Genetic Structure of *Rhizobium leguminosarum* biovar trifolii and viciae Populations Found in Two Oregon Soils under Different Plant Communities†

STEVEN R. STRAIN,^{1,2} KAMTIN LEUNG,¹ THOMAS S. WHITTAM,³ FRANS J. DE BRUIJN,⁴
AND PETER J. BOTTOMLEY^{1,2*}

Department of Microbiology¹ and Department of Crop and Soil Sciences,² Oregon State University, Corvallis, Oregon 97331-3804; Department of Biology, The Pennsylvania State University, University Park, Pennsylvania 16802³; and Michigan State University-Department of Energy Plant Research Laboratory, Department of Microbiology, and National Science Foundation Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824⁴

Received 18 January 1994/Accepted 2 June 1994

An investigation was carried out to determine the genetic structure in soil populations of *Rhizobium leguminosarum* bv. trifolii and viciae at each of two Oregon sites (A and C) that were 1 km apart. Although the soils were similar, the plant communities were quite different because grazing by domestic animals had been allowed (site A) or prevented (site C). Analysis of allelic variation at 13 enzyme-encoding loci by multilocus enzyme electrophoresis delineated 202 chromosomal types (ETs) among a total of 456 isolates representing two populations of *R. leguminosarum* bv. trifolii (AT and CT) and two populations of *R. leguminosarum* bv. viciae (AV and CV). Regardless of their site of origin or biovar affiliation, isolates of the same ET were confirmed to be more closely related to each other than to isolates of other ETs by repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus sequences and the PCR technique. Despite the wide range in densities of the *Rhizobium* populations ($<10^2$ to $>10^5$ /g of soil), their overall genetic diversities were similar (mean genetic diversity, 0.45 to 0.51), indicating that low-density populations of soil-borne bacterial species are not necessarily of little genetic diversity. Linkage disequilibrium analysis revealed significant multilocus structure (nonrandom associations of alleles) within each of the four populations. From a combination of cluster and linkage disequilibrium analyses, a total of eight distinct groups of ETs were defined in the four populations. Two groups (I and III) contributed significant numbers of ETs and isolates to each population. The two populations of *R. leguminosarum* bv. viciae (AV and CV) exhibited similar genetic structures despite existing at different densities, in different plant communities, and in the presence (CV) or absence (AV) of their local *Vicia* hosts. In contrast, total linkage disequilibrium was partitioned differently in two biovar populations occupying the same soil (AV and AT), with disequilibrium in the latter being due entirely to the presence of group V.

Substantial amounts of research have been directed at the population genetic structure of human- and animal-associated bacterial species (20, 22, 27, 28, 36, 37). Many of these bacterial species were shown to exhibit a clonal population structure, in which a species consists of a finite number of chromosomally distinct genotypes and exhibits nonrandom associations of alleles (linkage disequilibrium) at chromosomal loci. Clonal structure is thought to be maintained in these species because genetic exchange events of a magnitude sufficient to erode the individuality of the chromosome are rare relative to the rates of clonal expansion, migration, and random extinction of existing genotypes.

Recent evidence indicates, however, that the genetic structure of soil-borne species may differ from that considered the norm for the majority of animal-associated species. Istock et al. (14) analyzed a local population of *Bacillus subtilis* recovered from a soil microsite and found little evidence of linkage

disequilibrium. They concluded that genetic exchange was occurring often relative to asexual reproduction and that, in the absence of selective pressure, clonal structure had not developed. Souza et al. (30, 31) arrived at similar conclusions by demonstrating that several local populations of *Rhizobium etli* were in linkage equilibrium. In addition, they concluded that recombination was occurring between genotypes in a local population which exhibited linkage disequilibrium. Although these data support the possibility that the forces which drive the ecology and evolution of animal- and soil-associated species might be different, only a few population genetic studies have been conducted with soil-borne relative to animal-associated species.

Recently, we described the serotypic composition of a *Rhizobium leguminosarum* bv. trifolii soil population at a specific location in Oregon. We determined the intra- and interserotype relationships of the isolates by examining allelic variation at 13 chromosomally encoded enzyme loci by multilocus enzyme electrophoresis (MLEE) (18, 19). Because of our familiarity with this particular soil population and the ecology and management of the surrounding area, an opportunity existed to extend our studies to *R. leguminosarum* bv. trifolii and viciae populations at other sites in the vicinity. In our previous study, we showed the importance of using several clover (*Trifolium*) species to obtain a clear picture of the

* Corresponding author. Mailing address: Department of Microbiology, Oregon State University, Nash Hall, Rm. 220, Corvallis, OR 97331-3804. Phone: (503) 737-1844. Fax: (503) 737-0496. Electronic mail address: bottomlp@ucs.orst.edu.

