**Rhizobium Population Genetics: Enzyme Polymorphism in Isolates from Peas, Clover, Beans and Lucerne Grown at the Same Site**

By J. P. W. YOUNG

John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

(Received 11 February 1985; revised 9 May 1985)

Isolates of *Rhizobium meliloti* and of *R. leguminosarum* biovars *viceae*, *trifolii* and *phaseoli* were obtained from a single site in Norfolk, England, and examined by enzyme electrophoresis of glucose-6-phosphate dehydrogenase, superoxide dismutase, β-galactosidase and esterases. All the enzymes had mobility variants, but the variation of the different enzymes was highly correlated, so that only a restricted number of combinations (electrophoretic types, ETs) were found. Some ETs were confined to a single biovar, but others were shared amongst *R. leguminosarum* biovars *viceae*, *trifolii* and *phaseoli*. Most *R. meliloti* isolates were quite distinct from those of *R. leguminosarum*. Electrophoresis of total soluble protein in the presence of sodium dodecyl sulphate revealed variation that correlated with the ET rather than with the host range. It is suggested that the *Rhizobium* isolates from this site comprise a number of genetically distinct lineages, some of which may carry any of several different host-range determinants, and which therefore appear in more than one biovar.

**INTRODUCTION**

*Rhizobium* species, by definition, form nitrogen-fixing nodules on the roots of legumes. They show specificity for the host plant genus, and this has long been the basis for their taxonomy. Until very recently, the classification proposed by Baldwin & Fred (1929) has remained unchanged. They defined *R. leguminosarum* as the symbiont of peas (*Pisum satiium*), *R. trifolii* of clover (*Trifolium spp.*), *R. phaseoli* of beans (*Phaseolus vulgaris*) and *R. meliloti* of lucerne (*Medicago sativa*). On the whole, strains of these four types do indeed have distinct, non-overlapping host ranges, but is this sufficient to demarcate them as separate species?

Numerical taxonomy has repeatedly shown that symbionts of lucerne have many distinctive characteristics, but the symbionts of peas, of clover, and of beans can only be reliably distinguished on the basis of their host range (Graham, 1964; 't Mannetje, 1967; Heberlein et al., 1967; Moffett & Colwell, 1968; Crow et al., 1981). Genetic studies have confirmed these findings, since the chromosomes of certain strains from peas, clover and beans can be recombined freely, but lucerne strains do not readily recombine with these (Johnston & Beringer, 1977; Kondorosi et al., 1980). Furthermore, the genetic determinants of symbiotic functions, including host range, have been shown to be plasmid-borne, and these determinants remain fully functional (confering host range and effective symbiosis) when transferred amongst strains from peas, clover and beans (Johnston et al., 1978; Hooykaas et al., 1981; Lamb et al., 1982). These observations have led to a reappraisal of the taxonomy. *R. meliloti* is retained as a distinct species, but the former species *R. leguminosarum*, *R. trifolii* and *R. phaseoli* are brought together as biovars *viceae*, *trifolii* and *phaseoli*, respectively, of a single species, *R. leguminosarum* (Jordan, 1984).

This new classification is based on laboratory observations. To what extent does it reflect the genetic structure of a mixed population in the field? In other words, are these taxonomic species...
also biological species? A biological species has been defined as 'a reproductive community of populations (reproductively isolated from others) that occupies a specific niche in nature' (Mayr, 1982, p. 273). This definition was framed primarily with higher eukaryotes in mind, but it can usefully be applied to bacteria provided recombination mechanisms exist, and provided we recognize that different parts of the genome (chromosome, plasmids) may experience different barriers to recombination. The essence of the definition is permanent reproductive isolation, which will lead to discontinuities at the boundaries between species in the distribution of genetic variation at polymorphic loci, regardless of whether these loci are directly involved in ecological differentiation. Genetic polymorphism thus provides a powerful tool for the analysis of population structure, and modern biochemical methods allow us to discern polymorphism in virtually any organism.

