J. evol. biol. 11 (1998) 549–562
1010–061X/98/050549–14 $ 1.50 + 0.20/0

Journal of Evolutionary Biology

Geographic structure of lineage associations in a plant-bacterial mutualism

M. A. Parker* and J. M. Spoerke

Department of Biological Sciences, State University of New York, Binghamton, New York 13902, USA

Key words: Bradyrhizobium; coevolution; geographic variation; Leguminosae; mutualism.

Abstract

Two species linked by a mutualistic relationship may evolve correlated population differentiation if there is long-term continuity of interactions between specific partners. This phenomenon was analyzed by multilocus enzyme electrophoresis on the annual legume *Amphicarpaea bracteata* and its nitrogen-fixing bacterial symbionts (*Bradyrhizobium* sp.) sampled from >20 sites over a 1000 km area. Three analyses indicated that genetic differentiation was correlated in the two organisms. First, the genetic distance between bacterial populations at each pair of sites was significantly positively related to the genetic distance between their host plant populations, as evaluated by the Mantel test. Second, a cluster analysis revealed that several divergent lineages were present both among plants and among bacteria. Bacterial lineages showed a highly nonrandom distribution across plant lineages that was consistent in each of two regions sampled. Finally, there were numerous cases where populations of the same plant lineage 1000 km apart harbored bacterial isolates with an identical multilocus genotype. Thus, despite recurrent opportunities for partner switching, particular genotypes of these two organisms associate consistently across multiple habitats throughout their geographic range.

Introduction

The comparison of geographic variation in two symbiotically associated species can reveal a great deal about the process of coevolution shaping their interaction. An absence of correlated genetic variation implies that dispersal and population extinction have proceeded largely independently in each organism, so that any pair

* Author for correspondence, e-mail: mparker@binghamton.edu

of symbiotic partners have only an ephemeral relationship. By contrast, a pattern of spatially correlated differentiation suggests long-term historical continuity of symbiotic interactions. This may involve a combination of both nonadaptive processes and natural selection. For example, if the scale of gene flow is similarly restricted in both species (e.g., because one disperses by attachment to the other), then random population differentiation may evolve in parallel in both species according to an isolation by distance model (Slatkin, 1993). Alternatively, natural selection may generate nonrandom associations between particular genotypes of hosts and symbionts (Frank, 1994).

For mutualistic symbioses within natural communities, no studies to date have analyzed the extent or causes of correlated geographic differentiation. In this study, we compared the geographical population structure of the annual legume *Amphicarpaea bracteata* and its nitrogen-fixing bacteria (genus *Bradyrhizobium*). Several processes affect spatial genetic variation within legume-rhizobial symbioses. First, seedlings must reacquire bacteria from the environment at the start of their life cycle (Sprent and de Faria, 1988), so the opportunity exists for partner switching each generation. Second, some root-nodule bacteria show very broad host ranges (Law, 1985), implying that there will often be few constraints on bacterial transfer between unrelated hosts. In opposition to these factors that act to randomize genotypic associations, natural selection may commonly promote correlated spatial variation. Fitness of both legumes and rhizobia can vary greatly with different symbiotic partners (e.g., Chanway et al., 1991; Parker, 1995). Thus, selection at each site may favor organisms that are most compatible with locally prevalent partner genotypes. Little is currently known about how these conflicting processes shape the evolution of geographic variation within legume-rhizobial symbioses.

A. bracteata reproduces primarily by self-fertilization (Schnee and Waller, 1986), and like many inbred plants, local populations are often dominated by a small number of homozygous genotypes, which may be strongly differentiated between nearby sites (Parker, 1992). Isozyme studies have shown that populations of *A. bracteata* cluster into three distinct groups (lineages Ia, Ib, and II), each of which has a wide geographic distribution (Fig. 1). Lineages Ia and Ib are genetically similar, sharing at least some alleles in common at most loci. However, lineage II plants shared no alleles in common with lineages Ia or Ib at 7 out of 18 loci examined, and may thus represent a distinct, previously unrecognized species (Parker, 1996). In a study of one habitat where plants of lineages Ia and II coexisted, these hosts were found to harbor highly distinct *Bradyrhizobium* genotypes (Spoerke et al., 1996).

In the current study, we greatly extended our sampling in order to test whether particular genotypes of plants and bacteria also associated consistently on a geographic scale. We used multilocus enzyme electrophoresis to examine three questions. First, does variation within either organism conform to an isolation by distance model, where genetic divergence is positively correlated with geographic distance among sites? Second, do populations of plants and bacteria show parallel differentiation? Finally, when a plant lineage is distributed across many sites, does it associate with the same bacterial genotypes in each area?

