

## Species Limits in *Rhizobium* Populations That Nodulate the Common Bean (*Phaseolus vulgaris*)

BERTRAND D. EARDLY,<sup>1,2\*</sup> FU-SHENG WANG,<sup>2</sup> THOMAS S. WHITTAM,<sup>2</sup>  
AND ROBERT K. SELANDER<sup>2</sup>

*Penn State Berks Campus, Reading, Pennsylvania 19610,<sup>1</sup> and Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania 16802<sup>2</sup>*

Received 9 September 1994/Accepted 21 November 1994

**Evolutionary genetic relationships among 146 bean-nodulating *Rhizobium* strains, including 94 field isolates from three localities in Colombia and 36 strains from Mexico, were examined by multilocus enzyme electrophoresis and restriction fragment length polymorphism analysis of a PCR-amplified 260-bp segment of the 16S rRNA gene. Seventy-five electrophoretic types (ETs), corresponding to multilocus enzyme genotypes, were identified, including a genotypically diverse group of 18 ETs in Colombia that is strongly differentiated from the ETs of *R. etli*, which occur in Mexico, Colombia, and Brazil. Most strains of the distinctive Colombian ETs carried the same 16S rRNA allele as did strains of *R. etli*, but, surprisingly, 17 isolates of two of these ETs had the allele that is characteristic of *R. leguminosarum*, and strains of two other divergent groups of ETs were also polymorphic for the two alleles. No fully satisfactory explanation for the occurrence of the *R. leguminosarum* 16S rRNA allele in three distantly related groups of strains is available, but horizontal transfer and recombination of the gene, in whole or in part, would seem to be more plausible than convergence in nucleotide sequence.**

Bacteria that nodulate the roots of the common bean (*Phaseolus vulgaris*), pea (*Pisum sativum*), and clover (*Trifolium* spp.) have traditionally been classified as biovars of a single species, *Rhizobium leguminosarum*, under the names *R. leguminosarum* bv. phaseoli, viciae, and trifolii, respectively (8, 12). For strains recovered from bean nodules, two subdivisions have been distinguished on the basis of differences in symbiotic (Sym) plasmids and nodulation host range (11, 15). Type I strains carry multiple copies of the plasmid-borne *nifH* gene and have a relatively narrow host range, centering on beans, whereas type II strains have a single copy of the *nifH* gene and are also able to nodulate *Leucaena* spp. Strains of the two types are also reported to differ in the structure of their extracellular polysaccharides (7).

In 1988, a multilocus enzyme electrophoresis (MLEE) study of Mexican and South American strains then assigned to *R. leguminosarum* bv. phaseoli revealed a heterogeneous complex of strongly differentiated phylogenetic lineages and suggested that the recognition of several species would be warranted (16). Subsequently, one particularly deep lineage represented by type II strains was designated as *R. tropici* (15), and a second new species, *R. etli*, was described for type I organisms on the basis of a comparative analysis of MLEE allele profiles, DNA-DNA hybridization experiments (18), and nucleotide sequence variation in the 16S rRNA gene (19).

MLEE has also been employed in recent studies of the structure of Mexican populations of bean-nodulating rhizobia (18, 24, 25), including efforts to assess the frequency of horizontal gene transfer at local and regional levels.

The objective of the present research was to extend the population genetic and systematic analysis of bean-nodulating rhizobia by comparing a sample of field isolates from Colombia with a collection of Mexican and other strains previously stud-

ied by Piñero et al. (16). Chromosomal genotypes were determined by MLEE and by restriction fragment length polymorphism (RFLP) analysis of a 260-bp segment of the 16S rRNA gene (28). The Colombian isolates were further characterized by RFLP analysis of the *nifH* gene, serotyping, and assessment of symbiotic effectiveness.

### MATERIALS AND METHODS

**Sources of strains.** The geographic sources of the 147 isolates examined in this study were as follows: Colombia ( $n = 99$ ); Mexico ( $n = 36$ ); Brazil ( $n = 4$ ); Belize ( $n = 1$ ; OLIVIA-4); England ( $n = 1$ ; CIAT-676); and the United States ( $n = 6$ , including the type strain of *R. leguminosarum*, ATCC 10004<sup>T</sup>, which was isolated from *P. sativum*). Of the 51 isolates previously studied by Piñero et al. (16), 46 were included in the sample; of the 5 strains not examined, 2 (CFN-299 and CIAT-899) were recently reclassified as *R. tropici* (15), 1 (CIAT-895) had a 16S rRNA gene restriction profile identical to that of CFN-299, and 2 (F-20 and CFN-7) were not viable. Among the Mexican isolates examined was CFN-42<sup>T</sup>, the type strain of *R. etli*, from Guanajuato (19).

