Phylogenetic Relationships among *Rhizobium* Species Nodulating the Common Bean (*Phaseolus vulgaris L.*)

**PETER VAN BERKUM,**¹*² DESTA BEYENE,**² and BERTRAND D. EARDLY²*

Soybean and Alfalfa Research Laboratory, Agricultural Research Service, BARC-West, U.S. Department of Agriculture, Beltsville, Maryland 20705; Holetta Agricultural Research Center, Institute of Agricultural Research, Addis Ababa, Ethiopia; and Penn State Berks Campus, Reading, Pennsylvania 19610³

The phylogenetic relationships among *Rhizobium* species that nodulate *Phaseolus vulgaris* (common bean) were determined by directly sequencing the amplified 16S ribosomal DNA genes of these organisms. The bean strains formed four separate clusters. One cluster was composed of *Rhizobium leguminosarum* bv. trifolii, *R. leguminosarum* bv. viciae, and *R. leguminosarum* bv. phaseoli. Two other clusters comprised *Rhzobium etli* and *Rhzobium tropici*, and the fourth cluster contained a single bean-nodulating strain. Data for species identification were obtained from DNA-DNA reassociation experiments. The levels of DNA relatedness among strains belonging to the three biovars of *R. leguminosarum* ranged from 58 to 67%. The levels of DNA relatedness between *R. leguminosarum* and *R. etli* and *R. tropici* ranged from 43 to 45% and 13 to 16%, respectively. The levels of DNA relatedness between the strain belonging to the fourth cluster and strains of the other three *Rhizobium* species that nodulate beans were less than 10%.

*Phaseolus vulgaris* L. (common bean) is an agriculturally important legume crop which benefits from a symbiosis with bacteria belonging to the genus *Rhizobium*. The rhizobia that infect host legumes, such as peas, clovers, and common beans, have been placed in a single species, *Rhizobium leguminosarum*. This species has been subdivided into three biovars largely on the basis of specificity for host plant infection and nodulation (12). Jordan (12) noted that of the three biovars of *R. leguminosarum*, *R. leguminosarum* bv. phaseoli was more distinct than the other two. The reasons for this became apparent when the heterogeneity of strains classified as members of *R. leguminosarum* bv. phaseoli was described (3). Bean isolates obtained from English fields are only weakly polymorphic (26). In contrast, wide phenotypic and genotypic variations have been reported among strains originating in the Americas (16, 17, 19, 20).

Bean strains that originated in Mexico and South America were identified as members of a heterogeneous complex of strongly differentiated phylogenetic lineages, and the data indicated that several species should be recognized (19). One of the deep lineages was subsequently classified as *Rhizobium tropici* (17). On the basis of 16S ribosomal DNA (rDNA) gene sequences, Willems and Collins (25) showed that the phylogenetic position of *R. tropici* was distinct from that of *R. leguminosarum*. However, their analysis was with reference to a clover strain, strain ATCC 14480, because the 16S rDNA sequences of *R. leguminosarum* bv. phaseoli had not been determined and were not available.

Until recently, all bean strains other than *R. tropici* were classified as *R. leguminosarum*. However, there are a number of differences between the nucleotide sequences of a partial analysis of the 16S rDNA genes of bean strains of American origin and the sequence of *R. leguminosarum* bv. viciae type strain ATCC 10004. This finding led to the suggestion that the bean strains of American origin should be referred to as *Rhizobium* sp. type I strains rather than *R. leguminosarum* bv. phascolii (3).

Subsequently, a closer examination of several strains that originated from beans grown in the Americas led to reclassification of *Rhizobium* sp. type I as *Rhizobium etli* (20). This distinction was based primarily on the sequence differences in a 260-bp region of the 16S rDNA genes. The phylogenetic relatedness of the 16S rDNA alleles of *R. etli*, the bean-nodulating *R. leguminosarum* strains, and *R. tropici* has not been determined.

Since *Rhizobium* species that were isolated from many different genera and species of wild legumes nodulate *Phaseolus vulgaris* (1, 26), the phylogenetic relatedness of the different species of bean-nodulating rhizobia should be determined. In this study our objective was to determine the phylogenetic positions of *R. etli* and *R. leguminosarum* bv. phaseoli by using 16S rDNA sequencing and DNA reassociation analyses. We also determined the phylogenetic relatedness of *R. leguminosarum* bv. viciae type strain USDA 2370 (= ATCC 10004) to strains of other biovars of *R. leguminosarum*.