Materials and methods

The root nodule bacteria associated with *A. bracteata* are in the genus *Bradyrhizobium*, but appear to be distinct from currently recognized species in this genus (Spoerke et al., 1996). Isozyme data (Marr et al., 1997) and RFLP analyses (Parker, personal observations) suggest that these bacteria are more similar to *B. elkanii* (Kuykendall et al., 1992) than to *B. japonicum* (Young and Haukka, 1996), but are genetically divergent from both taxa. Both *B. elkanii* and *B. japonicum* are effective nitrogen-fixing symbionts of soybean (*Glycine max*), whereas bacterial isolates from *A. bracteata* can form nodules on soybean, but the nodules do not fix nitrogen (Marr et al., 1997). Apart from soybean, little is known about the nodulation host

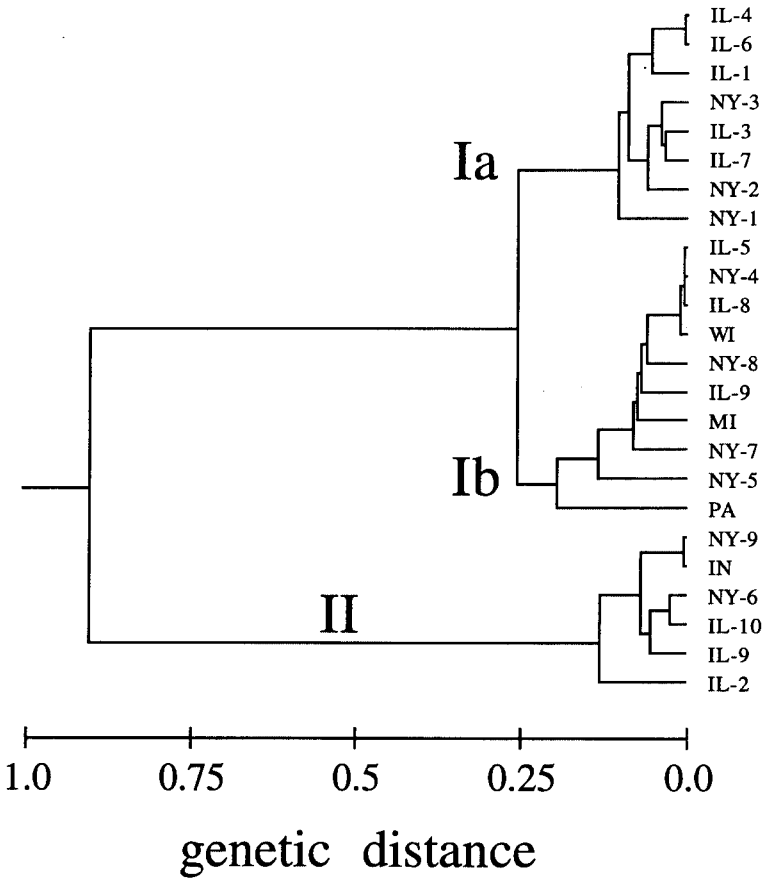

Fig. 1. Relatedness of *A. bracteata* populations estimated by enzyme electrophoresis (adapted from Parker [1996]). Approximately 30 plants per population were analyzed for variation at 18 enzyme loci, and populations were clustered by UPGMA (Sneath and Sokal 1973) after calculating pairwise genetic distances (Nei 1978).

range of bacteria associated with *A. bracteata*, although Wilson (1939) reported that some isolates from *A. bracteata* could form nodules on certain other species of native North American legumes.

Bacteria were sampled from root nodules collected in 21 sites across *A. bracteata's* range in eastern North America. Eleven sites in Wisconsin, Illinois, Indiana, and Michigan (0.07 km to 287 km apart) are subsequently termed the "midwestern region", while ten populations in New York and Pennsylvania (0.04 km to 133 km apart) are termed the "eastern region". Pairs of eastern and midwestern sites were separated by 850–1150 km.

In a previous study, plants from all of these sites were characterized for isozyme variation at 18 loci (Parker, 1996). At all but one site, only a single plant lineage was present. At a site in McHenry County, IL (the "IL-9" site), plants of lineages Ib and II coexisted with negligible hybridization (Parker, 1994). These two groups were thus treated as separate biological populations in the cluster analysis (Fig. 1), and nodules were sampled separately from each plant lineage at this site ($n = 10$ plants per lineage). A total of 10 plants were sampled in each of the other sites. A single bacterial isolate was purified from one haphazardly selected nodule per plant as described in Spoerke et al. (1996). One isolate was inadvertently lost from the Michigan site, so the total number of isolates obtained for electrophoretic analysis was 219 (10 for each plant lineage at IL-9, 9 for the MI site, and 10 each for the remaining 19 sites). For several analyses described below, we supplemented this data set with previously published results for 51 isolates sampled from two Illinois sites (Spoerke et al., 1996; designated IL-1 and IL-2 in Fig. 1), yielding a total sample size of 270 isolates from 24 host populations.