Included in the Colombian sample were 94 field isolates (provided by J. Kipe-Nolt, Bloomsburg State University, Bloomsburg, Pa.) that were recovered from surface-sterilized nodules of plants (23) grown axenically in soils from field sites at three localities: Palmira ( $n = 34$  isolates, numbers 1085 to 1123), Santander ( $n = 24$ , numbers 1124 to 1143), and Ipiales ( $n = 36$ , numbers 1321 to 1396). At both Palmira and Santander, the elevation is 1,000 m above sea level and the mean annual temperature is 23°C, whereas Ipiales lies at an elevation of 2,900 m and has a mean annual temperature of 12°C.

**MLEE.** Starch gel electrophoresis and enzyme staining have been described by Selander et al. (20). The electrophoretic buffer systems used and the 15 enzymes assayed, all of which are presumed to be encoded by chromosomal genes, were as follows: Tris citrate (pH 6.7) for 6-phosphogluconate dehydrogenase (6PG), hydroxybutyrate dehydrogenase (HBD), and malate dehydrogenase (MDH); Tris-citrate (pH 8.0) for phosphoglucose isomerase (PGI), isocitrate dehydrogenase (IDH), glucose 6-phosphate dehydrogenase (G6P), leucyl-alanine peptidase (PEP), and leucine aminopeptidase (LAP); borate (pH 8.2) for indophenol oxidase (IPO), hexokinase (HEX), adenylate kinase (ADK), and phosphoglucomutase (PGM); and Tris maleate, (pH 7.4) for leucyl-glycyl-glycine peptidase (LGG), phenylalanyl-leucine peptidase (PLP), and nucleoside phosphorylase (NSP).

Distinctive mobility variants (electromorphs) of each enzyme, numbered in order of decreasing anodal mobility, were equated with alleles at the corresponding structural gene locus, and allele profiles (electrophoretic types [ETs]) were equated with multilocus genotypes. Because the occurrence of secondary bands of activity of certain enzymes (IPO and HEX) has been associated with the presence of plasmids in strains of *R. leguminosarum* bv. trifolii (1), such bands were ignored in scoring gels. Of the 15 enzymes assayed in the present study, 12

\* Corresponding author. Mailing address: Penn State Berks Campus, Tulpehocken Road, P.O. Box 7009, Reading, PA 19610-6009. Phone: (610) 320-4832. Fax: (610) 320-4857. Electronic mail address: BDE1@PSUVM.PSU.EDU.

were also analyzed by Piñero et al. (16), but the allele designations used here are not cognate with those of the earlier study.

**Statistical analysis.** Genetic diversity at an enzyme locus was calculated as  $h = n(1 - \sum x_i^2)/(n - 1)$ , where  $x_i^2$  is the frequency of the  $i$ th allele and  $n$  is the number of ETs. The mean genetic diversity per locus ( $H$ ) is the arithmetic average of  $h$  values over the 15 enzyme loci.

The genetic distance between pairs of ETs was estimated as the proportion of loci at which dissimilar alleles occurred (mismatches), and clustering was performed from a matrix of pairwise genetic distances by the unweighted pair group method (22).

**RFLP analysis of a 16S rRNA gene segment.** All isolates were screened for restriction sites in a PCR-amplified 260-bp segment of the 16S rRNA gene. This segment is homologous to the part of the *Escherichia coli* 16S rRNA gene that includes codons 44 through 337 (3, 27). The flanking 24-bp PCR primers and the amplification procedure are described elsewhere (6, 28). Aliquots of the amplified PCR product from each strain were digested with the restriction enzymes *Hin*II, *Bsp*I286I, and *Fok*I, which were chosen on the basis of the published nucleotide sequences of *R. etli* (6), *R. leguminosarum*, and other *Rhizobium* species (26); the resulting fragments were identified electrophoretically in 15% nondenaturing acrylamide gels.

Prior to amplification, the individual template preparations of two representative strains (1103 and 1395) with unusual 16S rRNA alleles and multilocus enzyme ET combinations were serially diluted and digested with diagnostic restriction enzymes in an effort to detect heterology of 16S rRNA sequences. Mixed template preparations were used as controls. Following amplification, the presence of heterologous templates was evident in the control preparations but not in those of the two individual strains.

**RFLP analysis of the *nifH* gene.** The 94 Colombian field isolates were screened for variation in the *Bam*HI RFLP pattern of the *nifH* gene by protocols described previously (6, 11).

**Serotyping.** The serotypes of the Colombian field isolates were determined by direct immunofluorescence. Five non-cross-reactive fluorescent-antibody preparations and their homologous serotype strains (17) were provided by E. L. Schmidt, University of Minnesota, St. Paul, as follows: CIAT-134 (serogroup 1), OLIVIA-4 (serogroup 2a), CIAT-57 (serogroup 2c), CIAT-676 (serogroup 3), and VIKING-1 (serogroup 4). The five antibody preparations were diluted in phosphate-buffered saline (23) to the lowest concentration yielding maximum fluorescence with the homologous strain, i.e., 1:8, 1:4, 1:8, 1:32, and 1:8, respectively. Heat-fixed smears of isolates grown in yeast extract-maltose medium (17) were reacted independently with each of the antibodies, washed with phosphate-buffered saline, and examined under a UV microscope. The intensity of fluorescence was scored on a scale of 1 to 4, and an isolate was assigned to a serogroup if it consistently received a score of 2 or higher upon reaction with a particular antibody preparation. Several strains received scores of 2 or higher with two antisera (for example, strain 1131 was serotyped as 1, 2a), and some isolates failed to react with any of the antibodies and hence were nontypeable.

