DNA Sequence of the Common Nodulation Genes of *Bradyrhizobium elkanii* and Their Phylogenetic Relationship to Those of Other Nodulating Bacteria

Raymond C. Dobert,[1,2] Brenda T. Breil,[2,3] and Eric W. Triplett[1,2,3]

[1]Department of Agronomy, [2]Center for the Study of Nitrogen Fixation, and [3]Graduate Program in Cell and Molecular Biology, 1575 Linden Drive, University of Wisconsin, Madison 53706 U.S.A.

Received 18 March 1994. Accepted 24 May 1994.

A 6.6-kb *Bam*HI fragment containing the common nodulation genes of *Bradyrhizobium elkanii* USDA94 was identified by southern hybridization using the common *nod* genes of *B. japonicum* as a probe. This fragment was cloned and sequenced. Analysis of the sequence showed open reading frames highly homologous to *nolA*, *nodD2*, *nodD1*, and *nodKABC* from other bradyrhizobial sources. The sequence showed higher homology to the common *nod* genes of *Bradyrhizobium* sp. (*Parasponia*) than to those from *B. japonicum*. The open reading frame identified between *nodD1* and *nodA* in the *B. elkanii* sequence was far more similar to *nodK* from *Bradyrhizobium* sp. (*Parasponia*) than to *nodY* from *B. japonicum*. The molecular phylogeny of *nodD* and *nodAB* from many sources was analyzed. The genetic distance between the *nod* genes is far greater than the distance between the 16S rRNA and *nifH* genes. The differences between the *nod* genes among the species of *Rhizobium* is as great as that between *Bradyrhizobium* and *Rhizobium*. The host range of the microsymbiont was found to be a better predictor of the similarities of the common *nod* genes than the 16S rRNA or *nifH* genes. We propose two groups of *nod* genes among the rhizobia and bradyrhizobia, based on molecular phylogenetic analysis: those which nodulate legumes of temperate origin in the tribes Vicieae or Trifolieae and those which nodulate legumes of tropical origin in the tribe Phaseoleae.

While a great deal is known regarding the genetics of the *Glycine max* symbiont *Bradyrhizobium japonicum*, little is known regarding the genetics of the newly reclassified species *B. elkanii* (Anonymous 1993; Kuykendall *et al.* 1992), formerly referred to as *B. japonicum* group II. Although both species effectively nodulate soybean, they differ dramatically in their physiology and host range. In contrast to *B. japonicum*, strains of *B. elkanii* produce the foliar-chlorosis-inducing compound rhizobitoxine (Minamisawa 1989), make significant amounts of indoleacetic acid in nodules (Kaneshiro and Nicholson 1990) and in culture (Minamasawa and Fukai 1991), and fail to express uptake hydrogenase activity in soybean nodules (Keyser *et al.* 1982; Minamisawa 1989). Exopolysaccharide profiles of the two species also differ (Minamisawa 1989). The two species have been shown to be genetically distinct as demonstrated by DNA:DNA hybridization experiments (Hollis *et al.* 1981), restriction fragment length polymorphism of *nif* and *nod* regions (Minamisawa 1990; Stanley *et al.* 1985), and 16S rRNA sequence (Young *et al.* 1991).

Soybeans nodulated by *B. elkanii* have been shown to have depressed plant growth and yield compared to plants nodulated by *B. japonicum* (Vasilas and Fuhrmann 1993). The economic impact of this yield reduction may be significant, since strains of *B. elkanii* are predominant in soils of the southern United States and are common inoculum strains in Brazil (Keyser *et al.* 1984; Rumjanek *et al.* 1993). Yield data together with the ability of *B. elkanii* to produce rhizobitoxine points to an inefficient symbiotic relationship between *B. elkanii* and soybean.

Differences in the host ranges of *B. japonicum* and *B. elkanii* have also been demonstrated. Strains of *B. elkanii* are distinct from *B. japonicum* in their ability to nodulate at low frequency "nonnodulating" *rj1* soybeans and their ineffective nodulation of *Rj4* soybeans (Devine *et al.* 1990). In contrast, most *B. japonicum* strains effectively nodulate soybeans carrying the *Rj4* allele, and all those tested are completely excluded by lines containing both *rj1* alleles (Clark 1957; Devine *et al.* 1990; Devine and Breithaupt 1980; Payne and Pueppke 1985; Pueppke and Payne 1987). Unlike strains of *B. japonicum*, strains of *B. elkanii* are capable of initiating nodules on *Arachis hypogaea* (Devine *et al.* 1983). The inefficient nodulation of soybean and the ability to nodulate additional legume hosts has been used to argue that *B. elkanii* possesses a symbiotic homology for a legume host other than soybean (Devine *et al.* 1983).

In addition to the many physiological differences between *B. japonicum* and *B. elkanii* the two species also produce a different spectrum of lipooligosaccharide nodulation signals, often referred to as Nod factors, which are required for nodulation to occur (Carlson *et al.* 1993) and important in controlling host range in other symbiotic systems (Dénarié *et al.* 1992). Strains of *B. japonicum* produce up to five Nod factors in detectable quantities, all consisting of a modified chitin pentasaccharide containing a 2-*O*-methylfucose residue

Corresponding author: Eric W. Triplett.

Present address of R. C. Dobert: Biotechnology Information Center, National Agricultural Library, U.S. Department of Agriculture, 10301 Baltimore Blvd., Beltsville, MD 20705.

MPMI Vol. 7, No. 5, 1994, pp. 564-572

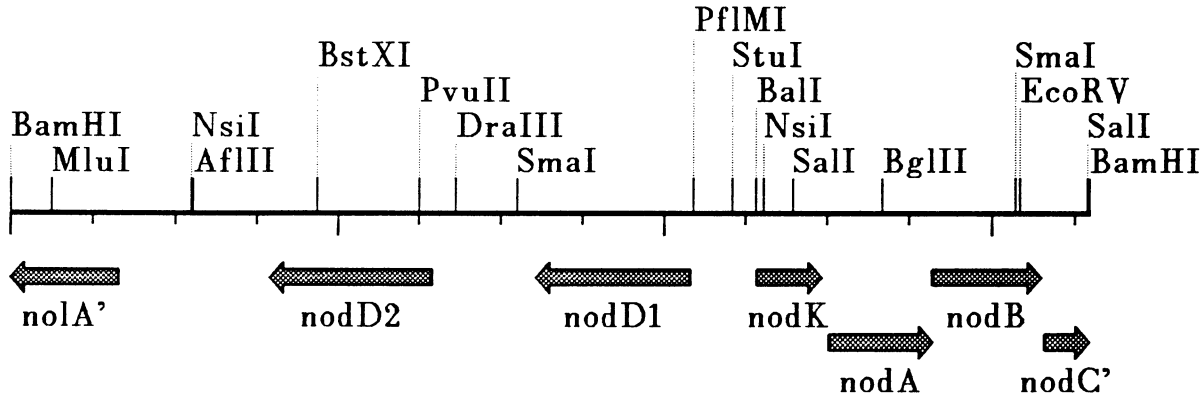

Fig. 1. Organization of common nodulation genes in *Bradyrhizobium elkanii*. Organization derived from nucleotide sequence analysis of a 6.59-kb *Bam*HI fragment from *B. elkanii* USDA94. Selected restriction sites derived from the nucleotide sequence are also shown. The *nolA* and *nodC* genes are labeled with a prime mark, as these genes are not complete in this *Bam*HI fragment. The distance between tick marks is 500 bp.

