APPLIED AND ENVIRONMENTAL MICROBIOLOGY, May 1995, p. 1791–1798 Vol. 61, No. 5
0099-2240/95/$04.00+0

Genetic Structure of a Lotic Population of *Burkholderia* (*Pseudomonas*) *cepacia*

MARK G. WISE,[1] LAWRENCE J. SHIMKETS,[1]* AND J VAUN MCARTHUR[2]

Department of Microbiology, University of Georgia, Athens, Georgia 30602-2605,[1] and Savannah River Ecology Laboratory, Aiken, South Carolina 29802[2]

Received 7 November 1994/Accepted 9 March 1995

The genetic structure of a population of *Burkholderia* (*Pseudomonas*) *cepacia* isolated from a southeastern blackwater stream was investigated by using multilocus enzyme electrophoresis to examine the allelic variation in eight structural gene loci. Overall, 213 isolates were collected at transect points along the stream continuum, from both the sediments along the bank and the water column. Multilocus enzyme electrophoresis analysis revealed 164 distinct electrophoretic types, and the mean genetic diversity of the entire population was 0.574. Genetic diversity values did not vary spatially along the stream continuum. From a canonical discriminant analysis, Mahalonobis distances (measurements of genetic similarity between populations) revealed significant differences among the subpopulations at the sediment sampling points, suggesting bacterial adaptation to a heterogeneous (or patchy) microgeographical environment. Multilocus linkage disequilibrium analysis of the isolates revealed only limited association between alleles, suggesting frequent recombination, relative to binary fission, in this population. Furthermore, the dendrogram created from the data of this study and the allele mismatch distribution are typical of a population characterized by extensive genetic mixing. We suggest that *B. cepacia* be added to the growing list of bacteria that are not obligatorily clonal.

The most widely used method of assessing genetic diversity and structure in bacterial populations has been multilocus enzyme electrophoresis (MLEE) (30). In this technique, bacterial isolates are characterized by the relative mobilities of a number of cellular enzymes. Each of these enzymes, when applied to a gel matrix and subjected to electrophoresis, migrates on the basis of its electrostatic charge which is, in turn, directly related to its amino acid sequence. Different mobility variants can be directly equated with alleles at the corresponding structural gene locus. Profiles of the variants (allozymes or electromorphs) for multiple enzymes constitute an electrophoretic type (ET) that can be considered a multilocus genotype. A large proportion (80 to 90%) of amino acid substitutions can be detected by this technique (30). MLEE data have been used to estimate the levels of single-locus and multilocus genotypic variations in populations, as well as the extent of genetic exchange within a population. The genetic distance between strains can also be calculated, resulting in a dendrogram for visualizing the phylogenetic relationships between isolates (30, 33).

The population genetics of *Escherichia coli* has been thoroughly examined by using MLEE, and results have revealed that natural populations of *E. coli* harbor extensive genetic diversity, but this diversity is organized into a limited number of genetically distinct clones (5, 6, 12, 31, 32, 37, 39). On the basis of the high levels of linkage disequilibrium (nonrandom association of alleles at different loci in a population) observed, recombination is not frequent enough to break up associations between loci on the chromosome and populations are basically clonal. The population genetic structure of a few other bacterial species has been examined, and most of these studies have described populations that are also basically clonal, including *Salmonella* spp., *Legionella pneumophilia*, *Haemophilus influenzae*, *Porphyromonas gingivalis*, *Serratia marcescens*, and *Bordetella* spp. (10, 17, 29, 33). These results have led to the proposal of a clonal bacterial paradigm characterized by the existence of strong linkage disequilibrium.

Not all bacteria appear to follow the clonal paradigm, however. Although bacterial reproduction is asexual, transfer of genes can be achieved through the processes of transduction, transformation, and conjugation. The frequency of such occurrences in *E. coli* is too low to disrupt the overall clonal structure, but that does not preclude any recombination at all. Dykhuizen and Green (8) showed that phylogenies constructed for different chromosomal genes (by using DNA sequence data) from the same set of *E. coli* isolates were significantly different from one another. This likely means that there is some localized recombination in *E. coli*. Maynard Smith and Dowson (19) showed that some individual bacterial genes have a mosaic structure that could likely have arisen by recombination.

Recent allozyme studies also suggest that not all bacteria are equally clonal. Istock et al. (14) found only slight linkage disequilibrium in a population of *Bacillus subtilis* taken from one site in an Arizona desert. O'Rourke and Stevens (25) reported that *Neisseria gonorrhoeae* has a nonclonal population structure, the first such structure described for a pathogen. The index of association (a measure of multilocus linkage disequilibrium) for this population of *N. gonorrhoeae* did not differ significantly from zero, the expected value for populations in panmixia, i.e., those characterized by random mating (20). Souza et al. (35) analyzed the genetic structure of *Rhizobium leguminosarum* biovar phaseoli populations, and their data suggest that recombination may be quite common. The population variations in few other environmental bacterial species have been examined.

Lotic ecosystems (rivers and streams) are molded by unique physical forces, specifically, the flow of water. Bacteria are essential for biogeochemical cycling in these ecosystems, and they provide the base of the food web (18). Yet, little is known about the species composition and variability of natural bacterial communities in streams. Even less is known about the

* Corresponding author. Phone: (706) 542-2681. Fax: (706) 542-2674. Electronic mail address: shimkets@bscr.uga.edu.

population structure of lotic bacteria. The continuous movement of water in streams has many effects on the bacterial inhabitants. Bacteria are constantly affected by the physical force of the water such that they must maintain location or be displaced downstream. Bacterial assemblages must be productive enough to replace members that are swept away or bacteria from somewhere outside the stream, like the floodplain, for example, must colonize the area (15). Displaced cells must be able to colonize new areas to which they are transported. Lotic ecosystems, however, can be very heterogeneous environments. Current stream theory describes extensive changes in the physical characteristics of a stream as one moves from the headwaters to the mouth (9, 26). The diversity of soluble organic matter and the amount of input fluctuate along a stream continuum (21). Successful bacterial populations must, therefore, be able to thrive in a variety of habitat conditions. Extensive genetic diversity and frequent genetic exchange may allow lotic bacterial populations to adapt to differing microhabitats (22).

