MPMI Vol. 12, No. 11, 1999, pp. 947–950. Publication no. M-1999-0823-01O.

Technical Advance

FISH Chromosome Mapping Allowing Karyotype Analysis in *Medicago truncatula* Lines Jemalong J5 and R-108-1

Malika Cerbah,[1] Zoltan Kevei,[2] Sonja Siljak-Yakovlev,[1] Eva Kondorosi,[2] Adam Kondorosi,[2] and Toan Hanh Trinh[2]

[1]Laboratoire d'Evolution et Systématique, Centre National de la Recherche Scientifique, URA 2154, Université Paris XI, Bât 362, 91405 Orsay, France; [2]Institut des Sciences Végétales, Centre National de la Recherche Scientifique, UPR40, Avenue de la Terrasse, 91198 Gif-sur-Yvette CEDEX, France
Accepted 14 July 1999.

Two lines of *Medicago truncatula*, R-108-1 and Jemalong J5, exhibit differences in their symbiotic properties, morphological markers, and capacity for regeneration and transformation. Using cytogenetic techniques, we detected differences in chromomycin A_3 staining on the two sides of the secondary constriction of the nucleolar chromosome pair of J5 and R-108-1, while DAPI (4′, 6-diamidino-2-phenylindole) staining revealed no differences between the two lines. FISH (fluorescent in situ hybridization) analysis showed a difference between J5 and R-108-1 for the same 5S rRNA genes; three loci were found in J5 but only two loci in R-108-1. However, for the 18S-5.8S-26S rDNA genes only a single locus was found in both lines. Based on these analyses, a karyotype for the two lines was established.

The tetraploid *Medicago sativa* (alfalfa) is a major forage crop all over the world. Due to its capacity to establish symbiosis with *Sinorhizobium meliloti* and mycorrhizal fungi, *M. sativa* can have a direct impact both on ecology and agriculture. Recently, a detailed genetic map has been constructed for alfalfa and about 1,300 markers (RFLP [restriction fragment length polymorphism], RAPD [random amplified polymorphic DNA], cDNA, isozyme, and morphological markers) have been mapped on the eight linkage groups representing the haploid chromosome set of the diploid *M. sativa* (Kiss et al. 1993). However, the out-crossing nature of *M. sativa* and the relatively large genome size make genetic studies difficult in this species. Therefore, *M. truncatula*, an annual autogamous, diploid species with a small genome size (0.98 to 1.15 pg; Blondon et al. 1994) was proposed as a model plant for the Leguminosae (Barker et al. 1990). Efficient and reproducible regeneration and transformation methods have been worked out for two lines of *M. truncatula*, Jemalong J5 and R-108-1 (Chabaud et al. 1996; Trieu and Harrison 1996; Trinh et al. 1998). These two lines exhibit differences in several morphological markers, such as the shape and pigmentation of leaves and flowers, the form of seed pods, and also in their symbiotic properties (Hoffmann et al. 1997). In addition, in various *Medicago* spp. only a few karyological data are available, like the C-banded karyotype in *M. sativa* subsp. *coerulea* (Bauchan and Hossain 1997). In the same species, cytological studies on the nucleolar organizer regions (NORs) have been conducted by chromomycin banding and in situ hybridization of ribosomal DNA (Calderini et al. 1996). However, until now no karyotype has been available for *M. truncatula*. Therefore, we carried out cytogenetical studies on the *M. truncatula* lines J5 and R-108-1 with fluorochrome banding and in situ hybridization to learn if there are differences in genome organization between the two lines. This molecular cytogenetic approach may serve as a basis for the integration of genetic and physical maps and for the fine analysis of the chromosome structure including mutations as well as individual genes.

Corresponding author: Hanh Trinh; Telephone: (33)-01-69 823 574; FAX: (33)-01-69 823 693; E-mail: trinh@isv.cnrs-gif.fr

For the cytological studies, young root meristems excised from plants cultivated in the greenhouse (Trinh et al. 1998) were pretreated with α-bromonaphtalene at 4°C for 2.5 h and fixed in 3:1 ethanol-acetic acid. Chromosome spreads were obtained by isolation of protoplasts after incubation in an enzyme mixture of 4% hemicellulase (Sigma, St. Quentin Fallavier, France), 1% pectolyase Y23 (Seishin, Tokyo), and 4% cellulase R10 Onozuka (Yakult Honsha, Tokyo), by the air-drying technique of Geber and Schweizer, with light modifications described in Cerbah et al. (1998). Protoplasts were obtained after sequential centrifugations of 5 min in microvolumes of 100 µl in 0.01 M citrate buffer and fresh fixative (3:1 ethanol-acetic acid).

For fluorochrome banding, chromomycin A_3 (CMA_3) and DAPI (4′, 6-diamidino-2-phenylindole) were used. The GC-rich heterochromatin was stained with CMA_3 according to Schweizer (1976). The AT-rich heterochromatin regions were revealed by DAPI counterstaining of chromosomes after in situ hybridization. CMA_3 staining revealed large GC-rich heterochromatin blocks in both sides of the secondary constriction (SC) of the nucleolar chromosome pair. In J5, the two CMA_3 bands were the same size (Fig. 1A), while in R-108-1 they were asymmetric—one band was smaller than the other (Fig. 1B). Another difference was that in J5 a thin band at the

they were found not only at the centromeric position but also on each side of the SC (Fig. 1C, D).

For in situ hybridization, two ribosomal DNAs (18S-5.8S-26S and 5S rDNA) were used because they present the highly conserved families of repeated sequences in plants (Lapitan 1992). Therefore, their loci can be visualized on the chromosomes with the use of heterologous probes. The 18S-5.8S-26S rDNA allows the localization of NORs in nucleolar chromosomes. The 5S rDNA was used as tandemly repeated units with a high copy number (1,000 to 50,000 copies per haploid genome for 200 to 900 bp long; Goldsbrough et al. 1982). To localize these two ribosomal genes (18S-5.8S-26S and 5S rDNA) on the chromosomes, fluorescent in situ hybridization (FISH) was carried out with two probes simultaneously according to the protocol of Heslop-Harrison et al. (1991) and with the modifications of Cerbah et al. (1998). One of the probes was a 4-kb *Eco*RI fragment containing a part of the 18S-5.8S-26S rDNA as well as intergenic spacers from *Arabidopsis thaliana*. The second probe was the pTA 794 clone (Gerlach and Dyer 1980), which contains a 410-bp *Bam*HI fragment of the 5S wheat rDNA, recloned in pUC. Both probes (18S-5.8S-26S and 5S rDNA) were labeled by the polymerase chain reaction (PCR) with digoxigenin-11-dUTP (Boehringer Mannheim, Meylan, France) for the 18S-5.8S-26S rDNA and Fluoro-Red-dUTP (Amersham, Les Ulis, France) for the 5S rDNA. The slides containing the chromosome spreads were incubated with 100 µg of DNase-free RNase per ml in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) for 1 h at 37°C, then with 5 µg of pepsin per ml in 0.01 M HCl for 15 min at 37°C, followed by a wash in 2× SSC and dehydration with an ethanol series. Approximately 0.5 µg/ml of the rDNA probes was mixed with 50% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, 0.1% (wt/vol) sodium dodecyl sulfate, 250 µg of salmon sperm per ml, and 2× SSC. The mixtures were denatured at 80°C for 10 min and immediately chilled on ice. After addition of the hybridization mix, the slides were also denatured at 80°C for 10 min (for chromosome denaturation), and hybridization was carried out overnight at 37°C. Post-hybridization washing consisted of a 5 min immersion in 2× SSC at 42°C followed by two stringent washes in 20% formamide in 0.1× SSC at 42°C. Slides were washed in 0.1× SSC for 5 min at 42°C, in 2× SSC for 3 × 5 min and in 4× SSC, 0.2% Tween 20 (4× SSCT) for 5 min at room temperature. For detection of the 18S-5.8S-26S signal, the slides were treated with 5% (wt/vol) bovine serum albumin (BSA) in 4× SSCT for 5 min, incubated for 1 h at 37°C in 20 µg of anti-digoxigenin-fluorescein per ml (Boehringer Mannheim) in the same buffer, and rinsed with 4× SSCT three times for 5 min at room temperature. Chromosomes were counterstained for 30 min with 2 µg of DAPI per ml in McIlvaine solution at pH 7 (0.11 M Na_2HPO_4 plus 0.006 M citric acid). Preparations were mounted in antifade solution (AF1 citifluor) and examined with a Zeiss Axiophot microscope. Forty metaphase chromosomes isolated from root tips of different individuals were analyzed for each line—J5 and R-108-1. These metaphase chromosomes were digitized with a highly sensitive CCD camera (Princeton Instruments, Evry, France) and an image analyzer system (Graftek; Fluograb, CITY, France), allowing superimposed images. Photographs were obtained with a Sony color printer (Sony, Paris). In situ hybridization of rRNA genes is shown in Figure 1E and F. The 18S-5.8S-26S rRNA genes (green) were found on the nucleolar chromosome pair in both lines and co-localized with the GC-rich heterochromatin at the SC. In contrast, the 5S rRNA gene (red) was detectable in three chromosome pairs of J5 while it was found on only two chromosome pairs of R-108-1. In R-108-1, the 5S and 18S-5.8S-26S rRNA genes were carried by different chromosomes while in J5 one of the three 5S rRNA gene loci was observed near the 18S-5.8S-26S rRNA genes in the chromosome pair 1. The remaining two 5S rRNA gene loci were carried by two metacentric chromosome pairs in both lines (pairs 6 and 8). The 5S loci were co-localized with DAPI bands.

As shown in Figure 1G and H, most of the *M. truncatula* chromosomes have similar morphology, especially the chromosome pairs 2, 3, 4, and 5, which were difficult to distinguish from each other. The karyotype of *M. truncatula* is constituted by metacentric (chromosome pairs 6 and 8) and submetacentric chromosomes. The length of the largest chromosome pair (1) was about 4.5 µm, while that of the smallest pair (8) was about 3 µm. An SC, which is the site of the NORs, was observed adjacent to the centromere on the short arm of chromosome pair 1. Despite the difficulties of the karyotype analysis, an ideogram was built for the two lines with the aid of an image analyzer by measuring chromosome length, centromere position, detection of DAPI and CMA_3 bands, and rDNA localization. Other cytogenetic markers, like C-bands and FISH experiments with other probes, are needed for further specific marking of the chromosome.

The major difficult step of in situ hybridization is the spreading of the chromosomes directly on slides. It is rare to obtain metaphases with all chromosomes well separated (Fig. 1E and F): in most cases, some chromosomes are always stuck together (Fig. 1A and B).

Our results show that in addition to differences in symbiotic properties and morphological markers, the lines Jemalong J5 and R-108-1 also exhibit some distinct cytogenetical properties as well as differences in the location of the 5S rDNA. However, no differences were found in the number and location of the 18S-5.8S-26S rDNA between the two lines, as was also observed in the diploid *M. sativa* subsp. *coerulea* and subsp. *falcata* (Calderini et al. 1996). The additional locus of 5S in J5 and the larger CMA_3 band observed in the same location might reflect a larger genome size in J5 (2C = 1.15 pg) than in R-108-1 (2C = 0.98 pg) and differences in base composition (38.6% GC and 38.1% GC, respectively; Blondon et al. 1994). The absence of the third 5S locus in the R-108-1 line seems to correlate with the small CMA_3 band observed in this chromosome pair. Nevertheless, from these results, it is difficult to know whether the chromosomal locations of 5S rDNA observed are due to differences in the copy number of rDNA sequences between the two lines, followed by translocation mechanisms. Comparison of the number and location of the 18S-5.8S-26S and 5S rDNA genes in related species of *Medicago* may shed some light on the chromosome evolution in this genus.

ACKNOWLEDGMENTS

We thank Odile Robin and Patricia Durand for technical assistance, and Nathalie Mansion for photographic aid. This work was supported by EU INCO-COPERNICUS contract no. ERBIC15CT960506.

Mol Genet Genomics (2002) 266: 1012–1019
DOI 10.1007/s00438-001-0628-3

ORIGINAL PAPER

G. Endre · P. Kaló · Z. Kevei · P. Kiss · S. Mihacea
B. Szakál · A. Kereszt · G.B. Kiss

Genetic mapping of the non-nodulation phenotype of the mutant MN-1008 in tetraploid alfalfa (*Medicago sativa*)

Received: 24 October 2001 / Accepted: 3 December 2001 / Published online: 23 January 2002

Abstract Roots of the non-nodulating *Medicago sativa* mutant MN-1008 neither undergo root-hair curling, cortical cell division nor any of the early molecular events that accompany nodule initiation and development following rhizobial infection or treatment with Nod factor. These observations suggested that the mutation(s) impaired a pivotal function in Nod factor perception or in the signal transduction pathway. In this paper we show that the genetic lesion conditioning the recessive non-nodulation phenotype in the tetraploid alfalfa mutant MN-1008 can be localized to a single region on LG5 of the *M. sativa* genetic map. This conclusion is based on genetic analyses conducted at the tetraploid level, involving both segregation analysis and genetic mapping of the trait with respect to molecular DNA markers. The genetic mapping of the Nod$^-$ phenotype was performed in a segregating tetraploid F2 population, taking advantage of the availability of an advanced genetic map for diploid alfalfa. Two tightly linked flanking markers have been identified which will facilitate the physical mapping and cloning of the gene(s) that underlie(s) the non-nodulation phenotype.

Keywords Symbiotic nitrogen fixation · Nodulation mutant · Genetic mapping · *Medicago sativa* · *Rhizobium* Nod factor

Introduction

The first exchange of signals between a leguminous plant and its microsymbiont is crucial for the formation of the nitrogen-fixing symbiotic nodule. Rhizobial nodulation genes are mainly induced by flavonoid-type molecules excreted by the host plant. This results in the production by the bacteria of specific lipochito-oligosaccharide compounds, called Nod factors (Lerouge et al. 1990; Schultze and Kondorosi 1996). The first morphological plant responses upon rhizobial inoculation are root-hair curling and cortical cell division, leading to bacterial invasion and nodule development. Purified Nod factors alone can also evoke root-hair deformation and cortical cell division in a host-specific manner. Several studies have begun to explore the subcellular events involved in this signal transduction cascade (for review, see Downie and Walker 1999; Stougaard 2000), but the genetic determinants that control them have not yet been identified. An increasing number of induced mutants is available in several legume species, which can be used to dissect the signal transduction pathway, and efforts to identify the genes responsible for the nodulation phenotype are underway (for review see Stougaard 2001).

The non-nodulating alfalfa (*Medicago sativa*) mutant MN-1008 was among the first symbiotic mutants (Caetano-Anolles and Gresshoff 1991) to be identified. Different alfalfa lines were crossed, and progeny were self-pollinated; nodulation screening then determined that MN-1008 was unable to form symbiotic nodules in the presence of the compatible rhizobial cells (Peterson and Barnes 1981). Based on segregation data for the non-nodulation phenotype in tetraploid F2 populations, Peterson and Barnes (1981) concluded that the trait was determined by two, unlinked, recessive genes. Further studies of MN-1008 revealed a lack of characteristic nodulation responses, like root-hair deformation, cortical cell division (Dudley and Long 1989; Endre et al. 1996), and calcium spiking (Ehrhardt et al. 1996). Changes were also detected in membrane depolarization activity (Felle et al. 1996). In addition, MN-1008 was unable to form a vesicular-arbuscular mycorrhizal symbiosis (Bradbury et al. 1991). However, an alfalfa-specific phenomenon, the infrequent appearance of spontaneous nodules on the root in the absence of

Communicated by A. Kondorosi

G. Endre (✉) · P. Kaló · Z. Kevei · P. Kiss · S. Mihacea
B. Szakál · A. Kereszt · G.B. Kiss
Institute of Genetics, Biological Research Center
of the Hungarian Academy of Sciences H-6701 Szeged,
P.O. Box 521. Hungary
E-mail: endre@nucleus.szbk.u-szeged.hu
Tel.: +36-62-432232
Fax: +36-62-433503

rhizobia (Truchet et al. 1989), was observed in this mutant (Caetano-Anolles et al. 1993). This suggests that nodule organogenesis is intact, but the plant is missing a key element in the perception or transduction of the bacterial signal molecules that normally induce the specific initiation of nodule formation.

We decided to use map-based cloning to identify the gene(s) in MN-1008 responsible for the mutant phenotype. Initially, we attempted to reduce the ploidy of the mutant line to the diploid level in order to facilitate genetic mapping. We identified a single diploid non-nodulating derivative (Endre et al. 1996); however, the reduced ploidy level impaired the viability of the diploid individual, so that no segregating population could be produced. We therefore took the only available option, and mapped the non-nodulation phenotype in a tetraploid segregating population, with the indispensable aid of the improved genetic map available for diploid *M. sativa* (Kiss et al. 1993; Kaló et al. 2000).