(Carlson *et al.* 1993). In contrast, *B. elkanii* strain USDA 61 produces 10 detectable Nod factors, including eight metabolites which are distinctly different from those produced by *B. japonicum* (Carlson *et al.* 1993). It is not known how these differences in Nod metabolite production influence host range in *B. elkanii*.

Much is known about the genetics of nodulation in several rhizobial and bradyrhizobial species, including the structures Nod factors produced by these systems. In rhizobia and bradyrhizobia, the common nodulation genes, *nodABC*, are transcribed divergently from *nodD1* under the control of a conserved *nod* box. In bradyrhizobia an additional gene of unknown function, designated *nodY* in *B. japonicum* (Nieuwkoop *et al.* 1987; Banfalvi *et al.* 1988) and *nodK* in *Bradyrhizobium* sp. (*Parasponia*) (Scott, 1986), has been found upstream of *nodA*. In addition, strains of *B. japonicum* possess *nolA*, which is downstream of *nodD2*. The *nolA* gene is essential for nodulation of soybean genotypes which exclude serocluster 123 strains of *B. japonicum* (Sadowsky *et al.* 1991). In many rhizobial systems nodulation genes are functionally and spatially linked.

To begin the genetic analysis of the genetics of nodulation by *B. elkanii*, a 6.6-kb region of the *B. elkanii* genome containing the common *nod* genes was isolated and sequenced. Detailed genetic information from *B. elkanii* might also help clarify its phylogenetic relationship to *B. japonicum* and other members of the genus *Bradyrhizobium*. Our results suggest that *B. elkanii* may be more closely related to *Bradyrhizobium* sp. (*Parasponia*) than to *B. japonicum*.

With our sequence data of the common *nod* genes of *B. elkanii*, the sequences of *nodD1* and *nodAB* from up to 11 species are now available. We present in this paper a molecular phylogeny of the common *nod* genes, which includes considerably more data than were available for previous *nodD* comparisons and shows a comparison of *nodAB* sequences. The molecular phylogeny of nodulation is compared to that of the 16S rRNA and *nifH* genes.

RESULTS

Organization of common *nod* genes in *B. elkanii*.

A 6.59-kb *Bam*HI fragment of DNA from *B. elkanii* USDA94 which hybridizes to the common *nod* genes of *B. japonicum* USDA110 was cloned and sequenced. Based on its high homology to analogous genes from either *B. japonicum* or *Bradyrhizobium* sp. (*Parasponia*), the *Bam*HI fragment contains most of *nolA*, all of *nodD2*, *nodD1*, *nodK*, *nodA*, and *nodB*, and a small portion of the 5′ end of *nodC* (Fig. 1). The organization of these genes is very similar to that found in *Bradyrhizobium* sp. (*Parasponia*) and *B. japonicum*. The *nolA* and *nodD2* genes in *B. elkanii* are upstream of *nodD1* and in the same direction of transcription as *nodD1*, as they are in *B. japonicum*.

```
  1 MARL...SIAIQGESKMHKTEVDLVPAACVLDELSHTDGQPSERVMPIAI 47
    ||||    |||||.||||:|||||||.:|||||||:.|| |.: . || |
  1 MARLFIAVIAIQGASKMHRTEVDLVPVGCVLDELSRIDGLPRDATAPILI 50
    .    .    .    .    .    .    .    .    .    .
 48 LGATELSQDADGVCTTRSPHWDVRAYVDHGIAGTTRASCAIVSPPTMAPL 97
    |:..|  :.|| : .. | .||||||.:|||: ..  .     ::. |::
 51 LDEAEPLHAADRLRAIGSLPWDVRACFGHGIGRKDAREMRDRLAANEAAI 100
    .    .    .    .    .    .    .    .    .    .
 98 RFSQLSGDLARIVERGCRMPTHSRLISGSANSECLT 133
    .:.   :  || .|||||||.| |||.|||||| ||
101 PIRSALARSARRAERGCRMPAHFRLITGSANSEFLT 136
```

Fig. 2. BESTFIT comparison of deduced amino acid sequences of NodK proteins from *Bradyrhizobium elkanii* (upper line) and from a *Bradyrhizobium* sp. that nodulates *Parasponia* (lower line). The two proteins are 63.2% similar and 51.0% identical.

Homology to common *nod* genes from *Bradyrhizobium* sp. (*Parasponia*).

The most striking feature of the analysis of these DNA sequence data is the high degree of homology between the common *nod* regions of *B. elkanii* USDA94 and *Bradyrhizobium* sp. (*Parasponia*) ANU289. Over 3,833 bp spanning from *nodD1* to *nodC*, the two strains share 86.7% identity at the nucleotide level. This high degree of homology is also evident at the amino acid level.

Presence of *nodK* in *B. elkanii*.

A unique feature of *nod* gene organization in *Bradyrhizobium* sp. (*Parasponia*) is the presence of *nodK*, which is 408 bp downstream of *nodD1* and 42 bp upstream of *nodA*. The *nod* region of USDA94 contains an open reading frame 407 bp downstream of *nodD1* and 42 bp upstream of *nodA* which is 51% identical and 63.2% similar at the protein level to *nodK* from *Bradyrhizobium* sp. (*Parasponia*) (Fig. 2). Because of its similarity to *nodK* from *Bradyrhizobium* sp.

(*Parasponia*) and its nearly identical location with respect to other *nod* genes, we have tentatively called this gene *nodK* in *B. elkanii*. However, the homology between the two *nodK* genes is far less than the homology observed between the *nodD1* and *nodABC* genes of *B. elkanii* and *Bradyrhizobium* sp. (*Parasponia*) (Table 1). The homology observed between the two *nodK* genes is far greater than that observed between *nodK* of *B. elkanii* and *nodY* of *B. japonicum* (Table 1). The extent of this homology is similar to that between *nodY* of *B. japonicum* and *nodK* of *Bradyrhizobium* sp. (*Parasponia*).

The *nolA* and *nodD2* sequences in *B. elkanii*.

Another interesting feature of this sequence is the very high homology of *nolA* from *B. japonicum* and *B. elkanii* (Table 1). The two NolA proteins show 93.6% similarity and 89.0% identity. This is almost as highly conserved as the NodA and NodD1 proteins from these two species (Table 1). The *nodD2* gene in *B. elkanii* has high homology to that in *B. japonicum* (Table 1). The *nodD2* genes shared 81.3% similarity and 70.4% identity between these two species. The *nolA* and *nodD2* genes have not been described in *Bradyrhizobium* sp. (*Parasponia*).

The *nod* boxes of *B. elkanii*.