**Symbiotic effectiveness.** The relative abilities of the Colombian field isolates to fix nitrogen were assessed by comparing the shoot dry weights of common bean plants (cultivar BAT 76) inoculated with individual strains and the shoot dry weights of control plants inoculated with the effective strain CIAT-144 (6).

## RESULTS

**Genetic diversity.** In the MLEE analysis of the 147 strains, all 15 enzyme loci assayed were polymorphic, with an average of 7.7 alleles per locus. A total of 75 distinctive ETs, representing multilocus genotypes, were identified; among these, the mean genetic diversity per locus ( $H$ ) was 0.669. Allele profiles for 40 of the ETs are shown in Table 1, and data for all ETs will be provided, upon request, by B.D.E. Twenty-four of the ETs were represented by more than one isolate, with a maximum of 16 isolates for ET 9, of which Colombian strain 1322 is representative.

**Relationships of ETs.** Cluster analysis placed the 75 ETs in five major groups (labelled A to E in Fig. 1), corresponding to lineages that diverged at genetic distances greater than 0.60. Group A was composed of 18 ETs that formed two clusters (labelled A<sup>1</sup> and A<sup>2</sup>), all isolates of which were from Colombia. Group B consisted of ETs 19, 21, and 22, which were represented by Mexican strains, CFN-3, F-4, and CFN-215, respectively, together with ET20, which was represented by NITRAGIN-8251, a strain of unknown geographic source. Group C consisted of six ETs, with the following strain assignments: ET 23, strain F-6 from Mexico; ET 24, strain F-9 from Mexico; ET-25, strains 1147 and 1149, both from Colombia, and strain TAL-182 from Hawaii; ET 26, strain 1103 from Colombia; ET

27, strain ATCC 10004<sup>T</sup> from the United States and strain CIAT-676 from England; and ET 28, strain 1350 from Colombia. The 42 ETs in group D were represented by approximately equal numbers of strains from Colombia and Mexico, including the type strain (CFN-42<sup>T</sup>) of *R. etli* (ET 63). Finally, group E, which was the most divergent of the five groups, consisted of five ETs, all of which were represented exclusively by strains from Mexico. Five of the strains were collected from soil in a field growing *Leucaena leucocephala*. Because these strains are inefficient in fixing nitrogen on the common bean and have distinctive Sym plasmids (16), they have been regarded as representing an unnamed species distinct from *R. etli*, *R. tropici*, and *R. leguminosarum* (10, 16).

For the strains examined earlier by MLEE by Piñero et al. (16) that were also included in the present study, the indicated relationships were generally similar, even though only 12 of the 15 enzymes assayed in each study were the same. However, there was a notable difference between the studies in the placement of strains F-6, F-9, and TAL-182 (one of three isolates associated with ET 25). In the present analysis, these strains formed a cluster (ET 23 to ET 25) in group C, whereas Piñero et al. (16) placed them in a cluster allied with CFN-42, VIKING-1, and DES-109, all of which fall in our group D (Fig. 1).

**Distribution of 16S rRNA alleles.** Variation in the nucleotide sequence of the 16S rRNA gene demonstrated by Eardly et al. (6), Segovia et al. (19), and Willems and Collins (26) was indexed by an RFLP analysis of a 260-bp segment of the gene (Table 2; Fig. 2). For purposes of the present study, the important point is that the nucleotide sequence of this segment in isolates of *R. etli*, including the type strain, CFN-42<sup>T</sup> (ET 63), differs from that in *R. leguminosarum*, as represented by ATCC 10004<sup>T</sup> (ET 27), at nine sites, five of which were sampled by the three restriction enzymes used in the RFLP survey. Complete sequencing of the 260-bp segment in four strains with the *R. etli* RFLP pattern (1342, CFN-3, 1147, and CFN-42<sup>T</sup>), and six strains with the *R. leguminosarum* RFLP pattern (1322, 1395, NITRAGIN-8251, F-4, CFN-215, and 1350) identified all nine diagnostic sites but revealed no additional polymorphic nucleotides (4).

The results shown in Fig. 1 demonstrate a lack of correspondence between 16S rRNA gene sequences and the relationships among strains estimated by MLEE. Most of the strains in groups A through D (118 of 141) carried the 16S rRNA allele that is characteristic of *R. etli* and is hereinafter designated as the *e* allele. It was present in almost all the Mexican isolates and in most of the Colombian field isolates, including all of those from Palmira and Santander and half of those from Ipiales. As shown in Fig. 1 (bold lines), the *l* (for *leguminosarum*) allele of the 16S rRNA gene was present in 23 strains associated with seven ETs, as follows: group A, strain 1322 and the other 15 isolates of ET 9 (in subgroup A<sup>1</sup>), all from Ipiales, and 1395 (ET 10, also in group A<sup>1</sup>), also from Ipiales; group B, NITRAGIN-8251 (ET 20), F-4 (ET 21), and CFN-215 (ET 22); group C, ATCC 10004<sup>T</sup> (ET 27), CIAT-676 (ET 27), and 1350 (ET 28).