The most popular tool for studying biochemical polymorphism has been enzyme electrophoresis, which separates allelic forms of enzyme molecules by differences in their surface charge. Identity of surface charge does not, of course, necessarily imply identity of the corresponding nucleotide sequence, so each mobility class may include undetected genetic variants. The possible selective value of different electrophoretic variants has been the subject of considerable debate (see Dykhuizen & Hartl, 1983, for a bacterial example). However, these two considerations do not invalidate the use of enzyme electrophoresis to investigate the genetic structure of populations, and the technique has been applied very successfully and widely in eukaryotes. Its use with bacteria has been very limited, however, being primarily restricted to taxonomic questions (for example Bowman et al., 1967; Baptist et al., 1969; Cocks & Wilson, 1972; Foissy, 1974; Goullet, 1980). The only major effort in bacterial population genetics has been in Escherichia coli, which has been shown by enzyme electrophoresis, corroborated by other methods such as serology and DNA homology, to comprise a small number of cosmopolitan lineages, well differentiated genetically from each other, and each represented by a cluster of closely related strains (Milkman, 1973, 1975; Selander & Levin, 1980; Caugant et al., 1981, 1983; Whittam et al., 1983; Ochman et al., 1983; Ochman & Selander, 1984a).

There is abundant evidence for genetic variation within Rhizobium species (reviewed by Young, 1985), and several authors have reported enzyme electrophoretic variants (Fottrell & O'Hora, 1969; Murphy & Masterson, 1970; Mytton et al., 1978; Engvild & Neilson, 1984), suggesting that, with sufficiently large samples systematically collected, it should be possible to use the technique to investigate population structure. The results of such an investigation are reported in this paper.

**METHODS**

*Location.* The sampling site was a field at Grove Farm, Hainford, Norfolk, England (map reference TG 233177). In the year of sampling (1982) the field was sown with a commercial crop of French beans (*Phaseolus vulgaris*). In 1978 this field also carried French beans, and was the location of yield trials that included inoculation and sampling of natural *Rhizobium leguminosarum* biovar *phaselii* isolates (Beynon, 1980; Beynon & Josey, 1980). The present study site was approximately 300 m from that used by Beynon. Non-legume crops were grown in 1979–1981. The field boundaries were hedges with a poor ground flora; no wild legumes were noted.

*Host plants.* The following test host plants were used for sampling: pea (*Pisum sativum* J1 1194), Afghan pea (*Pisum sativum* J1 1103), clover (*Trifolium repens* cv. Grasslands Huia), French bean (*Phaseolus vulgaris* cv. Fino), lucerne (*Medicago sativa* cv. Europe).

*Sampling layout.* The test site was approximately 20 m from the edge of the field. An area 9 m × 9 m was protected from the nitrogen fertilizer that was applied to the remainder of the crop before sowing. Sixteen days after sowing, the crop was cleared from an area 2 m × 2 m at the centre of the site, and replaced by seeds of test host plants. Seeds of the five test hosts were surface sterilized in hypochlorite and sown at 20 cm spacing in a 10 × 10 randomized Latin square pattern. Peas and beans were sown singly, clover and lucerne in small clusters.

*Strain isolation.* All nodules were harvested from the test plants after 7 weeks of growth. Each nodule was surface sterilized in hypochlorite, and crushed and streaked on TY agar (Beringer, 1974). From each nodule streak, one well-separated colony was chosen, grown up on TY medium, washed off in 20% (v/v) glycerol and stored at −20 °C.

*Nodulation tests.* Each isolate was tested for the ability to form nodules on the appropriate test plant, grown axenically. Peas and lucerne were grown on an agar medium (Young et al., 1982), beans on Leca (porous mineral
Enzyme polymorphism in Rhizobium

Enzyme extracts. Each strain was grown for 2-3 d on TY medium in two 8.5 cm Petri plates. Confluent growth from these plates was scraped off in 1% NaCl and combined in a 1.5 ml Eppendorf tube. After brief centrifugation (MSE Micro Centaur), the pellet was resuspended in 0.5 ml buffer (62.5 mM-TES, 5 mM-magnesium acetate, 15% (v/v) glycerol, 0.1% mercaptoethanol, pH 7.2), centrifuged again, and resuspended in 0.1 ml of the same buffer containing 1 mg lysozyme ml\(^{-1}\). After 10 min incubation at room temperature, samples were sonicated on ice (MSE Soniprep 150 with microprobe) to break the cells and then centrifuged at high speed for 10 min. Supernatants were stored at \(-20^\circ\)C.

