Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 5234–5238, May 1995
Microbiology

Types of *Listeria monocytogenes* predicted by the positions of *Eco*RI cleavage sites relative to ribosomal RNA sequences

(bacteria/ribotyping/classification/identification)

ROMEO J. HUBNER, EILEEN M. COLE, JAMES L. BRUCE, CHANNEARY I. MCDOWELL, AND JOHN A. WEBSTER*

Central Research and Development, E. I. du Pont de Nemours and Company, Wilmington, DE 19880

Communicated by H. E. Simmons, DuPont Central Research and Development, Wilmington, DE, February 6, 1995

ABSTRACT **By using taxonomic characters derived from *Eco*RI restriction endonuclease digestion of genomic DNA and hybridization with a labeled rRNA operon from *Escherichia coli*, a polymorphic structure of *Listeria monocytogenes*, characterized by fragments with different frequencies of occurrence, was observed. This structure was expanded by creating predicted patterns through a recursive process of observation, expectation, prediction, and assessment of completeness. This process was applied, in turn, to normalized strain patterns, fragment bands, and positions of *Eco*RI recognition sites relative to rRNA regions. Analysis of 1346 strains provided observed patterns, fragment sizes, and their frequencies of occurrence in the patterns. Fragment size statistics led to the creation of unobserved combinations of bands, predicted pattern types. The observed fragment bands revealed positions of *Eco*RI sites relative to rRNA sequences. Each *Eco*RI site had a frequency of occurrence, and unobserved fragment sizes were postulated on the basis of knowing the restriction site locations. The result of the recursion process applied to the components of the strain data was an extended classification with observed and predicted members.**

Classification is the arrangement of strains into taxonomic groups on the basis of observed similarities. Bacteria have been classified into genera, species, and types with a variety of phenotypic characteristics to provide a basis for identification (1). Patterns of DNA restriction fragments containing portions of the rRNA operons provide another means of description and classification (2, 3).

A bacterial genome contains numerous restriction enzyme recognition sites within and flanking the sequences that are highly conserved in related strains. By considering the mutational gains or losses of these sites as statistically independent events, we hypothesized that a species taxonomic structure incorporating all possible strain variation could consequently be defined. The conserved sequences and regional restriction sites inferred from our work in describing *Listeria monocytogenes* using *Eco*RI fragments containing sequences homologous to a rRNA operon from *Escherichia coli* (4) formed the basis of our analysis. Polymorphic fragments from different rRNA regions, each containing a given part of a given operon, could be combined into patterns, some of which would remain unobserved until the sample set became large enough to be truly representative of the natural population.

By using matrix analysis on the data, the observed polymorphisms of the different rRNA regions were combined into patterns that have not yet been observed. The polymorphisms were also used to suggest the positions of *Eco*RI sites relative to an rRNA region. A maximum-likelihood model developed for use in this context predicted the pairings of restriction sites that led to the observed fragments and suggested pairings that could form additional sizes of rRNA sequence-containing fragments. Through a recursive process of examining the observed patterns, the fragment bands, and the derived *Eco*RI sites, expected patterns and their computed frequencies of occurrence were added to the observed characteristics of the species. The inclusion of these predicted types in the species definition increases the probability that new isolates will be recognized as belonging to one of the established, named types.

MATERIALS AND METHODS

Bacterial Strains. The basis for this analysis was a study of 1346 strains of *L. monocytogenes*, acquired from diverse sources and described elsewhere (4). The 13 strains that produced patterns influenced by incubation temperature were not included in the statistical analysis of polymorphic fragments and restriction sites.

DNA Fragments. Strains were described by patterns of DNA-fragment bands that resulted from a single experimental method, using *Eco*RI and an rRNA operon from *E. coli*, which has been described in detail (4, 5). In the context of this paper, the term fragment is limited to *Eco*RI fragments containing rRNA sequences, those hybridizing with the rRNA information-containing probe.