The population structure of environmental bacteria (with an emphasis on gene flow) is an important consideration when assessing the risks associated with releasing genetically engineered microorganisms into the environment. Specifically, the extent to which genetic exchange occurs within a local population and the extent of migration between populations are significant factors when attempting to evaluate the potential persistence and spread of modified genes. To evaluate these factors, we decided to focus on aquatic forms of *Burkholderia* (*Pseudomonas*) *cepacia*, a common stream bacterium. In this work, our objectives were threefold. First, we wanted to assess the extent of genetic diversity in aquatic *B. cepacia* isolated from a blackwater creek and specifically to examine whether diversity is affected by the site of isolation. Second, we were interested in examining whether genetic distance between isolates correlates with the geographical distance between isolation sites. Finally, we were interested in describing the genetic structure of the population by evaluating the extent of linkage disequilibrium between loci and thereby making some inferences on the prevalence of recombination in the population (clonality versus sexuality).

MATERIALS AND METHODS

Sampling sites. Samples were collected on 4 November 1993 from sites in the Upper Three Runs Creek (UTR) drainage basin on the U.S. Department of Energy's Savannah River Site near Aiken, S.C. UTR is one of many blackwater streams, common in the southeastern United States, which are named for their tea-colored water. These water systems have a low content of suspended sediment but high concentrations of dissolved organic matter (23). The sampling scheme is shown in Fig. 1. Tinker Creek, a tributary of UTR, was sampled approximately 1 m from its bank along a transect at 0-, 1-, 4-, 8-, 12-, 16-, and 50-m points. Sediment was collected with sterile petri plates and placed in sterile whirl pack bags. Five independent replicate samples were taken from each site. Water samples were taken with sterile whirl pack bags from the top of the water column at the 0- and 50-m transect points. To examine spatial differences on a larger scale, sediment and water samples (with five replicates each) were taken at the Aquatic Ecology Laboratory (AEL) located on UTR, approximately 5 km downstream from the transect at Tinker Creek.

All samples were placed on ice in the field and transported back to the laboratory. For sediment samples, 4 ml of sterile saline (0.85% NaCl) was added to 10 g of sediment and vortexed for 30 s. The slurry was sonicated for 10 s at 4% maximum power (Heat Systems-Ultrasonics, Inc., model W-380) to disrupt bacterium-particle attachment and then vortexed again for 10 s, and 650 μl of the liquid was plated on TB-T agar (11) amended with cycloheximide at a final concentration of 100 μg/ml to inhibit fungal growth. TB-T agar utilizes a combination of trypan blue and tetracycline to provide a high degree of selectivity for *B. cepacia* biotypes. For water samples, 25 ml was subjected to centrifugation at 3,000 × *g* for 5 min. Twenty-four milliliters of the supernatant was removed, and the remaining 1 ml was resuspended and plated on TB-T agar amended with cycloheximide. Plates were incubated at 37°C.

Isolation of *B. cepacia*. To identify *B. cepacia* isolates, colony hybridization was performed with a 23S rRNA probe specific for this organism (28). Since both 16S and 23S rRNAs contain highly conserved nucleotide sequences, it is possible to design oligonucleotide probes that are species specific (40). The 23S probe has been shown to hybridize to *B. cepacia* subgroup A biotypes with a low frequency (7%) of false positives (16). Colonies from TB-T plates were transferred to *Pseudomonas* Isolation Agar (Difco) plates (all isolates acquired from the TB-T plates grew on *Pseudomonas* Isolation Agar) and subsequently attached to Hybond N^+ nylon filters (Amersham). The cells were lysed with 0.5 M NaOH, neutralized in 0.5 M Tris (pH 8), and immersed in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate [pH 7.2]) and then in 100% ethanol. Filters were baked at 80°C for 1 h and stored.

The *B. cepacia* specific probe (3′-TAACAATCTACGCTACCC-5′) which hybridizes to positions 1406 to 1423 of the 23S rRNA was labeled at the 3′ end with digoxigenin-11-ddUTP by using a terminal transferase enzyme as specified in the Genius nonradioactive DNA labeling and detection kit (Boehringer Mannheim). Briefly, the oligonucleotide, at a final concentration of approximately 5 pmol/μl, was incubated with 5 mM $CoCl_2$–0.2 M potassium cacodylate–25 mM Tris-HCl (pH 6.6)–0.25 mg of bovine serum albumin per ml–0.2 mM digoxigenin-11-ddUTP–2.5 U of terminal transferase per μl at 37°C. The reaction was stopped after 15 min by adding 1 μl of 0.5 M EDTA and placing the tube on ice. The

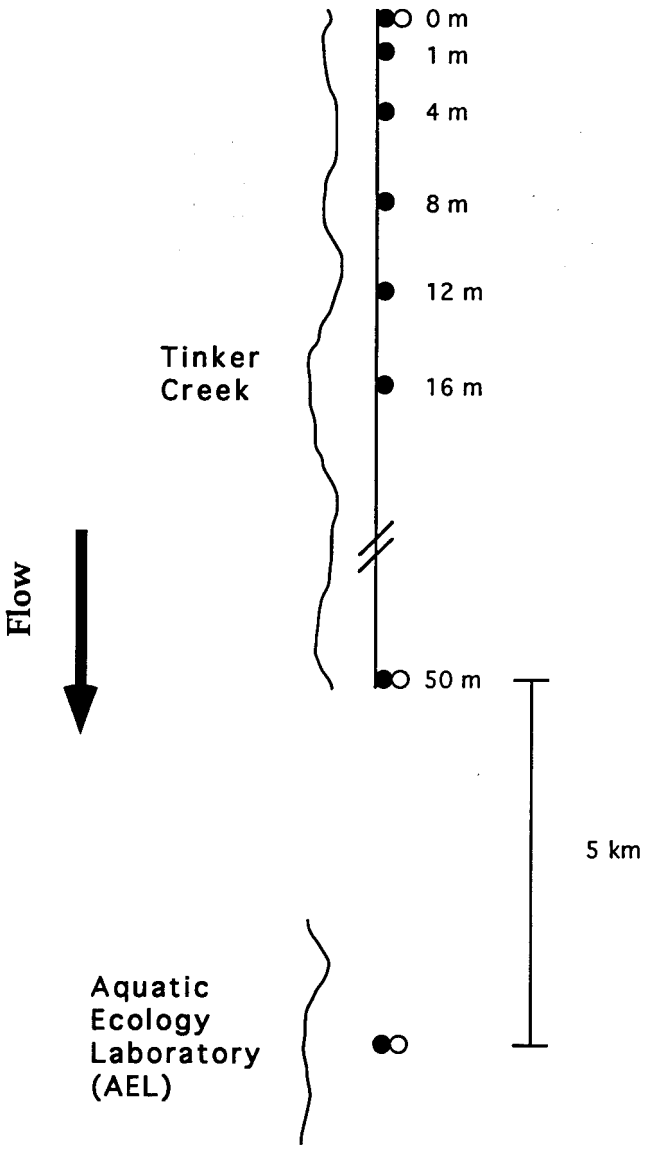

FIG. 1. Sampling scheme. Samples were collected at transect points 0 m, 1 m, 4 m, 8 m, 12 m, 16 m, and 50 m along the bank of Tinker Creek; the AEL, approximately 5 km downstream from the transect, was also sampled. Symbols: ●, sediment sampling sites; ○, water column sampling sites. Five replicate samples (a to e) were taken from each sampling site.