† Technical paper 10,375 of the Oregon Agricultural Experiment Station.

diversity within a soil population of *R. leguminosarum* bv. trifolii (12, 21, 25, 33, 35). In the case of *R. leguminosarum* bv. viciae, it is not clear if there is a need for multiple trap host species to estimate diversity, since the evidence for specific hosts selecting specific subtypes from soil populations is equivocal (13, 17, 32, 38, 39). With the latter concern in mind, we identified an uncultivated, ungrazed, open woodland site about 1 km from our original site, where four of the most common vetch species of the Pacific Northwest are found in abundance (*Vicia hirsuta*, *Vicia villosa*, *Vicia sativa*, and *Vicia americana*) and where the *R. leguminosarum* bv. viciae population occurs in high density ($>10^5$ /g of soil). Furthermore, since only a few *Trifolium dubium* individuals were found on the site perimeter, the *R. leguminosarum* bv. trifolii population density at the site was low ($<10^2$ /g of soil). In contrast, our original site is a grass-subclover (*Trifolium subterraneum*) pasture in which *Vicia* spp. do not exist because of heavy cattle grazing. As a result, soil population densities of the two biovars on this site are reversed relative to that of the vetch-dominated site ($>10^5$ /g for bv. trifolii; $<10^2$ /g for bv. viciae). Our objectives were (i) to compare genetic diversities within populations of the two different biovars existing at different densities in the same soil and (ii) to compare genetic diversity of different soil populations of the same biovar existing under different plant communities in which host plant species were either present or absent.

MATERIALS AND METHODS

Study site descriptions. Site A is an improved subclover (*T. subterraneum* L.)-orchard grass (*Dactylis glomerata* L.) pasture located on a northwest-facing toe slope of Soap Creek Valley, approximately five miles north of Corvallis, Ore. *Vicia* species do not occur on the site because of heavy grazing by cattle. The soil is a silty-clay loam of the Abiqua series (fine, mixed, mesic, Cumulic Ultic Haploxeroll), and the site and soil characteristics have been described in detail elsewhere (1, 3). Site C is an unimproved, ungrazed, open woodland area located mid-slope on the southeast-facing side of Soap Creek Valley. The soil is a silty-clay loam of the Dixonville series (fine, mixed, mesic, Pachic Ultic Haploxeroll). Although several *Vicia* species are abundant on the site, only a few specimens of *T. dubium* were identified on the perimeter of the site. The densities of the populations of *R. leguminosarum* bv. viciae and trifolii in surface soil samples (0 to 10 cm deep) recovered from sites A and C were determined by the soil dilution-plant infection most probable number procedure (34). *V. villosa* and *Trifolium hybridum* were used as hosts for *R. leguminosarum* bv. viciae and trifolii, respectively. On site C, *R. leguminosarum* bv. viciae and trifolii were present at respective densities of $>10^5$ and $<10^2$ g of oven-dry soil $^{-1}$. On site A, population densities of the two biovars showed a reciprocal relationship to those of site C, with *R. leguminosarum* bv. viciae and trifolii present at $<10^2$ and $>10^5$ g of oven-dry soil $^{-1}$, respectively.

Populations of *R. leguminosarum*. In an attempt to obtain isolates that represented a full spectrum of genetic diversity in the four soil populations, we used several host plant species for recovering isolates. In addition, we conducted a preliminary serological analysis of the isolates with antisera to 13 antigenically distinct serotypes of *R. leguminosarum* bv. trifolii recovered from site A (19). A total of 176 isolates of *R. leguminosarum* bv. viciae were recovered from field-grown *V. hirsuta*, *V. villosa*, *V. sativa*, and *V. americana* and from *Pisum sativum* and *V. villosa* grown in soil transported from site C to the laboratory (CV population). Isolates of *R. leguminosarum* bv. trifolii were recovered from field-grown *T. dubium* growing at site C

and from subclover, crimson clover (*Trifolium incarnatum*), and white clover (*Trifolium repens*) grown in soil transported from site C to the laboratory. From a preliminary serological analysis of 175 isolates, eight antigenically distinct serotypes were identified and 70 isolates were chosen from among these serotypes (CT population). A total of 112 isolates of *R. leguminosarum* bv. viciae were recovered from root nodules of four host plants, *P. sativum*, *V. hirsuta*, *V. villosa*, and *V. sativa*, grown in soil transported from site A to the laboratory (AV population). Two hundred isolates of *R. leguminosarum* bv. trifolii were analyzed from a larger collection of isolates recovered from various clover species grown under field or greenhouse conditions in site A soil. Isolates were chosen from among 13 serotypes so that as many of the chromosomal types in the soil population as possible were examined (AT population). The characterization of the AT population has been described in detail elsewhere (18, 19).

Recovery of rhizobial isolates from plants. Plants were transported from the field to the laboratory, with their root systems intact and surrounded by soil. Excess soil was removed by washing roots in 0.1% (vol/vol) Tween 20. Nodules were harvested and surface sterilized by standard procedures (34). Isolates were obtained from laboratory-grown plants as follows. Approximately 1,500-ml portions of sterile vermiculite were packed into plastic pots and moistened with 600 ml of a sterile plant nutrient solution devoid of mineral nitrogen (19). Surface soil (0 to 10 cm deep) was obtained from either site A or site C and amended with KH_2PO_4 (55 mg of P kg $^{-1}$), $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ (1 mg of Mo kg $^{-1}$), and K_2SO_4 (20 mg of S kg $^{-1}$) in sufficient water (350 g kg $^{-1}$ [oven-dry soil]) to raise the soil-water potential to approximately -30 kPa. Portions of amended soil (1.5 kg) were added to each of three replicate pots for each plant species. The soil layer was gently packed to reduce settling, and the weight of each pot with soil was recorded. Seeds of vetch, pea, and clover species were surface sterilized by standard procedures (34) and sown 1 to 2 cm deep into each of the three replicate pots. Upon germination, vetch and clover seedlings were thinned to approximately 10 per pot and peas were thinned to about 4 per pot. Plants were grown under greenhouse conditions. After 6 wk of growth, 20 nodules were recovered from each replicate pot of each species and rhizobial isolates were recovered by standard procedures (34).

