# Comparison of the Evolutionary Dynamics of Symbiotic and Housekeeping Loci: A Case for the Genetic Coherence of Rhizobial Lineages

*J. J. Wernegreen*[1] *and M. A. Riley*

Department of Ecology and Evolutionary Biology, Yale University

In prokaryotes, lateral gene transfer across chromosomal lineages may be mediated by plasmids, phages, transposable elements, and other accessory DNA elements. However, the importance of such transfer and the evolutionary forces that may restrict gene exchange remain largely unexplored in native settings. In this study, tests of phylogenetic congruence are employed to explore the range of horizontal transfer of symbiotic (*sym*) loci among distinct chromosomal lineages of native rhizobia, the nitrogen-fixing symbiont of legumes. Rhizobial strains isolated from nodules of several host plant genera were sequenced at three loci: symbiotic nodulation genes (*nodB* and *nodC*), the chromosomal housekeeping locus glutamine synthetase II (*GSII*), and a portion of the *16S rRNA* gene. Molecular phylogenetic analysis shows that each locus generally subdivides strains into the same major groups, which correspond to the genera *Rhizobium, Sinorhizobium,* and *Mesorhizobium.* This broad phylogenetic congruence indicates a lack of lateral transfer across major chromosomal subdivisions, and it contrasts with previous studies of agricultural populations showing broad transfer of *sym* loci across divergent chromosomal lineages. A general correspondence of the three rhizobial genera with major legume groups suggests that host plant associations may be important in the differentiation of rhizobial *nod* and chromosomal loci and may restrict lateral transfer among strains. The second major result is a significant incongruence of *nod* and *GSII* phylogenies within rhizobial subdivisions, which strongly suggests horizontal transfer of *nod* genes among congenerics. This combined evidence for lateral gene transfer within, but not between, genetic subdivisions supports the view that rhizobial genera are ''reproductively isolated'' and diverge independently. Differences across rhizobial genera in the specificity of host associations imply that the evolutionary dynamics of the symbiosis vary considerably across lineages in native settings.

## Introduction

Factors that restrict gene flow have been studied extensively for sexual eukaryotes, and are known to include ecological and spatial isolation among populations and reproductive isolation between distinct species lineages (Mayr 1963; Templeton 1989; Sharman, Close, and Maynes 1990; Gittenberger 1991). However, the evolutionary forces structuring the genetic divergence of bacterial populations and species remain unclear. Due to differences in the nature of gene transfer (Cohan 1994*a,* 1994*b,* 1995) and the potential for periodic selective sweeps in bacterial populations (Levin 1981; Dykhuizen and Hartl 1983; Hartl 1992), it is difficult to draw direct parallels between the evolutionary dynamics of sexual eukaryotes and those of prokaryotes. Lateral gene transfer among bacterial lineages may be mediated by accessory DNA elements, such as plasmids and phages (Campbell 1981; Datta 1985; Levin 1986; Eberhard 1990; Arber 1991; Simonsen 1991; Hartl 1992; Young and Levin 1992; Salyers and Shoemaker 1994), which may transfer across distinct chromosomal lineages and also shuttle chromosomal loci (Broda 1979; Holloway 1979). Evidence for recombination within bacterial species includes the incongruence of phylogenies of different loci (Normand and Bousquet 1989; Dykhuizen and Green 1993; Dolbert, Brell, and Triplett 1994; Ueda et al. 1995; Young and Haukka 1996), the patchy composition of individual genomes (Milkman and Stoltzfus 1988; Lawrence, Hartl, and Ochman 1991; Milkman 1996, 1997), and linkage equilibrium inferred from multilocus enzyme electrophoresis (MLEE) analyses (Caugant et al. 1987; Istock et al. 1992; Lenski 1993; Maynard Smith 1993; Maynard Smith et al. 1993; O'Rourke and Stevens 1993; Bottomley, Cheng, and Strain 1994; Wise, Shimkets, and MacArthur 1995).

Despite this potential for frequent and broad gene exchange in bacteria, prokaryotes often form distinct phenotypic and genetic clusters. The agreement of two or more loci in inferred species relationships (Olsen and Woese 1993; Maidak et al. 1997; Eisen 1995) supports the view that bacterial species constitute coherent genetic lineages across which broad lateral transfer is relatively rare. Several MLEE analyses highlight factors that may restrict recombination among bacterial populations by detecting groups of strains which undergo recombination and identifying the ecological, spatial, or genetic parameters that best correlate with such groups. For example, linkage equilibrium within but not between genetic clusters suggests genetic constraints to recombination (Maynard Smith 1993), such as limited host ranges of plasmids (Young and Wexler 1988) or sensitivity of homologous recombination to sequence divergence (Zawadski, Roberts, and Cohan 1995). Apparent recombination within but not between bacterial groups occupying distinct habitats suggests that local populations lack opportunities for recombination (Souza 1992; Maynard Smith et al. 1993; Souza et al. 1994). Such populations may diverge genetically if each experiences independent effects of drift or selection (Maynard Smith and Haigh 1974; Kaplan, Hudson, and Langley 1989; Cohan 1995). While previous studies point to

[1] Present address: Department of Ecology and Evolutionary Biology, University of Arizona.

Key words: plasmid, evolution, horizontal transfer, rhizobia, phylogenetic congruence, nodulation.

Address for correspondence and reprints: Jennifer Wernegreen, Department of Ecology and Evolutionary Biology, University of Arizona, Biological Sciences West, Room 310, Tucson, Arizona 85721. E-mail: werjen@u.arizona.edu.

*Mol. Biol. Evol.* 16(1):98–113. 1999
 ISSN: 0737-4038

factors that may constrain gene flow in bacteria, further studies of native strains are necessary to explore the relative importance of genetic, ecological, and spatial barriers to recombination.

The mapping of ecological factors onto genetic data has been limited for bacteria due to the inability to identify and sample from distinct ecological niches of microbes. Bacterial symbionts associated with specific hosts provide model systems in which one may identify and sample from potentially distinct niches: different host taxa. The nitrogen-fixing symbionts of legumes, the Rhizobiaceae (or rhizobia), are an ideal group in which to explore the importance of spatial, ecological, and genetic barriers to lateral gene transfer across bacterial lineages. The family consists of three main branches of symbionts within the alpha subdivision of the proteobacteria, including the slow-growing genus *Bradyrhizobium,* the stem-nodulating *Azorhizobium,* and the ''fast-growing'' rhizobia, which include *Rhizobium, Sinorhizobium,* and *Mesorhizobium* (but see Terefeword et al. 1998 for evidence that *Rhizobium galegae* constitutes a distinct genus). Polyphasic taxonomy and sequence variation at ribosomal RNA genes clearly distinguishes major groups and sets *Mesorhizobium* apart from the other fast-growing genera (Young and Haukka 1996). However, establishing species relationships will require additional data and/or taxa, as current species phylogenies differ and depend on the data set and model of phylogenetic reconstruction (Martinez-Romero 1994; Martinez-Romero and Caballero-Mellado 1996; Young 1996; Young and Haukka 1996; Terefeword et al. 1998).

Extensive biochemical and molecular characterization of the legume–rhizobia interaction in agricultural strains has provided a framework for predictions about the dynamics of the symbiosis in native populations. An effective symbiosis requires several bacterial symbiotic (*sym*) loci, including nitrogen-fixation (*nif*) loci and nodulation (*nod*) genes that encode Nod factors which trigger nodule formation (reviewed in Long 1989; Martinez, Romero, and Palacios 1990; Relic et al. 1994). The common *nod* genes, *nodA, nodB,* and *nodC,* occur in all symbiotic rhizobia, encode the backbone of the Nod factor, and are often clustered with several additional, host-specific *nod* genes. Studies of the genera *Rhizobium* and *Sinorhizobium* indicate that *nod* genes may be the primary determinants of host-specificity in the field (Brewin, Wood, and Young 1983; Schofield et al. 1987; Spaink et al. 1987; Young and Wexler 1988; Martinez, Romero, and Palacios 1990; Laguerre et al. 1992; Relic et al. 1994; Laguerre et al. 1996). The ability of these *sym* loci to transfer across divergent chromosomal lineages has been shown in the lab (Spaink et al. 1987; Martinez, Romero, and Palacios 1990; Relic et al. 1994) and in agricultural populations (Young and Wexler 1988; Laguerre et al. 1992; Sullivan et al. 1995; Laguerre et al. 1996). Such transfer may be facilitated by the clustering of *nod* and *nif* genes on transmissible plasmids in *Rhizobium* and *Sinorhizobium* species (Martinez, Romero, and Palacios 1990) or on transposon-like elements in *Mesorhizobium loti* (Sullivan et al. 1995; Sullivan and Ronson 1998). The loss and reacquisition of *sym* plasmids may occur frequently in the field. In *Rhizobium leguminosarum,* for instance, the *sym* plasmid is not essential for the survival of nonsymbiotic strains in the soil, where it is lost at high frequencies (Laguerre, Bardin, and Amarger 1993). *Sym* gene transfer is also supported by discordance between the genealogies of *nod* and *16S rRNA* loci, which points to broad lateral transfer during the diversification of the family (Young and Johnston 1989; Dolbert, Brell, and Triplett 1994; Ueda et al. 1995; Young 1996; Young and Haukka 1996).

Due to the involvement of *nod* loci in host-specificity, *sym* gene transfer has been viewed as important in decoupling the adaptation of rhizobial populations to different host plant species and soil microhabitats (Sprent 1994). For example, rhizobial populations may adapt to distinct soil habitats and diverge at chromosomal housekeeping loci, yet share a ''pool'' of *sym* plasmids and nodulate the same host legume species across habitats. In *R. leguminosarum,* the transfer of host-specific symbiotic genes may allow the same chromosomal genotype to nodulate distinct host plants. A strain of this species may associate with one of at least three legume genera, *Trifolium, Viceae,* or *Phaseolus,* depending on its particular type of *sym* plasmid (Martinez, Romero, and Palacios 1990). In population level studies (Young and Wexler 1988; Laguerre et al. 1992; Laguerre, Bardin, and Amarger 1993; Sullivan et al. 1995; Laguerre et al. 1996) and phylogenetic surveys (Dolbert, Brell, and Triplett 1994; Ueda et al. 1995), stronger correlations of host plant groups with *nod* genotypes than with rhizobial chromosomal genotypes suggest that host plant associations are important in structuring *nod* gene divergence.

Recent studies of native systems show that the fast-growing rhizobia vary in both aspects of the symbiotic dynamics: *sym* gene transfer across chromosomal lineages and specificity of rhizobia–legume interactions (reviewed in Hirsch 1996). First, levels of *sym* gene transfer may be reduced in native settings by ecological, physiological, or geographical constraints on gene exchange, or the inability of *nod* loci to function in distinct chromosomal backgrounds. Recent population level analysis (Wernegreen, Harding, and Riley 1997) and phylogenetic characterization (Haukka, Lindstrom, and Young 1998) of native isolates show correlations of *nod* and chromosomal genotypes that argue against broad *sym* gene transfer. Further studies of native systems are necessary to assess the breadth of *nod* gene transfer and the ecological or genetic factors that may restrict lateral transfer.

