Plant Physiol. (1997) 113: 1081–1090

Novel, Highly Expressed Late Nodulin Gene (*LjNOD16*) from *Lotus japonicus*[1]

Philipp Kapranov, Frans J. de Bruijn*, and Krzysztof Szczyglowski

Michigan State University-Department of Energy Plant Research Laboratory (P.K., F.J.d.B., K.S.), Genetics Program (P.K., F.J.d.B.), and Department of Microbiology (F.J.d.B.), Michigan State University, East Lansing, Michigan 48824

We have isolated a *Lotus japonicus* cDNA corresponding to a highly abundant, late nodule-specific RNA species that encodes a polypeptide with a predicted molecular mass of 15.6 kD. The protein and its corresponding gene were designated Nlj16 and *LjNOD16*, respectively. *LjNOD16* was found to be expressed only in the infected cells of *L. japonicus* nodules. Related DNA sequences could be identified in the genomes of both *Glycine max* and *Medicago sativa*. In the latter, a homologous mRNA species was detected in the nodules. Unlike *LjNOD16*, its alfalfa homologs appear to represent low-abundance mRNA species. However, the proteins corresponding to the *LjNOD16* and its alfalfa homolog could be detected at similar levels in nodules but not in roots of both legume species. The predicted amino acid sequence analysis of nodulin Nlj16 revealed the presence of a long α-helical region and a positively charged C terminus. The former domain has a very high propensity to form a coiled-coil type structure, indicating that nodulin Nlj16 may interact with an as-yet-unidentified protein target(s) in the nodule-infected cells. Homology searches revealed no significant similarities to any known sequences in the databases, with the exception of two related, anonymous Arabidopsis expressed sequence tags.

Symbiotic nitrogen fixation is a unique example of a complex and subtly regulated biological process that takes place in a specialized plant organ—the nodule. Upon infection with specific strains of symbiotic bacteria belonging to the genus *Rhizobium*, *Bradyrhizobium*, or *Azorhizobium*, root cortical cells of legume plants undergo a dedifferentiation process that eventually leads to the formation of a nodule meristem. A highly organized and controlled series of events thereafter culminates in the formation of a fully functional, nitrogen-fixing nodule. Nodule ontogeny appears to be predominantly controlled by a plant morphogenetic program (for a recent review, see Gresshoff, 1993). Multiple signals, derived from both the host plant and the symbiotic bacteria, specify the induction and coordination of this organogenic process, which involves the activation of specific genes in both symbiotic partners (for review, see Schultze et al., 1994). A collection of genes encoding plant-nodule-specific proteins, collectively referred to as nodulin genes (van Kammen, 1984), has been isolated from various species of legumes (for review, see Mylona et al., 1995). These genes have been traditionally classified as early or late nodulin genes, reflecting the developmental time point of their expression (Nap and Bisseling, 1990). Induction of the early nodulin genes has been correlated with early morphogenetic processes such as preinfection, infection, and cortical cell division (Nap and Bisseling, 1990; Franssen et al., 1992; Cook et al., 1995). However, the functions and requirements for nodulation (Csanadi et al., 1994) of the early nodulin genes have not been elucidated.

Significant progress has been made in unraveling the cascade of both bacterial and plant signals, which initiate and direct early morphogenic events. The discovery of a set of flavonoid inducers of plant origin and *Rhizobium*-derived lipochito-oligosaccharide signal molecules known as Nod factors has led to exciting progress in this area (Horvath et al., 1987; Spaink et al., 1987; Lerouge et al., 1990; Spaink, 1992). Although research on the early stages of nodulation has progressed rapidly, the developmental cues and molecular events responsible for the final steps of nodule formation and functioning remain largely unknown (de Bruijn and Schell, 1992). These late stages include important events such as central nodule tissue formation, bacterial release from infection threads and plant cell colonization, production of peribacteroid membranes, bacteroid differentiation, and eventually, commencement of nitrogen fixation (Sprent, 1989). The latter process is accompanied by major molecular and biochemical alterations, which create and support the physiological requirements for nitrogen fixation and ammonia assimilation (Nap and Bisseling, 1990).

In spite of the highly complex nature of the late developmental events, only a limited number of late nodulin genes have been identified and characterized. Typical members of this group include genes encoding enzymes involved in specific biochemical pathways, e.g. the γ subunit of Gln synthetase involved in ammonia assimilation (Lara et al., 1983; Gebhardt et al., 1986; Bennett et al., 1989; Boron et al., 1989; Boron and Legocki, 1993), uricase II (Legocki and Verma, 1979; Bergmann et al., 1983), xanthine dehydrogenase (Triplett, 1985), Suc synthase (Thummler and Verma, 1987), peribacteroid membrane proteins (Fortin et al., 1985, 1987; Verma, 1992; Miao and Verma, 1993), leghemoglobins (Brisson et al., 1982), and a number of

[1] The research described in this manuscript was supported by a grant from the Department of Energy (DE-FG02-91ER20021).

* Corresponding author; e-mail debruijn@msu.edu; fax 1–517–353–9168.

Abbreviation: EST, expressed sequence tag.

proteins, the functions of which remain to be identified (Delauney and Verma, 1988). Recently, a symbiotically induced MADS-box containing gene (*nmh7*) has also been identified in alfalfa root nodules (Heard and Dunn, 1995). It has been suggested that the protein encoded by this gene could be involved in cellular activities specific to the differentiation of the infected cells (Heard and Dunn, 1995).

Although a number of late nodulin genes have been isolated, it appears likely that additional genes remain to be identified. Further cloning and detailed characterization of genes coding for novel late nodulins is crucial for understanding the molecular and biochemical details of nodule functioning and late morphogenic events. Therefore, we have initiated a systematic search for novel, late-nodule-specific transcripts in the model legume *Lotus japonicus* (Handberg and Stougaard, 1992; Jiang and Gresshoff, 1993) using the RNA differential display technology described by Liang and Pardee (1992; and also K. Szczyglowski, D. Hamburger, P. Kapranov, and F.J. de Bruijn, unpublished data). We report here the isolation and molecular characterization of *LjNOD16*, a novel, late-nodulin gene from *L. japonicus*.

MATERIALS AND METHODS

Plant Material

Lotus (*Lotus japonicus* GIFU B-129-S9) seeds were kindly provided by Dr. Jens Stougaard, Aarhus University, Denmark. Seeds were surface-sterilized and germinated as described by Handberg and Stougaard (1992). One-week-old *L. japonicus* seedlings were inoculated with *Rhizobium loti* strain NZP2235 (Jarvis et al., 1982) and transferred to pots (30 plants per pot) containing a 6:1 mixture of vermiculite and sand. All plants were grown in cabinets with a controlled environment: 18-h/6-h day/night cycle, light intensity of 246 μE s^{-1} m^{-2}, 22°C/18°C day/night temperature regime, and 70% RH. Both inoculated and uninoculated control plants were watered using nutrient solution (Broughton and Dilworth, 1971) containing 0.5 mM KNO_3. This low concentration of combined nitrogen supported growth of the uninfected control plants, but did not affect nodule formation on the inoculated *L. japonicus* plants (data not shown). For the initial stages (3, 7, and 11 d) root segments were harvested, whereas only nodules were collected for the 21-d stage. *Arabidopsis thaliana* Landsberg plants were kindly provided by Jacqueline Chernys from the laboratory of Dr. Hans Kende (Plant Research Laboratory, Michigan State University). *Medicago sativa* cv Cardinal seeds were surface-sterilized and germinated on 1% agar in water. Three-day-old seedlings were infected with *Rhizobium meliloti* strain 1021 and transferred to soil (3 parts vermiculite [grade 2]: 3 parts vermiculite [grade 3]: 1 part sand). *M. sativa* plants were grown under the same controlled environment as *L. japonicus* plants and watered with nutrient solution or distilled water. *M. sativa* plants were grown for 21 d before harvesting. The plant tissues were immediately frozen in liquid nitrogen and stored at −80°C until use.

