JOURNAL OF BACTERIOLOGY, Sept. 1985, p. 1021–1037
0021-9193/85/091021-17$02.00/0

Vol. 163, No. 3

Genetic Structure of Populations of *Legionella pneumophila*

ROBERT K. SELANDER,[1]* ROGER M. McKINNEY,[2] THOMAS S. WHITTAM,[3] WILLIAM F. BIBB,[2] DON J. BRENNER,[2] FREDERICK S. NOLTE,[4] AND PATRICIA E. PATTISON[1]

Departments of Biology[1] and Microbiology,[4] University of Rochester, Rochester, New York 14627; Division of Bacterial Diseases, Centers for Disease Control, Atlanta, Georgia 30333[2]; and Department of Biology, Pennsylvania State University, University Park, Pennsylvania 16802[3]

Received 5 January 1985/Accepted 22 June 1985

The genetic structure of populations of *Legionella pneumophila* was defined by an analysis of electrophoretically demonstrable allelic variation at structural genes encoding 22 enzymes in 292 isolates from clinical and environmental sources. Nineteen of the loci were polymorphic, and 62 distinctive electrophoretic types (ETs), representing multilocus genotypes, were identified. Principal coordinates and clustering analyses demonstrated that isolates received as *L. pneumophila* were a heterogeneous array of genotypes that included two previously undescribed species. For 50 ETs of *L. pneumophila* (strict sense), mean genetic diversity per locus was 0.312, and diversity was equivalent in ETs represented by isolates recovered from clinical sources and those collected from environmental sources. Cluster analysis revealed four major groups or lineages of ETs in *L. pneumophila*. Genetic diversity among ETs of the same serotype was, on average, 93% of that in the total sample of ETs. Isolates marked by particular patterns of reactivity to a panel of nine monoclonal antibodies were also genetically heterogeneous, mean diversity within patterns being about 75% of the total. Both Pontiac fever and the pneumonic form of legionellosis may be caused by isolates of the same ET. The genetic structure of *L. pneumophila* is clonal, and many clones apparently are worldwide in distribution. The fact that *L. pneumophila* is only 60% as variable as *Escherichia coli* raises the possibility that isolates recovered from clinical cases and man-made environments are a restricted subset of all clones in the species as a whole.

Legionella pneumophila, a gram-negative bacterium isolated, characterized, and described after an epidemic of pneumonia at an American Legion Convention in Philadelphia in 1976 (1), is one of an increasing number of phenotypically similar species that have been identified or implicated as etiological agents of Legionnaires disease and Pontiac fever (36). *L. pneumophila* occurs widely in natural freshwater habitats (13, 14) and has secondarily taken up residence in air-conditioning cooling towers, water tanks, and similar man-made aquatic habitats in hospitals, hotels, and other buildings, where it forms reservoirs for human infection (19, 54).

Although many aspects of the biology of *L. pneumophila* have been studied (1, 55), virtually nothing is known of the nature and extent of genetic variation and structure in natural populations, apart from information provided by serotyping and monoclonal antibody typing, comparisons of small numbers of strains with respect to electrophoretic profiles of soluble proteins (30) and β-lactamases (32), and limited surveys of plasmid contents (7, 24). Estimates of overall genetic relatedness of strains from DNA hybridization experiments have been used to define species limits within the genus *Legionella* (2) but have not provided a basis for the classification of strains within species. Consequently, research on *L. pneumophila* has been conducted almost entirely within the framework of a classification based on serotyping (35).

We here report the results of an electrophoretic analysis of protein polymorphism in *L. pneumophila*. From comparisons of the electrophoretic mobilities of enzymes encoded by 22 structural genes, we have discovered that isolates currently classified as *L. pneumophila* are a strongly heterogeneous array of multilocus genotypes that includes two undescribed cryptic species. The species *L. pneumophila* (strict sense) is highly polymorphic and clonal in population structure, with certain cell lines having intercontinental if not worldwide distributions. Serotypes and monoclonal antibody reactivity patterns (MAPs) are shown to have little relationship to the underlying genetic structure. Our analysis provides a basic population genetic framework for studies of phenotypic and other characters in *L. pneumophila* and a high-resolution marker system for epidemiological and systematic research.

MATERIALS AND METHODS

Isolates. This study was based primarily on a sample of 170 isolates, from a wide variety of sources, received as *L. pneumophila* (Table 1). Most isolates were obtained from collections maintained by the Division of Bacterial Diseases, Centers for Disease Control, Atlanta, Ga., but some isolates were provided by the University of Rochester Medical Center, Rochester, N.Y.

Additional material examined included (i) 100 isolates that were collected at the Wadsworth Medical Center and several other hospitals in Los Angeles, Calif., and had been classified as *L. pneumophila* and (ii) 22 isolates of *L. pneumophila* collected in the course of an epidemiological investigation by the Centers for Disease Control of a recent outbreak of Pontiac fever in New York City.

Because of apparent confusion regarding the identity of the Dallas 1E isolate, we note that the complete Centers for Disease Control catalog designation for the culture we examined is "DA-17-G2-C1 (DA1E)"; the designation for Dallas 2E is "DA-4-G2-E2-F1-C3 (DA2E)."

Preparation of lysates for electrophoresis. Each isolate was either grown overnight (at 35°C) in 100 ml of buffered yeast extract broth, supplemented with 0.1% α-ketoglutarate, or cultured for 3 to 4 days on four buffered charcoal-yeast extract agar plates (17). Cells were harvested by centrifugation of broth cultures and by scraping from plate cultures,

* Corresponding author.

TABLE 1. Characteristics of 170 isolates of *Legionella* spp.

Species and ET[a]	Isolate	Location	Source[b]	Serotype	MAP[c]
L. pneumophila					
1	Davenport 1	U.S.A. (Iowa)	C	1	14
	Concord 4	U.S.A. (Calif.)	C	1	14
	Indianapolis 3	U.S.A. (Ind.)	C	1	6
2	B1793-79	U.S.A. (N.Y.)	C	2	
3	Binghamton 1	U.S.A. (N.Y.)	C	2	
4	D5401	France	E	1	2
5	Vasteras 2	Sweden	C	4-5	
6	Gloucester 1	U.S.A. (Mass.)	C	1	12
7	B360-80	U.S.A. (N.Y.)	C	4	
8	Portland 2	U.S.A. (Ore.)	C	1	4
9	Los Angeles 2	U.S.A. (Calif.)	C	1	12
	Orlando 1	U.S.A. (Fla.)	C	1	12
	Stockholm 1	Sweden	C	1	12
	Stockholm 2	Sweden	C	1	12
	Stockholm 3	Sweden	C	1	12
	Vasteras 1	Sweden	C	1	12
	Vasteras 3	Sweden	E	1	12
10	HHE 2	U.S.A. (N.Y.)	E	4	
11	Allentown 1	U.S.A. (Pa.)	C	1	9
12	Bloomington 3	U.S.A. (Ind.)	E	2	
13	Nagasaki 1	Japan	C	1	9
14	Togus 2	U.S.A. (Maine)	C	1	6
	Bellingham 1	U.S.A. (Wash.)	C	1	6
15	Lancaster 1	U.S.A. (Pa.)	C	1	12
16	Toronto 1	Canada	C	1	10
17	Pontiac 1	U.S.A. (Mich.)	E	1	8
	SK 267	U.S.A. (Mich.)	E	1	8
	SK 273	U.S.A. (Mich.)	E	1	8
	SK 304	U.S.A. (Mich.)	E	1	8
	Alameda 1	U.S.A. (Calif.)	C	1	12
	Berkeley 1	U.S.A. (Calif.)	C	1	9
	California 1	U.S.A. (Calif.)		1	9
	New London 1	U.S.A. (Conn.)	C	1	8
	SMH 4	U.S.A. (N.Y.)	C	1	7
	SMH 7	U.S.A. (N.Y.)	C	1	7
	Adelaide 2	Australia	C	1	10
18	Chamblee 1	U.S.A. (Ga.)	C	1	9
	Birmingham 1	U.S.A. (Ala.)	C	1	9
19	Naples 1	U.S.A. (Fla.)	C	1	12
20	Cambridge 4	England		1	4
21	OLDA	U.S.A.	C	1	4
	Tucson 1	U.S.A. (Ariz.)	C	1	4
	Bloomington 1	U.S.A. (Ind.)	E	1	4
	Burlington 2	U.S.A. (Vt.)	C	1	4
	Albuquerque 1	U.S.A. (N. Mex.)	C	1	9
	Flint 1	U.S.A (Mich.)	C	1	9
	Houston 1	U.S.A. (Tex.)	C	1	3
	RH 1	U.S.A. (Ohio)	E	1	6
	SMH 1	U.S.A. (N.Y.)	C	1	9
	SMH 2	U.S.A. (N.Y.)	C	1	9
	SMH 3	U.S.A. (N.Y.)	C	1	9
	SMH 5	U.S.A. (N.Y.)	C	1	9
	SMH 6	U.S.A. (N.Y.)	C	1	9
	HHE 1	U.S.A. (N.Y.)	E	1	4
	HHE 3	U.S.A. (N.Y.)	E	1	9
	HHE 4	U.S.A. (N.Y.)	E	1	9
	B1474-80	U.S.A. (N.Y.)	C	1	9
	Darby 1	U.S.A. (Pa.)	C	1	3
	Philadelphia 5	U.S.A. (Pa.)	C	1	12
	Muenchen 1	Germany	C	1	12
	Camperdown 1	Australia	C	1	6
	Johannesburg 1	South Africa	E	1	3
	Johannesburg 2	South Africa	E	1	4
	Johannesburg 3	South Africa	E	1	3
	Johannesburg 4	South Africa	E	1	3
	Oxford 3	England	E	NT[d]	
22	Anniston 1	U.S.A. (Ala.)	C	4-5	

Continued on following page

TABLE 1—*Continued*

Species and ET[a]	Isolate	Location	Source[b]	Serotype	MAP[c]
23	Nagasaki 2	Japan	E	1	4
	Osaka 1E	Japan	E	1	4
	Osaka 2E	Japan	E	1	13
	Lyon 1	France	C	1	9
24	San Francisco 1	U.S.A. (Calif.)	E	1	15
	San Francisco 9	U.S.A. (Calif.)	C	1	1
	Long Beach 3	U.S.A. (Calif.)	C	1	4
	Dallas 1[e]	U.S.A. (Tex.)	C	1	4
25	Philadelphia 1	U.S.A. (Pa.)	C	1	9
	Philadelphia 2	U.S.A. (Pa.)	C	1	9
	Philadelphia 3	U.S.A. (Pa.)	C	1	9
	Philadelphia 4	U.S.A. (Pa.)	C	1	9
	Portland 1	U.S.A. (Ore.)	C	4	
	Frederick 1	U.S.A. (Md.)	C	4	
	Miami Beach 1	U.S.A. (Fla.)	C	1	14
	West Palm Beach 1	U.S.A. (Fla.)	C	1	12
	Kingston 1	U.S.A. (N.Y.)	C	1	14
	B1543-79	U.S.A. (N.Y.)	C	1	14
	UH 1	U.S.A. (Ohio)	E	1	
	UPH 1	U.S.A. (Ohio)	E	1	4
	Burlington E3	U.S.A. (Vt.)		1	14
	Burlington 4	U.S.A. (Vt.)	C	3	
	Tennyson 1	U.S.A. (Ind.)	E	1	9
	Bonn 1	Germany	E	1	11
	D4999	France	E	1	12
	D5006	France	E	1	12
	D5050	France	E	1	12
	D5671	France	E	1	12
	D5811	France	E	1	12
26	Buffalo 1	U.S.A. (N.Y.)	C	1	9
27	Burlington 1	U.S.A. (Vt.)	C	1	14
28	WBH 1	U.S.A. (Ind.)	C	1	9
29	Adelaide 1	Australia	C	1	8
	Rome 1E	Italy	E	1	10
	Rome 2E	Italy	E	1	10
	Lyon Cedex 1	France	C	1	14
	Houston 3	U.S.A. (Tex.)	C	1	4
30	Concord 3	U.S.A. (Calif.)	C	8 (4-5)	
31	Leiden 1E	Netherlands	E	4-5	
	Cambridge 2	England	C	5	
	West Haven 4	U.S.A. (Conn.)	C	1	6
	West Haven 2	U.S.A. (Conn.)	E	1	6
	West Haven 3	U.S.A. (Conn.)	E	1	6
32	Leiden 2	Netherlands	C	1	12
	Leiden 3	Netherlands	C	4	
33	Stockholm 4	Sweden	E	1	6
	Stockholm 5	Sweden	E	1	6
34	Winnipeg 1	Canada	C	4-5	
35	Ottawa 6	Canada	C	1	6
36	Winnipeg 2	Canada	C	1	6
37	Ottawa 2	Canada	C	1	12
38	Berlin 1	Germany	C	1	12
39	Lyon 3	France	E	1	12
40	Togus 1	U.S.A. (Maine)	C	2	
	Togus 3	U.S.A. (Maine)	C	2	
	Atlanta 6	U.S.A. (Ga.)	C	2	
	Macon 1	U.S.A. (Ga.)	C	2	
	East Point 1	U.S.A. (Ga.)	C	2	
	Burlington 5	U.S.A. (Vt.)	C	4	
41	Atlanta 1	U.S.A. (Ga.)	C	2	
	Atlanta 2	U.S.A. (Ga.)	C	2	
	York 1	U.S.A. (Pa.)	C	2	
42	SRP 20	U.S.A. (S.C.)	E	4	
	SRP 22	U.S.A. (S.C.)	E	4	
	SRP 23	U.S.A. (S.C.)	E	4	
	SRP 26	U.S.A. (S.C.)	E	4	
	SRP 32	U.S.A. (S.C.)	E	4	
43	Bloomington 2	U.S.A. (Ind.)	E	3	
44	684 (ATCC 33733)	U.S.A. (Pa.)	E	5	
	687 (ATCC 33734)	U.S.A. (Pa.)	E	5	