Band Patterns. Computerized procedures developed by the authors were used to extract data from each lane in the images and normalize the positions and intensities of fragment bands in the pattern of each strain. The fragment sizes were derived from their electrophoretic mobility by using the following expression (6): $(L - L_0)(M - M_0) = K_0$, where L is the fragment length in bases and M is the migration distance in pixels. A reference scale was created by analysis of standards in >100 original images to determine the values of the subscripted symbols. M_0 was set to 0; L_0 was 3085 bases, and the K_0 was 1762139 pixel-bases. The lane patterns in every image were remapped to the reference scale.

As reported previously, when identical patterns were obtained from more than one strain, those patterns were averaged, and the average was stored as a pattern type. The observed strain variation was resolved into 50 pattern types, 34 representing multiple strains and 16 single strains. Each fragment of a given size in the *L. monocytogenes* patterns was considered to be a taxonomic character, and its frequency of occurrence in all patterns from *L. monocytogenes* strains was determined. Pattern type dd 0566 [including strain DD 0566 (ATCC 15313)] was composed of the most frequently occurring fragment sizes and was designated as the base type for comparison of suggested polymorphic fragments. The bands were assigned a letter sequentially according to their relative position in the base type. The pattern types were arranged into coherent subsets with a single fragment variable in size and the remaining fragments constant in size at 100% frequency of occurrence within the subsets (4).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*To whom reprint requests should be addressed.

RESULTS

Polymorphic Sets. The observed frequencies for each of the fragment sizes in the various polymorphic sets are shown in Table 1. The accuracy of the fragment sizes was assessed by analysis of the following two sources of computational error. The variance of the computed size of a given fragment across all pattern types was ≈1%. The sizes computed by the size-mobility function derived from known standards had an offset error varying regionally from −1.5% to +1.5% in the range of 1–12 kbp. Thus, the error in the computed sizes could be as much as 2.5%.

All combinations of the independently sorting polymorphic fragments, one from each polymorphic set, may potentially represent strains. The combinatorial of matrices produced from the polymorphic sets in Table 1 was analyzed, and the population in each combination was computed. The EG two-dimensional matrix was found to contain the largest population, accounting for 87% of the sample set. Table 2 contains the strain frequencies for patterns in all combinations of the variable G fragments and the E-fragment variants, corresponding to figure 1 in ref. 4; the other fragments remained fixed at their most frequently occurring sizes.

A full eight-dimensional matrix (one dimension for each fragment, A–H) served as a DNA-based definition and classification of the species according to *Eco*RI restriction sites and extended the classification to predicted pattern types. It was apparent from the EG matrix in Table 2 that unobserved combinations of the observed fragment sizes could be extant. Established at the time our sample set contained 225 strains of *L. monocytogenes*, this classification method was tested to verify the occurrence of matrix-predicted patterns. Although every variation may not be readily observed, several of the expected patterns have, in fact, been observed. Patterns dd 5633 in the E 11.2, G 4.5 matrix position, dd 7730 in the E 9.2, G 5.2 position, and dd 7696 in the E 9.2, G 5.8 position exemplified such occurrences. Additionally, new E and G polymorphisms have since been observed. For example, the new G-fragment size first seen in dd 5204 in the E 5.2, G 9.1 position in Table 2 allowed the prediction of an entire new row, including dd 7674 and dd 5558, and the new E fragment size first seen in dd 6296 in the E 5.3, G 3.8 position allowed the prediction of an entire new column, including dd 6481 and dd 6362. The eight-dimensional matrix predicts a total of 4080 pattern types based on observed fragment sizes. This extended classification was tested for specificity by comparing these patterns with those from 550 strains representing all other *Listeria* species. No other species shared the sets of fragments (4) found in the observed and predicted pattern types in the classification of *L. monocytogenes* (data not shown).