MLEE. The preparation of cell extracts, electrophoretic conditions, and enzyme assays have been described in detail elsewhere (18, 26). Each isolate was characterized by its combination of allelic variants over the 13 enzymes assayed. Each distinct allelic variant profile was termed an electrophoretic type (ET). Allelic profiles for each of the isolates in the populations are available from the authors upon request. Genetic diversity (h) at an enzyme locus is calculated as $h = (1 - \sum x_i^2) / [n / (n - 1)]$, where x_i is the frequency of the i th allele at the locus, n is the number of ETs in the sample, and $n / (n - 1)$ is a correction factor for bias in small samples (24). Mean genetic diversity (H) is the arithmetic average of the h values over all enzyme loci examined. Coefficients of genetic differentiation [$G_{ST} = (H_T - H_S) / H_T$] (23) were used to apportion total diversity (H_T) over several populations into within-population (H_S) and between-population components. Relationships among ETs were revealed with a cluster analysis program (ETCLUS) designed specifically for this purpose. The program was translated from FORTRAN to the C programming language with minor modifications prior to use. ETCLUS employs the UPGMA algorithm (29) to cluster from a matrix of pairwise distances (calculated as the proportion of mismatched loci) between ETs. The relative similarities among ETs were displayed in the form of a dendrogram.

To determine the extent to which the populations exhibit linkage disequilibrium (nonrandom combinations of alleles between loci), the observed allelic mismatch frequency distribution was obtained by comparing each ET with every other ET once [for a total of $n(n-1)/2$ comparisons, where n is the number of ETs], and for each paired comparison, the number of dissimilar alleles (mismatches) was recorded. An equation for computing the variance of this distribution (V_o) has been derived previously (5, 6) and the process has been described in detail elsewhere (18, 22, 30, 31). The inflation of the observed variance of the allelic mismatch distribution (V_o) over the variance of the mismatch distribution expected if there were random associations of alleles at loci (V_e) provides an indication of the extent of linkage disequilibrium among loci.

Comparison of MLEE with REP and ERIC PCR. The profiles of DNA fragments generated by PCR with repetitive extragenic palindromic (REP) and enterobacterial repetitive intergeneric consensus (ERIC) sequences as primers were determined as described elsewhere (8). Isolates were chosen from three ETs and included strains of both biovars and of different geographical origins. Similarities among the isolates were evaluated by comparing the banding patterns observed among the isolates combined over both REP- and ERIC-specific patterns. Banding patterns were coded by enumerating the total number of unique bands seen over all the isolates and, for each isolate, assigning each band position a 1 or a 0 to indicate the presence or absence of the band, respectively. For a pair of isolates, similarity was calculated as the sum of the number of bands present in both isolates and the number of bands absent from both isolates divided by the total number of bands observed for all isolates examined. Similarity among the isolates was revealed in the form of a dendrogram constructed from the REP and ERIC PCR data with the NTSYS-pc analysis package (version 1.50; Exeter Software, Setauket, N.Y.).

RESULTS

MLEE versus REP and ERIC PCR. The variability among isolates of the same ET was assessed by a PCR technique with REP and ERIC sequence-specific primers (8). Although no two isolates of the same ET gave identical patterns of DNA fragments produced by PCR, the maximum divergence between isolates within any of the three groups was at a relative dissimilarity of 0.12 (Fig. 1). It is clear from these results that isolates of the same ET were more closely related to one another than to isolates of other ETs. Moreover, this conclusion applies across biovars, because both ETs 9 and 74 contain isolates of *R. leguminosarum* bv. trifolii clustering with isolates of *R. leguminosarum* bv. viciae. Within an ET group, there was no trend for isolates from the same site or of the same biovar to be more closely related to each other than to others in the group. For example, in the group represented by ET11, one of the site A isolates was more similar to site C isolates than to others from site A.