**MATERIALS AND METHODS**

Bacterial strains. The strains used in this study are listed in Table 1. Strains CIAT 899, CFN 299, CFN 42, and USDA 2370 were used because they are the type strains of *R. tropici* type B, *R. tropici* type A, *R. etli*, and *R. leguminosarum*, respectively (3, 7, 17, 20, 23). Clover (*Trifolium pratense*) strain ATCC 14480 was included because it is the only *R. leguminosarum* strain whose complete 16S rDNA sequence has been determined (25). Since bean strains originating in England share enzyme genotypes with *R. leguminosarum* bv. viciae and *R. leguminosarum* bv. trifolii (26), we used two strains from the Rothamsted collection to represent *R. leguminosarum* bv. phascolii. Strain RCR 3644 is a recommended inoculant strain and was isolated from a bean plant growing in a field in Cambridge, England. Strain RCR 3618D is the only other strain available from the Rothamsted collection which originated from a different location. Strain TAL 182 was used as an additional representative of *R. etli* because of genetic differences with strain CNF 42 (T = type strain) (2, 23).

PCR amplification and sequence analysis of the 16S rDNA gene. Colonies of the bacteria grown on the surface of solid modified arabinose-glucuronate medium (22) were placed in 200-μl portions of 0.1% *Tween* 20 (polyoxyethylene sorbitan monolaurate), and the cells were lysed by incubating the suspensions at 95°C for 10 min. Samples (5 μl) of these preparations were used in 120-μl PCR mixtures containing primers fD1 and rD1 (25), which had been synthesized without linker sequences for cloning. The reaction mixtures, which contained 10 pmol of each primer, each deoxynucleoside triphosphate at a concentration of 10 mM, and 100 mM MgCl₂ in 16S rDNA PCR buffer (1% Triton X-100, 50 mM KC1, 10 mM Tris; pH 9.0) were pretreated at 95°C for 3 min before 3 U of Taq DNA polymerase (Promega Corp., Madison, Wis.) was added. The following conditions were used for DNA amplification: 35 cycles consisting of denaturation at

¹ Corresponding author. Mailing address: Soybean and Alfalfa Research Laboratory, Agricultural Research Service, HH-19, Bidg. 011, BARC-West, U.S. Department of Agriculture, Beltsville, MD 20705. Phone: (301) 504-7280. Fax: (301) 504-5728. Electronic mail address: PBERKUM@ASRR.ARUSDA.GOV.

² Copyright © 1996, International Union of Microbiological Societies.
Vol. 46, 1996

**RHIZOBIUM SPECIES NODULATING PHASEOLUS VULGARIS**

94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 2 min, followed by a final extension step for 3 min in a thermocycler (Perkin-Elmer Cetus, Norwalk, Conn.). The PCR products (100 μl) were purified by using QIAquick Spin columns (Qiagen, Inc., Chatsworth, Calif.) and were sequenced by using a model 370A DNA sequencer (Applied Biosystems, Inc., Foster City, Calif.) and a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems). The forward sequencing primers used spanned *Escherichia coli* positions 6 to 25, 339 to 357, 515 to 530, 785 to 805, 1097 to 1114, and 1391 to 1406, and the reverse sequencing primers used spanned positions 1449 to 1453, 1406 to 1392, 1242 to 1226, 1115 to 1100, 926 to 907, 802 to 785, 536 to 519, 358 to 342, and 126 to 110 (24).

**Analysis of the sequence data.** The aligned sequences (6) were analyzed by using the SEQUENCE, DNADIST, FITCH, and CONSENSE programs (4). Trees were constructed by using the programs DRAWTREE and RETREE (4).