Enzyme electrophoresis. Extracts were run on horizontal 7% (w/v) polyacrylamide gels with pH 7.0 sodium phosphate buffer at 5 mM in the gel and 200 mM in the electrode compartments. After electrophoresis, the gel was sliced in half horizontally, and the cut surfaces were stained for specific enzyme activities by incorporating the appropriate ingredients in the molten 1% agar used to form an overlay. Glucose-6-phosphate dehydrogenase (G6PD) was stained by the standard NADP-linked tetrazolium method (Harris & Hopkinson, 1976). This stain also revealed superoxide dismutase (SOD) activity as clear zones in the background formed on exposure to UV light. This enzyme is also known as tetrazolium oxidase or indophenol oxidase. The substrate 4-O-methylumbelliferyl-\(\beta\)-d-galactoside revealed \(\beta\)-galactosidase (BGAL) activity as fluorescent bands visible under UV illumination; similarly, 4-O-methylumbelliferyl acetate revealed esterases.

Protein electrophoresis (SDS-PAGE). Extracts of total soluble protein were prepared and run on polyacrylamide gels in the presence of sodium dodecyl sulphate as described by Dibb et al. (1984), but with 15% (w/v) acrylamide in the running gel.

RESULTS

Nodulation. Peas (JI 1194), beans, clover and lucerne were all well nodulated, yielding a substantial number of isolates of each of the four corresponding Rhizobium classes, although rabbits reduced the number of plants available. The Afghan pea (JI 103), which is resistant to many European strains of \(R.\ leguminosarum\) biovar viceae (Young et al., 1982), did not form any nodules in the field. All isolates formed nodules with nitrogenase activity when tested on the appropriate host.

Enzyme electrophoresis. Most isolates had one major band of activity for each of the three enzymes G6PD, SOD and BGAL, but the mobility of the band varied from one isolate to another (Fig. 1). For each enzyme, each distinguishable mobility class was assigned an arbitrary single-letter designation in alphabetical order of decreasing mobility. Thus each distinct combination of mobilities (called an 'electrophoretic type' or ET by Caugant et al., 1981) could be described by three letters encoding G6PD, SOD and BGAL respectively. Some isolates, mostly \(R.\ meliloti\), had two BGAL bands (P and V).

Each isolate had several esterase isoenzymes, variable in both mobility and number (Fig. 2). Presumably several genetic loci were involved, but since the effects of each locus could not be disentangled, the esterase patterns were not classified formally, but used to confirm the evidence from the other enzymes.

\(R.\ leguminosarum\) strain 8002 (biovar phaseoli) and its derivative 8401, which lacks the symbiotic plasmid (Lamb et al., 1982), both had the ET MFF. In these strains, therefore, the genes encoding the enzymes G6PD, SOD and BGAL are not on the symbiotic plasmid. Similarly, strain 300 (biovar viceae) and its derivative 6015, which has a deletion in the symbiotic plasmid (Hirsch et al., 1980), had the ET SSQ. In \(E.\ coli\) the structural genes for enzymes with the activities considered in this study, and for all the other polymorphic enzymes studied in \(E.\ coli\), have been shown to be located on the chromosome (Bachmann, 1983; Ochman & Selander, 1984b). It seems likely that the enzymes I have studied in \(Rhizobium\) are also chromosomal markers, particularly in view of their strong disequilibrium with each other and with SDS protein profiles (see below), but the possibility that one or more are coded on plasmids (other than the symbiotic plasmid) cannot be entirely discounted.