TABLE 1. Genetic diversity of *B. cepacia* isolated from UTR

Sample(s)	No. of isolates	Genetic diversity (h)[a]								
		FUM	G6PDH-1	G6PDH-2	IDH	MDH	6PGDH	PGI	PGM	Mean (SE)
Individual-site sediment										
T0	20	0.660	0.686	0.817	0.000	0.209	0.608	0.660	0.725	0.546 (0.101)
T1	20	0.830	0.673	0.804	0.209	0.000	0.542	0.627	0.765	0.556 (0.106)
T4	20	0.810	0.686	0.810	0.209	0.111	0.471	0.739	0.399	0.529 (0.097)
T8	20	0.830	0.680	0.771	0.366	0.000	0.660	0.621	0.739	0.583 (0.097)
T12	20	0.816	0.713	0.809	0.000	0.000	0.699	0.765	0.404	0.526 (0.124)
T16	20	0.753	0.616	0.779	0.337	0.000	0.658	0.653	0.647	0.555 (0.092)
T50	20	0.795	0.696	0.713	0.351	0.000	0.520	0.591	0.503	0.521 (0.089)
AEL	20	0.739	0.817	0.758	0.366	0.000	0.621	0.791	0.706	0.600 (0.099)
All sediment	160	0.778	0.735	0.793	0.271	0.047	0.647	0.700	0.652	0.578 (0.096)
Water										
T0	16	0.808	0.423	0.756	0.282	0.000	0.526	0.603	0.679	0.510 (0.095)
T50	17	0.792	0.775	0.733	0.000	0.125	0.517	0.725	0.692	0.545 (0.110)
AEL	20	0.795	0.500	0.821	0.337	0.100	0.279	0.700	0.558	0.511 (0.091)
All water	53	0.779	0.602	0.788	0.223	0.082	0.419	0.690	0.641	0.528 (0.092)
All sediment and water	213	0.780	0.729	0.795	0.268	0.060	0.613	0.701	0.649	0.574 (0.094)

[a] Genetic diversity is expressed as $h = n(1 - \Sigma x_i^2)/(n - 1)$, where x_i is the frequency of the ith allele and n is the number of isolates in the sample. FUM, fumarate hydratase; G6PDH-1 and G6PDH-2, glucose-6-phosphate dehydrogenases; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; 6PGDH, 6-phosphogluconate dehydrogenase; PGI, phosphoglucose isomerase; PGM, phosphoglucomutase.

unincorporated nucleotides were separated from the labeled probe by using a NucTrap push column (Stratagene).

Filters were prehybridized in 5× SSC–0.5% blocking reagent–0.1% sodium *N*-lauroylsarcosine–0.02% sodium dodecyl sulfate (SDS) at 47°C for 2 h. Hybridization was carried out overnight in the same solution with a probe concentration of approximately 20 ng/ml at 47°C. The filters were washed twice for 5 min each time in 2× SSC–1% SDS at room temperature and three more times for 20 min each time in 1× SSC–1% SDS at 47°C. Hybridization results were visualized by using the Lumi-Phos system as recommended by the manufacturer (Boehringer Mannheim). The optimal hybridization conditions were determined empirically by using *Pseudomonas putida* (ATCC 12633) and *B. solanacearum* (ATCC 11696) as negative controls and the *B. cepacia* type strain (ATCC 24561) as a positive control. The target sequence of *P. putida* differs from that of *B. cepacia* at 9 of the 18 nucleotides (28). *B. solanacearum* is a close relative of *B. cepacia*; both were members of the group II rRNA homology cluster of pseudomonads that was recently reclassified into the genus *Burkholderia* (1). The *B. cepacia*, *P. putida*, and *B. solanacearum* type strains were included on all filters to control for the accuracy of the hybridization procedure.

At each sediment sampling site, 20 positive *B. cepacia* strains were retained, 4 from each of the five subsamples. The water samples contained considerably fewer cells: only 16 strains were isolated from the 25 ml of water sampled at the 0-m transect point, and 17 were collected from the 50-m point. Twenty strains were isolated from the water collected at AEL. In total, 213 isolates were examined in this study. All strains were frozen at −70°C in nutrient broth containing 20% glycerol.

MLEE. *B. cepacia* isolates were inoculated into 2 ml of Nutrient Broth (Difco) and grown overnight. The culture was transferred to a 1.5-ml microcentrifuge tube and centrifuged for 5 min at 2,000 × *g* to pellet the cells. Cells were resuspended in 150 μl of 0.1 M Tris-HCl (pH 7.6). The suspension was then subjected to sonication on ice at 5% of maximum power (Heat Systems-Ultrasonics, Inc., model W-380) for two 10-s periods separated by a 10-s interval. The lysate was then centrifuged again for 2 min at maximum speed to pellet cell debris, and the supernatant was frozen at −70°C.

Electrophoresis was performed on a cellulose acetate medium as recommended by the manufacturer (13). Cell lysate was applied to cellulose acetate plates, and electrophoresis was carried out in Tris-glycine buffer at 200 V for 15 min. Gels were then immediately stained for enzyme activity as recommended by the manufacturer (13). Eight enzyme loci were examined: fumarate hydratase, glucose-6-phosphate dehydrogenases 1 and 2, isocitrate dehydrogenase, malate dehydrogenase, 6-phosphogluconate dehydrogenase, phosphoglucose isomerase, and phosphoglucomutase. Distinctive electromorphs were numbered in order of decreasing anodal migration. Many isolates exhibited a minor band with isocitrate dehydrogenase activity and inconsistent staining intensity that was hence ignored. Each gel run included a sample of the *B. cepacia* type strain. The accuracy of electromorph designations was confirmed by exhaustive comparison to the reference strain. Each isolate was characterized by its combination of eight electromorphs, and the distinctive profiles were designated as multilocus genotypes (also called ETs).