Second, native rhizobia represent a wide range of host-specificity. Specific legume interactions have been demonstrated within populations of native *Rhizobium* (Wernegreen, Harding, and Riley 1997) and the distantly related *Bradyrhizobium* (Parker 1995). However, even among agricultural isolates, host interactions may be less specific. Single rhizobial genotypes may cross-inoculate diverse legume species or genera (Hashem et al. 1997), and, conversely, single legume species may associate with several rhizobial genotypes (Kishinevsky,

**Table 1**
**Rhizobial Strains, Host Legume Species, Collecting Regions, and GenBank Accession Numbers**

| Strain ID[a] | Host Plant Species | Collecting Location | Source[b] | Strain Name[c] | GenBank Accession Numbers | | |
|---|---|---|---|---|---|---|---|
| | | | | | *nodBC* | *GSII* | *16S rRNA* |
| MED1 | *Medicago sativa* | Australia | USDA | 1953 | AF063452 | AF063941 | AF063994 |
| MED2 | *Medicago orbicularis* | Syria | USDA | 1955 | AF063453 | AF063942 | AF063995 |
| MED3 | *Medicago truncatula* | Jordan | USDA | 1970 | AF063454 | AF063943 | AF063996 |
| R.MEL (gb) | *Medicago* | | | | *M11268* | *X17523* | |
| MED4 | *Medicago littoralis* | Greece | D. Gordon | WSM870 | AF063458 | AF063947 | AF063999 |
| MED5 | *Medicago sativa* | Pakistan | USDA | 1944 | AF063465 | AF063954 | |
| MEL1 | *Melilotus alba* | Virginia | USDA | 1079 | AF063459 | AF063948 | AF063988 |
| MEL2 | *M. alba* | Florida | USDA | 1083 | AF063460 | AF063949 | AF063989 |
| MEL3 | *Melilotus officianalis* | Oklahoma | USDA | 1088 | AF063461 | AF063950 | AF063990 |
| TRIG1 | *Trigonella suavissima* | Australia | D. Gordon | CC2129 | AF063455 | AF063944 | |
| TRIG2 | *T. suavissima* | Australia | D. Gordon | CC2155 | AF063456 | AF063945 | AF063997 |
| TRIG3 | *T. suavissima* | Australia | D. Gordon | CC2160 | AF063457 | AF063946 | AF063998 |
| TRIG4 | *Trigonella balansae* | California | USDA | 1114 | AF063462 | AF063951 | AF063991 |
| TRIG5 | *Trigonella corniculata* | California | USDA | 1115 | AF063463 | AF063952 | AF063992 |
| TRIG6 | *Trigonella foenumgraecum* | North America | USDA | 1177 | AF063464 | AF063953 | AF063993 |
| TRIF1 | *Trifolium hybridum* | Turkey | USDA | 2061 | AF063486 | AF063973 | AF063977 |
| TRIF2 | *Trifolium tembense* | Australia | USDA | 2163 | AF063487 | | AF063978 |
| TRIF3 | *Trifolium repens* | Brazil | USDA | 2235 | AF063488 | | AF063979 |
| TRIF4 | *Trifolium resupinatum* | Tunisia | USDA | 2240 | AF063489 | | |
| TRIF5 | *Trifolium* (annual species) | North America | USDA | 2M16+4 | AF063490 | AF063974 | AF063980 |
| TRIF6 | *Trifolium semipilosum* | Kenya | USDA | 7104 | AF063491 | | AF063981 |
| PISU1 | *Pisum sativum* | Former Yugoslavia | USDA | 2386 | AF063492 | | AF063987 |
| VICEAE (gb) | *Pisum* | | | | *Y00548* | | |
| PISU2 | *Pisum hortense* | U.S.A. | K. Lindstrom | CIAM1115 | AF063494 | | AF063983 |
| PISU3 | *Pisum sativum* | Illinois | USDA | 2370 | AF063496 | | AF063982 |
| GAL1 | *Galega officianalis* | Prague | USDA | 3394 | AF063466 | AF063955 | AF064000 |
| GAL2 | *G. officianalis* | New Zealand | K. Lindstrom | CIAM1141 | AF063467 | AF063956 | AF064001 |
| GAL3 | *G. officianalis* | New Zealand | K. Lindstrom | CIAM1143 | AF063468 | AF063957 | |
| GAL4 | *G. officianalis* | New Zealand | K. Lindstrom | CIAM1146 | AF063469 | AF063958 | AF063984 |
| GAL5 | *Galega* (commercial innoculant) | North America | K. Lindstrom | CIAM490 | AF063470 | AF063959 | AF064002 |
| GAL6 | *Galega orientalis* | Finland | K. Lindstrom | CIAM540 | AF063471 | AF063960 | AF064003 |
| *R. galegae* (gb) | *Galega* | | | | *X87578* | | |
| GLYC1 | *Glycyrrhiza mulenses* | Shanxi | E. Wang | Sh19351 | AF063450 | | AF063986 |
| GLYC2 | *Glycyrrhiza lepigota* | South Dakota | USDA | 3620 | AF063493 | AF063975 | AF064013 |
| AST1 | *Astragalus hypoglottis* | South Dakota | USDA | 3143 | AF063472 | | AF064017 |
| AST2 | *Astragalus falcatus* | Afghanistan | USDA | 3353 | AF063473 | AF063961 | AF064018 |
| AST3 | *Astragalus adsurgens* | North Dakota | USDA | 3357 | AF063474 | AF063962 | AF064014 |
| AST4 | *Astragalus americanus* | Alaska | USDA | 3584 | AF063475 | AF063963 | AF064009 |
| AST5 | *Astragalus uliginosus* | Altai | K. Lindstrom | CIAM0128 | AF063480 | AF063968 | AF064012 |
| HEDY1 | *Hedysarum mongolicum* | China | E. Wang | N95 | AF063451 | | AF063985 |
| HEDY2 | *Hedysarum boreale* | Alaska | USDA | 3876 | AF063476 | AF063964 | AF064008 |
| HEDY3 | *H. boreale* | Alaska | USDA | 3883 | AF063477 | AF063965 | |
| HEDY4 | *Hedysarum alpinum* | Moscow | K. Lindstrom | CIAM2026 | AF063481 | AF063969 | AF064005 |

**Table 1**
**Continued**

| Strain ID[a] | Host Plant Species | Collecting Location | Source[b] | Strain Name[c] | GenBank Accession Numbers | | |
|---|---|---|---|---|---|---|---|
| | | | | | *nodBC* | *GSII* | *16S rRNA* |
| HEDY5 | *H. alpinum* | Moscow | K. Lindstrom | CIAM2032 | AF063482 | AF063970 | AF064004 |
| HEDY6 | *Hedysarum pallens* | Israel | D. Kishinevsky | N1 | AF063484 | | AF064015 |
| HEDY7 | *H. alpinum* | Moscow | K. Lindstrom | CIAM1416 | AF063495 | AF063976 | AF064006 |
| OXY1 | *Oxytropis ampanulata* | Altai | K. Lindstrom | CIAM2209 | AF063483 | AF063971 | AF064016 |
| OXY2 | *Oxytropis deflexa* | Alaska | USDA | 4007 | AF063479 | AF063967 | AF064007 |
| OXY3 | *O. deflexa* | Alaska | USDA | 4004 | AF063478 | AF063966 | AF064010 |
| OXY4 | *O. deflexa* | Alaska | USDA | 4003 | AF063485 | AF063972 | AF064011 |
| *R. phaseoli* (gb) | *Phaseolus* | | | | *M58626* | *X67296* | |

[a] In cases in which sequences were downloaded from GenBank, strains are labeled "gb" and accession numbers are given in italics.

[b] Collection or individual from which rhizobial strain was obtained. Most strains were provided by Dr. P. VanBerkum at the U.S. Department of Agriculture (USDA). Other strains were provided by D. Gordon at the Australian National University, B. D. Kishinevsky of the Agricultural Research Organization of the Volcani Center, K. Lindstrom at the University of Helsinki, and E. Wang at the Beijin Agricultural Center.

[c] Rhizobial strain name in original collection.

Sen, and Yang 1996; Van Berkum et al. 1998). The relatively broad host range of native *Mesorhizobium* has been shown in several recent studies, including studies of the symbionts of the legume genera *Astragalus*, *Oxytropis*, and *Onobrychis*, which often group independently of host plant genus and are able to cross-inoculate legume genera (Prevost, Bordeleau, and Antoun 1987; Prevost et al. 1987; Novikova et al. 1994; Laguerre et al. 1997), and *Mesorhizobium* nodulating *Lotus* (Sullivan et al. 1996) and *Cicer* (Nour et al. 1995), which represent a heterogeneous group of genotypes associated with single legume species.

Mechanisms for this type of broad host range include allelic variation at *nod* loci (Brom et al. 1991), the presence of several *nod* genes (Perret et al. 1998), and the convergence of different strains onto the same nodulation phenotype despite the fact that they have distinct *nod* loci (Haukka, Lindstrom, and Young 1998). These possibilities may be distinguished by characterizing legume associations at both rhizobial *sym* and housekeeping loci. For example, *nod* genes of *M. loti* occur on a transposon-like mobile element that transfers into diverse chromosomal genotypes in the field and confers on divergent strains the ability to nodulate the same host plant species (Sullivan et al. 1995; Sullivan and Ronson 1998). This rapid transfer suggests that the host range dynamics of *Mesorhizobium* may involve the acquisition of distinct sets of *sym* loci. However, a broad host range may also result from convergence of distinct *nod* genotypes on the same nodulation phenotype, as suggested by the overlapping host ranges for deep lineages of *Mesorhizobium* and *Sinorhizobium*, which are distinct at *16S rRNA*, *nod*, and *nif* loci (Haukka, Lindstrom, and Young 1998).

The first aim of this study was to explore the genetic plasticity of native rhizobia isolated from several legume genera in the subfamily Papilionoideae. Tests of phylogenetic congruence are useful for detecting recombination, as congruence indicates that different gene regions share a similar evolutionary history, while incongruence points to cases of horizontal gene transfer (Dykhuizen and Green 1993). In this study, phylogenies were compared across *nodB*, *nodC*, and two chromosomal loci, *16S rRNA* and glutamine synthetase II (*GSII*), the latter of which is critical for ammonium assimilation and is considered important in core metabolism (Taboada et al. 1996). These three loci were compared in order to test whether the two chromosomal housekeeping genes subdivide strains similarly, and to explore whether *nod* loci transfer broadly across chromosomal lineages. While studies of agricultural *Rhizobium* and *Sinorhizobium* predict that *nod* loci may transfer across divergent chromosomal lineages, recent work suggests that broad *sym* gene transfer may be constrained among native isolates. The second goal of this study was to assess the importance of host associations in shaping the genetic differentiation of native rhizobia by examining legume associations at rhizobial chromosomal and *nod* genotypes. In previous studies of agricultural *Rhizobium* and *Sinorhizobium* strains, host plant maps well across rhizobial *nod* phylogenies but not

across chromosomal phylogenies (Dolbert, Brell, and Triplett 1994; Ueda et al. 1995; Young and Haukka 1996). However, the potential for broad-host-range strains, particularly among the *Mesorhizobium*, suggests that the host plant may map poorly across both *nod* and *GSII* phylogenies. In addition, comparing host associations at both *sym* and housekeeping loci allows us to explore the genetic mechanisms that underlie host range dynamics: specifically, whether the broad host range of particular chromosomal genotypes may be attributed to lateral transfer of *nod* loci.

## Materials and Methods

### Rhizobial Strains

Forty-seven rhizobial strains, representing isolates from nodules of 13 genera of the temperate herbaceous Papilionoideae across several continents, were obtained from existing culture collections (table 1). Strains are considered native in that they were isolated from native host plants in the host's native geographical range. Strains were selected to represent a diverse range of host plants within the temperate Papilionoideae. In cases in which host plant genera have overlapping ranges, regions of geographical overlap were sampled whenever possible in order to distinguish the effects of host plant and geographic location. Sequences of *nodB, nodC, GSII,* and/or *16S rRNA* were obtained from GenBank for four additional rhizobial strains.