Table 1. Polymorphic sets of *Eco*RI fragments of *L. monocytogenes* hybridized with a labeled plasmid containing the *rrnB* rRNA operon of *E. coli*

Region	Size, kbp	Observation, %	Region	Size, kbp	Observation, %
C	2.3	0.2	G	3.2	4.8
	4.0	99.8		3.8	6.0
D	5.0	98.9		4.0	0.2
	9.2	1.1		4.5	0.4
E	5.2	44.6		5.2	1.3
	5.3	0.8		5.8	1.8
	5.8	0.2		6.2	42.0
	9.2	37.4		6.5	0.5
	11.2	17.0		7.4	0.3
F	3.6	0.8		7.6	0.1
	4.6	9.1		8.1	41.3
	4.8	0.1		9.1	0.4
	5.5	89.3		9.7	0.1
	7.2	0.8		10.3	0.2
	12.9	0.1		11.1	0.5
				11.7	0.2
				12.4	0.1
			H	7.1	1.2
				9.0	98.8

The fragment sizes were computed from pattern data and used in a maximum-likelihood algorithmic procedure to estimate positions of restriction sites.

The expected frequency of occurrence (P) of a pattern was computed as the product of the observed frequencies (p) of its component fragment bands:

$$P = \prod_{\alpha=a}^{h} p_\alpha = p_e p_g \prod_{\alpha \neq e\ \alpha \neq g} p_\alpha = 0.867\, p_e p_g. \quad [1]$$

For the EG combinations found in Table 2, the implicitly expressed product ($\alpha \neq e$, $\alpha \neq g$) assumes its maximum value of 0.867.

The observed frequency of the dd 0566 pattern type was 19.3%, and the expected frequency was computed as 16.4%. An assessment of completeness was computed as the sum of the expected frequencies of occurrence of the observed pattern types and was found to be 84%. The sampling error in the data was clearly revealed in the differences between the observed and expected frequencies. When the sampling error was minimized through an iterative procedure to correct for the number of strains represented by each pattern type, in the mathematical limit, the result was an equivalent sample set of 969 strains with the observed types accounting for 95% of the expectation.

Restriction Sites. Fragments A and B were constant in size throughout *L. monocytogenes*, as well as throughout the genus. Other sets of polymorphic fragments were either of low (C, D, and H rRNA regions), medium (E and F regions), or high (G region) number of size variants. Fragments in sets with low and medium number of size variants were considered to have one stable *Eco*RI recognition site, apparently within the rRNA operon, and to have the other flanking site variable in relative position with two to six variants. The high diversity of the G set of fragments could not be explained with variable recognition sites on only one side of the hybridizable region and was considered to have detectable mutations on both flanking regions.

Although mutations are inherently statistically independent random events, our observations suggest that the positions of the observed *Eco*RI sites are not randomized, remaining at preferred locations relative to an rRNA region within the genomes of strains of the species. In this context, the double mutation distinguishing a strain of dd 0566 (E 5.2, G 6.2) from a strain of dd 0653 (E 9.2, G 8.1) in Table 2 (and also figure 1 in ref. 4) minimally includes the intermediary mutation state defined by either the dd 1049 or the dd 1151 pattern type.

Possible G fragments are represented in Fig. 1, labeled according to the relative positions of the flanking restriction sites. Fragment 00 was the shortest one observed. Analysis of the size differences between fragments enables determination of the DNA distance between *Eco*RI sites. For example, the difference in length of fragment 00 and fragment 01 is equal to the difference of fragment 10 and fragment 11. This difference corresponds to the distance between the R_0 and R_1 *Eco*RI sites. Similarly, the difference in length of fragment 00 and 10 is equal to that between fragment 01 and fragment 11 and is the distance between L_0 and L_1.