In defining *R. etli*, Segovia et al. (19) sequenced the 16S rRNA gene in five strains representing the three main lineages of *R. leguminosarum* bv. phaseoli that had earlier been identified by Piñero et al. (16). Whereas Segovia et al. (19) reported that NITRAGIN-8251 has the *e* 16S rRNA allele, we found that this strain (grown from the freezer stock established by Piñero et al. [16]) actually has the *l* allele, as does strain F-4, which is related to NITRAGIN-8251 as shown by MLEE (Fig. 1; Table 1).

In our analysis of the 16S rRNA gene, we assayed variation at four restriction sites that were thought to provide an RFLP

TABLE 1. Allele profiles of 40 ETs

ET	Reference strain	Allele at indicated enzyme locus <sup>a</sup>														
		MDH	IDH	IPO	HBD	HEX	PGI	PEP	PLP	ADK	PGM	LGG	LAP	NSP	G6P	6PG
Group A																
1	1342	2	2	2	8	2	5	2	5	4	4	8	5	8	6	1
2	1343	2	2	2	8	2	5	3	4	4	4	8	5	8	6	1
3	1344	2	2	2	8	2	5	1	1	4	4	8	5	8	6	1
4	1369	2	2	2	8	2	5	5	6	4	4	10	5	7	6	1
5	1391	2	2	2	8	2	5	5	6	4	4	8	5	7	6	1
6	1362	2	2	2	8	2	2	5	6	4	4	10	5	6	6	1
7	1345	2	2	2	8	2	5	5	6	4	4	6	5	8	6	5
8	1131	2	2	2	9	2	6	5	6	4	2	10	3	2	6	1
9	1322 <sup>b</sup>	2	2	2	4	2	5	1	1	4	4	6	5	7	5	2
10	1395 <sup>b</sup>	1	2	2	4	2	5	1	1	4	4	6	5	7	5	2
11	1135	2	2	2	9	2	8	3	2	4	2	8	2	4	6	4
12	1144	2	2	2	5	2	8	2	2	4	2	6	2	4	6	4
13	1139	2	2	2	9	6	8	2	2	4	2	6	2	5	6	4
14	1148	2	5	2	9	2	8	1	1	4	2	8	2	4	6	4
15	1152	2	2	2	9	2	7	5	6	4	2	4	2	4	6	4
16	CIAT-893	2	2	2	6	2	8	2	2	4	2	4	2	3	6	6
17	1128	2	2	2	9	2	6	2	2	4	2	4	5	4	6	2
18	CIAT-134	2	5	2	9	3	7	2	2	4	2	10	5	2	6	6
Group B																
19	CFN-3	2	3	1	2	2	8	3	3	4	2	9	5	2	6	1
20	NITRAGIN-8251 <sup>b</sup>	2	2	2	2	2	8	2	2	5	3	1	5	3	5	7
21	F4 <sup>b</sup>	1	2	2	6	2	6	2	2	4	3	8	2	1	5	1
22	CFN-215 <sup>b</sup>	2	4	2	1	5	5	4	6	4	4	7	6	6	6	4
Group C																
23	F6	3	5	2	4	1	6	2	3	4	2	4	8	10	5	6
24	F9	3	5	2	4	1	6	2	3	4	2	4	7	10	5	6
25	1147	3	5	2	4	1	6	5	7	6	2	6	8	9	5	6
26	1103	3	5	2	0	1	3	4	4	4	2	8	6	8	5	6
27	ATCC-10004 <sup>Tb</sup>	2	4	2	7	1	3	4	6	4	2	6	8	5	5	6
28	1350 <sup>b</sup>	2	4	2	6	1	2	5	7	4	2	8	6	0	5	4
Group D																
31	OLIVIA-4	3	1	2	4	2	6	5	7	5	2	6	2	5	4	2
34	1145	3	1	2	4	4	6	5	7	5	2	3	5	5	4	1
40	FL-69	3	1	2	4	1	3	5	7	5	2	3	2	1	4	1
45	F-13	4	1	2	4	4	5	5	7	5	2	3	4	4	4	2
50	F-2	3	1	2	4	2	8	5	6	1	2	1	5	4	5	1
56	1396	3	5	2	4	2	2	5	6	5	2	2	5	4	0	6
61	DES-109	3	5	2	4	2	6	2	2	1	2	4	6	1	6	1
63	CFN-42 <sup>T</sup>	3	5	2	0	2	5	2	3	1	1	1	6	7	4	5
65	SIL-4	3	5	2	0	2	5	4	4	3	2	4	6	3	3	1
66	BRA-281	4	5	2	4	2	6	4	4	2	2	4	6	5	4	1
Group E																
71	FL-44	3	5	3	3	6	2	8	9	5	2	6	10	0	2	4
75	FL-27	4	7	3	9	2	1	7	7	3	1	8	9	6	2	7

<sup>a</sup> The enzymes assayed are listed in Materials and Methods.