### Molecular Methods

Cultures were streaked on yeast mannitol (YM) (Somasegaran 1994) plates for 3–5 days at 30°C. Single colonies were inoculated into 10 ml YM broth and grown shaking for 3–5 days at 28–30°C. Genomic DNA of each strain was prepared as described previously (Demezas et al. 1991). Regions of *nodBC, GSII,* and *16S rRNA* loci were amplified by PCR using the primers and reaction conditions described in table 2. PCR products were gel-purified (Quiagen gel-purification kit) and cycle-sequenced using the primers listed in table 2. A portion of each PCR product (approximately one third of the total length) was sequenced on both strands, and the remainder was sequenced on a single strand. All sequences obtained were submitted to GenBank, and accession numbers are given (table 1).

### Sequence Alignment

Multiple alignments of inferred protein sequences *nodB, nodC,* and *GSII* were performed with CLUSTAL W (Thompson, Higgins, and Gibson 1994), using the Gonnet distance matrix for estimating amino acid substitutions. For one region of each of *nodB* and *nodC* and the intergenic spacer (IGS) region between the loci, alignments were ambiguous across major subdivisions and were included only in within-group analyses (table 3).

The high variability of the *16S rRNA* region sequenced increases the likelihood of identifying distinct *16S rRNA* genotypes. However, this region alone is insufficient for phylogenetic analysis and has in fact been shown to give phylogenies that are discordant with other regions of the *16S rRNA* molecule (Eardly, Wang, and Van Berkum 1996). Therefore, the short region was not subjected to phylogenetic analysis, but was used only to identify the rhizobial species represented. To this end, the partial sequences were aligned against all existing prokaryotic *16S rRNA* sequences in the Ribosomal Database Project (RDP) (Maidak et al. 1997). A strain was assigned the same species ID as the sequence in the RDP that had the highest similarity to our submitted sequences and constituted more than 90% of the complete *16S rRNA* sequence.

### Phylogeny Estimations

Phylogenetic trees of rhizobial *nodB, nodC,* and *GSII* genes were inferred with maximum-parsimony (using PAUP, version 3.1.1; Swofford 1993) and neighbor-joining analyses (Saitou and Nei 1987) using MEGA (Kumar, Tamura, and Nei 1993) with Kimura's (1980) two-parameter nucleotide distances. Confidence in topologies was assessed using bootstrapping (100 replicates).

Genealogies of *nodB* and *nodC* were rooted with *R. leguminosarum* bv. *phaseoli* (or *Rhizobium phaseoli*). The *GSII* tree was left unrooted due to lack of an appropriate outgroup. Relationships among legume genera were obtained from combined morphological and molecular phylogenetic analysis of previous studies (see legend of fig. 5).

Native strains were placed into a phylogenetic framework by assigning species identifications based on similarity with type strains at the region of *16S rRNA* sequenced here (as described above). A maximum-likelihood-based phylogeny of the full *16S rRNA* of the type rhizobial species represented, as well as other related species, was downloaded from the RDP (Maidak et al. 1997). This phylogeny agrees with broad relationships among rhizobia published elsewhere (Martinez-Romero 1994; Martinez-Romero and Caballero-Mellado 1996; Young 1996; Young and Haukka 1996; Terefeword et al. 1998). Given the uncertain resolution of some species relationships (see *Introduction*), the *16S rRNA* data set was used only to identify the species represented and to estimate broad (generic) relationships for the purposes of this study.

### Character-Mapping Analysis

In order to determine the importance of host plant genus and geographical region in shaping genetic divergence of rhizobial loci, both factors were mapped as discrete characters across *nod* and *GSII* genealogies. Character fit was estimated by the length across topologies and the consistency index (CI; MacClade, version 3.05; Maddison and Maddison 1992). For the purpose of character mapping, ambiguous nodes in *Rhizobium* genealogies were treated in two ways: (1) Nodes lacking strong support (bootstrap value $< 80\%$) were treated as ''soft polytomies,'' which allow the shortest possible length of a character across the node. (2) Nodes with weak support (bootstrap value $< 50\%$) were randomly resolved, and character fit was calculated across each of 100 possible resolutions. For both methods, significance

**Table 2**
**Primers and Cycling Conditions Used to Amplify and Sequence Rhizobial *nodBC, GSII,* and *16S rRNA* Loci**

A. Primer Names, Concentrations of Magnesium Sulfate, Annealing Temperatures, and Primer Sequences for PCR Amplification

| Primer[a] | $MgSO_4$ Concentration | Annealing Temperature | Primer Sequence |
|---|---|---|---|
| Amplification of *GSII*, extension time 55 s | | | |
| GSF1A | 6 mM | 72°C | ATGACMAARTWTAAGCTCGAGTACATYTGGCTYGAT |
| GSR1A | | | AGRAYSWNNGAASSATCTGGTAGGGGTCGCC |
| GSF1B | 2 mM | 70°C | TAAGCTCGAGTACATTYTGGCTYGAT |
| GSR1B | | | GAASSATCTGGTAGGGTCGCC |
| Amplification of *nodB* and *nodC*, extension time 1 min 20 s | | | |
| nodB2F | 4 mM | 55°C | TGACGTTYGACGACGGTCCNAATCC |
| 566R2 | | | (Ueda 1995) |
| nodB3F | 2 mM | 67°C | TACCTGACSTTYGAYGACGGTCC |
| nodC3R | | | GGGCCGCAGCAACACATAACNGC |
| Amplification of *16S rRNA*, extension time 30 s | | | |
| Y1 | 4 mM | 62°C | Young et al. (1991) |
| Y2 | | | Young et al. (1991) |

B. Internal Primers Used for Cycle Sequencing of *GSII* and *nodBC,* and Rhizobial Isolates Sequenced with Each Primer[b]

| Internal Primer | | Isolates Sequenced with Specific Primers |
|---|---|---|
| *GSII* | | |
| GSR2seq | TGCATGCCSGAGCCGTTCCAGTC | |
| *nodBC* | | |
| seqnod.f2 | CTTTCAAGATCGATGA | TRIF |
| NDTRIFR2 | GGATAGTCCTGGTTGGC | TRIF (except TRIF1) |
| 2163nod2R | GGATAATCCTGGTTACGTAGG | TRIF1 |
| ND4F2386 | TTGGTCCGGCACTTACGAG | PISUM |
| NOD2386R | CATAGRCAGCGCGCTGAGCC | PISUM |
| nod4seqF | CCTTATGGTGTCTGGAGCGAGGAAGC | MED, MEL, TRIG, GLYC1, HEDY1 |
| nd4Rmed | CATCATCAACGACATAGACTCGC | MED, TRIG, TRIF |
| N95nodR2 | TCTCTGGGAGCAGAATGAA | HEDY1 |
| ShnodBR | TCTGGATCTCGCATCTTCGAG | GLYC1 |
| nodgal2F | CATGCGCTGGCGCTTCTAT | GAL |
| nodgal2R | AAGATTGCCTCACGATTGC | GAL |
| nd4Fciam | ACGTCAAATAGTCGAGGCAAGC | *Mesorhizobium* (except GLYC2, AST1, AST2, OXY1, HEDY6) |
| nd4Rciam | CGTCGCGATTTCCAGAACCGTCATCA | *Mesorhizobium* |
| nd4F3620 | CCTCGAGATTGGTCTCGC | GLYC2 |
| 3143nod2F | TCATGATCCGGTGTCCMCAGGC | AST1 |
| 3353nod2F | CATCAAAGTGGCTTGTGCTC | AST2 |
| N1nod4F | CCTATGGGATCTGGA | OXY1, HEDY6 |

[a] Primers with F's in their names are forward primers that match the coding strand (i.e., anneal to the noncoding strand), and those with R's in their names are reverse primers that match the noncoding strand (i.e., anneal to the coding strand). All primer sequences are listed 5′ to 3′. *PCR conditions*: Vent exo+ (proofreading) polymerase was used in 1 × buffer (supplied in 10 × concentration by New England Biological Co.). Primers were used at 1 μM final concentration, dntp's at 0.2 mM, and Vent at 1 U/100 μl reaction; 10–50 ng of genomic DNA was used as a template. Reactions were 50 or 100 μl final volume. PCR cycling conditions were identical for each reaction, except for annealing temperature and extension time, which are listed for each primer pair. General cycling conditions were as follows: 96°C for 1 min 30 s; 30 cycles of 96°C for 30 s, annealing temperature for 30 s, 72°C for extension time; and 72°C for 3 min.

[b] PCR products were gel-purified (Qiagen gel-purification kit) and diluted to approximately 30 ng/μl. Products were cycle-sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer), following the standard protocol for the Perkin Elmer 9600 PCR machine, and ethanol precipitation of sequencing reactions. PCR primers were used for sequencing the end regions of PCR products, and internal regions were sequenced with the primers listed.

was determined by comparing the length of the character across the empirical tree to the distribution of lengths across a pool of 100 random topologies of the same set of taxa (MacClade, version 3.05; JMP 3.1.5 program, SAS Institute, Cary, N.C.). *P* values represent the proportion of random topologies across which the character has the same or better fit. *P* values for method 2, above, were based on the maximum character length across random resolutions. Unless stated otherwise, *P* values are the same for methods 1 and 2.

### Tests of Phylogenetic Congruence

Congruence between *nod* genes and host plant phylogenies was explored using TREEMAP (Page 1994). Levels of similarity between topologies of the host legume tree and the *nodB* and *nodC* genealogies were tested by comparing the number of inferred cospeciation events shared by two trees with the number shared between one empirical tree and a set of randomly generated topologies. Congruence among rhizobial *nodB, nodC,* and *GSII* data sets was assessed by Templeton's

**Table 3**
**Rhizobial Gene Regions Sequenced in this Study, Numbers of Sites, and Numbers of Parsimony-Informative Positions**

| Gene Region | No. of Nucleotides Sequenced[a] | No. of Nucleotides Included in the Analysis[b] | No. of Parsimony-Informative Sites[b] |
|---|---|---|---|
| *nodB* | 555 | 522 | 304 |
| *nodC* | 549 | 480 | 277 |
| *GSII* | 891 | No ambiguous regions | 282 |
| IGS (intergenic spacer between *nodB* and *nodC*) | 35 | 0 | |
| *16S rRNA* | 284 | No ambiguous regions | 33 (48 variable sites) |

[a] Length of region sequenced includes gaps introduced in alignment.
[b] Sites with ambiguous alignment were excluded from the total analysis across all sequences.

test (as modified by Felsenstein 1985) and the incongruence length difference (ILD) test (Farris et al. 1985).

## Results

### Isolates Represent Several Distinct Species of Fast-Growing Rhizobia

Strains were affiliated with type rhizobial species based on their similarity at *16S rRNA* with other sequences in the RDP (Maidak et al. 1997). Several different species within the fast-growing rhizobia are represented, including: *R. leguminosarum, Rhizobium etli, R. galegae, Sinorhizobium meliloti, M. loti, Mesorhizobium huakuii,* and *Mesorhizobium ciceri.* The RDP *16S rRNA* phylogeny (Maidak et al. 1997) clearly shows that *Sinorhizobium* and *Rhizobium* form a group distinct from *Mesorhizobium* (fig. 1).