**Bacterial growth, DNA extraction, agarose gel electrophoresis, and DNA relatedness analysis.** Cultures were grown in 50-ml portions of modified arabines-gluconate medium to isolate genomic DNAs (22). Genomic DNA was extracted from washed cells and purified by CsCl centrifugation as described previously (18). Samples (10 μg) were digested with restriction endonuclease EcoRI (Beverton Research Laboratories, Gaithersburg, Md.) as recommended by the manufacturer at 37°C for 3 h. The membranes were autoradiographed so that the autoradiogram could be used as a guide to cut the lanes from the membranes. Each lane was immersed in 3 ml of scintillation fluid, and the radioactivity was measured with a scintillation counter (Packard model 2200CA Tri-Carb liquid scintillation analyzer). Each strain was analyzed twice with each probe, and the amount of radioactivity associated with each lane was determined twice. The levels of DNA relatedness between each strain and strains USDA 2370T, RCR 3644, CFN 42T, and RCR 3618D were expressed as percentages of the counts associated with the control lanes.

**RESULTS**

**Nucleotide sequence analysis.** None of the sequences was identical to any other sequence examined (Table 2). The minimum and maximum numbers of nucleotide differences were 3 and 41, which represented similarity values ranging from 99.8 to 97.1%. The 16s rDNA sequences of the two strains of *R. etli* were 99.6% similar. The levels of similarity between the 16s rDNA sequence of strain RCR 3618D and the 16s rDNA sequences of the other seven strains examined ranged from 98.7% with strain ATCC 14480 to 97.1% with strain CFN 299T. The levels of similarity between the strain RCR 3618D 16s rDNA sequence and the *R. etli* TAL 182 and CFN 42T 16s rDNA sequences were 98.2 and 97.9%, respectively.

The five sequences which we determined were aligned with the sequences of 25 reference strains belonging to the α-2 subclass of the *Proteobacteria*. A distance matrix obtained by using the Jukes-Cantor model of nucleotide substitution (13) measured a shuffling count (Packard model 2200CA Tri-Carb liquid scintillation analyzer). Each strain was analyzed twice with each probe, and the amount of radioactivity associated with each lane was determined twice. The levels of DNA relatedness between each strain and strains USDA 2370T, RCR 3644, CFN 42T, and RCR 3618D were expressed as percentages of the counts associated with the control lanes.

**Nucleotide sequence accession numbers.** The 16s rDNA sequences of strains CFN 42T, TAL 182, USDA 2370T, RCR 3618D, and RCR 3644 have been deposited in the GenBank database under accession numbers U28916, U28939, U29386, U29388, and U29388, respectively.

**RESULTS**

**Nucleotide sequence analysis.** None of the sequences was identical to any other sequence examined (Table 2). The minimum and maximum numbers of nucleotide differences were 3 and 41, which represented similarity values ranging from 99.8 to 97.1%. The 16s rDNA sequences of the two strains of *R. etli* were 99.6% similar. The levels of similarity between the 16s rDNA sequence of strain RCR 3618D and the 16s rDNA sequences of the other seven strains examined ranged from 98.7% with strain ATCC 14480 to 97.1% with strain CFN 299T. The levels of similarity between the strain RCR 3618D 16s rDNA sequence and the *R. etli* TAL 182 and CFN 42T 16s rDNA sequences were 98.2 and 97.9%, respectively.

The five sequences which we determined were aligned with the sequences of 25 reference strains belonging to the α-2 subclass of the *Proteobacteria*. A distance matrix obtained by using the Jukes-Cantor model of nucleotide substitution (13)

**TABLE 1. Strains used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Other designation(s)</th>
<th>Geographical origin</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>R. leguminosarum</em> bv. vicie USDA 2370T</td>
<td>ATCC 10004T</td>
<td>Illinois</td>
<td>USDA</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. trifoli ATCC 14480</td>
<td>USDA 2145, LMG 8820</td>
<td>Unknown</td>
<td>USDA</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. phaseoli RCR 3644</td>
<td>USDA 2671</td>
<td>England</td>
<td>RCR</td>
</tr>
<tr>
<td><em>R. etli</em> CFN 42T</td>
<td>USDA 9023T</td>
<td>Mexico</td>
<td>18</td>
</tr>
<tr>
<td><em>R. etli</em> TAL 182</td>
<td>USDA 9041</td>
<td>Hawaii</td>
<td>18</td>
</tr>
<tr>
<td><em>R. tropici</em> CFN 299T</td>
<td>USDA 9020T</td>
<td>Brazil</td>
<td>18</td>
</tr>
<tr>
<td><em>R. tropici</em> CIAT 899T</td>
<td>USDA 9020T</td>
<td>Colombia</td>
<td>18</td>
</tr>
<tr>
<td><em>Rhizobium</em> sp. (<em>Phaseolus</em>) strain RCR 3618D</td>
<td>USDA 2947</td>
<td>Unknown</td>
<td>RCR</td>
</tr>
</tbody>
</table>

* USDA, U.S. Department of Agriculture Research Service National Rhizobium Culture Collection; RCR, Rothamsted Collection of *Rhizobium*.