Continued on following page

TABLE 1—*Continued*

Species and ET[a]	Isolate	Location	Source[b]	Serotype	MAP[c]
45	Pratt 2	U.S.A. (Kans.)	E	5	
46	Oxford 4	England	E	4-5	
47	Mt. St. Helens 6	U.S.A. (Wash.)	E	4-5	
48	Mt. St. Helens 2	U.S.A. (Wash.)	E	1	4
49	Chicago 8	U.S.A. (Ill.)	E	7	
50	Mt. St. Helens 5	U.S.A. (Wash.)	E	1	4
Species 1					
51	Oak Ridge Y	U.S.A. (Tenn.)		4-5	
52	Indianapolis 1	U.S.A. (Ind.)	C	1	9
53	Detroit 2	U.S.A. (Mich.)	C	1	4
	Dallas 1E	U.S.A. (Tex.)	E	5	
	Dallas 2E	U.S.A. (Tex.)	E	5	
54	Los Angeles 1	U.S.A. (Calif.)	C	4	
	F1735	U.S.A. (Calif.)	C	4	
	F1769	U.S.A. (Calif.)	C	4	
	F1681	U.S.A. (Calif.)	E	4	
	F1694	U.S.A. (Calif.)	E	4	
	F1700	U.S.A. (Calif.)	E	4	
	F1706	U.S.A. (Calif.)	E	4	
	F1708	U.S.A. (Calif.)	E	1	
	F1738	U.S.A. (Calif.)	E	1	
	F1740	U.S.A. (Calif.)	E	1	
	F1742	U.S.A. (Calif.)	E	1	
55	Detroit 1	U.S.A. (Mich.)	C	1	4
	Detroit 3	U.S.A. (Mich.)	C	1	4
56	BL-555 G4-E2-C1	U.S.A. (Ind.)	E	4	
57	BL-522 G2-E3-C3A	U.S.A. (Ind.)	E	4	
58	San Francisco 3	U.S.A. (Calif.)	E	4	
	San Francisco 6	U.S.A. (Calif.)	E	4	
	San Francisco 8	U.S.A. (Calif.)	E	4	
59	Lansing 3	U.S.A. (Mich.)	C	NT	
Species 2					
60	U7W (ATCC 33736)	U.S.A. (Pa.)	E	5	
61	MICU-B (ATCC 33735)	U.S.A. (Pa.)	E	5	
62	U8W (ATCC 33737)	U.S.A. (Pa.)	E	5	

[a] For further information, see Table 2.
[b] C, Clinical; E, environmental.
[c] For further information, see Table 7. Only isolates of serogroup 1 are typeable.
[d] Not typeable.
[e] Not to be confused with Dallas 1E (ET-53).

suspended in 2 ml of buffer (10 mM Tris, 1 mM EDTA, 0.5 mM NADP [pH 6.8]), and sonicated (Branson Sonifier Cell Disrupter, model 200, with microtip) for 1 min, at 50% pulse, with ice bath cooling. After centrifugation at 20,000 × *g* for 20 min at 4°C, the clear supernatant fluid (lysate) was stored at −70°C.

Electrophoresis and specific enzyme staining. Techniques of horizontal starch gel electrophoresis and the demonstration of specific enzyme activity were similar to those described by Selander et al. (50) and Caugant et al. (8). The following 22 enzymes were assayed: isocitrate dehydrogenase, malate dehydrogenase, leucine dehydrogenase, hydroxybutyrate dehydrogenase, adenylate kinase, indophenol oxidase, aconitase, peptidase-1, peptidase-2, leucine aminopeptidase, esterase-2, glucose 6-phosphate dehydrogenase, unidentified dehydrogenase, esterase-3, esterase-6, threonine dehydrogenase, lysine dehydrogenase, aspartic acid dehydrogenase, esterase-5, alanine dehydrogenase, glutamate dehydrogenase, and glutamic oxaloacetic transaminase.

Leucine dehydrogenase, hydroxybutyrate dehydrogenase, indophenol oxidase, glutamic oxaloacetic transaminase, threonine dehydrogenase, lysine dehydrogenase, and aspartic acid dehydrogenase were electrophoresed in a Poulik buffer system (gel buffer, pH 8.7; borate tray buffer, pH 8.2), and for all other enzymes, except esterases, we used a Tris citrate buffer system (pH 8.0).

We detected six esterases by using the α and β forms of naphthyl acetate and propionate as substrates, but only four esterases could be consistently scored for all isolates. Esterase-2 was active only with the α substrates, whereas esterase-5 catalyzed only α-naphthyl propionate; both enzymes were electrophoresed in a Tris-EDTA-borate buffer system (pH 8.0). Esterase-3, which was active with both β substrates, and esterase-6, which was specific for β-naphthyl acetate, were electrophoresed in a Tris hydrochloride buffer system (gel buffer, pH 8.0; borate tray buffer, pH 8.2) or a lithium hydroxide buffer system (gel buffer, pH 8.3; tray buffer, pH 8.1).

The enzyme designated as unidentified dehydrogenase was prominent on gels stained for phosphoglucose isomerase and adenylate kinase and showed weaker activity on those stained for alanine aminotransferase. We have not determined the physiological substrate(s) of this enzyme.

Leucine dehydrogenase showed strong activity with valine and weak activity with isoleucine and methionine as substrates. Alanine dehydrogenase showed weak activity with asparagine and methionine. All amino acids used as substrates were the L forms.

For each enzyme, distinctive mobility variants were designated as electromorphs and numbered in order of decreasing rate of anodal migration. Electromorphs of an enzyme were equated with alleles at the corresponding structural gene locus, and an absence of enzyme activity was attributed to a null allele, designated as 0. Because virtually all isolates showed activity for all 22 enzymes assayed, we presume that the corresponding structural gene loci are located on the chromosome rather than on plasmids.

Each isolate was characterized by its combination of alleles at the 22 enzyme loci. Distinctive profiles of electromorphs (Table 2), corresponding to unique multilocus genotypes, were designated as electrophoretic types (ETs; 8).

Serotyping and monoclonal antibody typing. Serotyping and monoclonal antibody typing were performed at the Centers for Disease Control, the latter with a panel of nine antisera prepared by McKinney et al. (34).

RESULTS

In the collection of 170 isolates received as *L. pneumophila*, 19 of the 22 enzyme loci assayed were polymorphic for alleles encoding electrophoretically detectable variants, and 3 loci (those for alanine dehydrogenase, glutamate dehydrogenase, and glutamic oxaloacetic transaminase) were monomorphic (Table 3). The average number of alleles per locus was 4.3.

By comparing the electrophoretic enzyme profiles of isolates typed for all 22 loci, we identified 62 distinctive allele combinations or ETs. A list of isolates with their ET designations and information on geographic locality, source, serotype, and (for most of those of serogroup 1) MAP is presented in Table 1. For each ET represented by two or more isolates, the first-listed isolate was designated as a reference strain.

A comparison of the electromorph profiles of the 62 ETs (Table 2) revealed that the 27 isolates representing ETs 51 through 62 are genetically very different from all other isolates received as *L. pneumophila*. To analyze this variation, we compared each ET with itself and with every other one in turn and, for each comparison, tabulated the number of loci at which dissimilar alleles occurred in the pair of ETs. The relative frequency distribution of mismatches for the total of 3,844 (= 62 × 62) comparisons is shown in Fig. 1. Pairs of ETs differed, on average, at 9.08 loci. The inflation of the observed variance (s^2 = 17.49) in the number of mismatches over that expected (σ_E^2 = 4.02) if the alleles occurred randomly and independently in genotypes indicated strong statistical associations between alleles in multilocus combinations (60). Moreover, the bimodal character of the distribution of mismatches suggested that the collection of 62 genotypes, representing the 170 isolates, is composed of at least two distinct groups.

To further analyze genetic relatedness among ETs, we calculated a matrix of pairwise weighted distance coefficients. The coefficient employed is the proportion of loci at which paired ETs have dissimilar alleles, with the contribution of each mismatch to the coefficient being inversely weighted by the genetic diversity (see below) at the locus in question. For a mismatch at the ith locus (with $j = 1 \ldots m$ alleles), the weight is given by the following equation: $w_i = \Sigma_j x^2_{ij}/\Sigma_{ij} x^2_{ij}$, where x_{ij} is the frequency of the jth allele at the ith locus over all 62 ETs. In this manner, greater weight was given to differences at less variable loci than to those at highly polymorphic loci.

From the matrix of coefficients of weighted genetic distance, we performed a principal-coordinates analysis (52). A plot of the 62 ETs along the first two coordinates (Fig. 2) placed the multilocus combinations in three distinct groups. A dendrogram produced by the average-linkage method of clustering from the same matrix showed the same three groups of ETs (Fig. 3). The isolates represented by ETs 1 through 50 are considered to be *L. pneumophila*, and each of the two other groups is regarded as genetically distinctive enough to warrant specific status. For purposes of reference, we have designated the isolates assigned to ETs 51 through 59 as species 1 and the isolates assigned to ETs 60 through 62 as species 2.

***L. pneumophila*. (i) Single-locus variation.** In *L. pneumophila*, 17 enzymes were polymorphic and 5 were monomorphic. The average number of alleles per locus was 3.2 (Table 3). For the 143 isolates that were typed for all 22 enzymes, 50 distinctive ETs were identified (Table 2). Mean genetic diversity ($\bar{h}$) over all loci in the 50 ETs, calculated by the unbiased method of Nei (39), was 0.312, with an interlocus variance (s^2) of 0.076 (Table 4). For a sample of n ETs, $h_i = n(1 - \Sigma_j x_{ij}^2)/(n - 1)$. In this equation, $1 - \Sigma_j x_{ij}^2$ is an estimate of the probability that two ETs, chosen at random, will have dissimilar alleles (i.e., a mismatch) at the ith locus, and $n/(n - 1)$ is a correction for bias in small samples. Mean genetic diversity per locus ($\bar{h} = \Sigma h_i/L$, where L is the number of loci) is the probability of a mismatch between ETs at the average locus.

