APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Nov. 1988, p. 2825–2832
0099-2240/88/112825-08$02.00/0

Vol. 54, No. 11

Genetic Diversity and Relationships among Isolates of *Rhizobium leguminosarum* biovar *phaseoli*

DANIEL PINERO,[1] ESPERANZA MARTINEZ,[2] AND ROBERT K. SELANDER[3]*

Centro de Ecología, Universidad Nacional Autónoma de México, Mexico City 04510,[1] and Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos,[2] Mexico, and Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802[3]

Received 10 June 1988/Accepted 15 August 1988

Fifty-one isolates of *Rhizobium leguminosarum* biovar *phaseoli* from various geographic and ecological sources, largely in Mexico, were characterized by the electrophoretic mobilities of 15 metabolic enzymes, and 46 distinctive multilocus genotypes (electrophoretic types [ETs]) were distinguished on the basis of allele profiles at the enzyme loci. Mean genetic diversity per enzyme locus among the 46 ETs was 0.691, the highest value yet recorded for any species of bacterium. The occurrence of strong nonrandom associations of alleles over loci suggested a basically clonal population structure, reflecting infrequent recombination of chromosomal genes. Multilocus genotypic diversity was unusually high, with the most strongly differentiated pairs of ETs having distinctive alleles at all 15 loci and major clusters of ETs diverging at genetic distances as large as 0.89. This great diversity in the chromosomal genome raises the possibility that *R. leguminosarum* biovar *phaseoli* is a polyphyletic assemblage of strains. As other workers have suggested, the inclusion of all strains capable of nodulating beans in a single biovar or species is genetically unrealistic and taxonomically misleading. A biologically meaningful classification of *Rhizobium* spp. should be based on assessment of variation in the chromosomal genome rather than on phenotypic characters, especially those mediated for the most part or wholly by plasmid-borne genes, such as host relationships.

In a recent taxonomic revision of the nitrogen-fixing bacteria of the genus *Rhizobium*, the nodulating symbionts of lucerne (*Medicago sativa*) were considered to constitute a distinct species, *Rhizobium meliloti*, but the formerly recognized species *Rhizobium leguminosarum* (symbionts of peas [*Pisum sativum*]), *Rhizobium trifolii* (symbionts of clover [*Trifolium* sp.]), and *Rhizobium phaseoli* (symbionts of common beans [*Phaseolus vulgaris*]) were reduced to the status of biovars of a single species and designated *R. leguminosarum* bv. *viceae*, *R. leguminosarum* bv. *trifolii*, and *R. leguminosarum* bv. *phaseoli*, respectively (19). The primary bases for the new classification were as follows: (i) the symbionts of peas, clover, and beans cannot reliably be distinguished on the basis of phenotypic characters, whereas those of lucerne have several distinctive features (12, 25); (ii) under laboratory conditions, conjugationally mediated recombination of chromosomal genes can readily be obtained among strains of *R. leguminosarum* bv. *viceae*, *R. leguminosarum* bv. *trifolii*, and *R. leguminosarum* bv. *phaseoli*, but strains of *R. meliloti* do not readily exchange chromosomal genes with strains of the other symbionts (17, 20); (iii) the genetic determinants of host range and other symbiotic functions have proven to be largely, if not entirely, plasmid borne and to remain functional when transferred among strains from peas, clover, and beans (15, 18, 21); and (iv) on average, there is a much higher level of similarity of the nucleotide sequence of total DNA among strains of the three biovars of *R. leguminosarum* than between these biovars and either *R. meliloti* or *Rhizobium loti* (7).

Several lines of evidence arguing against the new taxonomic arrangement may be cited. First, DNA hybridization experiments have demonstrated marked heterogeneity of chromosomal nucleotide sequence among strains of each of the three designated biovars of *R. leguminosarum* (7). Second, two-dimensional polyacrylamide gel electrophoresis of total cellular proteins failed to differentiate between isolates of *R. leguminosarum* bv. *viceae* and *R. leguminosarum* bv. *trifolii* but showed that isolates of *R. leguminosarum* bv. *phaseoli* are distinctive and markedly diverse genetically (28). Native isolates recovered from bean nodules also vary in their nodulation host range and in nitrogenase (*nif*) gene copy number (24, 27). Consistent with the data from two-dimensional polyacrylamide gel electrophoresis of total proteins, Young (35) observed considerable sharing of electrophoretic mobility variants of three metabolic enzymes between isolates of *R. leguminosarum* bv. *viceae* and *R. leguminosarum* bv. *trifolii* collected in fields in England. However, he also found that isolates recovered from beans and assigned to *R. leguminosarum* bv. *phaseoli* were only weakly polymorphic and shared enzyme genotypes with strains of the other two biovars. Finally, it has been shown that *Rhizobium* strains from several DNA homology groups are able to nodulate beans (7); indeed, diverse strains of *Rhizobium* spp. from nodules on many different genera and species of wild legumes are capable of nodulating beans (22, 24).

In the present study, we used multilocus enzyme electrophoresis (30, 31) to assess allelic variation at 15 metabolic enzyme loci in a collection of 51 isolates of *R. leguminosarum* bv. *phaseoli* obtained from a variety of geographic and ecological sources, mostly in Mexico. The results of our analysis indicate that strains of this biovar are strongly heterogeneous in their chromosomal structural genes and show a level of genetic diversity and genotypic divergence greater than that yet reported for any single species of bacterium.

MATERIALS AND METHODS

Bacterial strains. Fifty-one isolates of *R. leguminosarum* bv. *phaseoli* were obtained from the collections of the

* Corresponding author.