Bacterial isolates were characterized by starch gel electrophoresis at the following 20 enzyme loci as described in Spoerke et al. (1996): acid phosphatase (ACP), alanine dehydrogenase (ALA), butyrate esterase (EST), β-hydroxybutyrate dehydrogenase (HBD), diaphorase (DIA), fumarase (FUM), fructose-1,6-diphosphatase (F16), glucose-6-phosphate dehydrogenase (G-6), glutamic-oxalacetic transaminase (GOT), glyceraldehyde-3-phosphate dehydrogenase type 1 (GP1), isocitrate dehydrogenase (IDH), indophenol oxidase (IPO), leucine aminopeptidase (LAP), leucine tyrosine peptidase (PEP), malate dehydrogenase (MDH), malic enzyme (ME), phosphoglucose isomerase (PGI), phosphoglucomutase (PGM), 6-phosphogluconate dehydrogenase (6PGD), and shikimate dehydrogenase (SDH). Each isolate was characterized by its allelic profile for the 20 enzymes, and each unique multilocus genotype was designated an electrophoretic type (ET). Pairwise genetic distances between ETs were estimated by the proportion of enzyme loci at which allelic differences occurred. ETs were then clustered by UPGMA (Sneath and Sokal, 1973).

The relationship between plant and bacterial divergence was analyzed by two complementary approaches. First, after identifying lineages from cluster analyses, chi-square tests were used to determine whether bacterial lineages were nonrandomly distributed across plant lineages. Because the number of lineages distinguished by cluster analysis is to some extent arbitrary, a second analysis was performed that used populations as units of observation. The degree of correlated

population differentiation was analyzed by a variant of the Mantel test (Douglas and Endler, 1982; Dietz, 1983). For each organism, we determined Nei's genetic distance between each pair of populations (24 populations yields 276 pairwise distances), and then calculated the Spearman rank correlation (r_s) between plant distances and bacterial distances. The distribution of r_s expected under the null hypothesis of no correlation was determined by calculating r_s 2000 times after randomly reordering entries in one of the data sets. The probability of obtaining a particular r_s value by chance alone was calculated by comparing the observed r_s to the distribution derived from random permutations of the original data. We also used this approach to analyze for each species whether a correlation existed between geographic distance among sample sites and genetic distance.

Results

Bacterial diversity

Among the 219 bacterial isolates sampled, variation was detected at all 20 enzyme loci examined, with a mean of 4.7 alleles per locus (range, 2 to 10). Five loci were only marginally polymorphic (GOT, GPI, HBD, MDH, SOD), with a common allele occurring in more than 214 out of the 219 isolates. A total of 50 distinct multilocus genotypes (ETs) was observed among the 219 isolates. However, nearly half of the ETs (24/50) were represented by only a single isolate which showed only one allelic difference from related more common ETs. For brevity, these minor variant genotypes were omitted from allele profile summary (Tab. 1) and dendrogram of ET relationships (Fig. 2; complete data on ET allelic profiles are available from M. Parker on request).

Clustering of ETs based on pairwise allelic differences revealed the presence of three divergent bacterial lineages (designated A, B, and C; Fig. 2). Each bacterial lineage had a wide geographic distribution, occurring in both the midwestern and eastern regions (Tab. 1). These correspond to the three lineages observed previously in two Illinois sites (Spoerke et al., 1996). Pooling data from these two sites with the current sample of 22 populations, most of the local plant populations examined (13/24) had more than one bacterial lineage present. At five of these sites, all three bacterial lineages co-occurred locally. Lineage C was most abundant overall, represented by 53% of all isolates (143/270), and was present in 19 of the 24 populations sampled. This group of isolates was relatively homogeneous genetically (mean pairwise allelic difference among ETs within lineage = 13%), and was associated with all three lineages of *A. bracteata* hosts (Fig. 2). Lineage B was the least common of the three main bacterial lineages (51/270 isolates), and was also the most heterogenous (mean within-lineage pairwise allelic difference among ETs = 31%).

Table 1. Allele profiles at 20 enzyme loci for 26 electrophoretic types (ETs) of *Bradyrhizobium* isolated from *A. bracteata*. For each locus, alleles are listed according to relative anodal migration speed (1 = fastest; null alleles are represented by zeros).

	Electrophoretic type (ET)																								
Enzyme	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
ACP	3	3	3	5	1	1	1	2	2	2	2	6	5	5	4	5	5	5	5	5	5	5	5	5	5	5
ALA	3	3	3	3	3	3	0	3	3	3	3	3	2	2	2	2	2	2	1	2	2	2	2	2	2	4
DIA	2	1	1	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3
EST	6	6	6	3	2	2	2	1	1	1	1	2	4	4	4	4	4	4	4	4	4	4	4	4	4	5
F16	4	4	4	2	1	1	1	5	3	5	5	5	2	2	2	2	2	2	2	2	2	2	2	2	2	2
FUM	1	1	1	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2
G-6	5	5	5	5	6	6	6	2	2	0	1	6	3	3	3	3	3	3	3	4	0	0	3	3	3	3
GOT	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2
GPI	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
HBD	2	2	2	2	2	2	2	2	2	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
IDH	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
IPO	2	2	2	2	2	2	1	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
LAP	2	2	0	1	2	2	2	2	2	2	2	4	2	3	2	2	2	2	2	2	2	2	2	2	3	2
PEP	0	4	4	4	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	0	2
MDH	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	2	2
ME	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	4	3	3	3	5	3	3	3	3
PGI	2	2	2	2	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
PGM	3	3	3	3	6	6	6	4	4	4	4	4	5	5	5	1	2	5	5	5	5	5	5	5	5	5
6PGD	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SDH	3	3	3	3	4	4	4	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Total isolates	30	4	3	4	4	3	1	11	2	3	2	2	62	16	9	4	7	2	2	4	4	2	7	2	4	1
Number of populations ET was detected in:																										
midwest*	4	1	1	2	2	2	1	2	0	1	1	1	8	5	0	2	0	0	1	0	1	2	2	0	2	1
east	3	2	0	0	0	0	0	3	1	0	0	0	8	2	3	1	3	1	0	2	0	0	3	2	0	0

