Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 8389–8393, September 1992
Population Biology

Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: Evidence for sex?

(population genetics/recombination/migration/enzyme polymorphism/*Escherichia coli*)

VALERIA SOUZA*†, TOAI T. NGUYEN‡, RICHARD R. HUDSON‡, DANIEL PIÑERO*, AND RICHARD E. LENSKI§

*Centro de Ecología, Universidad Nacional Autonóma de México, Coyoacán, Ciudad de México, 04510, México; ‡Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717; and §Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824-1325

Communicated by Michael T. Clegg, June 5, 1992 (received for review December 4, 1991)

ABSTRACT **Many bacterial species exhibit strong linkage disequilibrium of their chromosomal genes, which apparently indicates restricted recombination between alleles at different loci. The extent to which restricted recombination reflects limited migration between geographically isolated populations versus infrequent mixis of genotypes within populations is more difficult to determine. We examined the genetic structure of *Rhizobium leguminosarum* biovar *phaseoli* populations associated with wild and cultivated beans (*Phaseolus* spp.) over several spatial scales, ranging from individual host plants to throughout the Western Hemisphere. We observed significant linkage disequilibrium at scales at least as small as a cultivated plot. However, the amount of disequilibrium was much greater among isolates collected throughout the Western Hemisphere than among isolates from one area of Mexico, even when disequilibrium was quantified using an index that scales for allelic diversity. This finding suggests that limited migration between populations contributes substantially to linkage disequilibrium in *Rhizobium*. We also compared the genetic structure for *R. leguminosarum* bv. *phaseoli* taken from a cultivated plot with that for *Escherichia coli* obtained from one human host in an earlier study. Even at this fine scale, linkage disequilibrium in *E. coli* was very near the theoretical maximum level, whereas it was much less extreme in the local population of *Rhizobium*. Thus, the genetic structure for *R. leguminosarum* bv. *phaseoli* does not exclude the possibility of frequent mixis within local populations.**

Although bacteria reproduce asexually, the processes of conjugation, transduction, and transformation allow the transfer of chromosomal genes among otherwise clonal lineages (1). Despite the existence of these mechanisms, those bacterial species that have been studied typically exhibit strong linkage disequilibrium (2–16), which apparently indicates restricted recombination between alleles at different loci. [In principle, selection for epistatic combinations of alleles at different loci can maintain linkage disequilibrium in the face of frequent recombination, but such situations are believed to be uncommon (2, 5, 13).] Linkage disequilibrium has also been found in several species of parasitic protozoa (17–19).

Additional information is necessary to discern whether linkage disequilibrium reflects infrequent mixis of genotypes within local populations or whether it results instead from limited migration (i.e., gene flow) between geographically isolated populations. For *Escherichia coli*, the inference that the observed disequilibrium reflects infrequent mixis within local populations is supported by the finding that geographic variation accounts for only a small fraction of the observed genetic diversity (5). Moreover, certain multilocus genotypes are widely distributed and persist stably for several decades (6, 10). [Even so, analyses of DNA sequence data from different genes in *E. coli* yield different phylogenies, which suggests some mixis between otherwise clonal lineages over much longer intervals (20).] In many other bacterial species that have been studied, however, the possible role of geographical isolation in promoting linkage disequilibrium has received less attention.

Also, most studies of the genetic structure of bacterial populations have been concerned with pathogens and commensals of animals. The effective isolation of genotypes in different individual hosts could be important in maintaining linkage disequilibrium, especially if the ecological processes governing colonization are such that mixed infections are relatively uncommon. In these circumstances, recombination might frequently occur but go undetected because, in the absence of genetic variation, mixis leaves no signature. Therefore, it is of interest to know whether similar levels of linkage disequilibrium exist in bacterial species that are free-living or associated with plants.