Genetic diversity in the populations. For each of the four populations, genetic diversity values (h) at each of the 13 enzyme-encoding loci are presented (Table 1). Mean genetic diversities (H) are similar for each population, ranging from 0.45 (CT population) to 0.51 (AV population). Although genetic diversities at the majority of individual loci were generally similar across the four populations, diversities at three loci were greater within specific populations (phosphoglucosyltransferase in CV, glucose-6-phosphate dehydrogenase in AV, and isocitrate dehydrogenase in AT). Only 25 of the 82 alleles identified over all loci are restricted to either site A or

FIG. 1. Relationships between *Rhizobium* isolates of the same ET determined by PCR with REP and ERIC sequence-specific primers.

site C populations, with the majority (23 of 25) of these population-specific alleles recovered in only one to three ETs. However, one AT population-specific β -hydroxybutyrate dehydrogenase allele was identified in 12 ETs, and one phosphoglucosyltransferase allele was identified exclusively in seven ETs recovered from the CV population. Coefficients of genetic differentiation (G_{ST}) were calculated to evaluate the interpopulation heterogeneity among allele frequencies at each locus (Table 2). There was no evidence that either biological or geographical separation of the populations contributed substantially to the variation in allele frequencies among the four populations. When each of the four populations was treated as a distinct subpopulation (the All comparison in Table 2), the total G_{ST} was 0.042, indicating that 4.2% of the total genetic diversity in the four populations was due to variation in the allele frequencies between the populations.

Linkage disequilibrium within the Oregon populations. Although G_{ST} coefficients indicate that the level of genetic differentiation between the four populations at single loci is low, they provide no information about the extent to which the populations are composed of genotypically distinct lineages. To assess the multilocus structure and the extent of linkage disequilibrium, we calculated the observed (V_o) and expected (V_e) variances of the allele mismatch distribution and their ratios for each of the four populations of *R. leguminosarum* (Table 3). V_o/V_e ratios ranged from 1.6 (CV population) to 2.0 (AV population), with the two *R. leguminosarum* bv. trifolii populations (CT and AT) having virtually identical values (1.7). The extent of linkage disequilibrium in each of the populations was statistically significant at the 95% confidence level when calculated by the method of Brown et al. (6).

To elucidate which subpopulations contributed to the genetic structure of each population, ETs were provisionally grouped by the ETCLUS program. Analyses of linkage disequilibrium were carried out to define those clusters of ETs

TABLE 1. Genetic diversity (h) at each of 13 enzyme-encoding loci in four *R. leguminosarum* populations from two Oregon soils

Locus	Genetic diversity at a specific locus in population ^a :			
	<i>R. leguminosarum</i> bv. <i>viciae</i>		<i>R. leguminosarum</i> bv. <i>trifolii</i>	
	CV	AV	CT	AT
β -Galactosidase	0.76 (6)	0.79 (6)	0.80 (5)	0.80 (6)
Glucose-6-phosphates dehydrogenase	0.03 (2)	0.36 (4)	0.07 (2)	0.14 (3)
Isocitrate dehydrogenase	0.25 (3)	0.18 (4)	0.14 (2)	0.48 (5)
Malate dehydrogenase	0.00 (1)	0.09 (3)	0.00 (1)	0.04 (2)
β -Hydroxybutyrate dehydrogenase	0.68 (6)	0.71 (5)	0.83 (6)	0.80 (10)
6-Phosphogluconate dehydrogenase	0.35 (5)	0.53 (6)	0.42 (3)	0.43 (3)
Leucyl-glycine peptidase	0.71 (6)	0.77 (8)	0.79 (6)	0.73 (5)
Phosphoglucosomerase	0.67 (6)	0.67 (5)	0.52 (4)	0.48 (5)
Xanthine dehydrogenase	0.77 (6)	0.75 (6)	0.72 (5)	0.60 (6)
Nucleoside phosphorylase	0.72 (6)	0.56 (5)	0.70 (4)	0.64 (4)
Phosphoglucosomutase	0.32 (3)	0.05 (2)	0.07 (2)	0.00 (1)
Adenylate kinase	0.54 (4)	0.62 (4)	0.42 (2)	0.48 (2)
Superoxide dismutase	0.51 (2)	0.50 (3)	0.35 (2)	0.53 (3)
H^b	0.49 (4.3)	0.51 (4.7)	0.45 (3.4)	0.47 (4.2)
Isolates (n)	176	112	70	200
ETs (n)	78	42	28	54

^a Values in parentheses are number of alleles at each locus in each population. AV and CV represent the *R. leguminosarum* bv. *viciae* populations from sites A and C, respectively. AT and CT represent the *R. leguminosarum* bv. *trifolii* populations from sites A and C, respectively.

^b Mean genetic diversity in population.

whose alleles at the 13 loci are in linkage equilibrium (randomly assorted) but which are in disequilibrium with the other groups. By combining populations and subjecting them to cluster analysis followed by linkage disequilibrium analysis, we developed a consensus population structure which shows the overall relationships between major groups and the occurrence of each group in the four populations (Fig. 2). Groups I and III made substantial contributions to the percentages of ETs in all four of the populations (Table 4). Groups II and V were not found in all populations, and their contributions ranged from major to minor. Groups IV and VI were not found in all populations and contributed minimally to the gene pool.