Electrophoretic types of lucerne symbionts. The majority of strains isolated from lucerne (i.e. \(R.\ meliloti\) strains) were quite different in ET to those from the other host species (Table 1). There were only two major ETs amongst lucerne symbionts. These two types differed in mobility of

---

granules) wetted with Long Ashton nitrogen-deficient medium (Hewitt & Smith, 1975), and clover on slopes of Jensen's agar medium (Vincent, 1970). Nitrogenase activity was detected by acetylene reduction.
Fig. 1. (a) Variation of G6PD (dark bands) and SOD (light bands) in *R. leguminosarum* biovars *trifolii* (tracks 1 and 2) and *viceae* (tracks 3–5). Mobility classes of the strains in tracks 1–5 are G6PD-M, M, M, F, M and SOD-F, S, F, M, S respectively. (b) Variation of BGAL in biovar *viceae*. Tracks 6–10 are mobility classes BGAL-S, M, S, P, F respectively. Migration is anodal from the origin at the bottom.

Fig. 2. Esterase variation amongst *R. leguminosarum* isolates chosen to represent different biovars and different ETs (electrophoretic types for the enzymes G6PD, SOD and BGAL). Biovars: V, *viceae*; T, *trifolii*; P, *phaseoli*. ETs: 1, FMP; 2, MSS; 3, MFF; 4, SSQ. The far right-hand sample is John Innes strain 300, all others are from the Hainford site. Migration is anodal from the origin at the bottom.
Enzyme polymorphism in Rhizobium

Table 1. Numbers of Rhizobium strains with each enzyme electrophoretic type isolated from four different host plant genera

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>No. of strains of electrophoretic mobility class:</th>
</tr>
</thead>
<tbody>
<tr>
<td>G6PD</td>
<td>B D F M M M M M M M S S S S S S S S S</td>
</tr>
<tr>
<td>SOD</td>
<td>F F M F F F F F S S S S F S S S S S S S</td>
</tr>
<tr>
<td>BGAL</td>
<td>V P V P F K M P S K M Q S M M Q</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Host</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicago</td>
<td>14 13 1 28</td>
</tr>
<tr>
<td>Pisum</td>
<td>1 20 20 2 7 16 19 2 1 1 89</td>
</tr>
<tr>
<td>Trifolium</td>
<td>1 44 1 2 2 4 3 35 2 94</td>
</tr>
<tr>
<td>Phaseolus</td>
<td>40 1 41</td>
</tr>
</tbody>
</table>

G6PD, in the presence or absence of a second BGAL isoenzyme, and in the pattern of esterase isoenzymes (not shown), which was constant within but very different between the two types.

A single *R. meliloti* isolate had the ET MFF, which was frequent in the other species (see below). This isolate had some esterase bands in common with other MFF strains, but was not identical to any of them. It did nodulate and fix nitrogen on lucerne, but poorly in comparison with the other lucerne isolates.

Electrophoretic types of pea, clover and bean symbionts. For each enzyme, *R. leguminosarum* biovars *viceae*, *trifolii* and *phaseoli* had mobility variants in common (Table 1). Furthermore, variation of the different enzymes was highly correlated, so that the three biovars shared certain common combinations of alleles (ETs), while lacking other ETs that would result from recombination amongst these. Thus, MSS and MFF were frequent ETs in biovars *viceae* and *trifolii*, but MSF and MFS, which have the same variants in 'recombinant' arrangements, were rarer or even absent. While many ETs were found in both these biovars, some were confined to a single biovar, the clearest example being FMP in *viceae*.

The pattern of esterase isoenzymes confirmed the distinct identity of the ETs listed in Table 1. All MFF strains had a characteristic fast, sharp esterase band; SSQ strains had a similar but slower band; and most MSS, MFS and MFM strains had a slow, diffuse band that was absent in MFF, SSQ and FMP types (Fig. 2). Besides these constant features, however, there was substantial variation in esterase pattern within ETs. This variation bore no relation to host specificity (see Fig. 2). Amongst the major ETs, FMP was exceptional in being uniform for esterase pattern.

Interestingly, the *phaseoli* isolates had at least eight distinct esterase patterns, even though they were almost invariant at the other enzyme loci. At a second sampling site in the same field, some 200 m distant, two ETs were found in *phaseoli*: of 37 isolates (obtained in the same manner as at the main site), 23 were MFF and 14 were MFM.