TABLE 1. ETs and sources of isolates of *R. leguminosarum* bv. *phaseoli*

ET	Isolate	Geographic source	Host plant
1	CFN-1	Guanajuato, Mexico	*P. vulgaris*
2	OLIVIA-4	Minnesota, United States	*P. vulgaris*
3	CFN-2	Guanajuato, Mexico	*P. vulgaris*
4	F-3	Hidalgo, Mexico	*P. vulgaris*
5	COC-8	Morelos, Mexico	*P. coccineus*
6	FE-4	Hidalgo, Mexico	*P. vulgaris*
	FE-5	Hidalgo, Mexico	*P. vulgaris*
7	FL-69[a]	Puebla, Mexico	*P. vulgaris*[a]
8	F-13	Morelos, Mexico	*P. vulgaris*
9	F-14 (=CFN-285)	Morelos, Mexico	*P. vulgaris*
10	F-15 (=CFN-286)	Morelos, Mexico	*P. vulgaris*
11	FE-12	Morelos, Mexico	*P. vulgaris*
12	FL-51[a]	Puebla, Mexico	*P. vulgaris*[a]
13	FL-17	Puebla, Mexico	*P. vulgaris*
14	F-7	Hidalgo, Mexico	*P. vulgaris*
15	F-12	Hidalgo, Mexico	*P. vulgaris*
16	BRA-281	Brazil	*P. vulgaris*
17	BRA-8	Brazil	*P. vulgaris*
18	BRA-5	Brazil	*P. vulgaris*
	BRA-6	Brazil	*P. vulgaris*
19	CIAT-281	Colombia	*P. vulgaris*
20	SIL-4 (=CFN-307)	Morelos, Mexico	Wild *P. vulgaris*
21	ACU-3 (=CFN-227)	Morelos, Mexico	*P. acutifolius*
22	DES-105	Jalisco, Mexico	*Desmodium* sp.
23	F-2	Morelos, Mexico	*P. vulgaris*
	F-16	Hidalgo, Mexico	*P. vulgaris*
	F-19	Hidalgo, Mexico	*P. vulgaris*
24	F-17	Hidalgo, Mexico	*P. vulgaris*
	F-18	Hidalgo, Mexico	*P. vulgaris*
25	CFN-42	Guanajuato, Mexico	*P. vulgaris*
26	F-6	Jalisco, Mexico	*P. vulgaris*
27	F-9	Jalisco, Mexico	*P. vulgaris*
28	TAL-182	Hawaii	*P. vulgaris*
29	VIKING-1	Belize	*P. vulgaris*
30	DES-109	Jalisco, Mexico	*Desmodium* sp.
31	F-20	Morelos, Mexico	*P. vulgaris*
32	CFN-3	Guanajuato, Mexico	*P. vulgaris*
33	NITRAGIN-8251	United States	*P. vulgaris*
34	F-4	Jalisco, Mexico	*P. vulgaris*
35	CIAT-893	Colombia	*P. vulgaris*
36	CIAT-894	Colombia	*P. vulgaris*
37	CIAT-895	Colombia	*P. vulgaris*
38	CFN-7	Guanajuato, Mexico	*P. vulgaris*
39	CFN-299 (=UMR 1020)	Brazil	*P. vulgaris*
40	CIAT-899	Colombia	*P. vulgaris*
41	FL-27[a]	Puebla, Mexico	*P. vulgaris*[a]
42	FL-34[a]	Puebla, Mexico	*P. vulgaris*[a]
43	FL-44[a]	Puebla, Mexico	*P. vulgaris*[a]
44	FL-48[a]	Puebla, Mexico	*P. vulgaris*[a]
45	FL-53[a]	Puebla, Mexico	*P. vulgaris*[a]
46	COC-11	Morelos, Mexico	*P. coccineus*

[a] Collected in soil from a field grown to *L. leucocephala*; trapped on cultivated beans.

Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Mexico (Table 1). The 51 isolates were obtained from the following plant sources: 38 from nodules on the roots of cultivated *P. vulgaris*; 1 from wild *P. vulgaris*; 1 from cultivated tepary beans (*Phaseolus acutifolius* var. *latifolius*); 2 from cultivated scarlet runner beans (*Phaseolus coccineus*); and 2 from wild tick clover (*Desmodium* sp.). Seven additional isolates were recovered from soil in a field of cultivated huaje (*Leucaena leucocephala*) by being trapped on cultivated beans. Most of the isolates were collected in Mexico, but the sample also included five isolates from Colombia, five from Brazil, one from Belize, two from the continental United States, and one from Hawaii (Table 1).

Each of the 51 isolates had been identified as *R. leguminosarum* bv. *phaseoli* on the basis of ability to nodulate roots of the negro Jamapa cultivar of the common bean. However, only 43 of the isolates (those representing electrophoretic types [ETs] 1 to 38; Table 1) were judged to be typical *R. leguminosarum* bv. *phaseoli* by the criteria discussed by Martinez et al. (24) and Brom et al. (4). These criteria are the presence of reiterated *nifH* sequences on the Sym (for symbiotic) plasmid, ability to produce melanin pigment (a plasmid-mediated character), and ability to produce a gummy polysaccharide (a chromosome-mediated character).

In addition, the Sym plasmids of these isolates are of the narrow-host-range type (4).

Two of the isolates, CFN-299 (Brazil) and CIAT-899 (Colombia), lack the listed characteristics and carry the broad-host-range type of Sym plasmid, which shows little nucleotide sequence homology with the Sym plasmids of the typical strains of *R. leguminosarum* bv. *phaseoli*. These two strains further differ from the typical strains in having the capacity to nodulate *L. leucocephala* (24).

Two of the seven isolates (FL-51 and FL-69) collected in soil from a field grown to *L. leucocephala* were typical *R. leguminosarum* bv. *phaseoli*. Because the other five isolates (FL-27, FL-34, FL-44, FL-48, and FL-53) from this field were able to nodulate and fix nitrogen (albeit poorly) in common beans, they technically may be assigned to *R. leguminosarum* bv. *phaseoli* (19), but they have none of the special characteristics of typical *R. leguminosarum* bv. *phaseoli*. Moreover, their Sym plasmids are distinctive, being unlike either the narrow-host-range or the broad-host-range type.

