Novel Aspects of Symbiotic Nitrogen Fixation Uncovered by Transcript Profiling with cDNA Arrays

Gillian Colebatch,¹ Sebastian Kloska,² Ben Trevaskis,¹ Susanne Freund,¹ Thomas Altmann,¹ and Michael K. Udvardi¹

¹Max Planck Institute for Molecular Plant Physiology, Am Mühlenberg 1, 14476 Golm, Germany; ²Scienion AG, Volmerstr. 7a, 12489 Berlin, Germany

Submitted 11 February 2002. Accepted 14 February 2002.

An array of 2,304 cDNA clones derived from nitrogen-fixing nodules of *Lotus japonicus* was produced and used to detect differences in relative gene transcript abundance between nodules and uninfected roots. Transcripts of 83 different genes were found to be more abundant in nodules than in roots. More than 50 of these have never before been identified as nodule-induced in any species. Expression of 36 genes was detected in nodules but not in roots. Several known nodulin genes were included among the nodule-induced genes. Also included were genes involved in sucrose breakdown and glycolysis, CO₂ recycling, and amino acid synthesis, processes that are known to be accelerated in nodules compared with roots. Genes involved in membrane transport, hormone metabolism, cell wall and protein synthesis, and signal transduction and regulation of transcription were also induced in nodules. Genes that may subvert normal plant defense responses, including two encoding enzymes involved in detoxification of active oxygen species and one that may prohibit phytoalexin synthesis, were also identified. The data represent a rich source of information for hypothesis building and future exploration of symbiotic nitrogen fixation.

Symbiotic nitrogen fixation (SNF) by the legume-rhizobium couple is a process of both ecological and agricultural importance. The establishment of an effective symbiosis between legumes and rhizobia is a complex process, involving signaling and recognition by both partners from the outset. Rhizobial attachment to root hairs, penetration of the epidermis, and invasion of cortical tissue via the infection thread is accompanied by initiation of meristematic activity in root cortical and pericycle cells and by suppression of plant defense responses. Release of rhizobia from infection threads into individual cortical cells is achieved via endocytosis, which leaves the rhizobia enclosed in a plant membrane called the peribacteroid or symbiosome membrane that isolates them from the host cell cytoplasm. The resulting 'organelle' is called a symbiosome. Rhizobia continue to divide until infected cells are packed with thousands of bacteria, which are surrounded, either individually or in small groups, by the symbiosome membrane. Other profound changes occur during nodule development, including the establishment of a vascular network that delivers photosynthate, mostly in the form of sucrose, to

the nodule tissues and that ultimately facilitates the export of nitrogen-containing compounds from the nitrogen-fixing nodule. Before nitrogen fixation can take place, however, a microaerobic environment is established inside nodules, helping to trigger differentiation of rhizobia into the nitrogen-fixing bacteroid form. Over the past 20 years, significant insights into many aspects of SNF have been gained: critical early signaling events have been uncovered, numerous bacterial genes essential for nodule development or function have been identified, and important aspects of bacteroid and legume nodule metabolism have been elucidated (Day and Copeland 1991, Denarie et al. 1996, Downie and Walker 1999, Kahn et al. 1998, Long 2001, Udvardi and Day 1997, Vance et al. 1994). On the other hand, although many plant genes are expected to play important or even essential roles in nodule development and function, very few of these have been identified. Plant geneticists have discovered at least 40 genetic loci that are essential for different aspects of SNF in different legumes species (Caetano-Anolles and Gresshoff 1991) and, although many of these have been mapped, only one, *Nin*, which is essential for nodule development (initiation), has been isolated to date (Schauser et al. 1999, Stougaard 2000).

Although many thousands of plant genes are expressed in developing and mature nodules, it is likely that relatively few of these are absolutely essential for SNF. One way to identify plant genes that may play critical roles during SNF is to identify those that are induced during nodule development. Perhaps the most interesting of these are the so called nodulin genes, which are expressed only in developing or mature nodules. Different approaches have been taken to identify nodulin genes, including cold-plaque screening (Jiménez-Zurdo et al. 2000), differential display (Szczygłowski et al. 1997), subtractive hybridization (Kouchi and Hata 1993, Gamas et al. 1996), and differential hybridization (Perlick and Puhler 1993).

Nodulins have been subdivided into two major classes, early and late, based on when the gene transcripts are first detected (Bladergroen and Spaink 1998, Crespi and Galvez 2000, Pawlowski 1997, Verma et al. 1992). Early nodulin genes are activated in roots within hours or a few days of inoculation with rhizobia and are suspected to be involved in nodule development. Late nodulin genes are activated in mature nodules around the time that nitrogen fixation commences. Some of these encode enzymes involved in primary carbon and nitrogen metabolism. Others may help maintain the structural integrity of nodules, including the compartmentalization of bacteroids in symbiosomes.

In the past few years, very large numbers of expressed sequence tags (ESTs) (sequences obtained from the 5' or 3' ends of isolated cDNA clones) from legumes have been deposited in public databases. For instance, at the time of writing, there were 218,000 ESTs from soybean, 137,000 ESTs from *Medi-*

Corresponding author: M. K. Udvardi; Telephone: +49 331 567 8149; Fax: +49 331 567 8250; E-mail: Udvardi@mpimp-golm.mpg.de

*The Spotlight logo represents articles that, in the opinion of the senior editor and editor-in-chief, are of special interest to a broad readership.

cago truncatula, and 32,000 ESTs from *Lotus japonicus* in GenBank. ESTs and the corresponding cDNA clones are a valuable resource for gene discovery and analysis. One way in which EST clones are being put to use is in the production of DNA arrays for transcriptome analysis (Kehoe et al. 1999, Richmond and Somerville 2000). We have obtained over 5,000 ESTs from a *L. japonicus* nodule cDNA library and have constructed an array from the first 2,304 of these for expression analysis. Here, we describe how the *L. japonicus* cDNA array was used to discover novel nodule-induced genes, which are considered in the context of current models of symbiotic nitrogen fixation. The advantages and disadvantages of DNA arrays over other technologies for transcript profiling are also discussed.

RESULTS

A total of 5,495 clones from two *L. japonicus* nodule cDNA libraries, representing approximately 3,000 nonredundant ESTs, were sequenced from the putative 5' end. Sequences were deposited in the NCBI database under accession numbers AW719193 to AW720698, BI416290 to BI416425, BI417211 to BI417245, BI417411 to BI417515, BI417753 to BI417776, BI418035 to BI418205, BI419334 to BI419349, BI419601 to BI41962, BI419739 to BI419775, and BI419869 to BI419924.

The first 2,304 cDNA clones, representing approximately 1,500 different genes, were PCR-amplified and spotted in duplicate onto nylon membranes for expression analysis. Almost all (95%) of the clones spotted had homologues at the DNA level, and 78% had homologues at the protein level in the GenBank databases. Some (13%) of the clones were homologous to characterized enzymes with EC numbers. Clones encoding many glycolytic enzymes, as well as many of the enzymes involved in ammonia assimilation and amino acid biosynthesis, were included on the arrays. Furthermore, 17% of the clones spotted had at least one predicted transmembrane region, and 4% had two or more predicted transmembrane domains. Among these were clones encoding putative transporters for sugars, amino acids, nucleotides, peptides, and inorganic anions (NO_3^- , SO_4^{2-} , PO_4^{3-}) and cations (NH_4^+ , K^+). Also included were clones encoding putative aquaporins and ATPase subunits.

The cDNA array was used to compare gene transcript levels in nodules and uninfected roots from 7-week-old *L. japonicus* plants. Probes were prepared from total RNA by reverse transcription in the presence of [α - ^{32}P]dCTP. A total of five replicate hybridizations were performed for each organ. Data measurements that were more than twice the empty local background spot were regarded as detectable and were normalized to compensate for probe activity and for the amount of spotted cDNA. An expression ratio was calculated for clones whenever at least three out of the five replicate hybridizations with both nodules and roots produced detectable signals. The significance of this difference was tested with a standard *t* test. While little technical variation was apparent when replicate filters were hybridized with probes prepared using RNA from the same organ (Fig. 1A), there were marked and significant differences in gene expression between nodules and roots (Fig. 1B).