Table 5 presents estimates of mean genetic diversity among ETs assigned to the following three source groups: (i) ETs represented only by isolates recovered from clinical cases; (ii) ETs represented by isolates obtained only from environmental sources; and (iii) ETs represented by isolates recovered from both clinical and environmental sources. (Except as noted, environmental sources of isolates were man-made aquatic habitats.) Mean diversity was high in each group, and for all but four of the loci, diversity within the three source groups was closely similar to that for the total sample of ETs. Diversity among ETs isolated from clinical sources alone was virtually the same as that for the species as a whole, and there is no evidence that clinical isolates are, as a group, significantly less variable than those recovered from environmental sources. The only locus with a large proportion of diversity among the three source groups was that for adenylate kinase, with a coefficient of genetic differentiation among groups, G_{ST}, of 0.287, which may be compared with a mean value for $\bar{G}_{ST}$ of 0.028 for the 22 loci. (G_{ST} estimates the ratio of the probability of a mismatch at a locus for two ETs chosen randomly from different groups to the average probability of a mismatch in the pooled sample; it is, therefore, a measure of the relative magnitude of genetic differentiation among groups [38].) The coefficient for adenylate kinase is large because of the exclusive occurrence of the 2 and 3 alleles of adenylate kinase and alleles in the seven environmental isolates representing ETs 44 through 49 (Table 2).

(ii) Relationships among multilocus genotypes. The dendrogram in Fig. 3 summarizes estimates of the genetic relationships of the various ETs. At a genetic distance of about 0.25, there are four groups or lineages of *L. pneumophila*, each represented by a single ET or a cluster of ETs. Most ETs belong to a large cluster bounded by ET-9 and ET-39, which in turn is divided into several subgroups.

Because a variety of dendrograms can be generated with different measures of genetic distance or methods of clustering (52), the dendrogram in Fig. 3 should be regarded as a hypothesis rather than a definitive scheme of genetic rela-

TABLE 2. Electromorph profiles of 62 ETs of *Legionella* spp.

Species and ET	Reference isolate	N[a]	Alleles at locus for enzyme[b]																				
			IDH	MDH	LDH	HBD	ADK	IPO	ACO	PE1	PE2	LAP	ES2	G6P	UDH	ES3	ES6	TDH	LYD	ASD	ES5	ALD	GLD	GOT
L. pneumophila																								
1	Davenport 1	3	3	3	2	1	1	3	2	6	5	5	4	4	0	2	2	1	2	1	6	1	1	1
2	B1793-79	1	3	3	2	1	1	3	2	6	5	3	9	4	0	2	2	1	2	1	4	1	1	1
3	Binghamton 1	1	3	3	2	1	1	3	2	6	5	2	9	4	0	2	2	1	2	1	4	1	1	1
4	D5401	1	3	3	2	1	1	3	3	6	5	2	6	4	0	2	2	1	2	1	5	1	1	1
5	Vasteras 2	1	3	3	2	1	1	3	2	4	5	5	9	2	0	2	0	1	2	1	8	1	1	1
6	Gloucester 1	1	3	1	2	1	1	3	3	6	3	3	4	4	0	2	2	1	2	1	6	1	1	1
7	B360-80	1	2	4	2	1	1	3	2	6	5	5	4	4	0	2	2	1	2	2	6	1	1	1
8	Portland 2	1	3	3	2	1	1	3	3	1	5	5	6	3	0	1	2	1	3	1	5	1	1	1
9	Los Angeles 2	7	3	4	2	1	1	3	3	3	5	2	8	2	0	1	1	1	2	2	2	1	1	1
10	HHE 2	1	3	4	2	1	1	3	3	3	5	3	5	2	0	1	1	1	2	2	2	1	1	1
11	Allentown 1	1	2	4	2	1	1	3	3	3	5	3	8	2	0	1	1	1	2	2	2	1	1	1
12	Bloomington 3	1	1	4	2	1	1	3	3	3	5	3	6	2	0	1	1	1	2	2	8	1	1	1
13	Nagasaki 1	1	1	4	2	1	1	3	3	3	5	3	7	2	0	1	1	1	2	2	4	1	1	1
14	Togus 2	2	3	4	2	1	1	3	3	1	5	3	9	2	0	1	2	1	2	2	3	1	1	1
15	Lancaster 1	1	3	4	2	1	1	3	3	3	5	5	9	2	0	1	2	1	2	2	3	1	1	1
16	Toronto 1	1	3	4	2	1	1	3	3	3	5	5	8	2	0	1	2	1	2	2	4	1	1	1
17	Pontiac 1	11	3	4	2	1	1	3	3	1	5	5	8	2	0	1	2	1	2	2	4	1	1	1
18	Chamblee 1	2	3	4	2	1	1	3	3	1	5	5	4	2	0	1	2	1	2	2	4	1	1	1
19	Naples 1	1	3	4	2	1	1	3	3	1	5	2	8	2	0	1	2	1	2	2	4	1	1	1
20	Cambridge 4	1	3	4	2	1	1	3	3	1	5	3	6	3	0	1	2	1	2	2	5	1	1	1
21	OLDA	26	3	4	2	1	1	3	3	1	5	5	6	3	0	1	2	1	2	2	5	1	1	1
22	Anniston 1	1	3	4	2	1	1	3	3	1	5	5	6	3	0	1	2	1	2	2	0	1	1	1
23	Nagasaki 2	4	3	4	2	1	1	3	3	1	5	5	6	4	0	1	2	1	2	2	5	1	1	1
24	San Francisco 1	4	3	4	2	1	1	3	3	1	5	5	6	3	0	1	2	1	3	2	5	1	1	1
25	Philadelphia 1	21	3	4	2	1	1	3	2	3	5	1	3	2	0	1	2	1	2	2	7	1	1	1
26	Buffalo 1	1	3	4	2	1	1	3	3	3	5	1	3	2	0	1	2	1	2	2	7	1	1	1
27	Burlington 1	1	3	4	2	1	1	3	2	3	5	1	3	1	0	1	2	1	2	2	7	1	1	1
28	WBH 1	1	3	4	2	1	1	3	2	6	5	1	3	1	0	1	2	1	2	2	7	1	1	1
29	Adelaide 1	5	3	4	1	1	1	3	3	1	5	5	8	2	0	1	2	1	2	2	4	1	1	1
30	Concord 3	1	3	4	1	1	1	3	3	3	5	2	9	2	0	1	2	1	2	2	1	1	1	1
31	Leiden 1E	5	3	4	2	1	1	2	3	1	5	3	9	2	0	1	2	1	2	2	3	1	1	1
32	Leiden 2	2	3	4	2	1	1	2	2	1	5	3	9	2	0	1	2	1	2	2	3	1	1	1
33	Stockholm 4	2	3	4	2	1	1	2	2	1	5	2	9	2	0	1	2	1	2	2	3	1	1	1

34	Winnipeg 1	1	1	4	2	1	1	2	3	1	5	3	9	2	0	1	2	1	2	2	3	1	1	1
35	Ottawa 6	1	3	4	2	1	1	2	3	3	5	2	9	2	0	0	2	1	2	2	3	1	1	1
36	Winnipeg 2	1	3	4	2	1	1	2	3	1	5	2	9	2	0	1	2	0	2	2	3	1	1	1
37	Ottawa 2	1	3	4	2	1	1	2	2	3	2	2	9	2	0	1	2	1	2	2	3	1	1	1
38	Berlin 1	1	3	3	2	1	1	2	2	1	5	5	9	2	0	1	3	1	2	2	0	1	1	1
39	Lyon 3	1	3	4	2	1	1	3	2	1	5	2	9	2	0	1	3	1	2	2	0	1	1	1
40	Togus 1	6	3	4	2	1	1	3	2	3	3	2	3	2	0	2	2	1	2	2	7	1	1	1
41	Atlanta 1	3	3	4	2	1	1	3	2	3	3	3	3	2	0	2	2	1	2	2	7	1	1	1
42	SRP 20	5	3	4	2	1	1	3	3	1	3	2	3	2	0	2	0	1	2	2	7	1	1	1
43	Bloomington 2	1	2	4	2	1	1	1	3	1	5	2	9	2	0	1	2	1	2	1	3	1	1	1
44	684	2	3	4	1	1	3	3	2	3	5	2	8	2	0	1	2	1	2	2	1	1	1	1
45	Pratt 2	1	3	4	2	1	3	3	2	3	5	2	8	2	0	1	2	1	2	2	1	1	1	1
46	Oxford 4	1	3	4	1	1	3	3	2	3	5	2	8	2	0	1	1	1	2	2	1	1	1	1
47	Mt. St. Helens 6	1	3	4	1	1	3	3	2	3	5	1	8	2	0	1	1	1	2	2	1	1	1	1
48	Mt. St. Helens 2	1	3	4	1	1	3	3	1	3	5	2	9	2	0	1	1	1	2	2	1	1	1	1
49	Chicago 8	1	3	3	2	1	2	1	2	2	5	3	8	2	0	1	3	1	2	2	0	1	1	1
50	Mt. St. Helens 5	1	1	6	2	1	1	3	3	6	1	2	2	2	0	2	1	1	2	2	5	1	1	1
Species 1																								
51	Oak Ridge Y	1	2	2	3	1	3	3	5	7	5	2	0	2	2	2	2	2	1	2	8	1	1	1
52	Indianapolis 1	1	2	5	2	2	5	3	5	1	5	3	1	2	2	3	2	1	1	3	8	1	1	1
53	Detroit 2	3	2	5	2	2	3	3	5	5	5	2	1	1	2	3	2	2	1	3	8	1	1	1
54	Los Angeles 1	11	2	5	2	2	3	3	5	5	5	2	1	3	2	3	2	2	1	3	8	1	1	1
55	Detroit 1	2	2	5	2	2	3	3	5	1	5	2	1	2	2	3	2	2	1	3	8	1	1	1
56	BL-555	1	2	5	2	2	3	3	5	5	5	2	1	2	2	3	2	2	1	3	8	1	1	1
57	BL-522	1	2	5	2	2	3	3	5	1	5	2	1	2	2	1	2	2	1	3	8	1	1	1
58	San Francisco 3	3	2	5	2	2	3	3	5	3	5	2	1	2	2	1	2	2	1	3	8	1	1	1
59	Lansing 3	1	2	5	2	2	3	4	5	5	5	2	1	1	2	0	2	2	1	3	8	1	1	1
Species 2																								
60	U7W	1	2	4	2	3	5	3	4	5	4	4	5	1	1	4	1	1	1	2	8	1	1	1
61	MICU-B	1	2	4	2	3	5	3	4	5	4	4	4	1	1	4	1	1	1	2	8	1	1	1
62	U8W	1	2	4	2	3	4	3	4	5	4	4	4	1	1	4	1	1	1	2	8	1	1	1

[a] *N*, Number of isolates.

[b] Abbreviations: IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; LDH, leucine dehydrogenase; HBD, hydroxybutyrate dehydrogenase; ADK, adenylate kinase; IPO, indophenol oxidase; ACO, aconitase; PE1, peptidase-1; PE2, peptidase-2; LAP, leucine aminopeptidase; ES2, esterase-2; G6P, glucose 6-phosphate dehydrogenase; UDH, unidentified dehydrogenase; ES3, esterase-3; ES6, esterase-6; TDH, threonine dehydrogenase; LYD, lysine dehydrogenase; ASD, aspartic acid dehydrogenase; ES5, esterase-5; ALD, alanine dehydrogenase; GLD, glutamate dehydrogenase; GOT, glutamic oxaloacetic transaminase.