Preparation of lysates. Isolates were grown overnight on an orbital shaker at 30°C in 150 ml of tryptone-yeast broth supplemented with 5.3 mM calcium chloride. Cells were harvested by centrifugation at 6,000 × *g* for 10 min at 4°C. After suspension in 2 ml of 50 mM Tris hydrochloride buffer containing 5 mM EDTA (pH 7.5), the bacteria were sonicated for 2 min (Branson model 200 Sonifier with microtip), at 50% pulse, with ice cooling. After centrifugation of the sonicated suspension for 30 min at 20,000 × *g*, the clear lysates (supernatants) were poured off and stored at −70°C.

Multilocus enzyme electrophoresis. Techniques of starch-gel electrophoresis and selective staining of enzymes as applied to the population genetics and systematics of bacteria have been described by Selander et al. (30). Fifteen enzymes were assayed. The electrophoretic buffer systems and the enzymes assayed were as follows: Tris citrate, pH 6.7 (6-phosphogluconate dehydrogenase, hydroxybutyrate dehydrogenase, aconitase, NAD-malate dehydrogenase); Tris citrate, pH 8.0 (phosphoglucose isomerase, isocitrate dehydrogenase, glucose 6-phosphate dehydrogenase, glyceraldehyde 3-phosphate dehydrogenase, leucylalanine peptidase, leucine aminopeptidase); borate, pH 8.2 (indophenol oxidase, hexokinase, adenylate kinase, phosphoglucomutase); and Tris maleate, pH 7.4 (glutamate dehydrogenase).

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 electromorph profiles for the 15 enzymes (ETs) were equated with multilocus genotypes (30). Because all or nearly all isolates showed activity for all 15 enzymes assayed, we presume that the corresponding structural genes are chromosomal rather than plasmid borne. For the enzymes glucose 6-phosphate dehydrogenase and indophenol oxidase, Young (35) demonstrated activity in a strain of *R. leguminosarum* bv. *viceae* that had been cured of the Sym plasmid.

From allele frequencies for ETs, genetic diversity for an enzyme locus was calculated as $h = (1 - \Sigma x_i^2)[n/(n - 1)]$, where x_i^2 is the frequency of the ith allele and n is the number of ETs. Mean genetic diversity per locus (H) is the arithmetic average of h values for the 15 loci (30). Genetic distance between each pair of ETs was estimated as the proportion of loci at which dissimilar alleles occurred (mismatches), and clustering from a matrix of pairwise genetic distances was performed by the average-linkage method (34).

TABLE 2. Genetic diversity at 15 enzyme loci among ETs of *R. leguminosarum* bv. *phaseoli*

Enzyme locus[a]	Characteristics of:			
	46 ETs[b]		38 ETs[c]	
	No. of alleles	Genetic diversity (h)	No. of alleles	Genetic diversity (h)
PGM	7	0.655	7	0.580
ADK	13	0.813	11	0.822
HEX	10	0.770	7	0.701
MDH	6	0.410	3	0.359
PEP	9	0.755	5	0.649
G6P	8	0.789	6	0.714
IDH	7	0.759	6	0.726
HBD	9	0.781	5	0.701
6PG	9	0.808	7	0.774
ACO	6	0.776	6	0.728
LAP	12	0.861	9	0.818
G3P	6	0.775	5	0.728
GLD	5	0.544	4	0.400
PGI	7	0.493	6	0.376
IPO	4	0.374	3	0.152
Mean	7.9	0.691	6.0	0.615

[a] Abbreviations: PGM, phosphoglucomutase; ADK, adenylate kinase; HEX, hexokinase; MDH, NAD-malate dehydrogenase; PEP, leucylalanine peptidase; G6P, glucose 6-phosphate dehydrogenase; IDH, isocitrate dehydrogenase; HBD, hydroxybutyrate dehydrogenase; 6PG, 6-phosphogluconate dehydrogenase; ACO, aconitase; LAP, leucine aminopeptidase; G3P, glyceraldehyde 3-phosphate dehydrogenase; GLD, glutamate dehydrogenase; PGI, phosphoglucose isomerase; IPO, indophenol oxidase.

[b] Represented by the total sample of 51 isolates examined.

[c] Represented by 43 typical *R. leguminosarum* bv. *phaseoli* isolates (see text).

RESULTS

Genetic diversity among isolates. In the collection of 51 isolates of *R. leguminosarum* bv. *phaseoli* studied, all 15 enzymes assayed were polymorphic for 4 to 13 electromorphs (i.e., alleles; mean number per locus, 7.9; Table 2). There were 46 distinctive multilocus genotypes, or ETs (Table 3), among which mean genetic diversity per locus (H) was 0.691 (range in h from 0.374 for indophenol oxidase to 0.861 for leucine aminopeptidase).

The 38 ETs to which the 43 typical *R. leguminosarum* bv. *phaseoli* isolates were assigned (ETs 1 through 38) were somewhat less variable; H was 0.615 (range in h from 0.152 for indophenol oxidase to 0.822 for adenylate kinase), and the mean number of alleles per locus was 6.0.

Only four of the 46 ETs (ETs 6, 18, 23, and 24) were represented by more than one isolate. In each of two cases (ETs 6 and 24), two isolates of identical multilocus genotype were recovered from bean plants in the same field; ET 23 was represented by two isolates from a field in Hidalgo and also by an isolate from a field in Morelos; we have no information as to whether the two strains of ET 18 were collected in the same field.