Materials and methods

Plant material, nodulation testing and growth conditions

The non-nodulating tetraploid alfalfa mutant line designated as MnNC-1008(NN) (Peterson and Barnes 1981; Barnes et al. 1988) is referred to throughout the paper as MN-1008. Tetraploid *M. sativa* cv. Nagyszénási (Michaud et al. 1988) was used as the wild-type nodulating parent. Plants were maintained in pots in the greenhouse at constant temperature (22–25°C), with a 16-h photoperiod. The tetraploid mapping population was initiated by applying pollen from the nodulating *M. sativa* cv. Nagyszénási (MsNA/5, male parent) to the stigma of the mutant alfalfa MN-1008 (MN-1008/17, female parent). The F2 progeny were generated by self-pollinating the F1 hybrid plants. Seeds were collected after maturation and kept at 4°C for at least 2 weeks. To test the progeny in a nodulation assay, seeds were sterilized prior to germination by rinsing in ethanol and treating with 0.1% $HgCl_2$ solution for 5 min, then rinsing in sterile distilled water at least six times. For germination, the imbibed seeds were placed on the surface of plates containing 1% agar in distilled water, and stored upside down overnight. The germinated seedlings were then placed on the surface of 1% agar slants made with nitrogen-free Gibson medium (Gibson 1980). The seedlings were inoculated with 0.2–0.4 ml of a suspension (10^8–10^9 cells/ml) of *Sinorhizobium meliloti* AK631 in 0.9% NaCl solution (Putnoky et al. 1990) 2–4 days later. The nodulation phenotype was scored 6 weeks after inoculation, after which plants were potted. Plants scored as non-nodulating were evaluated once more for the presence or absence of nodules 2 months later.

The diploid mapping population used for the construction of the basic (Kiss et al. 1993) and the improved (Kaló et al. 2000) genetic maps of alfalfa was described previously.

DNA isolation and hybridization

Total DNA was isolated from young leaves according to Kiss et al. (1993). Aliquots (15 µg) of total DNA were digested with the restriction enzymes *Eco*RI, *Eco*RV, *Hin*dIII, *Dra*I and *Bam*HI (Amersham or Fermentas) according to the suppliers' instructions. The DNA fragments were fractionated on a 1.1% agarose gel and transferred by the capillary method (Southern 1975) to nylon membranes (Hybond-N+, Amersham), in accordance with the supplier's protocol. For probe preparation, PCR-amplified DNA fragments were isolated from agarose gels using the QIAEX Gel Extraction Kit (Qiagen), and were labelled with [α-^{32}P]dCTP by random priming (Feinberg and Vogelstein 1983), using the Pharmacia Ready-to-go labelling kit. Hybridization experiments were performed at 55–60°C, and the washes were carried out as described by Kiss et al. (1993).

PCR amplification

PCR amplification with random amplified polymorphic DNA (RAPD) primers (Operon Technologies) was based on a modification of the method of Williams et al. (1990). The 25-µl reaction mix consisted of 25 ng of total plant DNA template, 5 pmol of 10-mer primer, 200 µM of each dNTP, 2.4 mM $MgCl_2$, and 1 U of *Taq* polymerase in 1× *Taq* polymerase buffer (Promega). PCR cycling conditions were as follows: 5 s at 94°C, 1 min at 37°C, 1 min at 72°C for 40 cycles in a Programmable Thermal Controller (MJ Research).

For specific PCR amplifications the $MgCl_2$ concentration was reduced to 1.5 mM, and the PCR cycling conditions were: 30 s at 94°C, 1 min at the annealing temperature (depending on the primers) and 1 min at 72°C for 35 cycles, with a final extension step at 72°C for 4 min.

Genetic markers

Primers SHMTu1 (5′-GAGAAAGCTAGACAATGGAAG-3′) and SHMTd1 (5′-GCTTGCATAACAGATAGAGA-3′) were used to amplify the Shmt fragment, which was then used for labelling and hybridization.

The primers CADu3 (5′-CTAAGAAAGGCTTTGGTTGG-3′) and CADd3 (5′-TCACGAATCGATATTTGACA-3′) were used to identify the presence of the N5 allele of the Cad marker, and the primer pair CADu2 (5′-GTCAAATATC-GATTCGTGAT-3′) and CADd1 (5′-TATTCGACTCATCATGGTTA-3′) identified the allele N6. Both alleles are derived from the Nod$^+$ male parent.

The primers Comau1 (5′-CTCAACTCACATGATCTA-3′) and Comad1 (5′-CTCAACTCACATGATCTA-3′) were used to amplify the Coma fragment, which was then used for labelling and hybridization.

For OT18B, RAPD PCR was done with Operon Technologies primer T18 (5′-GATGCCAGAC-3′), the fragment linked to the Nod region was isolated and used for labelling and hybridization.

All other markers were described previously by Kaló et al. (2000).

Genetic nomenclature, genotyping and genetic analysis

In the present study, the Nod$^-$ phenotype was localized to a single region on the genetic map of alfalfa (see below). However, the nature of the genetic alteration and the number of genes affected in this region are not known; therefore, the chromosomal region responsible for the non-nodulation trait will be referred to as "the genetic locus", "the genetic region", "the Nod locus", "the nod region" or "the Nod mutation(s)". This region covers the part of Linkage Group 5 (LG5) between the markers U71 and U492 (see Fig. 1B). Molecular cloning and sequence analysis will reveal whether the genetic alteration that conditions this Nod$^-$ trait was caused by a single mutation in a single gene or by multiple events affecting more than one gene.

Genotyping and genetic analysis in the diploid population were performed as described previously (Kiss et al. 1993, 1998; Kaló et al. 2000). In the tetraploid population, dominant-type scoring (presence or absence of band) was performed by identifying Nod$^+$ male parental (MsNA/5) alleles N5 and N6, and Nod$^-$ female parental (MN-1008/17) alleles n1, n2 in the Nod region. Genetic analysis and mapping were conducted by the non-mathematical colormapping procedure described previously (Kiss et al. 1998). Genetic maps are shown as the classical chromosome-like structure.

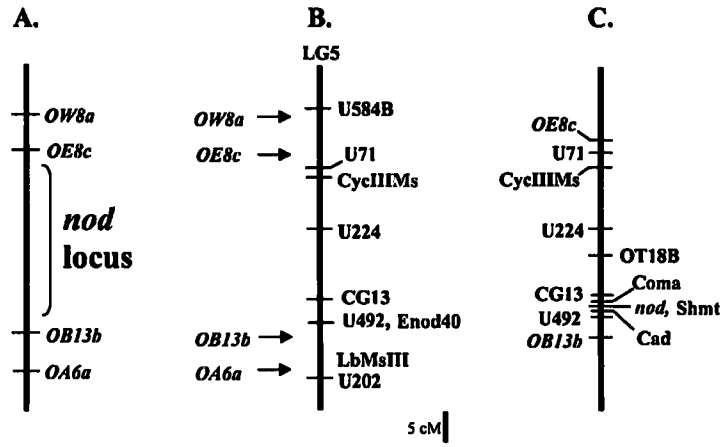

Fig. 1A–C Genetic maps of the nod region. A The four RAPD markers showing linkage to the Nod⁻ phenotype in the tetraploid population. B The RAPD markers mapped on the LG5 of the published map of the diploid *M. sativa* population, between the two RFLP markers U584B (at 6 cM) and U202 (at 39 cM) (Kaló et al. 2000). C RFLP and PCR markers mapped in the tetraploid population showing the inferred position of the nod locus

Results

Generation of an F2 mapping population segregating for the Nod⁻ phenotype of MN-1008

In order to map the genetic determinant of the non-nodulation (Nod⁻) trait in the tetraploid *M. sativa* mutant MN-1008, segregating F2 progeny were produced from a hybrid F1 generation. Flowers of a single individual of MN-1008 (MN-1008/17) as female parent were fertilized with pollen originating from the flowers of a single individual (MsNa/5) of the nodulation-competent (Nod⁺) *M. sativa* cv. Nagyszénási (male parent). F1 seeds were collected, germinated and the seedlings were tested for symbiotic nodulation. The nodulation-competent phenotype of the progeny confirmed the hybrid nature of the F1 plants and the recessive character of the non-nodulating trait of the maternal parent MN-1008. After being replanted in pots, F1 individuals were grown and self-pollinated by tripping the flowers as they developed. F2 seeds from 44 F1 plants were collected and each family was examined for seed production, germination ability, viability and vigour of the progeny. In addition, nodulation tests were conducted using as many individuals as possible from each family to estimate the Nod⁻/Nod⁺ segregation ratios. This survey revealed distinct differences in viability and fitness among individuals of the different F2 families, but the Nod⁻/Nod⁺ segregation ratios in independent families were similar (see Table 1 for the survival rates and nodulation phenotype of the F2 families). Since many seedlings of the F2 progeny exhibited slow growth and development under the conditions used to test for the symbiotic nodulation phenotype, we extended the time of evaluation of nodulation by an additional 2 months, compared to the 6-week period used by Peterson and Barnes (1981; see also Materials and methods). The range of values (1:25–1:57; Table 1) found for the segregation ratios of the Nod⁻ vs. Nod⁺ phenotypes in the F2 families showed the closest correlation with the theoretical segregation ratio for a single, recessive gene in the tetraploid F2 population (1:35), with a duplex configuration for the nodulation trait in the F1 parents (see below). Based on their viability, vigour and other characteristics, two families (NAB and NBW) were selected for further studies, in which several thousands of F2 seeds were investigated.

Identification and mapping of RAPD markers linked to the non-nodulation phenotype

The mapping of the genetic alteration that conditions the non-nodulation phenotype in MN-1008 was begun by identifying linked RAPD markers using Bulked Segregant Analysis (BSA, Giovannoni et al. 1991; Michelmore et al. 1991). For this, the self-pollinated F2 progeny of two F1 hybrids, NAB and NBW, were used. One bulk representing the mutant plants and another representing the wild-type nodulating plants were created with the members of the NAB and NBW families, respectively. DNA samples from five individuals were combined in each bulk for use as the template in RAPD PCR amplification. Several polymorphic amplification products were identified after testing more than 500 different 10-mer primers (Operon Technologies). These putative linked RAPD fragments were re-tested by PCR amplification using genomic DNA templates from the individual F2 plants to confirm the linkage of the polymorphic amplification products. Eventually four RAPD markers (OA6a, OB13b, OE8c and OW8a) were identified that showed unambiguous co-inheritance with the nodulation trait (Fig. 1A). In other words, the dominant Nod⁺ phenotype co-segregated with the presence of the amplification product, while the recessive Nod⁻ character associated with the absence of the amplification product (Table 2). Among the RAPD markers, OA6a (~650 bp) was identified only in the NBW family, OW8a (~300 bp) was identified only in the NAB family, whilst OB13b (~850 bp) and OE8c (~480 bp) could be detected in both F2 populations. This variance reflects the inheritance of different chromosomes from the heterozygous wild-type parent in the two F1 hybrids

Table 1 Germination, viability and nodulation abilities in the F2 generation

F2 family[a]	Number of sterilized F2 seeds	Number of germinated F2 seeds	Number of F2 plants lost	Number of Nod^- F2 plants	Number of Nod^+ F2 plants	Nod^-/Nod^+
NAA	773	748	387	8	353	1/45
NAB	5571	4988	2412	50	2526	1/52
NAC	482	446	266	4	176	1/45
NAK	400	392	213	6	173	1/30
NAX	646	575	230	14	331	1/25
NBI	202	189	68	4	117	1/30
NBO	120	118	12	3	103	1/35
NBW	1285	1199	460	13	726	1/57

[a]Only those families are listed in which at least 100 plants survived and were scored in the nodulation test

Table 2 Genotypic categories of the plants in the NAB and NBW families

Marker	Genotypic categories							
	1	2	3	4	5	6	7	8
	NAB family							
OW8a	Y	N	N	N	N	Y	Y	Y
OE8c	Y	Y	N	N	N	N	Y	Y
Nod	+	+	+	−	−	−	−	+
OB13b	Y	Y	Y	Y	N	N	N	N
Total numbers of plants	773	5	1	3	8	2	1	4
	NBW family							
OE8c	Y	N	N	N	N	Y	Y	Y
Nod	+	+	−	−	−	−	+	+
OB13b	Y	Y	Y	N	N	N	N	Y
OA6a	Y	Y	Y	Y	N	N	N	N
Total numbers of plants	432	3	0	2	6	0	3	1

[a]Genotype scores: Y, presence, and N, absence (highlighted in *bold type* and *underlined*) of a PCR product originating from the Nod^+ parental alleles; +, Nod^+ phenotype; –, Nod^- phenotype (highlighted in *bold face*)

and thereby in the two segregating populations. The maximum parsimony principle was used to determine marker order by identifying the minimum number of recombination events (transition from normal to boldface type in Table 2). Based on the recombination detected, plants were classified into eight genotypic categories as shown in the columns of Table 2. All categories were represented among the individuals of the NAB family, but two categories (3 and 6) were not represented in the NBW family. The total numbers of plants tested were 797 and 447, respectively.

The next phase in determining the location of the nodulation trait on the genetic map of alfalfa was to map the linked RAPD amplification products as RFLP markers on the genetic map of diploid alfalfa (Kaló et al. 2000). For this purpose the amplified fragments (OA6a, OB13b, OE8c and OW8a) were used as hybridization probes on filters bearing genomic DNA samples from individual members of the diploid mapping population, digested with different restriction enzymes (*Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII). At least one of the digests resulted in a polymorphic hybridization pattern that could be used for genotyping and mapping of these markers on the diploid alfalfa map. The linkage analysis revealed that all four RAPD markers mapped to linkage group (LG) 5, between RFLP markers U584B and U202 (Fig. 1B), which define a genetic distance of about 30 cM (Kaló et al. 2000).

Confirmation of the map position of the nodulation trait in the tetraploid population

To ensure that the non-nodulating phenotype in the original tetraploid plants mapped to the same region of LG5, we analyzed the segregation of selected RFLP markers (U71, CycIIIMs, U224, CG13, and U492) identified previously in the diploid mapping population (Kaló et al. 2000, and Fig. 1B). For this purpose the NAB family was selected, and the hybridization experiments with the appropriate RFLP probes were carried out on filters containing genomic DNA samples of selected Nod^- and Nod^+ F2 individuals. Total DNA was digested with the restriction enzymes *Eco*RI, *Eco*RV, *Dra*I, *Bam*HI and *Hin*dIII. The main aim of this thorough restriction analysis was to try to visualise all four alleles of a marker in the population (two alleles, designated as n1 and n2, originating from the mutant MN-1008, two, N5 and N6, from the nodulating *M. sativa* parent). Identifying as many alleles as possible increases the probability of detecting recombination events between the Nod^- phenotype and the molecular markers in the region. Figure 2 shows an example of the hybridization experiments, in which the four alleles of marker CycIIIMs were identified.

To position a particular marker relative to the Nod^- locus, genotypic scoring was done by determining the

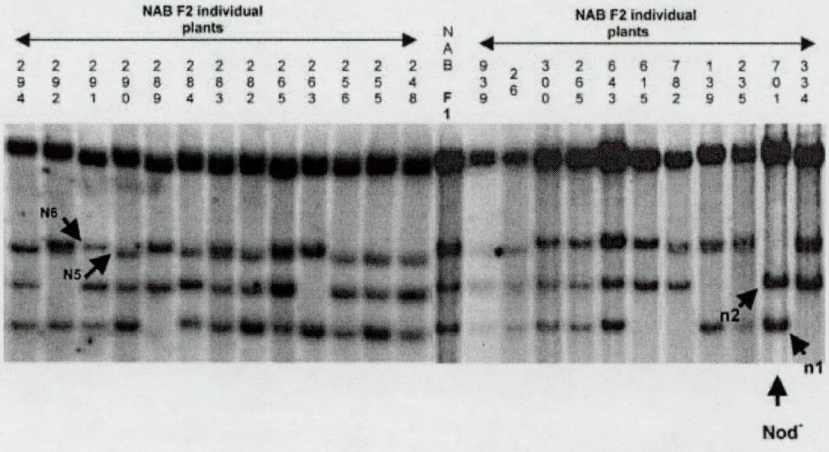

Fig. 2 Visualization of the four alleles of CycIIIMs. Total DNA from NAB F1 and F2 individuals (their numbers are indicated *above* the lanes), was digested with *Eco*RV and hybridized with a CycIIIMs-specific probe. Plant No. 701 contains hybridizing bands n1 and n2 originating from the Nod^- maternal plant. N5 and N6 indicate the alleles inherited from the paternal Nod^+ parent

presence or absence of any Nod^+ parental alleles. Following the inheritance of every allele helped to order the molecular markers unambiguously relative to each other. Genotypic data were stored and analyzed according to the colormap procedure (Kiss et al. 1998) adapted for the tetraploid situation. A recombination event between the molecular markers and the Nod locus could be observed only in the rare cases in which an individual plant was Nod^- but carried at least one Nod^+ allele for the marker, or when the plant was Nod^+, but did not carry any Nod^+ allele for the marker. Based on the recombination events detected, plants were classified into five genotypic categories as shown in columns 1–5 in Table 3. No other genotype configuration was observed (e.g. double recombination category). The total number of plants tested was 797. This mapping experiment showed that the non-nodulation phenotype was indeed linked to RFLP markers selected from LG5. Moreover, the order of the markers was conserved as in the diploid alfalfa map (Fig. 1B, C), indicating a high degree of synteny between the two *M. sativa* subspecies, which have different ploidy levels. Specifically, two recombination events between the non-nodulation locus and marker CycIIIMs, and a single recombination event between the locus and marker U492, were used to place the non-nodulation locus between these markers. The recombination events were detected in the primary population of the NAB family, consisting of about 800 individuals. No recombination was detected between the nodulation locus and the markers U224 and CG13.

