Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 9621–9624, December 1988
Ecology

Genetic diversity in natural populations of a soil bacterium across a landscape gradient

(*Pseudomonas cepacia*/enzyme polymorphism/habitat variability/genetically engineered microorganisms)

J. VAUN MCARTHUR*, DAVID A. KOVACIC†, AND MICHAEL H. SMITH*

*Savannah River Ecology Laboratory, Drawer E, University of Georgia, Aiken, SC 29801; and †Department of Landscape Architecture, University of Illinois, Urbana, IL 61801

Communicated by Eugene P. Odum, September 19, 1988

ABSTRACT **Genetic diversity in natural populations of the bacterium *Pseudomonas cepacia* was surveyed in 10 enzymes from 70 clones isolated along a landscape gradient. Estimates of genetic diversity, ranging from 0.54 to 0.70, were higher than any previously reported values of which we are aware and were positively correlated with habitat variability. Patterns of bacterial genetic diversity were correlated with habitat variability. Findings indicate that the source of strains used in genetic engineering will greatly affect the outcome of planned releases in variable environments. Selection of generalist strains may confer a large advantage to engineered populations, while selection of laboratory strains may result in quick elimination of the engineered strains.**

Although natural selection is the process that affects allele frequency, promotes genetic diversity, and ultimately results in environmental adaptation, few studies have shown a relationship between genetic diversity and environmental variability (1). There is considerable circumstantial evidence indicating a positive relationship of genetic polymorphisms to environmental heterogeneity (2). It is difficult to demonstrate the relationship of physiological, morphological, and behavioral attributes of higher organisms to enzyme polymorphisms; however, this may be less difficult for bacteria. While relationships between genetic diversity and environmental variables (e.g., soil ion concentrations) probably exist in larger organisms, a more direct relationship should be exhibited by microbes, for which the physical-chemical characteristics of the environment are the direct basis of selection (3).

The homeostatic mechanisms that reduce perceived environmental extremes are lacking or are not very effective in bacteria compared to those of higher organisms. Environmental differences in nutrient concentration, nutrient availability, temperature, and pH directly affect the kinetics of bacterial enzymes. These factors act as selective agents on genes controlling enzyme systems; therefore, a relationship should exist between various components of genetic diversity among populations of bacteria and variability associated with their environment.

Microorganisms, perhaps the most adaptable of all organisms, are physiologically active at temperatures from −5°C to 105°C and possibly 250°C, and at pressures ranging from <1 to 40 atmospheres (4–7). This range of adaptability results, in part, from (i) their ability to evolve rapidly compared to higher organisms, (*ii*) short generation times, (*iii*) large effective populations, (*iv*) high surface-to-volume ratios that require immediate response to changes in environmental conditions, and (*v*) DNA exchange in bacterial populations (8). Evidence of the most direct effects of natural selection due to environmental differences should be manifested in microbes.

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.

Researchers using immunological and serological methods have shown that there exists considerable variability within bacterial taxonomic groups (9–12). However, only two attempts have been made to estimate genetic diversity in natural populations of bacteria (13, 14). Both of these studies utilized *Escherichia coli* from a variety of sources. Although genetic variability was found, there was no attempt to correlate the observed diversity with environmental variables. Selander and Levin (14) concluded that the observed variability was much less than expected. This observation is not surprising since *E. coli* inhabits relatively large, mobile, homeostatic organisms that provide similar constant environments for the bacteria.

Bacteria living in more variable environments—e.g., soil—are exposed to wide fluctuations in temperature, moisture, and nutrient availability and should exhibit greater genetic diversity than *E. coli*. We hypothesized that genetic diversity of populations of a bacterial species would differ along an environmental gradient. Specifically, changes in the genetic diversity of alleles controlling metabolic enzymes in *Pseudomonas cepacia*, a common soil microbe, should be positively correlated with environmental variability.