### Phylogenies of Rhizobial Loci

For each rhizobial locus sampled (*nodB, nodC,* and *GSII*), groups that were strongly supported by parsimony methods were not contradicted by distance meth-

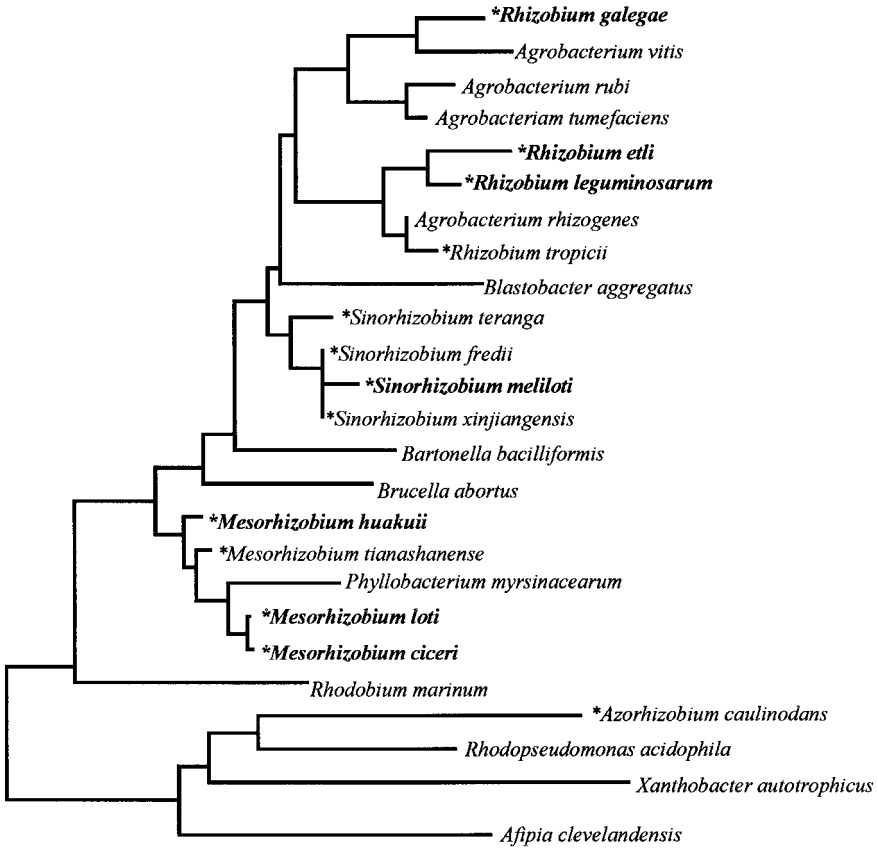


FIG. 1.—Phylogeny of several type rhizobial species (*) and other related bacteria, based on *16S rRNA.* Relationships are based on a maximum likelihood of full (>90% complete) *16S rRNA* sequences (Felsenstein 1981; Olsen et al. 1994), and branch lengths are proportional to the number of nucleotide substitutions per base pair. The tree was downloaded from the Ribosomal Database Project (Maidak et al. 1997). Although species relationships within genera are not well resolved, the relationships among genera generally agree among published phylogenies (see text). Based on similarity at a portion of *16S rRNA,* strains in this study were affiliated with the type species in boldface.

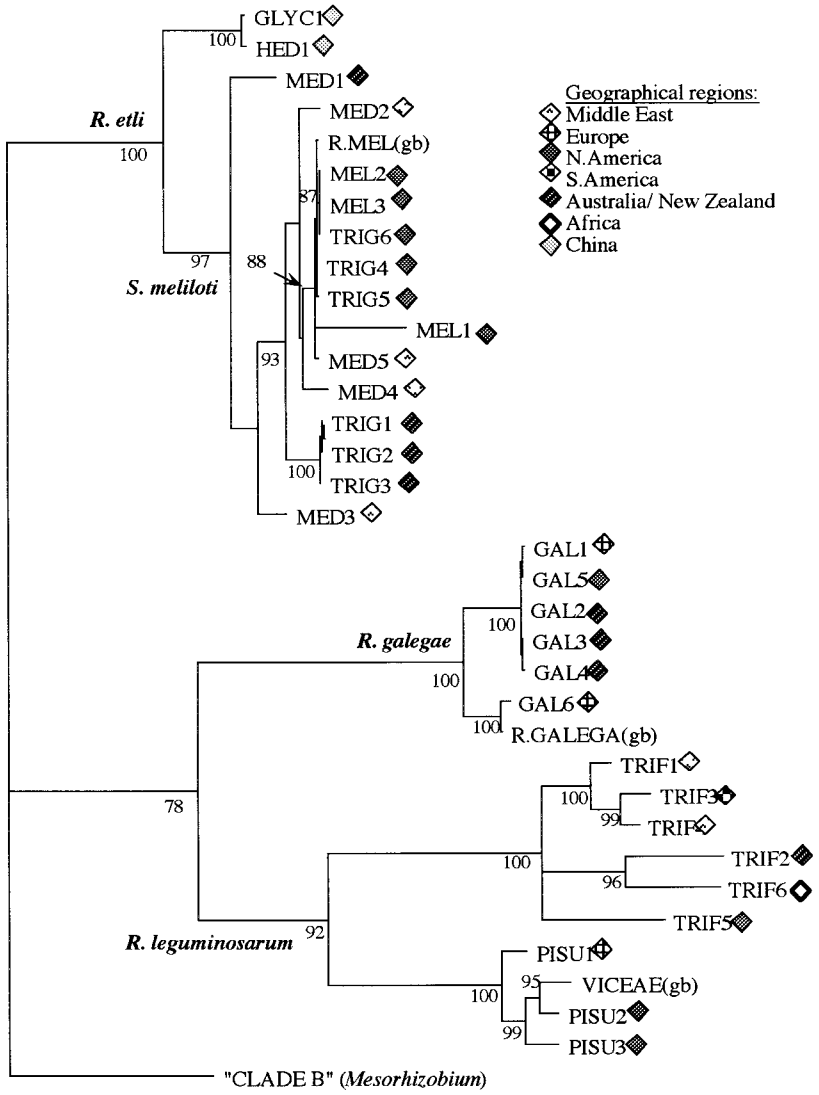


FIG. 2.—Phylogeny of combined data of *nodBC*, and the intergenic spacer (IGS) region across *Rhizobium* and *Sinorhizobium* isolates. Both *nodB* and *nodC* divide strains into two major clades: clade A, presented, and clade B (*Mesorhizobium*), for which *nodB* and *nodC* were treated separately (see text and fig. 3). Strains are labeled by geographical regions and *16S rRNA* groups (in boldface at nodes). Strains are named according to the host plant genus from which they were isolated (table 1). The tree is based on 619 parsimony-informative sites, is rooted with members of the *Mesorhizobium*, and is the strict consensus of two most-parsimonious trees (tree length = 1,747, consistency index [CI] = 0.6027, homoplasy index [HI] = 0.3973, retention index [RI] = 0.8806, rescaled consistency index [RC] = 0.5308). Branch lengths are proportional to the number of nucleotide changes, and bootstrap values above 80% are given at nodes.

ods, and vice versa. Only parsimony-based trees are presented.

Both *nodB* and *nodC* divide strains into two large clades (*nod* clades A and B) one branch from the root (figs. 2 and 3). *Nod* clade A includes only *Rhizobium* and *Sinorhizobium*, and *nod* clade B includes only *Mesorhizobium*. For *Rhizobium* and *Sinorhizobium*, *nodB* and *nodC* data sets are not significantly heterogeneous by either Templeton's test or the ILD test ($P > 0.05$ for each test) and are therefore combined in a single phylogenetic analysis (fig. 2). Within clade A, *nod* genes of *Rhizobium* are generally distinct from those of *Sinorhizobium*. Across *Mesorhizobium* (*nod* clade B), *nodB* and *nodC* data sets are significantly heterogeneous ($P < 0.01$ for each test); consequently, the *nod* loci are considered separately for this genus (fig. 3). Despite their chromosomal divergence, *M. loti* are virtually identical at *nodB* and *nodC* (*nod* clade ''B2,'' fig. 3). Phylogenetic analysis at *GSII* also subdivides strains into three well-supported clades (1, 2, and 3, corresponding to *Sinorhizobium*, *Rhizobium*, and *Mesorhizobium*, respectively).

### *NodB*, *nodC*, *16S rRNA*, and *GSII* Subdivide Strains into the Same Major Groups

*GSII*, *16S rRNA*, and *nod* loci divide strains into the same three major groups, corresponding to the three genera of fast-growing rhizobia (figs. 2–4). This correspondence suggests that transfer of *nod* genes and *GSII* among genera is restricted. The exceptions are *R. leguminosarum* bv. *phaseoli*, which falls outside both of the major groups at *nodB* and *nodC* but groups closely

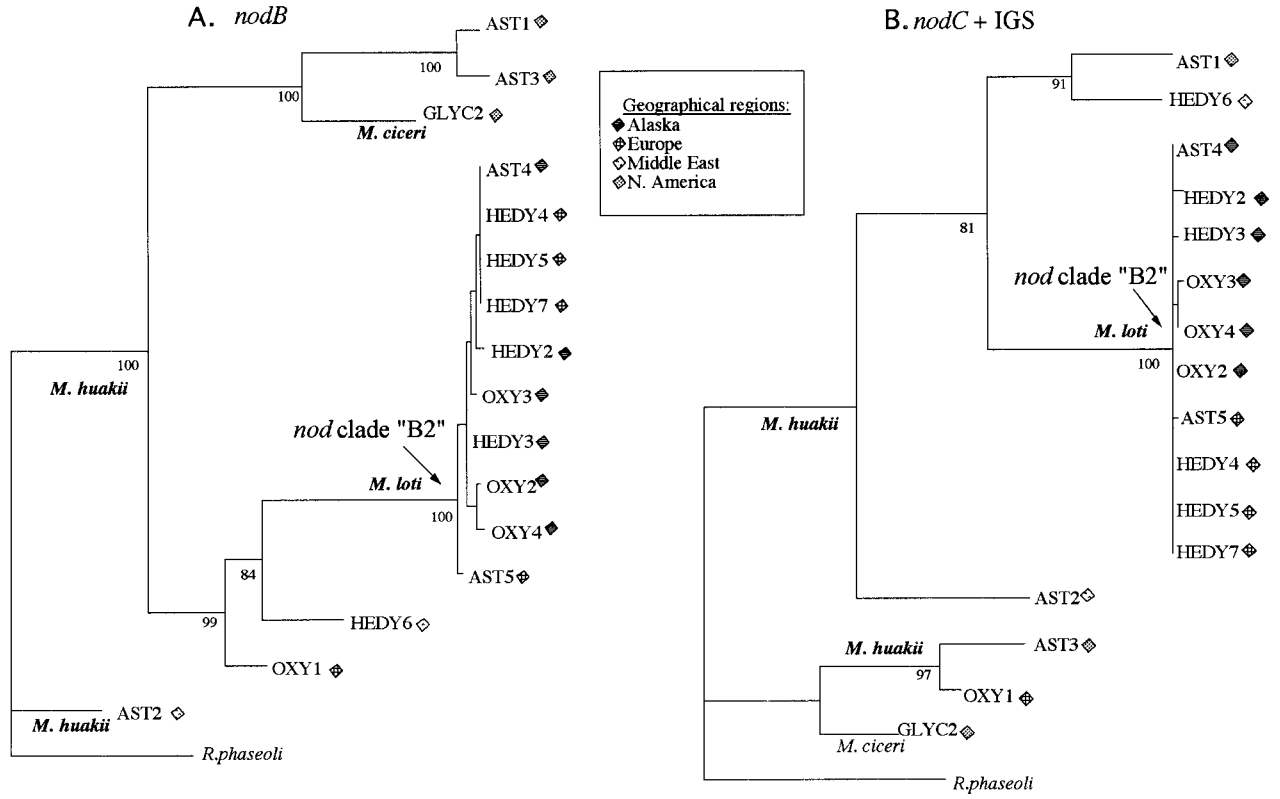


FIG. 3.—Relationships among *Mesorhizobium* isolates at *nodB* (left) and combined data of *nodC* and the intergenic spacer region (right). The two *nod* data sets are significantly heterogeneous (see text) and were analyzed separately. *Mesorhizobium* strains are labeled by geographical regions and *16S rRNA* groups (in boldface at nodes), and are named according to the host plant genus from which they were isolated (table 1). One closely related *nod* group, marked ''B2,'' is virtually identical at *nodB* and *nodC. The nodB* and *nodC* trees are based on 164 and 144 parsimony informative sites, respectively, and represent the single most-parsimonious tree for *nodB* (tree length = 304, consistency index [CI] = 0.7533, homoplasy index [HI] = 0.2467, retention index [RI] = 0.8348, rescaled consistency index [RC] = 0.6288) and the consensus of two most-parsimonious trees for *nodC* (tree length = 306, CI = 0.6667, HI = 0.3333, RI = 0.7323, RC = 0.4882). Branch lengths reflect the number of nucleotide changes, and bootstrap values above 80% are given at nodes.

with another *R. leguminosarum* isolate (TRIF5) at *GSII,* and the strains GLYC1 and HEDY1, which group with *Sinorhizobium* at *nodBC* but are identical to *R. etli* at the *16S rRNA* region sequenced. In addition, the fact that two *Rhizobium* species (*R. galegae* and *R. leguminosarum*) are distinct at *GSII* and *nod* loci suggests that transfer between them is restricted.

