Journal of Clinical Microbiology

# Comparative Analysis of Multilocus Sequence Typing and Pulsed-Field Gel Electrophoresis for Characterizing *Listeria monocytogenes* Strains Isolated from Environmental and Clinical Sources

**Tamara Revazishvili, Mamuka Kotetishvili, O. Colin Stine, Arnold S. Kreger, J. Glenn Morris Jr. and Alexander Sulakvelidze**

***J. Clin. Microbiol.*** **2004, 42(1):276. DOI: 10.1128/JCM.42.1.276-285.2004.**

JOURNAL OF CLINICAL MICROBIOLOGY, Jan. 2004, p. 276–285 Vol. 42, No. 1
0095-1137/04/$08.00+0 DOI: 10.1128/JCM.42.1.276–285.2004


# Comparative Analysis of Multilocus Sequence Typing and Pulsed-Field Gel Electrophoresis for Characterizing *Listeria monocytogenes* Strains Isolated from Environmental and Clinical Sources

Tamara Revazishvili,† Mamuka Kotetishvili,† O. Colin Stine, Arnold S. Kreger, J. Glenn Morris, Jr., and Alexander Sulakvelidze*

*Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201*

Received 27 June 2003/Returned for modification 12 August 2003/Accepted 18 September 2003

**One hundred seventy-five *Listeria monocytogenes* strains were characterized by serotyping, pulsed-field gel electrophoresis (PFGE), and multilocus sequence typing (MLST) based on loci in *actA*, *betL*, *hlyA*, *gyrB*, *pgm*, and *recA*. One hundred twenty-two sequence types (STs) were identified by MLST based on allelic profiles of the four housekeeping genes (*betL*, *gyrB*, *pgm*, and *recA*), and 34 and 38 alleles were identified for *hlyA* and *actA*, respectively. Several *actA* and *hlyA* alleles appeared to be predominantly associated with clinical isolates. MLST differentiated most of the *L. monocytogenes* strains better than did PFGE, and the discriminating ability of PFGE was better than that of serotyping. Several strains with different serotypes were found, by MLST and PFGE, to have very closely related genetic backgrounds, which suggested possible "antigen switching" among them. MLST can be a useful typing tool for differentiating *L. monocytogenes* strains (including strains undistinguishable by PFGE typing and serotyping), and it may be of value during investigations of food-borne outbreaks of listeriosis.**

*Listeria monocytogenes* is an important food-borne pathogen which causes ca. 2,500 listeriosis cases/year, and it is responsible for ca. 500 deaths/year in the United States (28). Because of the high fatality rate (20 to 30%) associated with *L. monocytogenes* infection, the U.S. Food and Drug Administration and the Food Safety and Inspection Service of the U.S. Department of Agriculture have established a "zero tolerance" policy (i.e., no detectable levels permitted) for the species in ready-to-eat foods (37). However, despite the zero tolerance policy and multimillion dollar food recalls due to *L. monocytogenes* contamination, several outbreaks of listeriosis have occurred during the 1980s and 1990s, with a variety of contaminated foods, including coleslaw (34), unpasteurized cheese (24), butter (26), pasteurized milk (11), and other foods (10, 37), implicated as sources of the etiologic agent. Thus, improving the ability to identify outbreak-causing strains rapidly and to trace them to sources of contamination is important for addressing many of the epidemiological, clinical, and legal issues associated with listeriosis outbreaks.

Several phenotypic and genotypic typing methodologies have been used to subtype *L. monocytogenes*, including serotyping (35), electrophoretic typing of esterases (zymotyping) (20), multilocus enzyme electrophoresis (MLEE) (32), ribotyping (12), random amplification of polymorphic DNA (RAPD) (4), and pulsed-field gel electrophoresis (PFGE) (14). The methods differ in their discriminatory abilities and reproducibility, and some of them, e.g., phage-typing and serotyping, are limited to a few laboratories that have access to specific typing phages and sera, respectively. In general, DNA-based typing approaches are recognized as having better discriminatory power than phenotypic approaches and being better suited for investigating *L. monocytogenes* outbreaks (25). Among molecular typing approaches, ribotyping and PFGE typing have been reported (20) to be superior to other approaches, with the discriminatory power of PFGE superior to that of ribotyping (1). However, even those superior approaches do not always adequately distinguish between strains of *L. monocytogenes*, including strains of serotypes 1/2b and 4b primarily associated with human listeriosis (10, 37), and they are poorly suited for long-term epidemiological studies and for determining the evolution and phylogenetic relationships among various *L. monocytogenes* strains and serotypes.

Sequencing the entire bacterial genome is likely to differentiate *Listeria* strains extremely well and to yield information about their phylogeny. However, that approach is a major undertaking which is not yet technically feasible to use for investigating listeriosis outbreaks. Sequencing a single *L. monocytogenes* gene (7) also may not be optimal for accurate comparison of the isolates, because the data may be either noninformative, if the gene selected for analysis is highly conserved, or misleading, if the gene selected for analysis is highly variable. Also, because evolution occurs by a net-like process, gene trees based on a single gene are not likely to permit accurate determination of the genetic relatedness among various isolates and the evolutionary history of the species (16). Thus, a technique involving simultaneous, sequence-based analysis of several genes should provide the ideal balance between sequence-based resolution and technical feasibility. Such an approach, called multilocus sequence typing (MLST), was recently used by Maiden et al. (27) to characterize meningococci and has rapidly been applied to the study of many

* Corresponding author. Mailing address: Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, MSTF Bldg., 10 South Pine St., Baltimore, MD 21201. Phone: (410) 706-4587. Fax: (410) 706-4581. E-mail: asulakve@epi.umaryland.edu.

† T.R. and M.K. contributed equally to the research described in this paper.

TABLE 1. Primers used for MLST, numbers of alleles and polymorphic sites identified per gene, and $d_N/d_S$ ratios for various genes determined for a subset of 157 *L. monocytogenes* isolates

| Gene | Primers (5′→ 3′) | Maximum length of amplified fragments (bp) | Length of fragments analyzed by START (bp) | No. of alleles | No. of polymorphic sites | Mean GC content (%) | $d_N/d_S$ | Sequence variation(%) |
|---|---|---|---|---|---|---|---|---|
| *actA* | CACAGATGAATGGGAAGAAGAA<br>CTTGTAAAACTAGAATCTAGCGA | 963 | 415 | 24 | 42 | 40.05 | 0.23 | 10.1 |
| *betL* | GCTGGAATGGGAATTGG<br>GAAAGCCACCAAGCCCAATA | 650 | 400 | 17 | 31 | 36.99 | 0.03 | 7.8 |
| *hlyA* | GCAAAATTTGGTACGGCATTTAAA<br>ACCGTTCTCCACCATTCCCAAGC | 830 | 453 | 15 | 17 | 37.25 | 0.00 | 3.8 |
| *gyrB* | GGGCGTTGTGGCTTCTCGTGCACG<br>CGGCATCAATAGCGTCTTTGATAT | 708 | 406 | 15 | 20 | 39.55 | 0.00 | 4.9 |
| *pgm* | GGCGCTGGCGGCTTTTTTGCTGA<br>GAACCAACCTTCACCAAGCGCACC | 768 | 409 | 14 | 18 | 39.57 | 0.05 | 4.4 |
| *recA* | GTGAATGATCGTCAAGCGGC<br>AGTAGAATAGAATTTAAGCGCACG | 655 | 466 | 12 | 34 | 41.06 | 0.00 | 7.3 |

other bacterial species. Because of its robustness and its ability to differentiate a vast number of genotypes in which genetic variations accumulate relatively slowly (in the evolutionary sense), MLST has recently been suggested (6, 36) to be a valuable primary tool for phylogenetic analysis of various bacteria.