The amount of DNA amplified and spotted was too low to be detected in the reference hybridization for 458 clones. Transcripts for 949 clones were scored as not detected in both roots and nodules because the signals associated with them were less than twice the local background level. Transcripts for 348 clones were detected only in roots, while transcripts for 65 clones were detected only in nodules. Of the remaining 484 clones, 308 showed no difference in expression levels between roots and nodules.

Transcripts of 83 different genes, represented by 130 clones on the array, were more abundant in nodules than in roots (Table 1). Of these, 36 genes appeared to be expressed only in nodules. Based on protein sequence homology, putative functions could be assigned to 67 nodule-upregulated genes, while 10 genes encoded products with homology to hypothetical proteins, other ESTs, or genomic sequence, and 6 genes had no significant match in the GenBank database (including non-mammalian ESTs).

Confirmation of the array data was performed using real-time reverse transcription-polymerase chain reaction (RT-PCR). Specific primer pairs were designed for 12 of the putative nodule-upregulated genes. The signal for ubiquitin was used to normalize for differences in the total amount of first-strand cDNA present in each reaction. In contrast to the results of the array analysis, transcripts from all 12 genes, including the 8 putative nodulin genes, were detected in roots. However, threshold cycle (C_T) values for these eight genes were generally high, indicating low abundance of these transcripts in roots (C_T indicates the fractional cycle number at which the amount of amplified target reaches a fixed threshold). The ratios of gene activity in nodules vs. gene activity in roots, as determined by both the array and the real-time RT-PCR experiments, are shown in Table 2. Real-time RT-PCR confirmed that all of the genes tested were expressed at higher levels in nodules than in uninfected roots. The ratios of expression levels determined by the two methods, however, differed markedly in some cases.

DISCUSSION

cDNA array versus real-time RT-PCR.

cDNA array analysis identified 83 genes that were upregulated in nodules compared with uninfected roots of 7-week-old *L. japonicus* plants (Table 1). Almost half (36) of these genes exhibited a nodule-specific pattern of expression, with transcripts detected in nodules but not in roots. Twelve of the genes identified by cDNA array analysis as upregulated in nodules, including eight that appeared to be nodule-specific, were subjected to real-time RT-PCR analysis, confirming that all were more highly expressed in nodules than in roots (Table 2). Not unexpectedly, real-time RT-PCR proved to be a more sensitive technique to detect low-abundance transcripts. Thus, transcripts of all twelve genes, including the eight putative nodule-specific genes, were detected in roots as well as in nodules. However, in all but one case (coproporphyrinogen oxidase), transcripts of the putative nodule-specific genes were greater than 100-fold more abundant in nodules than in roots, based on real-time RT-PCR analysis.

In three of the four cases in which a nodule to root expression ratio could be calculated from the array data as well as the real-time RT-PCR data, the latter method yielded a higher ratio. In the fourth case (LjNEST2h5, a putative lysophospholipase gene), the reverse was true. A possible explanation for both these results is that cDNA arrays are less specific in identifying transcripts than real-time RT-PCR. Thus, a given cDNA may hybridize not only with transcripts derived from the same gene, but also those derived from closely related genes, even under high-stringency conditions such as those used in our experiments. Most of the genes that were tested with real-time RT-PCR are members of larger families. Certainly the transporter genes, but also GA 2-oxidase, nitrilase, and carbonic anhydrase are usually present in more than one copy in plant genomes. Cross hybridization with a probe derived from a related and constitutively expressed gene would lead to an overestimate of transcript levels for the spotted gene in roots as well as in nodules. This would push the nodule to root transcript ratio

closer to a value of 1 and could result in an underestimate of the true ratio. On the other hand, if the spotted cDNA cross hybridizes to transcript from a related gene that is strongly upregulated in nodules, this would lead to an overestimate of the level of nodule transcript for the spotted gene and an overestimate of the nodule to root transcript ratio. In contrast, real-time RT-PCR with carefully designed, gene-specific primers is less likely to be affected by the presence of related transcripts.

Nodule-induced genes.

'*Nodulins*.' Among the genes that were upregulated in nodules relative to roots were some that have previously been identified as nodulins, or nodule-specific, in *L. japonicus*. These include LjNOD21 (Szczyglowski et al. 1997), NLj16 (Kapranov et al. 1997), and two different leghemoglobins corresponding to LjN77 and LjN132 (Szczyglowski, et al. 1997). *L. japonicus* homologues of nodulins from other species were

Fig. 1. Scatter plots of gene activity (relative transcript level) in *Lotus japonicus* nodules and roots. Gene activities were determined by a reverse-Northern approach using cDNA arrays. Lines represent activity ratios of 1 (center line), 2, and 4. **A**, Comparison of transcript profiles from two independent reverse transcription reactions performed from nodule RNA and hybridized to different filters. Points represent normalized values for a single experiment. **B**, Comparison of transcript profiles from nodules and roots of 7-week-old *L. japonicus* plants. Points represent the average of five experiments.

also identified, including MtENOD16 (Greene et al. 1998) and MtN24 (Gamas et al. 1996) from *M. truncatula*, GmN36A from soybean (Kouchi and Hata 1993), and ENOD12 from pea (Scheres et al. 1990). Leghemoglobins, which are among the most abundant proteins in nodules, bind oxygen with high affinity and are believed to facilitate oxygen diffusion in infected cells in which the free oxygen concentration is in the nanomolar range (Appleby 1984). Rapid delivery of oxygen to the terminal oxidases of bacteroids and mitochondria by leghe-moglobins may be essential to keep the steady-state free oxy-gen concentration below the level that inactivates nitrogenase. The roles of the other nodulins listed remain unknown.

Primary metabolism. The functional category containing the largest number of nodule-induced genes was that of primary metabolism. Conspicuous among the genes of this category were those encoding enzymes involved in sucrose catabolism and glycolysis. These included genes for sucrose synthase (EC 2.4.1.13) and ADP-glucose pyrophosphorylase (EC 2.7.7.27), which are involved in sucrose breakdown. Nodule-induced genes for glycolysis included glyceraldehyde-3-phosphate de-hydrogenase (EC 1.2.1.12), phosphoglycerate kinase (EC 2.7.2.3), and phosphopyruvate hydratase (EC 4.2.1.11). Su-

crose is delivered to nodules via the phloem and is the nodule's primary source of reduced carbon. Sucrose is required, not only as a source of energy for nitrogen fixation and associated metabolism, but also as a source of carbon skeletons for ammonium assimilation and export from nodules. Our results indicate that the increased glycolytic capacity of nodules com-pared with the capacity of roots (Copeland and Hong 1989) is due, at least in part, to increased transcription of some of the genes encoding enzymes of this pathway. Transcription of su-crose synthase is induced in nodules of soybean (Thummler and Verma 1987), broadbean (Küster et al. 1993, Perlick and Puhler 1993), *Alnus glutinosa* (van Ghelue et al. 1996), and *M. truncatula* (Hohnjec et al. 1999). The importance of sucrose for nodule metabolism was underlined by physiological studies of the *rug4* mutant of pea, which is deficient in sucrose syn-thase activity and severely impaired in symbiotic nitrogen fixa-tion (Gordon et al. 1999).