Toward these ends, we have analyzed the population genetic structure of *Rhizobium leguminosarum* biovar *phaseoli*, which lives in the soil as well as in nodules of wild and cultivated beans (*Phaseolus* spp.), where it fixes nitrogen. Host specificity in *Rhizobium* is determined by symbiotic plasmids, some of which are self-transmissible (21–23). *Rhizobium* species, like most other bacteria, are also infected by bacteriophages (24, 25), and these extrachromosomal elements may serve as vehicles for the exchange of chromosomal genes (26). [*Rhizobium* species can also be transformed in the laboratory (27), but we are unaware of any evidence for transformation under natural conditions in this genus.] Thus, there exists the potential for plasmid- and phage-mediated exchange of chromosomal genes in *Rhizobium*, but the extent to which this actually occurs in natural populations is much less clear. Piñero *et al.* (14) found very high chromosomal genetic diversity among isolates of *R. leguminosarum* bv. *phaseoli* collected throughout the Western Hemisphere, and they also observed significant linkage disequilibrium at this geographical scale. However, Young (9) detected much less chromosomal genetic diversity among isolates of this biovar from one field in England. To assess the relative importance of limited migration between populations versus infrequent mixis within populations of *R. leguminosarum* bv. *phaseoli*, we estimate linkage disequilibrium over spatial scales ranging from individual host plants to throughout the Western Hemisphere. We also compare the levels of linkage disequilibrium that exist in local populations of *Rhizobium* and *E. coli*.

MATERIALS AND METHODS

Bacterial Isolates. A total of 276 isolates of *R. leguminosarum* bv. *phaseoli* were obtained from nodules of 77 bean plants growing in five populations at three sites in the state of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†Present address: Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824-1325.

Morelos, Mexico, during 1988 (28). Of these isolates, 190 were from a cultivated plot of *Phaseolus vulgaris* in Santiago Tepetlapa and 20 were from a population of wild *Phaseolus coccineus* located about 45 km to the northeast. At a third site, 7 km south of the first, 33 isolates were obtained from wild *P. vulgaris*, 15 from wild *P. coccineus*, and 18 from wild *P. coccineus* in the vicinity of a cultivated plot. Each isolate was confirmed to nodulate *Phaseolus* spp.

Multilocus Enzyme Electrophoresis. All 276 bacterial isolates were characterized by starch-gel electrophoresis for enzymes encoded by nine polymorphic loci. Lysates were obtained for cultures derived from single colonies using the methods described by Piñero *et al.* (14). The procedures for starch-gel electrophoresis were those described by Selander *et al.* (29). The following nine enzymes, having already been identified as polymorphic in a preliminary survey, were assayed: isocitrate dehydrogenase, 6-phosphogluconate dehydrogenase, malate dehydrogenase, glucose-6-phosphate dehydrogenase, 3-hydroxybutyrate dehydrogenase, peptidase, xanthine dehydrogenase, and esterases 1 and 2. At least three electrophoretic assays were performed on each strain for each enzyme tested. Each distinctive multilocus combination of mobility variants is designated an electrophoretic genotype. A table that lists these genotypes, the corresponding number of isolates, and their source locations (by population and by individual host plant) is available from the first author.

Statistical Analysis. The coefficient of linkage disequilibrium between the ith allele at locus j and the kth allele at locus l is given by $D_{ik,jl} = g_{ik,jl} - p_{ji}p_{lk}$, where $g_{ik,jl}$ is the frequency of the two-locus haploid genotype and $p_{ji}p_{lk}$ is the product of the frequencies of the corresponding alleles. This coefficient is frequently standardized by dividing by D_{max}, the maximum value for $D_{ik,jl}$ that is possible given the allelic frequencies (30, 31). The absolute value of the standardized coefficient, $|D'|$, ranges from 0, at linkage equilibrium, to 1, when one or two of the four possible two-locus haploid genotypes are completely absent.