The nucleotide sequences of *nodD1* and *nodA nod* boxes of *B. elkanii* USDA94 are 5′-CGTACGACCAATAGCCATCGTGGCATACCTAATTCATTTTCCAAAC-3′ and 5′-ATCCATCGTGTGGATGTGTTCTATCGAAACAATCGATTTTACCAAAC-3′, respectively.

As in *B. japonicum*, a conserved *nod* box sequence was found upstream of *nodK* in *B. elkanii*. Alignment of the *nod* box sequences immediately upstream of *nodA*, *nodK*, or *nodY* (hereafter referred to as *nodA nod* boxes) shows that these sequences are more highly conserved than the *nodD1 nod* boxes (Wang and Stacey 1991). The *nodA* box showed very high homology among the bradyrhizobia. The *nodA nod* box of *B. elkanii* is 93.6 and 100% homologous to the *nodA* boxes of *Bradyrhizobium* sp. (*Parasponia*) and *B. japonicum*, respectively. The high homology observed between the putative *nodA nod* box in *B. elkanii* and the classical 47-bp *nod* box consensus (Wang and Stacey 1991) is good evidence that this sequence serves as a promoter function in other rhizobia and bradyrhizobia. As in *Bradyrhizobium* sp. (*Parasponia*), the *nodA nod* box is 349 bp upstream of *nodK*, which is somewhat more than the 181 bp between the start codon of *nodY* and the *nodA nod* box in *B. japonicum*.

As is the case with the entire sequence of the common *nod* region, the *nodD nod* box from *B. elkanii* is more similar to *Bradyrhizobium* sp. (*Parasponia*) than it is to *B. japonicum*. The *nodD1* box is 89.4 and 72.3% homologous to the *nodD1* boxes of *Bradyrhizobium* sp. (*Parasponia*) and *B. japonicum*, respectively.

Table 1. Sequence similarity and identity between the proteins coded by the common *nod* genes of *Bradyrhizobium elkanii* and those of *B. japonicum* (Bj) and *Bradyrhizobium* sp. (*Parasponia*) (BspP)

	Similarity (%)		Identity (%)	
***nod* gene**	**Bj**	**BspP**	**Bj**	**BspP**
A	95.7	95.7	91.0	91.0
B	83.9	93.2	75.6	89.5
C	...[a]	87.8	...	81.1
D1	95.1	96.1	90.1	94.5
D2	81.3	...	70.4	...
K	49.2[b]	63.2	31.1[b]	51.0
nolA	93.6	...	89.0	...

[a] Not determined because insufficient sequence data are available from *B. japonicum* or *Bradyrhizobium* sp. (*Parasponia*) to make the comparison.

[b] Comparison between NodY of *B. japonicum* and NodK of *B. elkanii*.

Phylogeny of common *nod* genes from rhizobial, azorhizobial, and bradyrhizobial sources.

Phylogenetic trees of *nodD* (including *nodD1*, *nodD2*, and *nodD3*), *nodA*, and *nodB* were constructed (Figs. 3–5) from the sequences in the databases. The evolutionary distance of these genes among the species shown is far greater than that of the 16S rRNA and *nifH* genes among the same species (Figs. 3–6). Also, the phylogenetic relationships of the *nod* genes differ from those discerned from 16S rRNA and *nifH* sequences. For example, 16S rRNA gene sequences show that *R. fredii* is very closely related to *R. meliloti* (Fig. 6A). The *nifH* comparison also shows that *R. fredii* is more closely related to other rhizobia than it is to the bradyrhizobia (Fig. 6B). However, the *nod* genes of *R. fredii* are more homologous to the bradyrhizobia than they are to *R. meliloti* (Figs. 4B and 5). Also, the common *nod* genes of *R. leguminosarum* bv. *viceae* and *R. leguminosarum* bv. *trifolii* are as distant from each other as are *nod* genes between other species within the same genus (Figs. 3–5).

We examined the phylogeny of the available *nodD* sequences with respect to the domains described by Györgypal *et al.* (1991). Our data show that the phylogeny of the 3′ end of *nodD*, which includes approximately 650 bp, closely matches the host range of the bacteria. In contrast, the 300 bp at the 5′ end of *nodD* is intermediate between the host range of the bacteria and the phylogenetic pattern of their 16S rRNA genes (Figs. 4 and 6A).

DISCUSSION

Similarity of *nod* gene organization among bradyrhizobia.

The common nodulation genes of *B. elkanii* are organized in a manner very similar to that of *B. japonicum* and *Bradyrhizobium* sp. (*Parasponia*). The *nodD2* and *nolA* genes are found downstream of *nodD1* and are in the same orientation as *nodD1*. This is identical to the organization of these genes in *B. japonicum*. The distances between the *nolA*, *nodD2*, and *nodD1* genes in *B. elkanii* are also very similar to those found in *B. japonicum*. The organization of the *nodD1* and *nodKABC* genes in *B. elkanii* is identical to that of *Bradyrhizobium* sp. (*Parasponia*) and is very similar to that of *B. japonicum*. To date, *nolA* and *nodD2* have not been described in *Bradyrhizobium* sp. (*Parasponia*). All bradyrhizobia examined to date possess an open reading frame upstream of *nodA* which is presumably part of the common *nod* gene operon. In *B. elkanii* and *Bradyrhizobium* sp. (*Parasponia*) this gene is referred to as *nodK*.

Homology of *nod* genes in *Bradyrhizobium*.

The sequence data presented in this paper demonstrate the close phylogenetic relationship between the putative *nod* genes of *B. elkanii* and *Bradyrhizobium* sp. (*Parasponia*).

Since the 16S rRNA sequence of the *Parasponia* bradyrhizobia is not available and only a partial 16S rRNA sequence is available in *B. elkanii*, it is not known whether these two groups are closely linked phylogenetically.

The common *nod* genes of three bradyrhizobia, *B. japonicum*, *B. elkanii*, and *Bradyrhizobium* sp. (*Parasponia*), have been sequenced, and all show an open reading frame of unknown function between *nodD1* and *nodA*. In *B. japonicum*, this gene is referred to as *nodY*. The *nodK* gene, which has very little homology to *nodY*, has been found only in *Parasponia* bradyrhizobia and *B. elkanii*. However, the *nodK* sequences of these two groups also differ.

The *nodD1* genes among the bradyrhizobia are very similar, which is not surprising, because of their essential role in nodulation. Göttfert *et al.* (1992) have shown that *nodD2* is not essential for nodulation in *B. japonicum*. As mutations in *nodD2* have no effect on nodulation, it is not surprising that the *nodD2* sequence is less highly conserved than the *nodD1* and *nodABC* sequences. However, it is possible that *nodD2* plays a role in nodulation of host species not examined by Göttfert *et al.* (1992).

The discovery of *nolA* in *B. elkanii* was surprising, since Sadowsky *et al.* (1991) observed little or no hybridization homology between the *B. japonicum nolA* and total DNA from several strains of *B. elkanii*, including USDA94. Given the very high sequence homology between *nolA* of *B. elkanii* and *nolA* of *B. japonicum*, hybridization between these genes would be expected.