Nucleic Acid Isolation, Southern and Northern Analyses

Plant genomic DNA was isolated following the procedure described by Rogers and Bendich (1988). Total RNA isolation was performed according to the method of Verwoerd et al. (1989), except that the extraction buffer was as described in Hall et al. (1978). The poly(A)$^+$ fraction of total RNA was isolated using oligo(dT)-cellulose spin column kits (5 Prime→3 Prime, Inc., Boulder, CO) following the manufacturer's instructions.

For Southern blot analysis, 10 μg of genomic DNA was completely digested using the appropriate restriction enzymes, separated on 0.8% agarose gels, and transferred to nylon filters according to standard procedures (Sambrook et al., 1989). Membranes were prehybridized and hybridized under high- (2× SSC) or low-stringency conditions (4× SSC), 0.5% SDS, 5× Denhard's solution, and 100 μg mL^{-1} of denatured sheared salmon sperm DNA at 65°C. Washes were carried out under low or high stringency, as described by Sambrook et al. (1989).

For northern analysis, 10 μg of total RNA was separated on 1.2% agarose-formaldehyde gels in Mops buffer, as described by Sambrook et al. (1989), and transferred to nitrocellulose membranes (Fisher Scientific). Prehybridization and hybridization reactions were performed in 0.5 M phosphate buffer, pH 7.2, 7% SDS, and 1% BSA at 65°C, according to the procedure described by Church and Gilbert (1984). Filters were washed twice for 15 min in 2× SSC, 0.1% SDS, and once for 15 min at 0.3 × SSC, 0.1% SDS at 65°C. Probes were labeled with [α-^{32}P]dATP using a random primer kit (Boehringer Mannheim) following the manufacturer's instructions.

RNA Differential Display

Differential display was carried out using mapping kits (RNAmap, GenHunter Corp., Brookline, MA), essentially following the manufacturer's instructions. cDNA synthesis was performed using 0.5 μg of total RNA isolated from 21-d-old *L. japonicus* roots, or nodules harvested 21 d after rhizobial infection. Differentially expressed bands were purified from the polyacrylamide gel and reamplified using the procedure described in the manual of the mapping kit. The DNA fragments were purified on a 1% agarose gel, blunt-ended using the Klenow fragment of DNA polymerase I, and cloned into the *Sma*I-digested vector pK18 (Pridmore, 1987).

cDNA Library Screening

The cDNA library from mature nodules of *L. japonicus* was kindly provided by Dr. Jens Stougaard, Aarhus, Denmark. The library was constructed with oligo-dT primer in lambda vector (UniZAP, Stratagene). Screening for full-copy cDNAs corresponding to the *PCR-5* product was performed following standard procedures (Sambrook et al., 1989; Stratagene manual).

DNA Sequencing and Computer Analyses

DNA sequencing was performed using a kit (Sequenase 2.0, United States Biochemical), according to the manufac-

turer's instructions. Computer analysis of DNA sequences was carried out using SeqEd software (Applied Biosystems). Analysis of predicted protein sequences was performed using the GCG (Genetics Computer Group, Madison, WI), PHDsec (Rost and Sander, 1993, 1994), and COILS version 2.2 (Lupas, 1996) programs. Homology searches were performed using BLAST software (Altschul et al., 1990).

In Situ Hybridization

Twenty-one-day-old nodules were fixed, dehydrated, and embedded into paraffin according to procedures described by Van de Wiel et al. (1990). Nodule sections of 7 μm were hybridized with digoxigenin-UTP labeled antisense and sense RNA probes, using the conditions reported by Engler et al. (1994). The full-copy cDNA encoding nodulin Nlj16 and a cDNA representing a *L. japonicus* leghemoglobin gene were used to prepare RNA probes using a labeling kit (DIG-RNA, Boehringer Mannheim) following the manufacturer's instructions. Washing and detection conditions were as described by Engler et al. (1994). Dry slides were mounted (Polymount, Polysciences Inc., Warrington, PA) and examined by dark- and bright-field microscopy (Axiophot microscope, Zeiss).

Production of Recombinant Nlj16 Protein in *Escherichia coli*

The 470-bp *Bgl*II-*Hin*cII fragment of the full-length *LjNOD16* cDNA, which contained the entire coding region of the nodulin *LjNOD16* gene, was blunt-ended using Klenow enzyme and cloned into dephosphorylated expression vector (pET 15B, Novagen, Madison, WI), which had been digested with *Xho*I and blunt-ended with Klenow enzyme. The construct with the insert cloned in the sense orientation was identified by digestion with *Xho*I and *Bam*HI. In-frame fusion of the coding region of *LjNOD16* to the N-terminal His tag of pET 15B was confirmed by DNA sequencing. The construct was transformed into *E. coli* strain BL21 (DE3, Novagen) for expression of the recombinant protein. The His-tag-Nlj16 fusion protein was purified under denaturing conditions with resin (His-Bind Resin, Novagen) following the manufacturer's instructions. The His tag was cleaved off by digestion with thrombin (Novagen) and the Nlj16 recombinant protein was further purified by passing it through the resin column after digestion. The flow-through and wash fractions were collected. Nlj16 was refolded by dialyzing against PBS (Sambrook et al., 1989) overnight. The refolded protein was mixed with adjuvant (TiterMax, CytRx Corp., Norcross, GA) and used for immunization of rabbits following the instructions in the manual.

Western Blot Analysis

The total protein extracts from *L. japonicus* and *M. sativa* roots and nodules were obtained by grinding tissues in liquid nitrogen and boiling for 5 min in extraction buffer (100 mM Tris, pH 6.8, 5% SDS, 0.5% β-mercaptoethanol). For western blot analysis, 50 μg of total protein was separated by SDS-PAGE (Laemmli, 1970) using 15% acrylamide gels, and electroblotted overnight onto nitrocellulose membranes (Protran, Schleicher & Schuell) in Towbin buffer (Towbin et al., 1979) supplemented with 0.05% SDS. Blocking, binding, and washes were performed in PBS supplemented with 0.3% Tween 20 (Sigma). Preimmune and immune sera from a rabbit immunized with Nlj16 recombinant protein were used in a 1:2000 dilution. Antibody detection was performed using goat anti-rabbit antibodies conjugated to alkaline phosphatase (Kirkegaard and Perry Laboratories, Gaithersburg, MD).