Table 3 Genotypic categories of the plants in the primary tetraploid mapping (NAB) population

Marker	Genotypic categories[a]				
	1	2	3	4	5
U71	Y	N	N	N	Y
CycIIIMs	Y	N	N	N	Y
U224	Y	Y	N	N	N
CG13	Y	Y	N	N	N
Nod	+	+	−	−	−
U492	Y	Y	Y	N	N
Total numbers of plants	782	1	1	12	1

[a]Genotype scores are as indicated in Table 2

Fine mapping in the nodulation region

In order to identify tightly linked molecular markers for use in chromosomal walking experiments, we continued to saturate the region between markers CycIIIMs and U492 with molecular markers using BSA. However, in this phase of the experiments, the diploid alfalfa mapping population was used to create maternal and paternal homozygote bulks, and more than 300 RAPD primers were tested with them. This screen resulted in the identification of 20 new RAPD markers that mapped between the markers CycIIIMs and U492 (data not shown). In addition to the RAPD markers, a new RFLP marker (Shmt, Turner et al. 1992) and two specific (sequence-based) PCR markers (Coma, Cad) were also located in this segment of the diploid alfalfa map. Afterwards, four selected markers (Shmt, Coma, Cad, and RAPD marker OT18B) were transferred from the diploid map onto the tetraploid map to determine their positions with respect to the nodulation locus. Markers OT18B and Shmt were mapped as RFLP markers using their labelled fragments as hybridization probes. Coma and Cad were genotyped as specific PCR amplification products. The inheritance of the polymorphic alleles showed linkage with the nodulation trait in the primary NAB population; no recombination event was detected, suggesting tight linkage.

To determine the order and the relative position of these closely linked markers, a larger population with more individuals was needed in the tetraploid segregating population. Therefore the NAB F2 family was extended by continuously tripping the flowers of the NAB F1 plant, and the new F2 individuals were tested for symbiotic nodulation. A final population of more than 2500 individuals (among them 50 with the mutant phenotype; see Table 1) was used for fine-scale mapping. The results of a thorough analysis for markers U224,

OT18B, Coma, Shmt, Cad, and U492 in this extended population are summarized in Table 4. Based on the recombination events detected, plants were classified into 11 genotypic categories as shown in the columns of Table 4. Three of the fourteen theoretical recombination categories were not represented in the population. The total number of plants tested was 2576. No other genotypic configuration (e.g. a double recombination category) was observed. As a result of the fine-scale mapping, the location of the non-nodulation trait could be narrowed down further (Fig. 1C).

Based on the linkage analysis of these markers, the marker most closely linked to the Nod$^-$ locus was the RFLP marker Shmt, with no recombination event being detected in the population consisting of more than 2500 plants. Two other markers, Coma and Cad, exhibited one and two recombinations, respectively, flanking both sides of the Nod locus. The order of all tested molecular markers was conserved in both the tetraploid and diploid alfalfa populations. Based on the number of the recombination events in the two populations, the genetic distance between the two markers flanking the locus responsible for the non-nodulation trait is estimated to be about 0.4 cM.

Discussion

We have shown in this paper that the mutation(s) that condition the recessive non-nodulation phenotype in the tetraploid alfalfa mutant MN-1008 can be localized to a single region on LG5 of the *M. sativa* genetic map. This conclusion is based on genetic analyses conducted at the tetraploid level, involving both segregation analysis and genetic mapping of the trait with respect to molecular DNA markers. The genetic mapping of the Nod$^-$ phenotype was performed in a tetraploid F2 segregating population with the aid of the advanced genetic map of diploid alfalfa. The transfer of genetic markers back and forth between the maps for the two subspecies (diploid and tetraploid) of *M. sativa* was made possible by the high degree of synteny shared between genomes. Screening both populations for linked markers resulted in the saturation of this region with molecular markers and the identification of two tightly linked flanking markers, which are well suited for physical mapping and cloning of the gene(s) that underlie(s) the non-nodulation phenotype.

To map the non-nodulation trait in MN-1008, a tetraploid mapping strategy was adopted, because our initial efforts to produce diploid segregating population was unsuccessful. We had originally tried to isolate suitable diploid derivatives from a 4x-2x cross (Bingham 1969). This would simplify the analysis of segregation, and thus facilitate genetic mapping. However, although diploid progeny of MN-1008 could be isolated (Endre et al. 1996), the viability of the derived diploid plants was dramatically diminished, and appropriate mapping populations could not be established. As a consequence, a tetraploid genetic mapping strategy was implemented, even though the non-nodulating phenotype of MN-1008 was previously associated with two unlinked recessive genes (Peterson and Barnes 1981), which would make the genetic analysis much more difficult.

Genetic analyses of the mutation(s) responsible for the non-nodulation phenotype were started by determining the segregation ratio of the Nod$^-$ trait. This investigation was conducted on the tetraploid F1 hybrid plants and their F2 progeny. Determination of the segregation ratios would indicate the number of loci conditioning the Nod$^-$ phenotype. In the present study the Nod$^+$ phenotype of all F1 plants obtained from the cross of the mutant (MN-1008/17) with a nodulation-competent *M. sativa* plant (MsNa/5) confirmed the recessive nature of the Nod$^-$ mutation(s), and also indicated that the parental plant MsNa/5 could not carry two or more recessive alleles. When the segregation of the Nod$^-$ phenotype was tested in the F2 families, the segregation ratio ranged from 1/25 to 1/57 (Table 1). The two extended F2 families (NAB and NBW) which were investigated in more detail gave an average ratio of 1/52. These experimental segregation data correlated best with the theoretical ratio (1/36) expected for the Mendelian segregation of a single recessive allele in a tetraploid F2 population originating from a duplex ($n_in_iN_iN_i$) F1 hybrid. Since all families tested showed similar inheritance, these F1 plants must have originated from a nulliplex ($n_in_in_in_i$) × quadruplex ($N_iN_iN_iN_i$) parental cross. In a case where two unlinked recessive loci determine a trait, the expected segregation ratios in

Table 4 Genotypic categories of the plants in the extended tetraploid mapping (NAB) population

Marker	Genotypic categories[a]										
	1	2	3	4	5	6	7	8	9	10	11
U224	Y	N	N	N	N	N	N	Y	Y	Y	Y
OT18B	Y	Y	N	N	N	N	N	N	Y	Y	Y
Coma	Y	Y	Y	N	N	N	N	N	N	Y	Y
SHMT	Y	Y	Y	Y	N	N	N	N	N	N	Y
Nod	+	+	+	+	−	−	−	−	−	−	+
CAD	Y	Y	Y	Y	Y	N	N	N	N	N	Y
U492	Y	Y	Y	Y	Y	Y	N	N	N	N	N
Number of plants	2516	4	2	1	2	1	41	3	2	1	3

[a]Genotype scores are as indicated in Table 2

the F2 populations can be 1/1296 (if the F1 plant was $n_in_iN_iN_i$ $n_jn_jN_jN_j$), 1/144 (if the F1 was $n_in_in_iN_i$ $n_jn_jN_jN_j$), 1/36 (if the F1 was $n_in_in_in_i$ $n_jn_jN_jN_j$), 1/16 (if the F1 was $n_in_in_iN_i$ $n_jn_jn_jN_j$), or 1/4 (if the F1 was n_in_i- n_in_i $n_jn_jn_jN_j$), depending on the allelic configuration of the Nod^+ parent. We have not detected any F2 families showing the extremely high (1/4 or 1/16) or low (1/144 or 1/1296) incidence of individuals with the Nod^- phenotype. A segregation ratio of 1/36 is also possible in the progeny of an F1 plant with an allele configuration $n_in_in_in_i$ $n_jn_jN_jN_j$. This situation, however, can be considered as single-locus segregation since the second locus does not segregate.

The tetraploid segregation data and the conclusion reached in this study concerning the number of segregating loci differ from the those originally published by Peterson and Barnes (1981). The contradiction might originate from the fact that the two genetic analyses differed in several aspects. The main differences were: (1) the evaluation of the nodulation phenotype; i.e. the determination of the number of Nod^- individuals in the F2 families; and (2) the use of different nodulation-competent parents, which could have different effects on segregation. In their original paper, Peterson and Barnes (1981) scored the plants for their nodulation ability 6 weeks after inoculation with rhizobia. In our experiments we found that in many cases the evaluation of nodulation phenotype at this time point resulted in some false Nod^- scores in several F2 families; hence we repeated the scoring 2 months later. No data were provided for the number of individuals lost in the populations studied by Peterson and Barnes (1981), but some kind of inbreeding depression effect could probably result in different types of distortion in the segregation of the non-nodulation phenotype in their F2 populations as well.

Based on our detailed segregation studies, the non-nodulation trait is likely to be due to a single genetic locus. However, the distorted segregation, and the loose correlation between the theoretical value (1/36) and the experimental value (1/52), meant that a final conclusion could only be drawn after mapping the non-nodulation trait.

The genetic mapping was started by collecting linked RAPD markers using BSA with two families (NAB and NBW). Each of the bulks contained five individuals from the Nod^- and Nod^+ plants, respectively. The Nod^- bulk represented the homozygote recessive genotype and was consequently equivalent to 20 chromosomes carrying Nod^- alleles only. On the other hand the Nod^+ bulk represented heterozygote genotype containing Nod^+ dominant and Nod^- alleles as well (the exact ratio of Nod^- to Nod^+ chromosomes was not determined). The identification and verification of the linked markers (OA6a, OB13b, OE8c and OW8a) was straightforward, irrespective of the genetic configuration of the chromosomes in the Nod region of the two F1 parents (NAB and NBW), which was unknown. The purpose of the BSA was to identify markers linked to the phenotype which could be transferred to the diploid map to determine the map position. The molecular markers linked with the non-nodulation mutation were all mapped to the same region of LG5 of the alfalfa genetic map, supporting the idea that a mutation in a single locus was responsible for the phenotype. The existing detailed diploid genetic map and the mapping population used to construct it (Kiss et al. 1993; Kaló et al. 2000) were indispensable for mapping. Besides facilitating the mapping of the RAPD markers linked to the Nod^- mutation, it provided RFLP markers that had previously been mapped in the region, and was used to enrich the region with additional markers. The genetic analysis of more than ten molecular markers in this ~30-cM region of the LG5 in both diploid and tetraploid populations showed overall synteny.

The mapping procedure in the tetraploid population was more complex. On the one hand, we had to follow the segregation of four alleles instead of two, and on the other hand only a fraction of the F2 progeny was suitable for mapping the recessive/dominant trait. Several methods were used to detect polymorphism among all four alleles of a gene, of which RFLP hybridization proved to be the most efficient. As shown in Fig. 2, the four segregating alleles in the NAB family could be clearly distinguished: those inherited by the two chromosomes from the mutant parent were labelled as n1 and n2, while those carried by the chromosomes originating from the normal nodulating parent were designated as N5 and N6. The fact that we could distinguish and follow all four alleles for most of the molecular markers enabled us to identify the recombination events between them, and hence determine their order in the tetraploid alfalfa. Detection of a recombination event between closely linked markers and the non-nodulation phenotype was more difficult. Only two subpopulations of the F2 individuals were appropriate for this: those with a Nod^- phenotype (1/36 of the F2 progeny), i.e. with homozygous configuration at the Nod^- locus (quadruplex configuration, $n_kn_kn_kn_k$, where n_k could be either n1 or n2), and those Nod^+ individuals with triplex ($n_kn_kn_kN_l$, where N_l could be either N5 or N6) configuration (1/9 of the F2 progeny). Thus, only 5/36 of the progeny provided a genetic background in which a recombination event could be positioned in relation to the Nod phenotype and the molecular marker. Therefore the tetraploid mapping population was extended to over 2500 plants and all were screened for their chromosomal configurations. The genotypes of those individuals belonging to the "useful" 5/36 set were carefully scored for all markers in the region to localize the recombination events precisely.

The genetic mapping of the Nod^- trait supported the inference drawn from the segregation analysis – that only one recessive locus was responsible for the nonnodulation phenotype. Following the original nomenclature used by Peterson and Barnes (1981), we designated this mapped region as *nn1* locus. With the help of the tightly linked molecular markers identified in this study, the localization and eventual isolation, by

chromosomal walking, of the gene *nn1* that conditions the non-nodulation phenotype seems feasible.

Acknowledgements We are particularly indebted to S. Jenei, K. Lehoczky, Z. Liptay and P. Somkúti for skillful technical assistance, to M. Graham for critical reading of the manuscript, and to N. Ellis for the pea Shmt cDNA clone. This work was supported by the BRC (Szeged, Hungary), the Bástyai-Holczer Foundation, and by the following grants: AKP 96-360/62 and AKP 00-246/35 (MTA, Hungarian Academy of Sciences), OTKA F030408 and T025467 (Hungarian Scientific Research Fund), OMFB EU-97-D8-063 (National Committee for Technical Development), NKFP Medicago Genomics Grant No.: 4/023/2001 (Ministry of Education), EuDicotMap Grant No. BIO 4CT97 2170, and Medicago Grant No. QLG2-CT-2000-30676 (European Union). A.K. was the recipient of a post-doctoral fellowship from the Hungarian Scientific Research Fund (OTKA D32691).