Statistical analysis. Electromorphs and ETs are equated with alleles and allelic combinations, respectively, for statistical analysis. Computer programs written by T. S. Whittam were used to analyze the data for ET designations, genetic diversity calculations, and ET clustering. The genetic diversity (h) per locus, j, is expressed as $h_j = n(1 - \Sigma x_i^2)/(n - 1)$, where x_i is the frequency of ith allele and n is the number of individuals in the sample. The expression $n/(n - 1)$ is a correction for bias in small population sizes (24, 30). Mean genetic diversity (H) is the arithmetic average of h over all of the loci examined.

A Mantel analysis was used to examine the relationship between matrices of genetic and geographic distances (matrix comparison procedure, NTSYS-pc, version 1.80). The ET designations were also subjected to a canonical discriminant analysis (27). Mahalonobis distances (measures of dissimilarity between populations) were calculated for sediment sampling sites. A dendrogram showing the genetic relatedness between *B. cepacia* isolates was produced by using the average-linkage clustering algorithm from a matrix of coefficients of pairwise genetic distances (34). One hundred sediment isolates, those collected from transect points T0 to T12, were used in this cluster analysis.

Multilocus linkage disequilibrium in haploid populations is calculated on the basis of the distribution of allelic mismatches between pairs of bacterial isolates over all of the loci examined. The ratio of the observed variance in mismatches, V_o, to the expected variance in a corresponding population at linkage equilibrium, V_e, provides a measure of multilocus linkage disequilibrium that can be expressed as follows as the index of association (I_A): $I_A = V_o/V_e - 1$ (2, 3, 20, 35). To determine whether V_o is statistically different from V_e in any sample, i.e., I_A is significantly greater than zero, a Monte Carlo procedure can be iterated which randomly scrambles the alleles to eliminate any effect of linkage disequilibrium (35). Significance is calculated as the probability of obtaining, by chance alone, a variance in the number of mismatches as extreme as or more extreme than that actually observed. Computer programs written by T. Nguyen were used to calculate V_o and V_e and to perform the Monte Carlo simulations.

RESULTS

ETs and genetic diversity. All of the enzyme loci studied were polymorphic, and the number of alleles ranged from two (isocitrate dehydrogenase) to six (glucose-6-phosphate dehydrogenase 2). The average number of alleles per locus was 4.38. Among the 213 *B. cepacia* isolates examined, 164 unique ETs were identified. The most common clone was found at a frequency of 4.7% (10 of the 213 isolates). This ET was found

TABLE 2. Mahalonobis distances[a] between subpopulations of *B. cepacia* isolated from the sediments of UTR

Site	Mahalonobis distance							
	T0	T1	T4	T8	T12	T16	T50	AEL
T0	0.00	0.40	0.92	0.44	1.23	0.75	1.52	1.04
T1		0.00	1.13	0.72	1.96	1.63	1.52	1.08
T4			0.00	0.55	0.88	0.80	0.73	0.36
T8				0.00	1.55	0.82	0.91	0.59
T12					0.00	1.13	2.23	1.51
T16						0.00	0.88	1.68
T50							0.00	1.70
AEL								0.00

[a] Mahalonobis distances were generated from the discriminant analysis. Lower values indicate genetically more similar subpopulations.

in the sediments at transect points T0, T4, T8, T12, T16, and T50 and in the water collected from the AEL site.

A survey of the genetic diversity of the individual samples is given in Table 1. The most diverse locus was glucose-6-phosphate dehydrogenase 2 ($H = 0.795$), and the least diverse was malate dehydrogenase ($H = 0.060$). The mean genetic diversity of all samples was 0.574. Isolates collected from the water column were less diverse ($H = 0.528$) than those collected from the sediments ($H = 0.578$). Among the individual sampling sites, the most diverse population was the sediment sample collected at AEL ($H = 0.600$), but differences between the individual sites were not statistically significant.

Genetic distance analysis. Mantel analysis of the 160 sediment samples showed that, overall, geographic and genetic distances were not correlated regardless of whether the data were log transformed ($P = 0.4670$) or not ($P = 0.4470$). Mahalanobis distances did, however, reveal significant differences in genetic distance between groups of isolates at different sites. A matrix showing the distance between the eight groups of 20 isolates at the sediment sampling sites is given in Table 2. In this type of analysis, more similar subpopulations have lower Mahalonobis distance values. Surprisingly, the most similar sites were the T4 transect point and AEL, which are separated by >5 km. A schematic representation of the Mahalonobis distance values is shown in Fig. 2. The isolates from transect points T0, T1, T4, and T8 are all quite similar to each other and to isolates from the sediments at AEL. On the other hand, *B. cepacia* isolated from transect points T12, T16, and T50 are not very similar to each other and only slightly similar to some of the upstream sampling points.

Linkage disequilibrium analysis. The complete set of isolates and subsets of the population were analyzed for multilocus linkage disequilibrium (Table 3). Among the 213 isolates, 22,578 pairwise comparisons were possible. On average, the isolates differed at 4.39 of the 8 loci examined. The ratio of the observed variance in the number of mismatches (V_o) to that expected under the null hypothesis of linkage equilibrium (V_e) was 1.25; therefore, the I_A was 0.25. The Monte Carlo procedure indicated that the difference between the observed and expected values was significant ($P < 0.001$ on the basis of 10,000 iterations). This indicates that there is some level of linkage disequilibrium between the eight loci studied.

Analysis of disequilibrium among isolates collected from sediment and water revealed I_A values of 0.28 ($P < 0.001$) and 0.21 ($P < 0.01$), respectively. Both of these I_A values differ significantly from zero. There is, therefore, linkage disequilibrium in both waterborne cells and those isolated from sediments.

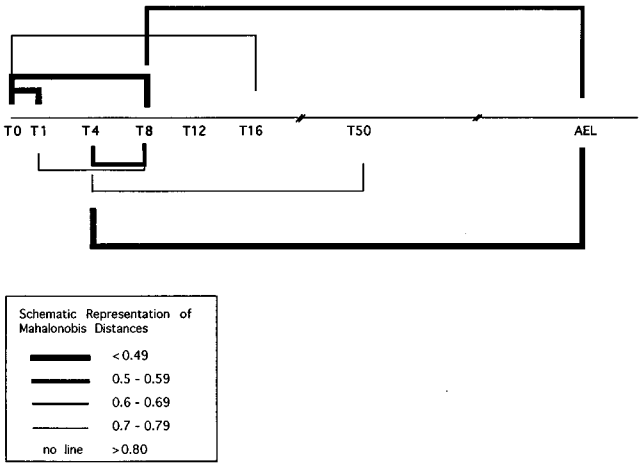

FIG. 2. Schematic representation of Mahalonobis distances. The genetic distances between subsets of the *B. cepacia* population are shown by lines connecting the eight sediment sampling sites. The thicknesses of the connecting lines represent ranges in the Mahalonobis distance values as given in the distance matrix (Table 2). Subpopulations connected by thicker lines (lower Mahalonobis distance values) indicate genetically more similar groups.

