bacteriol* 1996, **178:**7159-7166.
30. Quail MA, Guest JR: **Purification characterization and mode of action of PdhR, the transcriptional repressor of the *pdhR-aceEF-lpd* operon of *Escherichia coli*.** *Mol Microbiol* 1995, **15:**519-529.
31. Ogawa J, Long SR: **The *Rhizobium meliloti groELc* locus is required for regulation of early *nod* genes by the transcription activator NodD.** *Genes Dev* 1995, **9:**714-729.
32. Ercolano E, Mirabella R, Merrick M, Chiurazzi M: **The *Rhizobium leguminosarum glnB* gene is down-regulated during symbiosis.** *Mol Gen Genet* 2001, **264:**555-564.
33. Djordjevic MA, Natera SH, Chen HC, Weiller G, Menzel C, van Noorden G, Weinman J, Taylor S, Guo J, Rolfe BG: **Proteome analysis of *Sinorhizobium meliloti*.** In *Nitrogen Fixation, Global Perspective 2002*. Edited by Finan TM, O'Brian MR, Layzell DB, Vessey JK, Newton W. Oxford: CABI; 2002: 50-54.
34. Hentschel U, Steinert M, Hacker J: **Common molecular mechanisms of symbiosis and pathogenesis.** *Trends Microbiol* 2000, **8:**226-231.
35. LeVier K, Phillips RW, Grippe VK, Roop RM, Walker GC: **Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival.** *Science* 2000, **287:**2492-2493.
36. Barnett MJ, Fisher RF, Jones T, Komp C, Abola AP, Barloy-Hubler F, Bowser L, Capela D, Galibert F, Gouzy J, *et al.*: **Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid.** *Proc Natl Acad Sci USA* 2001, **98:**9883-9888.
37. Chen L, Chen Y, Wodd DW, Nester EW: **A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*.** *J Bacteriol* 2002, **184:**4838-4845.
38. Ratledge C, Dover LG: **Iron metabolism in pathogenic bacteria.** *Annu Rev Microbiol* 2000, **54:**881-941.
39. Santos R, Herouart D, Sigaud S, Touati D, Puppo A: **Oxidative burst in alfalfa-*Sinorhizobium meliloti* symbiotic interaction.** *Mol Plant Microbe Interact* 2001, **14:**86-89.
40. Hérouart D, Sigaud S, Moreau S, Frendo P, Touati D, Puppo A: **Cloning and characterization of the *katA* gene of *Rhizobium meliloti* encoding a hydrogen peroxide-inducible catalase.** *J Bacteriol* 1996, **178:**6802-6809.
41. Imlay KR, Imlay JA: **Cloning and analysis of *sodC* encoding the copper-zinc superoxide dismutase of *Escherichia coli*.** *J Bacteriol* 1996, **178:**2564-2571.
42. Won SH, Lee BH, Lee HS, Jo J: **An *Ochrobactrum anthropi* gene conferring paraquat resistance to the heterologous host *Escherichia coli*.** *Biochem Biophys Res Commun* 2001, **285:**885-890.
43. Nienaber A, Hennecke H, Fischer HM: **Discovery of a haem uptake system in the soil bacterium *Bradyrhizobium japonicum*.** *Mol Microbiol* 2001, **41:**787-800.
44. Lynch D, O'Brien J, Welch T, Clarke P, Cuiv PO, Crosa JH, O'Connell M: **Genetic organization of the region encoding regulation biosynthesis and transport of rhizobactin 1021, a siderophore produced by *Sinorhizobium meliloti*.** *J Bacteriol* 2001, **183:**2576-2585.
45. Johnston AW, Yeoman KH, Wexler M: **Metals and the rhizobial-legume symbiosis - uptake utilization and signalling.** *Adv Microb Physiol* 2001, **45:**113-156.
46. Bearden SW, Staggs TM, Perry RD: **An ABC transporter system of *Yersinia pestis* allows utilization of chelated iron by *Escherichia coli* SAB11.** *J Bacteriol* 1998, **180:**1135-1147.
47. Zhou D, Hardt WD, Galan JE: ***Salmonella typhimurium* encodes a putative iron transport system within the centisome 63 pathogenicity island.** *Infect Immun* 1999, **67:**1974-1981.
48. Ehrhardt DW, Wais R, Long SR: **Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals.** *Cell* 1996, **85:**673-681.
49. McCue LA, McDonough KA, Lawrence CE: **Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*.** *Genome Res* 2000, **10:**204-219.
50. Archdeacon J, Talty J, Boesten B, Danchin A, O'Gara F: **Cloning of the second adenylate cyclase gene (*cya2*) from *Rhizobium meliloti* F34: sequence similarity to eukaryotic cyclases.** *FEMS Microbiol Lett* 1995, **128:**177-184.
51. Sharypova LA, Yurgel SN, Keller M, Simarov BV, Puhler A, Becker A: **The eff-482 locus of *Sinorhizobium meliloti* CXM1-105 that influences symbiotic effectiveness consists of three genes encoding an endoglycanase, a transcriptional regulator and an adenylate cyclase.** *Mol Gen Genet* 1999, **261:**1032-1044.
52. Mock M, Ullmann A: **Calmodulin-activated bacterial adenylate cyclases as virulence factors.** *Trends Microbiol* 1993, **1:**187-192.
53. Ladant D, Ullmann A: ***Bordetella pertussis* adenylate cyclase: a toxin with multiple talents.** *Trends Microbiol* 1999, **7:**172-176.
54. ***Sinorhizobium meliloti* Genome Project** [http://sequence.toulouse.inra.fr/meliloti.html]
55. Vincent JM: *A Manual for the Practical Study of Root Nodule Bacteria.* Oxford: Blackwell; 1970.
56. Maillet F, Debellé F, Dénarié J: **Role of the *nodD* and *syrM* genes in the activation of the regulatory gene *nodD3* and of the common and host-specific *nod* genes of *Rhizobium meliloti*.** *Mol Microbiol* 1990, **4:**1975-1984.
57. Gallusci P, Dedieu A, Journet EP, Huguet T, Barker DG: **Synchronous expression of leghaemoglobin genes in *Medicago truncatula* during nitrogen-fixing root nodule development and response to exogenously supplied nitrate.** *Plant Mol Biol* 1991, **17:**335-349.
58. Arfin SM, Long AD, Ito ET, Tolleri L, Riehle MM, Paegle ES, Hatfield GW: **Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor.** *J Biol Chem* 2000, **275:**29672-29684.
59. Jackson AD, Larkins BA: **Influence of ionic strength pH and chelation of divalent metals on isolation of polyribosomes from tobacco leaves.** *Plant Physiol* 1976, **57:**5-10.
60. Tao H, Bausch C, Richmond C, Blattner FR, Conway T: **Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media.** *J Bacteriol* 1999, **181:**6425-6440.
61. **Génopôle Toulouse** [http://genopole.toulouse.inra.fr]
62. Alizadeh AA, Eisen MB, Davis RE, Lossos IS, Rosenwald A, Boldrick, JC, Sabet, H, Tran T, Yu X *et al.*: **Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling**. *Nature* 2000, **403:**503-511.