Other aspects of *L. japonicus* nodule carbon metabolism ap-peared to be more dramatically upregulated than glycolysis. Genes encoding two different carbonic anhydrases (CA) (EC 4.2.1.1) were found to be expressed in a 'nodule-specific' man-ner using cDNA arrays (Table 1). One clone (LjN20a7)

Table 1. Nodule-upregulated genes^a

Gene	Best database match	Nodule/root ratio ^b
Cell wall and cell growth		
LjN21E11	Actin interacting protein, <i>Arabidopsis thaliana</i>	Sp
LjN28B4	Pectinesterase, <i>Vitis vinifera</i>	Sp
LjN210H10	Proline-rich protein, <i>Daucus carota</i>	Sp
LjN211G4	Extensinlike protein, <i>A. thaliana</i>	210
LjN23G5	Proline-rich protein PRP2 precursor, <i>A. thaliana</i>	120
Defense		
LjN210A2	Chitinase homologue, <i>Sesbania rostrata</i>	1.7
Gene expression and signal transduction		
LjN21H12	<i>Glycine max</i> EST, possible ring-finger protein	Sp
LjN210D8	MADS-box protein homologue, <i>A. thaliana</i>	Sp
LjN211A7	Zinc finger protein, <i>Saccharomyces cerevisiae</i>	Sp
LjN213D6	ER6 gene, putative receptor kinase, <i>Lycopersicon esculentum</i>	Sp
LjN217C3	GTP binding protein GRPF1, <i>Fagus sylvatica</i>	Sp
LjN211A12	Calmodulin (HaCaM), <i>Helianthus annuus</i>	7.2
LjN213E10	Leucine zipper containing protein, <i>A. thaliana</i>	3.6
LjN21D2	Calmodulin, <i>Spinacia oleracea</i>	3.2
LjN212B1	Histone deacetylase, <i>A. thaliana</i>	2.4
LjN21F3	Translationally controlled tumor protein (TCTP), <i>Oryza sativa</i>	2.1
Hormone metabolism		
LjN216B9	GA 2-oxidase, <i>Phaseolus coccineus</i>	22
LjN212D12	Ent-kaurene synthase A (gibberellin biosynthesis), <i>Pisum sativum</i>	6.1
LjN212G9	Indole-3-acetate beta-glucosyltransferase, <i>A. thaliana</i>	5.3
Membrane transport		
LjN21G2	Multidrug resistance protein,	Sp
LjN28D4	Potassium transporter (KUP 1), <i>A. thaliana</i>	Sp
LjN212C4	Sulfate transport protein, <i>A. thaliana</i>	Sp
LjN221A2	Sulfate transporter (ATST1), <i>A. thaliana</i>	Sp
Nodulins		
LjN23G12	Enod 16, <i>Medicago truncatula</i>	Sp
LjN24B11	Leghemoglobin (Iljlb), <i>Lotus japonicus</i>	Sp
LjN27C12	MtN24, <i>Medicago truncatula</i>	Sp
LjN221H6	Early nodulin 12A, <i>Pisum sativum</i>	Sp
LjN21H4	Leghemoglobin (mljlb), <i>Lotus japonicus</i>	41
LjN212H12	Nlj21, <i>L. japonicus</i>	32
LjN211G8	Nlj16, <i>L. japonicus</i>	6.7
LjN21G11	Early nodulin 36A, <i>G. max</i>	6.1
Primary and secondary metabolism		
JN21A8	Monofunctional aspartokinase, <i>G. max</i>	Sp
LjN23E1	Carbonic anhydrase, <i>G. max</i>	SP
LjN213C5	Coproporphyrinogen oxidase, <i>G. max</i>	Sp
LjN215A11	Isoliquiritigenin 2'-O-methyltransferase, <i>Medicago sativa</i>	Sp
LjN218C12	Cytochrome P450, <i>L. japonicus</i>	Sp
LjN219G10	Oxidase like protein, <i>Lycopersicon esculentum</i>	Sp

(continued on next page)

^a Genes upregulated in *L. japonicus* nodules compared to uninfected roots as determined by cDNA array analysis.

^b Sp = signal detected only in nodule samples and not in root samples

represents a class of CA not previously shown to be upregulated in nodules. The second clone (LjNEST3e1) was most similar to CA genes isolated from soybean nodules (Kavroulakis et al. 2000) and *Medicago sativa* (de la Pena et al. 1997). Subsequent real-time RT-PCR analysis showed that the transcript of LjNEST3e1 was 500-fold more abundant in nodules than in roots (Table 2). A gene for phosphoenolpyruvate carboxylase (PEPC) (EC 4.1.1.31) was 4.2-fold upregulated in *L. japonicus* nodules compared to roots (Table 1). Genes encoding PEPC have been found to be upregulated in nodules of other legumes in the past (Hata et al. 1998, Pathirana et al. 1992, Sukanuma et al. 1997). Together, CA and PEPC enable the recycling of CO₂ liberated during respiration. Recycling of CO₂ presumably increases carbon-use efficiency in nodules. It would be interesting to test whether such recycling is important for SNF, using plant mutants with reduced CA or PEPC activity, or both.

The product of PEPC, oxaloacetate, can be used to produce malate, via malate dehydrogenase (MDH) (EC 1.1.1.37), or aspartate, via aspartate amino transferase (AAT) (EC 2.6.1.1). Interestingly, genes encoding both of these enzymes were found to be upregulated in *L. japonicus* nodules (Table 1). Transcripts

for MDH were fivefold more abundant in nodules than in roots, while those of AAT appeared from cDNA array analysis to be nodule specific. Similar results have been observed in other species (Gregerson et al. 1994, Miller et al. 1998). Up regulation of MDH and AAT would ensure rapid assimilation of oxaloacetate, which may avoid feedback inhibition of the CO₂ recycling pathway mentioned above. Up regulation of these enzymes may also ensure rapid carbon supply in the form of malate for bacteroid respiration and nitrogen fixation, as well as rapid assimilation of ammonium for amide export from nodules.

Several genes encoding enzymes involved in ammonia assimilation and amino acid metabolism were upregulated in *L. japonicus* nodules compared to roots. In addition to AAT mentioned above, transcripts for asparagine synthetase (AS) (EC 6.3.5.4) and glutamine synthetase (GS) (EC 6.3.1.2) were 7.4-fold and 2.3-fold, respectively, more abundant in nodules than in roots. Homologues of all three of these genes have been found to be induced in nodules of other legume species (Cullimore et al. 1984, Dunn et al. 1988, Gregerson et al. 1994, Hirel et al. 1987, Marsolier et al. 1995, Padilla et al. 1987, Perlick and Puhler 1993, Reynolds et al. 1992, Roche et al.

Table 1. (continued from preceding page)

Gene	Best database match	Nodule/root ratio ^b
LjNEST20A7	Carbonic anhydrase, <i>A. thaliana</i>	Sp
LjNEST20A9	Aspartate aminotransferase, <i>A. thaliana</i>	Sp
LjNEST22B1	Aspartate kinase, <i>G. max</i>	Sp
LjNEST2H5	Lysophospholipase, <i>A. thaliana</i>	26
LjNEST19F10	Carbamoyl phosphate synthetase, small subunit, <i>A. thaliana</i>	20
LjNEST2F5	Methyltransferase, <i>Amycolatopsis mediterranei</i>	9.4
LjNEST13B3	Chorismate mutase 3, <i>A. thaliana</i>	9.3
LjNEST11B7	UDP-glucose 4-epimerase, <i>Pyrococcus horikoshii</i>	8.7
LjNEST10B9	Asparagine synthase (glutamine-hydrolyzing), <i>L. japonicus</i>	7.4
LjNEST10B11	Nitrilase, <i>Nicotiana tabacum</i>	6.8
LjNEST12E12	Glutathione peroxidase, <i>A. thaliana</i>	6.8
LjNEST1H5	Formamidase, <i>A. thaliana</i>	6.0
LjNEST2H6	Dimethylaniline monooxygenase, <i>A. thaliana</i>	5.6
LjNEST9C2	Malate dehydrogenase, <i>Medicago sativa</i>	5.0
LjNEST3A7	GDP-mannose pyrophosphorylase, <i>Solanum tuberosum</i>	4.2
LjNEST9G10	Phosphoenolpyruvate carboxylase, <i>G. max</i>	4.2
LjNEST7B10	Sucrose synthase, <i>Vicia faba</i>	4.0
LjNEST10F12	Thiamine biosynthetic protein thi, <i>Citrus sinensis</i>	3.7
LjNEST13E2	Phosphopyruvate hydratase, <i>Oryza sativa</i>	2.5
LjNEST13F2	ADP-glucose pyrophosphorylase (aggS1), <i>Pisum sativum</i>	2.3
LjNEST16A7	Phosphoglycerate kinase cytosolic, <i>Nicotiana tabacum</i>	2.3
LjNEST16A10	Glutamate-ammonia ligase chloroplast, <i>Phaseolus vulgaris</i>	2.3
LjNEST9F2	Glyceraldehyde-3-phosphate dehydrogenase, <i>Pisum sativum</i>	2.0
Protein synthesis and processing		
LjNEST2A1	<i>G. max</i> EST, possible DNAJ chaperonelike protein	Sp
LjNEST2F7	Ribosomal protein S9, <i>A. thaliana</i>	Sp
Miscellaneous		
LjNEST22A12	Flavonol-induced cDNA, <i>Petunia x hybrida</i>	Sp
LjNEST2D2	Cycloidea-like protein family, <i>A. thaliana</i>	Sp
LjNEST10A9	Ethylene response protein, <i>A. thaliana</i>	10
LjNEST14H12	Chloroplast nucleoid DNA binding protein, <i>A. thaliana</i>	6.5
LjNEST1D5	No match	Sp
LjNEST8G7	No match	Sp
LjNEST12C2	No match	Sp
LjNEST13H8	No match	Sp
LjNEST2E3	No match	5.9
LjNEST4A4	No match	2.2
LjNEST1H3	Genomic sequence, <i>L. japonicus</i>	Sp
LjNEST3A10	Genomic sequence, <i>L. japonicus</i>	Sp
LjNEST20H11	Hypothetical protein, <i>A. thaliana</i>	Sp
LjNEST11H5	Hypothetical protein, <i>A. thaliana</i>	Sp
LjNEST15H9	Hypothetical protein, <i>A. thaliana</i>	8.4
LjNEST16H7	Hypothetical protein, <i>Cicer arietinum</i>	6.4
LjNEST11E1	Genomic sequence, <i>A. thaliana</i>	5.9
LjNEST12G1	Hypothetical protein, <i>Cicer arietinum</i>	4.0
LjNEST1C10	EST, <i>Medicago truncatula</i>	3.2
LjNEST5B3	Genomic sequence, <i>L. japonicus</i>	3.0