* 12 total populations sampled in midwestern region and 10 in eastern region.

Bacterial distribution across host lineages

In both geographic regions sampled, there was a highly nonrandom distribution of bacterial lineages across host plant lineages (Fig. 3; χ^2 [4 d.f.] = 71.4 and 95.3, respectively, for the midwestern and eastern regions, $P < 0.001$). Bacterial lineage A was the most common group on plants of lineage Ia, but was rarely found on the other two host lineages. Instead, bacterial lineage C dominated on plant lineages Ib and II in both geographic regions.

In the IL-9 site where plant lineages Ib and II coexisted, there was considerable overlap of bacterial genotypes associated with the two types of hosts. Three of the six ETs detected at this site were shared in common by both plant lineages, and all of the remaining bacterial genotypes were rare ETs that together represented only 5 of the 20 isolates sampled. This is consistent with the overall pattern of high resemblance between the bacteria on these two lineages across all sites (Figs. 2 and 3).

Eight bacterial ETs were detected in both geographic regions, and these widely distributed bacterial genotypes were often associated with the same host lineage in each region (Tab. 2). For example, ET1 (bacterial lineage A, Fig. 2) was present in every lineage Ia host population sampled, and was never found in any other type of host population. The closely related ET2 was also found in each of two separate lineage Ia host populations both in IL and NY. ET8 (bacterial lineage B) was also restricted to host lineage Ia, and occurred in both IL and NY. The most common ET found in lineage Ib and lineage II host populations (ET13, bacterial lineage C) occurred in all 16 sites where these host lineages were sampled. Four other lineage C ETs were also shared in common between eastern and midwestern plant populations of lineages Ib and II (Tab. 2). Considering all 24 combinations of plant lineage and ET within Table 2 together, the fraction of each plant lineage's