<sup>b</sup> Strains with the *l* 16S rRNA allele.

pattern diagnostic of *R. etli*. Although strains of group E also exhibited the common *R. etli* RFLP pattern, sequencing of the 260-bp segment of the gene in strain FL-27 has shown that it differs from that of the *R. etli* type strain, CFN-42<sup>T</sup>, at two nucleotide sites (6). Also, Laguerre et al. (9) have identified additional nucleotide differences in other parts of the gene between CFN-42<sup>T</sup> and a strain that is genotypically similar if not identical to FL-27.

**Association of characters.** Among the Colombian field isolates, there was no relationship between the occurrence of the *e* and *l* 16S rRNA alleles, *nifH* RFLP types, or symbiotic effectiveness. For example, the mean symbiotic effectiveness of

the 18 Ipiales isolates with the *e* 16S rRNA allele was 78.1% ± 3.3% of that of the effectively nodulated controls, while the corresponding value for the 18 isolates with the *l* allele from the same locality was 80.8% ± 4.7%. Of the 94 Colombian field isolates, only 5 (1088, 1115, 1157, 1107, and 1097), all of which represent ETs of group D, did not have a reiterated (type I) *nifH* RFLP pattern; these isolates were also symbiotically ineffective. With the exception of ET 38 (isolate 1107), the ETs of these strains were also represented by other isolates that showed reiterated *nifH* RFLP patterns typical of *R. etli* and were symbiotically effective.

There were nonrandom associations between serotype and

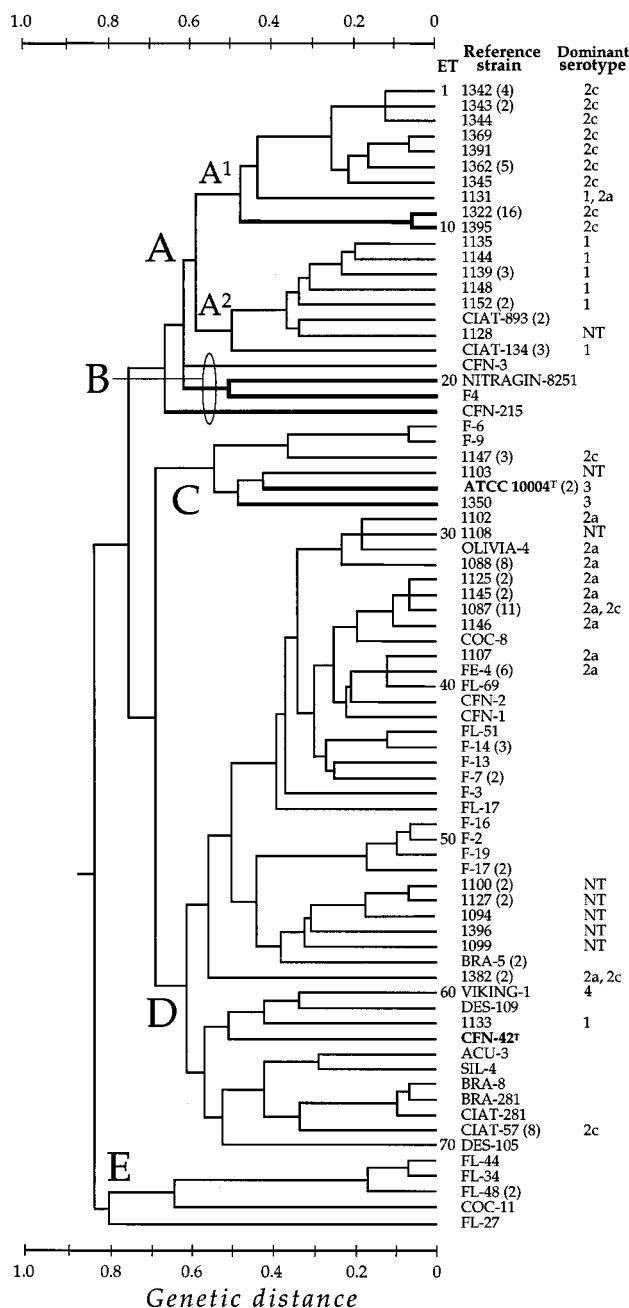


FIG. 1. Evolutionary tree for 75 ETs of *Rhizobium* spp. constructed by the unweighted pair group method from a matrix of pairwise genetic distances (allele mismatches) based on an MLEE analysis of 15 loci. A representative strain is listed for each ET; and for ETs with more than one isolate, the number is indicated in parentheses. Dominant serotypes of ETs represented by Colombian field isolates are also indicated. Strains from Colombia are denoted by four-digit numbers or the prefix CIAT (with the exception of CIAT-57, which is from the United States), and isolates from Brazil have the prefix BRA. Other strains are from Mexico, with the following exceptions: OLIVIA-4 (Belize), VIKING-1 (United States), ATCC 10004<sup>T</sup> (United States), and NITRAGIN-8251 (geographic origin unknown). With the exception of group E, fine-lined branches denote lineages represented by strains of *R. etli* and bold-lined branches indicate lineages of ETs represented by strains carrying 16S rRNA alleles with the *R. leguminosarum* RFLP pattern. All ETs except ET 22, ET 27, ET 38, and ET 71 through ET 75 are represented by strains of type 1 (16), as defined by criteria discussed by Martínez et al. (11) and Brom et al. (2). Strains associated with ET 22 and ET 27 were not analyzed for their *nifH* RFLP pattern.

