The *Bradyrhizobium japonicum* nolA gene and its involvement in the genotype-specific nodulation of soybeans

(nodulation gene/gene-for-gene interactions/symbiosis/plant-bacteria interaction)

**ABSTRACT** Several soybean genotypes have been identified which specifically exclude nodulation by members of *Bradyrhizobium japonicum* serocluster 123. We have identified and sequenced a DNA region from *B. japonicum* strain USDA 110 which is involved in genotype-specific nodulation of soybeans. This 2.3-kilobase region, cloned in pMJ12, allows *B. japonicum* serocluster 123 isolates to form nodules on plants of serogroup 123-restricting genotypes. The nodules, however, were ineffective for symbiotic nitrogen fixation. The nodulation-complementing region is located approximately 590 base pairs transcriptionally downstream from nodD2. The 5′ end of pMJ12 contains a putative open reading frame (ORF) of 710 base pairs, termed nolA. Transposon Tn3-HoHo mutations only within the ORF abolished nodulation complementation. The N terminus of the predicted nolA gene product has strong similarity with the N terminus of MerR, the regulator of mercury resistance genes. Translational lacZ fusion experiments indicated that nolA was moderately induced by soybean seed extract and the isoflavone genistein. Restriction fragments that hybridize to pMJ12 were detected in genomic DNAs from both nodulation-restricted and -unrestricted strains.

In soybean production areas of the midwestern United States, members of *Bradyrhizobium japonicum* serocluster 123 are the dominant indigenous competitors for soybean nodulation (1–3). Cregan and Keyser (4) identified several soybean genotypes [including plant introductions (PIs) and cultivars] which restricted nodulation and reduced the competitiveness of *B. japonicum* strain USDA 123. These genotypes nodulate normally with strain USDA 110. Three of the genotypes, PIs 377578, 371607, and 417566, were subsequently shown to restrict nodulation by different serocluster 123 isolates (5, 6).

In *B. japonicum*, the essential common nodulation (nod) genes (nodA, nodB, nodC, and nodD1), as well as nodY, nodD2, and nodSUIJ, are closely linked on the chromosome (7–9). Host-specific nodulation (hsn) determinants have also been reported in *Rhizobium meliloti* (10), *Rhizobium leguminosarum* biоварs trifoli (11) and viceae (12), in the broad host-range *Rhizobium* sp. strain MPIK3030 (13), and in *B. japonicum* (7, 14, 15). In *Rhizobium* and *Bradyrhizobium*, nod, hsn, and other symbiosis genes are induced by flavonoid exudates from the host plant (16–18).

While hsn genes have been localized in *Rhizobium* and *Bradyrhizobium*, there have been no reports of the identification of genotype-specific nodulation (GSN) genes in bradyrhizobia. The GSN genes would refer to those bacterial sequences which allow nodulation of specific plant genotypes within a given legume species. One example of a GSN-like gene is nodX. This gene, in *R. leguminosarum* biovar *viceae* strain TOM, is not needed for nodulation of commercial pea cultivars, but it is essential for nodulation of the more primitive Afghanistan pea (19). Another GSN-like locus has also been reported in *Rhizobium fredii* strain USDA 257 (20).

In this report, we describe the identification and sequence analysis** of a gene, *nolA*, from *B. japonicum* strain USDA 110, that allows serocluster 123 isolates to nodulate soybean plants having USDA 123-restricting PI genotypes. We show that sequences in this region are induced by the isoflavone genistein and that the predicted amino acid sequence of the deduced open reading frame (ORF) contains a motif that may be similar to the helix-turn-helix motif associated with DNA-binding regulatory proteins.

**MATERIALS AND METHODS**

**Bacterial Strains, Plasmids, and Growth Conditions.** The *B. japonicum* strains used in this study were obtained from the U.S. Department of Agriculture–Agricultural Research Service (Beltsville, MD). The nodulation phenotypes of *B. japonicum* strains on soybean hosts are as previously described (5, 6, 21). Cosmids prJUT10, pr32, and pr29 were as previously described (22). The *B. japonicum* strains were maintained and mated on AG medium (23). The *Escherichia coli* strains were grown on LB medium (ref. 24, p. 440). As required, media were supplemented with tetracycline (Tc; 25–100 μg/ml), kanamycin (50 μg/ml), spectinomycin (50 μg/ml), carbenicillin (25 μg/ml), or ampicillin (30 μg/ml).