TABLE 3. Frequencies of alleles at 22 enzyme loci in ETs of three *Legionella* species

Locus and allele[a]	Frequency of alleles[b] in:		
	L. pneumophila (n = 50)	Species 1 (n = 9)	Species 2 (n = 3)
IDH			
3	**0.86**	0.00	0.00
1	**0.08**	0.00	0.00
2	0.06	1.00	1.00
MDH			
4	0.80	0.00	1.00
3	**0.16**	0.00	0.00
1	**0.02**	0.00	0.00
6	**0.02**	0.00	0.00
2	0.00	**0.11**	0.00
5	0.00	**0.89**	0.00
LDH			
2	0.88	0.89	1.00
1	**0.12**	0.00	0.00
3	0.00	**0.11**	0.00
HBD			
1	1.00	0.11	0.00
2	0.00	**0.89**	0.00
3	0.00	0.00	**1.00**
ADK			
1	**0.88**	0.00	0.00
3	0.10	0.89	0.00
2	**0.02**	0.00	0.00
5	0.00	0.11	0.67
4	0.00	0.00	**0.33**
IPO			
3	0.80	0.89	1.00
2	**0.16**	0.00	0.00
1	**0.04**	0.00	0.00
4	0.00	**0.11**	0.00
ACO			
3	**0.58**	0.00	0.00
2	**0.40**	0.00	0.00
1	**0.02**	0.00	0.00
5	0.00	**1.00**	0.00
4	0.00	0.00	**1.00**
PE1			
3	0.40	0.11	0.00
1	0.40	0.33	0.00
6	**0.16**	0.00	0.00
2	**0.02**	0.00	0.00
4	**0.02**	0.00	0.00
7	0.00	**0.11**	0.00
5	0.00	0.44	1.00
PE2			
5	0.88	1.00	0.00
3	**0.08**	0.00	0.00
1	**0.02**	0.00	0.00
2	**0.02**	0.00	0.00
4	0.00	0.00	**1.00**
LAP			
2	0.36	0.89	0.00
5	**0.28**	0.00	0.00
3	0.26	0.11	0.00
1	**0.10**	0.00	0.00
4	0.00	0.00	**1.00**
ES2			
9	**0.34**	0.00	0.00
8	**0.22**	0.00	0.00
6	**0.16**	0.00	0.00
3	**0.14**	0.00	0.00
4	0.08	0.00	0.67
2	**0.02**	0.00	0.00
5	0.02	0.00	0.33
7	**0.02**	0.00	0.00
0	0.00	**0.11**	0.00
1	0.00	**0.89**	0.00
G6P			
2	0.72	0.67	0.00
4	**0.14**	0.00	0.00
3	0.10	0.11	0.00
1	0.04	0.22	1.00
UDH			
0	**1.00**	0.00	0.00
2	0.00	**1.00**	0.00
1	0.00	0.00	**1.00**
ES3			
1	0.76	0.22	0.00
2	0.22	0.11	0.00
0	0.02	0.11	0.00
3	0.00	**0.56**	0.00
4	0.00	0.00	**1.00**
ES6			
2	0.72	1.00	0.00
1	0.18	0.00	1.00
3	**0.06**	0.00	0.00
0	**0.04**	0.00	0.00
TDH			
1	0.98	0.11	1.00
0	**0.02**	0.00	0.00
2	0.00	**0.89**	0.00
LYD			
2	**0.96**	0.00	0.00
3	**0.04**	0.00	0.00
1	0.00	1.00	1.00
ASD			
2	0.84	0.11	1.00
1	**0.16**	0.00	0.00
3	0.00	**0.89**	0.00
ES5			
3	**0.20**	0.00	0.00
4	**0.16**	0.00	0.00
5	**0.14**	0.00	0.00
7	**0.14**	0.00	0.00
1	**0.12**	0.00	0.00
0	**0.08**	0.00	0.00
2	**0.06**	0.00	0.00
6	**0.06**	0.00	0.00
8	0.04	1.00	1.00
ALD			
1	1.00	1.00	1.00
GLD			
1	1.00	1.00	1.00
GOT			
1	1.00	1.00	1.00

[a] For abbreviations, see Table 2, footnote *b*.

[b] Alleles are arranged in order of decreasing frequency in *L. pneumophila*. Frequencies of alleles recorded in only one species are indicated by boldface numbers.

tionships. For example, clustering by the same average-linkage method used to construct the dendrogram in Fig. 3, but from a matrix of coefficients of unweighted genetic distance (proportion of mismatches, with each locus equally weighted), did not affect the species trifurcation, but it produced several changes within both *L. pneumophila* and species 1, as follows. (i) The cluster shown in Fig. 3 composed of ET-1 through ET-8 lost ET-8 (Portland 2), which joined cluster ET-14–ET-18. (ii) Cluster ET-19–ET-24 lost ET-19, which joined the large cluster ET-31–ET-39. ET-43 also joined this cluster, which was otherwise unchanged in composition. (iii) Cluster ET-44–ET-48 was

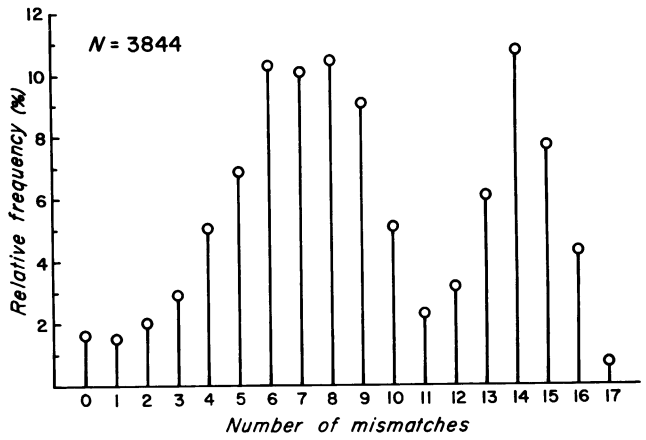

FIG. 1. Distribution of number of enzyme loci at which unlike alleles (mismatches) occurred in pairwise comparisons of 62 ETs of *Legionella* spp.

joined by ET-30. Clusters ET-9–ET-13, ET-25–ET-28, and ET-40–ET-42 and single lines ET-49 and ET-50 remained unchanged.

In addition to these relatively minor changes in composition of the clusters, there were several differences between the dendrograms in the pattern of branching in the genetic-distance range of 0.10 to 0.25. However, one feature that remained the same is the deep separation of the Davenport 1 family (ET-1–ET-7 or ET-8) and the ET-49 and ET-50 lines, which occurs in the range 0.25 to 0.35 (Fig. 3).

In sum, our analysis revealed the following four major divisions in *L. pneumophila*: (i) the Davenport 1 group (ET-1–ET-8), (ii) ET-49, (iii) ET-50, and (iv) a complex of clusters that includes most of the ETs. The large fourth group is composed of relatively closely related multilocus genotypes. Several subclusters are evident within this group,

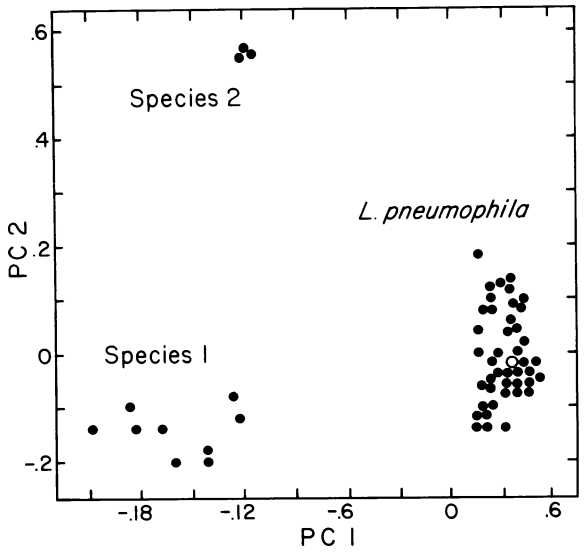

FIG. 2. Distribution of 62 ETs of *Legionella* spp. on the first two axes (PC1 and PC2) derived from a principal-coordinates analysis based on a matrix of coefficients of weighted genetic distance (22 enzyme loci; see reference 52, p. 248–249). The open circle represents seven ETs.

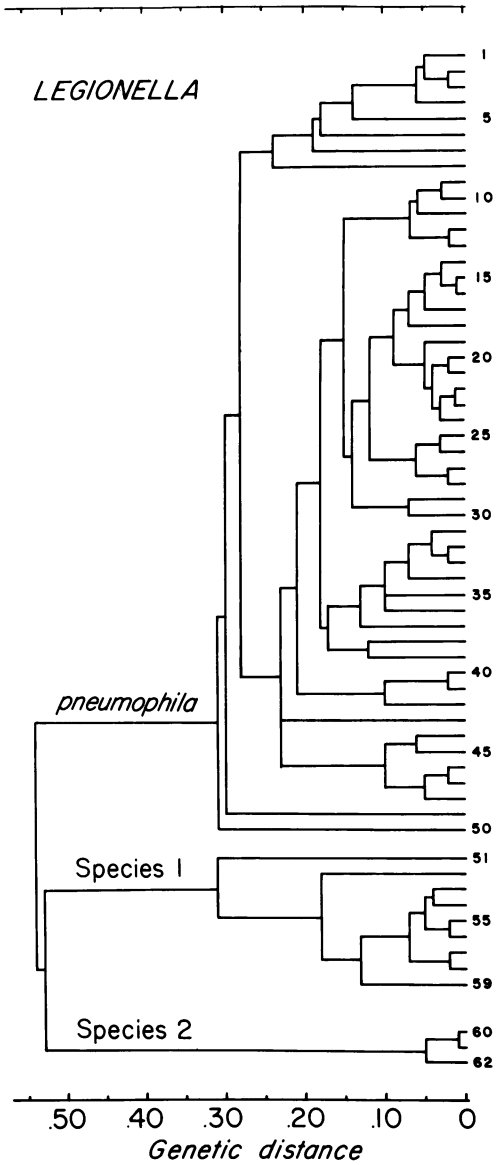

FIG. 3. Genetic relationships among 62 ETs of *L. pneumophila* (ETs 1 through 50), species 1 (ETs 51 through 59), and species 2 (ETs 60 through 62). The dendrogram was generated by the average-linkage method of clustering from a matrix of coefficients of weighted genetic distance, based on 22 enzyme loci. ETs are numbered sequentially from top to bottom in the order of their listing in Table 1. ET assignments of some of the better-known isolates are as follows: for *L. pneumophila*, Los Angeles 2 (ET-9), Bellingham 1 (ET-14), Pontiac 1 (ET-17), OLDA (ET-21), San Francisco 1 (ET-24), Philadelphia 1 and Burlington 4 (ET-25), Leiden 1E (ET-31), Togus 1 and Burlington 5 (ET-40), and Bloomington 2 (ET-43); for species 1, Dallas 1E (ET-53) and Los Angeles 1 (ET-54).

but as noted, their composition depends to some degree on the measure of genetic distance employed.

(iii) Geographic distribution of multilocus genotypes. Of the 50 ETs, 13 (26%) are represented by isolates from two or more geographic localities, and in seven cases (ETs 9, 17, 21, 23, 25, 29, and 31), the distribution is intercontinental (Table 1). For example, ET-21 (OLDA) has been recovered at 17 different localities in North America, Australia, South Africa, continental Europe, and England. The type isolate of

TABLE 4. Genetic diversity at 22 enzyme loci in ETs of three *Legionella* species

Enzyme[a]	Genetic diversity in[b]:			Diversity statistics[c]		
	L. pneumophila (n = 50; 0.312 [0.076])	Species 1 (n = 9; 0.182 [0.051])	Species 2 (n = 3; 0.061 [0.038])	Within species (0.281 [0.056])	Total (0.420 [0.064])	G_{ST} (0.330 [0.096])
IDH	0.256	0.000	0.000	0.206	0.464	0.556
MDH	0.341	0.222	0.000	0.307	0.493	0.378
LDH	0.216	0.222	0.000	0.206	0.207	0.003
HBD	0.000	0.222	0.000	0.032	0.309	0.896
ADK	0.220	0.222	0.667	0.242	0.457	0.471
IPO	0.340	0.222	0.000	0.306	0.310	0.014
ACO	0.513	0.000	0.000	0.414	0.664	0.377
PE1	0.667	0.750	0.000	0.647	0.729	0.113
PE2	0.223	0.000	0.000	0.180	0.267	0.326
LAP	0.729	0.222	0.000	0.620	0.725	0.145
ES2	0.799	0.222	0.667	0.709	0.850	0.166
G6P	0.460	0.556	0.000	0.451	0.515	0.123
UDH	0.000	0.000	0.000	0.000	0.332	1.000
ES3	0.381	0.694	0.000	0.408	0.545	0.251
ES6	0.453	0.000	0.000	0.365	0.439	0.169
TDH	0.040	0.222	0.000	0.065	0.256	0.748
LYD	0.078	0.000	0.000	0.063	0.368	0.828
ASD	0.274	0.222	0.000	0.253	0.423	0.401
ES5	0.883	0.000	0.000	0.712	0.877	0.188
ALD	0.000	0.000	0.000	0.000	0.000	0.000
GLD	0.000	0.000	0.000	0.000	0.000	0.000
GOT	0.000	0.000	0.000	0.000	0.000	0.000

[a] For abbreviations, see Table 2, footnote *b*.
[b] n, Number of ETs. Other numbers in parentheses are means, and numbers in brackets are variances.
[c] Within-species diversity is the mean of the diversities for the three species, weighted by sample size (n). Total diversity is calculated by the equation for n in the text from the weighted mean allele frequencies for the entire sample of 62 ETs. G_{ST} is the coefficient of genetic differentiation between species (see the text).

the species, Philadelphia 1, is a common genotype (ET-25) that has been isolated in eight states and in Germany and France. Isolates from the first-recorded outbreak of Pontiac fever, in Pontiac, Mich., in 1968 (Pontiac 1, SK 267, SK 273, and SK 304), belong to ET-17 and are indistinguishable by our technique from seven other isolates collected in California, Connecticut, New York, and Australia.