A particularly useful index to describe multilocus linkage disequilibrium in haploid populations is based on the distribution of allelic mismatches between pairs of isolates over all loci (3, 4, 14, 32, 33). For example, if one isolate has multilocus genotype $A_1B_1C_1D_1E_1$ and another isolate has genotype $A_2B_1C_2D_2E_1$, then the number of mismatches for this pair of isolates is three. For an infinite population, the mean number of mismatches equals $\Sigma_j h_j$, where $h_j = 1 - \Sigma_i p_{ji}^2$ is the allelic diversity at the jth locus. The corresponding variance in the number of mismatches is given by

$$\sum_j h_j - \sum_j h_j^2 + 2 \sum_j \sum_{l>j} \sum_i \sum_k (2p_{ji}p_{lk}D_{ik,jl} + D_{ik,jl}^2)$$

(32, 33). If there is no disequilibrium among loci (i.e., all $D_{ik,jl} = 0$), then the expected variance in mismatches reduces to $\Sigma_j h_j - \Sigma_j h_j^2$.

For a sample of n isolates, we calculate the observed mean and variance in the number of mismatches using the $n(n-1)/2$ pairs of isolates. The mean number of mismatches, for a given set of loci, provides a measure of allelic diversity. The ratio of the variance in mismatches observed for a particular sample, V_o, to the expected variance in a corresponding population at linkage equilibrium, V_e, provides a measure of multilocus linkage disequilibrium in which the denominator scales for the effects of the single-locus genetic diversities (32). One can determine for any sample whether V_o is statistically different from V_e, which would indicate significant linkage disequilibrium, by iterating a Monte Carlo procedure in which alleles in the sample are repeatedly scrambled (without replacement) to eliminate any effect of linkage disequilibrium. The nominal significance is computed as the probability of obtaining, by chance alone, a variance in the number of mismatches as or more extreme than actually observed.

RESULTS

The 276 isolates of *R. leguminosarum* bv. *phaseoli* obtained from several different populations in the state of Morelos, Mexico, comprised 100 distinct nine-locus electrophoretic genotypes. A total of 276(275)/2 = 37,950 pairwise comparisons are possible using this complete set of isolates (Table 1, row C). Pairs of isolates differed from one another, on average, at 3.03 of the nine loci (34%). The ratio of the observed variance in numbers of mismatches to the variance expected under the null hypothesis of linkage equilibrium, V_o/V_e, was 1.23. The Monte Carlo procedure described above indicated a significant difference between the observed and expected variances ($P < 0.001$, based on 10,000 iterations), thus demonstrating linkage disequilibrium among the bacteria within the state of Morelos.

Among bacterial isolates taken from the same population of beans, there are 18,931 possible pairwise comparisons (summing the within-population comparisons over all five populations). Pairs of isolates taken from the same host population differed from one another, on average, at 2.56 of the nine loci (28%), and the ratio V_o/V_e was 1.11 (Table 1, row B). The Monte Carlo procedure was performed using the single largest population (190 isolates, comprising almost 95% of the within-population comparisons) and yielded a significant deviation of observed and expected variances ($P < 0.01$, based on 10,000 iterations). Thus, linkage disequilibrium is also evident among the bacteria residing within a single cultivated plot.

Only 808 pairwise comparisons are possible among bacterial isolates taken from the same individual bean plant (summing the within-plant comparisons over all of those plants from which two or more isolates were obtained). These pairs of isolates differed from one another, on average, at 2.28 of the nine loci (25%), and the ratio V_o/V_e was 1.19 (Table 1,

Table 1. Hierarchical analysis of linkage disequilibrium for 276 isolates of *R. leguminosarum* bv. *phaseoli* in the state of Morelos, Mexico

Geographical scale	No. of pairs of isolates	Pairwise mismatches* $\bar{X}$	V_o	V_e	V_o/V_e	$P^\dagger$
Pairs of isolates						
A From same individual host plant	808‡	2.28	1.21	1.01	1.19	<0.05
B From same host population	18,931§	2.56	1.41	1.28	1.11	<0.01
C From Morelos	37,950	3.03	1.89	1.54	1.23	<0.001

*$\bar{X}$ = mean number of allelic mismatches among pairs of isolates (out of nine loci); V_o = observed variance in number of mismatches among pairs of isolates; V_e = expected variance assuming linkage equilibrium. In rows A and B, variances are based on deviations from means obtained for individual host plants and host populations, respectively.
†Probability of rejecting by chance alone the null hypothesis that $V_o = V_e$; see text for details.
‡Summed over all of those plants from which two or more isolates were obtained.
§Summed over all five host populations.