RESULTS

Identification and Cloning of the *L. japonicus* PCR-5 cDNA

The recent development of arbitrarily primed RNA fingerprinting techniques, also known as RNA differential display, has resulted in an experimental tool with exciting potential for the detection of differential gene expression during complex biological processes (Liang and Pardee, 1992; Welsh et al., 1992). We have employed this technique to construct a library of *L. japonicus* late-nodulin ESTs (K. Szczyglowski, D. Hamburger, P. Kapranov, and F.J. de Bruijn, unpublished data). Our experiments involved a simple comparison of RNA profiles from two samples: mature, 21-d uninfected roots and fully developed nodules harvested 21 d after infection. Thirteen different combinations of PCR primers were initially used. Five bands appeared to be differentially expressed on display gels. These were excised, reamplified by PCR, and analyzed using northern blot analysis and DNA sequencing. Three out of five differentially displayed cDNA products corresponded to nodule-specific mRNAs, as determined by northern blot analysis (data not shown). One of the differentially expressed cDNAs, designated PCR-5, was found to be 607 bp long and showed no significant homology to DNA sequences in the databases (data not shown). We chose this cDNA for further detailed molecular characterization.

Northern blot analysis was used to correlate specific phases in nodule development with the expression pattern of the gene corresponding to the PCR-5 cDNA. This experiment revealed that the PCR-5 product corresponded to a highly abundant late-nodulin mRNA species. The mRNA corresponding to PCR-5 was found to be detectable between 7 and 11 d after infection, and to accumulate gradually to a high level in fully developed, 21-d-old nodules (see Fig. 1). The observed kinetics of mRNA accumulation resembled very closely the pattern of *L. japonicus* leghemoglobin gene expression (data not shown), justifying the categorization of this gene as a late-nodulin gene.

PCR-5 cDNA Corresponds to the Gene Encoding *L. japonicus* Nodulin Nlj16

PCR-5 cDNA was used as a probe to screen a *L. japonicus* nodule-specific cDNA library. Approximately 40 hybridizing cDNA clones were purified. DNA sequence analysis of the six longest cDNA clones (ranging from 790 to 1010 bp) showed that they all carried identical DNA sequences, but

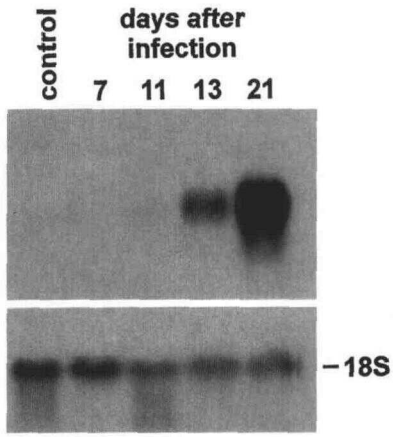

Figure 1. Northern blot analysis of PCR5 expression. Ten micrograms of total RNA isolated from 21-d-old uninfected roots (control) and root segments and nodules harvested 7, 11, 13, and 21 d after infection was analyzed using the radiolabeled PCR-5 product as a probe (top). The blot was reprobed with 18S rDNA as an RNA loading control (bottom).

had different lengths of 3′-nontranslated regions. Six different sites of poly(A)$^+$ addition were identified in the cDNAs. The sequence of the longest cDNA clone is shown in Figure 2A (accession no. U64964). The different polyadenylation sites, derived from the sequences of the corresponding cDNA clones, are as indicated. Further analysis of the DNA sequence of the longest cDNA clone revealed that it contained a 41-bp, 5′ untranslated region, followed by an open reading frame 423 nucleotides long, and a 546-bp, 3′ untranslated region. The stop codon (TAA) was found to be located 6 bp upstream of the putative AUG initiation triplet, indicating that the cDNA contains the entire coding region (Fig. 2A). The deduced protein sequence was found to correspond to a polypeptide of 141 amino acid residues and a molecular weight of approximately 15.6 kD. Therefore, we propose that our cDNA represents a late *L. japonicus* nodule-specific gene, designated *LjNOD16*, encoding a 15.6-kD nodule-specific protein (nodulin Nlj16).

Nodulin Nlj16 *Contains a Long* α*-Helical Domain and Is Homologous to Two Anonymous ESTs from* ***A. thaliana***

A hydropathy profile of nodulin Nlj16 was derived using the algorithm of Kyte and Doolittle (1982) and LASERGENE software (DNASTAR, Madison, WI). This analysis indicated that the protein was mostly hydrophilic, except for a hydrophobic stretch of 16 amino acids, with tetrad repeats of Pro residues at the N-terminal end (Fig. 2B; see also Fig. 2A). Secondary structure analysis using the PHDsec program (Rost and Sander, 1993, 1994) predicted with a high probability (7–9 on a scale from 0 to 10) the presence of two adjacent α-helical regions divided by a loop or turn (Fig. 2A). The relatively long α-helical domain, spanning 80 out of 141 amino acid residues of the Nlj16 protein, would be expected to form a coiled-coil-type protein structure, as predicted by the COILS version 2.2 program (Lupas, 1996). A homology search using the BLAST algorithm (Altschul et al., 1990) revealed no significant similarities to any gene with known function, except for a limited homology to two anonymous Arabidopsis ESTs: EST168K8 (accession no. R64923) and 110G16 (accession no. T42081). The corresponding *A. thaliana* cDNA clones were obtained from Dr. Tom Newman's laboratory (Michigan State University-Department of Energy Plant Research Laboratory) and from the Arabidopsis Biological Resource Center (Ohio State University), respectively, and their entire DNA sequence was determined. The complete DNA sequences for ESTs 168K8 and 110G16 have been deposited in GenBank under accession nos. U64965 and U64966, respectively. The deduced amino acid sequences of EST 168K8 (751 bp; *A. thaliana* EST1; Fig. 3) and EST 110G16 (644 bp; *A. thaliana* EST2; Fig. 3), which represent 200 and 150 amino acids, respectively, from the C-terminal ends of the proteins, showed a high level of homology to each other throughout the whole amino acid sequence (Fig. 3). In addition, a significant similarity between the *A. thaliana* ESTs and the amino acid sequence of the Nlj16 protein, especially in its α-helical domain, was found (Fig. 3). The region of the 168K8 EST with the highest similarity to the α-helical domain of Nlj16 also had a high probability to form two α-helices and to participate in generating coiled-coil structures, as predicted by the PHDsec (Rost and Sander, 1993,