Undescribed species. (i) Species 1. Species 1, which was represented by 24 isolates (Table 1) of nine multilocus genotypes (ETs 51 through 59; Table 2), differs from *L. pneumophila* in allele frequencies at 19 loci (Table 3). All available isolates of species 1 can be distinguished from those of *L. pneumophila* by possession of distinctive alleles at seven loci (malate dehydrogenase, hydroxybutyrate dehydrogenase, aconitase, esterase-2, unidentified dehydrogenase, lysine dehydrogenase, and aspartic acid dehydrogenase), with the single exception of Oak Ridge Y (ET-51; see below).

Mean genetic diversity for the ETs of species 1 was 0.182 (s^2 = 0.051), which is only 58% of that recorded for *L. pneumophila* (Table 4).

TABLE 5. Mean genetic diversity in groups of ETs of *L. pneumophila* isolated from clinical, environmental, and both clinical and environmental sources

Source	No. of ETs[a]	Mean genetic diversity (variance)
Clinical	29 (39)	0.310 (0.089)
Environmental	13 (30)	0.327 (0.086)
Clinical and environmental	8 (73)	0.245 (0.087)
Total	50 (132)	0.313 (0.076)

[a] Number of isolates in parentheses.

ETs 53 through 58, which include isolates from Indiana, Michigan, Texas, and California, are closely similar in electromorph profile (Table 2; Fig. 3). ET-59, represented by Lansing 3, is somewhat divergent from this group and is further distinctive in being untypeable with the eight standard antisera currently available. ET-52, represented by Indianapolis 1, is even more divergent, but the most atypical isolate is Oak Ridge Y (ET-51), which splits off from other isolates of species 1 at a genetic distance level of 0.30. In some respects, the Oak Ridge Y isolate (ET-51) is intermediate between other representatives of species 1 and those of *L. pneumophila*. It has the distinctive alleles of species 1 at the loci for isocitrate dehydrogenase, aconitase, unknown dehydrogenase, and lysine dehydrogenase but shares alleles with *L. pneumophila* at the loci for aspartic acid dehydrogenase and hydroxybutyrate dehydrogenase. In addition, it

TABLE 6. Mean genetic diversity ($\bar{h}$) at 22 enzyme loci in ETs of *Legionella* spp. classified by serogroup

Serogroup	Sample size (n), mean, and variance in:					
	3 Species			*L. pneumophila*		
	n	Mean	Variance	n	Mean	Variance
1	36	0.359	0.066	32	0.286	0.076
2	5	0.309	0.094	5	0.309	0.094
3	2	0.364	0.242	2	0.364	0.242
4	10	0.456	0.070	6	0.294	0.109
5	7	0.500	0.076	3	0.242	0.108
4-5	7	0.426	0.091	6	0.355	0.106
Diversity[a]						
Within		0.380	0.058		0.286	0.070
G_{ST}		0.073	0.003		0.055	0.008

[a] Diversity statistics include single isolates of serogroups 7 and 8.

TABLE 7. Mean genetic diversity ($\bar{h}$) in relation to MAP

MAP no.	Reactivity pattern[a] to antibody no.									No. of ETs	Mean genetic diversity[b] (variance)
	1	2	3	4	5	6	7	8	9		
1				+	+				+	1	0.000 (0.000)
2			+	+	+	+		+		1	0.000 (0.000)
3				+	+	+		+	+	1	0.000 (0.000)
4	+			+	+	+	+	+	+	10	0.412 (0.058)
5	+	+		+		+				0	
6	+			+		+				7	0.240 (0.071)
7	+	+		+	+				+	1	0.000 (0.000)
8	+	+		+					+	2	0.045 (0.045)
9	+	+		+	+					10	0.292 (0.085)
10	+	+		+						3	0.061 (0.038)
11	+	+			+					1	0.000 (0.000)
12	+	+								11	0.269 (0.092)
13	+			+						1	0.000 (0.000)
14	+	+	+							4	0.295 (0.122)
15	+			+	+	+			+	1	0.000 (0.000)

[a] Monoclonal antibodies prepared by McKinney et al. (34). +, Reactivity.
[b] For genetic diversity within MAPs, the mean was 0.243 and the variance was 0.037. The total mean was 0.294, and total variance was 0.064.

has unique alleles at the loci for malate dehydrogenase, leucine dehydrogenase, peptidase-1, and esterase-2.

(ii) Species 2. Species 2 is known only from three environmental isolates, MICU-B, U7W, and U8W (ATCC 33735 through 33737), collected at the Pittsburgh Veterans Administration Medical Center and extensively studied by Garrity et al. (16). Each isolate has a distinctive electromorph profile (Table 2), but all three ETs are very closely related (Fig. 2 and 3); mean genetic diversity was only 0.061 (s^2 = 0.038). Isolates of species 2 are distinct from *L. pneumophila* and from species 1 at nine loci, and they have unique alleles at seven loci.

Genetic variation in relation to serotype. Our analysis revealed several cases of variation in serotype among isolates of the same ET. As is shown for *L. pneumophila* in Table 1, ET-25 includes isolates of serogroups 1, 3, and 4; ET-31 is represented by isolates of serogroups 1, 4-5, and 5; ET-32 is represented by isolates of serogroups 1 and 4; and ET-40 is represented by isolates of serogroups 2 and 4. In species 1 (Table 1), ET-53 is represented by isolates of serogroups 1 and 5, and ET-54 is represented by isolates of serogroups 1 and 4. However, the serotype may be uniform or nearly so in some ETs, as, for example, in ET-21, which is represented by 26 isolates of *L. pneumophila* from 17 localities, all but 1 of which are serogroup 1; the exception (Oxford 3) was untypeable (Table 1).

Estimates of the extent of genetic variation among ETs of the same serogroup are presented in Table 6, which is based on serotypes of isolates belonging to 61 ETs (Table 1). For the 6 ETs represented by isolates of two or three different serogroups, each ET was assigned to two or three serogroups, thus yielding a total sample size of 69.

TABLE 8. Mean genetic diversity within groups of ETs represented by isolates reacting with each of nine monoclonal antibodies

Monoclonal antibody no.	No. of ETs	Mean genetic diversity[a] (variance)
1	51	0.295 (0.064)
2	32	0.252 (0.068)
3	5	0.318 (0.139)
4	38	0.297 (0.056)
5	26	0.326 (0.063)
6	20	0.344 (0.053)
7	11	0.388 (0.054)
8	13	0.364 (0.051)
9	16	0.315 (0.045)

[a] For genetic diversity within antibody categories, the mean was 0.309, and variance was 0.054. For genetic diversity of the total sample, the mean was 0.311, and variance was 0.057.

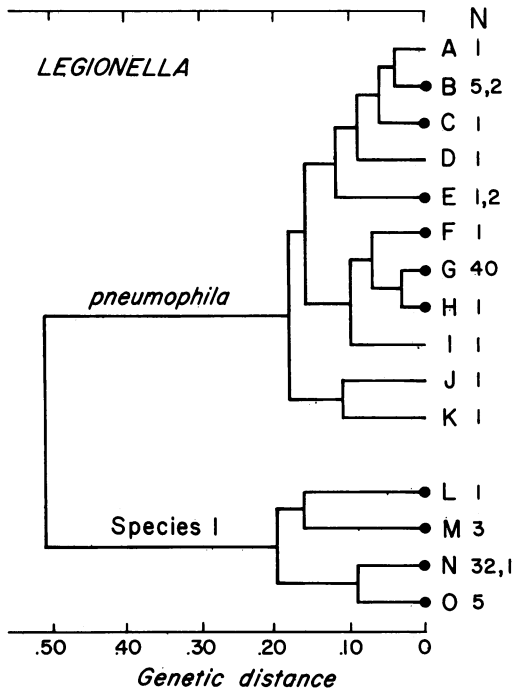

FIG. 4. Genetic relationships among 15 ETs of *L. pneumophila* and species 1 represented by isolates from hospitals in the Los Angeles region, Calif. Solid circles indicate ETs recovered at the Wadsworth Medical Center. N, Number of isolates. For ETs B, E, and N, the two numbers indicate isolates from the Wadsworth Medical Center and other hospitals, respectively.

Mean genetic diversity among ETs of the same serogroup was, on average, 0.380, which is 93% ($\overline{G}_{ST}$ = 0.073) of that in the total sample of 61 ETs. In comparison with the genetic differences between species, for which $\overline{G}_{ST}$ = 0.330 (Table 4), the extent of differentiation among ETs of different serogroups is small.

Mean genetic diversities for serogroups 4, 5, and 4-5 are larger than average (Table 6) because ETs assigned to these serogroups represent in about equal proportion *L. pneumophila* and species 1 or, in the case of serogroup 5, all three species. Hence, the genetic differences between the species account for most of the increased diversity among ETs of these serogroups. For serogroup 1, which is also represented by ETs of two species (*L. pneumophila* and species 1), the diversity is not severely inflated because most (87%) of the ETs assigned to serogroup 1 are *L. pneumophila*. However, in this case and even among ETs of *L. pneumophila* assigned to a single serogroup (1, 2, or 3), there is considerable genetic diversity (Table 6).

The average genetic diversity within serogroups of *L. pneumophila* alone (Table 6) is 0.286, which is 95% of the diversity within the species as a whole. Isolates of the same serotype clearly are not homogeneous with regard to multilocus genotype.

Genetic variation in relation to MAP. McKinney et al. (34) prepared a panel of nine monoclonal antibodies to cell surface antigens of Knoxville 1 (antibodies 1, 2, and 3) and OLDA (antibodies 4 through 9), both of serogroup 1. These antibodies react only with isolates of serogroup 1. Testing 130 serogroup 1 isolates of *L. pneumophila*, McKinney et al. (34) identified 13 distinctive MAPs, and several additional patterns have since been recorded. For purposes of analysis, we have numerically designated the MAPs represented by isolates analyzed in this paper according to the scheme shown in Table 7.

There is considerable genetic diversity among ETs assigned to each of several MAPs for which sample size is greater than five. On average, mean genetic diversity per locus within MAPs is about 75% of that recorded in the total sample of ETs. As is shown by this analysis and the distribution of various MAPs among the ETs (Table 1), a given MAP may occur in isolates having a large variety of multilocus genotypes.

Table 8 shows estimates of mean genetic diversity per locus within groups of ETs assigned to nine categories on the basis of the reactions of their isolates to each of the antibodies. Thus, for example, group 1 includes the 51 ETs represented by 1 or more isolates that gave a positive reaction in tests with antibody 1. The total number of assignments was 212 ETs. As shown, the mean diversity among ETs within single antibody categories ($\overline{h}$ = 0.309) is not significantly less than that of the total sample of ETs ($\overline{h}$ = 0.311).

Genetic diversity within a local region. Phenotypic diversity has often been detected among isolates of *L. pneumophila* collected in individual hospitals, especially where there has been an extended history of nosocomial infection. One such institution is the Wadsworth Medical Center, in Los Angeles, Calif. (19), from which we have obtained 90 clinical and environmental isolates of *Legionella* spp. collected in the period 1978 to 1984. For the 22 enzyme loci examined, electrophoretic profiles indicate that 49 of the isolates are *L. pneumophila* and 41 are species 1 (Fig. 4). The isolates were serotyped as either serogroup 1 or 4, but serotyping does not consistently distinguish either the species or the ETs within species because of the following distribution of serogroups: *L. pneumophila* isolates were 94% serogroup 1 and 6% serogroup 4, and species 1 isolates were 32% serogroup 1 and 68% serogroup 4 (P. H. Edelstein, in preparation).