**TABLE 2. Numbers of nucleotide differences in the aligned 16s rDNA sequences of *R. leguminosarum*, *R. etli*, *R. tropici*, and strain RCR 3618D isolated from *Phaseolus vulgaris***

<table>
<thead>
<tr>
<th>Strain</th>
<th>No. of nucleotide differences*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>R. leguminosarum</em> bv. phaseoli RCR 3644</td>
<td>3</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. trifoli ATCC 14480</td>
<td>3</td>
</tr>
<tr>
<td><em>R. leguminosarum</em> bv. vicie USDA 2370T</td>
<td>3</td>
</tr>
<tr>
<td><em>R. etli</em> TAL 182</td>
<td>12</td>
</tr>
<tr>
<td><em>R. etli</em> CFN 42T</td>
<td>12</td>
</tr>
<tr>
<td><em>R. tropici</em> type B strain C-O5</td>
<td>6</td>
</tr>
<tr>
<td><em>R. tropici</em> type A strain CFN 299T</td>
<td>6</td>
</tr>
<tr>
<td><em>Rhizobium</em> sp. (<em>Phaseolus</em>) strain RCR 3618D</td>
<td>30</td>
</tr>
</tbody>
</table>

* The number of nucleotides in the aligned sequences was 1,533. Gap sites and missing information were removed before the numbers of nucleotide differences were determined.
FIG. 1. Optimal unrooted phylogenetic tree obtained by using the Fitch-Margoliash tree selection criteria when Jukes-Cantor distances were estimated from 16S rDNA sequence data. This tree shows the relationships of R. leguminosarum, R. etli, R. tropici and Rhizobium sp. (Phaseolus) with several related taxa belonging to the α-2 subgroup of the Proteobacteria. The taxa used in this analysis were Bradyrhizobium japonicum (B. japonicum) (GenBank accession number Z35330), Blastobacter denitrificans (Bd. denitrificans) (S46917), Rhodopsseudomonas palustris (Rp. palustris) (D25312), Afipia felis (Af. felis) (M65248), Afipia clevelandensis (Af. clevelandensis) (M69186), Azorhizobium caulinodans (Az. caulinodans) (X67221), Sinorhizobium fredii (Sf. fredii) (X67231), Sinorhizobium meliloti (Sm. meliloti) (X67222), Sinorhizobium saheli (Ss. saheli) (X68390), Sinorhizobium terangii (St. terangii) (X68387), Brucella neotomae (Bn. neotomae) (L26167), Ochrobactrum anthropi (Oo. anthropi) (D12794), Bartonella bacilliformis (Bb. bacilliformis) (M65249), Rochalimaea henselae (Rh. henselae) (M73229), Mycoplasma dimorpha (M. dimorpha) (D12796), Phyllobacterium myrsinacearum (Ph. myrsinacearum) (D12789), Agrobacterium rubi (Ag. rubi) (X67228), Agrobacterium tumefaciens biovar 1 (Ag. tumefaciens bv. 1) (D14500), Agrobacterium vitis (Ag. vitis) (X67225), Rhizobium huakuii (Rh. huakuii) (D13431), Rhizobium loti (Rl. loti) (X67229), Rhizobium tropici (Rt. tropici) (X67233 and X67234), Rhizobium leguminosarum bv. trifolii (Rl. leguminosarum) (X67227), and Rhizobium galegae (R. galegae) (X67226). The fine branches in the box are resolved in Fig. 2.

was used to construct an unrooted tree by the Fitch-Margoliash method (5). The bean-nodulating strains clustered together and along with R. leguminosarum diverged beyond the branch points of three Agrobacterium species and Rhizobium galegae (Fig. 1). The strains representing the three biovars of R. leguminosarum formed a tight cluster (Fig. 2). The strain RCR 3618D branch was located in between the branch points of R. etli and R. tropici (Fig. 2).