Table 2. Linkage disequilibrium in *R. leguminosarum* bv. *phaseoli* isolated from individual plants

Plant no.	No. of isolates	No. of pairs of isolates	Pairwise mismatches* $\bar{X}$	V_o	V_e	V_o/V_e	$P^\dagger$
2	12	66	2.50	1.25	0.93	1.35	0.03
6	10	45	3.16	1.78	1.33	1.33	0.09
12	11	55	2.20	1.00	0.93	1.07	0.44
19	12	66	2.99	2.26	1.57	1.43	0.03
21	11	55	2.53	0.79	1.25	0.64	0.95
23	15	105	2.22	1.20	1.21	0.99	0.50

All six plants that yielded 10 or more bacterial isolates were used.
*$\bar{X}$ = mean number of allelic mismatches among pairs of isolates (out of nine loci); V_o = observed variance in number of mismatches among pairs of isolates; V_e = expected variance assuming linkage equilibrium.
†Probability of rejecting by chance alone the null hypothesis that $V_o = V_e$.

row A). Ten or more isolates were obtained from each of six plants (Table 2), which accounted for almost half of the within-plant comparisons. The Monte Carlo procedure gave significant deviations of observed and expected variances in two of these cases ($P < 0.05$, based on 10,000 iterations for each), and Fisher's method for combining probabilities from independent tests of significance (34) rejected the null hypothesis of linkage equilibrium over all six cases combined ($P < 0.05$). Hence, we conclude that there may exist significant linkage disequilibrium even among isolates of *R. leguminosarum* bv. *phaseoli* obtained from the same host plant.

Note that the mean number of allelic mismatches ($\bar{X}$), which measures genetic diversity, increases with the spatial scale (Table 1). But at these relatively fine scales, there is no apparent trend in the ratio V_o/V_e, which reflects the extent of linkage disequilibrium. (The associated *P* values become progressively smaller due to the increased statistical power associated with larger sample sizes and greater genetic diversity.)

We also compared the 276 isolates of *R. leguminosarum* bv. *phaseoli* collected in the state of Morelos, Mexico, with 51 isolates obtained from throughout the Western Hemisphere and characterized genetically by Piñero *et al.* (14) using similar methods (Table 3). Because the number of mismatches and the ratio V_o/V_e depend on the number of loci scored (32), our comparison uses only those six loci that were surveyed in both studies. The proportion of mismatches was 38% (2.30/6) for pairs of isolates collected in the state of Morelos but increased to 69% (4.15/6) for pairs of isolates obtained from throughout the Western Hemisphere. The ratios of V_o/V_e for the samples from the state of Morelos and throughout the Western Hemisphere were 1.23 and 2.36, respectively. To determine if the difference in these ratios was statistically significant, we performed a "bootstrap" test (36). In this test, the two actual samples were treated as statistical populations, and 1000 new samples (each one of the same size as the original sample) were drawn with replacement from the two populations. The ratio of V_o/V_e was then calculated for each of the new samples in order to generate the approximate sampling distributions of this ratio for the two different geographical scales. There was no overlap whatsoever in these sampling distributions, thus indicating that the difference in linkage disequilibrium between these two geographical scales is highly significant ($P < 0.001$).

Two of the 51 isolates identified by Piñero *et al.* (14) as *R. leguminosarum* bv. *phaseoli* have recently been proposed for inclusion in a new species, *R. tropici* (35). Excluding these two isolates from the analysis has only a slight effect on the ratio of V_o/V_e, which decreases from 2.36 to 2.30 (Table 3). Furthermore, a repetition of the bootstrap test described above, but using the reduced set of isolates from throughout the Western Hemisphere, still gives no overlap in the sampling distributions of V_o/V_e at the two geographic scales ($P < 0.001$). This result confirms that linkage disequilibrium among genotypes from throughout the Western Hemisphere is much greater than among genotypes from the several populations in Morelos, Mexico.