Linkage disequilibrium. As a measure of association of alleles at loci in multilocus genotypes, we compared the observed variance in the distribution of mismatches among all possible pairs of ETs with the variance expected if multilocus genotypes were formed by random combination

TABLE 3. Allele profiles at 15 enzyme loci in 46 ETs of *R. leguminosarum* bv. *phaseoli*

ET	Alleles at indicated enzyme locus[a]														
	PGM	ADK	HEX	MDH	PEP	G6P	IDH	HBD	6PG	ACO	LAP	G3P	GLD	PGI	IPO
1	5	5	10	7	5	5	3	5	3	5	3	1	5	4	5
2	5	4	5	7	5	5	3	5	3	5	1.5	3.5	5	6.2	5
3	5	5	5	7	5	5	3	5	2.7	6	1.5	3	5	6	5
4	5	4	7	7	5	5	3	5	4	5	2	3.5	5	4	5
5	5	5	7	7	5	5	3	5	4	5	1	3	5	7	5
6	5	5	3	7	5	5	3	5	4	2	2	2	5	6.2	5
7	5	5	3	7	5	5	3	5	4	3	2	3	5	4	5
8	5	5	7	7	5	5	3	0	0	5	3	2	5	6	5
9	5	5	7	7	5	5	3	5	3	3	3	2	5	6	5
10	5	5	7	7	5	5	3	5	4	2	3	2	5	6	5
11	5	5	7	7	5	5	3	5	4	2	3	4	5	6	5
12	5	5	7	7	5	5	2.5	5.5	4	2	1	2	5	6	5
13	3	5	7	7	5	5	3	5	3	2	2.5	3.5	5	6	5
14	5	5	3	7	5	5	3	0	0	0	2	1	5	6	5
15	5	0	7	7	5	0	3	0	0	0	3	1	5	6	5
16	5	1	5	7	2	5	6	5.5	4	5	5	3	5	6	5
17	5	1	5	7	2	5	6	5	3	5	5	3	5	6	5
18	5	1	5	7	5	5	6	5	3	5	3	3	5	6	5
19	3.5	1	5	7	2	5	6	5.5	4	3	5	3	5	6	5
20	2.5	1.5	5	7	2	4	6	0	4	5	5	3	5	6	5
21	5	0.7	5	7	2	4	6	0	2.7	6	5	3	5	6	5
22	5	1.5	5	6	3	4	5	0	3	5	3	3	5	6	5
23	5	0.3	5	7	5	7	3	5	4	5	3	3	4	6	5
24	5	0.3	5	7	5	7	3	0	4	5	3	3	4	6	5
25	2	0.3	5	7	1	4	6.2	0	5	6	5	4	7	6	5
26	5	3	3	7	1	7	6	5	5	6	4	4	5	6	5
27	5	3	3	7	1	7	6	5	5	6	6	4	5	6	5
28	5	5.2	3	7	5	7	6	5	5	7	5	4	5	6	5
29	2.5	3	5	7	1	8	6	6	3	6	4	3	5	11	5
30	3	0.2	6	7	1	8	6	0	4	6	4	2	5	6	5
31	3	3	3	7	5	7	5	0	6	6	5	3	7	3.7	5
32	6	3	5	6	2	7.5	4	3	3	2	3	3	7	6	3
33	6	5	5	6	1	7	4	3	5	5	3	3	7	6	5
34	6	3	5	5	1	7	4	5	4	5	2	3	7	6	5
35	3	3	5	6	1	8	4	6	5	5	2	4	0	6	5
36	3	3	5	6	1	7.5	4	6	5	5	2	4	7	6	5
37	3	7	0.5	5	7	7	4	5.5	8	5	6.5	4	5	6	7
38	5.5	3	8	6	2	7.5	5	3	5	6	5	2	5	6	7
39	5	3	2.2	3.5	7.5	8	6	2	8	6	5	2	5	6	7
40	2	3.2	1	3	12	8	5	5.5	6	7	4	4	10	6	5
41	5	2	5	8	7.5	3	6.5	9	7	2	7	6	10	2.5	10
42	5.5	5	10	7	11	3	6.2	0	5	3	8	6	10	4	10
43	5.5	5	10	7	11	3	6	4	5	3	8	6	10	4	10
44	5.5	5	10	7	11	3	6	4	5	3	7	6	10	4	10
45	5.5	5	10	7	11	3	6	4	3.5	3	7	6	10	4	10
46	2	1.5	6.5	7	6	2	6.2	9.5	5	2	10	3	10	4	10

[a] Abbreviations are as in Table 2.

of alleles (5) (Fig. 1). The observed variance in proportion of mismatches was 10.785, and the expected variance was 2.858, with a 95% confidence limit of 4.030. Hence, the observed variance was significantly larger than expected, indicating that the observed combinations of alleles over loci were strongly nonrandom.

Genetic relationships of ETs. Genotypic diversity among the 46 ETs was very high (Fig. 2), with some pairwise genetic distances between ETs being 1.00, which means that the most strongly differentiated pairs of strains shared no alleles. ETs 1 through 15 formed a cluster at a level of about 0.45, which was joined by a second cluster of nine ETs (ETs 16 through 24) at a level of about 0.55. The remaining 22 ETs fell in seven single lineages or small clusters diverging at depths from about 0.65 to 0.89.

Note that ETs 39 through 46, represented by eight isolates that did not have the typical *R. leguminosarum* bv. *phaseoli* characteristics, belonged to lineages at the bottom of the dendrogram and that all of these except ET 39 were distantly related to ETs 1 through 38, to which the typical isolates were assigned.