The six ETs of *L. pneumophila* from Wadsworth Hospital fall into two major clusters, and the four ETs of species 1 also belong to two clusters (Fig. 4). For each of the two species, most isolates are of a single ET. The electromorph profiles of most of the Wadsworth isolates have not yet been compared with the profiles identified in our larger study, but we have determined that ET-N (reference isolate, F1700) is ET-54 (reference isolate, Los Angeles 1; Tables 1 and 2).

Mean genetic diversity for the ETs of *L. pneumophila* from Wadsworth Hospital was 0.173 (s^2 = 0.074), or roughly half that recorded for the species in our larger survey (Table 4). For the ETs of species 1, $\overline{h}$ = 0.177 (s^2 = 0.079), which is only a little less than the value ($\overline{h}$ = 0.182) recorded for the nine ETs of species 1 listed in Table 2. Total mean diversity over the 10 ETs of *Legionella* spp. from Wadsworth Hospital was $\overline{h}$ = 0.398 (s^2 = 0.077).

Nine isolates of *L. pneumophila* and one isolate of species 1 from several other hospitals in the Los Angeles area were also examined, five of which are of ETs not recovered from Wadsworth Hospital. For the total of 11 ETs of *L. pneumophila* from the Los Angeles region, $\overline{h}$ = 0.210 (s^2 = 0.070), and for the total of 15 ETs of *L. pneumophila* and species 1, $\overline{h}$ = 0.365 (s^2 = 0.063).

Species 1 is not to be confused with *L. wadsworthii* (ATCC 33877), a very different form described from isolates recovered from a pneumonia patient at the Wadsworth Medical Center (10).

Pontiac fever isolates. We have analyzed variation in two series of environmental isolates of *L. pneumophila* that have been epidemiologically linked with outbreaks of the nonpneumonic, self-limiting form of legionellosis. The first of these two series includes Pontiac 1, which was isolated in 1977 from a frozen sample of water collected in 1968 in the building in Pontiac, Mich., where the first-recorded outbreak of the disease occurred (17), and SK 267, SK 273, and SK 304, which were recovered from frozen lung tissue of sentinel guinea pigs (27). All four isolates are serogroup 1 MAP 8 representatives of ET-17, an apparently common genotype that also includes isolates recovered from clinical cases of the pneumonic disease (Legionnaires disease) at six localities in California, New York, and Connecticut and in Adelaide, Australia (Table 1). All of these isolates are serogroup 1 but were variously typed as MAP 7, 8, 9, 10, or 12.

The second series consists of 22 isolates collected from water by epidemiologists from the Centers for Disease Control investigating an outbreak of Pontiac fever in a bank in New York City in 1984. All of these isolates are ET-21 (OLDA) and were typed as serogroup 1 MAP 9. ET-21 is a widespread genotype represented by both clinical and environmental isolates from four continents and England (Table 1). All but one of the ET-21 isolates are serogroup 1 (the exception was untypeable), but they have been variously typed as MAP 3, 4, 6, 9, or 12.

ET-17 and ET-21 are related, differing, according to our analysis, only at the highly polymorphic esterase-2, esterase-5, and glucose 6-phosphate dehydrogenase loci (Table 2).

DISCUSSION

Estimates of genetic relatedness among isolates. Because the multilocus electrophoretic technique attempts to determine overall genetic relatedness among isolates from examination of a small fraction of the structural gene loci of the

genome, it is significant that estimates obtained for isolates of *Escherichia coli* and *Shigella* spp. have been shown to correlate (r = about 0.70) with comparable estimates for the same isolates from DNA hybridization experiments (42). More extensive data are now available for *Legionella* spp., and the correlation is even stronger (R. K. Selander, unpublished data). Values of genetic similarity (1 − weighted genetic distance) and the relative binding rate of DNA at 75°C with Philadelphia 1 as a reference are available for 33 isolates of *L. pneumophila*, species 1, and species 2, together with comparable values for the same isolates hybridized with Los Angeles 1 (species 1). For the combined data (n = 66), Spearman's rank correlation coefficient, r_S, is 0.827 ($P < 0.001$). Although the overall relationship is highly significant, relatively little contribution is made by correlation within species.

DNA hybridization has been widely applied in bacterial taxonomy to assess overall genetic relatedness at the species level, but it is of limited value in measuring relationships among cell lines within species, owing to the large experimental error involved, which is on the order of 10% (3). Even replicate experiments in the same laboratory may yield different estimates of nucleotide sequence similarity, and values reported from different laboratories may be widely discordant. For example, reported estimates of the relative binding rate at 75°C for Flint 1 (ET-21) to Philadelphia 1 (ET-25) range from 71 to 82% (3–5, 33), and estimates of the relative binding rate for Pontiac 1 (ET-17) with Philadelphia 1 as reference are variously reported as 74 (5), 79 (33), and 90% (21).

Relationships among species. When clustered from coefficients of genetic distance, *L. pneumophila*, species 1, and species 2 are equidistant from one another, the tripartite divergence occurring at a level of approximately 0.54 (Fig. 3). Our estimates of genetic relatedness are consistent with those obtained in DNA hybridization experiments by D. J. Brenner (unpublished data). Each species shows a relative binding rate of about 50% at 75°C when it is hybridized to the other two. Species 1 and 2 are being taxonomically described by D. J. Brenner (unpublished data).

Inasmuch as Dallas 1E has been widely used in research as a reference strain for serogroup 5 (11), a discrepancy regarding its identity must be accounted for. Garrity et al. (16) reported relative binding rates for Dallas 1E with Philadelphia 1 of 94 and 93% at 64 and 75°C, respectively, and the reciprocal experiment yielded values of 94 and 95% at 64 and 75°C, respectively. With *L. pneumophila* isolate 684 (ATCC 33733; ET-44) as reference, Dallas 1E showed relative binding rates of 77 and 69% at 64 and 75°C, respectively. In initial tests, D. J. Brenner (unpublished data) also obtained high relatedness values (85%) between Philadelphia 1 and Dallas 1E. However, using a culture derived from the original stock culture, he obtained values for Dallas 1E with Philadelphia 1 that are much lower, i.e., 64 and 47% at 60 and 75°C, respectively. Dallas 2E, which is genotypically identical to Dallas 1E (ET-53), showed relative binding rates of 69 and 47% at 60 and 75°C, respectively, with Philadelphia 1 as reference. For Detroit 2, which also is ET-53, the comparable values were 65 and 45% at 60 and 75°C, respectively. (Data for isolate 684 are not available from Brenner's laboratory.)

Our results clearly agree with those of D. J. Brenner (unpublished data) in identifying Dallas 1E as a typical representative of species 1. Dallas 1E is closely related to ET-54 (Los Angeles 1 and 10 other isolates from Los Angeles), ET-55 (Detroit 1 and Detroit 3), and ET-56 (BL-555) (Fig. 3). Hybridized with three other members of species 1 as references, i.e., Los Angeles 1 (ET-54), Lansing 3 (ET-59), and Indianapolis 1 (ET-52), Dallas 1E gave mean values of 82.7% (range, 77 to 94%) at 64°C and 82.3% (range, 68 to 93%) at 75°C.

We conclude that both the DNA hybridization experiments of Garrity et al. (16) and the initial experiments of D. J. Brenner (unpublished data) were based on a strain other than the original Dallas 1E.

DNA hybridization experiments by Garrity et al. (16) had earlier indicated that the three isolates of species 2 (U7W, U8W, and MICU-B) are distantly related to other isolates classified as *L. pneumophila*. They were shown to have a relative binding rate of 95 to 99% with one another, and comparable experiments in which these three isolates were compared with Philadelphia 1, as a reference, yielded only 64 to 67% relative annealing at 64°C and 35 to 42% relative annealing at 75°C. Reciprocally, with U7W and U8W as references, the comparable values for Philadelphia 1 were 59 and 63% at 64°C and 16 and 15% at 75°C. The estimates of D. J. Brenner (unpublished data) of relative binding rate at 75°C for these three isolates with Philadelphia 1 are significantly higher (mean, 54.7%; range, 41 to 63%) than those reported by Garrity et al. (16; mean, 37.7%; range, 35 to 42%), but in any event, these estimates are sufficiently low to justify specific distinction.

Notwithstanding the considerable genetic difference between isolates of species 2 and *L. pneumophila*, Garrity et al. (16) reported that serotyping and tests of 10 phenotypic characters failed to distinguish these isolates. These authors concluded that U7W, U8W, and MICU-B "are probably members of another species" but advised against making a formal taxonomic change until "additional serological or biochemical methods become available to differentiate these strains from *L. pneumophila*" (16). Our analysis confirms the genetic distinctness of these three isolates from *L. pneumophila*.

Clonal population structure. Enzyme variants detected by electrophoresis have been extensively used as chromosomal genetic markers to measure genetic relatedness among strains and to identify clones in studies of the genetic structure of *E. coli* (8, 42, 60) and other bacteria (37, 48). Because evolutionary convergence to the same multilocus genotype is highly improbable (49), isolates of identical ET are considered members of the same clone or cell line.

The repeated recovery of isolates of the same multilocus genotype at many different localities and at different times strongly suggests that the genetic structure of *L. pneumophila* is clonal. Although laboratory experiments have shown that *L. pneumophila* is capable of plasmid-mediated recombinational exchange of chromosomal material (9), we conclude that chromosomal recombination occurs very infrequently in natural populations. The recovery of seven of the ETs on two or more continents further suggests that many of the more common clones have achieved worldwide distribution. Hence, it follows that for *L. pneumophila* there must be a means of rapid dispersal over wide areas. Although legionellosis apparently is not transmitted person to person, the possibility that humans are agents of its dispersal cannot be dismissed. An alternative hypothesis is that long-distance dispersal is achieved primarily by wind transport, and, significantly, legionellae have been found in rain water samples collected in Pennsylvania (57) and Puerto Rico (C. B. Fliermans, personal communication). Laboratory experiments indicate that *L. pneumophila* can survive for many months in tap water or distilled water (51).

Further evidence of clonal structure is provided by the grouping of ETs shown in Fig. 3, which reflects nonrandom associations of alleles over loci. Strong linkage disequilibrium among loci may be generated in species having low rates of recombination as a result of the repeated loss of variation caused by the stochastic extinction of lines during episodes of periodic selection (also called nonspecific selection or indirect selection; 29, 31).

On the assumption that the electrophoretically demonstrable enzyme variation observed in bacteria is selectively neutral or nearly so (20), the considerable single-locus genetic diversity in *L. pneumophila* indicates that the species has maintained a large effective population size for a long time. For neutral alleles, the amount of genetic variation maintained in a species at equilibrium is a simple function of the effective population size and the mutation rate (28). Therefore, we can rule out the possibility that *L. pneumophila* has recently evolved, a hypothesis entertained soon after the bacterium was discovered in 1976, when it was known only from clinical cases and man-made habitats (see discussions by Starr [53] and Fliermans [13]). Because the species is worldwide in distribution and apparently ubiquitous in freshwater environments (14), we might have expected it to be at least as genetically diverse as *E. coli*. However, our estimate of mean genetic diversity for *L. pneumophila* ($\bar{h}$ = 0.313) is only 60% of the comparable estimate (0.52) obtained for ETs of a diverse collection of naturally occurring strains of *E. coli* (including *Shigella* spp.; 60). (Even when the ETs of species 1 and 2 are pooled with those of *L. pneumophila*, forming a sample of 62 ETs, diversity increases only to 0.413, or 79% of that for *E. coli*.) This circumstance raises the question of whether isolates of *L. pneumophila* recovered from clinical cases and man-made environments are only a limited, nonrandom subset of all genotypes existing in nature and perhaps represent a group of clones especially adapted for man-made environments. And the following observation suggests that this is indeed the case. All three isolates from natural environments that we have studied (MSH 2, MSH 5, and MSH 6, from springs and a pond on Mt. St. Helens, Wash.) have unusual multilocus genotypes. MSH 5 (ET-50) is, in fact, the most divergent isolate of *L. pneumophila* examined (Fig. 3). MSH 2 (ET-48) and MSH 6 (ET-47) are somewhat less unusual, but it may be significant that they fall within the 684 (ATCC 33733) clone family (ETs 44 through 48; Fig. 3), all other members of which were recovered from man-made environments and are as yet unknown clinically.