**DNA Biochemistry.** Total genomic *B. japonicum* DNA was isolated and Southern hybridizations were done as described (23). Plasmid DNA was isolated by using the alkaline rapid lysis procedure (ref. 24, pp. 90–91).

**Cloning and Identification of Nodulation and GSN Genes.** The HindIII fragments of cosmid pRK32 were subcloned in pVK102 as described (25, 26). Cosmids containing fragments corresponding to the nodD1YABC, nodSUIJ, nodZ, nodD2, or other gene regions (7–9, 25, 27) were used to transform E. coli strain S17-1 (ref. 24, pp. 254–255) (see Fig. 1). Conjugal transfer of pr29, pr32, and pMJHSIII subclones of pRK32 to *B. japonicum* strains USDA 162 and SD6-1c (serogroup 123) was done by using helper plasmid pKK2073 (28). Trans-conjugants were selected on AG medium containing Tc (60 μg/ml). A SalI partial-digestion genomic DNA library from

Abbreviations: PI, plant introduction; GSN, genotype-specific nodulation; ORF, open reading frame; Tc, tetracycline.

†To whom reprint requests should be addressed at: Soil Science Department, University of Minnesota, 246 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108.

**The sequence reported in this paper has been deposited in the GenBank data base (accession no. M38360).**
strain DE3-la was constructed in cosmid pVK102 as described (26). Cosmid clones corresponding to the GSN and nod genes were identified by hybridization to the HindIII fragments of PMJS12 and pMJS18 and mated into B. japonicum strain SD6-1c as described above.

**DNA Sequence Analysis.** The 1.9-kilobase (kb) Smal-I-Bgl II fragment (see Fig. 1) was digested with the appropriate restriction enzymes and fragments were inserted into the corresponding restriction sites of phage vectors M13mp18 and M13mp19 (29). The DNA sequence was determined by using the chain-termination method of Sanger et al. (30) and a DNA sequencer (Applied Biosystems model 373A). The UWGCC (University of Wisconsin Genetics Computer Group) software package (release 6.1) was used for computer-aided DNA and protein sequence analysis and the amino acid alignment was done by using the MULTALIN program (31). Data base searches were done by using the TFASTA program in conjunction with GenBank and EMBL databases.

**Construction of lacZ Fusions.** Transcriptional/translational lacZ fusions were made in the 1.7- and 2.3-kb HindIII fragments of pR32 (see Fig. 1) using Tn3-HoHol as described by Stachel et al. (32). The location and orientation of inserts were determined by restriction mapping. Fusions were transferred to B. japonicum strains SD6-1c and USDA 110 as described above.

**β-Galactosidase Assays.** Assays to monitor the induction of lacZ fusions by soybean seed extract and genistin were done as described (8). Strains containing lacZ fusions were induced for 14–16 hr with 2.0 μM genistin (ICN) or 10% (vol/vol) Glycine max cv. Essex seed extract (8). Duplicate cultures were assayed for each treatment and values are in β-galactosidase units (8, 16).

**Nodulation Studies.** Plant assays were done in a Monmouth fine sandy loam soil, containing less than one B. japonicum cell per g of soil (5, 6, 25). Pots were planted with four seeds each of Glycine max PI 377578 (serogroup 123 restricting (4)) and a nonrestrictive host, cv. Williams. Transconjugants (about 10 per mating) and nodulation-positive (MN1-1c, DE3-1a, and USDA 110) and -negative (USDA 162, SD6-1c, and USDA 123) controls were inoculated onto at least four replicate plants. Plants were watered with nitrogen-free nutrient solution (5) and harvested 35 days after inoculation, and nodule numbers were determined (4). Nodulation-complementing transconjugants were isolated from surface-sterilized nodules (ref. 33, p. 7–8) on AG medium containing Tc (100 μg/ml) or kanamycin (100 μg/ml). The identity of B. japonicum transconjugants and nodule isolates was verified by using strain-specific fluorescent antibodies (34) or spot-blot immunoassays (21). The USDA 162 and SD6-1c transconjugants isolated from nodules were retested for nodulation ability as described above.