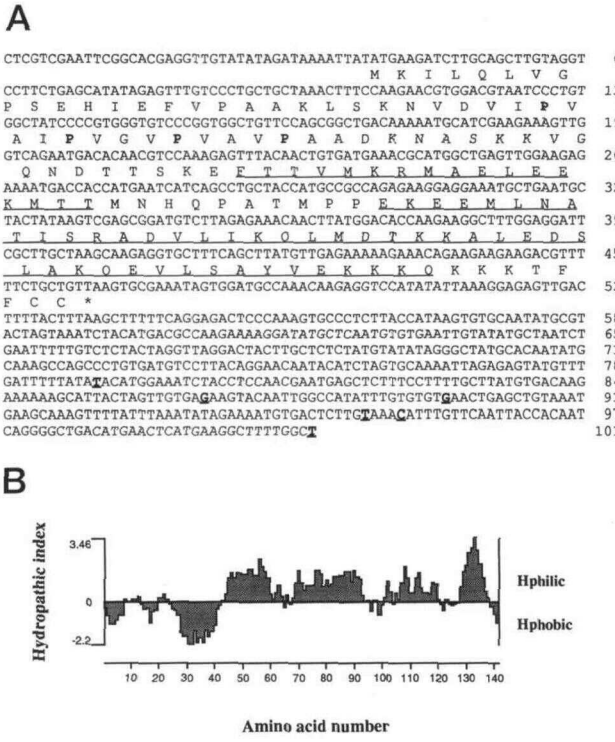

Figure 2. Nucleotide and deduced amino acid sequences of the *LjNOD16* cDNA. A, The positions of the multiple polyadenylation sites are indicated by the bold underlined nucleotides. Predicted α-helical domains of nodulin Nlj16 are underlined, and the four prolines of the N-terminal hydrophobic stretch are indicated in bold. The asterisk denotes the predicted stop codon. B, Hydropathy profile of nodulin Nlj16 determined by the method of Kyte and Doolittle (1982) with a window of seven residues.

```
          1                                                    50
Nlj16     ..........  .......MKI LQLVGPSEHI EFVPAAKLSK NVDVIPVAIP
                              ..  :| |:..: :   ...:|  . .:.. .:.:
A.t.EST1  DEYVPMVDKA VDATWKVKPA IQRVASRGAL MSPTVPKDHE GIKARVLVMF
                                |.|.:|:||  |.|.|||||| ::.|||||
A.t.EST2  .......... .......... INRAPSKGAH MPPNVPKDHE SFSARVLVTF

          51                                                   100
Nlj16     VGVPVAV.....................PAADK NASKKVGQND TTSKE.....
          : :. :||                     ||... ... .:| |: .. ||
A.t.EST1  MAFLMAVFTF FRTVTKKLPA TTTSSPAETQ GNAIELGSNG EGVKEECRPP
          |||:||::|| ||||..::   .|.  |:..  . .|| :..: |
A.t.EST2  MAFVMAILTF FRTVSNRV.. VTKQLPPPPS QPQIEGSAAA E.........

          101                                                  150
Nlj16     .......... .FTTVMKRMA ELEEKMTTMN HQPATMPPEK EEMLNATISR
                     :..| |::. |||:|:.|:.  .|..|| || ||:|||.::|
A.t.EST1  SPVPDLTETD LLNCVTKKLT ELEGKIGTLQ SKPNEMPYEK EELLNAAVCR
                |.| |||:| |||| |||:|||.|| |||.|||||| ||||||||||
A.t.EST2  .......EAD LLNSVLKKLT ELEEKIGALQ SKPSEMPYEK EELLNAAVCR


          151                                                  200
Nlj16     ADVLEKQLMD TKKALEDSLA KQEVLSAYVE K......... KKQKKKTFFC C
          .|.|| :|:. ||||| :.|  :|| | ||:: :          ||.||| :||
A.t.EST1  VDALEAELIA TKKALYEALM RQEELLAYID RQEEAQFQKM KKKKKKHLFC F
          |||||||||| |||||||||| |||||||||| |||.|| |    ||.|:|::|| |
A.t.EST2  VDALEAELIA TKKALYEALM RQEELLAYID RQEAAQHQ.. KKNKRKQMFC F
```

Figure 3. Alignment of Nlj16 with the deduced amino acid sequences of Arabidopsis ESTs 168K8 (*A. thaliana* EST1) and 110G16 (*A. thaliana* EST2). The alignment was performed using the GCG package (Genetics Computer Group). Vertical bars indicate identical amino acids, semicolons indicate conservative substitutions, and dots indicate semiconservative substitutions.

1994) and COILS (Lupas, 1996) programs. A computer search of the PROSITE and BLOCKS databases did not reveal any distinct conserved amino acid motifs in nodulin Nlj16 or in the protein sequences derived from the Arabidopsis ESTs.

Tissue-Specific Expression of the *LjNOD16* Gene

To gain further insight into the tissue specificity of *LjNOD16* gene expression, total RNA isolated from *L. japonicus* flower, leaf, stem, root, and fully developed nodule tissues was analyzed using northern blot hybridization (Fig. 4). The *LjNOD16* gene was found to be expressed exclusively in the nodule tissues, as revealed after a brief, 7.5-h exposure of the blot (data not shown). However, prolonged exposure of the blot over a period of 7 d revealed the presence of a very low level of hybridizing mRNA species in uninfected roots and flowers. In the latter case, an mRNA species of approximately 2500 nucleotides could be detected, as opposed to the 900- to 1200-nucleotide-long transcripts consistently found in *L. japonicus* nodules. Therefore, we conclude that the *LjNOD16* gene represents a nodulin gene, and that a related gene(s) with a different expression pattern may exist in the *L. japonicus* genome. Northern blot analysis was also performed using total RNA isolated from various organs of *A. thaliana* plants, and *A. thaliana* EST1 or *A. thaliana* EST2 cDNAs were used as probes. An example of this analysis using *A. thaliana* EST1 is shown in Figure 5. Both ESTs hybridized to an approximately 3000-nucleotide-long mRNA in all tissues tested.

Organization of *LjNOD16* Gene in Genomes of *L. japonicus* and Other Legume Plants

Southern blot analysis was performed to determine the number of *LjNOD16* genes in the *L. japonicus* genome. Genomic DNA isolated from *L. japonicus* plants was digested to completion using four different restriction enzymes and hybridized with an [α-^{32}P]dATP-labeled probe, representing the PCR-5 cDNA sequence. Under high-stringency conditions (see "Materials and Methods"), a single hybridizing band was observed for three out of four different restriction digests (note that a *Hin*dIII site is present in the PCR-5 cDNA sequence), indicating that the *LjNOD16* gene is represented by a single- or low-copy gene number in the *L. japonicus* genome (Fig. 6A). To determine whether a gene(s) homologous to *L. japonicus LjNOD16* was present in the genomes of other legumes, a Southern hybridization was performed using the full-length cDNA for *LjNOD16* as a probe. Several hybridizing bands were observed in genomic DNA isolated from *Lotus corniculatus*, *Glycine max*, and *M. sativa* plants. Three additional hybridizing fragments were detected in the *L. japonicus* genome under moderate stringency conditions using the full-length cDNA as a probe (Fig. 6B). The latter results indicate the presence of *LjNOD16*-related genes in the *L. japonicus* genome, and support the conclusion reached above based on the northern blot analysis (see Fig. 4).