Analysis of the observed frequencies of different fragment sizes derived from a specific hybridizable region enables further insight. The probability of occurrence of a fragment (for example, 00) is the product of the probabilities of

Table 2. Observed and expected frequencies of occurrence for observed and predicted combinations of fragment sizes containing the E and G rRNA regions of *L. monocytogenes* DNA

G-region size (kbp)	Frequency	Combination frequency, % (PT) E-region size (kbp) 5.2	5.3	5.8	9.2	11.2
3.2	Observed	4.80 (dd 1067)	0	0	0	0
	Expected	1.87	0.03	—	1.57	0.71
3.8	Observed	3.75 (dd 1966)	0.15 (dd 6296)	0	0.90 (dd 1288)	1.05 (dd 0678)
	Expected	2.33	0.04	—	1.96	0.89
4.0	Observed	0.08 (dd 6439)	0	0	0.08 (dd 7745)	0
	Expected	0.06	—	—	0.05	0.02
4.5	Observed	0.15 (dd 3391)	0	0	0.15 (dd 3579)	0.08 (dd 5633)
	Expected	0.15	—	—	0.12	0.06
5.2	Observed	0.83 (dd 1070)	0.23 (dd 6481)	0	0.23 (dd 7730)	0
	Expected	0.50	0.01	—	0.42	0.19
5.8	Observed	0.23 (dd 3232)	0	0	1.20 (dd 7696)	0
	Expected	0.70	0.01	—	0.59	0.27
6.2	Observed	19.28 (dd 0566)	0.38 (dd 6362)	0	2.63 (dd 1049)	10.13 (dd 0647)
	Expected	16.34	0.27	0.05	13.70	6.23
6.5	Observed	0.23 (dd 3549)	0	0	0.30 (dd 3392)	0
	Expected	0.20	—	—	0.17	0.08
7.4	Observed	0.30 (dd 3295)	0	0	0	0
	Expected	0.12	—	—	0.10	0.04
7.6	Observed	0	0	0	0.08 (dd 5449)	0
	Expected	0.03	—	—	0.02	0.01
8.1	Observed	2.33 (dd 1151)	0	0	31.58 (dd 0653)	5.63 (dd 1962)
	Expected	16.08	0.27	0.05	13.48	6.13
9.1	Observed	0.23 (dd 5204)	0	0	0.08 (dd 7674)	0.08 (dd 5558)
	Expected	0.15	—	—	0.12	0.06
9.7	Observed	0.08 (dd 3408)	0	0	0	0
	Expected	0.03	—	—	0.02	0.01
10.3	Observed	0.08 (dd 3287)	0	0	0	0.08 (dd 3607)
	Expected	0.06	—	—	0.05	0.02
11.1	Observed	0.45 (dd 1148)	0	0	0	0
	Expected	0.18	—	—	0.15	0.07
11.7	Observed	0	0	0	0.23 (dd 6301)	0
	Expected	0.09	—	—	0.07	0.03
12.4	Observed	0.08 (dd 6449)	0	0	0	0
	Expected	0.03	—	—	0.02	0.01

Other fragments' sizes match dd 0566. Observed combinations also list the corresponding pattern type (PT) number. The following conventions are used in this table: A dash (—) indicates a finite percentage <0.01%. A zero indicates that the corresponding combination has never been observed.

observation of the two corresponding *Eco*RI sites. Therefore, the probability of occurrence (P) of a specific fragment size containing the isolated hybridizable region can be calculated as follows, where the lower-case l and r indicate