MLST has rapidly gained recognition as one of the best molecular typing approaches available today, and it has been used to characterize several bacteria, including *L. monocytogenes* (5, 33). However, several important technical issues related to MLST of *L. monocytogenes* have remained to be addressed, including identifying the genes and the number of gene loci that provide optimal resolution and comparing MLST with other currently available typing approaches for differentiating strains of the bacterium. Thus, the goals of this study were (i) to determine which of several housekeeping and virulence-associated gene loci are most useful in the MLST-based differentiation of *L. monocytogenes* isolates and (ii) to compare the discriminatory ability of MLST with that of PFGE and serotyping. The data generated during the course of our studies also enabled us to determine the genetic relatedness among various clinical and environmental *L. monocytogenes* strains and serotypes and to gain preliminary insight into the evolutionary mechanisms responsible for *L. monocytogenes* diversity.

## MATERIALS AND METHODS

**Bacterial strains.** A total of 175 *L. monocytogenes* strains (74 environmental isolates, 94 clinical isolates, and 7 isolates of unknown origin) were examined in this study. All strains were isolated in the United States. Clinical strains were identified to species level by, and obtained from, various public health laboratories (information regarding which clinical isolates were isolated from blood specimens, stool specimens, etc., was not available). Environmental strains were obtained from the National Food Processors Association (food isolates) and a major U.S. poultry producer's research laboratory. The Centers for Disease Control and Prevention (CDC) standard *L. monocytogenes* strain H2446 was included in the strain collection and was used as a reference strain during all experiments. Serotyping was performed by a commercial laboratory specializing in serotyping of *L. monocytogenes* by the standard tube agglutination method.

**PFGE.** Plugs were prepared, and PFGE was performed, according to the CDC PulseNet standardized procedure for typing *L. monocytogenes* (15), by using the CHEF DR II apparatus (Bio-Rad Laboratories, Hercules, Calif.). The DNA in agarose plugs was digested by incubating the plugs (at 37°C for 4 h) with *Asc*I, and electrophoresis was performed in a 1% agarose gel (in 0.5× Tris-borate-EDTA buffer). The following electrophoresis conditions were used: voltage, 180 V; initial switch time, 4.0 s; final switch time, 40.01 s; run time, 16 h. Lambda ladder pulsed-field grade (PFG) and low-range PFG molecular weight markers were loaded on all gels (at least two markers per gel). *Asc*I-digested DNA from *L. monocytogenes* H2446 was included, as a reference, in all PFGE gels.

**MLST.** Six gene loci were selected for MLST analysis, including regions from *actA* (encoding an actin recruitment and polymerization protein), *betL* (encoding a glycine betaine transporter), *hlyA* (encoding listeriolysin O), *gyrB* (encoding the DNA gyrase B subunit), *pgm* (encoding phosphoglucomutase), and *recA* (encoding a DNA repair and recombination protein). Primers (Table 1) were designed by aligning corresponding *Listeria* GenBank (http://www.ncbi.nlm.nih.gov) sequences by use of ClustalX (19) and selecting conserved regions flanking potentially variable internal fragments of the targeted genes. The same primers were used for PCR amplification and sequencing.

Bacterial DNA for PCR was prepared by extracting chromosomal DNA from the PFGE plugs. (Since we already had PFGE plugs, it was convenient to use them to extract the DNA required for the MLST analysis; DNA for MLST can be extracted directly from bacteria by any other conventional means.) Briefly, plugs containing bacterial DNA were frozen and thawed (at −80 and 60°C, respectively) twice in TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]), the supernatant fluids were collected after clarification of the samples by centrifugation (at 7,000 × *g* for 1 min), and aliquots (1 μl) of the resulting supernatant fluids (containing approximately 50 ng of bacterial DNA) were used as templates for PCR amplification. For most samples, the PCR amplification conditions were as follows: an initial cycle at 94°C for 5 min; 35 amplification cycles, each consisting of sequential incubation at 94°C for 45 s, 55°C for 45 s, and 72°C for 1 min; and a final incubation at 72°C for 5 min. When the primers for *betL* and *gyrB* were used, different annealing temperatures (51 and 56°C, respectively) were sometimes required for optimal amplification. PCR was performed using a RoboCycler Gradient 96 machine (Stratagene, La Jolla, Calif.). Amplified fragments were sequenced in both directions using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems, Inc., Foster City, Calif.), and the labeled fragments were separated using an ABI 3700 DNA analyzer (Applied Biosystems, Inc.).

**Data analysis.** PFGE patterns were compared by means of the Dice coefficient, by using Fingerprinting DST Molecular Analyst software (Bio-Rad Laboratories). A dendrogram (see Fig. 1) was constructed by the unweighted pair group method using averages (UPGMA); a tolerance of 3% in the band position was applied. Computer-assisted analyses were performed according to the manufacturer's instructions.

The internal fragment sequences of *actA*, *hlyA*, *betL*, *gyrB*, *pgm*, and *recA* were determined during MLST. Reading of trace files and assembly of contigs were performed by using the Phred (8, 9) and Phrap (available at http://www.washington.edu) programs, respectively. The sequences were trimmed, and they were aligned by the ClustalX program (19). Each unique sequence was randomly assigned a distinct allele number. The sequence type analysis and recombinational test (START) program (http://outbreak.ceid.ox.ac.uk) was used to construct an MLST-based dendrogram and to determine the GC content, numbers of alleles and polymorphic sites, and proportions of nonsynonymous and synonymous base substitutions ($d_N$ and $d_S$, respectively). The split decomposition method, as implemented in the Splitstree program (17), was used to test for recombination.