Geographic variation. The 42 ETs represented by isolates from Mexico and South America (Colombia and Brazil) were analyzed, and the total genetic diversity at the 15 enzyme loci was apportioned geographically. Mean genetic diversity per locus for the 33 ETs from Mexico was 0.676, and the comparable value for the 9 ETs represented by isolates from Colombia and Brazil was 0.657. The total diversity (H_T) for the 42 ETs was 0.691 (coincidentally the same value as for the total sample of 46 ETs), and mean diversity within the two geographic regions (H_S) was 0.666. Consequently, the interregional component of diversity, $F_{ST} = (H_T - H_S)/H_T$, was 0.035, indicating only a mild degree of regional differentiation.

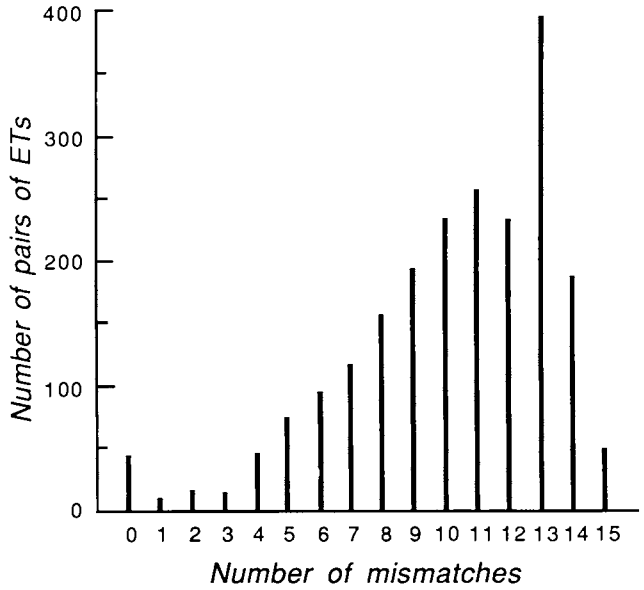

FIG. 1. Distribution of numbers of mismatches (numbers of loci at which dissimilar alleles occurred) for all possible pairs of 46 ETs of *R. leguminosarum* bv. *phaseoli*.

DISCUSSION

Levels of variation at metabolic enzyme loci. Electrophoretically detectable polymorphisms in several enzymes have been described for *Rhizobium* spp. (e.g., reference 11) and used to some extent in species identification (9), but only recently has multilocus enzyme electrophoresis been used to study the genetic population structure of these bacteria (35, 36). The results of these population genetics studies may be summarized as follows. Analysis of variation in three enzymes (glucose 6-phosphate dehydrogenase, superoxide dismutase [equivalent to indophenol oxidase], and β-galactosidase) in 252 isolates recovered from nodules on lucerne, peas, clover, and beans in a single field in Norfolk, England, yielded a mean genetic diversity per locus (H) of 0.50. Because variation in the three enzymes was correlated, only 15 distinctive combinations of alleles (ETs) were distinguished. Most of the 28 isolates from lucerne (*R. meliloti*) belonged to two ETs; one of these ETs was not represented by any isolate of *R. leguminosarum*, and the other was represented by one isolate each of *R. leguminosarum* bv. *viceae* and *R. leguminosarum* bv. *trifolii*. Further evidence of sharing of ETs was the occurrence of one strain of *R. meliloti* that had the common genotype of all three biovars of *R. leguminosarum*. These findings, together with those reported by Eardly et al. (8), suggest that there is no absolute genetic boundary between strains nodulating lucerne and those nodulating other host plants.

For 89 isolates of *R. leguminosarum* bv. *viceae* recovered from peas, Young (35) identified 10 ETs, 4 of which were common, and among 94 isolates of *R. leguminosarum* bv. *trifolii*, he distinguished 9 ETs, 2 of which were common. There was extensive sharing of ETs between these two biovars. For 41 isolates obtained from *Phaseolus* beans, Young (35) identified only two ETs, one of which was represented by 40 of the *R. leguminosarum* bv. *phaseoli* isolates and was also a common ET of *R. leguminosarum* bv. *trifolii* and *R. leguminosarum* bv. *viceae* isolates. At a second site 200 m distant in the same field, he collected 37 isolates of *R. leguminosarum* bv. *phaseoli*; again only two