References

Barnes DK, Vance CP, Heichel GH, Peterson MA, Ellis WR (1988) Registration of a non-nodulation and three ineffective nodulation alfalfa germplasms. Crop Sci 28:721–722

Bingham ET (1969) Haploids from cultivated alfalfa, *Medicago sativa* L. Nature 221:865–866

Bradbury SM, Peterson RL, Bowley SR (1991) Interaction between three alfalfa nodulation genotypes and two *Glomus* species. New Phytol. 119:115–120

Caetano-Anollés G, Gresshoff PM (1991) Plant genetic control of nodulation Annu Rev Microbiol 45:345–382

Caetano-Anollés G, Joshi PA, Gresshoff PM (1993) Nodule morphogenesis in the absence of *Rhizobium* In: Palacios R, Mora J, Newton WE (eds) New horizons in nitrogen fixation. Kluwer Academic, Dordrecht, pp 297–302

Downie JA, Walker SA (1999) Plant responses to nodulation factors. Curr Op Plant Biol 2:483–489

Dudley ME, Long SR (1989) A non-nodulating alfalfa mutant displays neither root hair curling nor early cell division in response to *Rhizobium meliloti*. Plant Cell 1:65–72

Ehrhardt DW, Wais R, Long SR (1996) Calcium spiking in plant root hairs responding to *Rhizobium* nodulation signals. Cell 85:673–681

Endre G, Kaló P, Hangyel Tárczy M, Csanádi G, Kiss GB (1996) Reducing the tetraploid non-nodulating alfalfa (*Medicago sativa*) MnNC-1008(NN) germplasm to the diploid level. Theor Appl Genet 93:1061–1065

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Felle HH, Kondorosi É, Kondorosi Á, Schultze M (1996) Rapid alkalinization in alfalfa root hairs in response to rhizobial lipochitooligosaccharide signals. Plant J 10:295–301

Gibson AH (1980) Methods for legumes in glasshouse and controlled environment cabinets. In: Bergersen FJ (ed) Methods for evaluating biological nitrogen fixation. Wiley, New York, pp 139–184

Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic Acids Res 19:6553–6558

Kaló P, Endre G, Zimányi L, Csanádi G, Kiss GB (2000) Construction of an improved linkage map of diploid alfalfa (*Medicago sativa*). Theor Appl Genet 100:641–657

Kiss GB, Csanádi G, Kálmán K, Kaló P, Ökrész L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol Gen Genet 238:129–137

Kiss GB, Kereszt A, Kiss P, Endre G (1998) Colormapping: a non-mathematical procedure for genetic mapping. Acta Biol Hung 49:125–142

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

Michaud R, Lehman WF, Rumbaugh MD (1988) World distribution and historical development. In: Hanson AA, Barnes DK, Hill RR (eds) Alfalfa and alfalfa improvement. American Society of Agronomy, Madison, Wis., pp 25–124

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Peterson MA, Barnes DK (1981) Inheritance of ineffective nodulation and non-nodulation traits in alfalfa. Crop Sci 21:611–616

Putnoky P, Petrovics G, Kereszt A, Grosskopf E, Ha DTC, Banfalvi Z, Kondorosi A (1990) *Rhizobium meliloti* lipopolysaccharide and exopolysaccharide can have the same function in the plant-bacterium interaction. J Bacteriol 172:5450–5458

Schultze M, Kondorosi Á (1996) The role of Nod signal structures in the determination of host specificity in the *Rhizobium*-legume symbiosis. World J Microbiol Biotechnol 12:137–149

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Stougaard J (2000) Regulators and regulation of legume root nodule development. Plant Physiol 124:531–540

Stougaard J (2001) Genetics and genomics of root symbiosis. Curr Op Plant Biol 4:328–335

Truchet G, Barker DG, Camut S, De Billy F, Vasse J, Huguet T (1989) Alfalfa nodulation in the absence of *Rhizobium*. Mol Gen Genet 219:65–68

Turner SR, Ireland R, Morgan C, Rawsthorne S (1992) Idntification and localization of multiple forms of serine hydroxymethyltransferase in pea (*Pisum sativum*) and characterization of a cDNA encoding a mitochondrial isoform. J Biol Chem 267:13528–13534

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

A receptor kinase gene regulating symbiotic nodule development

Gabriella Endre, Attila Kereszt, Zoltán Kevei, Sorina Mihacea, Péter Kaló & György B. Kiss

Institute of Genetics, Biological Research Center of the Hungarian Academy of Sciences, H-6701 Szeged, PO Box 521, Hungary

Leguminous plants are able to establish a nitrogen-fixing symbiosis with soil bacteria generally known as rhizobia. Metabolites exuded by the plant root activate the production of a rhizobial signal molecule, the Nod factor, which is essential for symbiotic nodule development[1,2]. This lipo-chitooligosaccharide signal is active at femtomolar concentrations, and its structure is correlated with host specificity of symbiosis[3], suggesting the involvement of a cognate perception system in the plant host. Here we describe the cloning of a gene from *Medicago sativa* that is essential for Nod-factor perception in alfalfa, and by genetic analogy, in the related legumes *Medicago truncatula* and *Pisum sativum*. The identified 'nodulation receptor kinase', NORK, is predicted to function in the Nod-factor perception/transduction system (the NORK system) that initiates a signal cascade leading to nodulation. The family of 'NORK extracellular-sequence-like' (*NSL*) genes is broadly distributed in the plant kingdom, although their biological function has not been previously ascribed. We suggest that during the evolution of symbiosis an ancestral NSL system was co-opted for transduction of an external ligand, the rhizobial Nod factor, leading to development of the symbiotic root nodule.

The tetraploid alfalfa non-nodulation mutant MN-1008 (ref. 4) is known to lack all symbiotic responses upon inoculation with compatible *Sinorhizobium meliloti* or treatment with the corresponding cognate Nod factor[5–7]. But spontaneous nodulation in the absence of *Sinorhizobium*, a Nod-factor-independent phenotype reported so far only on alfalfa[8], could be detected in the mutant[9]. Consequently, the perception of the extrinsic signal was abolished in MN-1008, but the capacity for nodule initiation and development, as an internal developmental programme of the plant, remained intact. Like several other non-nodulation mutants, MN-1008 is also resistant to colonization by vesicular-arbuscular mycorrhizal fungi (the Myc^- phenotype)[10]. This genetic intersection between mycorrhizal and rhizobial associations has attracted significant interest because it suggests that symbiotic nitrogen fixation may be derived, in part, from the more ancient and widely distributed mycorrhizal symbiosis[11]. Thus, determining the molecular basis of mutations

such as in MN-1008 is central to understanding two of the most important symbiotic associations of plants, and their genetic overlap in legumes.

Map-based cloning of the genetic determinant responsible for the non-nodulation trait of *M. sativa* MN-1008 was initiated by mapping the Nod$^-$ phenotype in a tetraploid segregating population of alfalfa[12]. As this study demonstrated a high level of synteny between the *M. sativa* and *M. truncatula* genomes, a bacterial artificial chromosome (BAC) library of *M. truncatula*[13] was used to construct a contig spanning the mutation. Two tightly linked restriction-fragment length polymorphism (RFLP) markers[12], SHMT and CAD, were used to isolate primary BAC clones, and a 600-kilobase (600-kb) contig (the 'Nod contig') was assembled based on a combination of chromosome walking from BAC end sequences and restriction endonuclease fingerprinting of the clones. Genetic markers internal to the Nod contig were used to position the Nod$^-$ mutation within a fine-structure linkage map (Fig. 1a). Genotypes were scored in a segregating tetraploid *M. sativa* F_2 population (NAB population[12]) with the help of recombinant plants. This analysis revealed correspondence between the *M. truncatula* physical map and the genetic map positions in *M. sativa*. Moreover, two recombinant plants, NAB 4156 and 4443, served to delimit the MN-1008 non-nodulation mutation (*nn1*) to a ~160-kb interval (Fig. 1a, b). BAC clones containing the minimum tiling path of the Nod region (part of BAC 67A11, and the entirety of BACs 2D11 and 28I12; Fig. 1b) were subcloned and sequenced (data available at ⟨http://www.szbk.u-szeged.hu/ ~ alfi/⟩). A total of 15 open reading frames were identified and putative functions were assigned based on sequence homologies (see Fig. 1b and Supplementary Information). The predicted gene content of the Nod contig shared limited microsynteny with regions of the *Arabidopsis thaliana* genome (Fig. 1c), but functional information for the syntenic counterparts in *Arabidopsis* did not enhance our ability to identify candidate genes.

The non-nodulation phenotype of MN-1008 could be explained by an alteration in several candidate genes in this region, including genes predicted to code for an ATP-binding cassette (ABC) transporter, a receptor kinase, a MADS-box protein and lectin proteins. We first chose the ABC transporter and receptor kinase, membrane-spanning proteins to analyse for genetic alterations by comparing the sequence of alleles from wild-type nodulating alfalfa and the mutant MN-1008 plant material. To reduce confusion from different F_2 alleles, F_3 individuals that were homozygous throughout the Nod contig region were identified and used for sequencing. Based on the *M. truncatula* BAC sequences, exon-specific primers were designed to amplify both genomic and complementary DNA templates. At least three clones from independent amplifications of overlapping portions of both genes were analysed from the wild-type and mutant backgrounds of *M. sativa*. Based on this analysis, only synonymous changes were detected in the ABC transporter coding genes, whereas an unambiguous mutation, an 'in frame' stop codon, was identified in the receptor kinase gene from the mutant. This mutation was found consistently in independent amplification products from both genomic and cDNA templates of MN-1008 mutant. We have named this receptor kinase gene 'nodulation receptor kinase', *NORK*.

In addition to the analyses of *Medicago* species, we sequenced the candidate *NORK* gene from related legume species *Melilotus alba*, *P. sativum*, *Vicia hirsuta* and *Lotus japonicus*, and in each case the deduced proteins exhibited high overall homology. As shown in Fig.

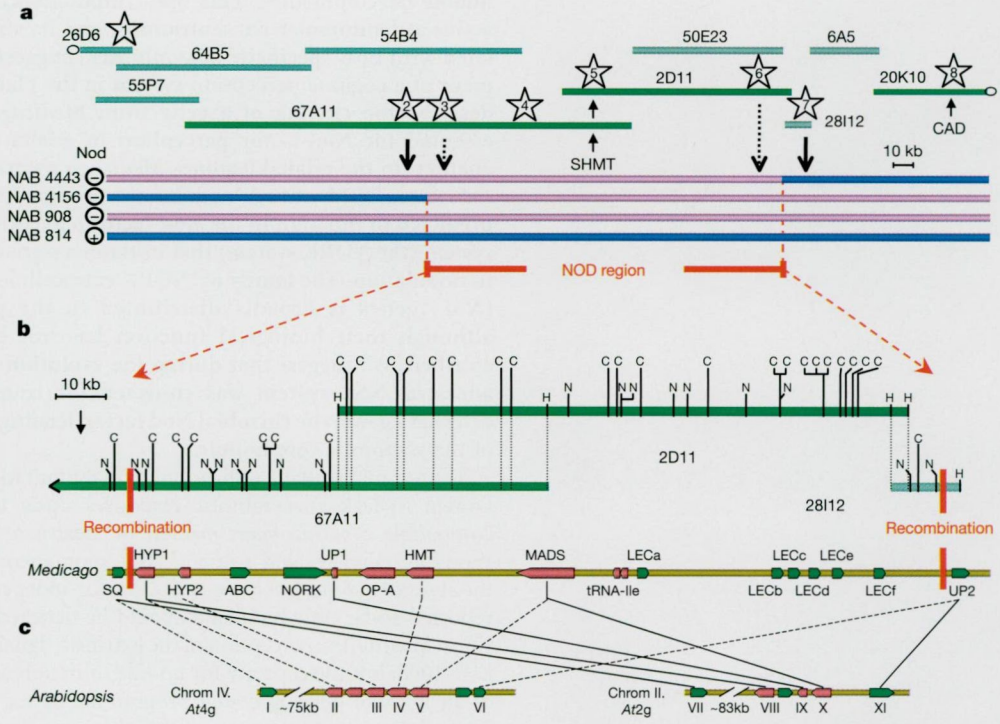

Figure 1 Genetic and physical map around the Nod locus. **a**, Primary BAC clones (67A11, 2D11 and 20K10) are coloured dark green, overlapping BAC clones (26D6, 55P7, 64B5, 54B4, 50E23, 28I12 and 6A5) are coloured light green. Fragments of the BAC clones (indicated as numbers in stars) were used for fine mapping. The Nod$^-$ and one Nod$^+$ genotypes are shown as pink and blue bars, respectively. **b**, Physical map and the gene content of the sequenced region. Restriction cut sites for *Cla*I (C) and *Nhe*I (N), and the *Hin*dIII (H) ends of the overlapping BACs are shown in the upper portion. Orientations of the open reading frames are indicated with green and red colours. **c**, Comparison of gene content with *A. thaliana* chromosomes 2 and 4 showing similarities to the sequenced region of *M. truncatula* I, *At*4g37760; II, *At*4g37900; III, *At*4g37920; IV, *At*4g37930; V, *At*4g37940; VI, *At*4g37960; VII, *At*2g22830; VIII, *At*2g22660; IX, *At*2g22640; X, *At*2g22630; XI, *At*2g22620.

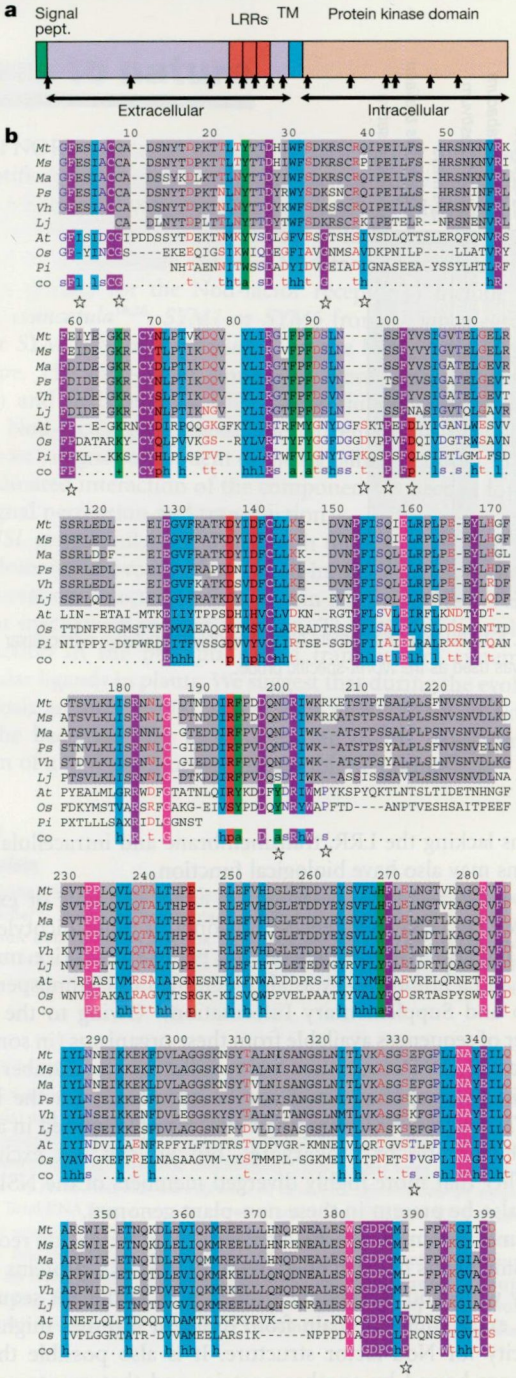

Figure 2 The NORK proteins. **a**, Schematic diagram of the NORK protein. The different domains are predicted on the basis of the deduced AA sequence. Vertical arrows indicate the location of the introns in the genomic DNA sequence. LRR, leucine-rich repeat; TM, transmembrane. **b**, Multiple alignment of the extracellular part of legume NORKs and three representatives of non-legume NSL proteins. *Mt*, *M. truncatula*; *Ms*, *M. sativa*; *Ma*, *Melilotus alba*; *Ps*, *Pisum sativum*; *Vh*, *Vicia hirsuta*; *Lj*, *Lotus japonicus*; *At*, *A. thaliana* LRRPK (accession number X97774); *Os*, *Oryza sativa* (accession number AAK92623); *Pi*, *Pinus taeda* (accession number BE758674); co, consensus. AA residues are coloured according to an 80% consensus (calculation included more NSL sequences, see Supplementary Information): hydrophobic (h: ACFGHIKLMRTVWY) and their aliphatic subset (l: ILV) with turquoise shading; aromatic (a: FHWY) and positive (+: HKR) with green shading; polar (p: CDEHKNQRST) with red shading; turn-like (t: ACDEGHKNQRST) in red; small (s: ACDGNPSTV) in blue. Dark violet background with white letters shows the position of 100% conserved AA, whereas pink background features the 80% consensus AAs. Grey background highlights conservation within legume NORKs, stars indicate the positions where AAs of the legume domain differ from the consensus of other NSL sequences.

2a, the predicted proteins possess an amino-terminal signal peptide motif, a putative extracellular domain containing three leucine-rich repeat (LRR) motifs[14], as well as a transmembrane segment and an intracellular domain with typical serine/threonine protein kinase signatures[15,16]. Several putative phosphorylation sites (see Supplementary Information) suggest a possible role also for (auto)phosphorylation in the protein's function. The predicted LRR motifs may indicate interaction with other protein(s)[17]. Although the N-terminal extracellular domain is well conserved among the legume proteins (Fig. 2b), this stretch, roughly 400 amino acids (AA) long, did not show homology to any known domain.

To further substantiate that in fact the mutation in the *NORK* gene was responsible for the non-nodulation phenotype, a similar sequencing strategy was undertaken on different Nod$^-$ mutants. In *M. truncatula* and *P. sativum*, mutagenesis programmes have yielded non-nodulation mutants with phenotypes and syntenic map positions that are suggestive of orthology with the alfalfa *nn1* gene[18,19]. Each sequenced *NORK* gene from the Nod$^-$ alleles of *M. truncatula* and *P. sativum* contained single base-pair alterations that are predicted to abolish NORK function. The nucleotide changes and their possible effects are shown in Table 1. Among the seven mutant alleles of *NORK*, six resulted in a Nod$^-$/Myc$^-$ phenotype. It is interesting to note that a non-synonymous substitution (see Table 1) in the mutant *M. truncatula* R38 resulted in Nod$^-$/Myc$^+$ phenotype. This suggests a composite role for NORK, branching the signal towards nodule formation and mycorrhiza colonization. The identification of molecular lesions in seven alleles, obtained from independent non-nodulation mutants in three legume species, provides strong circumstantial evidence of a role for NORK in nodulation and mycorrhizal association.