Homology of the *nod* boxes among rhizobia, bradyrhizobia, and azorhizobia.

In contrast to the genetic organization in *Rhizobium* species, the *nodD1* gene in *B. japonicum* is preceded by a *nod* box which functions to regulate its own expression (Banfalvi *et al.* 1988; Goethals *et al.* 1992; Wang and Stacey 1991). In *B. elkanii*, a putative *nod* box is present which shares exact homology to that found in *B. japonicum* and is located nearly the same distance from the putative *nodD1* start codon. Although not identified before, a putative *nodD nod* box is also present in *Bradyrhizobium* sp. (*Parasponia*), as shown by sequence homology. The presence of a *nodD nod* box in all three bradyrhizobial groups examined to date suggests that inducible expression of *nodD* is a common characteristic in bradyrhizobia. As the *nodA nod* box of *B. elkanii* is 100% homologous to that of *B. japonicum*, it contains the 9-bp consensus repeat four times, as described by Wang and Stacey (1991). The *nodA nod* box sequence of *B. elkanii* is also in agreement with the proposal by Goethals *et al.* (1992) that the A-T-C-N_9-G-A-T sequence is the binding site for the NodD protein.

Phylogeny of nodulation genes.

We present phylogenetic trees of *nifH* and 16S rRNA genes in addition to phylogenetic trees of four common *nod* genes so that the reader can readily see the differences in evolutionary paths of these genes among the species for which *nod* genes have been sequenced. More extensive phylogenetic trees of *nifH* and 16S rRNA genes, which show relationships in many bacterial species, are available in the literature (Yanagi and Yamasato 1993; Young 1992, 1993).

Earlier analyses of the molecular phylogeny of *nod* genes included *nodD* comparisons (Györgypal *et al.* 1991; Schlaman *et al.* 1992; Young and Johnston 1989). The analyses by Györgypal *et al.* (1991) and Young and Johnston (1989) included only *nodD* sequences from *Rhizobium* species. The analysis by Schlaman *et al.* (1992) included *nodD1*, *nodD2*,

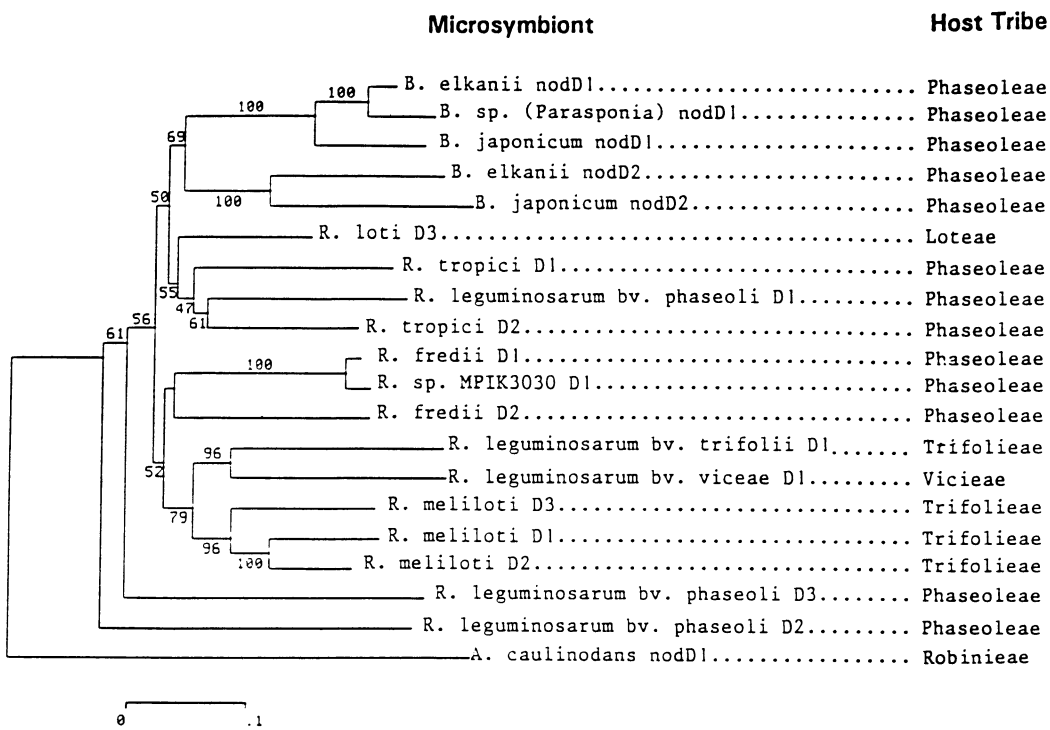

Fig. 3. Phylogenetic trees of *nodD1*, *nodD2*, and *nodD3* from species of *Azorhizobium*, *Bradyrhizobium*, and *Rhizobium*. Each branch is labeled with the species name of a bacterium and the name of the legume tribe nodulated by that species. Confidence levels (%) above each node were generated from 500 bootstrap trees. The distance scale represents the number of changes per sequence position with a median rate of change.

nodD3, and *syrM* from many sources, but this was done in the context of describing relationships between many transcriptional regulator proteins. This paper describes a molecuslar phylogenetic comparison of *nodD* and the common *nod* genes *nodAB* from azorhizobial, rhizobial, and bradyrhizobial sources with respect to the phylogeny of the host plant.

As described previously by Young and Johnston (1989) for *nodD*, the common *nod* genes mutated at a faster rate than the 16S rRNA gene. This may be reflected by the essential function of a chromosomal gene versus a set of genes which are