## Evidence for Transfer of *nod* Genes Within Major Groups

Within each genus, *nod* and *GSII* genes are significantly heterogeneous by both Templeton's test ($P < 0.01$ for all mappings of *nod* data sets across the *GSII* phylogenies; table 4) and the ILD test ($P < 0.01$ for each test; table 4). This heterogeneity of *nod* and *GSII* data sets suggests that the two loci have different evolutionary histories within major clades. For example, heterogeneity of *nod* and *GSII* data sets within each of the genera *Rhizobium* and *Sinorhizobium* argues for *nod* gene transfer within each, and the grouping of *R. etli* strains GLYC1 and HEDY1 with *Sinorhizobium* at *nodB* and *nodC* argues for one case of transfer across the two genera. Significant data set heterogeneity also suggests *sym* gene transfer within the *Mesorhizobium.* The clearest case of *nod* gene transfer is found for *M. loti.* The *GSII* phylogeny suggests that *M. loti* is basal, diverse, and paraphyletic; however, *M. loti nod* genes form a tight cluster of virtually identical genotypes (fig. 3 vs. fig. 4). This discrepancy strongly suggests recent lateral transfer of *nod* genes into diverse *M. loti* strains.

## Mapping of Host Plant Genus and Geographical Region Across Rhizobial Phylogenies

The relationship among the rhizobial genera represented generally reflects that among the host plants from which they were isolated. *Sinorhizobium* and *Rhizobium,* which group apart from *Mesorhizobium,* were isolated primarily from *Medicago* and *Trigonella* (*Sinorhizobium*) and from *Trifolium, Galega,* and *Pisum* (*Rhizobium*) (fig. 2). *Mesorhizobium* were isolated exclusively from the basal legume genera *Glychyrrhiza, Hedysarum, Astragalus,* and *Oxytropis* (fig. 3). The exceptions are the *R. etli* strains HEDY1 and GLYC1, which were isolated from *Glycyrrhiza* and *Hedysarum,* respectively.

Strict associations between *nod* genotypes and host legume groups are apparent for *R. leguminosarum* and *R. galegae,* as *nod* genes of *Trifolium, Galega,* and *Pi-*

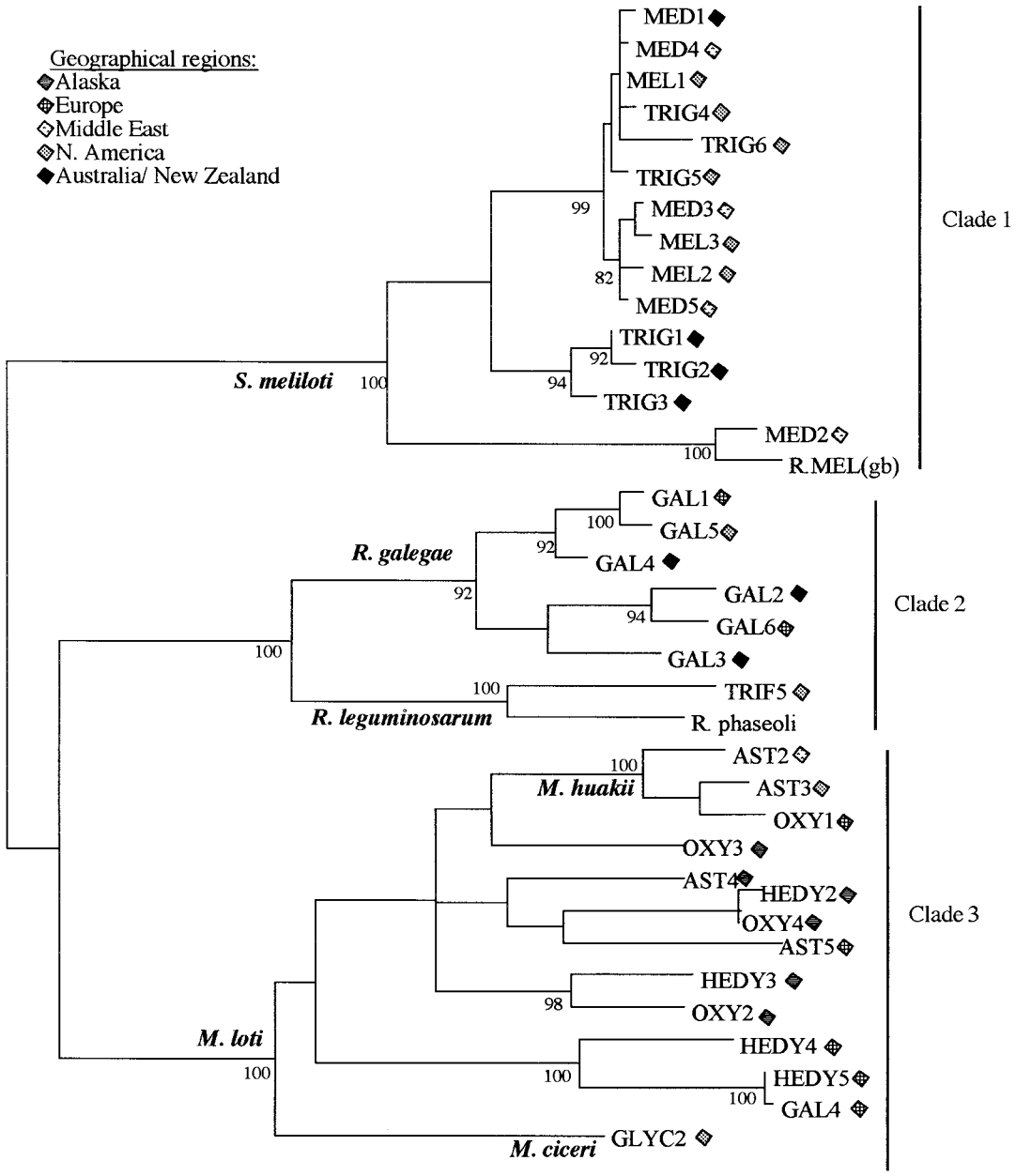


FIG. 4.—Phylogeny of rhizobial *GSII* loci. Analysis was based on 282 parsimony-informative sites, and the tree presented is the 50% majority-rule consensus tree of 368 MP trees (tree length = 770, consistency index [CI] = 0.5091, homoplasy index [HI] = 0.4909, retention index [RI] = 0.8371, rescaled consistency index [RC] = 0.4261). Nodes defining the three major groups (clades 1, 2, and 3) are marked and correspond with rhizobial genera *Sinorhizobium, Rhizobium,* and *Mesorhizobium,* respectively. Strains are labeled by geographical region and *16S rRNA* groups, and strain names reflect host plant genus (table 1). The tree presented is unrooted (but was rooted at the midpoint for illustrative purposes only). Branch lengths reflect the number of nucleotide changes, and bootstrap values above 80% are given at nodes. Bootstrap analysis for each of the three major clades was performed by resampling five members of each clade; in each resampling, 100% bootstrap support was obtained for each of the three clades. Bootstrap support was estimated by including all strains within a given clade, as well as representatives of the other two clades as outgroups.

*sum* form three distinct clades. Likewise, *Sinorhizobium* strains isolated from the closely related legume genera *Trigonella, Melilotus,* and *Medicago* form a single clade at *nodBC* (fig. 2). This strong association between host plant genus and *nodBC* alleles is reflected in a high CI (0.78) and the fact that character fit is better than that across random topologies ($P < 0.01$). The mapping of host plant genus across the *Mesorhizobium nod* phylogeny is less clear (fig. 3). Host plant genus also maps well when ambiguous nodes are collapsed to soft polytomies (method 1; $P < 0.01$ for both *nodB* and *nodC*), but it maps poorly when ambiguous nodes are randomly resolved (method 2; $P = 0.2$ for *nodB*, $P = 0.53$ for *nodC*). Similarly, for *GSII,* host plant maps well across *Sinorhizobium* and *Rhizobium* but relatively poorly across *Mesorhizobium.* Host plant genus has a shorter

**Table 4**
**Comparison of Rhizobial *nodB, nodC,* and *GSII* Loci Within the Major Clades Detected Using Templeton's Test and the Incongruence Length Difference Test**

| Strains Compared[a] | Data Set | Distin. Char. | Frequency of Characters Shorter Across: | | *P* Values[b] | |
|---|---|---|---|---|---|---|
| | | | *nod* tree | *GSII* tree | Temp.[c] | ILD[d] |
| *nodB* (clade B) vs. *GSII* (clade 3) | *nodB* | 89 | 0.89 | 0.11 | <0.01 | <0.01 |
| | *GSII* | 62 | 0.13 | 0.87 | <0.01 | |
| *nodC* (clade B) vs. *GSII* (clade 3) | *nodC* | 67 | 0.87 | 0.13 | <0.01 | <0.01 |
| | *GSII* | 72 | 0.58 | 0.41 | >0.05 | |
| *nodBC* (MED, MEL, TRIG) vs. *GSII* (clade 1) | *nodBC* | 50 | 0.98 | 0.02 | <0.01 | <0.01 |
| | *GSII* | 61 | 0 | 1 | <0.01 | |
| *nodBC* (GAL, TRIF5) vs. *GSII* (clade 2) | *nodBC* | 26 | 0.85 | 0.15 | <0.01 | <0.01 |
| | *GSII* | 22 | 0.59 | 0.41 | >0.05 | |

Note.—Data Set = the data set mapped across the *nod* and *GSII* topologies, Distin. Char. = the number of nucleotide sites in that data set that are shorter across one of the two topologies, Temp. = Templeton's test, and ILD = the incongruence length difference test.

[a] Within *nod* clade B, the loci *nodB* and *nodC* are treated separately in the comparison with *GSII* (see text).

[b] Significant *P* values for either Templeton's test or the ILD test indicate that the two molecular data sets are heterogeneous.