Multilocus linkage disequilibrium values of subpopulations present at the individual sampling sites varied considerably. For example, for the 20 strains isolated from the sediment at the T0 transect point, 190 pairwise comparisons were possible and the mean number of allelic mismatches was 4.26. The I_A was 0.00 ($P < 0.49$), indicating that the alleles at the eight loci are randomly assorted. At the other extreme, the 190 pairs of comparisons between cells isolated from T12 yielded an average of 3.89 mismatches and the I_A was 0.86 ($P < 0.001$). This site yielded some evidence of clonality, as all four bacterial strains collected from one of the replicate samples (T12e) had

TABLE 3. Multilocus linkage disequilibrium analysis of *B. cepacia* isolated from UTR[a]

Sample(s)	No. of pairs of isolates	X	V_o	V_e	I_A	P
Individual-site sediment						
T0	190	4.26	1.44	1.44	0.00	<0.49
T1	190	4.54	1.75	1.41	0.24	<0.03
T4	190	4.06	2.19	1.47	0.49	<0.01
T8	190	4.63	2.01	1.45	0.38	<0.01
T12	190	3.89	2.32	1.25	0.86	<0.001
T16	190	4.44	2.13	1.50	0.42	<0.01
T50	190	4.08	2.17	1.55	0.40	<0.01
AEL	190	4.77	1.56	1.39	0.12	<0.03
All sediment	12,720	4.43	1.88	1.47	0.28	<0.001
Water						
T0	120	3.81	2.64	1.50	0.77	<0.01
T50	136	4.29	1.78	1.31	0.35	<0.02
AEL	190	4.09	1.86	1.53	0.21	<0.10
All water	1,378	4.15	1.82	1.50	0.21	<0.01
All sediment and water	22,578	4.39	1.86	1.48	0.25	<0.001

[a] X, mean number of allelic mismatches between pairs of isolates among eight loci; V_o, observed variance in number of mismatches among pairs of isolates; V_e, expected variance assuming linkage equilibrium (the null hypothsis); $I_A = V_o/V_e - 1$; P, probability of rejecting by chance alone the null hypothesis that V_o equals V_e (on the basis of 10,000 iterations of the Monte Carlo procedure).

identical ETs. This repeated isolation of the same clone contributed substantially to the relatively high I_A value (especially since the sample size, at 20 isolates, was relatively small).

Dendrogram. The genetic relationship among the isolates collected from the sediments at sites T0, T1, T4, T8, and T12 is shown in the dendrogram in Fig. 3. Among this subset of the population (100 isolates), 79 unique ETs were identified. The dendrogram, generated by the average-linkage method of clustering, shows the extensive branching typical of highly recombinant populations (38). There appears to be no obvious relationship between clustering and the site of collection.

DISCUSSION

The study of allelic variation by MLEE has yielded significant insights into the genetic diversity, systematics, and population structure of bacteria. One of the most fundamental discoveries of a general nature relating to all bacteria thus far examined is that the amount of genetic variation carried by bacteria is significantly greater than that of higher eukaryotes. For example, the mean genetic diversity per enzyme locus (H) for *E. coli* hovers around 0.50, whereas it is 0.07 for humans (33). *B. cepacia* is one of the most commonly cultivated bacteria in the UTR study site (15a). A previous study with a smaller sample size focusing on environmental *B. cepacia* has shown it to have a mean genetic diversity per locus of between 0.54 and 0.70 (21). The value calculated from the isolates collected in this study, 0.574, confirms the relatively high degree of diversity in this species.

Genetic diversity values at the individual sampling sites did not vary significantly along the stream continuum. McArthur et al. (21) found that genetic diversity increased with environmental variability in soil-borne forms of *B. cepacia*. This finding contradicts the generally held belief that allelic variation in bacteria is selectively neutral (41) and suggests a pattern of microgeographical adaptation due to selection. For heterotrophic environmental bacteria, higher organic matter diversity is proportional to increased habitat variability (22). The river continuum concept of Vannote et al. (36) predicts that organic matter diversity will be higher in the headwaters of a stream than in the mouth because of the increased linkage between terrestrial and aquatic systems. If habitat variability affects allozymes, then the most diverse populations should be those sampled farthest upstream. In this study, the most diverse sampling site was the sediment collected from AEL (H = 0.600), the site farthest downstream, but it was not significantly different from the other sampling sites. This suggests three possibilities: (i) habitat variability is not significantly different along the length of the stream (at least in the area sampled), (ii) habitat variability was not manifested as genetic diversity in the eight enzymes examined (i.e., the enzyme polymorphisms in this case are indeed selectively neutral), or (iii) the physical characteristics of the stream, especially the constant flow, consistently move the bacteria downstream in such a way that the cells are only transiently associated with one particular area. If this is the case, bacterial cells may not have time to adapt to any selective pressure exerted by environmental variability at any one spatial area (unlike the relatively stationary soil-borne bacteria in the study of McArthur et al.).

Considering the dynamics of lotic ecosystems, it is not surprising that a Mantel analysis showed no overall correlation between genetic and geographic distances. The movement of bacteria and their genes in the water column must effectively mix the population in a way unlike terrestrial or lentic ecosystems. Furthermore, lateral inputs of bacteria from the banks along the stream continuum would likely disrupt any discern-