**RESULTS**

**Identification of Genes Involved in the GSN of Soybeans.** To determine if DNA regions within and clustered around nodDIYABC were involved in the GSN of soybeans, we conjugated pR29 and pR32 (both derived from strain USDA 110) (Fig. 1) into the PI 377578 nodulation-restricted strain SD6-1c (serogroup 123). When the SD6-1cp(R29) and -p(R32) transconjugants were inoculated onto PI 377578, only the pR32 transconjugants nodulated. The SD6-1cp(R32) transconjugants nodulated PI 377578 to an extent equal that of the nodulation-unrestricted strains, MN1-1c and USDA 110 (Table 1) and produced about 30–45 times the number of nodules (range of 90–138 nodules per plant) on the PI genotypes that the control strains SD6-1c or SD6-1c(pVK102) produced. The transconjugants, however, were unstable, since only about 4 out of 10 plate transconjugants were able to nodulate the PI genotypes. While all of the nodule isolates grew on AG medium with Tc and reacted with antisemum specific for serogroup 123 strains, only 7 out of 10 could form nodules when retested on the PI genotypes. Moreover, the SD6-1cp(R32) transconjugants formed ineffective symbioses with PI 377578. Nodules were white, suggesting a lack of leg-hemoglobin production. While some nodules were large and had surface patterning and color typical of soybean nodules, others were small and lacked surface ribs. Since the SD6-1cp(R29) transconjugants did not nodulate PI 377578 (data not shown), our results suggested that the DNA region in the area of nonoverlap (about 8.5 kb) was involved in the genotype-specific nodulation of soybeans. To test this, HindIII fragments from pR32 were subcloned in pVK102 (Fig. 1) and individually mated to SD6-1c, and transconjugants were tested for nodulation ability. Only the genomic DNA region contained in plasmid PMJS12 (2.3 kb) allowed SD6-1c to nodulate the restrictive soybean genotype (Table 1). The SD6-1c(pMJS12) transconjugants produced about 35 times more nodules (range of 90–120 nodule/plant) than the uncomplemented parent strain, SD6-1c(pVK102), and were also ineffective for nitrogen fixation on the PI genotypes. The GSN region is located about 350 bp transcriptionally downstream from the end of nodD (1). Sequences present on PMJS12 did not hybridize to B. japonicum nodDIYABC (PMJS18), R. meliloti hsnABCD, or Nod box consensus sequence gene probes (7, 35). The other SD6-1c

**FIG. 1. Physical–genetic map of B. japonicum strain I110 nodulation loci.** The locations of the GSN gene noI, nodD1YABC, nodD2, nodSU1, and nodZ on pR32 are shown. The relationship of subclones PMJS9, PMJS12, PMJS18, PMJS22, and pMJS24 used in complementation and hybridization studies to the physical–genetic map of nodulation loci on a 17-kb portion of pR32 are indicated. Arrows indicate proposed transcriptional orientation of genes. Numbers below subclone lines are HindIII fragment sizes in kb. The locations and orientations of Tn3-lacZ fusions in the 1.7- and 2.3-kb regions of pR32 are shown by stalked arrows. The region sequenced is denoted by the broken line. Restriction enzyme sites: H, HindIII; B, BamHI; B, Bgl II; R, EcoRI; and S, Smal I.
Table 1. Nodulation of Glycine max PI 377578 by the serogroup 123 strain SD6-1c and its pR32 and pMSJ12 transconjugants

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Nodules on PI 377578</th>
<th>Nodules on cv. Williams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninoculated</td>
<td>1.0</td>
<td>7.0</td>
</tr>
<tr>
<td>SD6-1c</td>
<td>4.0</td>
<td>26.0</td>
</tr>
<tr>
<td>SD6-1c(pVK102)</td>
<td>2.8</td>
<td>9.0</td>
</tr>
<tr>
<td>SD6-1c(pR32)</td>
<td>113.5*</td>
<td>139.0*</td>
</tr>
<tr>
<td>SD6-1c(pMSJ12)</td>
<td>109.8*</td>
<td>194.0*</td>
</tr>
<tr>
<td>MN1-1c</td>
<td>116.7</td>
<td>290.0</td>
</tr>
<tr>
<td>USDA 110</td>
<td>107.75</td>
<td>215.0</td>
</tr>
</tbody>
</table>

Nodule number and dry weight are per plant. Values for transconjugants are means of 8–24 replicates.