MATERIALS AND METHODS

The bacterium *P. cepacia* was isolated from 5-cm soil cores taken along an environmental gradient. Cores were randomly taken from 40 permanent plots established on four distinct habitat types located in the Meyers Branch drainage of the Savannah River Plant, Aiken County, South Carolina. The Meyers Branch study area encompassed approximately 16×10^5 m^2. Ten additional permanent plots were established in an adjacent sand hill site. The five habitat types were as follows: longleaf pine plantation (*Pinus palustris*); upland deciduous forest dominated by red oak (*Quercus falcata*) and mockernut hickory (*Carya tomentosa*); bottomland hardwood forest dominated by white oak (*Quercus alba*), loblollypine (*Pinus taeda*), tulip poplar (*Liriodendron tulipifera*), red maple (*Acer rubrum*), dogwood (*Cornus florida*), and holly (*Ilex opaca*); cypress swamp dominated by bald cypress (*Taxodium distichum*), black gum (*Nyssa sylvatica*), and red maple (*A. rubrum*); and sandhills dominated by turkey oak (*Quercus laevis*) and longleaf pine (*Pinus palustris*).

At each sampling location soil pore water was sampled biweekly from 10 porous cup lysimeters and analyzed for concentrations of dissolved organic carbon, NO_3^- nitrogen, NH_4^+, P, K, Ca, Mg, Fe, and Mn. Soil cores removed biweekly from the locations were analyzed for percent water, percent organic content, and pH. Weather stations located within a habitat type recorded daily air temperature, soil maximum and minimum temperature, and surface maximum and minimum temperature. Samples were collected for 18 months.

The bacteria were isolated from soil cores on a minimal salts/glucose medium supplemented with 2 g of nutrient broth per liter. Seventy clones were isolated from the habitat

types. Clones were identified by using a rapid-identification strip (API, Rapid NFT System, Analytab Products) that utilizes 21 biochemical tests. All 70 isolates were determined to be *P. cepacia* with >99% accuracy. The isolates were streaked on plates and the resulting growth was scraped from the plates into 2 ml of sterile grinding solution and sonicated (Fisher Sonic Dimembrator, model 300) at 200 W three times for 10 s in an ice bath. These clones were assayed for genetic variation in 10 enzymes by using starch gel electrophoresis. Multiple banding patterns were seen for most enzymes. Enzymes stained for were malic enzyme, malate dehydrogenase, isocitrate dehydrogenase, glucose-6-phosphate dehydrogenase, glucose phosphate isomerase, menadione phosphorylase, nucleoside phosphorylase, aconitase, 6-phosphogluconate dehydrogenase, and peptidase.

Genetic diversity estimates (H) were calculated for each enzyme by the following formula: $H = 1 - \Sigma x_i^2$, where x_i is the frequency of the ith electromorph (10). Mean genetic diversity was estimated from the diversity estimates of each enzyme by habitat. Coefficients of variation were calculated for each environmental variable within a habitat. Coefficients of variation were calculated for each environmental variable within a habitat over the sampling period (18 months). Correlation analyses were performed between mean genetic diversity and the 16 environmental variables. Habitat variability was estimated by summing the coefficients of variation for each habitat variable that was significantly correlated with mean genetic diversity (22). Spearman rank correlations between mean genetic diversity and habitat variability at each location were estimated. A canonical discriminate analysis was performed on the electromorph scores for each enzyme system by clone to determine relationships to habitat. Statistical analyses were performed by using SAS software (15).

RESULTS

Genetic diversity of *P. cepacia* was significantly correlated with coefficients of variation for 5 of the 16 environmental variables (percent organic matter, dissolved organic carbon, Fe, NO_3, and Mg) (Table 1). All of the significant correlations were positive. Of the 8 inorganic or organic variables measured, the coefficients of variation for 6 were highly correlated with each other. Only the coefficients of variation for P and Ca showed no correlation with either genetic diversity or the other chemical variables measured.