1993, Shi et al. 1997, Stanford et al. 1993, Temple et al. 1995, Tingey et al. 1987, Tsai and Coruzzi 1990). *Lotus* spp., like other temperate legumes, export nitrogen from nodules principally in the form of amides, especially asparagine (Vance et al. 1987). The importance of AAT, GS, and AS in nodule asparagine synthesis has been noted previously (Vance et al. 1994).

Five other genes involved in nitrogen metabolism or amino acid biosynthesis that have not been identified as nodule-induced in the past were found to be upregulated in *L. japonicus* nodules (Table 1). A gene for aspartate kinase (EC 2.7.2.4) was expressed in nodules but not roots, while a gene for chorismate mutase (EC 5.4.99.5) was 9.3-fold more abundant in nodules than roots. Genes for carbamoyl-phosphate synthetase (CPS) (EC 6.3.5.5.), nitrilase (EC 3.5.5.1), and formamidase (EC 3.5.1.49) were all significantly upregulated in nodules compared with roots. CPS consumes CO₂ and therefore, like CA and PEPC, may contribute to the recycling of this gas in nodules. Although only glutamine-hydrolyzing forms of CPS have been described in plants (Zhou et al. 2000), a form that uses ammonia instead of glutamine is present in some mammals. Given the abundance of ammonium in nodules, it would be interesting to test the substrate specificity of the *L. japonicus* nodule CPS. Ammonia is a product of both formamidase and nitrilase. However, if the reactions catalyzed by these enzymes are reversible, then they may be alternative sinks for ammonium, incorporating the nitrogen into nonpeptide C-N bonds. Measurement of formamide and formate concentrations in *L. japonicus* nodules may provide some insight into the potential role and significance of formamidase in this organ.

One gene encoding an enzyme of heme biosynthesis, coproporphyrinogen oxidase (EC 1.3.3.3), was found by cDNA array analysis to be expressed specifically in nodules. The heme moiety is essential for the production of leghemoglobin and other proteins, such as cytochromes. Coproporphyrinogen oxidase has been shown to be upregulated in nodules of soybean and pea (Madsen et al. 1993, Santana et al. 1998). The gene for glutamate 1-semialdehyde amino transferase, also involved in heme biosynthesis, was also found to be upregulated in soybean nodules (Sangwan and O'Brian 1993). Thus, a general induction of heme biosynthesis appears to be important for SNF.

Secondary metabolism and suppression of defense. In contrast to primary metabolism, relatively few genes involved in secondary metabolism were identified as nodule upregulated. One such gene encoded isoliquiritigenin 2'-O-methyltransferase, or chalcone OMT, an enzyme that, in *M. truncatula*, is involved in the biosynthesis of a potent flavonoid inducer of rhizobia nodulation genes (Maxwell et al. 1993). The methylation reaction catalyzed by chalcone OMT precludes subsequent use of the chalcone in phytoalexin biosynthesis. Thus, high levels of expression of this gene in *L. japonicus* nodules may be a mechanism to disarm part of the normal plant defense response against intruders. Antisense-inhibition of this 'nodule-specific' gene would be one way to test its importance in SNF.

Two genes for enzymes involved in scavenging of active oxygen species, glutathione peroxidase (EC 1.11.1.9) and guanosine 5'-diphosphate-mannose pyrophosphorylase (EC 2.7.7.13), which is involved in ascorbate biosynthesis (Conklin et al. 1999), were found to be 6.8-fold and 4.2-fold, respectively, induced in nodules. Although neither gene has been identified as nodule-induced in the past, a gene for ascorbate peroxidase is induced in alfalfa nodules (Dalton et al. 1998). Removal of peroxides in nodules may prevent oxygen toxicity and contribute to the maintenance of functional leghemoglobin (Dalton et al. 1986, Moreau et al. 1995). In view of the role that H₂O₂ appears to play in triggering the hypersensitive response in plants (Levine et al. 1994), it is also tempting to

speculate that induction of H₂O₂-consuming enzymes may be necessary to avoid additional defense responses against rhizobia in nodules.

Membrane transport. Homologues of four transporters, two different sulfate transporters, one potassium transporter, and a member of the multidrug resistance transporter family, were found to be 'nodule-specific' in *L. japonicus* (Table 1). Real-time RT-PCR analysis of two of these transporters showed that transcript levels were several hundred-fold higher in nodules than in roots. A soybean homologue of the putative sulfate transporters, GmN70, was also found to be 'nodule-specific' (Kouchi and Hata 1993). It will be interesting to confirm or determine the substrates of these transporters, their intracellular location, and their physiological roles. We are particularly interested in the possibility that one or more of the transporters may be located on the symbiosome membrane, where they could play important roles in nutrient exchange between the plant and the bacteroids.

Hormone metabolism. We found that several genes involved in hormone metabolism were induced in nodules (Table 1). Two of these, ent-kaurene synthase A and GA-2 oxidase, are involved in gibberellic acid (GA) biosynthesis and GA inactivation, respectively (Ait-Ali et al. 1997, Thomas et al. 1999). This result is somewhat perplexing, given the potential antagonistic effects of the two gene products. In situ RNA analysis of nodule cross-sections to test for spatial separation of gene expression may help to make sense of this result. Perhaps high levels of ent-kaurene synthase A produce high concentrations of active GA in some nodule cell types, while high levels of GA-2 oxidase ensure low levels of active GA in neighboring cell types. A similarly perplexing result was obtained for two genes involved in indole-3-acetic acid (IAA) metabolism. Genes encoding nitrilase (discussed above), an enzyme involved in IAA synthesis (Bartling et al. 1992), and indole-3-acetate beta-glucosyltransferase, an enzyme involved in producing inactive conjugates of IAA (Szczeszen et al. 1994), were both induced in nodules. Once again, it would be interesting to test whether there is any spatial separation in expression of these two genes in nodules.

Cell Wall. The functional category of cell wall structure-related and cell growth-related genes contained four nodule upregulated genes, including extensin and two proline-rich proteins. Several early nodulin genes, including ENOD12, encode proline-rich proteins, which are probably structural proteins of the cell wall (Löbner and Hirsch 1993, Scheres et al. 1990). During the infection process, remarkable changes in cell structure occur, including the formation of additional cell wall layers and of infection threads composed of cell wall material (Scheres et al. 1990).