*Significantly different from the uninoculated or negative controls (P < 0.05) by Duncan’s new multiple range test.

transconjugants (containing pMSJ9, pMSJ18, pMSJ22, or pMSJ24) failed to nodulate the restricting PI genotype (data not shown). Strains SD6-1c(pVK102), SD6-1c(pR32), and SD6-1c(pMSJ12) nodulated and fixed nitrogen normally with the nonrestrictive commercial soybean genotype Williams (Table 1). A majority (6 out of 10) of SD6-1c(pMSJ12) nodule-isolated transconjugants retained the ability to nodulate PI 377578, suggesting that pMSJ12 most likely contributed to the formation of these nodules. Plasmid pMSJ12 also allowed another serogroup 123 isolate, USDA 162, to nodulate PI 377578 (range of 126–145 nodules per plant), suggesting that its involvement in GSN is not limited to a particular serogroup strain.

To determine whether a functional nodulation complementing GSN DNA region was unique to strain USDA 110, we made a genomic DNA library from strain DE3-1a [nodulation unrestricted, serogroup 127 (5, 25)] and mated cosmid clones which hybridized to the pMSJ12 gene probe to strain SD6-1c. One of the transconjugants, SD6-1c(pFR27), was found to nodulate PI 377578 (mean of 22 nodules per plant) to the same extent as DE3-1a. Cosmid pFR27 contains sequences analogous to those extending from the USDA 110 nodD2 gene to the end of the GSN gene region (data not shown). A pFR27 subclone, pFR2501, containing only the DE3-1a GSN region, was later found to allow SD6-1c to nodulate the PI genotypes (mean of 21 nodules per plant), suggesting that other unrestricted strains contain functional GSN gene regions. The DE3-1a GSN region also failed to complement nitrogen fixation effectiveness in the SD6-1c transconjugants. In several cases, SD6-1c(pFR2501) transconjugants were also ineffective on both permissive (cv. Williams) and restrictive soybean hosts. The DE3-1a-derived GSN region was more unstable than the homologous 110 region, and cosmids containing the GSN region mated to recipients at frequencies 1/10th to 1/100th of those with DNA from other regions (10^-5 to 10^-7 per recipient), and 60–100% of transconjugants contained cosmids with deletions in the GSN gene region. The instability phenomenon also apparently prevented us from making site-directed mutations in the GSN gene region.

While strain SD6-1c is restricted for nodulation of PI 377578, it still forms a small number (Table 1) of crown nodules on this PI genotype (5). To determine if these nodules were produced by a nodulation-competent subpopulation of serogroup 123 strains, nodule isolates (16 and 19 from cv. Williams and PI 377578, respectively) of strain USDA 123 were inoculated onto PI 377578. Nodulation by these isolates was not significantly different (P < 0.05) than by the USDA 123 parent. Thus, the observed “leakiness” of the restrictive nodulation phenotype is not due to a phenotypically stable nodulation-competent subpopulation of serogroup 123 isolates.

Expression of the GSN Region. To determine if sequences in and around pMSJ12 are induced by host plant-derived signal compounds, we constructed Tn3-lacZ fusions in the 1.7- and 2.3-kb fragments of pR32 (Fig. 1) and studied their induction (in strain USDA 110) with soybean seed extract and genistein. Only fusion DG62 was induced by soybean seed extract (95 ± 5 β-galactosidase units) or genistein (89 ± 6 units). This fusion is in the 2.3-kb HindIII fragment contained in pMSJ12 and maps within the GSN ORF described below. Fusion DG62 also had a small amount of β-galactosidase activity (35 ± 7 units) in the absence of inducer, suggesting that this region is constitutively transcribed at a low level. None of the other lacZ fusions examined (DG1, DG6, or DG73) were induced by seed extract or genistein (mean of 5.6 ± 1 units). These results suggest that transcription of pMSJ12 is in the same direction as nodD2 and is independent of transcripts originating from adjacent upstream regions.