**Nucleotide sequence accession numbers.** The DNA sequences of the loci in *actA*, *betL*, *hlyA*, *gyrB*, *pgm*, and *recA* have been deposited in GenBank under accession numbers AY228804 through AY228978 (*actA*), AY228979 through AY229153 (*betL*), AY229154 through AY229328 (*gyrB*), AY229329 through AY229503 (*hlyA*), AY229504 through AY229678 (*pgm*), and AY229679 through AY229853 (*recA*).

## RESULTS

**Serotyping.** Eleven of the 175 isolates were not typeable (NT) with the standard *L. monocytogenes* antisera, and the remaining 164 strains were grouped into eight serotypes. The predominant serotypes were 1/2a (99 strains [ca. 60% of the serotypeable strains], 43 of which were clinical isolates), 4b (32 strains [ca. 20%], including 24 clinical isolates), and 1/2b (20 strains [ca. 8%], including 15 clinical isolates). The remaining 13 strains were grouped into serotypes 1/2c, 3a, 3b, 4a, and 4d.

**PFGE.** Fifty-seven PFGE types were identified among the 175 strains analyzed. PFGE type P1 predominated and included 19 strains (11 environmental isolates and 8 clinical isolates), followed by type P2 (11 environmental isolates and 2 clinical isolates) and type P3 (10 environmental isolates and 2 clinical isolates). Twenty-seven PFGE types (e.g., PFGE types P38, P39, P40, etc.) were represented by only a single isolate. The CDC standard strain H2446 was grouped in PFGE type P4, together with 10 other isolates. Clinical isolates predominated in some PFGE types (e.g., P4 included eight clinical isolates and three environmental isolates), while environmental isolates predominated in others (e.g., P3 included nine environmental isolates and two clinical isolates). However, genetic clusters encompassing predominantly clinical isolates or environmental isolates were not identified by the PFGE-based dendrogram (Fig. 1).

**MLST.** Sequence types (STs) were assigned based on the allelic profiles of the four housekeeping genes (*betL*, *gyrB*, *pgm*, and *recA*), which were determined based on their full-length sequences submitted to GenBank. The same allelic profiles were used to construct the MLST-based dendrogram (Fig. 2) by using the START program. One hundred twenty-two sequence types (STs) were identified by MLST, and 34 and 38 allele types were identified for *hlyA* and *actA*, respectively. Most of the STs (ca. 79%; 97 of 122 STs) contained only one *L. monocytogenes* isolate; however, the remaining 25 STs contained more than one isolate (e.g., seven *L. monocytogenes* isolates were grouped in ST55, and four isolates were grouped in ST56). The MLST-based dendrogram (Fig. 2) revealed two major clusters: cluster A (72 strains) and cluster B (97 strains). Six *L. monocytogenes* strains were not grouped in either of the two clusters.

The START program requires that sequences of equal length be analyzed in order to determine the numbers of alleles and polymorphic sites, mean G+C content, and $d_N/d_S$ ratios. Therefore, the sequences were trimmed for those analyses. Also, 18 strains were excluded from the START analysis because their sequences were less than 400 bp, the minimum sequence length we included in the analysis. The lengths of the trimmed DNA fragments varied from 400 bp for *betL* to 466 bp for *recA* (Table 1), sizes comparable to those of the fragments recently analyzed (33) for *L. monocytogenes* strains isolated in Spain (one of the six loci [*pgm*] we analyzed also was examined during that study). Among the six loci we analyzed by START, the number of alleles varied from 24 (for *actA*) to 12 (for *recA*) (Table 1). The *actA* gene region was the most variable (42 polymorphic sites and ca. 10.1% sequence variation), followed by the *betL* gene fragments (ca. 7.8% sequence variation) and the *recA* gene fragments (ca. 7.3% sequence variation) (Table 1).

The Splitstree program was used to detect recombination among various STs. Allelic profile data (based on the four housekeeping genes analyzed) were converted into distant matrix values by using START, and the resulting nexus file was analyzed by the split decomposition method, as implemented in the Splitstree program. Splitstree analysis of our 175 *L. monocytogenes* strains yielded a very low fit value (fit = 16) (data not shown), which may have resulted from the program's inability to analyze the large amount of information (21). Therefore, we also analyzed several small subsets of randomly selected strains, which considerably improved the fit values. One example of a Splitstree analysis displaying 10 parallelograms and a high fit value of ca. 89, based on 18 randomly selected *L. monocytogenes* isolates, is shown in Fig. 3.

The G+C contents varied slightly among the loci we analyzed. The *hlyA* region had the lowest G+C content (ca. 37%), and the *recA* region had the highest (ca. 41%) (Table 1). The $d_N/d_S$ ratios were $<1$ for all six loci, and they were zero for the *hlyA*, *gyrB*, and *recA* loci (Table 1).

**Correlation between serotyping, PFGE typing, and MLST.** In most cases, strains with the same ST and PFGE type belonged to the same serotype. However, there were exceptions. For example, *L. monocytogenes* strains 84 and 57 had ST121 and PFGE type P16 (both strains also had the same *actA* and *hlyA* alleles [17 and 10, respectively]), but they had different serotypes: 1/2b and 3b, respectively (Fig. 2). Also, although seven *L. monocytogenes* strains were grouped in ST28 and PFGE type P1, five of them had serotype 1/2a and two (strains 128 and 132) had serotype 1/2b (Fig. 2).

In some cases, strains with identical genetic backgrounds (i.e., the same ST and PFGE type) also had the same *actA* and *hlyA* alleles. For example, the seven strains in ST28 (PFGE type P1) had the same *actA* and *hlyA* alleles (allele 6 and allele 9, respectively), as did the two strains in ST108 (PFGE type P29; allele 33 and allele 29, respectively). However, some strains with identical genetic backgrounds did not have the same *actA* and *hlyA* alleles (e.g., strains 66 and 166 had the same ST30 and PFGE type P5, but their *hlyA* alleles were not the same) (Fig. 2).

Several strains with the same PFGE type (i.e., they were undistinguishable by PFGE-typing) were differentiated and assigned distinct STs by MLST (e.g., PFGE type P6 included eight isolates which were further subdivided into eight STs [Fig. 2]). In some cases, strains with the same PFGE type but distinct STs were not closely related based on the MLST dendrogram (i.e., the START-constructed dendrogram based on the allelic profiles of the four housekeeping genes). For example, strains 68 and 89 in PFGE type P5 were grouped in distinct STs (ST76 and ST27, respectively), and they were not closely related according to MLST, because the linkage distance between them was ca. 0.65 (on the START tree, the linkage distance between identical strains was zero, and the linkage distance between two most closely related, but not identical, isolates was 0.25) (Fig. 2). In some instances, strains with the