Signaling and transcription regulation. Ten genes encoding proteins that may have roles in signal transduction and gene transcription were found to be upregulated in *L. japonicus* nodules (Table 1). Five of these appeared to be 'nodule-specific,' including three encoding putative transcription factors: a MADS-box protein, a zinc finger protein, and a putative ring-finger protein. A leucine zipper-containing protein gene was also found to be upregulated in nodules. Although little is known about the role of specific transcription factors during SNF, they will almost certainly be found to be at the heart of many regulatory processes controlling nodule development and function. The only putative transcription factor with a clear role in SNF is the *Nin* gene in *L. japonicus*, which is essential for a very early step in nodule development (Schauser et al. 1999). The genes identified here are obvious candidates for reverse genetics approaches to identify those that are essential for SNF. We are currently mapping these genes to see if any are linked to known symbiotic loci.

Unknown. In addition to the clones for which putative functions could be assigned, ten nodule-enhanced clones were homologous only to genomic sequences or hypothetical proteins from other plant species and, thus, may prove to be novel genes important in nodule function. Perhaps even more interesting, though, are the three clones that match *L. japonicus* genome sequence only and the five for which no significant match could be found, all of which could represent legume-specific genes.

Missing nodulins?

Homologues of some nodulin genes, e.g., the gene encoding Nodulin 26, which is one of the most abundant proteins of the symbiosome membrane, were conspicuous by their absence in Table 1. Several explanations could account for such absences. First, the 1,500 genes represented on our arrays probably account for only a small fraction of all genes that are expressed in *L. japonicus* nodules. Thus, genes with very low transcript abundance in nodules are less likely to be present in this set of clones. Second, in the case of nodulin genes that are members of a multigene family, including Nodulin 26, cross hybridization between family members could have masked real differences in the abundance of specific transcripts, as discussed above. This may explain why Nodulin 26 gene homologues, which were present on our arrays, did not appear to be induced in nodules. Alternatively, the timing of the experiment may affect which genes are identified. These experiments were performed with RNA collected at one time point from mature nodules and may not contain transcripts of some early nodulins or of genes expressed at other stages of development. A final explanation that could account for differences in nodulin gene profiles between species is the existence of genuine differences in nodule development and nodule metabolism when species are compared, especially tropical and temperate legume species and species producing determinate and indeterminate nodules. Such differences will almost certainly reflect differences in gene expression between species.

cDNA arrays: pros and cons.

Transcript profiling using cDNA arrays, as exemplified by the work presented here, has a number of advantages and disadvantages compared with other techniques. One disadvantage compared with methods like cDNA-amplified fragment length polymorphism (AFLP) analysis (Bachem et al. 1996) that can, in theory, assay every gene that is expressed in a particular biological sample is that transcript profiling with DNA arrays is limited to the genes that are represented on the filter or slide. In practice, this is not a great drawback, especially for those model species, including several legumes, for which tens of thousands of nonredundant EST clones are now available for arraying. Perhaps a greater long-term problem for experiments based on cDNA arrays is the cross hybridization that can occur between immobilized clones and the transcripts of related genes. Cross hybridization could account for some of the differences that we observed between data obtained with arrays and with real-time RT-PCR (Table 2). One widely marketed solution to this problem is to use short, gene-specific oligonucleotides on arrays. The cost of such a solution may be prohibitive to many groups, however. Alternatively, shorter more specific cDNA fragments, such as from 3' untranslated regions, could be spotted. Another limitation of DNA arrays in general is their sensitivity. Approximately 40% of clones on our array gave no detectable signal, presumably because of low transcript abundance for these genes. Although it has been estimated that DNA arrays can detect a single transcript per cell (Ruan et al. 1998, Wodicka et al. 1997), many genes in complex, multicellular organisms like plants are not expressed in

all cells and many may be expressed at a very low level in one or a few cells only. It is unlikely that transcripts of such genes will be detectable using current DNA or oligonucleotide array technology unless transcripts are amplified before probe synthesis. This limitation was apparent in the work presented here; transcripts for some genes were not detected in roots using cDNA arrays although real-time PCR did detect such transcripts. Thus, real-time RT-PCR is a more sensitive alternative to DNA arrays for expression analysis. However, the costs of real-time RT-PCR limit its use to moderate throughput rather than the high, genome-scale throughput that is achievable with DNA arrays. In fact, the scale on which genes can be analyzed is one of the biggest advantages of DNA arrays at present. Whole genomes of many prokaryotes, as well as of many yeasts, are now represented on DNA arrays that can be used for transcriptome analysis. In the near future, whole-genome arrays for *Arabidopsis thaliana* will also be available. We are currently pooling our resources with other groups to produce a cDNA array that represents 10,000 different *L. japonicus* genes. Another significant advantage of DNA arrays over techniques such as cDNA-AFLP is the speed with which a completely new analysis can be made. In this study, we simply compared the transcript profile of nodules with that of roots. However, with the same filters we could quickly, easily, and inexpensively address questions about large-scale gene expression in different mutants of *L. japonicus*, or in wild-type plants subjected to different environmental conditions, to gain molecular insight into the phenotype of plants and physiological responses to the environment.

Summary.

In summary, we have used cDNA arrays representing 1,500 different *L. japonicus* genes to identify genes that are induced in nodules compared with roots. The data indicate that genes involved in sucrose breakdown and glycolysis, together with genes involved in ammonia assimilation and asparagine synthesis, are upregulated at the transcriptional level during SNF. The results are consistent with current models of nodule carbon and nitrogen metabolism. In our study, over 50 genes were identified for the first time as nodule-induced genes. These include genes involved in transport, hormone metabolism, and signal transduction, as well as several that may help to avoid plant defense responses against invading rhizobia. Included in the latter category were two genes involved in detoxification of active oxygen species and one that could interfere in the production of phytoalexins. Although only a fraction of all genes that are likely to be expressed in nodules have been surveyed

Table 2. Validation of putative nodule-upregulated genes by real-time PCR

Clone	Best blast match	Relative expression	
		Real-time RT-PCR	cDNA array ^a
LjNEST3e1	Carbonic anhydrase	530	Sp
LjNEST7c12	MtN24	310	Sp
LjNEST8d4	Potassium transporter	450	Sp
LjNEST11h5	Hypothetical protein	6900	Sp
LjNEST12c4	Sulfate transport protein	390	Sp
LjNEST13c5	Coproporphyrinogen oxidase	27	Sp
LjNEST20a7	Carbonic anhydrase	140	Sp
LjNEST21a2	Sulfate transporter (ATST1)	510	Sp
LjNEST12h12	Nlj21	140	32
LjNEST2h5	Lysophospholipase	2	26
LjNEST16b9	GA 2-oxidase	230	22
LjNEST10b11	Nitrilase	18	6.8

^a Sp = signal detected only in nodule samples and not in root samples.

in our study, the wealth of data that has been produced provides a valuable starting point for much future work.

MATERIALS AND METHODS

Biological materials.

L. japonicus GIFU (B-129) seeds were germinated and grown in sand in a controlled environment (16 h day and 8 h night, 21°C day and 17°C night temperature, and 60% relative humidity). Plants were inoculated when 7 days old with *Mesorhizobium loti* strain R7A and were provided with 1 mM KNO₃ for the first 2 to 3 weeks of growth. Roots from control noninoculated plants and nodules from infected plants were harvested 7 weeks after germination, snap-frozen in liquid nitrogen, and stored at -80°C.

Construction of a nodule cDNA library. Two *L. japonicus* nodule cDNA libraries were used to generate clones for array analysis. The first library was provided by J. Stougaard (Aarhus, Denmark) and constructed in the λ ZAP-XR vector (Stratagene, La Jolla, CA, U.S.A.). The second library was constructed from RNA from 3- and 5-week-old nodules in the pZL1 vector using the Superscript Lambda system for cDNA synthesis and λ cloning (Invitrogen, Groningen, The Netherlands). A total of 5,495 clones from *L. japonicus* nodule cDNA were sequenced from their 5' ends (AGOWA, Berlin, Germany) and deposited into GenBank.

cDNA arrays.