DNA Sequence Analysis. To determine the coding capacity of the GSN gene region and to analyze the DNA region downstream of nodD2, we sequenced the 1.9-kb Sma I–Bgl II fragment of pR32 (see Fig. 1). This fragment contains the 5' end of the 2.3-kb HindIII fragment cloned in pMSJ12 (beginning at nucleotide 630) and areas 3' of the end of the nodD2 transcript (27). We identified one ORF (nucleotides 959–1669) within the sequenced region on the basis of potential start codons (Fig. 2). To determine if the 710-bp ORF was involved in genotype-specific nodulation, we tested Tn3-HoHo insertion mutations within or downstream of the ORF (fusions DG62 and 131, respectively; see Fig. 1) for the production of nodules on PI 377578. Only SD6-1c transconjugants with an insertion mutation within the ORF of pR32 (DG62) failed to nodulate, suggesting that the ORF gene product is most likely involved in the extended host-range phenotype. We have named this sequence nola.

Amino Acid Sequence of the Deduced nola Gene Product. The putative translational start site at the first available ATG codon of the ORF, at position 956, results in a protein with a predicted molecular weight of 26,871. The predicted Nola protein product would have a fairly high arginine content (14.3%), which would result in a net positive charge. Since this positive charge might facilitate an interaction of the protein with the negatively charged DNA backbone, we looked for a helix-turn-helix motif, which is characteristic of many DNA-binding proteins. Using the method of Doolittle and Egan (36), we identified such a motif in amino acids 13–32 (Fig. 2), and we suggest that nola may code for a DNA-binding regulatory protein. A data base search revealed that the N terminus of Nola is similar to the N terminus of MerR. MerR has been shown to be a metalloregulatory DNA-binding protein that activates or represses transcription of mercury resistance genes (37). The alignment of Nola to the MerR proteins of Bacillus sp. strain RC607 and Staphylococcus aureus (38, 39) is shown in Fig. 3. While the homology between the two MerR proteins extends to the very end of both proteins, the homology between Nola and MerR is restricted to the first 70 amino acid residues. However, since this is the region in which the helix-turn-helix motif is located, Nola and MerR could share a mode of regulating gene expression.

Conservation of GSN Gene Region in B. japonicum Strains. Since results from previous studies indicated that pMSJ12 hybridized to genomic DNA from nodulation-restricted and unrestricted serocluster 123 strains (25), it was of interest to determine if the GSN gene region was present in other B. japonicum isolates. The pMSJ12 gene probe hybridized to genomic DNA from several serologically distinct B. japonicum strains. While pMSJ12 hybridized predominantly to 6.1-kb EcoRI fragments from B. japonicum strains in DNA
homology (40) (or fatty acid (41)) groups I and IA (USDA 4, 6, 38, 62, 110, 122, and 124), little or no hybridization was observed with DNAs from strains USDA 31, 46, 61, 76, 94, 130, or 135 (data not shown). With these strains, however, there was no apparent relationship between hybridization profiles and nodulation restriction. A GSN probe did not hybridize to genomic DNA from *R. leguminosarum* biovars *trifolii* and *viceae* strains or from *Agrobacterium tumefaciens* (not shown).

**DISCUSSION**

Using a nodulation-complementation approach, we have identified a *B. japonicum* USDA 110 gene, *nola*, which is involved in the GSN of soybeans. The *nola* gene, cloned in pMJS12, allows serogroup 123 isolates to nodulate soybean plants with restrictive genotypes. Since sequences present on this fragment allow for the nodulation of plants with a specific genotype within a legume species, we have designated this a GSN (genotype-specific nodulation) locus. This locus is conceptually different than *Bradyrhizobium* and *Rhizobium* *hsn* genes (9–13, 18), which are involved in host-range specificity at the legume genus level.