To directly test this hypothesis, we analysed the ability of the wild-type *NORK* gene of *M. truncatula* to complement the non-nodulation mutant, TR25, following transformation by *Agrobacterium rhizogenes*[20]. Roots appearing on the seedlings of the TR25 Nod$^-$ mutants after *Agrobacterium* infection were tested for nodulation by inoculating with *S. meliloti*. Plants transformed by *A. rhizogenes* carrying an empty vector control tested positive for β-glucuronidase (GUS) staining of transgenic roots, but were unable to form nodules. By contrast, TR25 plants infected by *A. rhizogenes* carrying the *NORK* gene on the same plasmid tested positive for GUS staining and developed nodules containing bacteroids (see Supplementary Information). The combination of the positive complementation tests and the genetic evidence provided by sequencing of multiple independent non-nodulation mutants, suggest that NORK is an essential component of the host plant nodulation signalling pathway.

To determine the range of legume and non-legume species in which the NORK protein is conserved, we used the DNA region coding for the postulated extracellular domain of NORK as a hybridization probe on Southern blots of total genomic DNA. Specific hybridizing bands could be detected (Fig. 3) in each of the legume genera tested (*Sesbania*, *Cassia*, *Trifolium*, *Desmodium*, *Vicia*, *Melilotus*, *Vigna*, *Macroptilium*, *Lotus*, *Glycine*, *Pisum*, *Phaseolus*, *Medicago*). Homology in *Cassia emerginata* is particularly interesting because this Caesalpinoid legume species lacks symbiotic nodule formation[21]. Genetic mapping of the *NORK* gene in *Medicago* demonstrated only one segregating locus (data not shown). Similarly, in most other legume species that we tested, only one strong hybridizing band was detected, suggesting the presence of a single copy gene. In a few cases, however, more than one equally hybridizing band could be detected, indicating either a single copy gene with an internal restriction site, or the presence of a small *NORK* gene family. Representative members of non-legume plants (maize, tobacco, wheat, rice and *Arabidopsis*) did not show a specific hybridization signal, indicating that high homology to *NORK* is restricted primarily to legumes.

Despite the absence of high nucleotide similarity outside of

33–57 (1998).
3. Schultze, M. & Kondorosi, A. What makes nodulation signals host-plant specific? *Trends Microbiol.* **3,**

in the EMBL/GenBank/DDBJ databases for supporting data sets are as follows: AJ418368; AJ418377; AJ418369; AJ418370; AJ418371; AJ418372; AJ418373; AJ418374; AJ418375; AJ418376; AJ418367; AJ428990; AJ428991.

MPMI Vol. 15, No. 9, 2002, pp. 922–931. Publication no. M-2002-0729-01R.

Glycine-Rich Proteins Encoded by a Nodule-Specific Gene Family Are Implicated in Different Stages of Symbiotic Nodule Development in *Medicago* Spp.

Zoltán Kevei,[1,2] José María Vinardell,[1] György B. Kiss,[2] Adam Kondorosi,[1,2] and Eva Kondorosi[1]

[1]Institut des Sciences du Végétal, CNRS UPR 2355, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France; [2]Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences, 6701 Szeged, Hungary

Submitted 17 December 2001. Accepted 23 May 2002.

Four genes encoding small proteins with significantly high glycine content have been identified from root nodules of *Medicago sativa*. All of these proteins as well as their *Medicago truncatula* homologues carried an amino terminal signal peptide and a glycine-rich carboxy terminal domain. All except nodGRP3 lacked the characteristic repeat structure described for cell wall and stress response-related glycine-rich proteins (GRP). Expression of these *GRP* genes was undetectable in flower, leaf, stem, and hypocotyl cells, whereas expression was highly induced during root nodule development, suggesting that *GRP* genes act as nodulins. Moreover, none of these nodule-expressed *GRP* genes were activated by hormones or stress treatments, which are inducers of many other *GRPs*. In *Rhizobium*-free spontaneous nodules and in nodules induced by a noninfective mutant strain of *Sinorhizobium meliloti*, all these genes were repressed, while they were induced in Fix⁻ nodules, unaffected in bacterial infection, but halted in bacteroid differentiation. These results demonstrated that bacterial infection but not bacteroid differentiation is required for the induction of the nodule-specific *GRP* genes. Differences in kinetics and localization of gene activation as well as in the primary structure of proteins suggest nonredundant roles for these GRPs in nodule organogenesis.

Additional keyword: leguminous plants.

Glycine-rich proteins (GRP) exhibit high structural diversity in plants. This large gene family is characterized by quasi-repetitive glycine-rich domains, most frequently with GGGX, GGXXXGG or GXGX repeats, resulting in proteins of distinct sizes (Sachetto-Martins et al. 2000). In most proteins, amino terminal signal peptides as well as other specific structures, such as cold shock domain (Kingsley and Palis 1994), cysteine-rich pattern (Rohde et al. 1990), or RNA-binding motifs (RNP-1 and RNP-2) (Dreyfuss et al. 1988) have also been identified.

The diverse expression pattern and subcellular localization of various *GRPs* suggest their implication in different physiological processes. Most GRPs are developmentally regulated and induced by biotic and abiotic factors. Salicylic acid (Hammond-Kosack and Jones 1996), abcisic acid (de Oliveira et al. 1990), methyl jasmonate (Molina et al. 1997), or ethylene induce many *GRP* genes (Memelink et al. 1990), while auxin has a negative effect on the expression of certain *GRPs* (Reddy and Poovaiah 1987). Pathogenic attacks such as viral (Linthorst et al. 1990) or fungal infections (Molina et al. 1997) also modulate *GRP* expression. Root knot nematode *Meloidogyne incognita* infection in alfalfa results in the induction of a GRP possessing an RNA-binding motif (Potenza et al. 2001).

Among the abiotic factors, regulation of *GRPs* can be elicited by osmotic stress (Gómez et al. 1988; Xu et al. 1995). Cold shock (Carpenter et al. 1994) and wounding (Showalter et al. 1991) activate *GRP* genes, while light (Kaldenhoff and Richter 1989) and circadian rhythm modulate *GRP* expression (Cretin and Puigdomenech 1990).

At the subcellular level, many GRPs were localized in the cell wall or were membrane-associated, but they were found also in the nucleus and cytosol (Sachetto-Martins et al. 2000). Organ- and tissue-specific expression studies showed the presence of *GRP* transcripts in the protoxylem (Keller et al. 1989), xylem, (Harrak et al. 1999) and phloem, as well as in epidermal tissues (Condit 1993). Moreover, anther-specific GRPs were identified from flowers (Mousavi et al. 1999), while other GRPs were detected during fruit development (Santino et al. 1997).

Based on their cell wall localization, several highly expressed GRP were reported as structural components of the plant cell wall, similar to proline-rich or hydroxyproline-rich proteins (Cassab 1998). Various studies on the regulation of the *GRP* genes suggest multiple roles in plant development. Functional studies revealed that GRPs might be implicated in cell lignification (Condit 1993) and mediating membrane-cytoskeleton or membrane-cell wall interconnections (Marty et al. 1996). GRPs with RNA-binding motifs might be involved in RNA processing or in the control of gene expression (Heintzen et al. 1994). Other GRP may be implicated in the stabilization of lipid-containing structures (Ross and Murphy 1996). So far, only the *Arabidopsis thaliana* AtGRP3 function has been demonstrated. This protein plays a role in the WAK1 (wall-associated receptor kinase) signaling pathway, where its binding to the WAK1 receptor kinase is necessary for the activation of KAPP, a kinase-associated phosphatase (Anderson et al. 2001; Park et al. 2001).

GRPs were also isolated from nitrogen-fixing root nodules of *Vicia faba* (Küster et al. 1995; Schröder et al. 1997). This unique plant organ develops in leguminous plants under com-

Corresponding author: E. Kondorosi; E-mail: Eva.Kondorosi@isv.cnrs-gif.fr.

Current address for J. M. Vinardell: Departamento de Microbiologia, Facultad de Biologia, Universidad de Sevilla, avda. Reina Mercedes, 41012 Sevilla, Spain.

bined nitrogen limitation in a symbiotic interaction with the rhizobial partner. Bacterial signal molecules (Nod factors) are required for nodule induction and different surface polysaccharides (EPS, KPS, and LPS) are involved in the invasion of plant cells. The indeterminate nodules developing in temporal legumes contain a persistent apical meristem (zone I), an invasion zone (zone II), in which cells arrested in division become infected with rhizobia and differentiate gradually along a few cell layers to reach the terminal phase for nitrogen fixation in zone III. In addition to housekeeping genes, consecutive activation of specific gene sets producing nodule-specific plant proteins (nodulins) are required for the various stages of nodule development.

Nodulin genes are classified as early and late nodulins, depending whether their expressions precede or accompany nitrogen fixation (Crespi and Gálvez 2000; Schultze and Kondorosi 1998). The *GRP* genes isolated from *Vicia faba* (*Vfnod-GRP1*, *Vfenod-GRP2*, *Vfenod-GRP3*, *Vfnod-GRP4*, and *Vfenod-GRP5*) exhibited nodule-specific expression, and tissue print hybridization localized their transcripts, mainly in the interzone II-III and the nitrogen-fixing zone III of the indeterminate nodules. Nodulin 24, a constituent of the peribacteroid membrane in soybean (Cheon et al. 1994) was also classified as GRP. In the nitrogen-fixing symbiosis between *Frankia* and the actinorhizal plant *Alnus glutinosa*, nodule-specific GRP were also detected in the cortical cells of infected nodules (Pawlowski et al. 1997). In *Medicago* spp., *NMs22* (Ganter et al. 1998) appears to encode a nodule-specific alfalfa glycine-rich protein. Screenings and sequencing of *Medicago* cDNA libraries in our laboratory also resulted in the identification of homologues of the broad bean *Vfenod-GRP5* from *M. sativa* (Jiménez-Zurdo et al. 2000) and from *M. truncatula* (Györgyey et al. 2000).

Our goal in this work was to isolate *GRP* genes from *M. sativa* and *M. truncatula* nodules and to get an insight into their function during nodule development. We addressed the following questions. i) Is the expression of the *GRP* genes nodule specific? ii) What are the requirements and signals for gene activation? iii) At what stage or stages of nodule development are the genes switched on? iv) Are the gene expressions specific for the different nodule zones or cell layers?

Our study revealed that all the isolated GRPs expressed exclusively in the nodules and encoded small proteins with a hydrophobic signal peptide that shared no homology with other GRPs in the databases. They represented nodulins as well as a novel subfamily of GRP that showed structural similarity only to *Vicia* GRP nodulins. The studied *GRP* genes were induced at different time points after inoculation and also exhibited differences in the kinetics and localization of gene expression. Moreover, we demonstrated that expression of these *GRP* genes was dependent on the infection of nodule cells by the microsymbiont *Sinorhizobium meliloti* but was not induced by abiotic or biotic factors. Analysis of the kinetics of gene expression and the localization of their transcriptions in nodules suggests that the nodule-specific GRPs might play nonredundant roles required at specific stages of nodule development.

RESULTS

Identification of cDNAs coding for GRP from *Medicago* nodules.

For the isolation of genes encoding GRP, a cDNA library made of young nodules of *Medicago sativa* subsp. *varia* A2 was hybridized with the *V. faba Vfnod-GRP1*, *Vfenod-GRP2*, *Vfenod-GRP3*, *Vfnod-GRP4*, or *Vfenod-GRP5* cDNA clones (Schröder et al. 1997) under low stringency to obtain both orthologue and related genes. These screenings resulted in the isolation of 37 cDNAs that represented four distinct genes encoding GRP. *MsnodGRP1* (accession number AF498986), represented by 11 cDNAs, was isolated with the *Vfnod-GRP1* probe that displayed limited homology (66% identity at nucleic acid level) with the probe. Screening with *Vfnod-GRP4* led to the isolation of 16 clones corresponding to two types of cDNAs; *MsnodGRP2A* and *MsnodGRP2B* (accession numbers AF498987 and AF498988, respectively) that were highly homologous to each other but distinct from *Vfnod-GRP4*. The *MsnodGRP3* (accession number AF498989) cDNAs (10 clones) were obtained with *Vfenod-GRP5* and exhibited 72% identity with the probe. In some cases due to the low stringency of hybridization, the screening also resulted in the isolation of non-nodule specific GRPs, like the homologue of polyadenylate-binding protein (Le et al. 1997) isolated from several plants and the recently described alfalfa GRP induced by root knot nematode (Potenza et al. 2001), which were not, however, studied further.

The *MsnodGRP1*, *MsnodGRP2A*, and *MsnodGRP2B* cDNA clones were partial. A homology search in the *M. truncatula* expressed sequence tag (EST) database has led, however, to the identification of the full-length sequences of these clones. All of these cDNAs coded for small putative proteins. MsnodGRP1 was 112, MsnodGRP2A was 101, MsnodGRP2B was 114, and MsnodGRP3 was 217 amino acids long (Fig. 1). All of them possessed a putative hydrophobic amino terminal secretory signal peptide predicted by the algorithm of von Heijne (1986) and a glycine-rich carboxy terminal part. Although, glycine-rich repeats are the major structural features of GRPs, in the MsnodGRP1, MsnodGRP2A, and MsnodGRP2B gene products, no glycine-rich repetitive blocks were present, in contrast to MsnodGRP3, which contained four repeats of 42, 40, 42, and 42 amino acids. Within these repeats, a WRDWG-

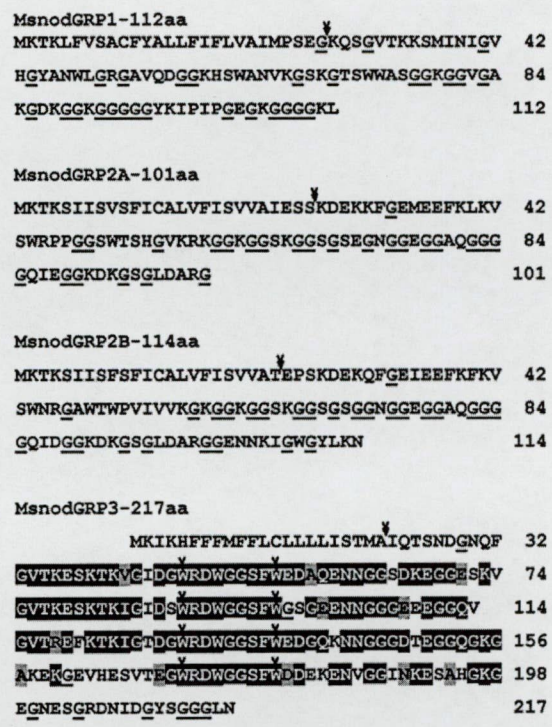

Fig. 1. Primary structure of glycine-rich proteins from *Medicago sativa* nodules. Glycine residues are in bold; signal peptides are underlined. In the four repeats of MsnodGRP3 (33-74, 75-114, 115-156, and 157-198), identical amino acid residues present in at least two repeats are on black background, while similar amino acids are shown on gray. The start and the end of WRDWGGSFW motifs are indicated with vertical arrows.