[c] As modified by Felsenstein (1985). Any ambiguities in rhizobial genealogies were treated to make Templeton's test more conservative (i.e., to minimize the chance of erroneously disproving the null hypothesis that the two data sets are congruent). In cases in which more than one most-parsimonious tree was obtained for a given data set, the most-parsimonious tree most similar (i.e., with the shortest symmetric difference; PAUP 3.1, Swofford 1993) to the alternative topology was selected, as suggested by Cunningham (1997).

[d] The ILD test (Farris et al. 1985) was conducted using the partition homogeneity test of PAUP4* (version 40d64, written by David L. Swofford).

length over the entire *GSII* phylogeny than it does across random topologies ($P < 0.01$).

### *Nod* Versus Legume Phylogenies

Given the close association of *nod* clades with host plant groups, topologies of the rhizobial *nod* phylogenies were compared with known relationships among host plant genera in order to test the historical stability of the plant–bacteria association. The topology of the relationship among the legume genera sampled was compiled from the literature, and *nod* phylogenies were pruned to preserve all evidence for host switching (see legend to fig. 5). A very broad agreement between the *nod* and legume phylogenies is apparent in the association of strains with derived *nod* alleles (*Rhizobium* and *Sinorhizobium*) with relatively derived legume genera, and in the association of *Mesorhizobium nod* loci with basal legume genera. The fit of the legume tree to *nodB* is significantly higher than expected by chance ($P < 0.002$), but the general congruence between the legume and *nodC* trees is not statistically significant at the 5% level ($P = 0.054$).

### Mapping of Geographical Region Across Rhizobial Phylogenies

Compared with host plant genus, geographical region has a relatively poor fit across rhizobial genealogies, as closely related strains often nodulate the same host plant genus but represent several distinct geographical regions (table 1). The CI of geographical region across the *nod* genealogy of *Rhizobium* and *Sinorhizobium* is relatively low (0.46), although this fit is better than that across random topologies ($P < 0.01$). Across the *Mesorhizobium nod* phylogeny, the mapping of geographical region depends on the treatment of poorly resolved nodes (data not presented). Likewise, clades at *GSII* associated with a particular legume genus often represent several geographical regions (fig. 4); nevertheless, geographical region has a shorter length across the *GSII* phylogeny than expected by chance ($P < 0.01$).

## Discussion

Empirical studies of genetic differentiation in bacteria primarily focus on clinical or agricultural isolates, which may be subject to strong human-mediated selective pressures. In contrast, this study tests for phylogenetic congruence among loci in order to test for lateral gene transfer among native isolates of rhizobia associated with temperate legumes. Studies of agricultural strains suggest that specificity in the legume–bacterial association may be largely due to host-specific nodulation loci (Martinez, Romero, and Palacios 1990). The occurrence of these genes on transmissible plasmids or elements in several rhizobial species may facilitate their transfer across divergent groups and allow distinct chromosomal lineages to nodulate the same host plant (Sprent 1994; Laguerre et al. 1996; Young and Haukka 1996; Sullivan and Ronson 1998). However, recent studies suggest that patterns of lateral gene transfer and host-specificity may differ considerably between agricultural and native isolates and may vary across rhizobial lineages (Parker 1995; Wernegreen, Harding, and Riley 1997; Haukka, Lindstrom, and Young 1998). Given the complex dynamics of the interaction in native settings, the goals of this study were to investigate the breadth of lateral transfer of *nod* loci across native rhizobial chromosomal lineages and to explore the host legume associations at rhizobial *sym* and housekeeping loci.

### Coherence of Major Subdivisions

The degree of congruence among genealogies of symbiotic *nod* genes and the chromosomal locus *GSII* was found to depend on the taxonomic level considered. With the exception of an agricultural isolate included

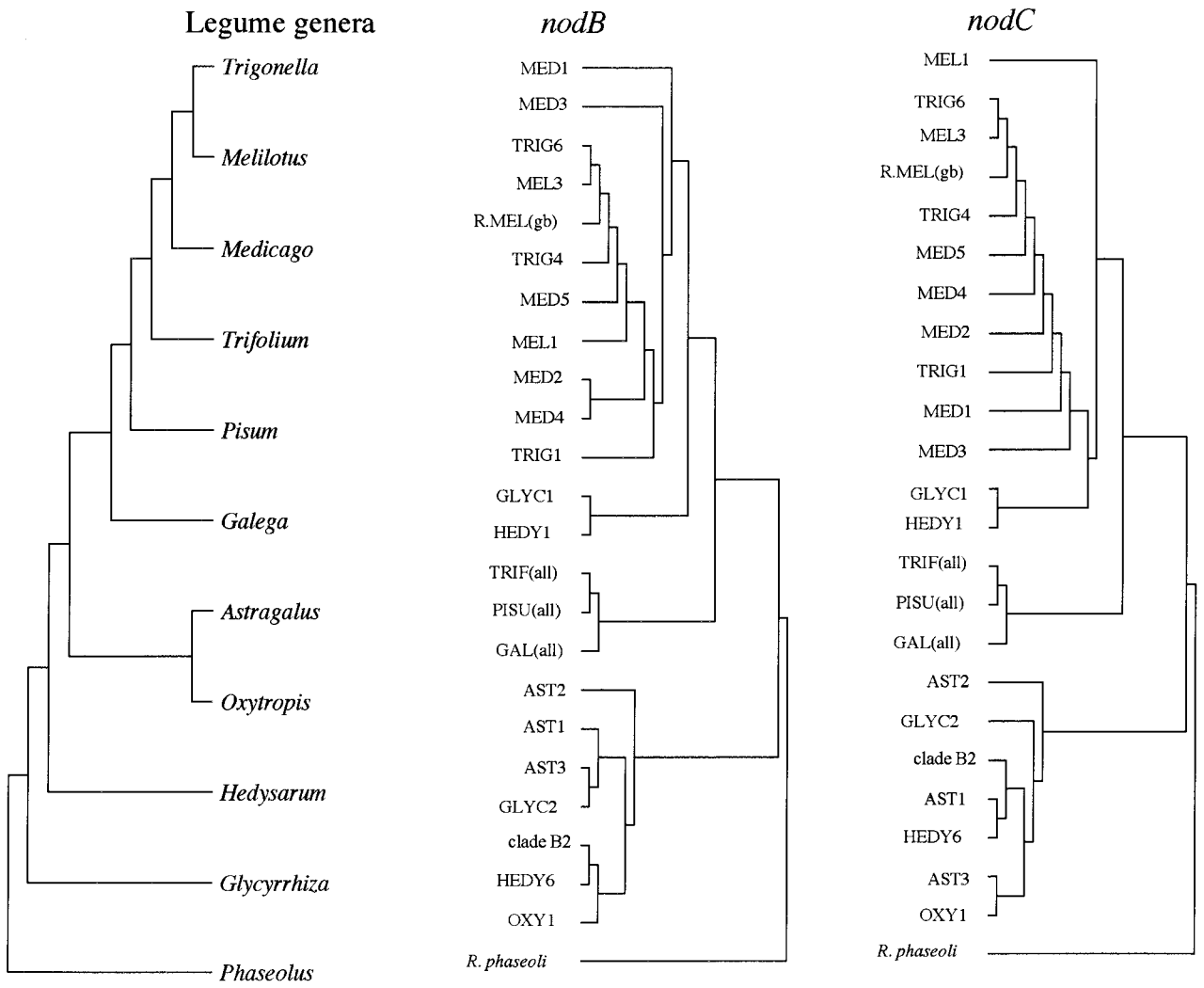


FIG. 5.—Phylogenies of legume genera (left) versus those of rhizobial *nodB* (middle) and *nodC* (right) loci. *NodB* and *nodC* were considered separately in the analysis, since they give different topologies for the *Mesorhizobium* (see text). Relationships among legume genera are based on combined morphological and molecular phylogenetic analyses of previous studies. The four tribes, along with genera included in this study and references for previous studies on which phylogenetic relationships are based, are as follows: Hedysareae (genus sampled in this study: *Hedysarum*; Chappill 1995; Sanderson and Liston 1995), Vicieae (*Pisum*; Sanderson and Liston 1995), the paraphyletic tribe Galegeae (*Astragalus, Oxytropis, Galega, Glychyrrhiza*; Chappill 1995; Sanderson and Liston 1995), and Trifolieae (*Melilotus, Medicago, Trigonella, Trifolium*; Small 1987). *Nod* clades were pruned to a single representative taxon only if (1) they formed single clades at both *nodB* and *nodC,* and (2) they were associated with just one host plant genus. This pruning method preserves all evidence for paraphyletic or polyphyletic groups at *nod* loci associated with single host plant genera and thus retains all evidence of ''switching'' of *nod* alleles among different hosts. Two groups had ambiguous relationships at *nod* loci: *Sinorhizobium,* for which the most parsimonious resolution was retained (despite bootstrap values less than 80%), and *Mesorhizobium loti* (*nod* clade ''B2''), which was pruned to a single ''polymorphic'' taxon associated with three host genera due to the scarcity of parsimony-informative sites at either *nodB* or *nodC.* Similarity between phylogenies of the host legumes and *nod* loci was assessed using TREEMAP, which compares the maximum number of cospeciation events between host and symbiont trees to the distribution of cospeciation events between the symbiont tree and 100 randomized host trees. The *nodB* and legume trees were more similar than expected by chance (seven cospeciation events; $P < 0.002$). While similar, the *nodC* and legume tree are not significantly similar at the 5% level (nine cospeciation events; $P < 0.054$).

for comparison (*R. leguminosarum* bv. *phaseoli*) and two isolates of *R. etli,* strains were divided into the same major clades at *GSII, 16S rRNA,* and *nod* loci. Based on type rhizobial species at *16S rRNA,* the three major subdivisions detected correspond with the genera *Rhizobium, Sinorhizobium,* and *Mesorhizobium.* This broad agreement of *nod* loci with two housekeeping genes, *GSII* and *16S rRNA,* argues for the stability of symbiotic loci within major chromosomal lineages and the independent genetic differentiation of rhizobial genera. A similar pattern of genetic coherence was found among native *Mesorhizobium* and *Sinorhizobium* nodulating leguminous trees, for which *nod, nif,* and *16S RNA* subdivided strains similarly (Haukka, Lindstrom, and Young 1998).

Associations with distinct ecological niches may constrain gene flow among native bacterial strains (Maynard Smith et al. 1993; Cohan 1994*a*, 1994*b*, 1995; Souza et al. 1994). Among symbionts, associations with distinct host taxa may restrict gene transfer and contrib-

ute to the genetic divergence of populations and species lineages. In this study, a broad correspondence between rhizobial clades and host legume genera suggests that host plant associations may be important in structuring genetic divergence of nodulation and housekeeping loci. Likewise, a general similarity of *nod* and legume phylogenies suggests that the association between the symbionts has been historically stable. Including geographically diverse strains from single host plant genera allowed a comparison of host and geographic effects. In several cases, *Rhizobium* or *Sinorhizobium* were similar at *nod* genes despite their collection from distant geographic locations. While our data suggest that host associations may restrict lateral transfer of *nod* genes, they are not required for the type of genetic coherence observed. Haukka, Lindstrom, and Young (1998) found that deep rhizobial lineages, distinct at chromosomal and *sym* loci, may associate with the same legume species. These authors suggest that *sym* gene transfer may be restricted by the inability of *nod* genotypes to function in divergent rhizobial chromosomal backgrounds.