Construction of arrays. The cDNA array procedure was adapted from Thimm et al. (2001). Plasmid DNA from 2,304 bacterial clones was PCR-amplified in 40 μ l reactions in 384 well plates. For amplification of clones in the λ ZAP-XR vector, standard T3 and T7 primers were used. For amplification of clones in the pZL1 vector, specific primers were designed (forward primer, 5'-GTACGTAAGCTTGGATCCTCTAG-3'; reverse primer, 5'-TGCAGGTACCGGTCCGGAATTCC-3'). Each reaction included 1.5 mM magnesium acetate, 0.2 mM each dNTP, 0.1 μ M each primer, 4 μ l of 10 $\times$ PCR buffer (0.5 M potassium acetate, 1% Tween 20, 15 mM magnesium acetate, 0.35 M Tris, pH 10.4, 0.15 M Tris, pH 4.8), and 0.5 units Taq DNA polymerase. The amplification conditions were 95°C for 5 min, followed by 25 cycles of 95°C for 1 min, 67°C for 1 min, and 72°C for 1 min, followed by one cycle of 72°C for 1 min.

For each clone, three independent reactions were performed and then pooled to reduce variation in PCR efficiency. Pooled PCR products were concentrated from 120 μ l to approximately 50 μ l by evaporation before being spotted onto Nytran Super-Charge nylon transfer membranes (Schleicher & Schuell, Dassel, Germany) using the BioGrid robotic system (BioRobotics Ltd, Cambridge, U.K.) with a 384er gridding tool (radius 0.4 mm, 10 transfers per spot). The spotting scheme followed a 4 $\times$ 4 secondary grid pattern, with each secondary grid containing seven clones spotted in duplicate, plus an empty local background spot. DNA on membranes was denatured with 0.5 M NaOH and was baked at 80°C for 2 h.

Reference hybridization of arrays. To estimate the amount of DNA spotted for each clone, filters were hybridized to a reference oligonucleotide that was complementary to the vector sequence. For the λ ZAP-XR vector, the oligonucleotide sequence was 5'-GCTGCAGGAATTC-3' and for the pZL1 vector, 5'-ACGCGTGGGTCTGA-3'. Oligonucleotides were end-labeled with γ^{32} P-ATP (Amersham Pharmacia, Freiburg, Germany) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA, U.S.A.). Filters were preincubated in SSarc (4 $\times$ SSC [1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 2 H₂O, pH 7.0, 24% Sarcosyl NL30, 4mM EDTA) for 2

h at 5°C, before addition of the probe and hybridization overnight at 5°C. Filters were washed in SSarc twice for 30 min at 5°C and exposed to a phosphor screen (Fuji, Tokyo) for 4 to 6 h, after which the imaging plate was scanned with the BAS-1800 II phosphor imager (Fuji) at a resolution of 50 μ m per pixel. Bound reference oligonucleotides were removed by washing the filters twice in diluted SSarc (1:10 in 1 mM EDTA) for 30 min at 80°C.

Complex hybridization of arrays. Total RNA was isolated from nodules and uninfected roots of 7-week-old *L. japonicus* plants as described by Jacobsen-Lyon and associates (1995). Five replicate complex hybridizations were performed for each organ using probes derived from two separately grown batches of plants (three replicates from one batch and two from the other). Complex probes were prepared from 10 μ g of total RNA by reverse transcription using SuperScript II reverse transcriptase (Invitrogen) in the presence of [α^{32} P]-dCTP. Filters were preincubated for 2 h at 65°C in Church buffer (0.5 M sodium phosphate, pH 7.2, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA) plus 100 μ g of denatured salmon sperm DNA per ml. Complex probes in fresh Church buffer were added to filters and hybridized for at least 24 h at 65°C. Filters were then washed successively for 20 min in 1 $\times$ SSC containing 0.1% SDS and 4 mM sodium phosphate (pH 7.2), for 20 min in 0.2 $\times$ SSC containing 0.1% SDS and 4 mM sodium phosphate (pH 7.2), and for 20 min in 0.1 $\times$ SSC containing 0.1% SDS and 4 mM sodium phosphate (pH 7.2), all at 65°C. Air-dried filters were exposed to a phosphor screen overnight, which was then scanned as described above.

Data analysis. Detection and quantification of signal intensities were performed using AIS software (4.0 rev. 1.7; Imaging Research Inc., St. Catharines, Ontario, Canada), resulting in raw data representing the pixel intensity within a defined area around each spot. An empty spot within each 4 $\times$ 4 subgrid was used for measurements of the local background intensity. Signal intensity values that were not more than twice the local background were regarded as undetectable and set to zero. An average gene activity from the five replicate experiments was accepted when nonzero values were obtained from at least three of the filters.

Raw signals were normalized in a three-step process to compensate for the total probe activity and the amount of DNA spotted for each clone. In the first step of the normalization procedure, the signals from the reference hybridization were internally normalized by subtracting the local background value and then dividing the remainder by the average of all signals for the given filter. If the measurement from the reference hybridization was not high enough (i.e., twice as high as the local background), the clone was flagged as undetectable and discarded from the analysis. In the second step, the signals from the complex hybridization were internally normalized by subtracting the local background and dividing the remainder by the background-subtracted value for the housekeeping gene ubiquitin. In the final step, the value from step 2 was divided by that from step 1 to obtain the normalized gene activity.

In order to detect significant changes in expression level between roots and nodules, we established the following procedure. Measurements for all clones in which the reference hybridization failed to detect enough PCR product were discarded from the dataset. All other clones were divided into two categories. The first group contained clones for which it was possible to detect a signal in at least three of the five replicates for one organ but for which the transcript was undetectable in the other organ. In these cases, it was not possible to calculate an expression ratio or perform a *t* test, and the clone was regarded as specifically expressed. The second group contained clones for which a transcript could be detected in at least three

of the five replicates in both nodules and roots. In these cases, it was possible to calculate an expression ratio based on the average of the valid measurements. The significance of this difference was tested with a standard *t* test.

Real-Time RT-PCR. Total RNA (40 µg) extracted from nodules and uninfected roots of 7-week-old *L. japonicus* plants was treated with DNase I, as described previously (Ausubel et al. 1994). DNase-treated RNA (10 µg) was used to make first strand cDNA using Moloney murine leukemia virus reverse transcriptase (Epicentre Technologies Madison, WI, U.S.A.), in a 50 µl reaction, according to the manufacturer's instructions.

Gene-specific primers were designed using the Primer Express Software, version 2.0 (Applied Biosystems, Foster City, CA, U.S.A.) and synthesized by Invitrogen. Primers were used at an initial concentration of 500 nM in a 25 µl reaction containing 12.5 µl SYBR Green PCR Mastermix (Applied Biosystems) and 2 µl of template diluted 1:20 or 1:40. Each assay included duplicate reactions for both dilutions of the cDNA, for a total of four replicates per clone. As controls, RNA that had not been reverse-transcribed and a negative control without any nucleic acid template were run with each assay. Separate reactions containing primers to three putative constitutively expressed genes encoding ubiquitin, elongation factor 1, and glyceraldehyde-3-phosphate dehydrogenase were also included in each assay to enable normalization and comparison between experiments.

PCR reactions were run with the ABI Prism 5700 sequence detection system (Applied Biosystems) using the following program: 50°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C for 15 s, followed by 60°C for 1 min. Following amplification, the dissociation kinetics of the amplicons, measured as change in fluorescence as a function of temperature, were measured for the detection of nonspecific products. The dissociation protocol was 95°C for 15 s, 60°C for 20 s, followed by a 20 min slow ramp from 60 to 95°C.

Before proceeding with real-time RT-PCR, we routinely verified that each set of primers produced a single amplicon during PCR, using gel electrophoresis. To calculate the relative abundance of gene transcripts in nodules versus in roots from real-time RT-PCR data, we assumed that each PCR cycle resulted in an exact doubling in the number of amplicons. To approach 100% efficiency in PCR, we designed primer pairs that produced very short amplicons, typically less than 100 bp in length.

ACKNOWLEDGMENTS

We thank B. Essigmann and S. Fischer for expert assistance in producing the cDNA arrays and M. McKenzie for useful comments on the manuscript. The Max Planck Society, Deutsche Forschungs Gemeinschaft, and Alexander von Humboldt Society are also thanked for their support of this work.