Localization of *LjNOD16* Transcripts in the *L. japonicus* Nodule

Cellular localization of the *LjNOD16* mRNA in nodules was analyzed using nonradioactive in situ hybridization. Semithin sections (7 μm) of *L. japonicus* 21-d-old nodules were hybridized with digoxigenin-UTP-labeled RNA probes transcribed from the *LjNOD16* and *L. japonicus* leghemoglobin cDNAs. Both sense and antisense tran-

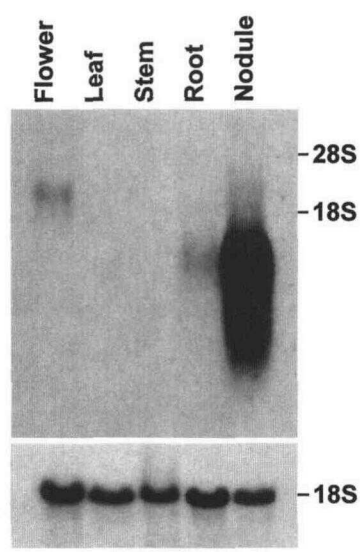

Figure 4. Expression of *LjNOD16* in different tissues of *L. japonicus*. Ten micrograms of total RNA from *L. japonicus* flowers, leaves, stems, uninfected roots, and nodules was probed with radiolabeled *LjNOD16* cDNA. The blot was exposed for 7 d to visualize the less intensely hybridizing mRNA species from flowers and uninfected roots, respectively (top). Equal RNA loading was verified by probing the same filter with the 18S DNA probe (bottom). Positions of 28S and 18S rRNAs are as indicated.

scripts were used as probes. Upon examination of several consecutive sections derived from different *L. japonicus* nodules, we observed that expression of *LjNOD16* was restricted to the infected cells of the nodule (Fig. 7, B and C), which closely resembled the expression pattern of the *L. japonicus* leghemoglobin mRNA (Fig. 7D). No expression of *LjNOD16* was detected in uninfected interstitial cells, nor in peripheral tissues such as the cortex and endodermis. In addition, no hybridization signals were observed when digoxigenin-UTP-labeled sense RNA transcripts corresponding to *LjNOD16* were used as a probe (Fig. 7A).

The *M. sativa LjNOD16* Homolog Is Expressed Specifically in Nodules

Because the *L. japonicus LjNOD16* cDNA probe hybridized to DNA sequences in the *M. sativa* genome, we investigated the tissue-specific expression of the corresponding gene(s) using northern blot analysis. Ten micrograms of the total RNA isolated from different *M. sativa* tissues was hybridized with the radiolabeled *L. japonicus LjNOD16* cDNA insert. Unexpectedly, no detectable signal could be found in any of the *M. sativa* tissues analyzed when total RNA was used (data not shown). We hypothesized, therefore, that the corresponding gene(s) might be expressed at a low level in the indeterminate nodules of *M. sativa* plants. We tested this assumption using 4 μg of poly(A)$^+$ RNA derived from leaves, stems, roots, and nodules of the same 21-d-old *M. sativa* plant (Fig. 8). Two hybridizing bands, corresponding to mRNA species of approximately 800 and 1800 nucleotides in length, respectively, could be detected specifically in nodules, but not in any other *M. sativa* tissues analyzed.

The *L. japonicus* Nlj16 Protein and Its *M. sativa* Homolog Can Be Detected in Nodules but Not in Roots

Polyclonal antibodies were raised against the recombinant *L. japonicus* Nlj16 protein and used to carry out a western blot analysis using total protein extracts from roots and nodules of *L. japonicus* and *M. sativa* plants (Fig. 9). The recombinant Nlj16 protein was loaded on the SDS-PAGE gel as a positive control. Five to six protein bands in the range of 16 to 19 kD could be consistently detected in the extract derived from *L. japonicus* nodules, but not from *L. japonicus* roots (Fig. 9A). In comparison, a 10-kD protein was specifically detected in *M. sativa* nodule extract (Fig. 9A). Intense protein bands were also observed in *M. sativa* root extracts. Because they could also be detected using the preimmune serum, it is highly unlikely that they reflect specific reactions with the anti-LjN16 antibody (Fig. 9B). The preimmune serum failed to detect the control recombinant Nlj16 protein. In addition, neither the *L. japonicus* 16- to 19-kD protein bands nor the 10-kD protein from *M. sativa* nodule extracts were recognized by preimmune serum (Fig. 9B). The same nodule proteins were, however, specifically detected with the antibodies purified from the anti-Nlj16 serum by affinity chromatography over a column with a recombinant Nlj16 protein (data not shown). On the other hand, the immunodepleted fraction of the serum could no longer efficiently recognize the 16- to 19-kD protein bands in *L. japonicus* nodules (data not shown). Based on these results, we conclude that the proteins detected specifically in the nodule extracts of both plants analyzed are likely to correspond to nodulin Nlj16 and its *M. sativa* homolog, respectively.

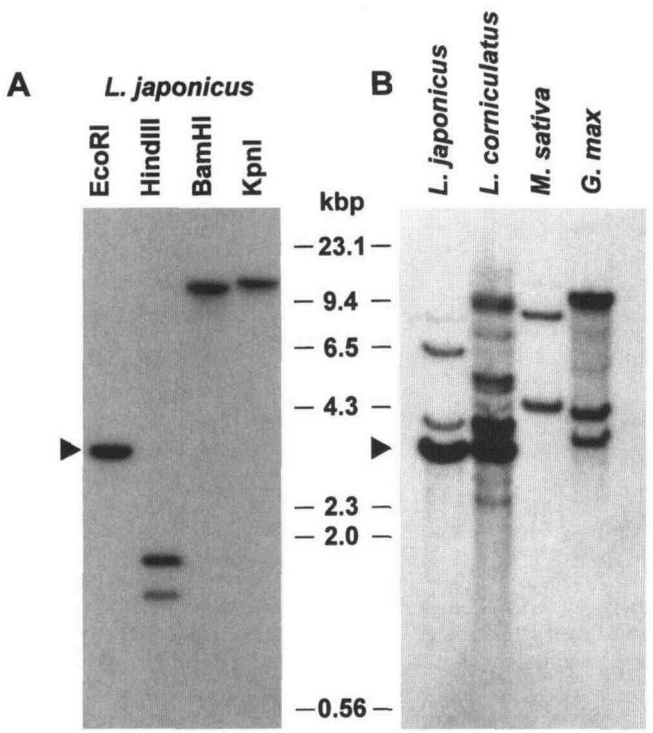

Figure 6. Organization of the *LjNOD16* gene in the genomes of *L. japonicus* and other legume species. A, Ten micrograms of *L. japonicus* genomic DNA was completely digested with the restriction enzymes indicated, separated on 0.8% agarose gel, and probed with [α-^{32}P]dATP labeled PCR-5 DNA. Hybridization and washes were performed under high-stringency conditions (see "Materials and Methods"). B, Southern blot hybridization with *Eco*RI digested DNA (10 μg) from *L. japonicus* cv Gifu, *L. corniculatus* cv Rodeo, *M. sativa* cv Cardinal, and *G. max* cv Dimon. A full-length *LjNOD16* cDNA was used as a probe. Hybridization was performed under low-stringency conditions (see "Materials and Methods"). The *Eco*RI fragment (approximately 4.0 kb) corresponding to the *LjNOD16* gene is marked with an arrowhead.