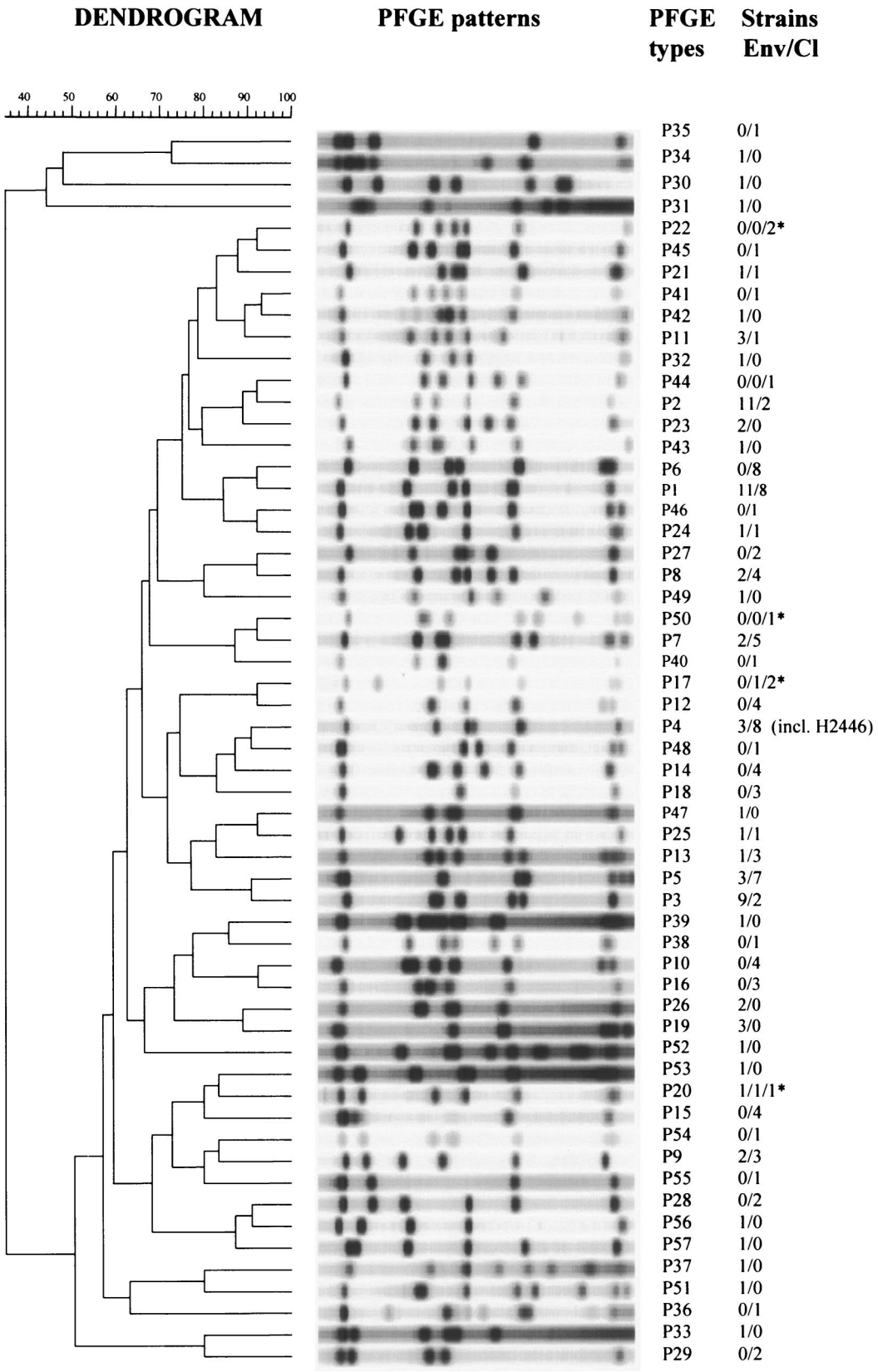


FIG. 1. PFGE-based dendrogram demonstrating the genetic diversity of various *L. monocytogenes* strains. Asterisks mark strains of unknown origin.

*Continued on following page*

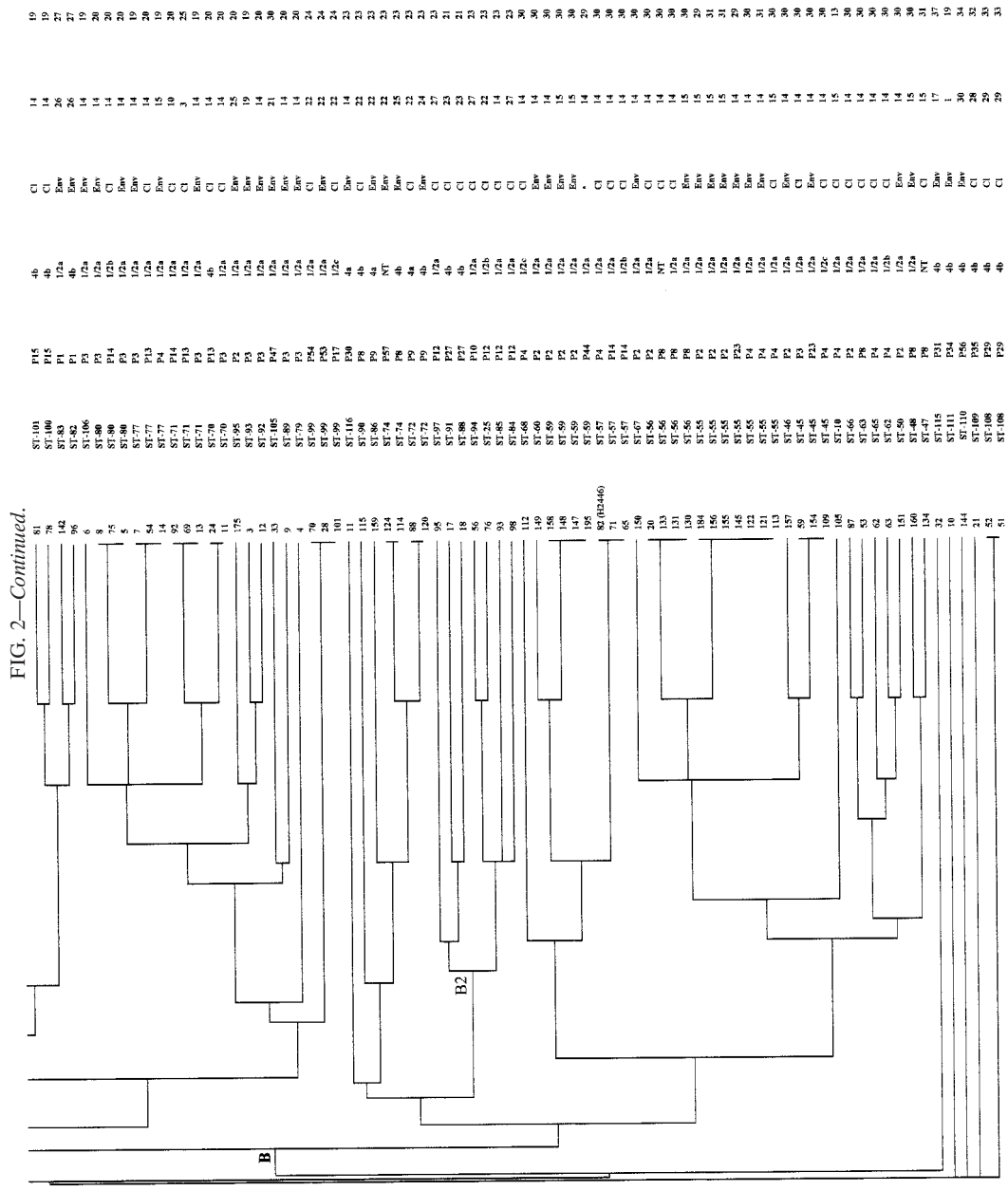


FIG. 2—*Continued.*

FIG. 2. MLST-based dendrogram of *L. monocytogenes* strains.