LITERATURE CITED

Ait-Ali, T., Swain, S. M., Reid, J. B., Sun, T., and Kamiya, Y. 1997. The LS locus of pea encodes the gibberellin biosynthesis enzyme entkaurene synthase A. *Plant J.* 11:443-54.
 Appleby, C. A. 1984. Leghaemoglobin and Rhizobium respiration. *Annu. Rev. Plant Physiol.* 35:443-478.
 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. 1994. *Current Protocols in Molecular Biology*. John Wiley & Sons Inc., New York.
 Bachem, C. W., van der Hoeven, R. S., de Bruijn, S. M., Vreugdenhil, D., Zabeau, M., and Visser, R. G. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *Plant J.* 9:745-753.
 Bartling, D., Seedorf, M., Mithofer, A., and Weiler, E. W. 1992. Cloning and expression of an Arabidopsis nitrilase which can convert indole-3-

acetonitrile to the plant hormone, indole-3-acetic acid. *Eur. J. Biochem.* 205:417-424.
 Bladergroen, M. R., and Spaink, H. P. 1998. Genes and signal molecules involved in the rhizobia-Leguminosae symbiosis. *Curr. Opin. Plant Biol.* 1:353-359.
 Caetano-Anolles, G., and Gresshoff, P. M. 1991. Plant genetic control of nodulation. *Annu. Rev. Microbiol.* 45:345-382.
 Conklin, P. L., Norris, S. R., Wheeler, G. L., Williams, E. H., Smirnoff, N., and Last, R. L. 1999. Genetic evidence for the role of GDP-mannose in plant ascorbic acid (vitamin C) biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 97:4198-4203.
 Copeland, L. J. V., and Hong, Z. Q. 1989. Enzymes of carbohydrate metabolism in soybean nodules. *Phytochemistry*. 28:57-61.
 Crespi, M., and Galvez, S. 2000. Molecular mechanisms in root nodule development. *J. Plant Growth Regul.* 19:155-166.
 Cullimore, J. V., Gebhardt, C., Saarelainen, R., Mifflin, B. J., Idler, K. B., and Barker, R. F. 1984. Glutamine synthetase of *Phaseolus vulgaris* L.: Organ-specific expression of a multigene family. *J. Mol. Appl. Genet.* 2:589-599.
 Dalton, D. A., Russell, S. A., Hanus, F. J., Pascoe, G. A., and Evans, H. J. 1986. Enzymatic reactions of ascorbate and glutathione that prevent peroxide damage in soybean root nodules. *Proc. Natl. Acad. Sci. U.S.A.* 83:3811-3815.
 Dalton, D. A., Joyner, S. L., Becana, M., Iturbe-Ormaetxe, I., and Chatfield, J. M. 1998. Antioxidant defenses in the peripheral cell layers of legume root nodules. *Plant Physiol.* 116:37-43.
 Day, D. A., and Copeland, L. 1991. Carbon metabolism and compartmentation in nitrogen-fixing legume nodules. *Plant Physiol. Biochem.* 29:185-201.
 de la Pena, T. C., Frugier, F., McKhann, H. I., Bauer, P., Brown, S., Kondorosi, A., and Crespi, M. 1997. A carbonic anhydrase gene is induced in the nodule primordium and its cell-specific expression is controlled by the presence of *Rhizobium* during development. *Plant J.* 11:407-420.
 Denarie, J., Debelle, F., and Prome, J. C. 1996. Rhizobium lipo-chitoooligosaccharide nodulation factors: Signaling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* 65:503-535.
 Downie, J. A., and Walker, S. A. 1999. Plant responses to nodulation factors. *Curr. Opin. Plant Biol.* 2:483-489.
 Dunn, K., Dickstein, R., Feinbaum, R., Burnett, B. K., Peterman, T. K., Thoidis, G., Goodman, H. M., and Ausubel, F. M. 1988. Developmental regulation of nodule-specific genes in alfalfa root nodules. *Mol. Plant-Microbe Interact.* 1:66-74.
 Gamas, P., de Carvalho Niebel, F., Lescure, N., and Cullimore, J. V. 1996. Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol. Plant-Microbe Interact.* 9:233-242.
 Gordon, A. J., Minchin, F. R., James, C. L., and Komina, O. 1999. Sucrose synthase in legume nodules is essential for nitrogen fixation. *Plant Physiol.* 120:867-878.
 Greene, E. A., Erard, M., Dedieu, A., and Barker, D. G. 1998. MtENOD16 and 20 are members of a family of phytoeyanin-related early nodulins. *Plant Mol. Biol.* 36:775-783.
 Gregerson, R. G., Miller, S. S., Petrowski, M., Gantt, J. S., and Vance, C. P. 1994. Genomic structure, expression, and evolution of the alfalfa aspartate aminotransferase genes. *Plant Mol. Biol.* 25:387-399.
 Hata, S., Izui, K., and Kouchi, H. 1998. Expression of a soybean nodule-enhanced phosphoenolpyruvate carboxylase gene that shows striking similarity to another gene for a house-keeping isoform. *Plant J.* 13:267-273.
 Hirel, B., Bouet, C., King, B., Layzell, D., Jacobs, F., and Verma, D. P. S. 1987. Glutamine-synthetase genes are regulated by ammonia provided externally or by symbiotic nitrogen-fixation. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:1167-1171.
 Hohnjec, N., Becker, J. D., Puhler, A., Perlick, A. M., and Küster, H. 1999. Genomic organization and expression properties of the MtSucS1 gene, which encodes a nodule-enhanced sucrose synthase in the model legume *Medicago truncatula*. *Mol. Gen. Genet.* 261:514-522.
 Jacobsen-Lyon, K., Jensen, E. Ø., Jørgensen, J., Marcker, K. A., Peacock, W. J., and Dennis, E. S. 1995. Symbiotic and non-symbiotic hemoglobin genes of *Casuarina glauca*. *Plant Cell* 7:213-223.
 Jiménez-Zurdo, J. I., Frugier, F., Crespi, M. D., and Kondorosi, A. 2000. Expression profiles of 22 novel molecular markers for organogenetic pathways acting in alfalfa nodule development. *Mol. Plant-Microbe Interact.* 13:96-106.
 Kahn, M. L., McDermott, T. R., and Udvardi, M. K. 1998. Carbon and nitrogen metabolism in the Rhizobiaceae. In: *The Rhizobiaceae*. H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas (eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands.