A linkage disequilibrium analysis was conducted with each soil population to identify which groups contributed to the overall multilocus structure (Table 5). In both of the *R.*

leguminosarum bv. *viciae* populations (CV, AV), disequilibrium was significant between all combinations of groups contributing significant numbers of ETs to the AV (groups I, II, III, IV, and VII) and CV (groups I, II, and III) populations. This finding indicates that ETs within each group are sufficiently similar to each other and yet dissimilar enough from ETs in other groups to be the source of the linkage disequilibrium in these populations. In the AT population, however, the presence of group V strains created most of the linkage disequilibrium, because exclusion of these isolates from the analysis caused the remainder of the population (groups I and III) to be in linkage equilibrium. Interestingly, there was almost a parallel situation in the AV population, involving groups I and III, because linkage disequilibrium was barely significant when they were combined. Despite the common

TABLE 2. Coefficients of genetic differentiation (G_{ST}) for Oregon populations of *R. leguminosarum* bv. *viciae* and *trifolii*^a

Locus	Comparison between populations ^b :				
	All	Site C	Site A	<i>R. leguminosarum</i> bv. <i>viciae</i>	<i>R. leguminosarum</i> bv. <i>trifolii</i>
β -Galactosidase	0.028	0.026	0.025	0.012	0.000
Glucose-6-phosphate dehydrogenase	0.053	0.000	0.025	0.071	0.000
Isocitrate dehydrogenase	0.040	0.020	0.042	0.003	0.042
Malate dehydrogenase	0.003	0.000	0.000	0.003	0.003
β -Hydroxybutyrate dehydrogenase	0.060	0.009	0.067	0.028	0.038
6-Phosphogluconate dehydrogenase	0.026	0.000	0.040	0.004	0.022
Leucyl-glycine peptidase	0.016	0.000	0.013	0.010	0.005
Phosphoglucosomerase	0.013	0.004	0.008	0.005	0.000
Xanthine dehydrogenase	0.063	0.016	0.075	0.000	0.096
Nucleoside phosphorylase	0.057	0.013	0.063	0.029	0.018
Phosphoglucosomutase	0.054	0.033	0.000	0.044	0.000
Adenylate kinase	0.026	0.055	0.006	0.000	0.000
Superoxide dismutase	0.081	0.092	0.000	0.002	0.092
Total	0.042	0.021	0.035	0.012	0.029

^a G_{ST} is calculated as the ratio of between-population diversity to total diversity in all populations (see Materials and Methods for details).

^b All, four populations (CV, AV, CT, and AT); site C, CV versus CT; site A, AV versus AT; *R. leguminosarum* bv. *viciae*, CV versus AV; *R. leguminosarum* bv. *trifolii*, CT versus AT.

TABLE 3. Linkage disequilibrium statistics for Oregon populations of *R. leguminosarum* bv. *viciae* and *trifolii*

Population	No. of ETs	Mismatch parameter			V_o/V_e	Significance ^c
		χ^2	V_o^b	V_e^b		
CV	78	6.30	3.86	2.38	1.6	$1.64 < V_e < 3.11$
CT	28	5.83	3.74	2.15	1.7	$1.03 < V_e < 3.27$
AV	42	6.59	5.05	2.49	2.0	$1.43 < V_e < 3.54$
AT	54	6.14	4.18	2.41	1.7	$1.52 < V_e < 3.31$

^a χ^2 , mean number of mismatched loci (out of 13) over $n(n-1)/2$ paired ET comparisons.

^b V_o and V_e , observed and expected variances of allele mismatch distribution, respectively.

^c Values represent the upper and lower limits of the 95% confidence intervals around the V_e values calculated by the method of Brown et al. (6).

occurrence of group overlap among the four populations, there were only 15 ETs that were found in more than one population. Eleven ET overlaps occurred between the AV and CV populations, with 10 of them in group I. It is unlikely that the same ETs arose independently on each site from assortative recombination events, because their frequencies of occurrence are severalfold greater than random chance predicts from the product of their particular allele frequencies in the population (data not presented).

Distribution of the groups among the plant species. Three groups (I, II, and III) accounted for 92 and 98% of the isolates of *R. leguminosarum* bv. *viciae* recovered on sites A and C, respectively. χ^2 analysis of the distribution of isolates from the various hosts among these groups revealed a highly significant degree of heterogeneity on both site A ($\chi^2 = 21.94$, df = 6, $P < 0.005$) and site C ($\chi^2 = 31.50$, df = 8, $P < 0.001$). However,

FIG. 2. Consensus dendrogram illustrating the overall relationships between the major groups of ETs and their distribution among the four populations of *R. leguminosarum*. An asterisk denotes that the group in the specific population is represented by only one ET.

TABLE 4. Distribution of isolates and ETs among the major groups in the four populations of *R. leguminosarum* bv. *viciae* and *trifolii*