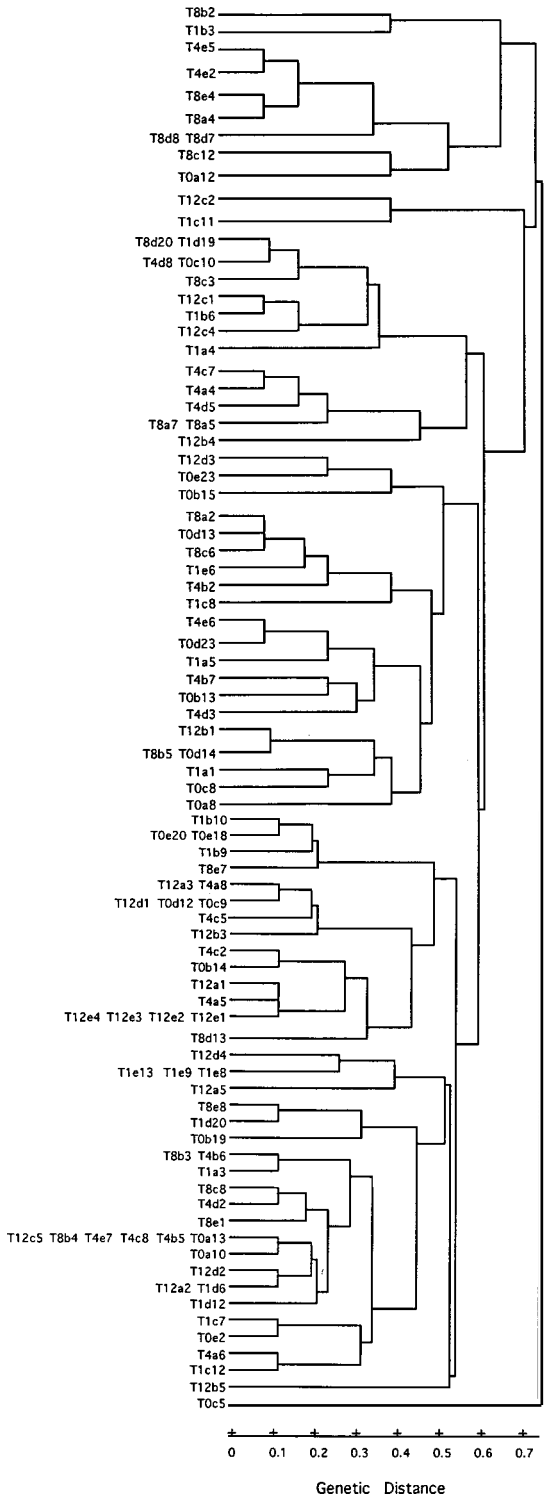

FIG. 3. Dendrogram showing the genetic relationships among 100 *B. cepacia* isolates collected from the sediments at transect points T0, T1, T4, T8, and T12. Clustering was performed by the average-linkage algorithm (33) from a matrix of pairwise genetic distances. Among the 100 isolates, 79 unique ETs were discovered. Each isolate name designates the transect point and replicate from which the isolate was collected; i.e., T8b2 is isolate 2 from replicate sample b from the sediment at the 8-m point at Tinker Creek.

ible correlation if newly input bacteria can easily proliferate and colonize new areas. However, information about the genetic dynamics of bacteria in streams was obtained by measuring the dissimilarity of individual subpopulations. Mahalonobis distances revealed significant differences between subpopulations isolated from the sediments. As one might expect, the transect points separated by just 1 m (T0 and T1) had two of the most similar subpopulations (Fig. 2 and Table 2). Indeed, the subpopulations of *B. cepacia* isolated between points T0 and T8 were all quite similar genetically. However, isolates from T12, T16, and T50 were very different from each other and from downstream isolates and only slightly similar to upstream isolates. Isolates from AEL, however, over 5 km downstream from the transect, showed a genetic makeup similar to that of cells collected from T4 and T8. This finding that populations are localized in niches along the stream suggests a certain patchiness in lotic ecosystems. If ecotypic factors affect allozymes, then there should be similarities between the habitat at AEL and that at T4 and T8. T12, T16, and T50 must be somehow different. The environmental parameters that may affect this genetic variability are unknown.

Many investigators using MLEE as a tool to assess bacterial population structure have concluded that bacteria have a clonal population structure. This may be expected, since bacteria reproduce asexually, by binary fission. Apparently, the prevalence of the parasexual mechanisms of genetic transfer (transformation, conjugation, and transduction) is not high enough to disrupt the overall clonal structure of most bacterial populations. Yet, most of the studies that have reported clonal population structures have involved pathogens and commensal bacterial species. The ecology of environmental bacteria (like lotic *B. cepacia* biotypes) is quite different from that of the typical pathogen. Habitat quality is probably significantly more variable for most environmental bacteria. In a highly variable habitat, recombination could be a considerable selective advantage. Bacterial populations that could draw from a common gene pool would be able to develop successful genotypic combinations (instead of depending on mutational events) in the face of selective pressure exerted by habitat variability. This would certainly lead to a competitive advantage for these species.

Whittam (38) has described the contrast between clonal and freely recombining haploid populations by using both a dendrogram showing the genetic relatedness between strains and a histogram showing the distribution of allele mismatches (found by comparing the allozyme profile of all pairs of isolates in the population and calculating the number of allele mismatches between each pair). Typical clonal populations with limited genetic exchange (in which variation is produced only by mutation) are characterized by a deep but limited branching pattern of the dendrogram and an allele mismatch distribution that is bimodal in shape and exhibits extensive variance. Allozyme analysis of a freely recombining population typically results in a dendrogram that shows extensive branching (making it appear more bush-like than tree-like), reflecting the randomization of haploid genotypes. Recombination severely diminishes the occurrence of linkage disequilibrium, and the mismatch distribution becomes unimodal in shape with limited variance. In this way, recombination tends to limit the divergence of genotypes among cell lineages in a population, since genes are consistently exchanged among population members.

The population structure of *B. cepacia* resembles the pattern typical of a freely recombining population. The allele mismatch distribution for all 213 *B. cepacia* isolates (Fig. 4) is unimodal and exhibits limited variance. The dendrogram in Fig. 3, showing the genetic relationship between 100 *B. cepacia*

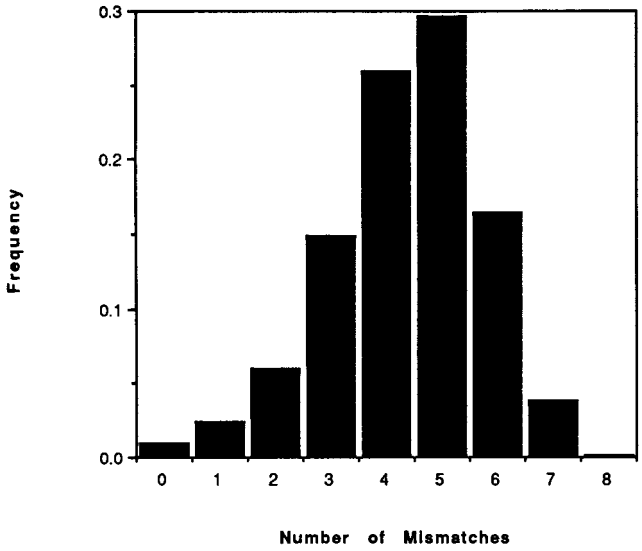

FIG. 4. Allele mismatch distribution among 213 lotic *B. cepacia* isolates. The frequency distribution of the number of mismatches between pairs of isolates at the eight enzyme loci studied is shown. Overall, 22,578 comparisons were possible and the mean number of mismatches was 4.39.