DISCUSSION

Nitrogen-fixing bacteria such as *Rhizobium* have tremendous agricultural value, and hence they will be subject to many efforts for improvement by genetic engineering. Therefore, it is of interest to understand the population genetic structure of these species in nature. The extent to which these species comprise a relatively few geographically widespread genotypes, as opposed to more numerous indigenous and locally adapted genotypes, is important for choosing strains for genetic modification so as to increase their efficacy. Also, the extent to which genetic exchange occurs within local populations and the extent of migration between populations are important considerations from the standpoint of evaluating the potential spread of introduced genetic modifications (37).

Our results demonstrate significant linkage disequilibrium among isolates of *R. leguminosarum* bv. *phaseoli* from the same cultivated plot and even from the same individual host plant. Linkage disequilibrium provides strong evidence for restricted genetic recombination, although the possibility that selection for epistatic combinations of alleles might maintain linkage disequilibrium in the face of frequent recombination cannot be excluded. This evidence for restricted recombination at such fine geographical scales was possible

Table 3. Comparison of linkage disequilibrium for *R. leguminosarum* bv. *phaseoli* between the state of Morelos, Mexico, and throughout the Western Hemisphere

Geographical scale	No. of isolates	No. of pairs of isolates	Pairwise mismatches* $\bar{X}$	V_o	V_e	V_o/V_e	$P^\dagger$
Morelos, Mexico‡	276	37,950	2.30	1.42	1.15	1.23	<0.001
Western Hemisphere§	51	1,275	4.15	2.73	1.16	2.36	<0.001
Western Hemisphere¶	49	1,176	4.02	2.72	1.18	2.30	<0.001

*$\bar{X}$ = mean number of allelic mismatches among pairs of isolates (out of six loci); V_o = observed variance in number of mismatches among pairs of isolates; V_e = expected variance assuming linkage equilibrium.
†Probability of rejecting by chance alone the null hypothesis that $V_o = V_e$.
‡Using all 276 isolates included in Table 1, but using only the six loci also scored in Piñero *et al.* (14).
§Using all 51 isolates included in Piñero *et al.* (14), but using only the six loci also scored in this study; these isolates were obtained from the continental United States ($n = 2$), Hawaii ($n = 1$), Mexico ($n = 37$, including 5, 9, 5, 10, and 8 from the states of Guanajuato, Hidalgo, Jalisco, Morelos, and Peubla, respectively), Belize ($n = 1$), Columbia ($n = 5$), and Brazil ($n = 5$).
¶Excluding two isolates included in Piñero *et al.* (14) that have recently been described as a new species, *Rhizobium tropici* (35).

only because there existed substantial genetic variation at these scales, since recombination in the absence of such variation leaves no signature.

But, on closer inspection, our results indicate that geographic separation contributes substantially to the observed linkage disequilibrium (even though we used an index, V_o/V_e, that adjusts for the effect of allelic diversity). We also observed that allelic diversity itself tended to increase with geographic scale. These findings suggest that effective migration between geographically distant populations of this species may be limited and they are consistent with the observations of Piñero *et al.* (14), who rarely collected identical electrophoretic genotypes and never at very distant sites. If restricted migration between populations is indeed a major cause of linkage disequilibrium in *R. leguminosarum* bv. *phaseoli*, then this leaves open the possibility of more frequent genetic recombination within local populations.

Comparison with the Population Structure of *E. coli*. The evidence for restricted migration between local populations of *R. leguminosarum* bv. *phaseoli* contrasts markedly with the situation that has been documented in *E. coli*. The same multilocus electrophoretic genotypes of *E. coli* are overrepresented in widely distant regions of the world, and rates of migration for this species have been inferred to be sufficient to prevent substantial geographical divergence in allele frequencies (2, 4, 5, 13). For *E. coli*, it appears that linkage disequilibrium is due primarily to limited recombination among otherwise clonal lineages rather than to restricted migration.