Table 1. Correlation matrix of the coefficients of variation for dissolved organic carbon (DOC), percent organic matter (% Org), Fe, Mg, K, Ca, NO_3, and genetic diversity

	Genetic diversity	Fe	% Org	Mg	NO_3	DOC	K	Ca
Fe	0.99*							
% Org	0.94*	0.88*						
Mg	0.92*	0.92*	0.91*					
NO_3	0.91*	0.94*	0.84	0.97*				
DOC	0.89*	0.89*	0.72	0.80	0.80			
K	0.69	0.72	0.58	0.85	0.87*	0.77		
Ca	0.11	−0.04	0.31	0.01	−0.18	0.06	−0.23	
P	0.01	0.03	0.15	0.06	0.12	−0.41	−0.22	0.75

*$P \leq 0.05$.

The sums of the 5 significantly correlated coefficients of variation by habitat were used as measures of environmental variability. In regression analysis, environmental variability accounted for 78% of the genetic diversity ($P = 0.047$) (Fig. 1). Greatest genetic diversity was found in isolates collected from the bottomland hardwood sites, the habitat with the highest environmental variability. Isolates from the pine and turkey oak habitats, those with the lowest environmental variability, had the lowest genetic diversity.

The patterns of enzyme diversity varied across the habitat types (Table 2). Of the 10 enzyme systems surveyed, isocitrate dehydrogenase and malic enzyme were the most variable (genetic diversity 0.83 and 0.69), and glucose-6-phosphate dehydrogenase, malate dehydrogenase, and nucleoside phosphorylase were the least variable (0.47, 0.46, and 0.37, respectively). The average genetic diversity across all habitats was 0.59.

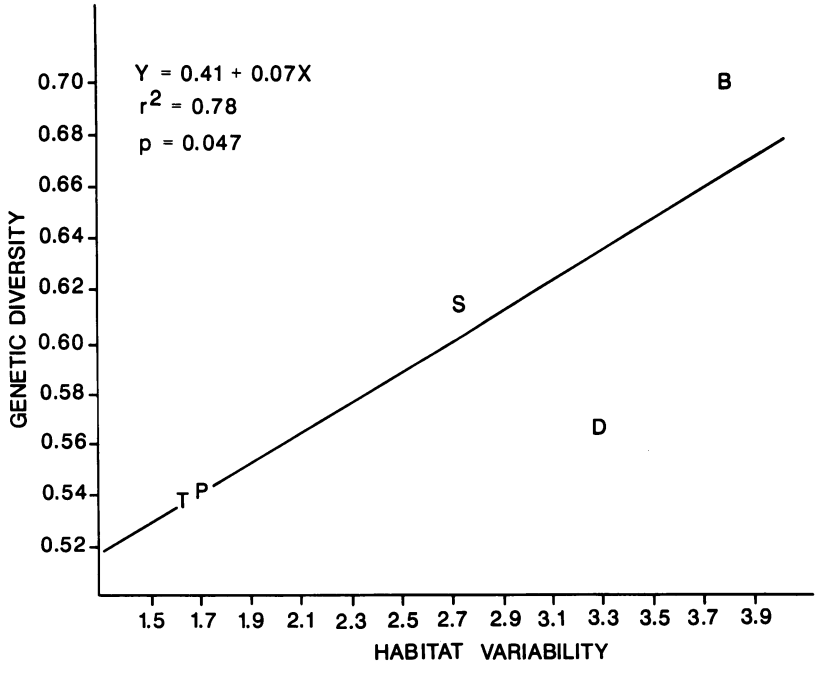

FIG. 1. Linear relationship between genetic diversity (y) in clones of *P. cepacia* and habitat variability (x) as measured by summing the coefficients of variation for five environmental variables (soil percent organic, dissolved organic carbon, Mg, Fe, NO_3). The equation of the regression line, the coefficient of determination (r^2), and the probability (P) of the regression are given. S, swamp; P, pine plantation; D, upland deciduous; B, bottomland deciduous; T, turkey oak.