In contrast to other *Bradyrhizobium* (7, 27) and *Rhizobium* *nod* and *hsn* (35) gene loci, the *nola* gene locus did not hybridize or have sequence similarity to the 25-mer *nod* box consensus sequence (7, 27), which has been implicated to have a role in the coordinated regulation of nodulation genes (35). This suggests that *nola* and several other *B. japonicum* nodulation genes, may be regulated by *nod*-box-independent promoters (7, 27).

While *nola* allows serogroup 123 strains to nodulate the PI genotypes, complementation is not complete and transconjugants fail to fix nitrogen. Also, while most SD6-1c transconjugants fix nitrogen with cv. Williams, the nodules produced are smaller than the wild-type parent. In some extreme cases, *nola* totally interfered with the nitrogen fixation ability of SD6-1c-pFR27) transconjugants on the permissive soybean host. We are not sure why the *nola* locus interacts in this manner, but there may be physical or functional incom-

**Fig. 2.** Nucleotide sequence of the *nola* gene region. The sequence starts with the *Sma* I site at the end of *nodD2* (see Fig. 1). The *HindIII* site at the start of pMJS12 (position 630) is underlined. The deduced amino acid sequence of the *nola* gene product is given in one-letter code. The putative helix-turn-helix motif at the N terminus of the *nola* protein is boxed.

**Fig. 3.** Amino acid sequence alignment of the *nola* protein of *B. japonicum* (BjNola), MerR of *Ba. circulans* sp. strain RC607 (BjMerR) (38), and MerR of *Staphylococcus aureus* (SaMerR) (39). Further details are given in the text. Identical amino acid residues are marked by asterisks. Dots at the end of the *nola* sequence indicate that amino acids extend farther but the sequence is not homologous. The putative helix-turn-helix motif at the N terminus is underlined.
patibility between nolA and the indigenous nodulation and nitrogen fixation genes (42). This instability may also be reflected in the low transfer frequencies we obtained with nolA-containing cosmids, the presence of deletions in the nolA region of transconjugants, and our inability to construct stable mutations in the strain USDA 110 nolA locus.

Though the function of the nolA gene is unknown, the N terminus of its product shows similarity to the N terminus of the MerR proteins from Bacillus (38, 43), Staphylococcus (31), and several Gram-negative bacteria. NolA and these proteins all possess helix-turn-helix motifs which are characteristic of many DNA-binding proteins. Thus, NolA and MerR may all be members of a unique family of gene regulators.

Our results indicate that the genetic components of both partners interact to influence nodulation and competition in the soybean/Bradyrhizobium symbiosis. Serogroup-specific nodulation restriction has been reported for several B. japonicum strain/A. strain combinations and could possibly be mechanistically related to race specificity systems found in several plant pathogens. In these systems, there is a gene-for-gene relationship between bacterial and plant genotypes. Djordjevic et al. (44) proposed that such a mechanism may be operating in the R. leguminosarum strain TOM/Afghanistan pea symbiosis, where nodulation restriction is determined by a single recessive host gene (45) and a single dominant bacterial gene (19). A conceptually similar gene-for-gene system is most likely functioning to control nodulation of serogroup 123 isolates with these specific soybean genotypes. In these soybean genotypes host-controlled nodulation restriction appears to be, at least partially, conditioned by a single dominant plant gene and a dominant bacterial locus.

We thank Pamela Prigg and Fuyuhioko Nishijima for technical assistance, Ray Tully for illustrations, and Renee Kossak, Carol Vance, and Adam Kondorosi for their helpful comments and suggestions. This study was supported in part by U.S. Department of Agriculture Competitive Research Grants 86-CRRCR-1-2029 and 88-37231-4003 (to M.J.S., P.B.C., and H.H.K.) and 88-37262-3894 and 62-600-1636 (to G.S.), by U.S. Public Health Service Grants GM31494-03 and 40183-01 (to G.S.), by the Department of Agriculture, and by the Swiss Federal Institute of Technology (to H.H.), by the Deutscher Akademischer Auslandsdienst (to M.G.), and by the Fundacion Juan March-Madrid (to F.R.-Q.). This is contribution no. 18,524 of the Minnesota Agriculture Experiment Station scientific publication series.