Further evidence for limited recombination within *E. coli* populations is that linkage disequilibrium coefficients are not only statistically significant but are also near their theoretical maxima given observed allele frequencies (i.e., $|D'| \approx 1$), even when the sample is geographically restricted (3, 31). We have calculated values of $|D'|$ for 542 isolates of *E. coli* that were obtained from one human host by Caugant *et al.* (3). (In that study, 10 or more isolates were analyzed from each of 22 samples over an 11-month period; two or more electrophoretically distinct genotypes were observed in 16 of the samples. In our analysis, we have pooled across sampling dates in order to increase sample sizes; in principle, temporal isolation of genotypes, like spatial isolation, can contribute to linkage disequilibrium. The fact that genetically mixed infections were evident in so many of the samples, however, mitigates this concern.) In our calculations, we have used only those loci that have two alleles each with a frequency of at least 20%, because estimates of $|D'|$ are highly dependent on sample size when the expected frequency of any one of the four two-locus haploid genotypes is low. Ten of the 13 polymorphic loci surveyed by Caugant *et al.* (3) meet this criterion, yielding $10(9)/2 = 45$ estimates of $|D'|$. All 45 values are >0.5, and 38 are >0.95 (Fig. 1), indicating nearly complete linkage disequilibrium that is consistent with only very limited recombination. The most common genotype differed from each of the other isolates by 6 or more of the 13 polymorphic loci surveyed, providing compelling evidence for the paucity of recombination (3).

By contrast, when we apply this same criterion to the 190 isolates of *R. leguminosarum* bv. *phaseoli* from the cultivated plot at Santiago Tepetlapa, four of the nine polymorphic loci qualify and none of the six resulting $|D'|$ values is >0.25 (Fig. 1). Thus, although multilocus linkage disequilibrium is statistically significant within this plot (Table 1, row B), recombination does not appear to be so infrequent as to cause severe shortages of any two-locus haploid genotypes. Evidently, the amount of migration between populations and the frequency of mixis within populations pictured for *E. coli* do not fit well with our observations on *R. leguminosarum* bv. *phaseoli*.

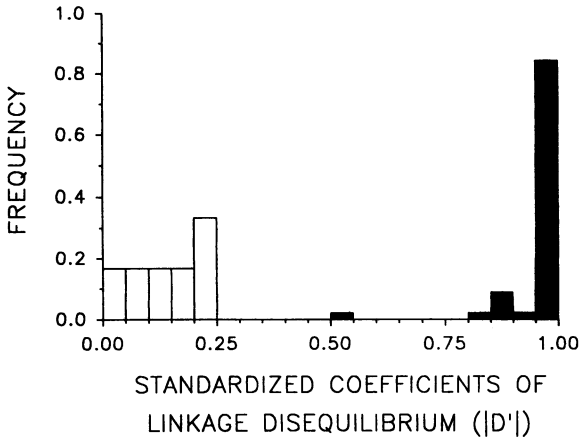

FIG. 1. Comparison of the distribution of standardized coefficients of linkage disequilibrium, $|D'|$, between local populations of *R. leguminosarum* bv. *phaseoli* (open bars) and *E. coli* (filled bars). The former was sampled from a cultivated plot and the latter was from an individual human. Values for $|D'|$ are shown only for those loci that have two alleles each with a frequency of at least 20%, because estimates of $|D'|$ are highly dependent on sample size when the expected frequency of any one of the four two-locus haploid genotypes is low. See text for additional details.

Limitations of the Present Analyses. Our results indicate geographic structuring of populations of *R. leguminosarum* bv. *phaseoli*: allelic diversity and linkage disequilibrium (adjusted for that diversity) increase with scale. Furthermore, our analyses show that the population genetic structure of this species is quite different from that observed for *E. coli*; in particular, standardized linkage disequilibrium coefficients appear to be consistent with frequent mixis in local populations of *R. leguminosarum* bv. *phaseoli*.