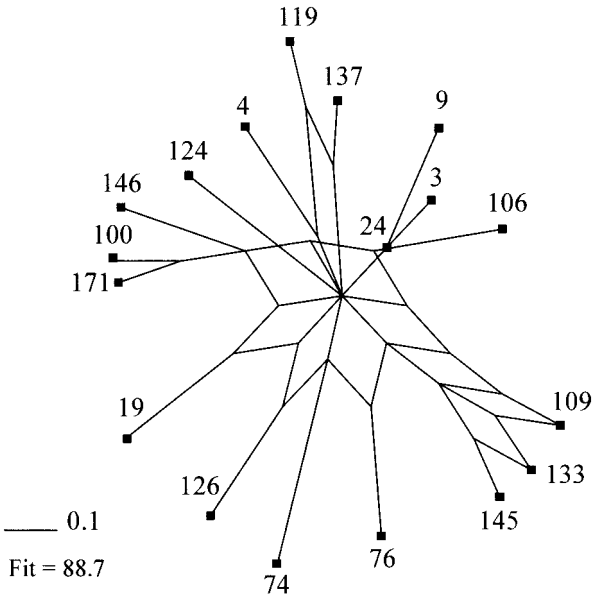


FIG. 3. Split decomposition analysis of 18 randomly selected *L. monocytogenes* strains.

same ST had distinct PFGE types. For example, isolates 5, 8, and 75 were assigned the same sequence type (ST80 [Fig. 2]), but they were differentiated into two closely related but distinct PFGE types: P3 and P14 (Fig. 1).

**Clustering of clinical and environmental isolates.** No clear-cut correlation between serotypes and the source of isolation of the strains (environmental versus clinical) was observed, although strains of the 1/2b and 4b serotypes seemed to predominate among the clinical specimens (75% of strains belonging to serotypes 1/2b and 4b were clinical isolates). We did not observe clusters (PFGE clusters were defined as groups of at least three distinct PFGE types) of clinical and environmental isolates in the PFGE-based dendrogram (Fig. 1). However, START analysis of the MLST data (Fig. 2) revealed three predominantly "clinical" subclusters (MLST subclusters were defined as groups containing a minimum of five nonidentical strains). The first subcluster (A1) contained 10 strains, of which 8 were clinical isolates (strains 15, 66, 67, 68, 74, 89, 165, and 166). All of the clinical strains in this subcluster were of the "pathogenic" serotypes 1/2a and 1/2b (strain 89 was not serotypeable), and they were grouped in two distinct PFGE-types (P5 and P36 [Fig. 1]). Subcluster B1 contained 17 strains (15 of which were clinical isolates) grouped in four PFGE types (P1, P6, P19, and P38). The third subcluster (B2) contained seven clinical strains (strains 17, 18, 56, 76, 93, 95, and 98) grouped in three PFGE types (P10, P12, and P27). As with subclusters A1 and B1, the seven strains in subcluster B2 were limited to "pathogenic" serotypes (1/2a, 1/2b and 4b). The strains in each of the three "clinical" clusters were not identical, i.e., they had distinct STs and PFGE types, and the linkage distance among strains in the same cluster was $\geq$0.25. Also, they were not clustered by PFGE, and many strains in the same MSLT subcluster were not closely related according to the PFGE-based dendrogram (Fig. 1).

*hlyA* alleles 10, 14, and 27 were most common in the A1 (5 of 10 strains), B1 (9 of 17 strains), and B2 (3 of 7 strains) "clinical" subclusters, respectively. Also, *actA* alleles 7, 31, and 23 were most common in the same subclusters: they were present in 8 of 10 strains in subcluster A1, 7 of 17 strains in subcluster B1, and 5 of 7 strains in subcluster B2, respectively. When our entire strain collection was analyzed, the same alleles (except for *hlyA* allele 10) also appeared to be associated with the clinical isolates. For example, 5 of the 175 isolates we analyzed had *hlyA* allele 27, and 4 of those 5 strains (80%) were clinical isolates; 11 of the 175 strains we analyzed had *actA* allele 31, and 8 (72%) of those strains were clinical isolates; 9 of the 175 isolates we analyzed had *actA* allele 7, and 6 (67%) of those strains were clinical isolates (Fig. 2). As noted above, in some strains, specific *actA* alleles were associated with specific *hlyA* alleles and vice versa; e.g., *hlyA* allele 3, in strains 136 (ST22), 139 (ST23), and 176 (ST42), was associated with *actA* allele 1 (Fig. 2). In general, however, no clear-cut correlation between the *hlyA* and *actA* alleles was observed. For example, strains with *hlyA* allele 10 had seven different *actA* alleles (alleles 7, 9, 11, 12, 17, 20, and 37), and strains with *actA* allele 31 had three different *hlyA* alleles (alleles 14, 15, and 24) (Fig. 2).

## DISCUSSION

**Genetic heterogeneity of *L. monocytogenes* strains and the role of recombination in *L. monocytogenes* diversity.** The genetic composition of the strain population we analyzed was moderately heterogeneous; for example, ca. 1 PFGE type was identified per 3 isolates, and ca. 1 ST was identified per 1.4 isolates. Although it is not possible to compare directly the MLST data we obtained with U.S. strains and the data obtained with the Spanish strains studied by Salcedo et al. (33), because different loci were analyzed during the two studies, our collection appeared to be slightly less heterogeneous. For example, in the study of Salcedo et al., the largest number of alleles was identified by the analyses of *cat* (ca. 1 allele per 3.9 strains), *dapE* (ca. 1 allele per 4.4 strains), and *abcZ* (ca. 1 allele per 5.2 strains). In our study, the largest number of alleles was identified by the analyses of *actA* (ca. 1 allele per 6.5 strains). Among the housekeeping genes we analyzed, the largest number of alleles was identified for *betL* (ca. 1 allele per 9.2 strains), followed by *gyrB* (ca. 1 allele per 10.5 strains) (Table 1). These observations should be taken into account when one is selecting loci for MLST of *L. monocytogenes*.