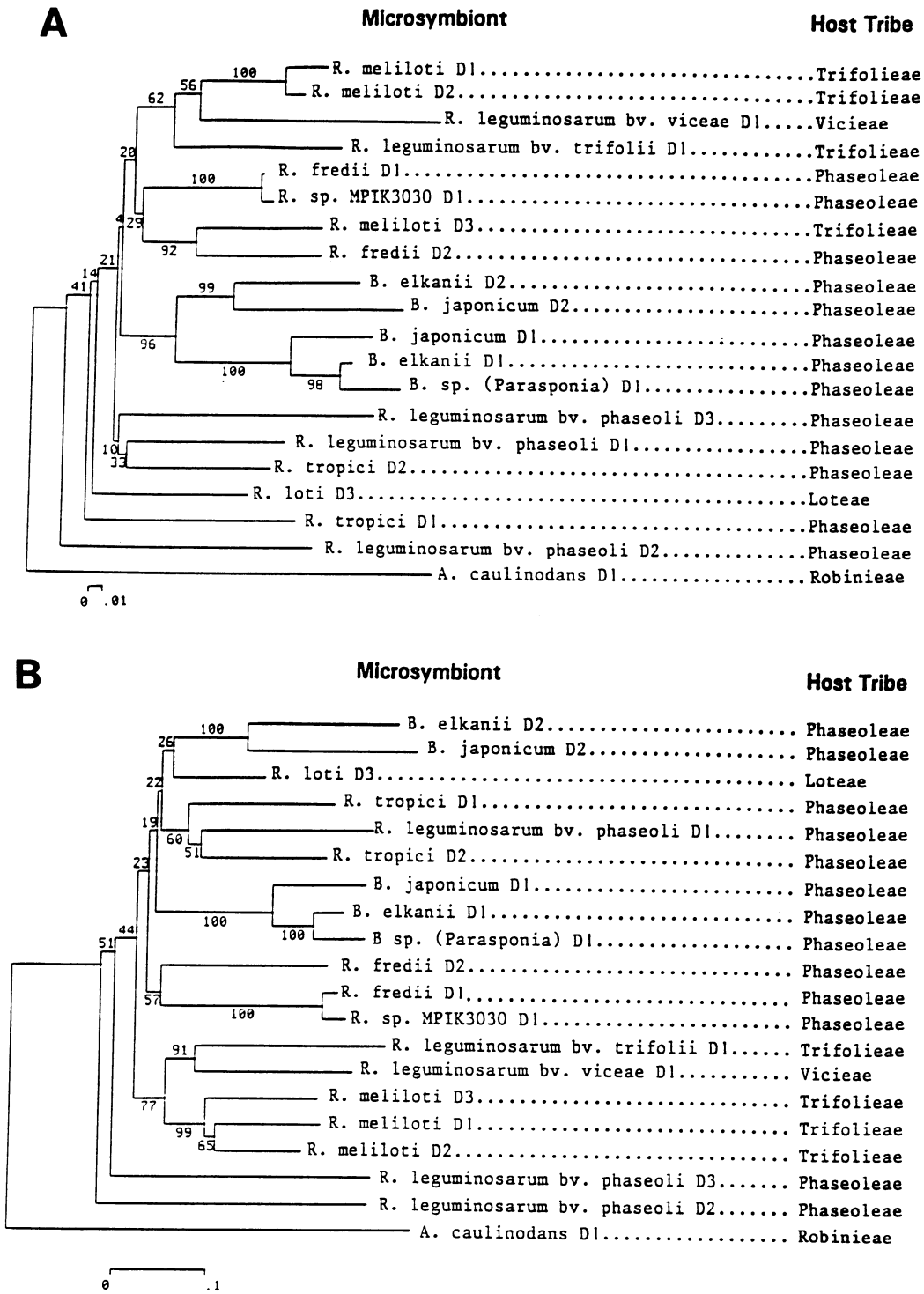

Fig. 4. Phylogenetic trees of the 5′ end (**A**) and the 3′ end (**B**) of *nodD1*, *nodD2*, and *nodD3* from species of *Azorhizobium*, *Bradyrhizobium*, and *Rhizobium*. Each branch is labeled with the species name of a bacterium and the name of the legume tribe nodulated by that species. Confidence levels (%) above each node were generated from 500 bootstrap trees. The distance scale represents the number of changes per sequence position with a median rate of change.

often plasmid-borne, subject to lateral transfer, and not essential for saprophytic growth.

Also, as previously described by Young (1992, 1993), the *nif* and *nod* genes have also evolved differently. The phylogeny of *nifH* more closely resembles the phylogeny of the 16S rRNA sequences than that of the *nod* genes (Figs. 3–6). Also, the *nifH* sequences are much more highly conserved than the *nod* gene sequences. Conservation of *nifH* is similar to that of the 16S rRNA genes. Presumably, the function of *nifH* is more sensitive to mutation than the *nod* genes are. Rhizobial *nifH* is distinct from bradyrhizobial *nifH*, whereas the *nod* gene phylogeny does not closely follow the taxonomic boundaries of the microsymbionts.

A complicating factor in the evolution of the *nif* and *nod* genes is the possible lateral transfer of Sym plasmids between similar chromosomal backgrounds during evolutionary history. Young (1992) has also suggested that gene duplication may be involved in the phylogeny of *nifH*. The differing paths of phylogeny of *nifH* and the common *nod* genes cannot be explained simply by lateral transfer of Sym plasmids, since both the *nif* and the *nod* regulons are often found nearby on Sym plasmids in *Rhizobium*.

Evolution of *nod* genes correlated with taxonomic divisions within Leguminosae.

The differences among the *nod* genes for the species of *Rhizobium* are as great as those between *Bradyrhizobium* and *Rhizobium*, despite the fact that 16S rRNA sequences clearly justify the distinction between *Bradyrhizobium* and *Rhizobium*. Striking examples of this are found in the relationships between *R. meliloti*, *R. fredii*, and *R. leguminosarum*. For example, the common *nod* genes of *R. leguminosarum* bvs. *viceae* and *trifolii* are far more distant than would be expected for homologous genes within the same species. Also, as shown by the 16S rRNA gene sequences (Fig. 6), *R. meliloti* and *R. fredii* are phylogenetically very similar. However, the nodulation genes of these two species are not very similar (Figs. 3–5). The *R. fredii nod* genes are more similar to those of the bradyrhizobia than to those of *R. meliloti*. The *R. meliloti nod* genes are more similar to those of *R. leguminosarum* than they are to *R. fredii*. These examples suggest that the phylogenetic relationships of the host legumes may be better predictors of the similarities of the *nod* genes than the 16S rRNA gene sequence, which is the most commonly used molecular taxonomic character for bacteria (Olsen *et al.* 1994). These data demonstrate that the evolutionary pressures on *nod* genes are very different from those imposed on *nifH* or on essential chromosomal genes such as the 16S rRNA gene.

Györgypal *et al.* (1991) identified two domains in NodD in *Rhizobium*. One of these domains is highly conserved and includes the 100 amino acids beginning at the amino terminus. The other domain includes the remaining 200 amino acids, is not as highly conserved, and is probably involved in flavonoid recognition. Györgypal *et al.* (1991) found that the phylogeny of the second domain matched the pattern of flavonoid specificity of the NodD proteins. Our data also show that the 3′ end of *nodD* which matches the pattern of flavonoid specificity also correlates well with the host range of the