Table 2. Genetic diversity of metabolic enzymes in *P. cepacia* within different habitat types and maximal number of allozymes found for each enzyme

Enzyme	No. of allozymes	Genetic diversity						
		Turkey oak	Pine	Upland deciduous	Swamp	Bottom-land	Mean	SD
MOD	8	0.66	0.67	0.65	0.64	0.85	0.69	0.08
ICD	13	0.84	0.87	0.83	0.78	0.85	0.83	0.30
ACON	4	0.64	0.65	0.70	0.56	0.57	0.62	0.05
G6PD	3	0.49	0.46	0.66	0.32	0.43	0.47	0.11
MNR	5	0.58	0.72	0.67	0.56	0.70	0.65	0.06
PEP	4	0.31	0.22	0.50	0.72	0.78	0.51	0.22
GPI	7	0.66	0.65	0.59	0.56	0.66	0.62	0.04
6PGD	10	0.61	0.51	0.52	0.80	0.78	0.64	0.12
MDH	5	0.27	0.34	0.20	0.78	0.69	0.46	0.23
NP	5	0.29	0.28	0.20	0.32	0.78	0.37	0.21
Mean	6.8	0.54	0.54	0.56	0.60	0.70	0.59	0.06

MOD, malic enzyme; ICD, isocitrate dehydrogenase; ACON, aconitase; G6PD, glucose-6-phosphate dehydrogenase; MNR, menadione phosphorylase; PEP, peptidase; GPI, glucose-phosphate isomerase; 6PGD, 6-phosphogluconate dehydrogenase; MDH, malate dehydrogenase; and NP, nucleoside phosphorylase.

Canonical discriminant analyses of the genetic information effectively separated individual isolates into groups concordant with distinct habitat types ($\lambda = 0.019, P < 0.05$) (Fig. 2). The overlap in the two-dimensional plot is effectively resolved by the third canonical variate. Placement of individual clones into their respective habitat groupings occurred with greater than 80% correct classification. Clones isolated from the bottomland hardwood sites were 100% correctly classified.

DISCUSSION

Measures of genetic diversity reported here are greater (2-fold) than those of Milkman (13) and Selander and Levin (14), who reported estimates of genetic diversity in *E. coli* 2 to 3 times greater than in any other oganism. Other soil bacteria may be as genetically diverse as *P. cepacia*; however, most microbial communities have not been studied in this regard (16).

The results of the canonical discriminant analysis and the relationship between genetic diversity and environmental variability suggest a pattern of microgeographical adaptation due to selection. Since dispersal in these bacteria is quite limited, differentiation due to selection among local populations is enhanced. Bacteria seem to respond to their environment as if it were coarse grained (17). Clones separated in space are likely to be in and stay within different habitat patches. Given these conditions and a short generation interval, adaptations to local conditions are promoted. Each site would tend to have clones adapted to its particular set of environmental conditions in a manner similar to that described in Wallace's (18) "bed of nails" model. In this model each "nail" represents an adaptive peak achieved by the individuals at that particular location, with the height of nails indicating the degree of adaptation. The way in which the nails are located in space—i.e., the pattern of adaptation—would depend upon the spacial distributions and temporal predictability of important environmental characteristics.