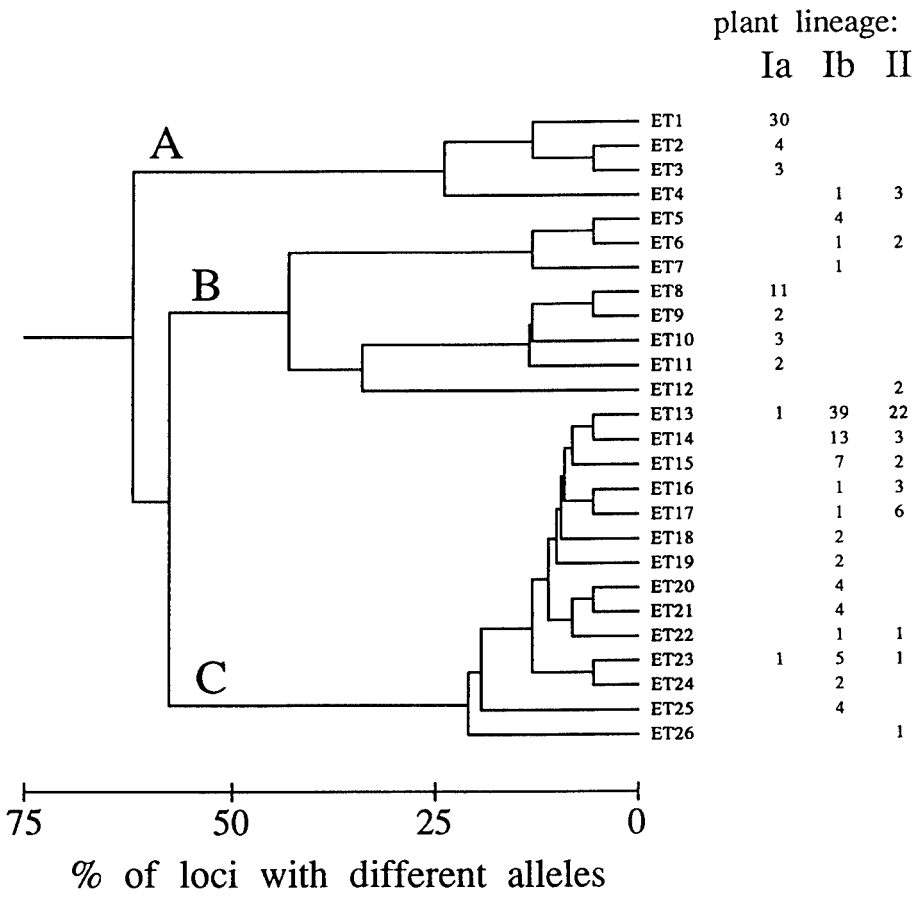

Fig. 2. Genetic relationships among 26 multilocus genotypes (ETs) of root nodule bacteria (*Bradyrhizobium* sp.) associated with *A. bracteata*. Listed on the right of the tree diagram is the number of bacterial isolates per ET observed from each plant lineage (for clarity, only non-zero values are presented).

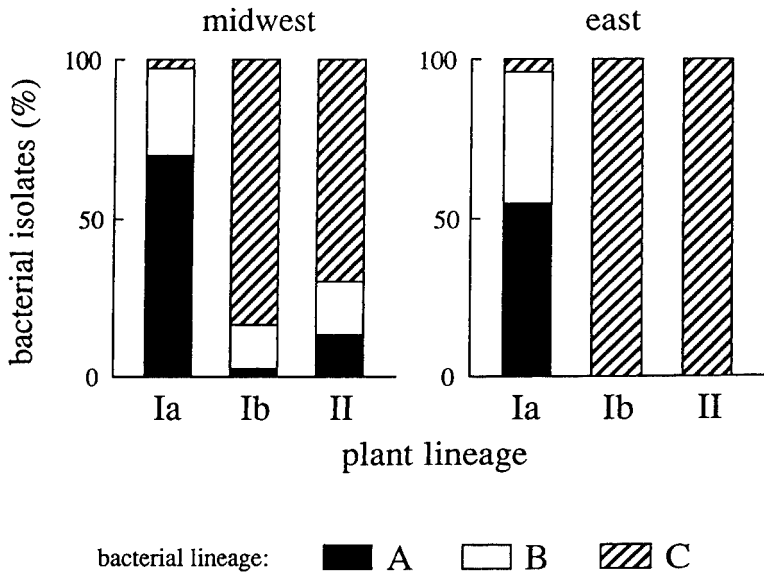

Fig. 3. Frequency of bacterial lineages on three plant lineages within two geographic regions. The total number of bacterial isolates sampled was 40, 49, and 30 respectively for the three host lineages in the midwestern region, and 30, 50, and 20 respectively for the eastern region.

populations occupied by these bacterial ETs was highly correlated in the midwestern vs. eastern regions (Spearman $r_s = 0.70$, $P < 0.0001$). The finding of identical ETs in both regions associated with each plant lineage suggests a high degree of persistence for specific plant/bacterial genotypic relationships.

Table 2. Bacterial ETs found in both eastern and midwestern North America. For each plant lineage, the percentage of populations within a region containing each bacterial ET is given. The total number of populations sampled for plant lineages Ia, Ib, and II was 5, 5, and 4 respectively in the midwestern region and 3, 5, and 2 in the eastern region.

	Populations where ET was detected (%)					
	Midwestern region			Eastern region		
	Plant lineage			Plant lineage		
Bacterial ET	Ia	Ib	II	Ia	Ib	II
ET1	100	0	0	100	0	0
ET2	40	0	0	67	0	0
ET8	40	0	0	100	0	0
ET13	0	100	100	33	100	100
ET14	0	60	50	0	40	0
ET15	0	0	25	0	40	50
ET16	0	20	50	0	0	50
ET23	20	20	0	0	40	50

Table 3. Genetic distance among pairs of plant populations and among pairs of bacterial populations grouped by spatial proximity (N = the number of pairs falling into each geographic distance category).

		Genetic distance (mean ±1 SE)	
	N	Plants	Bacteria
Within regions:			
Sites separated by:			
0 5 km	17	0.50 ± 0.11	0.28 ± 0.07
5– 25 km	17	0.55 ± 0.10	0.23 ± 0.08
25– 60 km	19	0.46 ± 0.09	0.33 ± 0.05
60–120 km	56	0.48 ± 0.05	0.41 ± 0.04
120–300 km	27	0.42 ± 0.07	0.31 ± 0.06
Between regions (>800 km)	140	0.44 ± 0.03	0.33 ± 0.03

Geographic structure

For both plants and bacteria, there was no correlation between genetic distance and the geographic distance between sample sites (plants: $r_s = -0.05$; bacteria: $r_s = 0.03$, $P > 0.35$). Indeed, pairs of sites within a region (less than 300 km apart) were as different from one another as pairs of sites > 800 km apart in the midwestern vs. eastern region (Tab. 3). Even among the closely spaced populations in our sample, there was little indication of a correlation between proximity and genetic similarity. For both organisms, pairs of sites 0 to 5 km apart did not have a noticeably lower mean genetic distance than pairs of more remote sites (Tab. 3). Thus, isolation by distance does not appear to be the primary factor structuring geographic variation within either organism.