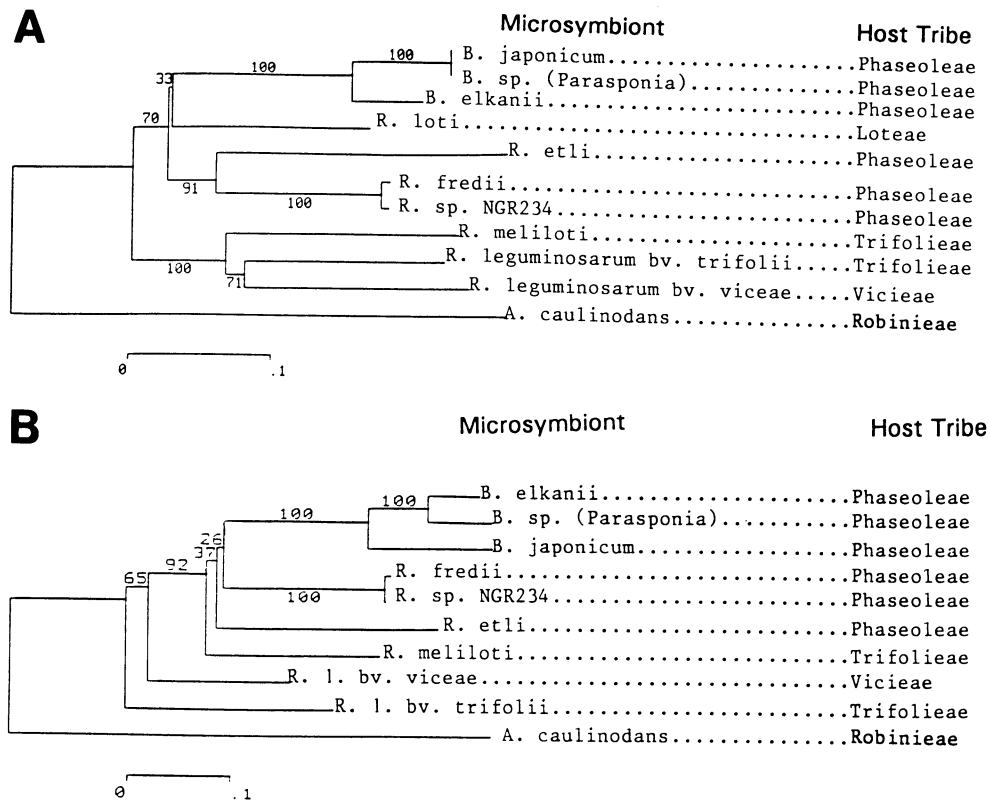

Fig. 5. Phylogenetic trees of *nodA* (**A**) and *nodB* (**B**) from species of *Azorhizobium*, *Bradyrhizobium*, and *Rhizobium*. Each branch is labeled with the species name of a bacterium and the name of the legume tribe nodulated by that species. Confidence levels (%) above each node were generated from 500 bootstrap trees. The distance scale represents the number of changes per sequence position with a median rate of change.

microsymbiont (Györgypal *et al.* 1991) (Fig. 4B). The 5′ end of *nodD* more closely resembles the phylogeny of the 16S rRNA gene than the host range of the bacteria.

The phylogenetic trees of the *nodA* and the 3′ end of *nodD* (Figs. 4B and 5A) show that there are two major branches among the rhizobia and bradyrhizobia. One of these branches includes *B. elkanii*, *Bradyrhizobium* sp. (*Parasponia*), *B. japonicum*, *R. fredii*, *Rhizobium* sp. NGR234, *R. etli*, and *R. loti*. With one exception, all of these species nodulate legumes in the tribe Phaseoleae. The exception is *R. loti*, which nodulates legumes within the tribe Loteae. This exception can be explained by lateral transfer of the *nod* genes from an ancestor in common with *R. etli*. The second branch includes *R. meliloti* and *R. leguminosarum* bvs. *trifolii* and *viceae*. These species nodulate legumes in the tribes Trifolieae and Vicieae. Thus the *nod* genes among the bradyrhizobia and rhizobia have evolved into two major groups. One group nodulates legumes of tropical origin (Phaseoleae). Legumes of the tribe Loteae can be found in temperate or tropical regions (Allen and Allen 1981). The other group nodulates legumes of temperate origin (Trifolieae and Vicieae). Representatives of the genus *Rhizobium* are found in both groups. We refer to these two *nod* gene groups as the Phaseol and Trifol nodulation groups, respectively.

Although the phylogeny of *nodB* does not show these two branches, it clearly shows that the *nodB* genes of *Bradyrhizobium*, *R. fredii*, and *R. etli* are much more closely related than would be expected from their 16S rRNA sequences (Fig. 5B). The *nodB* genes of *R. fredii* and *R. etli* are quite distant from those of *R. leguminosarum*. Thus, the *nodB* sequence phylogeny is well correlated with host range.

We propose that the *nod* genes of *Rhizobium* and *Bradyrhizobium* have evolved as two major groups which are well correlated with the taxonomy of the host legumes but uncorrelated with the taxonomy of the bacteria. The Phaseol and Trifol nodulation groups described above give strong support to the idea of symbiotaxonomy as proposed by Allen and Allen (1981). Although the cross-inoculation groups as originally defined by Fred *et al.* (1932) are no longer valid for the taxonomic positioning of rhizobia and bradyrhizobia, the nodulation phenotype has evolved with the host plants in a manner reminiscent of the cross-inoculation group concept and independent of the phylogeny of the microsymbionts. For example, *R. fredii* and *B. japonicum* are distant relatives within the α-proteobacteria, both of which nodulate *Glycine max*. However, the *nod* genes of *R. fredii* and *B. japonicum* are of the same lineage, which is quite distinct from that of *R. meliloti*.

Since the phylogenetic relationship of the *nod* genes seems to reflect the evolutionary pattern of hosts nodulated by the bacterium, the evolution of the host and/or changes in the host's environment may have been involved in applying the selection pressure that shaped the bacterial *nod* genes. More sequence data of the common *nod* genes from other root-nodulating bacteria are necessary to test the concept of distinct nodulation groups based on the taxonomy of the legume hosts.

MATERIALS AND METHODS

Bacteria and culture conditions.