Subpopulation group	% of isolates (ETs) in the group from population:			
	CV	AV	CT	AT
I	54.5 (48.7)	48.2 (45.2)	25.7 (21.4)	28.3 (25.9)
II	5.7 (5.1)	13.4 (9.5)	0.0 (0.0)	1.0 (1.9)
III	37.5 (43.6)	17.0 (14.3)	47.1 (67.9)	28.8 (55.6)
IV	0.6 (1.3)	7.1 (19.0)	1.4 (3.6)	0.0 (0.0)
V	1.7 (1.3)	0.0 (0.0)	25.7 (7.1)	41.9 (16.7)
VI	0.0 (0.0)	0.9 (2.4)	0.0 (0.0)	0.0 (0.0)
VII	0.0 (0.0)	13.4 (9.5)	0.0 (0.0)	0.0 (0.5)
VIII	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.0 (0.5)

if isolates from *V. villosa* are omitted from the analysis, statistical significance is lost on both sites (site A, $\chi^2 = 8.90$, df = 4, $P > 0.05$; site C, $\chi^2 = 10.05$, df = 6, $P > 0.05$). These effects are caused primarily by group III isolates being overrepresented in nodules of *V. villosa* on site C relative to the other hosts and by greater representation of group I isolates in *V. villosa* nodules on site A. Although previous studies from this laboratory showed that subclover and crimson clover grown at site A nodulated extensively with group V (19), both of these species were nodulated primarily by representatives of groups I and III in site C soil. Field-grown *T. dubium* on site C, however, was nodulated almost exclusively by isolates from group V.

DISCUSSION

The findings of this study have specific implications for *Rhizobium* biology and ecology and, more generally, for the

TABLE 5. Intergroup disequilibrium statistics within each of the four populations of *R. leguminosarum*

Comparison in population ^a	No. of ETs ^b	V_o^c	V_e^c	V_o/V_e	Significance ^d
CV					
Total	78	3.86	2.38	1.63	$1.64 < V_e < 3.11$
I + II + III	76	3.88	2.36	1.64	$1.60 < V_e < 3.12$
I + III	72	3.88	2.36	1.64	$1.60 < V_e < 3.12$
I + II	42	3.10	2.09	1.49	$1.20 < V_e < 2.97$
AV					
Total	42	5.05	2.49	2.03	$1.43 < V_e < 3.54$
I + II + III + IV	37	4.04	2.43	1.66	$1.39 < V_e < 3.48$
I + II + III	29	3.79	2.22	1.71	$1.16 < V_e < 3.27$
I + III	25	3.60	2.32	1.55	$1.06 < V_e < 3.59$
I + II	23	3.80	1.96	1.94	$0.93 < V_e < 2.98$
CT					
Total	28	3.74	2.15	1.74	$1.03 < V_e < 3.27$
I + III	25	3.92	2.06	1.90	$0.93 < V_e < 3.19$
AT					
Total	54	4.18	2.41	1.73	$1.52 < V_e < 3.31$
I + III	44	3.40	2.46	1.38	$1.44 < V_e < 3.47$

^a Only those intergroup comparisons which are directly comparable among the four populations are presented.

^b Number of ETs in the specific comparison.

^c V_o and V_e , observed and expected variances of allele mismatch distribution, respectively.

^d Values represent the upper and lower limits of the 95% confidence intervals around the V_e values calculated by the method of Brown et al. (6).

population biology of soil-borne bacterial species. Consequently, the goal of this discussion will be to expand upon the significance of this work to each of these fields. REP and ERIC PCR grouped strains of the same ET into well-defined clusters regardless of their biovar affiliation or geographical origin. While chromosomal instability, spontaneous mutation, and genetic exchange can explain the intra-ET variation, such mechanisms are obviously not so extensive that they can break down the multilocus structure revealed by examining allelic variation at the 13 enzyme-encoding loci. It is noteworthy that the mean overall genetic diversity ($H = 0.48$) for the four soil populations of *R. leguminosarum* bv. *viciae* and *trifolii* is similar to estimates obtained from analyses of *R. leguminosarum* bv. *trifolii* strains from culture collections (9, 18) and those obtained from analyses of many local and cosmopolitan populations of *Escherichia coli* (27). Presumably, there can be as much genetic diversity within local populations of *R. leguminosarum* bv. *viciae* and *trifolii* as in the species as a whole.

A key observation of our study is that genetic diversity in the low-density populations of *R. leguminosarum* bv. *viciae* and *trifolii* on sites A and C, respectively, were nearly identical to values in the high-density populations on the same sites. Because *Vicia* species are widespread throughout the areas of the Soap Creek valley that are neither grazed nor cultivated, we assume that *Vicia* spp. were once abundant on site A and that the AV *Rhizobium* population was larger than its present size. Presumably, the conditions conducive for subclover success (i.e., intense domestic animal grazing) have caused the demise of the upright-growing *Vicia* species and a concomitant decline in the AV population. It is apparent, however, that neither the decline of the host population nor the change in environmental conditions has created sufficient selective pressure to cause substantial loss of genetic diversity or change in the genetic structure of the AV *Rhizobium* population. In the case of site C, we favor the idea that clover species were never a major component of the plant community because of their inability to establish themselves and persist among tall grasses and woody shrubs. In this habitat, the climbing tendency of *Vicia* spp. makes them better-adapted legumes for these plant communities. We hypothesize that the high level of diversity measured in the CT population and the disequilibrium detected between groups I and III result from the random immigration of strains onto the site. In contrast to our findings, Harrison et al. (11) found that levels of genetic diversity in populations of *R. leguminosarum* bv. *trifolii* were significantly lower in acidic soils (pH 4.2 to 4.4), where rhizobial populations were present at $<10^2$ g of soil⁻¹, than in soils of higher pH, where densities were greater. Because it is well known that neither *Trifolium* species nor *R. leguminosarum* bv. *trifolii* will thrive under extremely acidic conditions (2, 10), it is possible that both plant absence and selective environmental pressure contributed to the low level of diversity observed in the United Kingdom studies. Obviously, further experimental studies of genetic diversity in soil-borne *Rhizobium* species are needed to gain understanding of the impact of host removal, environmental conditions, and migration on genetic diversity of *Rhizobium* populations.