The *Listeria* genome has a mosaic structure and contains several putative DNA uptake genes and prophages (13). Therefore, gene transfer by transformation has been proposed (13) to be primarily responsible for genomic differences between *L. monocytogenes* and other *Listeria* species (e.g., *Listeria innocua*), as well as some other bacteria closely related to *L. monocytogenes* (e.g., *Bacillus subtilis*). Also, although the mean G+C content of the recently sequenced *L. monocytogenes* strain EGD-e (serotype 1/2a) was estimated to be ca. 39%, several regions with different G+C contents have been identified on the chromosome, which suggests recent (on the evolutionary scale) horizontal gene acquisition(s) of those regions (13). The parallelograms calculated during our Splitstree analysis (Fig. 3) also support the idea that recombinational events played a role in the evolution of *L. monocytogenes.*

The G+C contents of the genes we analyzed (Table 1) were relatively consistent with the mean G+C content reported for the fully sequenced *L. monocytogenes* strain (13). Interestingly, *hlyA* and *actA* (two major virulence genes located in close proximity on the *L. monocytogenes* virulence gene cluster (23) had different G+C contents (37 and 40%, respectively). This observation is in agreement with the findings of Glaser et al. (13), in which several regions with different G+C contents were found in the genome of the fully sequenced *L. monocytogenes* strain, and it suggests that *hlyA* and *actA* genes from different ancestors may have been introduced into the bacterium's virulence gene cluster. Our observation of no correlation between *hlyA* and *actA* alleles also supports this hypothesis.

**Clustering of clinical strains by MLST.** *L. monocytogenes* virulence has traditionally been associated with the presence of several virulence genes, including *prfA*, *plcA*, *hlyA*, *mpl*, *actA*, and *plcB*, grouped in the *Listeria* virulence gene cluster (23). However, not all *L. monocytogenes* strains carry or express all of the above genes, and some strains have been found (30) to be as much as 100-fold more virulent in vivo than other strains. Several investigators (18, 32, 38) have suggested, based on ribotyping and allelic analyses of several virulence genes, that "clinical" (and presumably highly pathogenic) *L. monocytogenes* strains are of clonal origin and that they are grouped in distinct lineages that have various pathogenic potentials. During our PFGE analyses, we did not observe subclustering of clinical (and possibly more virulent) or environmental isolates. This observation agrees with a previous report (22) demonstrating no association between various PFGE types of *L. monocytogenes* and the strains' virulence potential for Caco-2 tissue culture cells. On the other hand, MLST identified three subclusters (A1, B1, and B2 [Fig. 2]) containing predominantly clinical isolates. Moreover, several *actA* and *hlyA* alleles in the subclusters (as well as among all of the 175 *L. monocytogenes* strains we analyzed) appeared to be predominantly associated with clinical isolates. At present, it is not clear whether the observed subclustering—or the presence of specific *hlyA* and *actA* alleles—is associated with a high pathogenic potential of the strains. Sequencing of additional loci (including loci from the *Listeria* virulence gene cluster) may identify additional clusters and the critical nucleotide substitutions responsible for clustering of *L. monocytogenes* strains that have high pathogenic potential. Since *L. monocytogenes* is acquired by humans primarily through consumption of contaminated foods (10, 37), we believe it is unlikely that a perfect division between "environmental" (including food) and "clinical" *L. monocytogenes* isolates can be identified. However, it may be possible to identify clusters containing a statistically significant prevalence of "clinical" versus "environmental" isolates. Additional studies in this area may have a profound impact on improving our understanding of the virulence traits of *L. monocytogenes* and on developing tools for rapid identification of *L. monocytogenes* strains of increased public health concern, i.e., strains with a high pathogenic potential.

**Correlation between the genetic backgrounds of strains and their serotypes.** PFGE- and ribotype-based genetic clustering of *L. monocytogenes* strains has been reported (3, 29) to correlate with their serotypes, and genotypic data generated by PFGE have been found (3) to be directly related to strain serotypes. Although we did not observe a similar clear-cut association of PFGE types with specific serotypes, most of the strains that had identical PFGE types and STs belonged to a single serotype. However, some genetically identical or very closely related *L. monocytogenes* strains with the same ST and PFGE type, e.g., strains 84 and 57 (ST121, P16, *hlyA* allele 10, *actA* allele 17), belonged to different serotypes (1/2b and 3b, respectively). This observation raises the possibility that the strains diverged into different serotypes, from the same ancestor, via an "antigen-switching" mechanism. Similar mechanisms are known to exist in some other bacterial pathogens; for example, strains of the recently emerged epidemic O139 serotype of *Vibrio cholerae* have been shown (2) to have diverged from a classical O1 strain(s) via O-antigen switching. The possibility that similar antigen switching may have occurred in *L. monocytogenes* suggests that conclusions about the pathogenic potential of *L. monocytogenes* strains (based solely on their serotypes) must be interpreted with caution and that strains belonging to "nonpathogenic" *L. monocytogenes* serotypes (i.e., serotypes other than 1/2a, 1/2b, and 4b) should not be considered to be avirulent or to have reduced virulence solely on the basis of their serotype classification. Our observation also suggests that conclusions about the genetic relatedness of strains based on their serotypes may be misleading and that genetic typing methodologies (e.g., MLST) must be used to determine the true genetic background of various *L. monocytogenes* isolates and the relationships among them.

**Discriminatory abilities of serotyping, PFGE, and MLST.** The *L. monocytogenes* strains we analyzed were grouped in 8 serogroups, 57 PFGE types, and 122 STs, which suggests that the discriminatory abilities of the three methodologies differ, i.e., that serotyping is the least discriminatory and MLST is the most discriminatory of the three approaches. Serotyping is one of the oldest approaches for typing *L. monocytogenes*, and it has been used extensively to differentiate strains and to trace disease-causing isolates to their sources of origin (35). However, the discriminatory power of serotyping is limited, and PFGE typing has been reported (3) to be superior for differentiating *L. monocytogenes* strains. Our data are in agreement with this observation.

Our observations that the number of STs generated by MLST was ca. 2-fold greater than the number of PFGE types (122 STs versus 57 PFGE types) and that several strains within the same PFGE type were differentiated by MLST support the idea that the discriminatory ability of MLST is greater than that of PFGE. In order to compare further the discriminatory abilities of MLST and PFGE, we determined the average linkage distances among the strains clustered in distinct PFGE types. In other words, all strains of randomly selected PFGE types were located on the "MLST tree" (Fig. 2), and the average linkage distances among the strains were calculated. If the linkage distances for all members of each of the PFGE groups were small (e.g., if the smallest linkage distance between any two "minimally different" isolates on the START-generated tree was 0.25 [Fig. 2]), PFGE would be considered to describe adequately the genetic variation in these strains. However, the linkage distances were found to be greater than 0.25 (e.g., the linkage distance between *L. monocytogenes* strains 35 and 97 [grouped in the same PFGE type, P1] was ca. 0.9), which further confirmed our conclusion that MLST's dis-

criminatory ability for subtyping *L. monocytogenes* is greater than that of PFGE.