The weak overall relationship of genetic divergence to geographic distance appears to be a consequence of the extensive geographic overlap of plant lineages (Fig. 1), together with the nonrandom distribution of bacterial lineages across plants (Fig. 3). Thus, for either organism, adjacent pairs of populations can be as genetically different as populations from widely separated sites. To discriminate within-lineage and between-lineage geographic structure, a further analysis was done by dividing the 24 populations into three groups according to plant lineage ($n = 8$ lineage Ia, 10 lineage Ib, and 6 lineage II sites; a parallel partition of sites according to bacterial lineage was not possible due to the erratic and overlapping distribution of bacterial lineages across sites [Tab. 1]). Mantel tests for a correlation between genetic distance and geographic distance were then done separately on plants and on bacteria within each group. All but one of the correlations were small and not significant; only for plants in lineage Ia did the relationship between genetic distance and geographic distance approach significance ($r_s = 0.28$, $P = 0.072$). These test have limited statistical power due both to the small number of sites sampled per group and the narrow range of genetic diversity within lineages. Nevertheless, the pattern suggested by current evidence is that genetic divergence is only weakly correlated with geography, both within plant lineages and for all sites pooled.

However, there was a significant overall correlation between population differentiation in plants and bacteria ($r_s = 0.14$, $P < 0.012$). This correlation was not unexpected, given that bacteria were nonrandomly distributed across plant lineages (Fig. 3), implying that divergent pairs of plant populations generally have dissimilar bacteria. Because the strength of this correlation was relatively weak, a further analysis was performed by classifying populations according to plant lineage. Pairs of plant populations from the same lineage harbored significantly more similar bacterial associates than population pairs involving separate plant lineages. The mean bacterial genetic distance among pairs of populations from the same host lineage was 0.09 (± 0.01, $n = 88$ pairs), compared to 0.45(± 0.02, $n = 188$) for pairs involving different plant lineages ($P < 0.002$ based on bootstrap resampling with 500 replicates [Efron and Gong, 1983]). However, there was little correspondence between divergence in the two organisms within the subset of comparisons involving separate plant lineages. For example, hosts from lineages Ia and Ib are closely related genetically (Fig. 1), yet these plants had highly divergent bacterial symbionts (Tab. 4). By contrast, plants of lineage Ib and lineage II are genetically very different, yet their bacterial mutualists were as similar as among population pairs from a single plant lineage. These results indicate that genetic divergence has not proceeded in a strictly parallel fashion in the two organisms.

Discussion

Despite a large body of research on agricultural species, many aspects of legume-rhizobial coevolution in natural communities remain very poorly understood. In particular, it is not yet clear whether these symbioses involve long-term relationships between specific partner genotypes that have highly correlated spatial distributions, or whether genotypic combinations are continually reshuffled across different habitats. This second scenario seems to describe certain agricultural populations, where plants introduced to a site may acquire a heterogeneous collection of symbiotic partners from the indigenous bacterial population. For example, Sullivan et al. (1995) argued that transfer of chromosomal nodulation genes from an inoculant strain to indigenous soil bacteria permitted a wide diversity of bacterial genotypes to become symbionts of a newly introduced host legume.

Table 4. Bacterial genetic distance across pairs of populations representing different plant lineages. Bootstrap resampling ($n = 500$ replicates) indicated that all three groups differed significantly in mean genetic distance ($P < 0.002$ for each comparison).

		Bacterial genetic distance		
Bacterial origin	N	Mean	SE	Range
Lineage Ia vs. Ib hosts	80	0.66	0.02	0.24–0.94
Lineage Ia vs. II hosts	48	0.55	0.02	0.25 0.88
Lineage Ib vs. II hosts	60	0.08	0.01	0–0.29

Our results, by contrast, provide strong evidence that plants do not acquire bacteria unselectively from populations indigenous to each habitat. If *A. bracteata* plants were sampling bacteria at random from a geographically varying pool, one would predict that nearby plant populations should by nodulated by genetically similar rhizobia, but there would be no reason to expect isolated populations of a plant lineage to share common bacterial genotypes. However, we found quite the opposite result. Bacterial genetic similarity was uncorrelated with spatial proximity, yet populations of each host lineage had similar bacteria in all areas sampled (Tab. 2, Fig. 3). Thus, this mutualism is characterized by strong repeatability of association between specific partner genotypes.