### Transfer Within Major Subdivisions

In contrast to the genetic stability of major lineages, significant incongruence (i.e., data set heterogeneity) of chromosomal and *nod* phylogenies within each of the three genera points to the transfer of *nod* loci among closely related strains. Perhaps the clearest case of lateral gene transfer is the near identity of *M. loti* strains at *nodB* and *nodC,* despite their divergence at *GSII.* While the pattern could be explained by accelerated rates of evolution of *GSII* within *M. loti, GSII* is known to be important in core metabolism and has been shown to evolve at generally constant rates in other taxa (Pesole et al. 1991). This high chromosomal diversity within *M. loti* agrees with previous taxonomic work suggesting that the species may actually represent several ''genomic species'' (Sullivan et al. 1996). *Sym* gene transfer across this diverse group may be facilitated by the occurrence of *nod* loci on a mobile element, as was recently demonstrated for an agricultural *M. loti* isolate (Sullivan et al. 1995; Sullivan and Ronson 1998).

Patterns of host plant associations differ considerably for *Mesorhizobium* versus *Rhizobium* and *Sinorhizobium.* Host plant maps relatively well across *Rhizobium* and *Sinorhizobium,* as *nod* and *GSII* clades generally correspond with single host plant groups. On the other hand, the wide distribution of *Hedysarum, Astragalus,* and *Oxytropis* isolates across *nod* and *GSII* phylogenies suggests that *Mesorhizobium* has a relatively broad host range, a clear example of which is the association of closely related *M. loti nod* alleles with each of the three legume genera. This lack of legume associations at *nod* loci argues that the broad host range of *Mesorhizobium* results from convergence of distinct strains onto the same nodulation phenotype, and the ability of the legume genera *Hedysarum, Astragalus,* and *Oxytropis* to associate with several distinct rhizobial genotypes (as suggested previously: Prevost, Bordeleau, and Antoun 1987; Prevost et al. 1987; Novikova et al. 1994; Nour et al. 1995; Sullivan et al. 1996; Laguerre et al. 1997). One alternative mechanism for the apparently broad host range is recombination between the *nodB* and *nodC* region sample here and any host-specificity genes. This type of recombination within the *sym* region is suggested by the incongruence of *nodB* and *nodC* phylogenies for *Mesorhizobium* (fig. 3) and presents the possibility that host-specificity determinants are unlinked to the regions sequenced here.

### Relationship of *nod* Gene Transfer to Host Range Dynamics

The transfer of niche-adaptive loci may expand the ecological range of bacterial isolates (Eberhard 1990; Young and Levin 1992; Cohan 1994*a*). This model is thought to apply to rhizobial symbiotic genes, since the transfer of *nod* loci may allow strains of a particular chromosomal genotype to associate with distinct host taxa (Young and Wexler 1988; Laguerre et al. 1992, 1996). However, among the isolates sampled here, *nod* gene transfer has not played a clear role in expanding the host range of chromosomal genotypes, since host plant groups did not map better across *nod* phylogenies than across chromosomal phylogenies. Within *Rhizobium* and *Sinorhizobium, nod* gene transfer has apparently occurred among strains nodulating the same host groups; across the *Mesorhizobium,* neither *nod* nor chromosomal loci group rhizobial strains by host plant genus. Similar to the results of Haukka, Lindstrom, and Young (1998), our results suggest that the broad host range of *Mesorhizobium* results from the convergence of genotypes on the same nodulation phenotype, rather than from the transfer of specific *nod* loci.

In summary, the observed patterns of sequence divergence of native rhizobia support two suggested criteria of ''reproductive isolation'' in bacteria (Dykhuizen and Green 1993): (1) similarity of relationships among loci sampled across major subdivisions, as shown for *nod* genes, *16S rRNA,* and *GSII,* and (2) incongruence of phylogenies of different loci within major lineages, as demonstrated for *nodB, nodC,* and *GSII.* The correspondence between these subdivisions and legume taxonomic divisions suggests that host associations may constrain gene flow among native rhizobia and that the broad transfer of symbiotic loci previously observed in agricultural strains may be an exception in native settings. Such broad transfer may be limited to systems which strongly select for novel *sym*-chromosome recombinants through the introduction of host legumes to new regions or new soil habitats (Sullivan et al. 1995).

## Acknowledgments

We thank Dr. Robert Dorit, Dr. Junhyong Kim, and Dr. Theodora Pinou for their suggestions on an early version of this manuscript, and two anonymous referees for their helpful comments. This work was supported by a National Institutes of Health First Award (GM47471) and a National Science Foundation Yong Investigator Award (DEB-9458247) to M.A.R., and a Howard Hughes Medical Institute Predoctoral Fellowship and a National Science Foundation Doctoral Dissertation Improvement Grant (DEB 963537) to J.J.W.

LITERATURE CITED


ARBER, W. 1991. Elements in microbial evolution. J. Mol. Evol. **33**:4–12.

BOTTOMLEY, P. J., H. CHENG, and S. R. STRAIN. 1994. Genetic structure and symbiotic characteristics of a *Bradyrhizobium* poulation recovered from a pasture soil. Appl. Environ. Microbiol. **60**:1754.

BREWIN, N., E. WOOD, and J. YOUNG. 1983. Contribution of the symbiotic plasmid to the competitiveness of *Rhizobium leguminosarum.* J. Gen. Microbiol. **129**:2973–2977.

BRODA, P. 1979. Plasmids. W. H. Freeman, San Francisco.

BROM, S., A. DE LOS SANTOS, M. GIRARD, G. DAVILA, R. PALACIOS, and D. ROMERO. 1991. High-frequency rearrangement of *R. leguminosarum* bv. *phaseoli* plasmids. J. Bacteriol. **173**:1344–1346.

CAMPBELL, A. 1981. Evolutionary significance of accessory DNA elements in bacteria. Annu. Rev. Microbiol. **35**:55–83.

CAUGANT, D. A., L. F. MOCCA, C. E. FRASCH, L. O. FROHOLM, W. D. ZOLLINGER, and R. K. SELANDER. 1987. Genetic structure of *Neisseria meningitidis* populations in relation to serogroup, serotype, and outer membrane protein pattern. J. Bacteriol. **169**:2781–2792.

CHAPPILL, J. A. 1995. Cladistic analysis of the Leguminosae: the development of an explicit phylogenetic hypothesis. Pp. 1–10 *in* M. D. CRISP and J. J. DOYLE, eds. Advances in legume systematics. Part 7: Phylogeny. Royal Botanic Gardens, Kew.

COHAN, F. M. 1994*a*. Genetic exchange and evolutionary divergence in prokaryotes. Trends Ecol. Evol. **9**:175–180.

———. 1994*b*. The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. Am. Nat. **143**:965–86.

———. 1995. Does recombination constrain neutral divergence among bacterial taxa? Evolution **49**:164–175.

CUNNINGHAM, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. **14**: 733–740.

DATTA, N. E. 1985. Plasmids as organisms. Pp. 3–16 *in* S. C. D. HELINSKI, D. CLEWELL, D. JACKSON, and A. HOLAENDER, eds. Plasmids in bacteria. Plenum Press, New York.

DEMEZAS, D. H., T. B. REARDON, J. M. WATSON, and A. H. GIBSON. 1991. Genetic diversity among *Rhizobium leguminosarum* bv. *trifolii* strains revealed by allozyme and restriction fragment length polymorphism analyses. Appl. Environ. Microbiol. **57**:3489–3495.

DOLBERT, R., B. BRELL, and E. TRIPLETT. 1994. DNA sequence of the common nodulation genes of *Bradyrhizobium elkanii* and their phylogenetic relationship to those of other nodulating bacteria. Mol. Plant Microbe Interact. **7**:564–572.

DYKHUIZEN, D. E., and L. GREEN. 1993. Recombination in *E. coli* and the definition of biological species. J. Bacteriol. **173**:7257–7268.

DYKHUIZEN, D. E., and D. L. HARTL. 1983. Selection in chemostats. Microbiol. Rev. **47**:150–168.

EARDLY, B. D., F. S. WANG, and P. VAN BERKUM. 1996. Corresponding *16S rRNA* gene segments in Rhizobiaceae and *Aeromonas* yield discordant phylogenies. Plant Soil **186**: 69–74.

EBERHARD, W. 1990. Evolution in bacterial plasmids and levels of selection. Q. Rev. Biol. **65**:3–22.

EISEN, J. A. 1995. The *RecA* protein as a model molecule for molecular systematic studies of bactier: comparison of trees of *RecA*'s and *16S rRNA*'s from the same species. J. Mol. Evol. **41**:1105–1123.

FARRIS, J. S., M. KALLERSJO, A. G. KLUGE, and C. BULT. 1985. Testing significance of incongruence. Cladistics **10**:315–319.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

———. 1985. Confidence limits on phylogenies with a molecular clock. Syst. Zool. **34**:152–161.

GITTENBERGER, E. 1991. What about non-adaptive radiation? Biol. J. Linn. Soc. **43**:263–272.

HARTL, D. L. 1992. Population genetics of microbial organisms. Curr. Opin. Genet. Dev. **2**:937–942.

HASHEM, F. M., L. D. KUKENDALL, G. EL-FADLY, and T. E. DEVINE. 1997. Strains of *Rhizobium fredii* effectively nodulate and efficiently fix nitrogen with *Medicago sativa* and *Glycine max.* Symbiosis **22**:255–264.

HAUKKA, K., K. LINDSTROM, and J. YOUNG. 1998. Three phylogenetic groups of *nodA* and *nifH* genes in *Sinorhizobium* and *Mesorhizobium* isolates from leguminous trees growing in Africa and Latin America. Appl. Environ. Microbiol. **64**: 419–426.

HIRSCH, P. R. 1996. Population dynamics of indigenous and genetically modified rhizobia in the field. New Phytol. **133**: 159–171.

HOLLOWAY, B. W. 1979. Plasmids that mobilize the bacterial chromosome. Plasmid **2**:1–19.

ISTOCK, C. A., K. E. DUNCAN, N. FERGUSON, and X. ZHOU. 1992. Sexuality in a natural population of bacteria—*Bacillus subtilis* challenges the clonal paradigm. Mol. Ecol. 1: 95–103.

KAPLAN, N. L., R. R. HUDSON, and C. H. LANGLEY. 1989. The ''hitchhiking effect'' revisited. Genetics **123**:887–899.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KISHINEVSKY, B. D., D. SEN, and G. YANG. 1996. Diversity of Rhizobia isolated from various *Hedysarum* species. Plant Soil **186**:21–28.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis. Version 1.01. Pennsylvania State University, University Park.

LAGUERRE, G., M. BARDIN, and N. AMARGER. 1993. Isolation from soil of symbiotic and nonsymbiotic *Rhizobium leguminosarum* by DNA hybridization. Can. J. Microbiol. **39**: 1142–1149.

LAGUERRE, G., E. GENIAUX, S. MARZURIER, R. CASARTELII, and N. AMARGER. 1992. Conformity and diversity among field isolates of *R. leguminosarum* bv. *viceae,* bv. *trifolii,* and bv. *phaseoli* revealed by DNA hybridization using chromosome and plasmid probes. Can. J. Microbiol. **39**: 412–419.

LAGUERRE, G., P. MAVINGUI, M. R. ALLARD, M. P. CHARNAY, P. LOUVRIER, S. I. MAZURIER, L. RIGOTTIER-GOIS, and N. AMARGER. 1996. Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: applications to *Rhizobium leguminosarum* and its different biovars. Appl. Environ. Microbiol. **62**:2029–2036.

LAGUERRE, G., P. VAN BERKUM, N. AMARGER, and D. PREVOST. 1997. Genetic diversity of rhizobial symbionts isolated from legume species within the genera *Astragalus, Oxytropis,* and *Onobrychis.* Appl. Environ. Microbiol. **63**: 4748–4758.