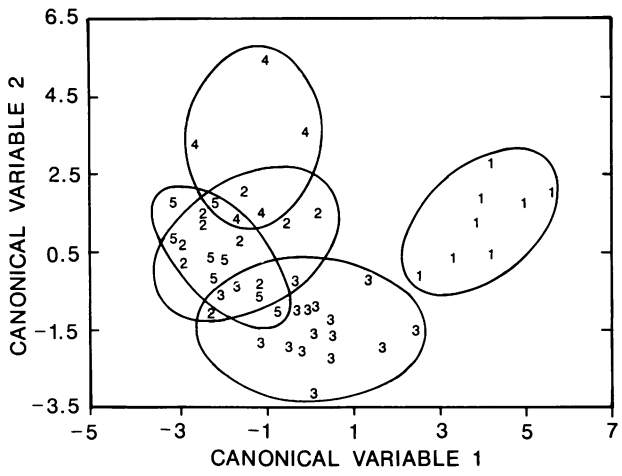

FIG. 2. Canonical discriminant plot of individual isolates of *P. cepacia* identified by habitat from which they were isolated based on the frequency of electromorphs of each enzyme. Ellipses represent groupings of habitat types: 1, bottomland deciduous; 2, pine plantation; 3, upland deciduous; 4, swamp; and 5, turkey oak.

Natural selective processes appear to be acting at two levels: (*i*) variation among habitats; and (*ii*) variation within habitats. Genetic diversity among clones in the major habitats was correlated with the characteristics of these habitats, pointing to the importance of directional selection at the landscape level. This selection is also the likely cause of correlation between environmental and genetic diversity. Mutation, dispersal, or stochastic effects do not adequately explain gene–habitat relationships.

At the habitat level, this species of bacteria must exhibit fitness peaks similar to those in Wright's shifting-balance model of evolution (19). Understanding the extent of genetic diversity of bacteria and the pattern of their adaptive topographies is a necessary part of discerning the evolutionary paradigm for this life form. Bacteria occupy an extreme position along the continuum of possible environmental grain and, thus, may be useful in delimiting the application of Levins' theory (17). Evolutionary insights gained from the study of large mobile organisms, which have a greater degree of homeostatic control over their internal biochemical environments, will probably not be useful in understanding bacterial evolution; however, aspects of certain general evolutionary dynamics observed for bacterial populations may assist in the development of models.

P. cepacia can utilize many different carbon compounds as its sole energy source (20). Growth experiments used laboratory strains and were performed under ideal conditions. Bacteria in natural environments would encounter a much more diverse suite of organic substrates. Since the nature (both qualitatively and quantitatively) of organic matter in a specific habitat will be dependent on the diversity of the vegetation, the bacteria living in a species-diverse plant community would encounter a higher relative diversity of

possible substrates. Additionally, the bacteria would encounter these molecules under constantly changing environmental conditions—e.g., temperature and pH.

Those variables whose coefficients of variation correlated significantly with genetic diversity and habitat variability were percent organic, Fe, NO_3, Mg, and dissolved organic carbon concentrations. The variability associated with concentrations of cations necessary for specific functions in metabolism or enzyme cofactors (21) would greatly alter metabolic efficiency (3). In nature, this variability may select for multiple isozymes of an enzyme, thus enabling the microbe to survive under constantly changing conditions. Although the multiple bands seen in most of the systems suggest alternative forms of the same enzyme, we cannot say that these multiple forms provide the advantages previously discussed. Hochachka and Somero (3) argue that multiple forms of an enzyme may be needed to restore metabolic or regulatory ability during changing conditions. Furthermore, the microenvironment around the enzyme (e.g., osmolytes, lipids, protons) must be continually adjusted to preserve enzyme function. While bacteria can control their microenvironment within the cell to some degree, they have little control outside the cell. Extreme variability in the external environment may change the internal conditions.

The patterns and magnitude of genetic variability demonstrated in *P. cepacia* have implications for the release of genetically engineered microorganisms. In light of present proposals to release genetically engineered bacteria into the environment, it is important that researchers or applied scientists understand the ecology and population genetics of natural microbial communities. Genetically engineered bacteria may not be able to respond to the variability of natural environments. This may be due, in part, to the potential sources of strains used to engineer new genomes. There are at least two distinct sources: (*i*) existing laboratory strains or (*ii*) new strains isolated from nature. Laboratory strains would not be expected to have genes that produce products that are functionally efficient in many different habitats. Alternatively, success of genetically engineered microorganisms in the environment may be dependent on the preselection of generalist strains, strains capable of utilizing a wide variety of substrates. Such preselection could circumvent natural selection processes in the environment and confer a large advantage to the engineered bacterial populations.