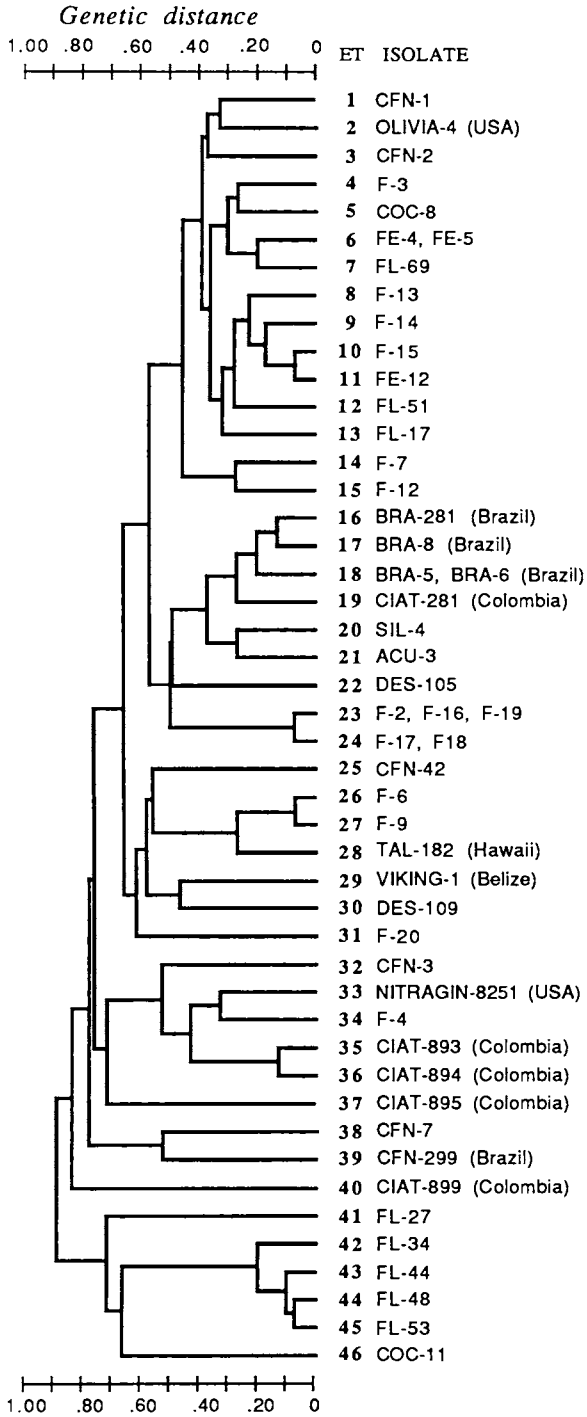

FIG. 2. Genetic relatedness among 46 ETs of *R. leguminosarum* bv. *phaseoli* based on electrophoretically detectable allelic variation at 15 enzyme loci (Table 2).

ETs were represented but, in this case, in nearly equal frequencies.

The low genotypic diversity observed by Young (35) among isolates from cultivated beans in Norfolk, England, stands in contrast to our findings for strains of *R. leguminosarum* bv. *phaseoli* from the Americas and to other evidence of considerable genetic heterogeneity among isolates of this biovar (7, 12, 14, 28). Young attributed the low level of

variation he observed to the narrow geographic base of his sample and noted that there are no native hosts for this biovar in Britain. He further suggested the possibility that strains that can nodulate beans have colonized or evolved since the introduction of cultivated *Phaseolus* beans to England.

To compare levels of genotypic diversity, we determined the number of ETs represented in two collections of isolates from Mexico on the basis of variation at only three enzyme loci, those encoding glucose 6-phosphate dehydrogenase, indophenol oxidase, and phosphoglucose isomerase. (Because β-galactosidase was not assayed in our study, we substituted phosphoglucose isomerase, which showed a level of genetic diversity similar to that recorded by Young [35] for β-galactosidase in England.) In one sample of nine isolates from a field in Hidalgo, there were four ETs; for six isolates from a field in Morelos, there also were four ETs. These numbers compare with the two ETs represented by 42 isolates from the study field in Norfolk, England, and two ETs for 37 isolates from the second locality reported by Young (35). Thus, populations of *R. leguminosarum* bv. *phaseoli* in single fields apparently are genetically more heterogeneous in Mexico than in England.

The number of ETs recognized in a sample of isolates obviously will depend on the number of loci analyzed. Considering only 6-phosphogluconate dehydrogenase, indophenol oxidase, and phosphoglucose isomerase, there were 16 ETs among our 51 isolates; for the total of 15 loci that we assayed, however, the number of ETs increased to 46.

Linkage disequilibrium and clonal population structure. Young (35) found evidence of linkage disequilibrium (nonrandom association of alleles in multilocus genotypes) in his populations, with some combinations of common alleles over the three loci being unrepresented. There was also evidence of strong linkage disequilibrium in our collection of isolates, as reflected by the inflated variance of the distribution of mismatches for pairs of ETs (Fig. 1). With linkage disequilibrium, the number of new ETs distinguished gradually decreases as the number of loci assayed increases, but it is clear that the analysis of only a few loci may seriously underestimate the full extent of genotypic variation in populations.

As previously noted by Young (35), the occurrence of strong linkage disequilibrium in natural populations of *Rhizobium* spp. implies a basic clonal population structure, in which the frequency of recombination of chromosomal genes is not high enough to randomize the genome. Clonal structure has been reported for many other types of bacteria, including *Escherichia coli* (1, 2, 31, 33), *Salmonella* spp. (P. Beltran, J. M. Musser, R. Helmuth, J. J. Farmer III, W. M. Frerichs, I. K. Wachsmuth, K. Ferris, J. G. Wells, A. Cravioto, and R. K. Selander, Proc. Natl. Acad. Sci. USA, in press), *Legionella pneumophila* (32), *Neisseria meningitidis* (6), and *Bordetella* spp. (26).

Association between plasmid and chromosomal markers. Two types of *R. leguminosarum* bv. *phaseoli* have been distinguished on the basis of the number of copies of the *nif* genes present on the Sym plasmid (24). The plasmids of most isolates carry two or three copies, but there is only one copy on the plasmids of some isolates from Colombia and Brazil. Moreover, six patterns (I through VI) of arrangement of reiterated *nif* genes have been distinguished.

We have analyzed some of the isolates studied by Martinez et al. (24). These included isolates with a single copy of the *nif* genes (ETs 39 and 40) and those representing pattern I (ETs 1, 3, 21, 25, 26, and 27), pattern II (ETs 8, 14, and 15), pattern III (ETs 20 and 29), and pattern IV (ETs 9 and 10). These data suggest that there is only a weak correlation between *nif* gene pattern and chromosomal genotype. At present, we do not know whether these results reflect the horizontal transfer of plasmid genes among strains of different lineages or convergence in the pattern of arrangement of plasmid DNA. However, under laboratory conditions, spontaneous rearrangements of both plasmid and chromosomal DNAs are not uncommon in *R. leguminosarum* bv. *phaseoli* (10). It is noteworthy that Schofield et al. (29) identified the same Sym plasmids in chromosomally unrelated strains of *R. leguminosarum* bv. *trifolii* as well as dissimilar Sym plasmids in chromosomally similar strains, which indicates the occurrence of interstrain exchange of plasmids within soil populations.