Analysis of much larger samples of isolates will be required to determine if the apparent lesser genetic variability of species 1 and 2, compared with that of *L. pneumophila*, is real or an artifact of inadequate sampling of populations. Perhaps both species 1 and 2 are abundant and highly polymorphic in natural habitats but have relatively few clones adapted for colonization of man-made environments. Further work on the natural populations of all three species is also needed to determine whether the species are, in fact, as genetically discrete as the isolates from clinical and man-made environmental sources indicate.

The paradox presented by the ubiquitous distribution of a nutritionally fastidious organism in freshwater habitats has recently been resolved by evidence that *L. pneumophila* obtains nutrients through association with green algae and cyanobacteria (46, 56) and is an intracellular parasite of protozoans (12, 47, 58). Host specialization and other ecological factors may also help account for the rich species diversity of the genus *Legionella* (2). It would be especially interesting to determine the relationships between clonal diversity in *L. pneumophila* and the distribution and abundance of the various microorganisms with which it is associated.

Variation in serotype in relation to population structure. In the virtual absence of information on the genetic relationships of isolates, research on *L. pneumophila* has heretofore been conducted within a framework provided by serotyping. Although heterogeneity of antigenic types has been detected by differential absorption and immunodiffusion of sera (7, 16, 62), only nine serogroups have been distinguished by conventional serological methods, and most isolates are serogroup 1 (35). Consequently, serotyping has been of limited use in distinguishing clones in epidemiological, systematic, and population genetics research. Our data show that serogroup 1 (and some other serogroups as well) occurs in association with a large variety of multilocus genotypes. The situation is similar to that in *E. coli*, in which genetic diversity among isolates of a given O (lipopolysaccharide) serogroup may be equivalent to that in the species as a whole (43) and in which otherwise unrelated clones may have the same O:K serotype (41). Moreover, serotyping does not distinguish *L. pneumophila* from either species 1 or 2, isolates of which have been variously serotyped as 1, 4, 4-5, and 5. The problem is that serotypes reflect variation in an unknown but probably small number of genetic loci encoding enzymes involved in the synthesis of antigenic moieties, the chemical structures of which are not well understood (15).

In sum, serotyping gives a grossly inaccurate picture of the genetic structure within *L. pneumophila* and species 1 and of the overall degree of relatedness among isolates of the three species we have studied. Indeed, it is precisely because isolates of species 1 are indistinguishable by conventional serotyping from those of *L. pneumophila* that they have in the past been mistakenly assigned to that species.

In the case of *Legionella* spp., as with *E. coli* and many other bacteria, there has been an unfortunate tendency for microbiologists to regard serogroups as genetically invariable units of population structure, when, in fact, they are only character states of isolates. Our analysis provides ample reason for discontinuing the practice of using serotype as the sole or primary criterion for the classification of strains.

These remarks should not be interpreted to mean that serogroups are distributed randomly among multilocus genotypes. Although only 36 (25%) of the 143 isolates of *L. pneumophila* typed for 22 enzyme loci are of serogroups other than 1, they represent 24 ETs (48%) of the total of 50 ETs recorded for the species. This fact indicates that genetic variance in the species is partitioned to a significant degree along lines of serotype. A second line of evidence of nonrandomness in the distribution of serogroups is the observation that isolates of ETs 40 through 47 and ET-49, which are genotypically rather unusual, were assigned to serogroups 2, 3, 4, 4-5, and 7 (ETs 48 and 50 are serogroup 1).

Our failure to examine isolates of serogroup 6 may have caused us to underestimate genetic diversity in *L. pneumophila* as a whole. Serogroup 6 isolates are not uncommon, and some significant part of the total genetic variance in *L. pneumophila* may be associated with them.

Variation in MAP in relation to population structure. On the basis of whether or not reactivity was obtained with antibody 2 of a panel of nine antibodies, McKinney et al. (34) divided 130 serogroup 1 isolates of *L. pneumophila* into two

subgroups, 1a and 1b. Isolates that we have designated in Table 7 as having MAPs 1, 2, 3, 6, 13, or 15 belong to subgroup 1b, and the other isolates for which MAPs are available belong to subgroup 1a.

We have examined 12 of 13 strains listed by McKinney et al. (34) as representatives of the two subgroups. One of the isolates assigned to subgroup 1b is D5401 (ET-4), which falls in the Davenport 1 cluster (ETs 1 through 8; Fig. 3), and the other 11 isolates are scattered through the large cluster bounded by ET-9 (Los Angeles 2) and ET-30 (Concord 3). Inasmuch as isolates assigned to each of the two subgroups are distributed, without apparent pattern, throughout the branches of the dendrogram containing serogroup 1 isolates (see MAP designations of isolates in Table 1), it is obvious that the classification by McKinney et al. (34) of MAPs bears no particular relationship to the underlying genetic structure of the species, apart from that associated with the distribution of serogroups. This conclusion is supported by the large proportion of genetic diversity observed within individual MAPs (Table 7).

Testing with a panel of 10 monoclonal antibodies, Watkins and Tobin (59) found that 45 of 50 serogroup 1 isolates of *L. pneumophila* could be placed in seven subgroups and that each of the remaining 5 isolates had a distinctive reactivity pattern. A similar degree of discrimination of isolates of serogroups 4 and 5 was obtained with another panel of 12 monoclonal antibodies. It is noteworthy that Dallas 1E and U7W, U8W, and MICU-B were placed together in a subgroup distinct from subgroups containing isolates of *L. pneumophila*. Division of 47 serogroup 1 isolates of *L. pneumophila* into five reactivity patterns with a panel of eight monoclonal antibodies made to Philadelphia 1 was reported by Joly et al. (25).

Although monoclonal antibodies have thus far been of little value in elucidating the genetic structure of *Legionella* species, they have been usefully employed in epidemiological research (45) and as a diagnostic tool (18).

Other studies of electrophoretic variation in proteins. Marre et al. (32) studied variation in β-lactamase in six isolates of *L. pneumophila* and in Dallas 1E and Dallas 2E, representing each of the then-known six serogroups, by isoelectric focusing. The enzymes of Philadelphia 1 (serogroup 1), Burlington 4 (serogroup 3), and JT (serogroup 6) were slow migrating; those of Togus 1 (serogroup 2), Burlington 5 (serogroup 4), and Marsh (serogroup 4) showed medium mobility; and the two isolates from Dallas (serogroup 5) had fast-migrating electromorphs. We have determined that Philadelphia 1 and Burlington 4 are members of the same clone (ET-25), as are Togus 1 and Burlington 5 (ET-40; Table 1). Moreover, Dallas 1E and Dallas 2E are members of the same clone (ET-53) of species 1.

The sodium dodecyl sulfate-polyacrylamide gel electrophoresis patterns of soluble protein extracts from a number of isolates of *L. pneumophila* and several other *Legionella* species were studied by Lema and Brown (30). Gels stained with either Coomassie blue or silver showed variation among isolates of *L. pneumophila* in the position and (with silver strain) the color of several minor bands, and there was a difference in banding pattern between Dallas 1E and Los Angeles 1 (both members of species 1, i.e., ET-53 and ET-54, respectively). Because the observed variation among strains and species was not analyzed quantitatively, we cannot relate the findings of Lema and Brown (30) to our estimates of genetic relationships among the isolates and species that we studied in common.

Disease syndromes in relation to genotype. Because the coefficient of genetic differentiation among ETs from clinical and environmental sources was small ($\overline{G}_{ST} = 0.028$), it is unlikely that these groups are genetically distinct. Consistent with this conclusion is the observation that mean genetic diversity among clinical isolates ($\overline{h} = 0.310$) was similiar to that among environmental isolates ($\overline{h} = 0.327$). These findings support the consensus opinion expressed in the literature that all strains of *L. pneumophila* and, indeed, most if not all *Legionella* species are capable of causing disease in humans. However, the relative frequencies of disease caused by the various species other than *L. pneumophila* have not been adequately determined (10, 61).

Our analysis has shown that organisms of the same multilocus genotype may cause both the pneumonic and nonpneumonic syndromes of legionellosis. The observation that isolates implicated as etiological agents of Pontiac fever in Michigan in 1968 and New York City in 1984 have similar genotypes (ET-17 and ET-21, respectively) could be interpreted as circumstantial evidence for the existence of special virulence factors in linkage disequilibrium with one or more of the enzyme loci assayed, were it not for the fact that ET-17 and ET-21 are among the more common clones of *L. pneumophila* and, hence, have relatively high probabilities of being associated with any outbreak of legionellosis. It is relevant that *L. feeleii*, a species showing only 10% DNA relatedness to *L. pneumophila*, has been strongly implicated as the cause of a severe outbreak of Pontiac fever in Windsor, Ontario, Canada (22).

Much further work with isolates known to have caused Pontiac fever will be required to determine whether, apart from their relative abundances, certain clones are more likely than others to cause this disease. Meanwhile, the fact that isolates of ET-17 are uniform in serogroup and MAP, as well as in multilocus genotype, lends some support to the hypothesis that the pneumonic and nonpneumonic forms of legionellosis are caused by viable and nonviable cells, respectively, of the same organism rather than by different strains or different virulence factors (see discussion in reference 6). Using the guinea pig model of airborne legionellosis, Huebner et al. (23) were unable to demonstrate a difference in virulence between Philadelphia 1 and Pontiac 1, and Katz and Matus (26) produced a Pontiac fever-like disease in guinea pigs by oral inoculation with Philadelphia 1.

Concluding comment. The primary objective of our research has been to define the genetic structure of *L. pneumophila* and related species to provide a genetic framework for analysis of the distribution and association of serogroups and other characteristics of these bacteria. In the past 20 years, the multilocus enzyme technique has been used to study variation and structure in hundreds of eukaryotic species (40), but only recently have serious attempts been made to apply this technique to bacteria. We trust that the present study demonstrates that this familiar technique has important contributions to make to the emerging field of bacterial population genetics (20), as well as to the more venerable areas of bacterial systematics and epidemiology.

ACKNOWLEDGMENTS

We thank P. H. Edelstein for supplying isolates from the Los Angeles, Calif., region.

This work was supported by Public Health Service grants DE 06031 and AI 22144 from the National Institutes of Health to R.K.S.