Two isolates, one from pea and one from clover, had the ET DF(PV), which was frequent in *R. meliloti*. These two isolates had the same esterase pattern, which differed from that of *R. meliloti* DF(PV) strains only in the mobility of one band.

Protein profiles on SDS-PAGE. Representative strains of various ETs showed many small differences in the pattern of protein bands (Fig. 3). A prominent feature of all profiles was a region with distorted and retarded protein bands. In most strains this corresponded to approximately 20 kDal, but in certain strains (all of ET MFF) it was much slower, affecting bands of about 50 kDal. This distortion is believed to be caused by lipopolysaccharide, which can vary widely between strains (N. J. Brewin, personal communication). Both patterns of distortion (20 kDal and 50 kDal) occurred in all three biovars of *R. leguminosarum*. For the protein bands themselves, there was some variation within ETs, but the more striking differences were between ETs. The ETs FMP, SSQ and MSS, for example, had distinctively different profiles, although MFF and MSS were very similar (apart from the lipopolysaccharide variation). Within each ET, no consistent differences were observed between the biovars as defined by host range.
DISCUSSION

Enzyme electrophoresis has revealed substantial genetic polymorphism among isolates of *Rhizobium*, as indeed it has in almost all organisms to which it has been applied. Much more interesting was the strong disequilibrium observed between loci. The number of distinct ETs, that is of multi-locus genotypes, was much less than would be expected from random combination of the variation at each locus. Thus, for example, the ETs MFF and MSS were common among *R. leguminosarum* biovar *trifolii* isolates, but the corresponding 'recombinant' types, MFS and MSF, were absent. Likewise, FMP was a common ET in *R. leguminosarum* biovar *viceae*, but these alleles were rarely found in other combinations. In *R. meliloti*, only two major ETs were found; these differed in G6PD, BGAL and esterase patterns. This pattern of strong multi-locus disequilibrium closely resembles that reported, using similar techniques, for *E. coli* and *Shigella* (Selander & Levin, 1980; Caugant et al., 1981; Whittam et al., 1983; Ochman et al., 1983; Ochman & Selander, 1984b). World-wide isolates of these enteric bacteria can be typed into a small number of genetic clusters; there is little variation between members of a cluster, but members of different clusters differ at a substantial fraction of loci examined.

If there is frequent genetic recombination between ETs then the strong disequilibrium between loci, which is the basis for describing the distinct ETs, must result from selection against the hybrids formed. In other words, the major ETs represent co-adapted genomes, and most recombinants between them are less fit than either parent. Because selection can operate throughout the many clonal generations that may intervene between recombination events, the opportunity to develop strong disequilibria may be much greater in a bacterial population than in higher organisms that have recombination in each generation. In this respect, bacteria resemble those animal species that exhibit cyclic parthenogenesis. In a monitored population of the crustacean *Daphnia*, for example, a bout of sexual reproduction generated high diversity, but...
the status quo was restored during subsequent clonal growth by the elimination of recombinant clones (Young, 1979). In the case of *Daphnia*, which is diploid, deleterious recessive alleles could be postulated to explain the loss of the (largely homozygous) recombinants, but the only explanation available for haploid species is the co-adaptation of alleles at many loci. Selander & Levin (1980) have pointed out that such a mechanism is unlikely to be effective in the face of substantial recombination. Levin (1981) argued that recombination in *E. coli* populations is likely to occur at very low frequency, so that the species comprises a number of more-or-less isolated lineages. These correspond to the discrete clusters of ETs described amongst *E. coli* isolates (Ochman et al., 1983; Ochman & Selander, 1984 a, b). Caugant et al. (1981) concluded that the differences between closely related strains were as likely to be due to mutation as to recombination, since in some cases the difference was due to a unique variant in one strain. The rare ETs in the present study did not carry unique alleles, and strains were in any case examined for only a few loci, so it is not possible to decide on the present evidence whether the rare ETs observed in *Rhizobium* arose by mutation or recombination, or whether they constitute lineages with long independent histories.