Given that bacteria are not hereditarily transmitted from parent plants to their offspring, independent dispersal processes in each species would tend to randomize any genotypic associations over time. Thus, mechanisms must exist to promote selective association of particular plants and bacteria. One such mechanism involves differential nodule formation ability. Laboratory inoculation studies have shown that certain bacterial ETs virtually lack the ability to form nodules on lineage Ib and lineage II plants. Interestingly, these specialized isolates occur within two separate bacterial lineages. Within lineage A, ET1 and ET2 are specialized for nodulation of lineage Ia hosts (Spoerke et al., 1996), and certain isolates from bacterial lineage B show identical nodulation specificity (Wilkinson et al., 1996). This may explain the low frequency of these two bacterial lineages on lineage Ib and lineage II hosts (Fig. 3). All other bacterial ETs that have been assayed (including ET4 of lineage A, most ETs from lineage B, and all lineage C ETs) show unspecialized phenotypes, forming abundant nodules on all three *A. bracteata* lineages. Differences among plants in the ability to interact with specialized bacterial isolates are controlled by alleles at a single locus (Parker and Wilkinson, 1997). The bacterial genes affecting nodulation specificity have not yet been investigated.

A second possible factor contributing to spatially correlated variation is that migration of the two organisms may not always occur independently. *A. bracteata* produces small seeds above ground, and also forms very large subterranean seeds on axillary shoots at the base of the plant (Schnee and Waller, 1986). The subterranean seeds are consumed by several species of birds and mammals (Martin et al., 1951), which may incidentally act as seed dispersal agents. Because subterranean fruits harbor rhizobia on their outer surface that are acquired from the soil in the vicinity of the parent plant, dispersal of subterranean seeds may result in establishment of new populations that preserve symbiotic relationships identical to those of the source site.

A final factor likely to promote spatially correlated variation is differential invasibility of sites due to competition. In the presence of specialized rhizobial isolates, lineage Ia plants show substantially higher growth and lifetime seed production than other *A. bracteata* lineages (Parker, 1995; Wilkinson and Parker, 1996). However, the competitive superiority of lineage Ia plants can be reversed by exposure to different bacterial genotypes (Parker, 1995; Wilkinson and Parker, 1996). Sites dominated by a mutually coadapted population of plants and bacteria

may therefore be resistant to invasion by other host genotypes. Among bacteria, genotypes also appear to vary in competitive fitness on different host genotypes (Wilkinson, 1995). Thus, the extant spatial distribution of plant genotypes may determine which bacteria can successfully colonize a site.

While we observed a strong relationship between variation in plants and bacteria, it is important to emphasize that we did not find strictly parallel divergence in the two organisms. The two most similar plant lineages (Ia and Ib) had very divergent bacterial symbionts, and two distantly related plant lineages (Ib and II) shared quite similar bacteria (Fig. 3, Tab. 4). Thus, our results rule out a simple model whereby bacteria evolve by continuously tracking host plant differentiation over time. Given the indiscriminate nodulation ability shown by certain bacterial isolates, the lack of strictly parallel differentiation is not unexpected. Indeed, the sharing of certain ETs by all three plant lineages (Fig. 2) implies that for some bacteria, there are few constraints limiting transfer between unrelated hosts. It is also possible that one or more of these bacterial lineages may not have differentiated directly in association with *A. bracteata*, but may instead have colonized this host after evolving on another legume species. We are currently testing this hypothesis by analyzing bacteria on related legume taxa occurring within *A. bracteata's* geographic range.

The frequent occurrence of bacterial isolates with an identical multilocus genotype at sites >1000 km apart was noteworthy (Tab. 2). Given the high overall genetic diversity that we observed (a total of 94 alleles across 20 enzyme loci), it is unlikely that two unrelated bacteria would convergently acquire precisely the same combination of alleles at every locus. Thus, it is reasonable to interpret pairs of isolates with the same ET as being clonal relatives. This inference is supported by recent studies indicating that isolates with the same ET usually show similar patterns of antibiotic sensitivity, regardless of geographic origin (M. Parker, unpublished data). If isolates with the same ET sampled 1000 km apart truly represent clonal descendants of a single ancestor, this would imply a substantial level of gene flow relative to the time scale of genetic change caused by mutation or recombination. The processes responsible for gene flow over this spatial scale are unclear. Symbiotic nitrogen-fixing bacteria have been recovered from nests of birds that employ soil as a nest material (Paschke and Dawson, 1993), but it is questionable how often such agents would move bacteria more than a kilometer or so. Understanding the scale and mechanisms of gene flow in natural populations of root-nodule bacteria remain important problems for future research.

In any case, the fact that particular ETs were not only found at distant locations, but also associated with the same plant lineage in both regions (Tab. 2), suggests that certain combinations of partners in this mutualism have sustained a long history of interaction. The extent to which specific partner genotypes preserve a long-term relationship is a key issue for the dynamics of coevolution (Bull and Rice, 1991; Law, 1991; Frank, 1994; Thompson, 1994; Hammerstein and Hoekstra, 1995). High continuity should facilitate the evolution of finely-tuned co-operation, while frequent partner switching disrupts the opportunity for specialized coadaptation. Plant growth rate experiments have in fact shown that bacteria associated with

each *A. bracteata* lineage are superior mutualist partners for all plants of that lineage, regardless of geographic origin (Wilkinson et al., 1996). Further comparative studies of other taxa are important to establish whether the patterns that we observe in *A. bracteata* are a prevalent feature of mutualisms within natural communities.