GSFW oligopeptide sequence was conserved (Fig. 1). An orthologue of the *MsnodGRP3* cDNA, *MtnodGRP3* (accession number AF498995), was found also in *M. truncatula* by systematic sequencing of a nodule cDNA library (*MtNo388*) in our laboratory (Györgyey et al. 2000). The MtnodGRP3 protein displayed 94% similarity and 88% identity to MsnodGRP3. The molecular structure of MsnodGRP3 resembled to that of NMs22 from alfalfa and that of Vfenod-GRP5 (Fig. 2A). Although the sizes of these proteins were different, all of them carried an N-terminal signal peptide and downstream repetitive glycine-rich blocks. In contrast to the four repeats in MsnodGRP3, NMs22 contained three repeats with

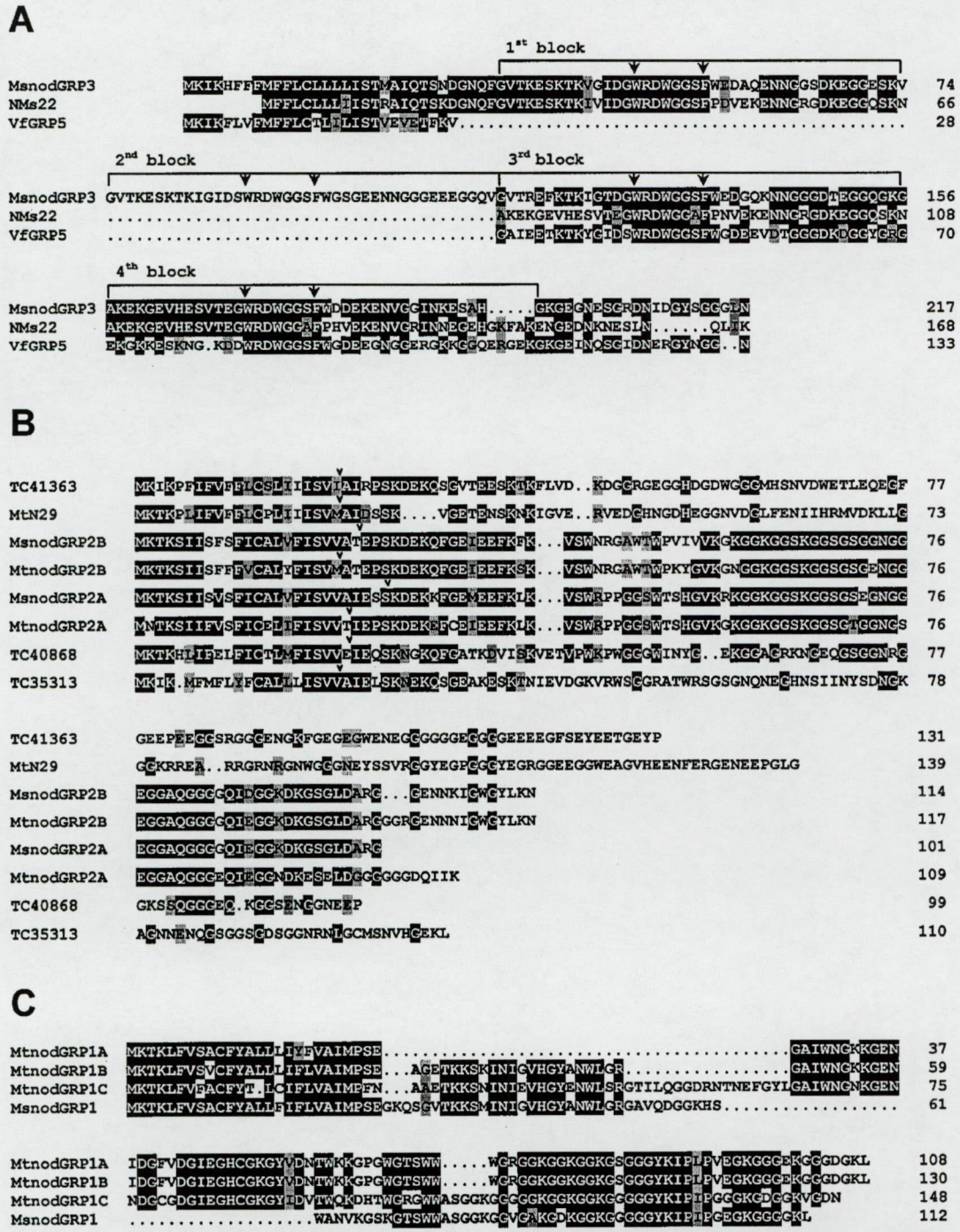

Fig. 2. Multiple alignments of homologous glycine-rich proteins. **A,** Comparison of the MsnodGRP3, Vfenod-GRP5 (VfGRP5), and alfalfa NMs22 proteins. Repetitive blocks of MsnodGRP3 are marked with horizontal arrows; conservative amino acid motifs are between vertical arrows. **B,** Multiple alignments of MsnodGRP2A, MsnodGRP2B, MtnodGRP2A, MtnodGRP2B, MtN29, and proteins deduced from the *Medicago truncatula TC40868*, *TC41363*, and *TC35313* expressed sequence tags. The end of putative signal peptides is marked with arrows in each sequence. **C,** Comparison of MsnodGRP1 and MtnodGRP1s. Identical amino acid residues are on black background, while similar amino acids are shown on gray. In **B,** and **C,** amino acid residues conserved in at least four sequences are indicated.

conservation of the WRDWGG oligopeptide, while Vfenod-GRP5 had only two repeats but with the same WRDWGGSFW oligopeptide sequence as MsnodGRP3. In contrast to MsnodGRP3, MsnodGRP1, lacking the characteristic glycine-rich repeats, shared no homology with known proteins.

MsnodGRP2A and MsnodGRP2B showed only a weak homology to Vfenod-GRP4 (62% similarity and 43% identity with MsnodGRP2A, 64% similarity and 46% identity with MsnodGRP2B). The MsnodGRP2A and MsnodGRP2B proteins exhibited 85% similarity and 80% identity to each other (Fig. 2B). To determine whether *MsnodGRP2A* and *MsnodGRP2B* represented allelic variants in the tetraploid, allogamous *M. sativa* genome or two distinct genes, a *M. truncatula* nodule cDNA library was screened with the *M. sativa* cDNA clones. This resulted in the isolation of both *MtnodGRP2A* and *MtnodGRP2B* (accession numbers AF498993 and AF498994, respectively) cDNA clones, revealing that these GRPs are encoded by different genes. The signal peptides of MtnodGRP2A and MtnodGRP2B proteins shared significant homology with the putative signal peptides of MtN29 (Gamas et al. 1996) and a few other GRP predicted from EST sequences (*TC40868*, *TC41363*, and *TC35313*) present in the *M. truncatula* databases (Fig. 2B). The respective genes also encoded small proteins that exhibited significantly high glycine content beside the putative hydrophobic signal peptide. The homology values between the signal peptides of these GRPs were around 80% or higher, while their glycine-rich domains displayed only 60 to 65% similarity and 30 to 35% identity. By screening the *M. truncatula* nodule cDNA library with the *MsnodGRP1* probe, we obtained three variants, *MtnodGRP1A*, *MtnodGRP1B*, and *MtnodGRP1C* (accession numbers AF498990, AF498991, and AF498992, respectively). The encoded proteins were homologous to MsnodGRP1, both at the N-terminal signal peptide region and in the C-terminal part. Downstream of the signal peptide, one oligopeptide block was conserved in MtnodGRP1B, MtnodGRP1C, and MsnodGRP1, but not in MtnodGRP1A; while in the central region a 30-amino-acid-long sequence was present in the MtnodGRPs but absent in MsnodGRP1 (Fig. 2C).

Expression of the four MsnodGRPs is nodule specific, requires rhizobial infection, and is unaffected by other GRP inducers.

Expression of the isolated *GRP* genes was investigated by Northern analysis using total RNA isolated from different organs of alfalfa (flower, root, hypocotyl, leaf, and stem) including young (7-day-old) and mature (21-day-old) nodules, as well as spontaneous nodules developed in the absence of *S. meliloti*. RNA samples (5 µg) were used for Northern hybridization, and equal loading was visualized by ethidium-bromide staining. The blots were also hybridized with the constitutive *Msc27* gene probe, which can be reliably used for normalization of the signals in roots and nodules in contrast to aerial plant organs, which express *Msc27* at different levels. No expression of any of the four *GRP* genes was detected in flowers, leaves, stems, hypocotyls, roots, or spontaneous nodules. In contrast, all were expressed in *S. meliloti*-induced nodules. Expression of the *MsnodGRP1* as well as the *MsnodGRP2A* and *MsnodGRP2B* genes was predominant in young nodules, while the *MsnodGRP3* transcript accumulation increased during nodule development (Fig. 3A). The *MsnodGRP1* probe revealed two transcripts, the smaller one corresponded to the size of the *MsnodGRP1* cDNA. The larger one might originate from a related gene; however, screening of the *M. sativa* subsp. *varia* A2 young nodule cDNA library resulted only in the isolation of *MsnodGRP1* in spite of the similar abundance of the two transcripts. Therefore, it may be that both transcripts derive from the *MsnodGRP1* gene but that they differ in their polyadenylation or alternative splicing during RNA maturation.

Northern analysis of GRP genes thus revealed that they were real nodulin genes. The lack of expression in spontaneous nodules demonstrated that bacterial infection is indispensable for the induction of *MsnodGRP1*, *MsnodGRP2A*, *MsnodGRP2B*, and *MsnodGRP3* genes. The possible requirement of bacteria for the expression of *MsnodGRP* genes was further studied with the help of two *S. meliloti* mutants that were wild type in respect to Nod factor production and initiation of nodule organogenesis but were defective in the invasion process or in their differentiation to nitrogen-fixing bacteroids. AK1492 is a dou-

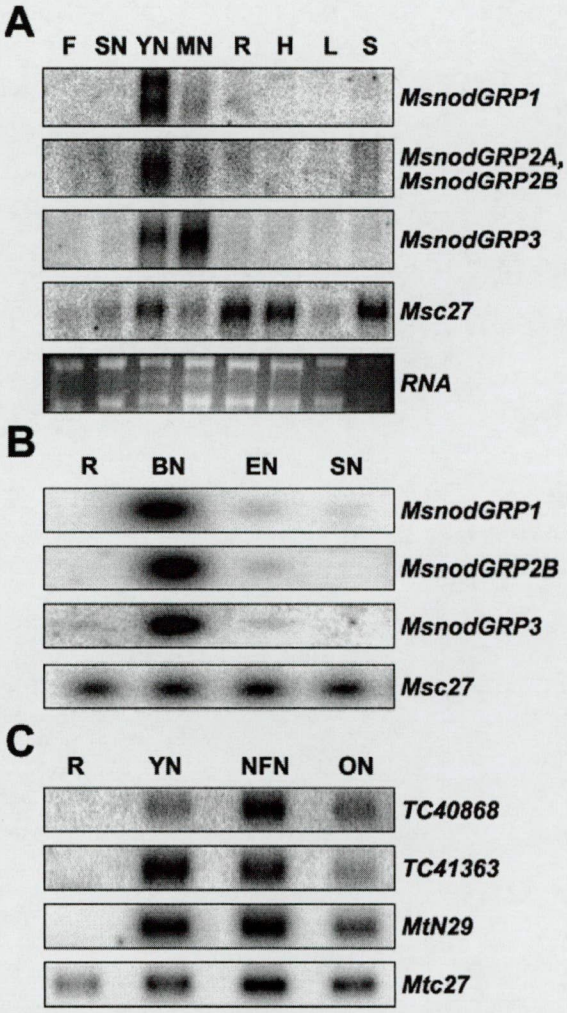

Fig. 3. Expression of *Ms/MtnodGRP* genes is nodule specific and requires *Sinorhizobium meliloti* infection. Hybridization probes are indicated on the right. Msc/Mtc27 and ethidium bromide stain of RNA gel served as controls for loading. **A,** Northern analysis of *nodGRPs* in alfalfa flowers (F), spontaneous nodules (SN), young nodules at 7 dpi (YN), mature nodules at 21 dpi (MN), roots (R), hypocotyls (H), leaves (L), and stems (S). **B,** Reverse transcription-polymerase chain reaction (RT-PCR) analysis of *MsnodGRP* expression in alfalfa roots (R), nodules induced by Bac$^-$ (BN) and Exo$^-$ (EN) *S. meliloti* mutants and by spontaneous nodules (SN) developed in the absence of rhizobia. **C,** RT-PCR results of the expression of *TC40868*, *TC41363*, and *MtN29* genes in *Medicago truncatula* R-108 roots (R), young nodules collected 9 dpi (YN), nitrogen-fixing nodules collected 20 dpi (NFN), and nodules at 29 dpi (ON).

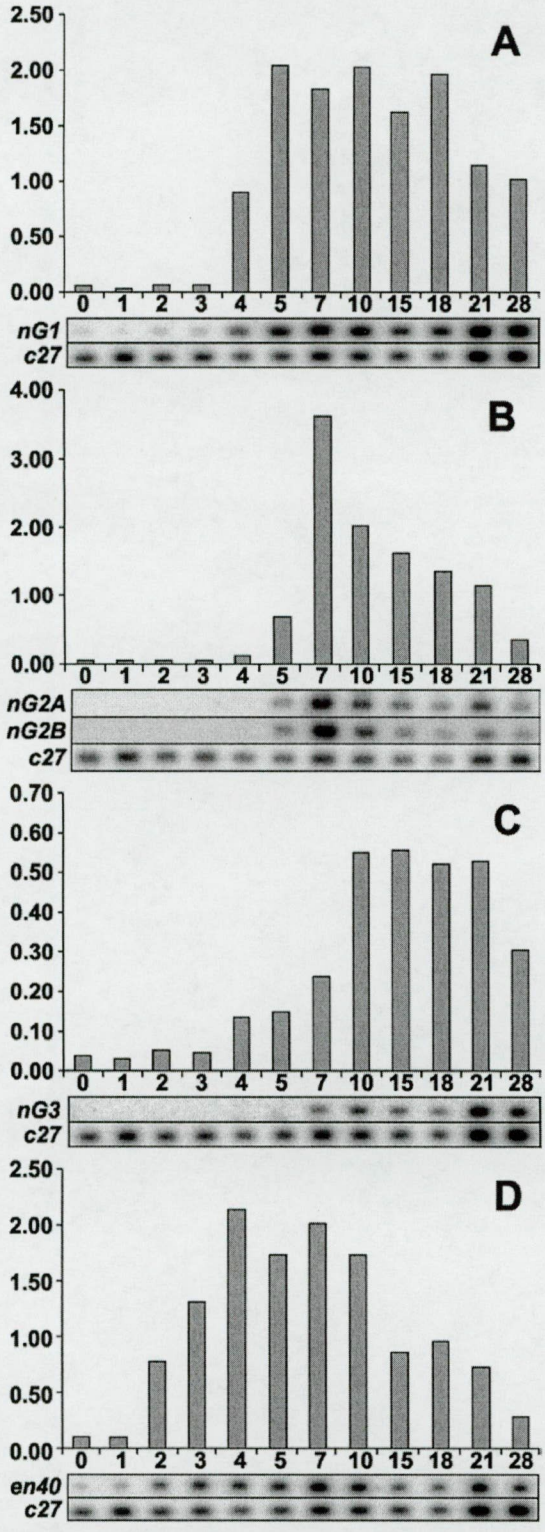

Fig. 4. Reverse transcription-polymerase chain reaction (RT-PCR) analysis of *MsnodGRP* and *Msenod40* genes in the course of nodule development. cDNAs were synthesized from total RNAs isolated from roots and nodules at 0, 1, 2, 3, 4, 5, 7, 10, 15, 18, 21, and 28 dpi and were used for PCR with gene-specific primers **A,** *nG1* for *MsnodGRP1*, **B,** *nG2A* and *B* for *MsnodGRP2A* and *B*, respectively, **C,** *nG3* for *MsnodGRP3*, and **D,** *en40* for *Msenod40*. PCR products were detected by Southern hybridization. Histograms show the normalized hybridization signal intensities calculated on the basis of *Msc27* (*c27*) signals.

ble mutant of *S. meliloti* 41, unable to produce wild-type exopolysaccharide (EPS) and capsular polysaccharide (KPS) and thereby impaired in the infection process ($Exo^-Kps^-Inf^-$ mutant). The other mutant, strain 8368, produced wild-type polysaccharides but was defective in bacteroid differentiation (Bac^-). In the nodules induced by strain AK1492, the expression of all four *GRP*s was drastically reduced, in contrast to the nodules induced by the Bac^- mutant in which expression of *GRP* genes was unaffected (Fig. 3B). These results indicated that the expression of these *GRP* genes might be controlled by bacterial polysaccharide signals or that they might require invasion of the nodule cell by *Rhizobium*, or both.

Since it was shown that expression levels of several *GRP* genes were regulated by hormones and osmotic stress, we investigated whether expression of these nodule-specific, rhizobia-induced, *GRP* genes was induced by factors other than rhizobia. Treatment of alfalfa roots with naphthalene acetic acid (NAA), absicic acid (ABA), kinetin and gibberelic acid (GA_3), or with exposure to heat (37°C) and cold (4°C) shocks, hypoxia ($+H_2O$), and drought ($-H_2O$) stresses, however, did not result in the expression of the nodule-specific GRPs tested by reverse transcription-polymerase chain reaction (RT-PCR) (data not shown). These results further strengthen the nodule-specific function of these GRPs.

MsnodGRP2A- and *MsnodGRP2B*-related *M. truncatula* genes are also nodule specific.

Although the proteins encoded by the *M. truncatula* ESTs *TC40868* and *TC41363* as well as MtN29 exhibited homology to Ms/MtnodGRP2 only in the signal peptide region, their relatively high glycine content prompted us to study their expression profiles. Using specific oligos, each gene was amplified from cDNAs prepared from mRNAs isolated from *M. truncatula* R108 nodules of different ages and roots treated with Nod factors (Fig. 3C). No expression of any of these genes was detected in roots, while all of them were induced during nodule development, albeit with different kinetics. The maximal transcript level for *TC40868* was detected in the nitrogen-fixing nodules, while for *TC41363*, it was in the young nodules. Expression of *MtN29* was similar in the young and the nitrogen-fixing nodules. As nodules aged, the transcript level of all the three genes declined.

MsnodGRP nodulins exhibit different temporal expression patterns.