- Kapranov, P., de Bruijn, F. J., and Szczyglowski, K. 1997. Novel, highly expressed late nodulin gene (LjNOD16) from *Lotus japonicus*. *Plant Physiol.* 113:1081-1090.
- Kavroulakis, N., Flemetakis, E., Aivalakis, G., and Katinakis, P. 2000. Carbon metabolism in developing soybean root nodules: The role of carbonic anhydrase. *Mol. Plant-Microbe Interact.* 13:14-22.
- Kehoe, D. M., Villand, P., and Somerville, S. 1999. DNA microarrays for studies of higher plants and other photosynthetic organisms. *Trends Plant Sci.* 4:38-41.
- Kouchi, H., and Hata, S. 1993. Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* 238:106-119.
- Küster, H., Frühling, M., Perlick, A. M., and Puhler, A. 1993. The sucrose synthase gene is predominantly expressed in the root nodule tissue of *Vicia faba*. *Mol. Plant-Microbe Interact.* 6:507-514.
- Levine, A., Tenhaken, R., Dixon, R., and Lamb, C. 1994. H₂O₂ from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell.* 79:583-593.
- Löbner, M., and Hirsch, A. M. 1993. A gene that encodes a proline-rich nodulin with limited homology to PsENOD12 is expressed in the invasion zone of *Rhizobium meliloti*-induced alfalfa root nodules. *Plant Physiol.* 103:21-30.
- Long, S. R. 2001. Genes and signals in the *Rhizobium*-legume symbiosis. *Plant Physiol.* 125:69-72.
- Madsen, O., Sandal, L., Sandal, N. N., and Marcker, K. A. 1993. A soybean coproporphyrinogen oxidase gene is highly expressed in root nodules. *Plant Mol. Biol.* 23:35-43.
- Marsolier, M. C., Debrosses, G., and Hirel, B. 1995. Identification of several soybean cytosolic glutamine synthetase transcripts highly or specifically expressed in nodules: Expression studies using one of the corresponding genes in transgenic *Lotus corniculatus*. *Plant Mol. Biol.* 27:1-15.
- Maxwell, C. A., Harrison, M. J., and Dixon, R. A. 1993. Molecular characterization and expression of alfalfa isoliquiritigenin 2'-O-methyltransferase, an enzyme specifically involved in the biosynthesis of an inducer of *Rhizobium meliloti* nodulation genes. *Plant J.* 4:971-981.
- Miller, S. S., Driscoll, B. T., Gregerson, R. G., Gantt, J. S., and Vance, C. P. 1998. Alfalfa malate dehydrogenase (MDH): Molecular cloning and characterization of five different forms reveals a unique nodule-enhanced MDH. *Plant J.* 15:173-184.
- Moreau, S., Puppo, A., and Davies, M. J. 1995. The reactivity of ascorbate with different redox states of leghaemoglobin. *Phytochemistry.* 39:1281-1286.
- Padilla, J. E., Campos, F., Conde, V., Lara, M., and Sanchez, F. 1987. Nodule-specific glutamine-synthetase is expressed before the onset of nitrogen-fixation in *Phaseolus vulgaris* L. *Plant Mol. Biol.* 9:65-74.
- Pathirana, S. M., Vance, C. P., Miller, S. S., and Gantt, J. S. 1992. Alfalfa root nodule phosphoenolpyruvate carboxylase: Characterization of the cDNA and expression in effective and plant-controlled ineffective nodules. *Plant Mol. Biol.* 20:437-450.
- Pawlowski, K. 1997. Nodule-specific gene expression. *Physiol. Plant.* 99:617-631.
- Perlick, A. M., and Puhler, A. 1993. A survey of transcripts expressed specifically in root nodules of broadbean (*Vicia faba* L.). *Plant Mol. Biol.* 22:957-970.
- Reynolds, P. H., Smith, L. A., Dickson, J. M., Jones, W. T., Jones, S. D., Rodber, K. A., Carne, A., and Liddane, C. P. 1992. Molecular cloning of a cDNA encoding aspartate aminotransferase-P2 from lupin root nodules. *Plant Mol. Biol.* 19:465-472.
- Richmond, T., and Somerville, S. 2000. Chasing the dream: Plant EST microarrays. *Curr. Opin. Plant Biol.* 3:108-116.
- Roche, D., Temple, S. J., and Sengupta-Gopalan, C. 1993. Two classes of differentially regulated glutamine synthetase genes are expressed in the soybean nodule: A nodule-specific class and a constitutively expressed class. *Plant Mol. Biol.* 22:971-983.
- Ruan, Y., Gilmore, J., and Conner, T. 1998. Towards *Arabidopsis* genome analysis: Monitoring expression profiles of 1,400 genes using cDNA microarrays. *Plant J.* 15:821-833.
- Sangwan, I., and O'Brian, M. R. 1993. Expression of the soybean (*Glycine max*) glutamate 1-semialdehyde aminotransferase gene in symbiotic root nodules. *Plant Physiol.* 102:829-834.
- Santana, M. A., Pihakski-Maunsbach, K., Sandal, N., Marcker, K. A., and Smith, A. G. 1998. Evidence that the plant host synthesizes the heme moiety of leghemoglobin in root nodules. *Plant Physiol.* 116:1259-1269.
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. 1999. A plant regulator controlling development of symbiotic root nodules. *Nature.* 402:191-195.
- Scheres, B., Van De Wiel, C., Zalensky, A., Horvath, B., Spaink, H., Van Eck, H., Zwartkruis, F., Wolters, A. M., Gloudemans, T., Van Kammen, A., and Bisseling, T. 1990. The ENOD12 gene product is involved in the infection process during the pea-*Rhizobium* interaction. *Cell.* 60:281-294.
- Shi, L., Twary, S. N., Yoshioka, H., Gregerson, R. G., Miller, S. S., Samac, D. A., Gantt, J. S., Unkefer, P. J., and Vance, C. P. 1997. Nitrogen assimilation in alfalfa: Isolation and characterization of an asparagine synthetase gene showing enhanced expression in root nodules and dark-adapted leaves. *Plant Cell.* 9:1339-1356.
- Stanford, A. C., Larsen, K., Barker, D. G., and Cullimore, J. V. 1993. Differential expression within the glutamine synthetase gene family of the model legume *Medicago truncatula*. *Plant Physiol.* 103:73-81.
- Stougaard, J. 2000. Regulators and regulation of legume root nodule development. *Plant Physiol.* 124:531-540.
- Suganuma, N., Okada, Y., and Kanayama, Y. 1997. Isolation of a cDNA for nodule-enhanced phosphoenolpyruvate carboxylase from pea and its expression in effective and plant-determined ineffective pea nodules. *J. Exp. Bot.* 48:1165-1173.
- Szczyglowski, K., Hamburger, D., Kapranov, P., and de Bruijn, F. J. 1997. Construction of a *Lotus japonicus* late nodulin expressed sequence tag library and identification of novel nodule-specific genes. *Plant Physiol.* 114:1335-1346.
- Szerszen, J. B., Szczyglowski, K., and Bandurski, R. S. 1994. *iaglu*, a gene from *Zea mays* involved in conjugation of growth hormone indole-3-acetic acid. *Science.* 265:1699-1701.
- Temple, S. J., Heard, J., Ganter, G., Dunn, K., and Sengupta-Gopalan, C. 1995. Characterization of a nodule-enhanced glutamine synthetase from alfalfa: Nucleotide sequence, *in situ* localization, and transcript analysis. *Mol. Plant-Microbe Interact.* 8:218-227.
- Thimm, O., Essigmann, B., Kloska, S., Altmann, T., and Buckhout, T. J. 2001. Response of *Arabidopsis* to iron deficiency stress as revealed by microarray analysis. *Plant Physiol.* 127:1030-1043.
- Thomas, S. G., Phillips, A. L., and Hedden, P. 1999. Molecular cloning and functional expression of gibberellin 2-oxidases, multifunctional enzymes involved in gibberellin deactivation. *Proc. Natl. Acad. Sci. U.S.A.* 96:4698-4703.
- Thummler, F., and Verma, D. P. 1987. Nodulin-100 of soybean is the subunit of sucrose synthase regulated by the availability of free heme in nodules. *J. Biol. Chem.* 262:14730-6.
- Tingey, S. V., Walker, E. L., and Coruzzi, G. M. 1987. Glutamine synthetase genes of pea encode distinct polypeptides which are differentially expressed in leaves, roots and nodules. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:1-9.
- Tsai, F. Y., and Coruzzi, G. M. 1990. Dark-induced and organ-specific expression of two asparagine synthetase genes in *Pisum sativum*. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:323-332.
- Udvardi, M. K., and Day, D. A. 1997. Metabolite transport across symbiotic membranes of legume nodules. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48:493-523.
- van Ghelue, M., Ribeiro, A., Solheim, B., Akkermans, A. D., Bisseling, T., and Pawlowski, K. 1996. Sucrose synthase and enolase expression in actinorhizal nodules of *Alnus glutinosa*: Comparison with legume nodules. *Mol. Gen. Genet.* 250:437-446.
- Vance, C. P., Reibach, P. H., and Pankhurst, C. E. 1987. Symbiotic properties of *Lotus pedunculatus* root nodules induced by *Rhizobium loti* and *Bradyrhizobium* sp. (Lotus). *Physiol. Plant.* 69:435-442.
- Vance, C. P., Gregerson, R. G., Robinson, D. L., Miller, S. S., and Gantt, J. S. 1994. Primary assimilation of nitrogen in alfalfa nodules: Molecular features of the enzymes involved. *Plant Sci.* 101:51-64.
- Verma, D. P. S., Hu, C.-A., and Zhang, M. 1992. Root nodule development: Origin, function, and regulation of nodulin genes. *Physiol. Plant.* 85:253-265.
- Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 15:1359-1367.
- Zhou, Z., Metcalf, A. E., Lovatt, C. J., and Hyman, B. C. 2000. Alfalfa (*Medicago sativa*) carbamoylphosphate synthetase gene structure records the deep lineage of plants. *Gene.* 243:105-114.