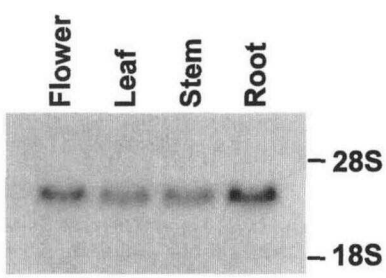

Figure 5. Expression pattern of *A. thaliana* EST 168K8. Ten micrograms of total RNA from flowers, leaves, stems, and roots of *A. thaliana* was separated by agarose-formaldehyde gel, transferred to nitrocellulose, and hybridized to 168K8 DNA labeled with [α-^{32}P]dATP. Positions of 28S and 18S rRNAs are as indicated.

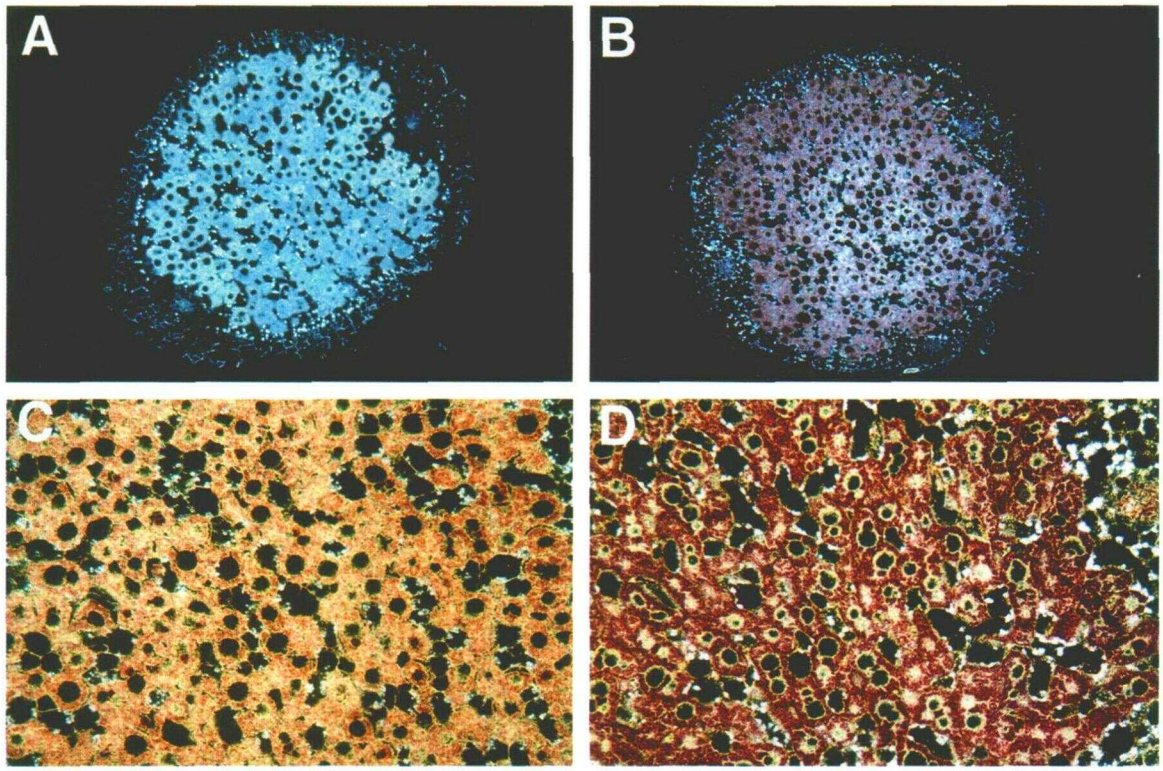

Figure 7. In situ localization of *LjNOD16* and leghemoglobin transcripts in sections of 21-d-old *L. japonicus* nodules. Dark-field micrographs are shown in which the hybridization signal appears as pink color. A, Section hybridized with sense *LjNOD16* RNA probe. B, Section hybridized with the anti-sense *LjNOD16* RNA probe. C, Detailed view of a region containing both infected and uninfected cells from the section shown in B. D, Detailed view of a region containing both infected and uninfected cells from a section hybridized with anti-sense probe corresponding to the *L. japonicus* leghemoglobin gene.

DISCUSSION

We have isolated and characterized a cDNA clone from the model legume *L. japonicus*, which represents a novel, late-nodulin gene designated *LjNOD16*. One intriguing feature of the *LjNOD16* is its very high level of nodule-specific transcriptional activity, which is similar to the transcriptional activity of the leghemoglobin genes in *L. japonicus* nodules. This property inspired us to perform a more detailed molecular analysis of the *LjNOD16* gene. The expression pattern of this gene, based on the developmental northern blot analysis, showed distinct characteristics of late-nodulin genes. The corresponding mRNA was first detectable in the *L. japonicus* roots around 11 d after *Rhizobium* infection and accumulated to a very high level in fully developed nodules. This expression pattern may suggest that the product of the *LjNOD16* gene is involved in relatively late stages of nodule ontogeny and/or nodule functioning. This assumption was further supported by the localization of transcripts corresponding to the *LjNOD16* gene in infected cells of *L. japonicus* nodules. Infected-cell-specific expression has also been demonstrated for leghemoglobin (*lb*) genes (Szczyglowski et al., 1994). Because in situ localization experiments were performed using fully developed *L. japonicus* nodules, it is not clear whether the induction of *LjNOD16* gene expression precedes bacterial colonization of the plant cells, or occurs concomitantly with the release of symbiotic bacteria from the infection threads and commencement of nitrogen fixation. The latter expression pattern is characteristic of the majority of late-nodulin genes identified thus far (Govers et al., 1987; Nap and Bisseling, 1990). Taking into account the time point after *Rhizobium* infection when the *LjNOD16* gene becomes activated, we propose that it is a late-nodulin gene coordinately induced just prior to or concomitant with the commencement of nitrogenase activity.

Northern blot analysis using different *L. japonicus* tissues revealed the presence of very low levels of mRNAs hybridizing to *LjNOD16* in uninfected roots and flowers, in addition to nodules. A faint band corresponding to the size of the *LjNOD16* mRNA was detected in the uninfected roots after a prolonged exposure of the blot. A different mRNA species of approximately 2500 nucleotides was detected in flowers, suggesting that there may be more than one gene related to *LjNOD16* in the *L. japonicus* genome. This assumption was supported by the data obtained from our Southern blot analyses. Several hybridizing bands of different intensities were observed under moderately stringent conditions. However, when high-stringency hybridization conditions were used, and the 3′-portion of the *LjNOD16* cDNA (PCR5) was employed as a probe, single

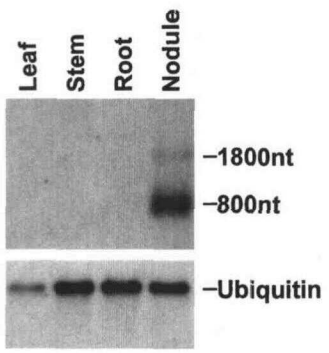

Figure 8. Expression of putative homolog(s) of *LjNOD16* in different tissues of *M. sativa*. Four micrograms of the poly $(A)^+$ fraction of total RNA isolated from leaves, stems, roots, and nodules of 21-d-old nodulated alfalfa plants was separated on 1.2% agarose-formaldehyde gel, blotted on nitrocellulose, and probed with radiolabeled *LjNOD16* cDNA fragment. Two hybridizing bands with approximate lengths of 800 and 1800 nucleotides (nt) detected under low-stringency conditions (see "Materials and Methods") are shown. The blot was stripped and rehybridized with a ubiquitin cDNA from *Sesbania rostrata* as a loading control.

hybridizing bands were detected, indicating that the *LjNOD16* gene is likely to be represented by a single- or low-copy gene in the *L. japonicus* genome. A full-copy *LjNOD16* cDNA hybridized with several genomic fragments derived from different legume species. This result suggests that the sequences related to the *L. japonicus LjNOD16* gene are present in the genomes of different legume plants, and may play a role in both determinate and indeterminate nodules. This assumption was supported by the results of the northern analysis using mRNAs derived from different *M. sativa* tissues. The *LjNOD16* cDNA probe hybridized specifically with two distinct, nodule-specific mRNA species, indicating that homologous transcripts are present in the indeterminate nodules of *M. sativa* plants.