By conducting MLEE analyses of strains obtained from various culture collections, we found that groups I and III of *R. leguminosarum* are distributed globally among both biovars. Several U.S. Department of Agriculture strains of *R. leguminosarum* bv. *viciae* (including ATCC 10004 and 10314) recovered from widely separate geographic locations belong to group III. In addition, isolates representing ATCC strains of *R. leguminosarum* bv. *trifolii* (10328, 14479, and 14484), a common ET (MFF) recovered from pea nodules at two United Kingdom

sites (38, 39), and the nodule-dominant serotype 2 recovered from peas grown in eastern Washington state (4) belong to group III. Laguerre et al. (17) showed that the dominant chromosomal type of *R. leguminosarum* bv. *phaseoli*, *trifolii*, and *viciae*, identified in nodules on plants growing in France, corresponded with the MFF chromosomal type. Group I contains ETs designated MSM, MSK, and MSS by Young and coworkers that were recovered from pea, clover, and bean nodules in the United Kingdom (38, 39); members of serotypes AS21 and WS1-01, which are common occupants of subclover nodules in Oregon (18); and strains WU95, CC275e, and 162X95, which are currently used as clover inoculants in Australia, New Zealand, and the United States, respectively (7, 15, 16).

In contrast to the recent studies with *B. subtilis* and *R. etli* (14, 30, 31), we detected linkage disequilibrium in each of the four soil populations of *R. leguminosarum*. Nevertheless, clusters of ETs that showed random associations of alleles were found in each population, indicating that recombination has occurred between certain genotypes at some time in their history. Because all groups contributed equally to the linkage disequilibrium in both populations of *R. leguminosarum* bv. *viciae*, the possibility exists that specific chromosomal types from within each lineage have gained a nodulation or saprophytic advantage over other genotypes in the same lineages. The situation could be analogous to what Maynard Smith et al. (22) described recently as an epidemic clonal population structure in which linkage disequilibrium occurs transiently until assortative recombination of genes eventually returns the population to linkage equilibrium. In the case of the *R. leguminosarum* bv. *trifolii* population on site A, however, disequilibrium is caused entirely by the presence of group V. In this case, the success of group V in occupying root nodules of subclover growing at the site (18, 19) has not had a deleterious impact upon the genetic diversity within the remainder of the population. Further studies are required to determine if linkage disequilibrium is a stable or transient phenomenon within each of these populations. A combination of field surveys and soil experiments involving specific isolates could be used to address these questions.

ACKNOWLEDGMENTS

The work described in this paper was supported by the Oregon Agricultural Experiment Station and an NSF predoctoral fellowship awarded to S.R.S.

We thank J. P. W. Young, University of York, York, United Kingdom, P. van Berkum, USDA-ARS, Beltsville, Md., and D. F. Bezdeck, Washington State University, Pullman, for strains of *R. leguminosarum* bv. *viciae*. C. A. Istock is recognized for personal insights into appropriate use of statistical procedures. C. Pelroy is warmly thanked for processing the final drafts of the manuscript.

REFERENCES

1. Almendras, A. S., and P. J. Bottomley. 1987. Influence of lime and phosphate on nodulation of soil-grown *Trifolium subterraneum* L. by indigenous *Rhizobium trifolii*. Appl. Environ. Microbiol. 53: 2090-2097.
2. Bottomley, P. J. 1992. Ecology of *Bradyrhizobium* and *Rhizobium*, p. 293-348. In G. Stacey, H. J. Evans, and R. H. Burris (ed.), Biological nitrogen fixation. Chapman and Hall, New York.
3. Bottomley, P. J., and M. H. Dughri. 1989. Population size and distribution of *Rhizobium leguminosarum* bv. *trifolii* in relation to total soil bacteria and soil depth. Appl. Environ. Microbiol. 55:959-964.
4. Brockman, F. J., and D. F. Bezdeck. 1989. Diversity within serogroups of *Rhizobium leguminosarum* biovar *viciae* in the Palouse region of eastern Washington as indicated by plasmid