TABLE 2. Presence or absence of four restriction sites in a 260-bp segment of the 16S rRNA gene

Species	Reference strain	Presence or absence of restriction site position (enzyme) <sup>a</sup>			
		89 ( <i>Hinf</i> I)	142 ( <i>Bsp</i> I286I)	173 ( <i>Fok</i> I)	177 ( <i>Bsp</i> I286I)
<i>R. etli</i>	CFN-42 <sup>T</sup>	-	+	-	-
	CIAT-57 <sup>b</sup>	-	-	-	-
<i>R. leguminosarum</i>	ATCC 10004 <sup>T</sup>	+	-	+	+
<i>R. tropici</i>	CIAT-899 <sup>T</sup>	+	-	-	+
	CFN-299 <sup>c</sup>	+	-	+	+
<i>R. meliloti</i>	Division A	+	+	+	+
	Division B	+	-	+	+
<i>R. galegae</i>	ATCC 43677 <sup>T</sup>	+	-	+	+
<i>R. loti</i>	ATCC 33669 <sup>T</sup>	+	-	+	+
<i>R. huakuii</i>	IAM 14158 <sup>T</sup>	+	-	+	+

<sup>a</sup> Sites are numbered according to their positions in the 16S rRNA segment shown in Fig. 2. The presence or absence of a restriction site is denoted by + or -, respectively.

<sup>b</sup> CIAT-57 and strain 127K12b, both of ET 69 (Fig. 1), were the only strains with this profile. Six other ET 69 strains carried the common allele of *R. etli*, as represented by strain CFN-42<sup>T</sup>.

<sup>c</sup> Only two strains (CFN-299 and CIAT-895) showed this profile.

chromosomal genotype. As shown in Fig. 1, serotype 2c was predominant among strains of subgroup A<sup>1</sup>, whereas most isolates of subgroup A<sup>2</sup> were of serotype 1. In group D, 2a was the common serotype of strains with ET 29 to ET 39, and the strains with ET 53 to ET 57 were uniformly nontypeable.

## DISCUSSION

Rhizobial strains in the Americas that have the common bean as their primary host are currently believed to represent two species, *R. etli* and *R. tropici* (14, 19, 24). However, our study identified an additional complex of lineages (group A) in Colombia that is roughly equal in phylogenetic depth and multilocus genotypic diversity to the populations named *R. etli*, which occur in Mexico, Colombia, and Brazil. Most strains with the group A ETs carry the same *e* 16S rRNA allele as *R. etli*, but, surprisingly, some strains (associated with ET 9 and ET 10) have the *l* allele of *R. leguminosarum*. Both *e* and *l* alleles are also represented in strains of groups B and C. The 17 strains of ETs 9 and 10 were not distinguishable from strains that carry the *e* allele in either symbiotic effectiveness or RFLP pattern of the *nifH* gene. However, like strains of *R. etli*, they have reiterated *nifH* genes and are therefore of type I.

The results of MLEE analysis alone would suggest the recognition of group A strains as a species distinct from *R. etli* (group D) and *R. leguminosarum* (group C) and the naming of several additional species for the deep lineages of group B. However, the distribution of 16S rRNA alleles does not fit such a scheme. The most puzzling aspect of the observed pattern of variation is the occurrence of both *l* allele strains and *e* allele strains in subgroup A<sup>1</sup> (ET 1 to ET 10), together with the association of *e* allele strain 1103 (ET 26) with *l* allele strains ATCC 10004<sup>T</sup>, CIAT-676 (both ET 27), and 1350 (ET 28) in group C. MLEE analysis indicates that the *l* allele strains in groups A, B, and C are only distantly related in chromosomal genotype. For example, group A strain 1395 (ET 10) and group B strain NITRAGIN-8251 (ET 20) differ from group C strain ATCC 10004<sup>T</sup> (ET 27) at 11 and 12 of the 15 loci assayed, respectively (Table 1).