The level of alfalfa mRNAs appeared to be significantly lower than the level of *LjNOD16* mRNA in *L. japonicus* nodules. In the former case, 4 μg of poly$(A)^+$ RNA and a long exposure time of 7 to 10 d produced detectable hybridization signals, clearly contrasting with the situation observed in the *L. japonicus* nodules. The apparent low abundance of alfalfa *NOD16* mRNAs is not likely to be due to the fact that a heterologous *L. japonicus LjNOD16* probe was used, because the same DNA insert gave a clearly detectable signal when hybridized with *M. sativa* genomic DNA (Fig. 6B). Although this remains to be proven, it is possible that a different mode of gene regulation may account for the observed difference in the abundance of homologous mRNAs in nodules of *L. japonicus* and *M. sativa* plants. It is noteworthy here that in the case of the *L. japonicus* nodule extract, an array of closely migrating proteins was consistently detected using anti-Nlj16 antibody. Although the basis for this phenomenon is not clear, it may reflect the presence of different forms of Nlj16 protein in the *L. japonicus* nodules, perhaps due to posttranslational modifications.

A detailed analysis of the amino acid sequence of the *L. japonicus* protein, deduced from the full-length cDNA sequence, revealed the presence of several interesting domains. Nlj16 was predicted to be a soluble protein, as suggested by Kyte-Doolittle hydropathy analysis. However, Nlj16 contains a hydrophobic domain, consisting of regularly spaced prolines, the function of which is not clear. Secondary structure predictions strongly indicated the presence of an extended α-helical domain, in which two α-helices are separated by a putative turn. Homology searches using the predicted amino acid sequence of Nlj16 revealed a significant similarity to two closely related anonymous Arabidopsis ESTs mostly in the predicted α-helical domain. Both *L. japonicus* and Arabidopsis α-helical sequences showed a very high propensity to form a coiled-coil-type structure, which may indicate that the nodulin Nlj16, as well as the Arabidopsis proteins, may interact with as yet unknown proteins in the plant cells.

With regard to the size of corresponding mRNAs, the Arabidopsis ESTs seem to represent a counterpart of the flower-specific *L. japonicus* mRNA. The mRNA corresponding to the nodule-specific gene is significantly smaller than either Arabidopsis or *L. japonicus* flower mRNA species. It is tempting, therefore, to speculate that the nodulin Nlj16 may represent a truncated version of the otherwise constitutively expressed proteins (as exemplified by the ubiquitous expression of the Arabidopsis ESTs), primarily harboring the conserved domain, which was adopted during evolution for specific symbiotic function(s). However, the exact role of nodulin Nlj16 in the infected cells of *L. japonicus* nodules remains to be determined.

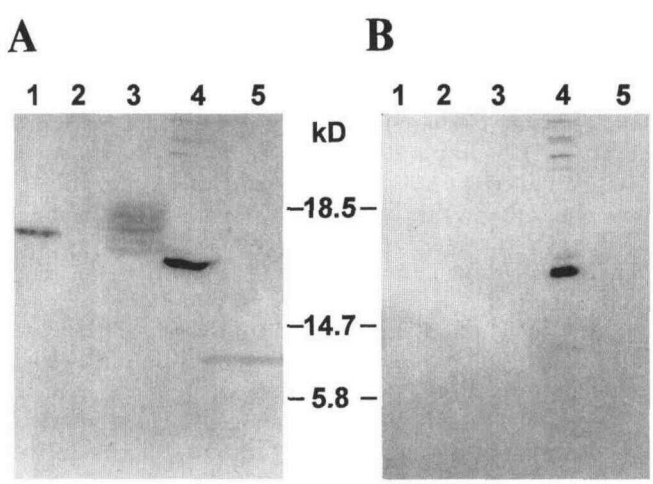

Figure 9. Detection of Nlj16 protein and its putative alfalfa homolog in nodules of *L. japonicus* and *M. sativa*. Lane 1, Recombinant Nlj16 protein (12.5 ng); lanes 2 and 4, total protein (50 μg) extracted from roots of *L. japonicus* and *M. sativa* plants, respectively; lanes 3 and 5, total protein extract (50 μg) from nodules of *L. japonicus* and *M. sativa* plants, respectively. A, Western blot analysis with anti-Nlj16 serum. B, Western blot analysis with preimmune serum.

ACKNOWLEDGMENTS

During the course of this work, K.S. visited the laboratory of Dr. Ton Bisseling. We thank Drs. Katharina Pawlowski and Ton Bisseling for their help and many useful suggestions with regard to the in situ hybridization procedure. We wish to thank Dr. Marcelle Holsters from the University of Gent, Belgium, for kindly providing the ubiquitin clone from *S. rostrata*. We thank Drs. Lee McIntosh, John Wilson, and Leslie Kuhn for helpful discussions. We thank Scott Shaw for the preparation of the plant material used in this work. We also thank Kurt Stepnitz and Marlene Cameron for their help in preparing the figures, Karen Bird for expert editorial assistance, and Susan Fujimoto, David Silver, Dirk Hamburger, and Rujin Chen from our laboratory for their advice and helpful comments on the manuscript.

Received October 15, 1996; accepted January 8, 1997.
Copyright Clearance Center: 0032–0889/97/113/1081/10.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Bennett MJ, Lighftoot DA, Cullimore JV (1989) cDNA sequence and differential expression of the gene encoding the glutamine synthetase γ-polypeptide of *Phaseolus vulgaris* L. Plant Mol Biol **12:** 553–565

Bergmann H, Preddie E, Verma DPS (1983) Nodulin-35: a subunit of specific uricase (uricaseII) induced and localized in the uninfected cells of soybean nodules. EMBO J **2:** 2333–2339

Boron L, Legocki AB (1993) Cloning and characterization of a nodule-enhanced glutamine synthetase-encoding gene from *Lupinus luteus*. Gene **136:** 95–102

Boron L, Szczyglowski K, Konieczny A, Legocki AB (1989) Glutamine synthetase in *Lupinus luteus*: identification and preliminary characterization of nodule-specific cDNA clone. Acta Biochim Polon **36:** 295–301

Brisson NA, Pombo-Gentile A, Verma DPS (1982) Organization and expression of leghemoglobin genes. Can J Bot **60:** 272–278