LITERATURE CITED

1. **Balows, A., and D. W. Fraser (ed.).** 1979. International symposium on Legionnaires' disease. Ann. Intern. Med. **90**:491–703.
2. **Brenner, D. J., A. G. Steigerwalt, G. W. Gorman, H. W. Wilkinson, W. F. Bibb, M. Hackel, R. L. Tyndall, J. Campbell, J. C. Feeley, W. L. Thacker, P. Skaliy, W. T. Martin, B. J. Brake, B. S. Fields, H. V. McEachern, and L. K. Corcoran.** 1985. Ten new species of *Legionella*. Int. J. Syst. Bacteriol. **35**:50–59.
3. **Brenner, D. J., A. G. Steigerwalt, G. A. Hebert, and J. E. McDade.** 1979. *Legionella pneumophila* sp. nov.: the Legionnaires' disease bacterium, p. 59–66. *In* G. L. Jones and G. A. Hebert (ed.), "Legionnaires'": the disease, the bacterium and methodology. U.S. Department of Health, Education, and Welfare publication no. (CDC) 79-8375. Center for Disease Control, Atlanta.
4. **Brenner, D. J., A. G. Steigerwalt, and J. E. McDade.** 1979. Classification of the Legionnaires' disease bacterium: *Legionella pneumophila*, genus novum, species nova, of the family Legionellaceae, familia nova. Ann. Intern. Med. **90**:656–658.
5. **Brenner, D. J., A. G. Steigerwalt, R. E. Weaver, J. E. McDade, J. C. Feeley, and M. Mandel.** 1978. Classification of the Legionnaires' disease bacterium: an interim report. Curr. Microbiol. **1**:71–75.
6. **Broome, C. V.** 1984. Current issues in epidemiology of legionellosis, 1983, p. 205–209. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
7. **Brown, A., R. M. Vickers, E. M. Elder, M. Lema, and G. M. Garrity.** 1982. Plasmid and surface antigen markers of endemic and epidemic *Legionella pneumophila* strains. J. Clin. Microbiol. **16**:230–235.
8. **Caugant, D. A., B. R. Levin, and R. K. Selander.** 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. Genetics **98**:467–490.
9. **Dreyfus, L. A., and B. H. Iglewski.** 1985. Conjugation-mediated genetic exchange in *Legionella pneumophila*. J. Bacteriol. **161**:80–84.
10. **Edelstein, P. H., D. J. Brenner, W. C. Moss, A. G. Steigerwalt, E. M. Francis, and W. L. George.** 1982. *Legionella wadsworthii* species nova: a cause of human pneumonia. Ann. Intern. Med. **97**:809–813.
11. **England, A. C., III, R. M. McKinney, P. Skaliy, and G. W. Gorman.** 1980. A fifth serogroup of *Legionella pneumophila*. Ann. Intern. Med. **93**:58–59.
12. **Fields, B. S., E. B. Shotts, Jr., J. C. Feeley, G. W. Gorman, and W. T. Martin.** 1984. Proliferation of *Legionella pneumophila* as an intracellular parasite of the ciliated protozoan *Tetrahymena pyriformis*. Appl. Environ. Microbiol. **47**:467–471.
13. **Fliermans, C. B.** 1984. Philosophical ecology: *Legionella* in historical perspective, p. 285–289. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
14. **Fliermans, C. B., W. B. Cherry, L. H. Orrison, S. J. Smith, D. L. Tison, and D. H. Pope.** 1981. Ecological distribution of *Legionella pneumophila*. Appl. Environ. Microbiol. **41**:9–16.
15. **Gabay, J. E., and M. A. Horwitz.** 1985. Isolation and characterization of the cytoplasmic and outer membranes of the Legionnaires' disease bacterium (*Legionella pneumophila*). J. Exp. Med. **161**:409–422.
16. **Garrity, G. M., E. M. Elder, B. Davis, R. M. Vickers, and A. Brown.** 1982. Serological and genotypic diversity among serogroup 5-reacting environmental *Legionella* isolates. J. Clin. Microbiol. **15**:646–653.
17. **Glick, T. H., M. D. Gregg, B. Berman, G. Mallison, W. W. Rhodes, Jr., and I. Kassanoff.** 1978. Pontiac fever: an epidemic of unknown etiology in a health department. I. Clinical and epidemiologic aspects. Am. J. Epidemiol. **107**:149–160.
18. **Guillet, J.-G., J. Hoebeke, C. Tram, S. Marullo, and A. D. Strosberg.** 1983. Characterization, serological specificity, and diagnostic possibilities of monoclonal antibodies against *Legionella pneumophila*. J. Clin. Microbiol. **18**:793–797.
19. **Haley, C. E., M. L. Cohen, J. Halter, and R. D. Meyer.** 1979. Nosocomial Legionnaires' disease: a continuing common-source epidemic at Wadsworth Medical Center. Ann. Intern. Med. **90**:583–586.
20. **Hartl, D. L., and D. E. Dykhuizen.** 1984. The population genetics of *Escherichia coli*. Annu. Rev. Genet. **18**:31–68.
21. **Hebert, G. A., A. G. Steigerwalt, and D. J. Brenner.** 1980. *Legionella micdadei* species nova: classification of a third species of *Legionella* associated with human pneumonia. Curr. Microbiol. **3**:255–257.
22. **Herwaldt, L. A., G. W. Gorman, T. McGrath, S. Toma, B. Brake, A. W. Hightower, J. Jones, A. L. Reingold, P. A. Boxer, P. W. Tang, C. W. Moss, H. Wilkinson, D. J. Brenner, A. G. Steigerwalt, and C. V. Broome.** 1984. A new *Legionella* species, *Legionella feeleii* species nova, causes Pontiac fever in an automobile plant. Ann. Intern. Med. **100**:333–338.
23. **Huebner, R. E., P. W. Reeser, and D. W. Smith.** 1984. Comparison of the virulence of the Philadelphia and Pontiac isolates of *Legionella pneumophila*, p. 123–124. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
24. **Johnson, S. R., and W. O. Schalla.** 1982. Plasmids of serogroup 1 strains of *Legionella pneumophila*. Curr. Microbiol. **7**:143–146.
25. **Joly, J. R., Y. Y. Chen, and D. Ramsay.** 1984. Subtyping of *Legionella pneumophila* serogroup 1 by using monoclonal antibodies, p. 264–265. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
26. **Katz, S. M., and J. P. Matus.** 1984. Febrile illness produced in guinea pigs by oral inoculation of *Legionella pneumophila*, p. 124–127. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
27. **Kaufmann, A. F., J. E. McDade, C. M. Patton, J. V. Bennett, P. Skaliy, J. C. Feeley, D. C. Anderson, M. E. Potter, V. F. Newhouse, M. B. Gregg, and P. S. Brachman.** 1981. Pontiac fever: isolation of the etiologic agent (*Legionella pneumophila*) and demonstration of its mode of transmission. Am. J. Epidemiol. **114**:337–347.
28. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.
29. **Koch, A. L.** 1974. The pertinence of the periodic selection phenomenon to prokaryote evolution. Genetics **77**:127–142.
30. **Lema, M., and A. Brown.** 1983. Electrophoretic characterization of soluble protein extracts of *Legionella pneumophila* and other members of the family *Legionellaceae*. J. Clin. Microbiol. **17**:1132–1140.
31. **Levin, B. R.** 1981. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. Genetics **99**:1–23.
32. **Marre, R., A. A. Medeiros, and A. W. Pasculle.** 1982. Characterization of the beta-lactamases of six species of *Legionella*. J. Bacteriol. **151**:216–221.
33. **McDade, J. E., D. J. Brenner, and F. M. Bozeman.** 1979. Legionnaires' disease bacterium isolated in 1947. Ann. Intern. Med. **90**:659–661.
34. **McKinney, R. M., L. Thacker, D. E. Wells, M. C. Wong, W. J. Jones, and W. F. Bibb.** 1983. Monoclonal antibodies to *Legionella pneumophila* serogroup 1: possible applications in diagnostic tests and epidemiologic studies. Zentralbl. Bakteriol. Mikrobiol. Hyg. Abt 1 Orig. A **255**:91–95.
35. **McKinney, R. M., H. W. Wilkinson, H. M. Sommers, B. J. Fikes, K. R. Sasseville, M. M. Yungbluth, and J. S. Wolf.** 1980. *Legionella pneumophila* serogroup six: isolation from cases of legionellosis, identification by immunofluorescence staining, and immunological response to infection. J. Clin. Microbiol. **12**:395–401.
36. **Meyer, R. D.** 1983. *Legionella* infections: a review of five years of research. Rev. Infect. Dis. **5**:258–278.
37. **Musser, J. M., D. M. Granoff, P. E. Pattison, and R. K. Selander.** 1985. A population genetic framework for the study of invasive diseases caused by serotype b strains of *Haemophilus influenzae*. Proc. Natl. Acad. Sci. U.S.A. **82**:5078–5082.

38. **Nei, M.** 1977. F-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet. **41**:225–233.
39. **Nei, M.** 1978. Estimation of average heterozygosity and genetic distance from a small sample of individuals. Genetics **89**:583–590.
40. **Nevo, E., A. Beiles, and R. Ben-Shlomo.** 1984. The evolutionary significance of genetic diversity: ecological, demographic and life history correlates, p. 13–213. *In* G. S. Mani (ed.), Evolutionary dynamics of genetic diversity. Lecture notes in biomathematics, 53. Springer-Verlag, Berlin.
41. **Ochman, H., and R. K. Selander.** 1984. Evidence for clonal population structure in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **81**:198–201.
42. **Ochman, H., T. S. Whittam, D. A. Caugant, and R. K. Selander.** 1983. Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. J. Gen. Microbiol. **129**:2715–2726.
43. **Ochman, H., R. A. Wilson, T. S. Whittam, and R. K. Selander.** 1984. Genetic diversity within serotypes of *Escherichia coli*, p. 202–224. *In* Proceedings of the Fourth International Symposium on Neonatal Diarrhea. VIDO Publications, University of Saskatchewan, Saskatoon, Canada.
45. **Para, M. F., and J. F. Plouffe.** 1983. Production of monoclonal antibodies to *Legionella pneumophila* serogroups 1 and 6. J. Clin. Microbiol. **18**:895–900.
46. **Pope, D. H., R. J. Soracco, H. K. Gill, and C. B. Fliermans.** 1982. Growth of *Legionella pneumophila* in two-membered cultures with green algae and cyanobacteria. Curr. Microbiol. **7**:319–322.
47. **Rowbotham, T. J.** 1980. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. J. Clin. Pathol. **33**:1179–1183.
48. **Schill, W. B., S. R. Phelps, and S. W. Pyle.** 1984. Multilocus electrophoretic assessment of the genetic structure and diversity of *Yersinia ruckeri*. Appl. Environ. Microbiol. **48**:975–979.
49. **Selander, R. K., and B. R. Levin.** 1980. Genetic diversity and structure in *Escherichia coli* populations. Science **210**:545–547.
50. **Selander, R. K., M. H. Smith, S. Y. Yang, W. E. Johnson, and J. B. Gentry.** 1971. Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*), p. 49–90. *In* Studies in genetics VI. University of Texas publication no. 7103. University of Texas Press, Austin.
51. **Skaliy, P., and H. V. McEachern.** 1979. Survival of the Legionnaires' disease bacterium in water. Ann. Intern. Med. **90**:662–663.
52. **Sneath, P. H. A., and R. R. Sokal.** 1978. Numerical taxonomy. W. H. Freeman & Co., San Francisco.
53. **Starr, M. P.** 1979. Plant-associated bacteria as human pathogens: disciplinal insularity, ambilateral harmfulness, epistemological primacy. Ann. Intern. Med. **90**:708–710.
54. **Stout, J., V. L. Yu, R. M. Vickers, J. Zuravleff, M. Best, A. Brown, R. B. Yee, and R. Wadowsky.** 1982. Ubiquitousness of *Legionella pneumophila* in the water supply of a hospital with endemic Legionnaires' disease. N. Engl. J. Med. **306**:466–468.
55. **Thornsberry, C., A. Balows, J. C. Feeley, and W. Jakubowski (ed.).** 1984. Legionella. American Society for Microbiology, Washington, D.C.
56. **Tison, D. L., D. H. Pope, W. B. Cherry, and C. B. Fliermans.** 1980. Growth of *Legionella pneumophila* in association with blue-green algae (cyanobacteria). Appl. Environ. Microbiol. **39**:456–459.
57. **Turner, W. E., P. Nash, G. F. Harada, and L. M. Iampietro.** 1984. Detection of *Legionella pneumophila* and *Legionella micdadei* in water collected from a rainfall in southeastern Pennsylvania, p. 321–322. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
58. **Tyndall, R. L., and E. L. Domingue.** 1982. Cocultivation of *Legionella pneumophila* and free-living amoebae. Appl. Environ. Microbiol. **44**:954–959.
59. **Watkins, I. D., and J. O. Tobin.** 1984. Studies with monoclonal antibodies to *Legionella* species, p. 259–262. *In* C. Thornsberry, A. Balows, J. C. Feeley, and W. Jakubowski (ed.), Legionella. American Society for Microbiology, Washington, D.C.
60. **Whittam, T. S., H. Ochman, and R. K. Selander.** 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **80**:1751–1755.
61. **Wilkinson, H. W., A. L. Reingold, B. J. Brake, D. L. McGiboney, G. W. Gorman, and C. V. Broome.** 1983. Reactivity of serum from patients with suspected legionellosis against 29 antigens of Legionellaceae and *Legionella*-like organisms by indirect immunofluorescence assay. J. Infect. Dis. **147**:23–31.
62. **Wong, K. H., W. O. Schalla, M. C. Wong, P. R. B. McMaster, J.C. Feeley, and R. J. Arko.** 1982. Biologic activities of antigens from *Legionella pneumophila*, p. 434–443. *In* J. Robbins, J. Hill, and J. Sadoff (ed.), Bacterial vaccines. Thieme-Stratton, New York.