Acknowledgments

We are grateful to H. H. Wilkinson for helpful discussions and laboratory assistance. This work was supported by NSF grant BSR-9020413 to M.P.

References

Bull, J. J. and W. R. Rice. 1991. Distinguishing mechanisms for the evolution of co-operation. J. Theor. Biol. 149: 63–74.

Chanway, C. R., R. Turkington and F. B. Holl. 1991. Ecological implications of specificity between plants and rhizosphere micro-organisms. Adv. Ecol. Res. 21: 121–169.

Dietz, E. J. 1983. Permutation tests for association between two distance matrices. Syst. Zool. 32: 21–26.

Douglas, M. E. and J. A. Endler. 1982. Quantitative matrix comparisons in ecological and evolutionary investigations. J. Theor. Biol. 99: 777–795.

Efron, B. and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Amer. Statistician 37: 36–48.

Frank, S. A. 1994. Genetics of mutualism: the evolution of altruism between species. J. Theor. Biol. 170: 393–400.

Hammerstein, P. and R. F. Hoekstra. 1995. Mutualism on the move. Nature 376: 121–122.

Kuykendall, L. D., B. Saxena, T. E. Devine and S. E. Udell. 1992. Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. Can. J. Microbiol. 38: 501–505.

Law, R. 1985. Evolution in a mutualistic environment. pp 145–170. *In* D. H. Boucher (Ed.), The Biology of Mutualism. Oxford University Press, New York.

Law, R. 1991. The symbiotic phenotype: origins and evolution. pp 57–71. *In* L. Margulis and R. Fester (Eds.), Symbiosis as a Source of Evolutionary Innovation. MIT Press, Cambridge, Massachusetts.

Marr, D. L., T. E. Devine and M. A. Parker. 1997. Nodulation restrictive genotypes of *Glyine* and *Amphicarpaea*: a comparative analysis. Plant Soil 189: 181–188.

Martin, A. C., H. S. Zim and A. L. Nelson. 1951. American Wildlife & Plants. Dover Publications Inc., New York.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Parker, M. A. 1992. Outbreeding depression in a selfing annual. Evolution 46: 873–841.

Parker, M. A. 1994. Evolution in natural and experimental populations of *Amphicarpaea bracteata*. J. evol. biol. 7: 567–579.

Parker, M. A. 1995. Plant fitness variation caused by different mutualist genotypes. Ecology 76: 1525–1535.

Parker, M. A. 1996. Cryptic species within *Amphicarpaea bracteata* (Leguminosae): evidence from isozymes, morphology, and pathogen specificity. Can. J. Bot. 74: 1640–1650.

Parker, M. A. and H. H. Wilkinson. 1997. A locus controlling nodulation specificity in *Amphicarpaea bracteata* (Leguminosae). J. Heredity 88: 449–453.

Paschke, M. W. and J. O. Dawson. 1993. Avian dispersal of *Frankia*. Can. J. Bot. 71: 1128–1131.

Schnee, B. K. and D. M. Waller. 1986. Reproductive behavior of *Amphicarpaea bracteata* (Leguminosae), an amphicarpic annual. Am. J. Bot. 73: 376–386.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47: 264–279.

Sneath, P. H. A. and R. R. Sokal. 1973. Numerical Taxonomy. W. H. Freeman, San Francisco.

Spoerke, J. M., H. H. Wilkinson and M. A. Parker. 1996. Nonrandom genotypic associations in a legume-*Bradyrhizobium* mutualism. Evolution 50: 146–154.

Sprent, J. I. and S. M. de Faria. 1988. Mechanisms of infection of plants by nitrogen fixing organisms. Plant Soil 110: 157–163.

Sullivan, J. T., H. N. Patrick, W. L. Lowther, D. B. Scott and C. W. Ronson. 1995. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc. Natl. Acad. Sci. USA 92: 8985–8989.

Thompson, J. N. 1994. The Coevolutionary Process. The University of Chicago Press, Chicago.

Wilkinson, H. H. 1995. Evolution of symbiotic compatibility in a legume-*Bradyrhizobium* mutualism. Dissertation, State University of New York, Binghamton, NY.

Wilkinson, H. H. and M. A. Parker. 1996. Symbiotic specialization and the potential for genotypic coexistence in a plant-bacterial mutualism. Oecologia 108: 361–367.

Wilkinson, H. H., J. M. Spoerke and M. A. Parker. 1996. Divergence in symbiotic compatibility in a legume-*Bradyrhizobium* mutualism. Evolution 50: 1470–1477.

Wilson, J. K. 1939. Leguminous plants and their associated organisms. Cornell Univ. Agric. Exp. Sta. Mem. 221: 1–48.

Young, J. P. W. and K. E. Haukka. 1996. Diversity and phylogeny of rhizobia. New Phytol 133: 87–94.

Received 3 March 1997;
revised 22 April 1997;
accepted 3 July 1997.