Broughton WJ, Dilworth MY (1971) Control of leghemoglobin synthesis in snake beans. Biochem J **125:** 1075–1080

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA **81:** 1991–1995

Cook D, Dreyer D, Bonnet D, Howell M, Nony N, VandenBosch K (1995) Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. Plant Cell **7:** 43–55

Csanadi G, Szecsi J, Kalo P, Kiss P, Endre G, Kondorosi A, Kondorosi E, Kiss GB (1994) ENOD12, an early nodulin gene, is not required for nodule formation and efficient nitrogen fixation in alfalfa. Plant Cell **6:** 201–213

de Bruijn FJ, Schell J (1992) Regulation of plant genes specifically induced in developing and mature nitrogen-fixing nodules: *cis*-acting elements and *trans*-acting factors. *In* DPS Verma, ed, Control of Plant Gene Expression. CRC Press, Boca Raton, FL, pp 241–258

Delauney AJ, Verma DPS (1988) Cloned nodulin genes for symbiotic nitrogen fixation. Plant Mol Biol Rep **6:** 279–285

Engler JA, Van Montagu M, Engler G (1994) Hybridization in situ of whole mount messenger RNA in plants. Plant Mol Biol Rep **12:** 321–331

Fortin MG, Morison NA, Verma DPS (1987) Nodulin-26, a peribacteroid membrane nodulin is expressed independently of the development of the peribacteroid compartment. Nucleic Acids Res **15:** 813–824

Fortin MG, Zalechowska M, Verma DPS (1985) Specific targeting of membrane nodulins to the bacteroid enclosing compartment in soybean nodules. EMBO J **4:** 3041–3046

Franssen HJ, Vijn I, Yang WC, Bisseling T (1992) Developmental aspects of the *Rhizobium*-legume symbiosis. Plant Mol Biol **19:** 89–107

Gebhardt C, Oliver JE, Forde BG, Saarelainen R, Miflin BJ (1986) Primary structure and differential expression of glutamine synthetase genes in nodules, roots and leaves of *Phaseolus vulgaris*. EMBO J **5:** 1429–1435

Govers F, Nap JP, van Kammen A, Bisseling T (1987) Nodulins in the developing root nodule. Plant Physiol Biochem **25:** 309–322

Gresshoff PM (1993) Molecular genetic analysis of nodulation genes in soybean. Plant Breeding Reviews **11:** 275–318

Hall TC, Ma Y, Buchbinder BU, Pyne JW, Sun SM, Bliss FA (1978) Messenger RNA for G1 protein of French bean seeds: cell-free translation and product characterization. Proc Natl Acad Sci USA **75:** 3196–3200

Handberg K, Stougaard J (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. Plant J **2:** 487–496

Heard J, Dunn K (1995) Symbiotic induction of a MADS-box gene during development of alfalfa root nodules. Proc Natl Acad Sci USA **92:** 5273–5277

Horvath B, Bachem CWB, Schell J, Kondorosi A (1987) Host-specific regulation of nodulation genes in *Rhizobium* is mediated by the plant signal, interacting with the *nodD* gene product. EMBO J **6:** 841–848

Jarvis BDW, Pankhurst CE, Patel JJ (1982) *Rhizobium loti*, a new species of legume root nodule bacteria. Int J Syst Bact **32:** 378–380

Jiang Q, Gresshoff PM (1993) *Lotus japonicus*: a model plant for structure-function analysis in nodulation and nitrogen fixation. Curr Topics Plant Mol Biol **2:** 97–110

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105–132

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680–685

Lara M, Cullimore JV, Lea PJ, Johnston AWB, Lamb JW (1983) Appearance of a novel form of plant glutamine synthetase during nodule development in *Phaseolus vulgaris*. L. Planta **157:** 254–258

Legocki RP, Verma DPS (1979) A nodule-specific plant protein (Nodulin-35) from soybean. Science **205:** 190–193

Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Prome JC, Dénarié J (1990) Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulfated and acylated glucosamine oligosaccharide signal. Nature **344:** 781–784

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science **257:** 967–971

Lupas A (1996) Prediction and analysis of coiled-coil structures. Methods Enzymol **266:** 513–525

Miao G-H, Verma DPS (1993) Soybean nodulin-26 gene encoding a channel protein is expressed only in the infected cells of nodules and is regulated differently in roots of homologous and heterologous plants. Plant Cell **5:** 781–794

Mylona P, Pawlowski K, Bisseling T (1995) Symbiotic nitrogen fixation. Plant Cell **7:** 869–885

Nap J-P, Bisseling T (1990) Developmental biology of a plant-prokaryote symbiosis: the legume root nodule. Science **250:** 948–954

Pridmore RD (1987) New and versatile cloning vectors with kanamycin resistance marker. Gene **56:** 309–312

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues. *In* SB Gelvin, RA Schilperoort, eds, Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, The Netherlands, **A6** 1–10

Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol **232:** 584–599

Rost B, Sander C (1994) Combining evolutionary information and neural networks to predict protein secondary structure. Proteins **19:** 55–77

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schultze M, Kondorosi E, Ratet P, Buire M, Kondorosi A (1994) Cell and molecular biology of *Rhizobium*-plant interaction. Int Rev Cytol **156:** 1–75

Spaink HP (1992) Rhizobial lipooligosaccharides: answers and questions. Plant Mol Biol **20:** 977–986

Spaink HP, Wijffelman CA, Pees E, Okker RJH, Lugtenberg BJJ (1987) *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. Nature **328:** 337–340

Sprent JI (1989) Which steps are essential for the formation of functional legume nodules? New Phytol **111:** 129–153

Szczyglowski K, Szabados L, Fujimoto S, Silver D, de Bruijn FJ (1994) Site-specific mutagenesis of the nodule-infected cell expression (NICE) element and the AT-rich element ATRE-BS2* of the *Sesbania rostrata* leghemoglobin *glb3* promoter. Plant Cell **6:** 317–332

Thummler F, Verma DPS (1987) Nodulin-100 of soybean is the subunit of sucrose synthetase regulated by the availability of free heme in nodules. J Biol Chem **262:** 14730–14736

Towbin H, Staehelin T, Gordon T (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Proc Natl Acad Sci USA **76:** 4350–4354

Triplett EW (1985) Intercellular nodule localization and nodule specificity of xanthine dehydrogenase in soybean. Plant Physiol. **77:** 1004–1005

Van de Wiel C, Scheres B, Franssen H, Van Lierop MJ, Van Lammeren A, Van Kammen A, Bisseling T (1990) The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. EMBO J **9:** 1–7

van Kammen A (1984) Suggested nomenclature for plant genes involved in nodulation and symbiosis. Plant Mol Biol Rep **2:** 43–45

Verma DPS (1992) Signals in root nodule organogenesis and endocytosis of *Rhizobium*. Plant Cell **4:** 373–382

Verwoerd TC, Dekker BMM, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Res **17:** 2362

Welsh J, Changa K, Dalal SS, Cheng R, Ralph D, McClelland M (1992) Arbitrarily primed PCR fingerprinting of RNA. Nucleic Acids Res **20:** 4965–4970