B. elkanii strain USDA94 was maintained on YM (Vincent 1970) with tetracycline (12.5 μg/ml) at 28° C. *Escherichia*

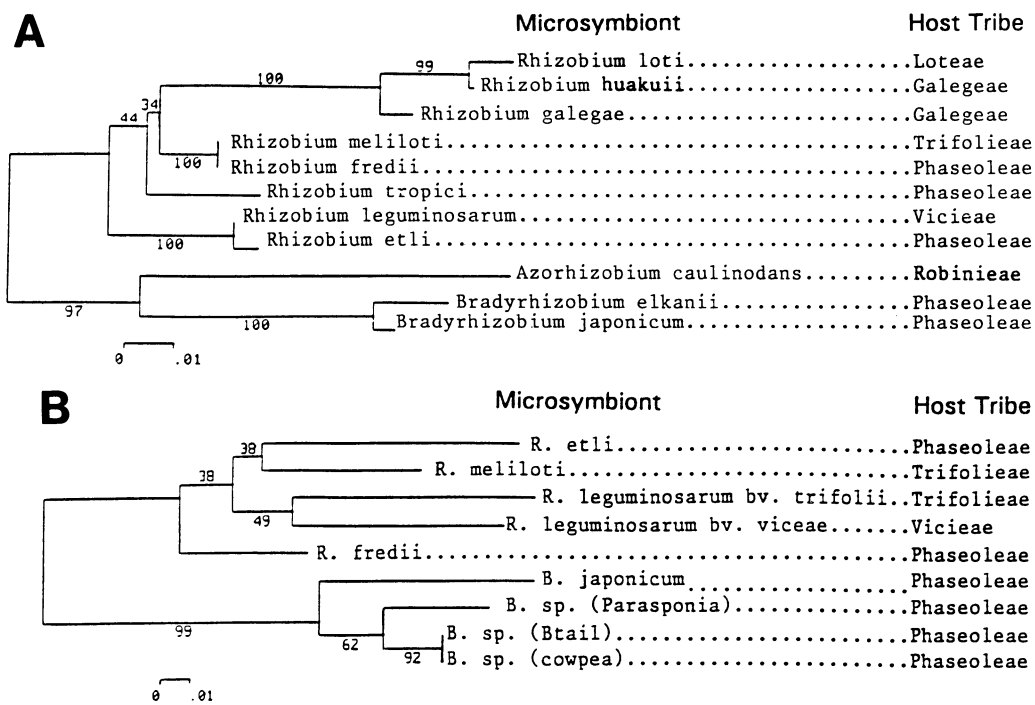

Fig. 6. Phylogenetic tree of 16S rRNA (**A**) and *nifH* (**B**) genes from species of *Azorhizobium*, *Bradyrhizobium*, and *Rhizobium*. Each branch is labeled with the species name of a bacterium and the name of the legume tribe nodulated by that species. Confidence levels (%) above each node are shown which were generated from 500 bootstrap trees. The distance scale represents the number of changes per sequence position with a median rate of change. All 16S rRNA sequences are full-length except those of *R. etli* and *B. elkanii*. Among the *nifH* sequences, those for *Bradyrhizobium* sp. (Btail), *Bradyrhizobium* sp. (cowpea), *R. fredii*, *R. leguminosarum* bv. *trifolii*, and *R. leguminosarum* bv. *viceae* are incomplete.

coli strains DH5α (Sambrook *et al.* 1989) and XL1-Blue (Stratagene, La Jolla, CA) were cultured on Luria-Bertani medium (Sambrook *et al.* 1989) at 37° C. Where appropriate, antibiotics were added at the following concentrations: ampicillin and kanamycin at 50 μg/ml, naladixic acid at 20 μg/ml, and tetracycline at 12.5 μg/ml.

DNA isolation and manipulations.

Except where otherwise indicated, all DNA isolations and manipulations were performed as described by Sambrook *et al.* 1989. To clone the common *nod* genes from USDA94, the DNA from this strain was totally digested with *Bam*HI and separated on an agarose gel. DNA fragments of roughly 6 kb were excised and released by treatement with β-agarase (New England Biolabs, Beverly, MA) according to the manufacturer's instructions. This DNA was then cloned into the *Bam*HI site of the broad-host-range vector pDSK519 (Keen *et al.* 1988). Colony hybridizations and Southerns were performed with the ECL kit, following the manufacturer's instructions (Amersham Corp., Arlington Heights, IL). One clone, referred to as pD9445, was found that hybridized to the 3.9-kb *Hin*dIII fragment of pMJS18 which contains the common *nod* genes of *B. japonicum* (Sadowsky *et al.* 1990).

DNA sequencing and sequence analysis.

The 6.6-kb *Bam*HI fragment of pD9445 containing the putative *nod* genes was cloned in both orientations into the *Bam*HI site of pBluescript II KS$^+$ by a cohesive-end ligation to yield pB94N3 and pB94N13. These were transformed into XL1-Blue competent cells for single-stranded DNA rescue. VCS-M13 was used as the helper phage and was propagated as described by the manufaturer (Stratagene). The single-strand inserts of pB94N3 and pB94N13 were sequenced by primer walking with the use of the Sequenase 2.0 dideoxy chain termination kit (United States Biochemical Corp., Cleveland, OH). All ambiguities were resolved using dIs in place of dGs according to the manufacturer's instructions. Computer sequence analysis was performed with the programs of the Genetics Computer Group, Inc. (Madison, WI), maintaining their defaults (Devereux *et al.* 1984). The 6,590-base sequence has been assigned GenBank accession number U04609. The nucleotide sequence data of the *B. elkanii nod* genes are not shown here.

Construction of phylogenetic trees.

The phylogenetic trees were generated by the PILEUP and MEGA programs (Devereux *et al.* 1984; Kumar *et al.* 1993). Pairwise distance matrices and phylogenetic trees were calculated using the Jukes-Cantor and Neighbor-Joining algorithms, respectively. Sequences from the following GenBank accession numbers were used in the construction of the phylogenetic trees for 1) the *nod* genes: M18972, M18971, M73699, X04473, X01649, L06241, M58625, M58626, L01272, L01273, X03721, J03671, X01650, X54215, X54214, M60872, L18897, M81825, X03720, X73362, and U04609; 2) *nifH:* M10587, J01781, X02478, M36435, L16503, K01620, K02676, and M15401; and 3) the 16S rRNA gene: D12791, D12797, D12793, X67222, D12792, D12798, D12782, L20762, X67221, A65000, X66024. The *nodAB* sequence from *B. japonicum* USDA123 was kindly provided by E. Applebaum. Accession number L16503 is an unidentified and unpublished sequence from *R. fredii* which has very high homology to other *nifH* sequences. Based on this homology, we refer to L16503 as *nifH* from *R. fredii.*

ACKNOWLEDGMENTS

We thank J. P. W. Young for his helpful suggestions in the generation of the phylogenetic trees presented in this manuscript. We also thank M. J. Sadowsky, of the University of Minnesota, for providing pMJS18. Support for this work was provided to EWT by the University of Wisconsin–Madison, College of Agricultural and Life Sciences, through Hatch Project no. D479 and by USDA competitive grant 9103443. BTB was supported by a 1967 Natural Sciences and Engineering Research Council of Canada Scholarhip.

LITERATURE CITED

Allen, O. N., and Allen, E. K. 1981. The Leguminosae. University of Wisconsin Press, Madison.

Anonymous. 1993. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. Int. J. Syst. Bacteriol. 43:398-399.

Banfalvi, Z., Nieuwkoop, A., Schell, M., Best, L., and Stacey, G. 1988. Regulation of *nod* gene expression in *Bradyrhizobium japonicum.* Mol. Gen. Genet. 214:420-424.

Carlson, R. W., Sanjuan, J., Bhat, U. R., Glushka, J., Spaink, H. P., Wijfjes, A. H. M., van Brussel, A. A. N., Stokkermans, T. J. W., Peters, N. K., and Stacey, G. 1993. The structures and biological activities of the lipo-oligosaccharide nodulation signals produced by type I and II strains of *Bradyrhizobium japonicum.* J. Biol. Chem. 268:18372-18381.

Clark, F. E. 1957. Nodulation responses of two near-isogenic lines of soybean. Can. J. Microbiol. 3:113-123.

Dénarié, J., Debellé, F., and Rosenberg, C. 1992. Signaling and host range variation in nodulation. Annu. Rev. Microbiol. 46:497-531.

Devereux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-395.

Devine, T. E., and Breithaupt, B. H. 1980. Phenotypic thermal stability of rhizobitoxine-induced chlorosis and the nodulation controlling gene, rj1. Crop Sci. 20:394-396.

Devine, T. E., Kuykendall, L. D., and Breithaupt, B. H. 1983. Nodule-like structures induced on peanut by chlorosis producing strains of *Rhizobium* classified as *R. japonicum.* Crop Sci. 23:394-397.

Devine, T. E., Kuykendall, L. D., and O'Neill, J. J. 1990. The Rj4 allele in soybean represses nodulation by chlorosis-inducing bradyrhizobia classified as DNA homology group II by antibiotic resistance profiles. Theor. Appl. Genet. 80:33-37.