The induction of GRP genes during nodule development was studied by RT-PCR experiments. Total RNA samples for cDNA preparation were isolated from roots of sterile seedlings grown on agar plates in nitrogen-free medium after inoculation of roots with *S. meliloti* at 0 to 3 days post inoculation (dpi), from nodule primordia (4 to 5 dpi), and from young (7 to 10 dpi) and mature nodules (15 to 28 dpi). Accumulation of the *GRP* mRNAs was measured during nodule development and compared to that of *Msc27* used as a control for the quantification of the cDNAs in the different samples (Fig. 4). As a control for nodule development, expression of early nodulin gene *Msenod40* was included (Fig. 4D). In the case of *MsnodGRP1*, a weak background signal was detectable in the root, while a significant increase in mRNA abundance was observed at 4 dpi, reaching its maximum at 5 dpi, which was maintained during the development and maturation of nodules (Fig. 4A). *MsnodGRP2A* and *MsnodGRP2B* had no background signal in the root and were strongly induced at 5 dpi (1 day later than *MsnodGRP1*), with a maximal transcript accumulation at 7 dpi that then decreased gradually (Fig. 4B). Induction of the *MsnodGRP3* gene started at day 7 and was maximal through the active, nitrogen-fixing period (Fig. 4C). In comparison with

nodGRP genes, expression of *Msenod40* was induced at 2 dpi, thus 2 days earlier than *MsnodGRP1*, but both of them before nitrogen fixation. Expression of *MsnodGRP2A* and *MsnodGRP2B* exhibited a three-day delay and *MsnodGRP3* a five-day delay compared with *Msenod40*. These differences in the expression kinetics indicated that *MsnodGRP*s might be involved in consecutive stages of nodule development with different functions.

Distinct spatial expression patterns of the GRP nodulin genes.

To determine the spatial expression patterns of *GRP* genes in nodules, in situ hybridizations were carried out with sense and antisense RNA probes that were synthesized from cDNAs encoding the glycine-rich domains and that were specific for *MsnodGRP1*, *MsnodGRP2*, and *MsnodGRP3* but did not distinguish *MsnodGRP2A* from *MsnodGRP2B*. The sections

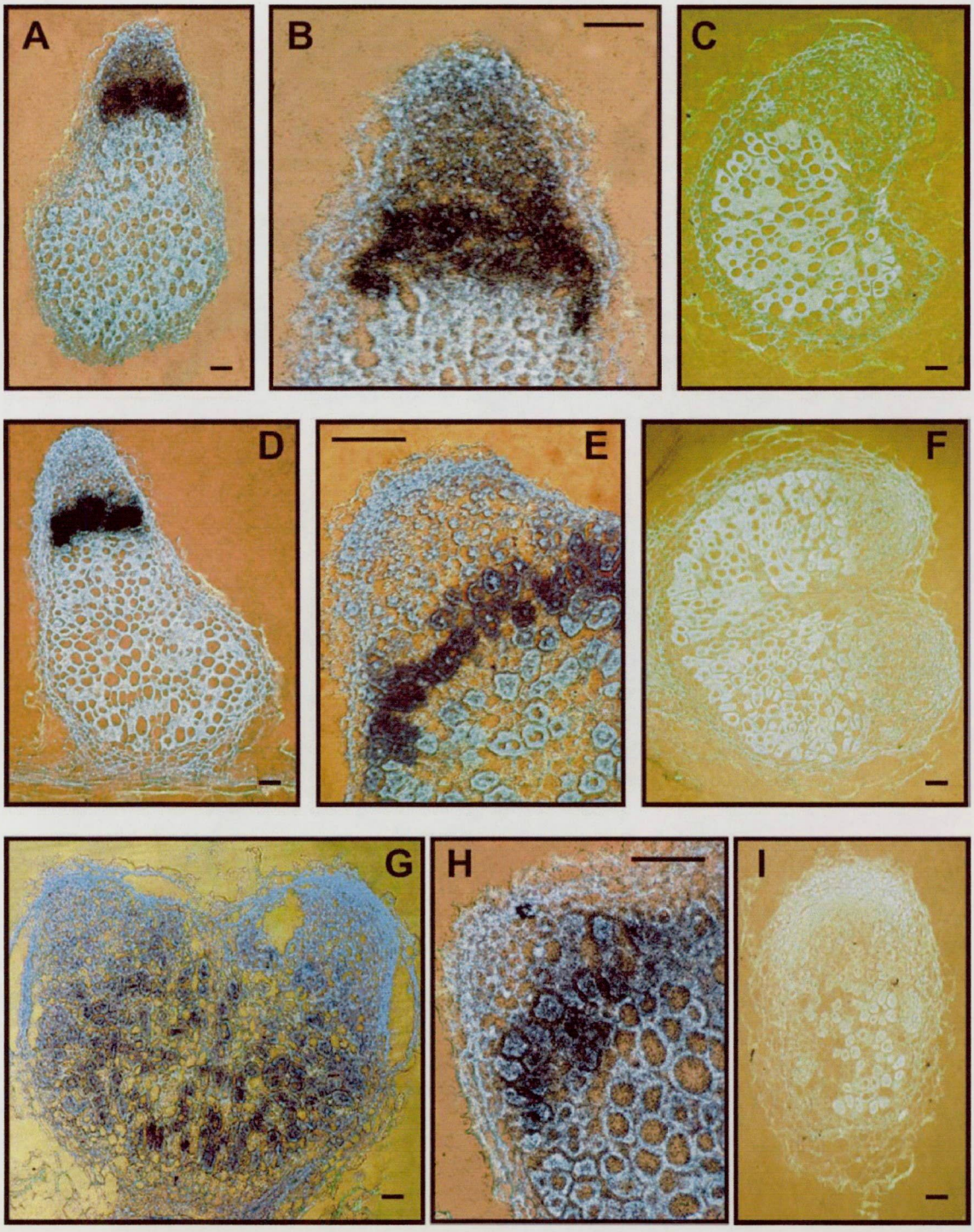

Fig. 5. Localization of *MsnodGRP* transcripts in *Medicago sativa* nodules by in situ hybridization. Hybridization of antisense riboprobes specific for **A** and **B**, *MsnodGRP1*, **D** and **E**, *MsnodGRP2*, and **G** and **H**, *MsnodGRP3*. Hybridization with sense probes of **C**, *MsnodGRP1*, **F**, *MsnodGRP2*, and **I**, *MsnodGRP3*. Bars correspond to 200 µm.

derived from alfalfa nodules of different ages that represented all stages of nodule development. The *MsnodGRP1* transcripts were localized in the nodule apex in zone II, while no signals were detected in the nitrogen fixation zone III (Fig. 5A). In zone II, the highest expression level was found in cells adjacent to interzone II-III, where *MsnodGRP1* transcripts were also present (Fig. 5B). The *MsnodGRP2* gene exhibited expression exclusively in interzone II-III (Fig. 5D) and predominantly in a single cell layer (Fig. 5E). In the case of *MsnodGRP3*, no hybridization signals were detected in the nodule apex, and the expression appeared to be associated with the nitrogen-fixing symbiotic cells in zone III (Fig. 5G). In a few cases, the hybridization signals were stronger in the early symbiotic zone than in the older cells of zone III (Fig. 5H). The lack of background signals on nodule sections hybridized with the sense RNA probes (Fig. 5C, 5F, 5I) and the restricted signal localizations obtained with the antisense probes confirmed that these hybridizations were specific for all three *nodGRPs*.

DISCUSSION

In this paper, we presented the isolation and characterization of cDNAs coding for different GRP from *M. sativa* and *M. truncatula* nodules. These proteins, like most GRP, were composed of a secretory signal peptide followed by a glycine-rich domain. However, their glycine content was lower (20 to 30%) than that of other GRPs (where it was around 80% or even higher) although still significantly higher than in other proteins not considered rich in glycine. In addition, these nodulin genes were exclusively expressed in *Medicago* nodules, confirming that they represent a novel subfamily of the *GRP* genes.

All nodule-specific GRPs described in this study were small proteins, and most of them lacked the glycine-rich repetitive motifs characteristic for non-nodule-specific GRPs (Sachetto-Martins et al. 2000). In these proteins, only few repetitions of a GGX sequence were present. MsnodGRP3 was the only one that contained four repeats of an oligopeptide with the WRDWGGSFW motif. This motif was present in two other nodule-specific proteins, NMs22 (Ganter et al. 1998) and Vfenod-GRP5 (Schröder et al. 1997) that differed from MsnodGRP3, however, in the number of repeats. The protein encoded by a *M. truncatula* EST, TC40954, and predicted as a homologue of NMs22 (NMt22) had only two oligopeptides of the four repeats, likewise Vfenod-GRP5. These differences suggest that the functions of these nodulins are distinct from that of MsnodGRP3. Moreover, database search carried out with the signal peptide of MsnodGRP3 revealed similarity to the signal peptide of nodulin DD17, a recently available sequence in the database (accession number CAC84521) from another leguminous plant, *Trifolium repens*, which also contained several glycine residues at carboxy terminal region.

The primary structures of MsnodGRP2A and MsnodGRP2B were highly homologous and were encoded by different genes, as it was revealed by isolation of the corresponding *M. truncatula* clones. These proteins showed a weak homology to Vfnod-GRP4, excluding the signal peptide that was present in alfalfa but not in broad bean. In contrast, the signal peptides of MsnodGRP2A and MsnodGRP2B exhibited significant homologies with three proteins predicted from *M. truncatula* EST sequences (TC40868, TC41363, and TC35313). These proteins also displayed glycine-rich domains but without homology to Ms/MtnodGRP2A and Ms/MtnodGRP2B or to each other. One of the signal peptide homologues was MtN29, identified as a nodulin (Gamas et al. 1996). Moreover, a similar signal peptide was present in the early nodulin Enod7 (accession number CAA63660) from *Pisum sativum*, but the glycine content of this putative protein was relatively low.

MsnodGRP1 was the only GRP that did not show homology to other proteins in the databases. Screening of a *M. truncatula* library with the *MsnodGRP1* probe resulted in three homologues but not the real orthologues of the alfalfa protein. The homology was extended to the signal peptide and the carboxy terminal part of the proteins, while the central regions were less conserved (Fig. 2C).

Relatedness of the GRP nodulins is presented in a phylogenetic tree (Fig. 6). As it was expected, MsnodGRP1 and the MsnodGRP1-related *M. truncatula* proteins MtnodGRP1A, MtnodGRP1B, and MtnodGRP1C formed one group. Interestingly, the signal peptide homologues of Ms/MtnodGRP2s, TC40868, TC41363, and TC35313, were on distinct branches from the Ms/MtnodGRP2 proteins. TC40868 and TC41363 proteins were in the group of Ms/MtnodGRP3 proteins. This group also included NMs22/NMt22, Vfenod-GRP5, DD17, and Vfenod-GRP3. TC35313 was more related to nodGRP1 proteins. The MsnodGRP2 and MtnodGRP2 proteins were related to Vfnod-GRP4. MtN29, the signal peptide homologue of Ms/MtnodGRP2s, exhibited distinct origin. Taking into account that most of these nodGRPs showed no homology to other GRP, these nodulation-related GRPs represent both a novel class of GRPs and a novel group of nodulins.

Expression of all *Medicago* nodGRPs was nodule specific and required infection of rhizobia; however, their activation and maximal transcript accumulation differed during nodulation. The *MsnodGRP1* gene was induced at 4 dpi, two days later than *Msenod40*, an early nodulin gene that is required for cell proliferation and for initiation of the nodule primordium (Crespi et al. 1994). At day 4, nodule primordium differentiates to a meristem and an invasion zone in which infection of plant cells with bacteria occurs. Consistently with the expression kinetics, *MsnodGRP1* transcripts were localized by in situ hybridization along all cell layers of zone II, where they were particularly abundant in distal cell layers and in interzone II-III but were absent in zone III. This expression profile suggests that nodGRP1 might be involved in differentiation of symbiotic tissue. This localized zone II-associated expression pattern also explains the decrease in the transcript level during nodule maturation detected by Northern analysis and RT-PCR. By nodule growth, the mass of zone III increases while that of zone II remains constant, thus, although the expression level of *MsnodGRP1* does not change, the concentration of their transcripts decreases with the increasing amount of mRNAs from zone III. From *M. truncatula*, we have isolated three homologous *nodGRP1* cDNAs (*MtnodGRP1A*, *MtnodGRP1B*, and *MtnodGRP1C*). This suggests that additional nodGRP1-related proteins might be produced in *M. sativa* nodules as well.

Transcripts of MsnodGRP2A and MsnodGRP2B were indistinguishable by Northern analysis, due to the cross-hybridization of the genes. Using specific oligos for each gene, RT-PCR experiments demonstrated that neither of the two genes expressed in roots and that both were activated at 5 dpi. At days 5 to 7, the distal cell layers of zone II are converted into nitrogen-fixing symbiotic cells. During this period, transcript accumulation of *MsnodGRP2*s increased fivefold and then decreased gradually. In situ hybridization revealed the exclusive presence of *MsnodGRP2* transcripts in interzone II-III in two or a maximum of three cell layers, where the most drastic and probably the most critical developmental changes occur prior to terminal differentiation of nitrogen-fixing symbiotic cells. This expression profile is in perfect agreement with transient expression of *MsnodGRP2* detected by RT-PCR analysis, revealing an illusive and more drastic decrease in *MsnodGRP2* transcript levels than in those of *MsnodGRP1* by nodule growth and maturation. As revealed by RT-PCR analysis of root and nodule RNAs, expression of the signal peptide homo-

logues of Ms/MtnodGRP2s (TC40868, TC41363, and MtN29) was nodule specific as well.

Northern analysis demonstrated that expression of *MsnodGRP3* was induced at 7 dpi in the young nitrogen-fixing nodules and was predominant during the period of active nitrogen fixation. By in situ hybridization, the *MsnodGRP3* transcripts were detected in the nitrogen-fixing cells in zone III that correlated with the relatively retarded induction of *MsnodGRP3* compared with *MsnodGRP1* and *MsnodGRP2A/B* and with the elevated transcript levels through the active nitrogen-fixing period.

The expression pattern of *nodGRP* genes in nodules induced by *S. meliloti* mutant strains demonstrated that activation of all the *nodGRP* genes we studied required bacterial infection but was indifferent of bacteroid differentiation. Differential activation kinetics of these genes and distinct spatial localization of the *nodGRP* mRNAs indicate that they might have distinct functions during nodule organogenesis. Their expression in particular zones and cell layers suggests that these nodGRPs may be involved in signaling processes, similar to AtGRP3, which is a binding partner of the cell wall-integrated WAK1 protein that is responsible for development of leaf cells (Kohorn 2001). During nodule differentiation, when bacteria invade the host cells, nodGRP1 and nodGRP2A/B might function or be involved as signals for the production of specific molecular structures (such as peribacteriod membranes) that are required for the development of symbiotic cells. The function of nodGRP2A/B appears to be, however, more specific and might be required for terminal differentiation of symbiotic cells. In contrast to nodGRP1 and nodGRP2A/B, nodGRP3 plays a role in a later stage of symbiosis associated with nitrogen fixation. The possible involvement of these nodGRPs in *Rhizobium*-elicited signaling processes is in line with their exclusive nodule-specific expression and with the finding that none of the *MsnodGRP*s were induced by biotic and abiotic regulators of the other, non-nodule-specific GRPs.

Elucidation of the real nodGRP functions would necessitate the analysis of knock out *M. truncatula* mutants. Although, construction of tagged *M. truncatula* mutant banks is in progress in the laboratory, their creation will take several years. Therefore, as alternative approaches for gene silencing, we have started the generation of antisense and RNAi transgenic *M. truncatula* plants to evaluate the importance of nodGRPs in nodule development. For transgenic studies in nodulation, isolation of the promoter regions of *nodGRP* genes could also be extremely useful to drive localized and nodule-specific gene expression. The identification of interacting protein partners of the nodGRPs is expected to give an insight into putative signaling pathways. Therefore, yeast two-hybrid screens will be performed for isolation of nodGRP partners. All these future studies will contribute to elucidation of the biological functions of nodGRPs, which appear to be specific for leguminous plants forming indeterminate nodules.

MATERIALS AND METHODS

Bacterial strains.

In the *M. sativa* and *M. truncatula* nodulation assays, the following *Rhizobium* strains were used: wild-type *S. meliloti* Rm41; Sm8368, a Bac$^-$ derivate of Sm1021 carrying a Tn*phoA* insertion in the *bacA* gene (Glazebrook et al. 1993); and AK1492, an Eps$^-$ Kps$^-$ Fix$^-$ mutant strain of Rm41 containing a deletion in *ExoB* and a Tn5 insertion in the *fix23* region encoding capsular polysaccharide production (Kondorosi et al. 1984; Putnoky et al. 1988). Bacteria were grown at 30°C in tryptone-agar medium (Orosz et al. 1973) that was supplemented for the mutant strains with kanamycin (100 μg ml^{-1}).