The central question raised by the results of our study is why

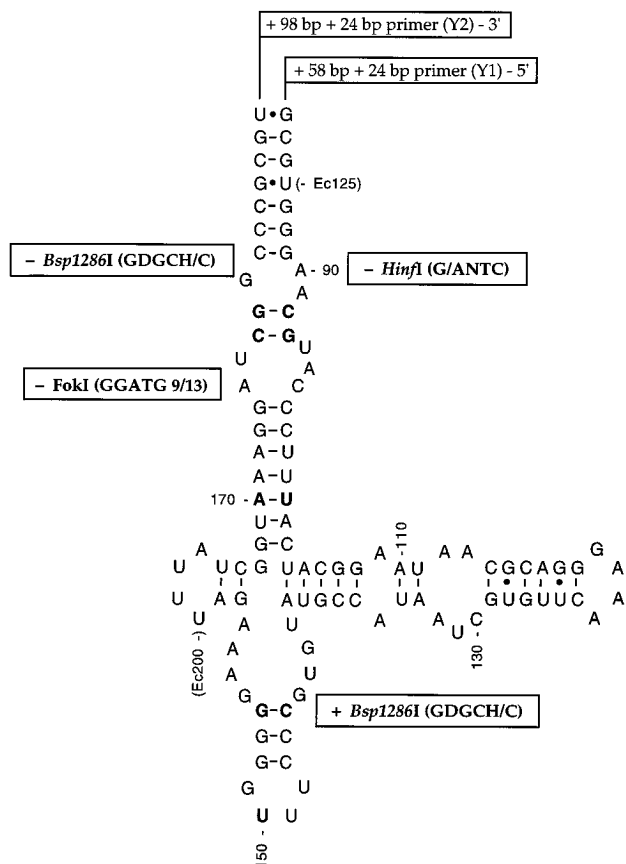


FIG. 2. Predicted secondary structure of part of the 16S rRNA molecule encoded by a 308-bp DNA segment of the 16S rRNA gene amplified from *R. etli* CFN-42<sup>T</sup> (derived from data of Young et al. [28]). The small numbers indicate nucleotide positions in the amplified segment, while those prefixed by Ec correspond to positions in the *E. coli* sequence (3, 27). The nine nucleotides at which the sequence of *R. leguminosarum* ATCC 10004<sup>T</sup> differs from that of *R. etli* (6) are shown in boldface type. The location of primers and the presence (+) or absence (-) of restriction sites at positions 89, 142, 173, and 177 in the amplified DNA segment are indicated in boxes.

MLEE analysis did not place all the strains carrying the *l* 16S rRNA allele in a single cluster. The simplest interpretation is that the *l* allele, in whole or in part, has been horizontally transferred and recombined among distantly related chromosomal lineages. Our invocation of recombination of 16S rRNA genes is not without precedent, for Sneath (21) has presented strong statistical evidence for the occurrence of intragenic recombination between *Aeromonas* species, thereby accounting for the marked disagreement between estimates of phylogenetic relationships based on 16S rRNA gene sequences and the results of DNA-DNA hybridization experiments reported by Martinez-Murcia et al. (13). Evidence for the recombination of a horizontally transferred segment of the 16S rRNA gene that is homologous to the segments examined in our study and that of Sneath (21) has been obtained through cluster analysis of strains of *R. galegae*, *R. loti*, and *Agrobacterium tumefaciens* (10a). Also relevant is the observation that 16S rRNA allele segments of both the *e* and *l* types occur in strains presently assigned to *R. leguminosarum* bv. *trifolii* (5) that diverge at a MLEE genetic distance of 0.55 (4).

The recombination hypothesis could be tested by a complete sequence analysis of the 16S rRNA genes and their flanking regions in strains of groups A, B, and C. Assortative (whole-

gene) or intragenic recombination of the enzyme loci assayed by MLEE is a less satisfactory explanation because of the number of genes involved, and convergence in ET would also seem unlikely for the same reason. Convergence in the 16S rRNA gene sequence, although possible, was not indicated in a comparative sequence analysis of a segment of the *e* and *l* alleles in 10 representative strains.

It is apparent that additional molecular genetic evidence will be required to clarify the evolutionary relationships of *Rhizobium* strains that nodulate the common bean and to determine the extent to which recombination of horizontally transferred chromosomal genes or gene segments have been involved in their evolution. Future work should include a comprehensive study of allelic variation in the 16S rRNA gene and other chromosomal genes, preferably by complete sequence determination, among strains of all lineages of rhizobia that are capable of nodulating beans, as well as those of lineages currently classified as *R. leguminosarum* bv. *trifolii* and *viciae*.

#### ACKNOWLEDGMENTS

Strains were furnished by J. Kipe-Nolt, E. L. Schmidt, R. S. Smith, and R. Wheatcroft. We also thank A. R. J. Eaglesham for the use of laboratory and greenhouse facilities and E. L. Schmidt for provision of fluorescent antibodies.

This research was supported in part by NIH grant AI-22144.