Our statistical analyses cannot, however, formally exclude explanations other than frequent mixis for the relatively low linkage disequilibrium in local populations of *R. leguminosarum* bv. *phaseoli*. For example, if this species has a very high mutation rate, then independent mutations to indistinguishable electrophoretic mobility states might be confused with mixis. We do not know of any reason to suspect that *Rhizobium* species have unusually high mutation rates. Moreover, if homoplasies were so common as to obscure linkage disequilibrium in local populations, then we would not expect to have discerned an increase in allelic diversity (which is limited by the number of distinguishable electrophoretic alleles) with geographic scale. Direct evidence bearing on the frequency of recombination within local populations of *Rhizobium* could be provided by experiments, similar to those performed by Graham and Istock (38, 39) using *Bacillus subtilis*, in which genotypes differing at several loci are introduced into a microcosm and the appearance of recombinant genotypes is monitored.

Gene Phylogenies and Species Concepts. Maynard Smith *et al.* (16) have recently discussed the apparent discrepancy between electrophoretic studies of bacterial populations, which demonstrate strong linkage disequilibrium, and DNA sequence data, which indicate a mosaic structure within certain genes. They suggest that the DNA studies provide evidence for "localized sex," that is occasional recombination involving small, localized regions of the genome (see also refs. 20 and 40). Our results suggest another way in which sex in bacteria may be regarded as localized. If there is limited migration between bacterial populations, then geographic isolation can give rise to strong linkage disequilibrium among isolates from different populations, even though recombination within local populations may be rather frequent.

Piñero *et al.* (14) hypothesized that *R. leguminosarum* bv. *phaseoli* is a polyphyletic assemblage, based on the very high

genetic diversity among isolates from throughout the Western Hemisphere as well as the fact that host-range specificity in *Rhizobium* (which is used to define biovars) is determined by horizontally transmissible plasmids. They called for a reclassification of the genus based on variation in chromosomal genes, which would better reflect evolutionary descent. The evidence that we have presented here suggests that patterns of chromosomal variation in *Rhizobium* may be reticulate rather than strictly clonal. Such reticulation may be useful in defining nonarbitrary boundaries between species; Dykhuizen and Green (20) have recently proposed operationally defining bacterial species on the basis of evidence for recombination of chromosomal genes. In particular, they propose that (20) "The phylogeny of different genes from individuals of the same species should be significantly different, whereas the phylogeny of genes from individuals of different species should not be significantly different." This criterion is attractive in that it reflects the notion of a shared gene pool within—but not between—species, and hence it conforms to the biological species concept (41).

We thank Francisco Ayala, Julia Bell, Michael Clegg, David Demezas, Daniel Dykhuizen, Luis Eguiarte, Bruce Levin, Michel Tibayrenc, and two anonymous reviewers for helpful discussion and comments on the manuscript. V.S. was supported by a scholarship from Universidad Nacional Autonóma de México; R.R.H. and R.E.L. were supported by grants from the National Institutes of Health and the National Science Foundation, respectively.