Plant material.

Seeds of alfalfa (*M. sativa* subsp. *sativa* cv. Sitel) and *M. truncatula* ecotype R-108 (Hoffman et al. 1997) were sterilized by gentle agitation in 20% Pasti-chlor (Argos, Bondoufle, France) for 30 min at room temperature, then washed four times with sterile water, and germinated in the dark on 1% water agar plates for 24 h. For the time course experiments, the seedlings were transferred to square petri dishes containing nitrogen-free solution (Crespi et al. 1994) with 1% agar and incubated in a growth chamber at 28°C. After 4 to 5 days, the roots were spot-inoculated with *S. meliloti* Rm41 and samples were collected from 1-, 2-, 3-, 4-, 5-, 7-, 10-, 15-, 18-, 21-, and 28-day-old nodule primordia or nodules, as well as from noninoculated roots as a control.

For large scale RNA isolation, seedlings were grown in aeroponic factories (Hoffman et al. 1997). For nodulation assay, seedlings were grown for one week in nitrogen-limiting medium and then inoculated with wild-type or mutant *S. meliloti* strains. Wild-type nodules were harvested at 7 and 21 dpi from *M. sativa* and at 9, 20, and 29 dpi from *M. truncatula*. In the case of Bac$^-$ and Exo$^-$ mutants, the nodules were collected 2 to 3 weeks after infection. Flowers, roots, hypocotyls, leaves, and stems of alfalfa were collected from 2- to 3-week-old plants grown in nitrogen-containing solution. Spontaneous nodules were collected from the highly embryogenic *M. sativa* subsp. *varia* A2 line after three weeks of nitrogen starvation.

Library screening and sequencing.

Alfalfa GRP cDNA clones were obtained by screening a cDNA library of young *M. sativa* subsp. *varia* A2 nodules (Crespi et al., 1994) using the *Vfnod-GRP1*, *Vfenod-GRP2*, *Vfenod-GRP3*, *Vfnod-GRP4*, and *Vfenod-GRP5* cDNA probes of *Vicia faba* (Küster et al. 1995; Schröder et al. 1997). *MtnodGRP2A* and *MtnodGRP2B* cDNA clones were isolated from a *M. truncatula* R-108 nodule cDNA library (Györgyey et al. 2000). The cDNA clones were sequenced by the chain terminator method (Pharmacia kit; Pharmacia SAS, Saint Quentin-en-Yveline, France) using an automatic laser sequencer 373A (Applied Biosystems, Foster City, CA, U.S.A.).

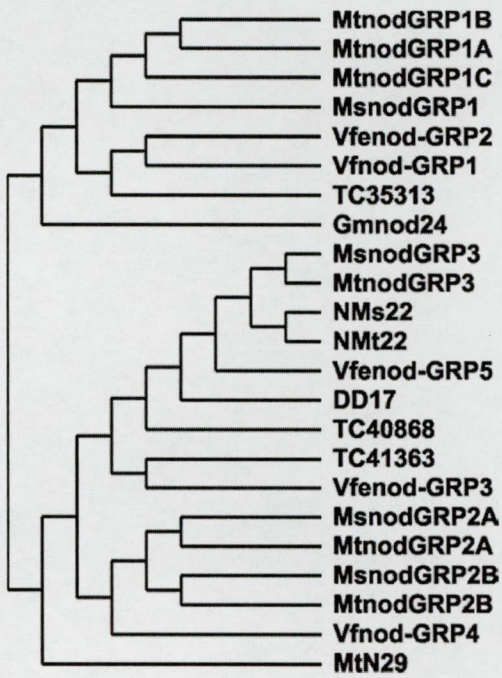

Fig. 6. Phylogenetical tree of MsnodGRPs and their homologues.

Northern and RT-PCR analysis

Total RNA was isolated from alfalfa by the RNeasy plant mini kit (Qiagen, Hilden, Germany) from nodules, roots, leaves, stems, flowers, and hypocotyls. For Northern blots, 8 μg of total RNA was used. For RT-PCR, cDNAs were synthesized by the treatment of 1 μg of total RNA with DNase (FPLC pure; Amersham Pharmacia Biotech, Uppsala, Sweden) that was retrotranscribed in the presence of M-MLV reverse transcriptase (Promega, Madison, WI, U.S.A.), RNase inhibitor (RNasin, Promega) and oligo-dT primers. *Msc27* (Györgyey et al. 1991) and *Mtc27* were amplified in 17 cycles (94°C for 30 s, 55°C for 30 s, and 72°C for 1 min) while *MsnodGRP1*, *MsnodGRP2A*, *MsnodGRP2B*, *MsnodGRP3*, *TC40868*, *TC41363*, and *MtN29* were amplified in 21 cycles. After agarose gel electrophoresis, RT-PCR products were blotted on nylon membranes (BioTrans; ICN, Costa Mesa, CA, U.S.A.). α[^{32}P]-dCTP radiolabeled DNA probes were generated by random priming (Amersham Pharmacia Biotech). Hybridizations were performed in 0.5 M Na-phosphate, pH 7.2, and 7% sodium dodecyl sulfate (SDS) buffer (Church and Gilbert 1984) for 36 h at 65°C; then the membranes were washed twice with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 0.1% SDS and with 1× SSC and 0.1% SDS. Hybridization signals were exposed on phosphor-imager screens. For hybridization signals, intensities were quantified and normalized by Phosphor-imager STORM (Molecular Dynamics, Sunnyvale, CA, U.S.A.) and specific computer packages.

In situ hybridization.

In situ hybridizations were carried out according to de Almeida Engler and associates (2001) using 200-bp-long ^{35}S-labeled sense and antisense riboprobes (corresponding to the glycine-rich domains). Slides were exposed to Kodak NTB-2 emulsion, and signals were detected with fluorescent-light microscope and CDD-camera.

ACKNOWLEDGMENTS

We are grateful to P. Mergaert and K. Kelemen for providing cDNAs for nodulation kinetics and analysis of stress and hormone effects. We thank M. Redondo-Nieto for the preparation of the figures in the manuscript. Z. K. was supported by a co-tutelle Ph.D. grant of the French (C.R.O.U.S.) and Hungarian governments and by the "Jumelage" program between the C.N.R.S. and the Hungarian Academy of Sciences, while J.M.V. was supported by the TMR-Marie Curie fellowship program.

LITERATURE CITED

Anderson, C. M., Wagner, T. A., Perret, M., He, Z., He, D., and Kohorn, B. D. 2001. WAKs: Cell wall-associated kinases linking the cytoplasm to the extracellular matrix. Plant Mol. Biol. 47:197-206.

Carpenter, C. D., Kreps, J. A., and Simon, A. E. 1994. Genes encoding glycine-rich *Arabidopsis thaliana* proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. Plant Physiol. 104:1015-1025.

Cassab, G. I. 1998. Plant cell wall proteins. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49:281-309.

Cheon, C. I., Hong, Z., and Verma, D. P. 1994. Nodulin-24 follows a novel pathway for integration into the peribacteroid membrane in soybean root nodules. J. Biol. Chem. 269:6598-6602.

Church, G. M. and Gilbert, W. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. U.S.A. 81:1991-1995.

Condit, C. M. 1993. Developmental expression and localization of petunia glycine-rich protein 1. Plant Cell 5:277-288.

Crespi, M. and Gálvez, S. 2000. Molecular Mechanisms in Root Nodule Development. J. Plant Growth Regul. 19:155-166.

Crespi, M. D., Jurkevitch, E., Poiret, M., d'Aubenton Carafa, Y., Petrovics, G., Kondorosi, E., and Kondorosi, A. 1994. *Enod40*, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO (Eur. Mol. Biol. Organ.) J. 13:5099-5112.

Cretin, C., and Puigdomenech, P. 1990. Glycine-rich RNA-binding proteins from *Sorghum vulgare*. Plant Mol. Biol. 15:783-785.

de Almeida Engler, J., De Groodt, R., Van Montagu, M., Engler, G. 2001. *In situ* hybridization to mRNA of *Arabidopsis* tissue sections. Methods 23:325-334.

de Oliveira, D. E., Seurinck, J., Inze, D., Van Montagu, M., and Botterman, J. 1990. Differential expression of five Arabidopsis genes encoding glycine-rich proteins. Plant Cell 2:427-436.

Dreyfuss, G., Swanson, M. S., and Pinol-Roma, S. 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trends Biochem. Sci. 13:86-91.

Gamas, P., de Carvalho Niebel, F., Lescure, N., and Cullimore, J. 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.

Ganter, G., Raja, S., and Dunn, K. 1998. The cDNA sequence of Nms-22 (accession no. AF030252), a gene whose message is specific to the infected cells of alfalfa (*Medicago sativa*) root nodules. Plant Physiol. 116:446.

Glazebrook, J., Ichige, A., and Walker, G. C. 1993. A *Rhizobium meliloti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. Genes Dev. 7:1485-1497.

Gómez, J., Sánchez-Martínez, D., Stiefel, V., Rigau, J., Puigdomenech, P., and Pages, M. 1988. A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. Nature 334:262-264.

Györgyey, J., Gartner, A., Németh, K., Magyar, Z., Hirt, H., Heberle-Bors, E., and Dudits, D. 1991. Alfalfa heat shock genes are differentially expressed during somatic embryogenesis. Plant Mol. Biol. 16:999-1007.

Györgyey, J., Vaubert, D., Jiménez-Zurdo, J. I., Charon, C., Troussard, L., Kondorosi, A., and Kondorosi, E. 2000. Analysis of *Medicago truncatula* nodule expressed sequence tags. Mol. Plant-Microbe Interact. 13:62-71.

Hammond-Kosack, K. E., and Jones, J. D. G. 1996. Resistance gene-dependent plant defense responses. Plant Cell 8:1773-1791.

Harrak, H., Chamberland, H., Plante, M., Bellemare, G., Lafontaine, J. G., and Tabaeizadeh, Z. 1999. A proline-, threonine-, and glycine-rich protein down-regulated by drought is localized in the cell wall of xylem elements. Plant Physiol. 121:557-564.

Heintzen, C., Melzer, S., Fischer, R., Kappeler, S., Apel, K., and Staiger, D. 1994. A light- and temperature-entrained circadian clock controls expression of transcripts encoding nuclear proteins with homology to RNA-binding proteins in meristematic tissue. Plant J. 5:799-813.

Hoffmann, B., Trinh, T. H., Leung, J., Kondorosi, A., and Kondorosi E. 1997. A new *Medicago truncatula* line with superior in vitro regeneration, transformation, and symbiotic properties isolated through cell culture selection. Mol. Plant-Microbe Interact. 10:307-315.

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.

Kaldenhoff, R. and Richter, G. 1989. Sequence of cDNA for a novel light-induced glycine-rich protein. Nucleic Acids Res. 17:2853.

Keller, B., Schmid, J., and Lamb, C. J. 1989. Vascular expression of a bean cell wall glycine-rich protein–β-glucuronidase gene fusion in transgenic tobacco. EMBO (Eur. Mol. Biol. Organ.) J. 8:1309-1314.

Kingsley, P. D. and Palis, J. 1994. GRP2 proteins contain both CCHC zinc fingers and a cold shock domain. Plant Cell 6:1522-1523.

Kohorn, B. 2001. WAKs: Cell wall associated kinases. Curr. Opin. Cell Biol. 13:529-533.

Kondorosi, E., Banfalvi, Z., and Kondorosi, A. 1984. Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: Identification of nodulation genes. Mol. Gen. Genet. 193:445-452.

Küster, H., Quandt, H. J., Broer, I., Perlick, A. M., and Pühler, A. 1995. The promoter of the *Vicia faba* L. VfENOD-GRP3 gene encoding a glycine-rich early nodulin mediates a predominant gene expression in the interzone II-III region of transgenic *Vicia hirsuta* root nodules. Plant Mol. Biol. 29:759-72.

Küster, H., Schröder, G., Frühling, M., Pich, U., Rieping, M., Schubert, I., Perlick, A. M., and Pühler, A. 1995. The nodule-specific VfENOD-GRP3 gene encoding a glycine-rich early nodulin is located on chromosome I of *Vicia faba* L. and is predominantly expressed in the interzone II-III of root nodules. Plant Mol. Biol. 28:405-421.

Le, H., Chang, S., Tanguay, R., and Gallie, D. 1997. The wheat poly(A)-binding protein functionally complements pab1 in yeast. Eur. J. Biochem. 243:350-357.

Linthorst, H. J., van Loon, L. C., Memelink, J., and Bol, J. F. 1990. Char-

acterization of cDNA clones for a virus-inducible, glycine-rich protein from petunia. Plant Mol. Biol. 15:671.
Marty, I., Monfort, A., Stiefel, V., Ludevid, D., Delseny, M., and Puigdomenech, P. 1996. Molecular characterization of the gene coding for GPRP, a class of proteins rich in glycine and proline interacting with membranes in *Arabidopsis thaliana*. Plant Mol. Biol. 30:625-636.
Memelink, J., Linthorst, H. J., Schilperoort, R. A., and Hoge, J. H. 1990. Tobacco genes encoding acidic and basic isoforms of pathogenesis-related proteins display different expression patterns. Plant Mol. Biol. 14:119-126.
Molina, A., Mena, M., Carbonero, P., and García-Olmedo, F. 1997. Differential expression of pathogen-responsive genes encoding two types of glycine-rich proteins in barley. Plant Mol. Biol. 33:803-810.
Mousavi, A., Hiratsuka, R., Takase, H., Hiratsuka, K., and Hotta, Y. 1999. A novel glycine-rich protein is associated with starch grain accumulation during anther development. Plant Cell Physiol. 40:406-416.
Orosz, L., Svab, Z., Kondorosi, A., and Sik, T. 1973. Genetic studies on Rhizobiophage 16-3. I. Genes and functions on the chromosome. Mol. Gen. Genet. 125:341-350.
Park, A. R., Somi, K. C., Yun, U. J., Jin, M. Y., Lee, S. H., Sachetto-Martins, G., and Park, O. K. 2001. Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. J. Biol. Chem. 276:26688-26693.
Pawlowski, K., Twigg, P., Dobritsa, S., Guan, C., and Mullin, B. C. 1997. A nodule-specific gene family from *Alnus glutinosa* encodes glycine- and histidine-rich proteins expressed in the early stages of actinorhizal nodule development. Mol. Plant-Microbe Interact. 10:656-664.
Potenza, C., Thomas, S. H., and Sengupta-Gopalan, C. 2001. Genes induced during early response to *Meloidogyne incognita* in roots of resistant and susceptible alfalfa cultivars. Plant Sci. 161:289-299.
Putnoky, P., Grosskopf, E., Ha, D. T., Kiss, G. B., and Kondorosi, A. 1988. *Rhizobium fix* genes mediate at least two communication steps in symbiotic nodule development. J. Cell Biol. 106:597-607.
Reddy, A. S., and Poovaiah, B. W. 1987. Accumulation of a glycine rich protein in auxin-deprived strawberry fruits. Biochem. Biophys. Res. Commun. 147:885-891.
Rohde, W., Rosch, K., Kroger, K., and Salamini, F. 1990. Nucleotide sequence of a *Hordeum vulgare* gene encoding a glycine-rich protein with homology to vertebrate cytokeratins. Plant Mol. Biol. 14:1057-1059.
Ross, J. H., and Murphy, D. J. 1996. Characterization of anther-expressed genes encoding a major class of extracellular oleosin-like proteins in the pollen coat of Brassicaceae. Plant J. 9:625-637.
Sachetto-Martins, G., Franco, L. O., and de Oliveira, D. E. 2000. Plant glycine-rich proteins: A family or just proteins with a common motif? Biochim. Biophys. Acta 1492:1-14.
Santino, C. G., Stanford, G. L., and Conner, T. W. 1997. Developmental and transgenic analysis of two tomato fruit enhanced genes. Plant Mol. Biol. 33:405-416.
Schröder, G., Frühling, M., Pühler, A., and Perlick, A. M. 1997. The temporal and spatial transcription pattern in root nodules of *Vicia faba* nodulin genes encoding glycine-rich proteins. Plant Mol. Biol. 33:113-123.
Schultze, M., and Kondorosi, A. 1998. Regulation of symbiotic root nodule development. Annu. Rev. Genet. 32:33-57.
Showalter, A. M., Zhou, J., Rumeau, D., Worst, S. G., and Varner, J. E. 1991. Tomato extensin and extensin-like cDNAs: Structure and expression in response to wounding. Plant Mol. Biol. 16:547-565.
von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683-4690.
Xu, D., Lei, M., and Wu, R. 1995. Expression of the rice Osgrp1 promoter-Gus reporter gene is specifically associated with cell elongation/expansion and differentiation. Plant Mol. Biol. 28:455-471.