Genetic variation and classification. We question the proposition that *R. leguminosarum* bv. *phaseoli* is a meaningful taxonomic unit, since the level of genotypic diversity it accommodates greatly exceeds that of conventionally defined species in other groups of bacteria. Similar conclusions were reached by Roberts et al. (28) on the basis of an analysis of two-dimensional polyacrylamide gel electrophoresis patterns of total proteins. Many of the strains we studied are very distantly related, sharing few if any alleles at the 15 loci; on the basis of their overall chromosomal genetic character, they should be assigned to separate species.

The data of Young (35) for populations in England suggested that there is little chromosomal gene variation in *R. leguminosarum* bv. *phaseoli*, but our analysis indicates an unusually high degree of diversity; also, consistent with our findings, Flores et al. (10) found that each of more than 50 native strains from Mexico had a unique genomic fingerprint and plasmid profile. It is likely that *R. leguminosarum* bv. *phaseoli* is represented in natural populations in the Americas by hundreds, if not thousands, of multilocus enzyme genotypes. Indeed, the level of genetic diversity recorded for our sample ($H = 0.691$) is greater than that previously reported for any well-studied single species of bacterium, including *E. coli* (31). Moreover, the genetic distance between many pairs of the strains was 1.00 (Fig. 1), and several clusters of ETs diverged at distances greater than 0.80, a level well beyond those reported for strains of any other species of bacterium. Our finding of great depth among strains identified as *R. leguminosarum* bv. *phaseoli* is consistent with the evidence from the DNA hybridization studies of Jarvis et al. (16) and Crow et al. (7). By the convention currently adopted for members of the family *Enterobacteriaceae* and certain other groups of bacteria (3), our estimates of genetic distance between pairs of ETs of *R. leguminosarum* bv. *phaseoli* indicate that as many as seven species could be distinguished among the strains analyzed.

There is an obvious need for a reclassification of the genus *Rhizobium* on the basis of variation in the chromosomal genome rather than of host specificity or other phenotypic characteristics that are mediated largely or entirely by genetic determinants borne on plasmids. The present classification is an ecological-physiological scheme, not a genetic-evolutionary classification reflecting the overall similarity of strains and their probable genealogical lines of descent. Perhaps through evolutionary time, numerous diverse clonal lineages associated with various species of legumes have been able to colonize *Phaseolus* spp. by the acquisition of plasmids conferring genes for nodulation. Graham and Parker (13) long ago demonstrated that *P. vulgaris* is a promiscuous host for *Rhizobium* species (see also reference

8), and nodulation of common beans by strains of *Rhizobium* spp. recovered from root nodules of a wide variety of genera and species of wild legumes has been reported by Lange (22), in Australia, and by Martinez et al. (24), in Mexico. On the basis of these findings and the observation that, on bean plants, nodule numbers and nitrogen-fixing performances of *Agrobacterium tumefaciens* transconjugants carrying the broad-host-range-type Sym plasmid of *R. leguminosarum* bv. *phaseoli* are similar to those of these nonspecific strains, Martinez et al. (23) suggested that *P. vulgaris* is "quite a permissible host for nodulation, allowing many different bacteria to establish symbiosis with it. That is, symbiotic plasmids with *P. vulgaris* nodulating genes may be contained in different bacterial backgrounds and direct the establishment of an effective symbiosis with beans." Brom et al. (4) have shown that a variety of types of Sym plasmids can confer the ability to establish an effective symbiosis with *P. vulgaris*.

What is presently called *R. leguminosarum* bv. *phaseoli* may well be a polyphyletic assemblage of genetic lineages rather than a monophyletic unit. Because the horizontal transfer of chromosomal genes is likely to occur much less frequently than the transfer of plasmid genes, the multilocus chromosomal genotype may be expected to better reflect the geneaological histories of these lineages.

ACKNOWLEDGMENTS

This research was supported by Public Health Service grant AI22144 from the National Institutes of Health.