Our finding that MLST had greater discriminatory ability than PFGE was not entirely unexpected. MLST detects all genetic variations within the amplified gene fragment, whereas PFGE only examines the variations that are in the cleavage sites for a particular restriction enzyme. Thus, one would expect MLST to detect more variations in the bacterial genome than PFGE and therefore to be more discriminatory. However, although this assumption has been shown to be true for some bacterial pathogens (21), one recent study (31) suggested that MLST was less discriminatory than PFGE in typing *Escherichia coli* O157:H7 isolates. That observation may be explained by the fact that *E. coli* O157:H7 is a highly clonal serotype. Thus, it is possible that critical strain-differentiating nucleotide substitutions were localized outside the seven loci analyzed by MLST during that study (31)—as a result of, for example, insertion or deletion of nucleotide sequences in strains that otherwise have identical or very similar genetic backgrounds. PFGE has the advantage of randomly "probing" the entire genome, whereas MLST only analyzes nucleotides within targeted genes. We observed a few examples of this phenomenon during our study, when strains of the same ST were differentiated into two or more PFGE types. For example, strains 70, 28, and 101 belonged to ST99, but they were separated into three distinct, albeit very closely related, PFGE types (P54, P53, and P17) (Fig. 2).

Notwithstanding the few exceptions noted above, we found that MLST utilizing loci from four housekeeping genes was more discriminatory than PFGE for typing of *L. monocytogenes.* Including additional loci in the analysis is likely to further increase the discriminatory power of MLST. In this context, although there are at present no established criteria for determining the minimal number of genes that must be analyzed to obtain a reliable MLST assignment, and although the number is likely to be different for different species (depending on the level of clonality of the species), MLST utilizing seven housekeeping genes has been proposed (6) to be a reasonable approach well-suited for most bacteria. Cai et al. (5) recently provided more insight into the rational design of DNA sequence-based strategies for *L. monocytogenes*.

We found that PFGE occasionally differentiated strains undistinguishable by MLST. This observation, together with the recent results obtained with *E. coli* O157:H7 (31), suggests that a side-by-side comparison of the two methodologies is warranted for each bacterial species before one makes a decision concerning which method has the superior discriminatory ability. This approach may be particularly relevant for highly clonal bacterial species because of the highly conserved nature of the housekeeping genes often analyzed by MLST.

The superior discriminatory ability of MLST, compared to that of PFGE and serotyping, may have important practical implications. PFGE (alone or in combination with serotyping) is currently the method of choice for investigating food-borne outbreaks of listeriosis and for tracing the outbreak-causing strain to the source of contamination, information which has important legal and financial ramifications. However, our data demonstrate that some *L. monocytogenes* strains undistinguishable by PFGE are not necessarily the same isolates and that, in some cases, they are not even closely related. Therefore, MLST may be of value for improving the differentiation of *L. monocytogenes* strains isolated during food-borne outbreaks of listeriosis and for tracing the outbreak-causing strains to their sources. Also, MLST may be used to determine the phylogenetic relatedness among *L. monocytogenes* strains, and its use should improve our understanding of the mechanisms involved in the emergence and divergence of various *L. monocytogenes* strains and serotypes, including serotypes primarily associated with human listeriosis.

## ACKNOWLEDGMENTS

We thank Clay Silas and the personnel of the National Food Processors Association for providing some of the food and environmental isolates of *L. monocytogenes* used in this study.

This study was supported by NRI/CSREES grant 2002-35201-11704 from the U.S. Department of Agriculture.

## REFERENCES

1. **Aarnisalo, K., T. Autio, A. M. Sjoberg, J. Lunden, H. Korkeala, and M. L. Suihko.** 2003. Typing of *Listeria monocytogenes* isolates originating from the food processing industry with automated ribotyping and pulsed-field gel electrophoresis. J. Food Prot. **66:**249–255.
2. **Bik, E. M., A. E. Bunschoten, R. D. Gouw, and F. R. Mooi.** 1995. Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. EMBO J. **14:**209–216.
3. **Brosch, R., J. Chen, and J. B. Luchansky.** 1994. Pulsed-field fingerprinting of listeriae: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. Appl. Environ. Microbiol. **60:**2584–2592.
4. **Byun, S. K., S. C. Jung, and H. S. Yoo.** 2001. Random amplification of polymorphic DNA typing of *Listeria monocytogenes* isolated from meat. Int. J. Food Microbiol. **69:**227–235.
5. **Cai, S., D. Y. Kabuki, A. Y. Kuaye, T. G. Cargioli, M. S. Chung, R. Nielsen, and M. Wiedmann.** 2002. Rational design of DNA sequence-based strategies for subtyping *Listeria monocytogenes.* J. Clin. Microbiol. **40:**3319–3325.
6. **Enright, M. C., and B. G. Spratt.** 1999. Multilocus sequence typing. Trends Microbiol. **7:**482–487.
7. **Ericsson, H., H. Unnerstad, J. G. Mattsson, M. L. Danielsson-Tham, and W. Tham.** 2000. Molecular grouping of *Listeria monocytogenes* based on the sequence of the *inIB* gene. J. Med. Microbiol. **49:**73–80.
8. **Ewing, B., and P. Greene.** 1998. Base-calling of automated sequencer traces using *Phred.* I. Accuracy assessment. Genome Res. **8:**175–185.
9. **Ewing, B., and P. Greene.** 1998. Base-calling of automated sequencer traces using *Phred.* II. Error probabilities. Genome Res. **8:**186–194.
10. **Farber, J. M., and P. I. Peterkin.** 1991. *Listeria monocytogenes*, a food-borne pathogen. Microbiol. Rev. **55:**476–511.
11. **Fleming, D. W., S. L. Cochi, K. L. MacDonald, J. Brondum, P. S. Hayes, B. D. Plikaytis, M. B. Holmes, A. Audurier, C. V. Broome, and A. L. Reingold.** 1985. Pasteurized milk as a vehicle of infection in an outbreak of listeriosis. N. Engl. J. Med. **7:**404–407.
12. **Gendel, S. M., and J. Ulaszek.** 2000. Ribotype analysis of strain distribution in *Listeria monocytogenes.* J. Food Prot. **63:**179–185.
13. **Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couve, A. de Daruvar, P. Dehoux, E. Domann, G. Dominguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K. D. Entian, H. Fsihi, F. G. Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gomez-Lopez, T. Hain, J. Hauf, D. Jackson, L. M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueno, A. Maitournam, J. M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J. C. Perez-Diaz, R. Purcell, B. Remmel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J. A. Vazquez-Boland, H. Voss, J. Wehland, and P. Cossart.** 2001. Comparative genomics of *Listeria* species. Science **294:**849–852.
14. **Graves, L. M., B. Swaminathan, M. W. Reeves, S. B. Hunter, R. E. Weaver, B. D. Plikaytis, and A. Schuchat.** 1994. Comparison of ribotyping and multilocus enzyme electrophoresis for subtyping of *Listeria monocytogenes* isolates. J. Clin. Microbiol. **32:**2936–2943.
15. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55–62.
16. **Holmes, E. C., R. Urwin, and M. C. Maiden.** 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis.* Mol. Biol. Evol. **16:**741–749.
17. **Huson, D. H.** 1998. SplitsTree: analysing and visualizing evolutionary data. Bioinformatics **14:**68–73.
18. **Jaradat, Z. W., G. E. Schutze, and A. K. Bhunia.** 2002. Genetic homogeneity among *Listeria monocytogenes* strains from infected patients and meat prod-