1. Levin, B. R. (1988) in *The Evolution of Sex*, eds. Michod, R. E. & Levin, B. R. (Sinauer, Sunderland, MA), pp. 194–211.
2. Selander, R. K. & Levin, B. R. (1980) *Science* **210**, 545–547.
3. Caugant, D. A., Levin, B. R. & Selander, R. K. (1981) *Genetics* **98**, 467–490.
4. Whittam, T. S., Ochman, H. & Selander, R. K. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1751–1755.
5. Whittam, T. S., Ochman, H. & Selander, R. K. (1983) *Mol. Biol. Evol.* **1**, 67–83.
6. Ochman, H. & Selander, R. K. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 198–201.
7. Musser, J. M., Granoff, D. M., Pattison, P. E. & Selander, R. K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5078–5082.
8. Selander, R. K., McKinney, R. M., Whittam, T. S., Bibb, W. F., Brenner, D. J., Nolte, F. S. & Pattison, P. E. (1985) *J. Bacteriol.* **163**, 1021–1037.
9. Young, J. P. W. (1985) *J. Gen. Microbiol.* **131**, 2399–2408.
10. Achtman, M. & Pluschke, G. (1986) *Annu. Rev. Microbiol.* **40**, 185–210.
11. Caugant, D. A., Mocca, L. F., Frasch, C. E., Frøholm, L. O., Zollinger, W. D. & Selander, R. K. (1987) *J. Bacteriol.* **169**, 2781–2792.
12. Musser, J. M., Bemis, D. A., Ishikawa, H. & Selander, R. K. (1987) *J. Bacteriol.* **169**, 2793–2803.
13. Selander, R. K., Caugant, D. A. & Whittam, T. S. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington), pp. 1625–1648.
14. Piñero, D., Martínez, E. & Selander, R. K. (1988) *Appl. Environ. Microbiol.* **54**, 2825–2832.
15. Harrison, S. P., Jones, D. G. & Young, J. P. W. (1989) *J. Gen. Microbiol.* **135**, 1061–1069.
16. Maynard Smith, J., Dowson, C. G. & Spratt, B. G. (1991) *Nature (London)* **349**, 29–31.
17. Tibayrenc, M., Ward, P., Moya, A. & Ayala, F. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 115–119.
18. Tibayrenc, M. & Ayala, F. J. (1988) *Evolution* **42**, 277–292.
19. Tibayrenc, M., Kjellberg, F. & Ayala, F. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2414–2418.
20. Dykhuizen, D. E. & Green, L. (1991) *J. Bacteriol.* **173**, 7257–7268.
21. Martínez, E., Romero, D. & Palacios, R. (1990) *Crit. Rev. Plant Sci.* **9**, 59–93.
22. Schofield, P. R., Gibson, A. H., Dudman, W. F. & Watson, J. M. (1987) *Appl. Environ. Microbiol.* **53**, 2942–2947.
23. Young, J. P. W. & Wexler, M. (1988) *J. Gen. Microbiol.* **134**, 2731–2739.
24. Buchanan-Wollaston, V. (1979) *J. Gen. Microbiol.* **112**, 135–142.
25. Bromfield, E. S. P., Sinha, I. B. & Wolynetz, M. S. (1986) *Appl. Environ. Microbiol.* **51**, 1077–1084.
26. Charles, T. C., Singh, R. S. & Finan, T. M. (1990) *J. Gen. Microbiol.* **136**, 2497–2502.
27. Bullerjahn, G. S. & Benzinger, R. H. (1982) *J. Bacteriol.* **150**, 421–424.
28. Souza, V. (1990) Ph.D. thesis (Universidad Nacional Autonóma de México, México City).
29. Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N. & Whittam, T. S. (1986) *Appl. Environ. Microbiol.* **51**, 873–884.
30. Lewontin, R. C. (1964) *Genetics* **49**, 49–67.
31. Hedrick, P. W. & Thomson, G. (1986) *Genetics* **112**, 135–156.
32. Brown, A. H. D., Feldman, M. W. & Nevo, E. (1980) *Genetics* **96**, 523–536.
33. Brown, A. H. D. & Feldman, M. W. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5913–5916.
34. Sokal, R. R. & Rohlf, F. J. (1981) *Biometry* (Freeman, San Francisco).
35. Martínez-Romero, E., Segovia, L., Mercante, F. M., Franco, A. A., Graham, P. & Pardo, M. A. (1991) *Int. J. Syst. Bacteriol.* **41**, 417–426.
36. Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Plans* (Soc. Ind. and Appl. Math., Philadelphia).
37. Tiedje, J. M., Colwell, R. K., Grossman, Y. L., Hodson, R., Lenski, R. E., Mack, R. N. & Regal, P. J. (1989) *Ecology* **70**, 298–315.
38. Graham, J. B. & Istock, C. A. (1978) *Mol. Gen. Genet.* **166**, 287–290.
39. Graham, J. B. & Istock, C. A. (1979) *Science* **204**, 637–639.
40. Milkman, R. & Bridges, M. M. (1990) *Genetics* **126**, 505–517.
41. Mayr, E. (1942) *Systematics and the Origin of Species* (Columbia Univ. Press, New York).