Perhaps the most striking feature of the population structure is the way in which ETs cut across the classical species boundaries defined by the host plant. *R. meliloti* had its own distinctive ETs, in agreement with the taxonomic and genetic evidence for its separate identity (cited in the Introduction). However, the existence of one strain of *R. meliloti* which was closely similar to the most common type in *R. leguminosarum*, and two strains of *R. leguminosarum* that resembled typical *R. meliloti* isolates, suggests that the boundary between strains nodulating lucerne and those nodulating the other hosts may occasionally be transgressed. The genetic overlap amongst the three biovars of *R. leguminosarum* is much more striking. For example, the ET MSS was common in *viceae* and * trifolii* isolates, and MFF was common in these biovars and also in *phaseoli*. That these were really the same ETs is confirmed by their distinctive esterase and total protein patterns. This finding is fully consistent with the known genetic properties of these three biovars in the laboratory. Thus, chromosomal recombination is possible between certain strains of these biovars (Johnston & Beringer, 1977), and symbiotic functions, including host range, are carried on plasmids (symbiotic plasmids) which can readily be moved between strains from different biovars (Johnston et al., 1978; Hooykaas et al., 1981; Lamb et al., 1982). The host range of such an artificial hybrid is determined by the symbiotic plasmid it carries and not by the background genotype. Strains nodulating peas, beans or clover may differ genetically only in their symbiotic plasmid, though classically they were defined as separate species. The present paper demonstrates that this is true of natural strains as well as of laboratory constructs.

Are the isolates of *R. leguminosarum* all members of a single biological species? The distinct ETs revealed in this study appear to represent chromosomal lineages that are genetically isolated; whether this isolation is sufficiently complete to merit subdivision into species must await further study. It is clear, at least, that strains of the same chromosomal type can, and do, harbour symbiotic plasmids conferring different host ranges. This may mean either that symbiotic plasmids are readily transferred between ETs, or that each ET has its own set of symbiotic plasmids for various host plants. The spread of symbiotic plasmids amongst ETs is certainly not without barriers, for some combinations were not found. For example, the ET FMP was never found with any host range but that of biovar *viceae*. This might indicate that other symbiotic plasmids are not transferred to, or are not maintained in, or are not functional in strains of the FMP type. Even when the FMP strains are excluded from consideration, the frequencies of the various ETs are different in biovars *viceae, trifolii* and *phaseoli* (chi-squared tests, $P < 0.01$). This could be because the rate of plasmid transfer between ETs is not sufficient to maintain uniformity in the face of genetic drift or of selective plasmid–chromosome interactions, or because each host species selects a different biased sample of ETs.

Using host plants to obtain isolates, we inevitably select only those bacterial strains that carry the appropriate symbiotic plasmid. If the same ET can carry any of several different host ranges, it is more than likely that it sometimes carries no symbiotic plasmid at all. Such strains would, on the strict symbiotic definition, not be *Rhizobium* at all, and by sampling nodules we would never
see them, even if they constituted a substantial fraction or the overwhelming majority of the soil population. Even within truly symbiotic *Rhizobium*, our choice of host species limits our vision. For example, fast-growing symbionts have been reported in the legume tribes Galegeae and Loteae (Allen & Allen, 1981; Jarvis et al., 1982), which have native British representatives. When these coexist, as they must, with the species studied here, do they share ETs with them? There may even be shared lineages between *Agrobacterium* and these *Rhizobium* species, for taxonomic (Heberlein et al., 1967) and genetic (Hooykaas et al., 1984) evidence indicates that *Agrobacterium* is no more distant from *R. meliloti* or *R. leguminosarum* than these are from each other.

If symbiotic plasmids are indeed transferred in nature, even at low frequency, then a practical consequence follows. When a crop plant is introduced to a new area, *Rhizobium* with the appropriate host range is sometimes absent from the soil and must be introduced by inoculation. In some such instances, however, wild *Rhizobium* strains may already be present which are capable of carrying the necessary symbiotic plasmid, and recombination with the introduced strain may soon produce a genetically diverse population. A possible example of this phenomenon was reported by Roughley et al. (1976). They introduced a single strain of *R. leguminosarum* biovar *trifolii* where no clover symbionts had previously been found. The following year, many nodules were occupied by strains that were serologically unrelated to the inoculum, and in some plots the inoculant strain had been substantially displaced by the third year. Jansen van Rensburg & Strijdom (1984) made a similar observation after inoculation of soybeans.