Correlative relationships between environmental heterogeneity and natural genetic variability of an organism provide support for evolutionary ecological models. Further experimental work that defines the environmental conditions that elicit the genetic responses are needed. The coupling of microbial ecology and evolutionary ecology should provide understanding of these relationships. The results of such couplings should show the mechanisms of the response(s) and provide a necessary background for environmental biotechnology.

We thank L. Leff, A. Leff, P. Johns, and S. Coleman for various aspects of data collection, manuscript preparation, and technical assistance. This research was supported by Contract DE-AC09-76SROO-819 between the U.S. Department of Energy and the University of Georgia, Institute of Ecology.

1. Nevo, E., Beiles, A. & Ben-Shlomo, R. (1984) in *Evolutionary Dynamics of Genetic Diversity*, ed. Mani, G. S. (Springer, Berlin), pp. 13–213.
2. Hedrick, P. W., Ginevan, M. E. & Ewing, E. P. (1976) *Annu. Rev. Ecol. Syst.* **7**, 1–32
3. Hochachka, P. W. & Somero, G. N. (1984) *Biochemical Adaptation* (Princeton Univ. Press, Princeton, NJ).
4. Corliss, J. B., Baross, J. A. & Hoffman, S. E. (1981) *Oceanol. Acta* No Sp. **59**, 56–69.
5. Baross, J. A. & Deming, J. W. (1983) *Nature (London)* **303**, 423–426.
6. Deming, J. W. & Becker, R. R. (1984) in *Current Perspectives in Microbial Ecology*, eds. Klug, M. J. & Reddy, C. A. (Am. Soc. Microbiol., Washington, D.C.), pp. 186–195.
7. Trent, J. D., Chastian, R. A. & Yayanos, A. A. (1984) *Nature (London)* **307**, 737–740.
8. Shapiro, J. A. (1985) in *Engineered Organisms in the Environment: Scientific Issues*, eds. Halvorson, H. O., Pramer, D. & Rogul, M. (Am. Soc. Microbiol., Washington, D.C.), pp. 63–69.
9. Woomer, P., Singleton, P. W. & Bohlool, B. B. (1988) *Appl. Environ. Microbiol.* **54**, 1112–1116.
10. Hankinson, T. R. & Schmidt, E. L. (1988) *Appl. Environ. Microbiol.* **54**, 1536–1540.
11. Sadowsky, M. J., Bohlool, B. B. & Keyser, H. H. (1987) *Appl. Environ. Microbiol.* **53**, 1785–1789.
12. Schmidt, E. L., Zidwick, M. J. & Abebe, H. M. (1986) *Appl. Environ. Microbiol.* **51**, 1212–1215.
13. Milkman, R. (1973) *Nature (London)* **182**, 1024–1026.
14. Selander, R. K. & Levin, B. R. (1980) *Science* **210**, 545–547.
15. SAS Institute (1985) *SAS User's Guide: Statistics* (SAS Inst., Cary, NC).
16. Margulis, L., Chase, D. & Guerrero R. (1986) *Bioscience* **36**, 160–170.
17. Levins, R. (1968) *Evolution in Changing Environments* (Princeton Univ. Press, Princeton, NJ).
18. Wallace, B. (1968) *Topics in Population Genetics* (Norton, New York).
19. Wright, S. (1931) *Genetics* **16**, 97–159.
20. Palleroni, N. J. (1984) in *Bergey's Manual of Systematic Bacteriology*, eds. Krieg, N. R. & Holt, J. G. (Williams & Wilkins, Baltimore), Vol. 1, pp. 140–199.
21. Gottschalk, G. (1979) *Bacterial Metabolism* (Springer, Berlin).
22. Hanski, I. (1978) *Ecology* **59**, 168–174.