DNA relatedness. The levels of DNA relatedness of the three strains of R. leguminosarum ranged from 58 to 67% (Table 3). The levels of DNA relatedness between R. leguminosarum bv. phaseoli RCR 3644 and R. etli and R. tropici were less than 60%. The level of DNA relatedness for the two strains of R. etli was 60%, while the levels of DNA relatedness between R. etli CFN 42T and R. tropici type A and B strains were less than 10%. The levels of DNA relatedness between strain RCR 3618D and the other seven strains were less than 10% (Table 3).

DISCUSSION

A combination of 16S rDNA sequence and DNA relatedness data indicated that Phaseolus vulgaris is nodulated by at
least four different Rhizobium species, including *R. leguminosarum*, *R. etli*, *R. tropici*, and an unnamed species represented by strain RCR 3618D. Although the levels of 16S rDNA sequence similarity among the strains were relatively high, the DNA relatedness data indicated that these organisms belong to separate species.

Young (26) observed the same chromosomal genotypes, as determined by multilocus enzyme electrophoresis, in isolates obtained from peas (*Pisum sativum*), clover (*Trifolium* species), and beans (*Phaseolus vulgaris*). The conclusion that *R. leguminosarum* nodulates beans was supported by the identical nucleotide sequences of a 260-bp fragment of the 16S rRNA gene of strain 8002 and the type strain of *R. leguminosarum* (3). Because the levels of DNA relatedness between bean-nodulating *Rhizobium* strains and *R. leguminosarum* were only 37 to 50%, Jarvis et al. (10) suggested that the bean strains were members of a species separate from *R. leguminosarum*. Subsequently, Crow et al. (1) concluded that the bean strains did not represent a sufficiently distinct population to warrant separation from *R. leguminosarum*, but also recognized the possibility that there are distinct genetic groups that nodulate beans. Both the 16S rRNA sequence data and the DNA relatedness data confirmed that one of these groups is *R. leguminosarum*, which was represented in our study by strain RCR 3618D.

We studied strain TAL 182 in addition to strain CFN 42 because these two strains of *R. etli* produce different electrophoretic patterns following PCR when primers for REP and ERIC sequences are used (23) and their genetic distance is approximately 0.6 as determined from variations in their allele profiles when multilocus enzyme electrophoresis is used (2, 19). Even though these two strains differ significantly, their 16S rDNA nucleotide sequences are similar and a DNA relatedness value of 60% indicates that they belong to the same species.

Of the three bean-nodulating *Rhizobium* species that we analyzed, only *R. leguminosarum* bv. *phaseoli* and *R. tropici* occur in European soils (1). However, French soils may contain two additional bean-nodulating *Rhizobium* species, which have been referred to as species H152 (14) and species R602 (8).

The levels of 16S rRNA nucleotide sequence similarity between RCR 3618D and the other bean strains ranged from 98.4 to 97.1% (22 to 41 base pair differences). These values are comparable to the 16S rDNA nucleotide similarity values obtained for the different *Sinorhizobium* species (99.2 to 97.6%) and the value obtained for *Rhizobium loti* and *Rhizobium huakui* (15) (98.3%). Therefore, the levels of 16S rDNA nucleotide similarity between RCR 3618D and the other bean rhizobia are similar to the levels of 16S rDNA similarity between other root nodule bacteria, indicating that RCR 3618D should be retained in the same genus as *R. leguminosarum*, *R. etli*, and *R. tropici*. However, these interspecies 16S rDNA nucleotide similarity values are greater than 97.0%, the value above which the usefulness of 16S rDNA sequences in species assignment has been questioned and DNA-DNA reassociation data are essential (21). The levels of DNA relatedness values between strain RCR 3618D and the other *Rhizobium* species examined were less than 10%, indicating that RCR 3618D represents a different species since the same or closely related species generally exhibit levels of DNA homology ranging from 60 to 100% (11). However, we decided not to name the species represented by RCR 3618D at this time and to wait until additional isolates are identified and compared with previously described *Rhizobium* species by the criteria outlined by Graham et al. (9).

**ACKNOWLEDGMENTS**

We especially thank William G. Weisburg, Gene-Trak Systems, for his help with choosing the primers used for 16S rDNA gene sequencing. We also thank Perry B. Cregan for synthesizing the primers and K. Lee Nash for expert technical assistance.

**REFERENCES**