ucts from two geographic locations determined by phenotyping, ribotyping and PCR analysis of virulence genes. Int. J. Food Microbiol. **76:**1–10.
19. **Jeanmougin, F., J. D. Thompson, M. Gouy, D. G. Higgins, and T. J. Gibson.** 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. **23:**403–405.
20. **Kerouanton, A., A. Brisabois, E. Denoyer, F. Dilasser, J. Grout, G. Salvat, and B. Picard.** 1998. Comparison of five typing methods for the epidemiological study of *Listeria monocytogenes.* Int. J. Food Microbiol. **43:**61–71.
21. **Kotetishvili, M., O. C. Stine, A. Kreger, J. G. Morris, Jr., and A. Sulakvelidze.** 2002. Multilocus sequence typing for characterization of clinical and environmental *Salmonella* strains. J. Clin. Microbiol. **40:**1626–1635.
22. **Larsen, C. N., B. Norrung, H. M. Sommer, and M. Jakobsen.** 2002. In vitro and in vivo invasiveness of different pulsed-field gel electrophoresis types of *Listeria monocytogenes.* Appl. Environ. Microbiol. **68:**5698–5703.
23. **Leimeister-Wachter, M., C. Haffner, E. Domann, W. Goebel, and T. Chakraborty.** 1990. Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *Listeria monocytogenes.* Proc. Natl. Acad. Sci. USA **87:**8336–8340.
24. **Linnan, M. J., L. Mascola, X. D. Lou, V. Goulet, S. May, C. Salminen, D. W. Hird, M. L. Yonekura, P. Hayes, R. Weaver, et al.** 1988. Epidemic listeriosis associated with Mexican-style cheese. N. Engl. J. Med. **13:**823–828.
25. **Louie, M., P. Jayaratne, I. Luchsinger, J. Devenish, J. Yao, W. Schlech, and A. Simor.** 1996. Comparison of ribotyping, arbitrarily primed PCR, and pulsed-field gel electrophoresis for molecular typing of *Listeria monocytogenes.* J. Clin. Microbiol. **34:**15–19.
26. **Lyytikainen, O., T. Autio, R. Maijala, P. Ruutu, T. Honkanen-Buzalski, M. Miettinen, M. Hatakka, J. Mikkola, V. J. Anttila, T. Johansson, L. Rantala, T. Aalto, H. Korkeala, and A. Siitonen.** 2000. An outbreak of *Listeria monocytogenes* serotype 3a infections from butter in Finland. J. Infect. Dis. **181:** 1838–1841.
27. **Maiden, M. C. J., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA **95:**3140–3145.
28. **Mead, P. S., L. Slutsker, V. Dietz, L. F. McCaig, J. S. Breeze, G. Shapiro, P. M. Griffin, and R. V. Tauxe.** 1999. Food-related illness and death in the United States. Emerg. Infect. Dis. **5:**607–625.
29. **Nadon, C. A., D. L. Woodward, C. Young, F. G. Rodgers, and M. Wiedmann.** 2001. Correlations between molecular subtyping and serotyping of *Listeria monocytogenes.* J. Clin. Microbiol. **39:**2704–2707.
30. **Nishibori, T., K. Cooray, H. Xiong, I. Kawamura, M. Fujita, and M. Mitsuyama.** 1995. Correlation between the presence of virulence-associated genes as determined by PCR and actual virulence to mice in various strains of *Listeria* spp. Microbiol. Immunol. **5:**343–349.
31. **Noller, A. C., M. C. McEllistrem, O. C. Stine, J. G. Morris, Jr., D. J. Boxrud, B. Dixon, and L. H. Harrison.** 2003. Multilocus sequence typing reveals a lack of diversity among *Escherichia coli* O157:H7 isolates that are distinct by pulsed-field gel electrophoresis. J. Clin. Microbiol. **41:**675–679.
32. **Norrung, B., and N. Skovgaard.** 1993. Application of multilocus enzyme electrophoresis in studies of the epidemiology of *Listeria monocytogenes* in Denmark. Appl. Environ. Microbiol. **59:**2817–2822.
33. **Salcedo, C., L. Arreaza, B. Alcala, L. De La Fuente, and J. A. Vazquez.** 2003. Development of a multilocus sequence typing method for analysis of *Listeria monocytogenes* clones. J. Clin. Microbiol. **41:**757–762.
34. **Schlech, W. F., P. M. Lavigne, R. A. Bortolussi, A. C. Allen, E. V. Haldane, A. J. Wort, A. W. Hightower, S. E. Johnson, S. H. King, E. S. Nicholls, and C. V. Broome.** 1983. Epidemic listeriosis—evidence for transmission by food. N. Engl. J. Med. **4:**203–206.
35. **Seeliger, H. P. R., and K. Hohne.** 1979. Serotyping of *Listeria monocytogenes* and related species. Methods Microbiol. **13:**31–49.
36. **Spratt, B. G.** 1999. Multilocus sequence typing: molecular typing of bacterial pathogens in an era of rapid DNA sequencing and the Internet. Curr. Opin. Microbiol. **2:**312–316.
37. **Swaminathan, B.** 2001. *Listeria monocytogenes*, p. 383–409. *In* M. P. Doyle, L. R. Beuchat, and T. J. Montville (ed.), Food microbiology: fundamentals and frontiers, 2nd ed. American Society for Microbiology, Washington, D.C.
38. **Wiedmann, M., J. L. Bruce, C. Keating, A. E. Johnson, P. L. McDonough, and C. A. Batt.** 1997. Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. Infect. Immun. **65:**2707–2716.