LAWRENCE, J. G., D. L. HARTL, and H. OCHMAN. 1991. Molecular considerations in the evolution of bacterial genes. J. Mol. Evol. **33**:241–250.

LENSKI, R. E. 1993. Assessing the genetic structure of microbial populations. Proc. Natl. Acad. Sci. USA **90**:4344–4336.

LEVIN, B. R. 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics **99**:1–23.

———. 1986. The maintenance of plasmids and transposons in natural populations of bacteria. Pp. 57–70 *in* S. B. LEVY, ed. Antibiotic resistance genes: ecology, transfer, and expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

LONG, S. 1989. *Rhizobium*-legume nodulation: life together in the underground. Cell **56**:203–214.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.0. Sinauer, Sunderland, Mass.

MAIDAK, B. L., G. J. OLSEN, N. LARSEN, R. OVERBEEK, M. J. MCCAUGHEY, and C. R. WOESE. 1997. The RDP (Ribosomal Database Project). Nucleic Acids Res. **25**:109–111.

MARTINEZ, E., D. ROMERO, and R. PALACIOS. 1990. The *Rhizobium* genome. Crit. Rev. Plant Sci. **9**:17–29.

MARTINEZ-ROMERO, E. 1994. Recent developments in Rhizobium taxonomy. Plant Soil **161**:11–20.

MARTINEZ-ROMERO, E., and J. CABALLERO-MELLADO. 1996. Rhizobium phylogenies and bacterial diversity. Crit. Rev. Plant Sci. **15**:113–140.

MAYNARD SMITH, J. 1993. The role of sex in bacterial evolution. J. Hered. **84**:326–327.

MAYNARD SMITH, J., and J. HAIGH. 1974. The hitchhiking effect of a favorable gene. Genet. Res. **23**:23–35.

MAYNARD SMITH, J., N. SMITH, M. O'ROURKE, and B. SPRATT. 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90**:4384–4388.

MAYR, E. 1963. Animal species and evolution. Harvard UniGversity Press, Cambridge, Mass.

MILKMAN, R. 1996. Recombinational exchange among clonal populations. Pp. 2663–2684 *in* F. C. NEIDHARDT, ed. *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D.C.

———. 1997. Recombination and population structure in *E. coli.* Genetics **146**:745–750.

MILKMAN, R., and A. STOLTZFUS. 1988. Molecular evolution of the *Escherichia coli* chromosome. II Clonal segments. Genetics **120**:359–366.

NORMAND, P., and J. BOUSQUET. 1989. Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. J. Mol. Evol. **29**:436–447.

NOUR, S. M., J. C. CLEYET-MAREL, P. NORMAND, and M. P. FERNANDEZ. 1995. Genotypic heterogeneity of strains nodulating chickpeas (*Cicer arietinum* L.) and description of *Rhizobium mediterraneum* sp. nov. Int. J. Syst. Bacteriol. **45**:640–648.

NOVIKOVA, N. I., E. A. PAVLOVA, N. I. VOROBJEV, and E. V. LIMESHCHENKA. 1994. Numerical taxonomy of *Rhizobium* strains isolated from temperate legumes. Int. J. Syst. Bacteriol. **44**:734–742.

OLSEN, G. J., H. MATSUDA, R. HAGSTROM, and R. OVERBEEK. 1994. FastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. CABIOS **10**:41–48.

OLSEN, G. J., and C. J. WOESE. 1993. Ribosomal RNA: a key to phylogeny. FASEB J. **7**:113–123.

O'ROURKE, M., and E. STEVENS. 1993. Genetic structure of *Neisseria gonorrhoeae* populations: a non-clonal pathogen. J. Gen. Microbiol. **139**:2603–2611.

PAGE, R. D. M. 1994. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. Syst. Biol. **43**:58–77.

PARKER, M. 1995. Plant fitness variation caused by different mutualist genotypes. Ecology **76**:1525–1535.

PERRET, X., V. VIPREY, C. FREIBERG, and W. J. BROUGHTON. 1998. Structure and evolution of NGRRS-1, a complex, repeated element in the genome of *Rhizobium* sp. strain NGR234. J. Bacteriol. **179**:7488–7496.

PESOLE, G., M. P. BOZZETTI, C. LANAVE, G. PREPARATA, and C. SACCONE. 1991. Glutamine synthetase gene evolution: a good molecular clock. Proc. Natl. Acad. Sci. USA **88**:522–526.

PREVOST, D., L. M. BORDELEAU, and H. ANTOUN. 1987. Symbiotic effectiveness of indigenous arctic rhizobia on a temperate forage legume: sainfoin (*Onobrychis viciifolia*). Plant Soil **104**:63–67.

PREVOST, D., L. M. BORDELEAU, S. CAUDRY-REZNICK, H. M. SCHULMAN, and H. ANTOUN. 1987. Characteristics of rhizobia isolated from three legumes indigenous to the Canadian high arctic: *Astragalus alpinus, Oxytropis maydelliana* and *Oxytropis arctobia.* Plant Soil **98**:313–324.

RELIC, B., X. PERRET, M. T. ESTRADA-GARCIA, J. KOPCINSKA, W. GOLINOWSKI, H. B. KRISHNAN, S. G. PUEPPKE, and W. J. BROUGHTON. 1994. Nod factors of *Rhizobium* are key to the legume door. Mol. Microbiol. **13**:171–178.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SALYERS, A. A., and N. B. SHOEMAKER. 1994. Broad host range gene transfer: plasmids and conjugative transposons. FEMS Microbiol. Ecol. **15**:15–22.

SANDERSON, M. J., and A. LISTON. 1995. Molecular phylogenetic systematics of *Galgeae,* with special reference to *Astragalus.* Pp. 331–350 *in* M. D. CRISP and J. J. DOYLE, eds. Advances in legume systematics. Vol. 7. Royal Botanical Gardens, Kew.

SCHOFIELD, P., A. GIBSON, W. DUDMAN, and J. WATSON. 1987. Evidence for genetic exchange and recombination of *Rhizobium* symbiotic plasmids in a soil population. Appl. Environ. Microbiol. **53**:2942–2947.

SHARMAN, G., R. CLOSE, and G. MAYNES. 1990. Chromosome evolution, phylogeny, and speciation of rock wallabies. Aust. J. Zool. **37**:351–364.

SIMONSEN, L. 1991. The existence conditions for bacterial plasmids: theory and reality. Microbiol. Ecol. **22**:187–205.

SMALL, E. 1987. Generic changes in Trifolieae subtribe Trigonellinae. Pp. 169–177 *in* C. H. STIRTON, ed. Advances in legume systematics. Vol. 3. Royal Botanic Gardens, Kew.

SOMASEGARAN, P. 1994. Handbook for rhizobia: methods in legume-rhizobium technology. Springer Verlag, New York.

SOUZA, V., L. EGUIARTE, G. AVILA, R. CAPPELLO, C. GALLARDO, F. MONTOYA, and D. PINERO. 1994. Genetic structure of *Rhizobium etli* biovar *phaseoli* associated with wild and cultivated bean plants (*Phaseolus vulgaris* and *Phaseolus coccineus*) in Morelos, Mexico. Appl. Environ. Microbiol. **60**:1260–1268.

SOUZA, V., T. T. NGUYEN, R. R. HUDSON, D. PINERO, and R. E. LENSKI. 1992. Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? Proc. Natl. Acad. Sci. USA **89**:8389–8393.

SPAINK, H. P., C. A. WIJFFELMAN, E. PEES, R. J. H. OKKER, and B. J. J. LUGTENBERG. 1987. *Rhizobium* nodulation gene *nodD* as a determinant of host-specificity. Nature **328**:337–340.

SPRENT, J. I. 1994. Evolution and diversity in the legume-rhizobium symbiosis: chaos theory? Plant Soil **161**:1–10.

SULLIVAN, J. T., B. D. EARDLY, P. VAN BERKUM, and C. W. RONSON. 1996. Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus.* Appl. Environ. Microbiol. **62**:2818–2825.

SULLIVAN, J. T., H. N. PATRICK, W. L. LOWTHER, D. B. SCOTT, and C. W. RONSON. 1995. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc. Natl. Acad. Sci. USA **92**:8985–8989.

SULLIVAN, J. T., and C. W. Ronson. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc. Natl. Acad. Sci. USA **95**: 5145–5149.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Illinois Natural History Survey, Champaign.

TABOADA, H., S. ENCARNACION, M. D. C. VARGAS, Y. MORA, E. MARTINEZ-ROMERO, and J. MORA. 1996. *GSII* constitutes a novel taxonomic marker in *R. etli* and other *Rhizobium* species. Int. J. Syst. Bacteriol. **46**:485–491.

TEMPLETON, A. R. 1989. The meaning of species and speciation: a genetic perspective. Pp. 3–27 *in* D. OTTE and J. A. ENDLER, eds. Speciation and its consequences. Sinauer, Sunderland, Mass.

TEREFEWORD, Z., G. NICK, S. SUOJALAINEN, L. PAULIN, and K. LINDSTROM. 1998. Phylogeny of *Rhizobium galegae* with respect to other rhiozobia and agrobacteria. Int. J. Syst. Bacteriol. **48**:349–356.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

UEDA, T., Y. SUGA, N. YAHIRO, and T. MATSUGUCHI. 1995. Phylogeny of *sym* plasmids of rhizobia by PCR-based sequencing of a *nodC* segment. J. Bacteriol. **177**:468–472.

VAN BERKUM, P., D. BEYENE, G. BAO, T. A. CAMPBELL, and B. D. EARDLY. 1998. *Rhizobium mongolense* sp. nov. is one of three rhizobial genotypes identified which nodulate and form nitrogen-fixing symbioses with *Medicago ruthenica.* Int. J. Syst. Bacteriol. **48**:13–22.

WERNEGREEN, J. J., E. E. HARDING, and M. A. RILEY. 1997. *Rhizobium* gone native: unexpected plasmid stability of indigenous *R. leguminosarum.* Proc. Natl. Acad. Sci. USA **94**: 5483–5488.

WISE, M. G., L. J. SHIMKETS, and J. V. MACARTHUR. 1995. Genetic structure of a lotic population of *Burkholderia cepacia.* Appl. Environ. Microbiol. **61**:1791–1798.

YOUNG, J. P. W. 1996. Phylogeny and taxonomy of rhizobia. Plant Soil **186**:42–52.

YOUNG, J. P. W., H. L. DOWNER, and B. D. EARDLY. 1991. Phylogeny of the phototrophic *Rhizobium* strain BTAI1 by polymerase chain reaction-based sequencing of a 16S ribosomal RNA gene segment. J. Bacteriol. **173**:2271–2277.

YOUNG, J. P. W., and K. HAUKKA. 1996. Diversity and phylogeny of rhizobia. New Phytol. **133**:87–94.

YOUNG, J. P. W., and A. W. B. JOHNSTON. 1989. The evolution of specificity in the legume-rhizobium symbiosis. Trends Ecol. Evol. **4**:341–349.

YOUNG, J., and B. LEVIN. 1992. Pp. 169–191 *in* R. J. BERRY, ed. Genes in ecology: 33rd symposium of the British Ecological Society. Blackwell Scientific Publications, Boston.

YOUNG, J., and M. WEXLER. 1988. *Sym* plasmid and chromosomal genotypes are correlated in field populations of *Rhizobium leguminosarum.* J. Gen. Microbiol. **134**:2731–2739.

ZAWADSKI, P., M. S. ROBERTS, and F. COHAN. 1995. The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. Genetics **140**:917–932.

DAN GRAUR, reviewing editor

Accepted October 6, 1998