LITERATURE CITED

1. **Achtman, M., M. Heuzenroeder, B. Kusecek, H. Ochman, D. Caugant, R. K. Selander, V. Vaisanen-Rhen, T. K. Korhonen, S. Stuart, F. Ørskov, and I. Ørskov.** 1986. Clonal analysis of O2:K1 *Escherichia coli* isolated from diseased humans and animals. Infect. Immun. **51**:268–276.
2. **Achtman, M., and G. Pluschke.** 1986. Clonal analysis of descent and virulence among selected *Escherichia coli*. Annu. Rev. Microbiol. **40**:185–210.
3. **Brenner, D. J.** 1983. Impact of modern taxonomy on clinical microbiology. ASM News **49**:58–63.
4. **Brom, S., E. Martinez, G. Davila, and R. Palacios.** 1988. Narrow- and broad-host-range symbiotic plasmids of *Rhizobium* spp. that nodulate *Phaseolus vulgaris*. Appl. Environ. Microbiol. **54**:1280–1283.
5. **Brown, A. H. D., and M. W. Feldman.** 1981. Population structure of multilocus associations. Proc. Natl. Acad. Sci. USA **78**: 5913–5916.
6. **Caugant, D. A., L. F. Mocca, C. E. Frasch, L. O. Frøholm, 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.
7. **Crow, V. L., B. D. W. Jarvis, and R. M. Greenwood.** 1981. Deoxyribonucleic acid homologies among acid-producing *Rhizobium*. Int. J. Syst. Bacteriol. **31**:152–172.
8. **Eardly, B. D., D. B. Hannaway, and P. J. Bottomley.** 1985. Characterization of rhizobia from ineffective alfalfa nodules: ability to nodulate bean plants [*Phaseolus vulgaris* (L.) Savi.]. Appl. Environ. Microbiol. **50**:1422–1427.
9. **Engvild, K. C., and G. Nielsen.** 1984. Distinguishing rhizobia by electrophoresis of isoenzymes, p. 333. *In* C. Veeger and W. E. Newton (ed.), Advances in nitrogen fixation research. Nijhoff, The Hague.
10. **Flores, M., V. González, M. A. Pardo, A. Leija, E. Martínez, D. Romero, D. Piñero, G. Dávila, and R. Palacios.** 1988. Genomic instability in *Rhizobium phaseoli*. J. Bacteriol. **170**:1191–1196.
11. **Fottrell, P. F., and A. O'Hora.** 1969. Multiple forms of D(−)3-hydroxybutyrate dehydrogenase in *Rhizobium*. J. Gen. Microbiol. **57**:287–292.
12. **Graham, P. H.** 1964. The application of computer techniques to the taxonomy of the root-nodule bacteria of legumes. J. Gen. Microbiol. **35**:511–517.
13. **Graham, P. H., and C. A. Parker.** 1964. Diagnostic features in the characterization of the root-nodule bacteria of legumes. Plant Soil **20**:383–396.
14. **Heberlein, G. T., J. De Ley, and R. D. E. Tijtgat.** 1967. Deoxyribonucleic acid homology and taxonomy of *Agrobacterium*, *Rhizobium*, and *Chromobacterium*. J. Bacteriol. **94**:116–124.
15. **Hooykaas, P. J. J., A. A. N. van Brussel, H. den Dulk-Ras, G. M. S. van Slogteren, and R. A. Schilperoort.** 1981. Sym plasmid of *Rhizobium trifolii* expressed in different rhizobial species and *Agrobacterium tumefaciens*. Nature (London) **291**: 351–353.
16. **Jarvis, B. D. W., A. G. Dick, and R. M. Greenwood.** 1980. Deoxyribonucleic acid homology among strains of *Rhizobium trifolii* and related species. Int. J. Syst. Bacteriol. **30**:42–52.
17. **Johnston, A. W. B., and J. E. Beringer.** 1977. Chromosomal recombination between *Rhizobium* species. Nature (London) **267**:611–613.
18. **Johnston, A. W. B., J. L. Beynon, A. U. Buchanan-Wollaston, S. M. Setchell, P. R. Hirsch, and J. E. Beringer.** 1978. High frequency transfer of nodulating ability between strains and species of *Rhizobium*. Nature (London) **276**:634–636.
19. **Jordan, D. C.** 1984. Family III. *Rhizobiaceae* Conn 1938, 321^{AL}, p. 234–254. *In* N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. Williams & Wilkins, Baltimore.
20. **Kondorosi, A., E. Vincze, A. W. B. Johnston, and J. E. Beringer.** 1980. A comparison of three *Rhizobium* linkage maps. Mol. Gen. Genet. **178**:403–408.
21. **Lamb, J. W., G. Hombrecher, and A. W. B. Johnston.** 1982. Plasmid-determined nodulation and nitrogen fixation abilities in *Rhizobium phaseoli*. Mol. Gen. Genet. **186**:449–452.
22. **Lange, R. T.** 1961. Nodule bacteria associated with the indigenous Leguminosae of South-Western Australia. J. Gen. Microbiol. **26**:351–359.
23. **Martínez, E., R. Palacios, and F. Sánchez.** 1987. Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. J. Bacteriol. **169**:2828–2834.
24. **Martinez, 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.
25. **Moffett, M. L., and R. R. Colwell.** 1968. Adansonian analysis of the Rhizobiaceae. J. Gen. Microbiol. **51**:245–266.
26. **Musser, J. M., D. A. Bemis, H. Ishikawa, and R. K. Selander.** 1987. Clonal diversity and host distribution in *Bordetella bronchiseptica*. J. Bacteriol. **169**:2793–2803.
27. **Palacios, R., M. Flores, S. Brom, E. Martinez, V. Gonzalez, S. Frenk, C. Quinto, M. A. Cevallos, L. Segovia, D. Romero, A. Garciarrubio, D. Pinero, and G. Davila.** 1987. Organization of the *Rhizobium phaseoli* genome, p. 151–156. *In* D. P. Verma and N. Brisson (ed.), Molecular genetics of plant-microbe interactions. Martinus Nijhoff Publishers, Montreal.
28. **Roberts, G. P., W. T. Leps, L. E. Silver, and W. J. Brill.** 1980. Use of two-dimensional polyacrylamide gel electrophoresis to identify and classify *Rhizobium* strains. Appl. Environ. Microbiol. **39**:414–422.
29. **Schofield, P. R., A. H. Gibson, W. F. Dudman, and J. M. Watson.** 1987. Evidence for genetic exchange and recombination of *Rhizobium* symbiotic plasmids in a soil population. Appl. Environ. Microbiol. **53**:2942–2947.
30. **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.
31. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger

(ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.

32. **Selander, R. K., R. M. McKinney, T. S. Whittam, W. F. Bibb, D. J. Brenner, F. S. Nolte, and P. E. Pattison.** 1985. Genetic structure of populations of *Legionella pneumophila*. J. Bacteriol. **163**:1021–1037.
33. **Selander, R. K., J. M. Musser, D. A. Caugant, M. N. Gilmour, and T. S. Whittam.** 1987. Genetic analysis of populations of pathogenic bacteria by multilocus enzyme electrophoresis. Microb. Pathog. **3**:1–7.
34. **Sneath, P. H. A., and R. R. Sokal.** 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco.
35. **Young, J. P. W.** 1985. *Rhizobium* population genetics: enzyme polymorphism in isolates from peas, clover, beans and lucerne grown at the same site. J. Gen. Microbiol. **131**:2399–2408.
36. **Young, J. P. W., L. Demetriou, and R. G. Apte.** 1987. *Rhizobium* population genetics: enzyme polymorphism in *Rhizobium leguminosarum* from plants and soil in a pea crop. Appl. Environ. Microbiol. **53**:397–402.