isolates collected within 12 m of each other, exhibits an extensive branching pattern, revealing the genotypic variation characterized by frequent recombination. This genetic exchange randomized the alleles at the eight structural genes examined to the extent that there was no clustering of the isolates based on site of collection. Since geographic isolation can give rise to linkage disequilibrium (35), this suggests that different *B. cepacia* genotypes come in contact with each other frequently enough to exchange genetic material (at least within the 12-m reach of the stream examined). Lotic ecosystems seem to provide an opportunity for bacterial cells of differing types to become associated with each other in a manner favorable for recombination.

Evidence of modest clonal proliferation is provided by the multilocus linkage disequilibrium calculations that reveal a significant level of association between alleles when the entire population sampled is subjected to analysis. The I_A for this population was 0.25, and the Monte Carlo procedure indicated that the value was significantly greater than zero (the value expected for absolute equilibrium). Given that binary fission is the primary mode of reproduction in bacteria, it is not unexpected to find some level of clonality. The subpopulations at the individual sampling sites showed a considerable range of I_A values: the I_A for isolates from T0 did not differ significantly from zero, the signature of a panmictic population; however, at site T12, the I_A was 0.86. The repeated isolation of one ET from one of the replicate samples (T12e) contributed substantially to the relatively high I_A value in the population at this site.

Figure 5 compares the I_A values for some of the bacterial species and genera studied to date. The large variation in the extent of linkage disequilibrium of bacteria in nature is evident. One pathogenic microorganism, *N. gonorrhoeae*, displays a completely panmictic population structure with no evidence of clonality. Recombination in *N. gonorrhoeae*, which is naturally competent, is most likely due to transformation, the only known mechanism for the transfer of chromosomal genes in this species (25). *B. subtilis*, another organism shown to have a reduced level of clonality, is also naturally transformable (14).

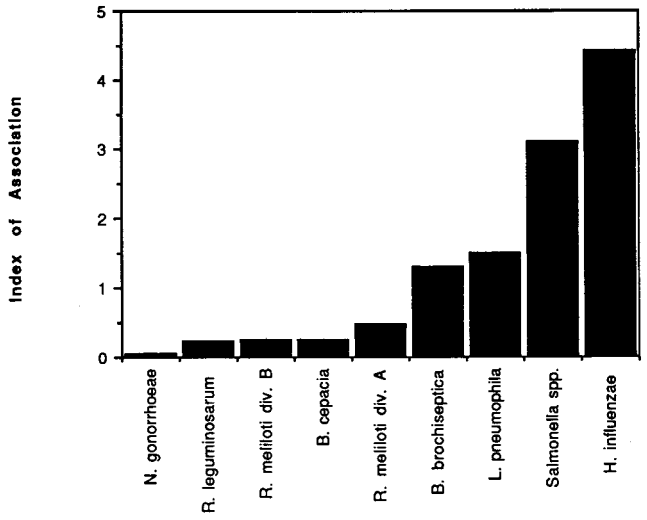

FIG. 5. Linkage disequilibrium in selected bacterial populations. The I_A values for some bacterial populations are shown. I_A values are those calculated by using all isolates (not ETs only). Values are from Maynard Smith et al. (20), except those for *R. leguminosarum* (34) and *B. cepacia* (this study).

Natural competency, however, doesn't necessarily lead to frequent recombination, since *H. influenzae*, also naturally transformable, exhibits a clonal population structure. The I_A value for the *B. cepacia* isolates (an organism not known to be naturally competent) examined in this study is at the low end of the scale, implying that this species, along with *Rhizobium* and *Neisseria* spp., has a highly promiscuous nature. The mechanism(s) of exchange in this population of lotic *B. cepacia* is unknown. Gene exchange via transduction and conjugation has been reported in certain *B. cepacia* strains (7).

Indeed, the relatively extensive amount of recombination in lotic *B. cepacia* implied by the results of this study may be typical of environmental bacteria. Bacteria adapted to live in the soil or sediments may encounter divergent strains in their particular niche more often than pathogenic bacterial forms, which may never occur as a mixed infection (with the notable exception of *N. gonorrhoeae*). Furthermore, any exchange between cells with identical genotypes would be undetectable. Interestingly, a previous study of 31 *B. cepacia* biotypes isolated in a nosocomial setting during an outbreak of infections linked to an intrinsically contaminated povidone-iodine solution revealed identical allozyme profiles (4). This indicates that single bacterial species may be able to exhibit differing types of population structures depending on environmental conditions. The balance struck between clonal proliferation and recombination probably varies widely in bacterial populations (14). A freely recombining population could, by stochastic processes, produce certain highly fit genotypes. During periods of stress, natural selection would favor such genetic combinations and under better conditions those genotypes may be able to propagate extensively in a clonal manner. In this sense, bacteria may "have the best of both worlds," as the genotypes of successful individuals are not obligatorily broken up by sex, yet recombination, in times of stress, would provide an extensive gene pool from which to draw. Although this work revealed some modest evidence of clonal proliferation and compelling evidence of recombination, further studies evaluating the temporal variation of *B. cepacia* genotypes are needed to draw final conclusions on the prevalence and frequency of genetic exchange in this species.

ACKNOWLEDGMENTS

We thank M. Moran for providing comments on the manuscript and A. Homsey and M. Williamson for the use of their computer.

This research was supported by contract DE-AC07-76SROO-819 between the U.S. Department of Energy and the University of Georgia's Savannah River Ecology Laboratory.