The low diversity amongst strains of *Phaseolus* in the present study is in contrast to previous reports of genetic heterogeneity in this biovar (Graham, 1964; Heberlein et al., 1967; Roberts et al., 1980; Crow et al., 1981), and probably results from the narrow geographic base of the sample. The reported diversity of isolates obtained from *Phaseolus* species may reflect the promiscuity of this host: it frequently nodulates with strains isolated from species in other cross-inoculation groups (e.g. Crow et al., 1981). It would appear that the Hainford site harbours only a small part of the broad range of *Rhizobium* strains that nodulate *Phaseolus* world-wide. This is hardly surprising, as there are no native hosts for this species in Britain, and it is plausible that these *R. phaseoli* strains have colonized and/or evolved since the introduction of cultivated *Phaseolus* beans.

When the three biovars of *R. leguminosarum* are combined, the average level of polymorphism of the three loci for G6PD, SOD and BGAL is $h = 0.50$ ($h = 1 - \sum x_i^2$, where $x_i$ is the frequency of the $i$th allele). This level is consistent with that found in *E. coli* (e.g. 0.47 in Selander & Levin, 1980). The number of alleles per locus is lower in this *Rhizobium* sample than in the *E. coli* studies, being 4, 3, and 7 for G6PD, SOD and BGAL respectively (based on 268 strains), against 5, 6 and 19 for the same three enzymes in *E. coli* (based on 109 strains), or an average of 7.3 over 20 enzymes (Selander & Levin, 1980). However, taking into account the much narrower geographic range of the *Rhizobium* sample, we may conclude that *R. leguminosarum* shows a level of genetic variation broadly comparable with that in *E. coli*. The uniformity of *E. coli* lineages over space and time is ascribed by Levin (1981) to the effect of periodic selection (Atwood et al., 1951), in which advantageous mutations are continually arising and carrying the genetic background of their clone (usually the most common one) to high frequency (i.e. genetic 'hitch-hiking'). The theoretical and experimental justifications for this idea are based on populations in chemostats or batch cultures, in which the competing genotypes are well mixed. These may be reasonable models for *E. coli* in an individual gut, but the outcome for the species as a whole could be a mosaic of differentiated populations, or a set of cosmopolitan lineages, if migration between populations were rare or frequent respectively. Direct observation of the immigration of new genotypes, and the widespread occurrence of some genotypes, indicate that the migration rate is relatively high in *E. coli* (Levin, 1981). We do not as yet have comparable data for *Rhizobium*, which has of course very different ecology.

So far as one can generalize from a single sampling site, the clonal structure of *Rhizobium* seems to be very similar to that found in *E. coli*. While the exact strains present and their relative abundance are probably specific to the single sample, the overall pattern of diversity and
Enzyme polymorphism in Rhizobium

relationship is probably a general feature of these species, since it is entirely consistent with our knowledge of Rhizobium genetics, DNA homology and numerical taxonomy.

Population genetics provides a viewpoint quite distinct from that of taxonomy. In this study, a substantial number of strains were sampled from a single population, whereas taxonomists usually examine one strain from each of many populations. The taxonomist gains a more complete picture of the range of phenotypes present in each species, but cannot draw conclusions about genetic exchange. It should be noted that the finding of polymorphism greatly strengthens the conclusions from the present study. If all strains from both peas and clover had been identical for all enzymes, the only conclusion would be that they were probably fairly closely related. It is the occurrence of the same polymorphisms in isolates from different hosts that confirms the genetic overlap of the biovars.

I thank Mr Henly Woods for permission and co-operation in setting up the sampling sites; Diane Rushbrook, Nora Demetriou, Jackie Roff, Annie Rowe and Graham Carter for assistance; and Nick Brewin, Roy Davies, Allan Downie and Andy Johnston for helpful commentary on the work and on this account of it.

REFERENCES