- profiles, intrinsic antibiotic resistance, and topography. *Appl. Environ. Microbiol.* **55**:109–115.
5. Brown, A. H. D., and M. W. Feldman. 1981. Population structure of multilocus associations. *Proc. Natl. Acad. Sci. USA* **78**:5913–5916.
 6. Brown, A. H. D., M. W. Feldman, and E. Nevo. 1980. Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics* **96**:523–536.
 7. Chatel, D. L., W. A. Shipton, and C. A. Parker. 1973. Establishment and persistence of *Rhizobium trifolii* in Western Australian soils. *Soil Biol. Biochem.* **5**:815–824.
 8. de Bruijn, F. J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* and other soil bacteria. *Appl. Environ. Microbiol.* **58**:2180–2187.
 9. Demezas, D. H., T. B. Reardon, J. M. Watson, and A. H. Gibson. 1991. Genetic diversity among *Rhizobium leguminosarum* bv. *trifolii* strains revealed by allozyme and restriction fragment length polymorphism analyses. *Appl. Environ. Microbiol.* **57**:3489–3495.
 10. Graham, P. H. 1992. Stress tolerance in *Rhizobium* and *Bradyrhizobium*, and nodulation under adverse soil conditions. *Can. J. Microbiol.* **38**:475–484.
 11. Harrison, S. P., D. G. Jones, and J. P. W. Young. 1989. *Rhizobium* population genetics: genetic variation within and between populations from diverse locations. *J. Gen. Microbiol.* **135**:1061–1069.
 12. Harrison, S. P., J. P. W. Young, and D. G. Jones. 1987. *Rhizobium* population genetics: effects of clover variety and inoculum dilution on the genetic diversity sampled from natural populations. *Plant Soil* **103**:147–150.
 13. Hynes, M. F., and M. P. O'Connell. 1990. Host plant effect on competition among strains of *Rhizobium leguminosarum*. *Can. J. Microbiol.* **36**:864–869.
 14. Istock, C. A., K. E. Duncan, N. Ferguson, and X. Zhou. 1992. Sexuality in a natural population of bacteria—*Bacillus subtilis* challenges the clonal paradigm. *Mol. Ecol.* **1**:95–103.
 15. 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.
 16. Jones, M. B., J. C. Burton, and C. E. Vaughn. 1978. Role of inoculation in establishing subclover in Californian annual grasslands. *Agron. J.* **70**:1081–1086.
 17. Laguerre, G., E. Geniaux, S. I. Mazurier, R. R. Casartelli, and N. Amarger. 1992. Conformity and diversity among field isolates of *Rhizobium leguminosarum* bv. *viciae*, bv. *trifolii*, and bv. *phaseoli* revealed by DNA hybridization using chromosome and plasmid probes. *Can. J. Microbiol.* **39**:412–419.
 18. Leung, K., S. R. Strain, F. J. de Bruijn, and P. J. Bottomley. 1994. Genotypic and phenotypic comparisons of chromosomal types within an indigenous soil population of *Rhizobium leguminosarum* bv. *trifolii*. *Appl. Environ. Microbiol.* **60**:416–426.
 19. Leung, K., K. Yap, N. Dashti, and P. J. Bottomley. 1994. Serological and ecological characteristics of a nodule-dominant serotype from an indigenous soil population of *Rhizobium leguminosarum* bv. *trifolii*. *Appl. Environ. Microbiol.* **60**:408–415.
 20. Levin, B. R. 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. *Genetics* **99**:1–23.
 21. Masterson, C. L., and M. T. Sherwood. 1974. Selection of *Rhizobium trifolii* strains by white and subterranean clovers. *Ir. J. Agric. Res.* **13**:91–99.
 22. Maynard-Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt. 1993. How clonal are bacteria? *Proc. Natl. Acad. Sci. USA* **90**:4384–4388.
 23. Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* **41**:225–233.
 24. Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small sample of individuals. *Genetics* **89**:583–590.
 25. Robinson, A. C. 1969. Host selection for effective *Rhizobium trifolii* by red clover and subterranean clover in the field. *Aust. J. Agric. Res.* **20**:1053–1060.
 26. 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.
 27. 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.
 28. Selander, R. K., and B. R. Levin. 1980. Genetic diversity and structure in *Escherichia coli* populations. *Science* **210**:545–547.
 29. Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman and Co., San Francisco.
 30. Souza, V., L. Eguarte, 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.
 31. 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.
 32. Turco, R. F., and D. F. Bezdicek. 1987. Diversity within two serogroups of *Rhizobium leguminosarum* native to soils in the Palouse of eastern Washington. *Ann. Appl. Biol.* **111**:103–114.
 33. Valdivia, B., M. H. Dughri, and P. J. Bottomley. 1988. Antigenic and symbiotic characterization of indigenous *Rhizobium leguminosarum* bv. *trifolii* recovered from root nodules of *Trifolium pratense* L. sown into subterranean clover pasture soils. *Soil Biol. Biochem.* **20**:267–274.
 34. Vincent, J. M. 1970. A manual for the practical study of root-nodule bacteria. In International biological program handbook 15. Blackwell Scientific Publications, Oxford.
 35. Weaver, R. W., D. Sen, J. J. Coll, C. R. Dixon, and G. R. Smith. 1989. Specificity of arrowleaf clover for rhizobia and its establishment in soil from crimson clover pastures. *Soil Sci. Soc. Am. J.* **53**:731–734.
 36. Whittam, T. S., H. Ochman, and R. K. Selander. 1983. Geographic components of linkage disequilibrium in natural populations of *Escherichia coli*. *Mol. Biol. Evol.* **1**:67–83.
 37. Whittam, T. S., H. Ochman, and R. K. Selander. 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **80**:1751–1755.
 38. 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.
 39. 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.