REFERENCES

1. **Anonymous.** 1993. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List no. 45. Int. J. Syst. Bacteriol. **43:**398–399.
2. **Brown, A. H. D., M. W. Feldman, and E. Nevo.** 1980. Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics **96:**523–536.
3. **Brown, A. H. D., and M. W. Feldman.** 1981. Population structure of multilocus associations. Proc. Natl. Acad. Sci. USA **78:**5913–5916.
4. **Carson, L. A., R. L. Anderson, A. L. Panlilio, C. M. Beck-Sague, and J. M. Miller.** 1991. Isoenzyme analysis of *Pseudomonas cepacia* as an epidemiologic tool. Am. J. Med. **91**(Suppl. 3B)**:**252S–255S.
5. **Caugant, D. A., B. L. Levin, and R. K. Selander.** 1981. Genetic diversity and temporal variation in the *Escherichia coli* population of a human host. Genetics **98:**467–490.
6. **Caugant, D. A., B. R. Levin, and R. K. Selander.** 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. J. Hyg. **92:**377–384.
7. **Cheng, H., and T. G. Lessie.** 1994. Multiple replicons constituting the genome of *Pseudomonas cepacia* 17616. J. Bacteriol. **176:**4034–4042.
8. **Dykhuizen, D. E., and L. Green.** 1991. Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. **173:**7257–7268.
9. **Forman, R. T. T., and M. Godron.** 1981. Patches and structural components for a landscape ecology. Bioscience **31:**733–740.
10. **Gargallo-Viola, D.** 1989. Enzyme polymorphism, prodigiosin production, and plasmid fingerprints in clinical and naturally occurring isolates of *Serratia marcescens*. J. Clin. Microbiol. **27:**860–868.
11. **Hagedorn, C., W. D. Gould, T. R. Bardinelli, and D. R. Gustavson.** 1987. A selective medium for enumeration and recovery of *Pseudomonas cepacia* biotypes from soil. Appl. Environ. Microbiol. **53:**2265–2268.
12. **Hartl, D. L., and D. E. Dykhuizen.** 1984. The population genetics of *Escherichia coli*. Annu. Rev. Genet. **18:**31–68.
13. **Hebert, P. D. N., and M. J. Beaton.** 1989. Methodologies for allozyme analysis using cellulose acetate electrophoresis: a practical handbook. Helena Laboratories, Beaumont, Tex.
14. **Istock, C. A., K. E. Duncan, N. Ferguson, and X. Zhou.** 1992. Sexuality in a natural population of bacteria—*Bacillus subtilis* challenges the clonal paradigm. Mol. Ecol. **1:**95–103.
15. **Leff, L. G.** 1994. Stream bacterial ecology: a neglected field? ASM News **60:**135–138.

15a. **Leff, L. G.** Unpublished data.

16. **Leff, L. G., R. M. Kernan, J V. McArthur, and L. J. Shimkets.** 1995. Identification of aquatic *Burkholderia* (*Pseudomonas*) *cepacia* by hybridization with species-specific rRNA gene probes. Appl. Environ. Microbiol. **61:**1634–1636.
17. **Loos, B. G., D. W. Dyer, T. S. Whittam, and R. K. Selander.** 1993. Genetic structure of populations of *Porphyromonas gingivalis* associated with periodontitis and other oral infections. Infect. Immun. **61:**204–212.
18. **Maltby, L.** 1992. Heterotrophic microbes, p. 165–194. *In* P. Calow and G. Petts (ed.), The rivers handbook. Blackwell Scientific Publications, Oxford.
19. **Maynard Smith, J., and C. G. Dowson.** 1991. Localized sex in bacteria. Nature (London) **349:**29–31.
20. **Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. Spratt.** 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA **90:**4384–4388.
21. **McArthur, J V., D. A. Kovacic, and M. H. Smith.** 1988. Genetic diversity in natural populations of a soil bacterium across a landscape gradient. Proc. Natl. Acad. Sci. USA **85:**9621–9624.
22. **McArthur, J V., L. G. Leff, and M. H. Smith.** 1992. Genetic diversity of bacteria along a stream continuum. J. N. Am. Benthol. Soc. **11:**269–277.
23. **Meyer, J. L.** 1990. A blackwater perspective on riverine ecosystems. Bioscience **40:**643–651.
24. **Nei, M.** 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89:**583–590.
25. **O'Rourke, M., and E. Stevens.** 1993. Genetic structure of *Neisseria gonorrhoeae* populations: a non-clonal pathogen. J. Gen. Microbiol. **139:**2603–2611.
26. **Pringle, C. M., R. J. Naiman, G. Bretschko, J. R. Karr, M. W. Oswood, J. R. Webster, R. L. Welcomme, and M. J. Winterbourn.** 1988. Patch dynamics in lotic systems: the stream as a mosaic. J. N. Am. Benthol. Soc. **7:**503–524.
27. **SAS Institute.** 1985. SAS users guide: statistics. SAS Institute, Cary, N.C.
28. **Schleifer, K. H., R. Amann, W. Ludwig, C. Rothemund, N. Springer, and S. Dorn.** 1992. Nucleic acid probes for the identification and in situ detection of pseudomonads, p. 127–134. *In* E. Galli, S. Silver, and B. Witholt (ed.), *Pseudomonas*: molecular biology and biotechnology. American Society for Microbiology, Washington, D.C.

29. **Selander, R. K., P. Beltran, and N. H. Smith.** 1990. Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. Infect. Immun. **58:**2262–2275.
30. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for bacteria population genetics and systematics. Appl. Environ. Microbiol. **51:** 873–884.
31. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
32. **Selander, R. K., and B. R. Levin.** 1980. Genetic diversity and structure in *Escherichia coli* populations. Science **120:**545–547.
33. **Selander, R. K., J. M. Musser, D. A. Caugant, M. N. Gilmour, and T. S. Whittam.** 1987. Population genetics of pathogenic bacteria. Microb. Pathog. **3:**1–7.
34. **Sneath, P. H. A., and R. R. Sokol.** 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco.
35. **Souza, V., T. T. Nguyen, R. R. Hudson, D. Pinero, and R. E. Lenski.** 1992. Hierarchical analysis of linkage disequilibrium in *Rhizobium* populations: evidence for sex? Proc. Natl. Acad. Sci. USA **89:**8389–8393.
36. **Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing.** 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. **37:** 130–137.
37. **Whittam, T. S.** 1983. Geographic components of linkage disequilibrium in natural populations of *Escherichia coli*. Mol. Biol. Evol. **1:**67–83.
38. **Whittam, T. S.** 1992. Sex in the soil. Curr. Biol. **2:**676–678.
39. **Whittam, T. S., H. Ochman, and R. K. Selander.** 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80:**1751–1755.
40. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.
41. **Young, J. P. W.** 1989. The population genetics of bacteria, p. 417–438. *In* D. A. Hopwood and K. F. Chater (ed.), Genetics of bacterial diversity. Academic Press, London.