#### REFERENCES

- Baldani, J. I., R. W. Weaver, M. F. Hynes, and B. D. Eardly. 1992. Utilization of carbon substrates, electrophoretic enzyme patterns, and symbiotic performance of plasmid-cured clover rhizobia. *Appl. Environ. Microbiol.* **58**:2308-2314.
- Brom, S., E. Martínez, G. Dávila, and R. Palacios. 1988. Narrow- and broad-host-range symbiotic plasmids of *Rhizobium* spp. strains that nodulate *Phaseolus vulgaris*. *Appl. Environ. Microbiol.* **54**:1280-1283.
- Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller. 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Mol. Biol.* **148**:107-127.
- Eardly, B. D. Unpublished data.
- Eardly, B. D. 1993. Restriction fragment length polymorphism in a 16S rRNA gene segment in *Rhizobium* isolated from 38 species of *Trifolium*, p. 610. In R. Palacios, J. Mora, and W. E. Newton (ed.), *New horizons in nitrogen fixation*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Eardly, B. D., J. P. W. Young, and R. K. Selander. 1992. Phylogenetic position of *Rhizobium* sp. strain Or 191, a symbiont of both *Medicago sativa* and *Phaseolus vulgaris*, based on partial sequences of the 16S rRNA and *nifH* genes. *Appl. Environ. Microbiol.* **58**:1809-1815.
- Gil-Serrano, A., A. Sánchez del Junco, P. Tejero-Mateo, M. Megias, and M. A. Caviedes. 1990. Structure of the extracellular polysaccharide secreted by *Rhizobium leguminosarum* var. *phaseoli* CIAT 899. *Carbohydr. Res.* **204**:103-107.
- Jordan, D. C. 1984. Family III. *Rhizobiaceae*. Genus I. *Rhizobium* Frank 1889, 338<sup>AL</sup>, p. 235-242. In N. R. Krieg and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 1. The Williams & Wilkins, Co., Baltimore.
- Laguette, G., M. Allard, F. Revoy, and N. Amarger. 1994. Rapid identification of rhizobia by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl. Environ. Microbiol.* **60**:56-63.
- Laguette, G., M. P. Fernandez, V. Edel, P. Normand, and N. Amarger. 1993. Genomic heterogeneity among French *Rhizobium* strains isolated from *Phaseolus vulgaris*. *L. Int. J. Syst. Bacteriol.* **43**:761-767.
- Lindström, K. Personal communication.
- Martínez, E., M. A. Pardo, R. Palacios, and M. A. Cevallos. 1985. Reiteration of nitrogen fixation gene sequences and specificity of *Rhizobium* in nodulation and nitrogen fixation in *Phaseolus vulgaris*. *J. Gen. Microbiol.* **131**:1779-1786.
- Martínez, E., D. Romero, and R. Palacios. 1990. The *Rhizobium* genome. *Crit. Rev. Plant Sci.* **9**:59-93.
- Martínez-Murcia, A. J., S. Benlloch, and M. D. Collins. 1992. Phylogenetic interrelationships of the genera *Aeromonas* and *Plesiomonas* as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. *Int. J. Syst. Bacteriol.* **42**:412-421.
- Martínez-Romero, E. 1994. Recent developments in *Rhizobium* taxonomy. *Plant Soil* **161**:11-20.
- Martínez-Romero, E., L. Segovia, F. M. Mercante, A. A. Franco, P. Graham, and M. A. Pardo. 1991. *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int. J. Syst. Bacteriol.* **41**:417-426.

16. Piñero, D., E. Martínez, and R. K. Selander. 1988. Genetic diversity and relationships among isolates of *Rhizobium leguminosarum* biovar *phaseoli*. *Appl. Environ. Microbiol.* **54**:2825–2832.
17. Robert, F. M., and E. L. Schmidt. 1985. Somatic serogroups among 55 strains of *Rhizobium phaseoli*. *Can. J. Microbiol.* **31**:519–523.
18. Segovia, L., D. Piñero, R. Palacios, and E. Martínez-Romero. 1991. Genetic structure of a soil population of nonsymbiotic *Rhizobium leguminosarum*. *Appl. Environ. Microbiol.* **57**:426–433.
19. Segovia, L., J. P. W. Young, and E. Martínez-Romero. 1993. Reclassification of American *Rhizobium leguminosarum* biovar *phaseoli* type I strains as *Rhizobium etli* sp. nov. *Int. J. Syst. Bacteriol.* **43**:374–377.
20. Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl. Environ. Microbiol.* **51**:873–884.
21. Sneath, P. H. A. 1993. Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. *Int. J. Syst. Bacteriol.* **43**:626–629.
22. Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco.
23. Somasegaran, P., and H. J. Hoben (ed.). 1985. Methods in legume-*Rhizobium* technology. NifTAL, University of Hawaii, Paia, Hawaii.
24. Souza, V., L. Eguiarte, G. Avila, R. Cappello, C. Gallardo, J. Montoya, and D. Piñero. 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.
25. Souza, V., T. T. Nguyen, R. R. Hudson, D. Piñero, and R. E. Lenski. 1992. Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? *Proc. Natl. Acad. Sci. USA* **89**:8389–8393.
26. Willems, A., and M. D. Collins. 1993. Phylogenetic analysis of rhizobia and agrobacteria based on 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* **43**:305–313.
27. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
28. 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 rRNA gene segment. *J. Bacteriol.* **173**:2271–2277.