Fred, E. B., Baldwin, I. L., and McCoy, E. 1932. Root Nodule Bacteria and Leguminous Plants. University of Wisconsin Press, Madison.

Goethals, K., Van Montagu, M., and Holsters, M. 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. Proc. Natl. Acad. Sci. (USA) 89:1646-1650.

Göttfert, M., Holzhäuser, D., Bäni, D., Hennecke, H. 1992. Structural and functional analysis of two different *nodD* genes in *Bradyrhizobium japonicum* USDA 110. Mol. Plant-Microbe Interact. 5:257-265.

Györgypal, Z., Kiss, G. B., and Kondorosi, A. 1991. Transduction of plant signal molecules by the *Rhizobium* NodD proteins. Bioessays 13: 575-581.

Hollis, A. B., Kloos, W. E., and Elkan, G. H. 1981. DNA:DNA hybridization studies of *Rhizobium japonicum* and related *Rhizobiaceae.* J. Gen. Microbiol. 123:215-222.

Kaneshiro, T., and Nicholson, J. J. 1990. *Bradyrhizobium japonicum* subspecies (USDA 110 and 26) characterized by fixed-nitrogen uptake in symbiotic indoleacetic acid production. Curr. Microbiol. 20:381-385.

Keen, N. T., Tamaki, S., Kobayashi, D., and Trollinger, D. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. Gene 70:191-197.

Keyser, H. H., Berkum, P. van, and Weber, D. F. 1982. A comparative study of the physiology of symbioses formed by *Rhizobium japonicum* with *Glycine max, Vigna unguiculata,* and *Macroptilium atropur-*

purem. Plant Physiol. 70:1626-1630.
Keyser, H. H., Weber, D. F., and Uratsu, S. L. 1984. *Rhizobium japonicum* serogroup and hydrogenase phenotype distribution in twelve states. Appl. Environ. Microbiol. 47:613-615.
Kumar, S., Tamura, K., and Nei, M. 1993. MEGA: Molecular Evolutionary Genetics Analysis. Version 1.0. Pennsylvania State University, University Park.
Kuykendall, L. D., Saxena, B., Devine, T. E., and Udell, S. E. 1992. Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. Can. J. Microbiol. 38:501-505.
Minamisawa, K. 1989. Comparison of extracellular polysaccharide composition, rhizobitoxine production, and hydrogenase phenotype among various strains of *Bradyrhizobium japonicum*. Plant Cell. Physiol. 30: 877-884.
Minamisawa, K. 1990. Division of rhizobitoxine-producing and hydrogen-uptake positive strains of *Bradyrhizobium japonicum* by *nifDKE* sequence divergence. Plant Cell Physiol. 31:81-89.
Minamisawa, K., and Fukai, K. 1991. Production of indole-3-acetic acid by *Bradyrhizobium japonicum:* A correlation with genotype grouping and rhizobitoxine production. Plant Cell Physiol. 32:1-9.
Nieuwkoop, A. J., Banfalvi, Z., Deshmane, N., Gerhold, D., Schell, M. G., Sirotkin, K. M., and Stacey, G. 1987. A locus encoding host range is linked to the common nodulation genes in *Bradyrhizobium japonicum*. J. Bacteriol. 169:2631-2638.
Olsen, G. J., Woese, C. R., and R. Overbeek. 1994. The winds of (evolutionary) change: Breathing life into microbiology. J. Bacteriol. 176:1-6.
Payne, J. H., and Pueppke, S. G. 1985. *Rj1* and *rj1* soybean isolines: Adsorption of rhizobia to roots and distribution of primary root nodules. Symbiosis 1:163-176.
Pueppke, S. G., and Payne, J. H. 1987. Responses of *Rj1* and *rj1* soybean isolines to inoculation with *Bradyrhizobium japonicum*. Plant Physiol. 84:1291-1295.
Rumjanek, N., Dobert, R. C., Berkum, P. van, and Triplett, E. W. 1993. Common soybean inoculant strains in Brazil are members of *Bradyrhizobium elkanii*. Appl. Environ. Microbiol. 59:4371-4373.
Sadowsky, M. J., Cregan, P. B., Gottgert, M., Sharma, A., Gerhold, D., Rodriguez-Quinones, F., Keyser, H. H., Hennecke, H., and Stacey, G. 1991. The *Bradyrhizobium japonicum nolA* gene and its involvement in the genotype-specific nodulation of soybeans. Proc. Natl. Acad. Sci. (USA) 88:637-641.
Sadowsky, M. J., Cregan, P. B., and Keyser, H. H. 1990. DNA hybridization probe for use in determining restricted nodulation among *Bradyrhizobium japonicum* serocluster 123 field isolates. Appl. Environ. Microbiol. 56:1768-1774.
Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schlaman, H. R. M., Okker, R. J. H., and Lugtenberg, B. J. J. 1992. Regulation of nodulation gene expression by NodD in rhizobia. J. Bacteriol. 174:5177-5182.
Scott, K. F. 1986. Conserved nodulation genes from the non-legume symbiont *Bradyrhizobium* sp. (*Parasponia*). Nucleic Acids Res. 14: 2905-2919.
Stanley, J., Brown, G. G., and Verma, D. P. S. 1985. Slow-growing *Rhizobium japonicum* comprises two highly divergent symbiotic types. J. Bacteriol. 163:148-154.
Vasilas, B. L., and Fuhrmann, J. J. 1993. Field response of soybean to nodulation by a rhizobitoxine-producing strain of *Bradyrhizobium*. Agron. J. 85:302-305.
Vincent, J. M. 1970. A Manual for the Practical Study of Root-Nodule Bacteria. International Biological Programme Handbook 15. Blackwell Scientific Publications, Oxford.
Wang, S.-P., and Stacey, G. 1991. Studies of the *Bradyrhizobium japonicum* $nodD_1$ promoter: A repeated structure for the *nod* box. J. Bacteriol. 173:3356-3365.
Yanagi, M. and Yamasato, K. 1993. Phylogenetic analysis of the family Rhizobiaceae and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer. FEMS Microbiol. Lett. 107:115-120.
Young, J. P. W. 1992. Phylogenetic classification of nitrogen-fixing organisms. Pages 43-86 in: Biological Nitrogen Fixation. G. Stacey, R. H. Burris, and H. J. Evans, eds. Chapman and Hall, New York.
Young, J. P. W. 1993. Molecular phylogeny of rhizobia and their relatives. Pages 587-592 in: New Horizons in Nitrogen Fixation. R. Palacios, J. Mora, and W. E. Newton, eds. Kluwer Academic, Dordrecht, Netherlands.
Young, J. P. W., Downer, H. L., and Eardly, B. D. 1991. Phylogeny of the phototrophic *Rhizobium* strain BTAil by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. J. Bacteriol. 173: 2271-2277.
Young, J. P. W., and Johnston, A. W. B. 1989. The evolution of specificity in the legume-rhizobium symbiosis. Trends Evol. Ecol. 4:341-348.