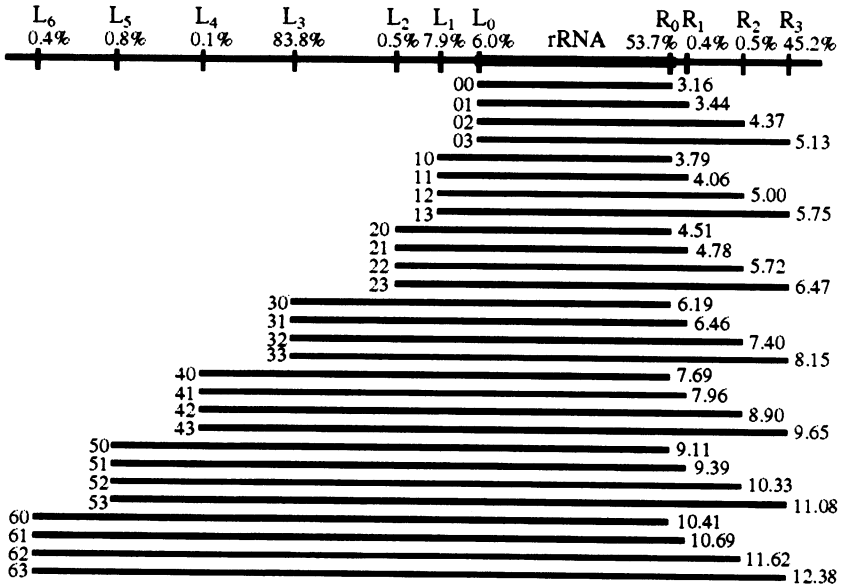

FIG. 1. Representation of the G rRNA region of DNA from *L. monocytogenes* strains and fragments resulting from its *Eco*RI digestion. The *Eco*RI sites are labeled L or R, depending on their position relative to the rRNA region. The percentages associated with each site indicate the observed frequency of cleavage at that site. Resulting fragments are represented by bars and have been grouped based on the left *Eco*RI site. On the left of the fragment bars are labels indicating the active *Eco*RI sites on the left and right of the hybridizable region. On the right of the bars, the model-computed fragment lengths in kbp are listed.

the probabilities of the conservation of the *Eco*RI site indicated by the subscript:

$$P_{00} = l_0 r_0 \qquad P_{10} = l_1(1 - l_0)r_0$$

$$P_{01} = l_0(1 - r_0)r_1 \qquad P_{11} = l_1(1 - l_0)(1 - r_0)r_1,$$

and, in general,

$$P_{mn} = l_m \prod_{i=0}^{m-1}(1 - l_i) r_n \prod_{j=0}^{n-1}(1 - r_j). \qquad [2]$$

Note that the following probability relationship is also true.

$$\frac{P_{00}}{P_{10}} = \frac{l_0 r_0}{l_1(1 - l_0)r_0} = \frac{l_0(1 - r_0)r_1}{l_1(1 - l_0)(1 - r_0)r_1} = \frac{P_{01}}{P_{11}}. \qquad [3]$$

Using these fundamental relationships, the information about the fragments can be represented in two matrices, a fragment-size matrix and a probability matrix. In the fragment-size matrix, the length difference between any two elements in the same column (row) in two designated rows (columns) is the same and corresponds to the size of the fragment between *Eco*RI sites. In the probability matrix, the ratio of probabilities in corresponding rows (columns) is constant.

These fundamental relationships were implemented in a maximum-likelihood computer procedure to build a matrix for each hybridizable region by inserting frequency of observation and fragment-size data in the appropriate row–column positions and computing values to fill in additional spaces in the matrix. Only the G polymorphic set produced a two-dimensional matrix, shown in Table 3, corresponding to the restriction sites shown in Fig. 1. The probability of activation of a potential restriction site was computed as the sum of the frequencies f_n of the observed fragments, starting or ending at the given restriction site. For example, the probability (P) for a specific left-side (L) restriction site (m) is given as follows:

$$P_{L_m} = l_m \prod_{i=0}^{m-1}(1 - l_i)$$

$$= \sum_{n=0}^{N_R} l_m \prod_{i=0}^{m-1}(1 - l_i) r_n \prod_{j=0}^{n-1}(1 - r_j) = \sum_{n=0}^{N_R} f_n, \qquad [4]$$

where N_R is the number of restriction sites observed on the right side of the rRNA hybridizable region.

The six G-fragment sizes, accounting for 96% of the observed population (Table 1), resulted from the combinatorial activation of five *Eco*RI sites, R_0 (54%) and R_3 (45%) on the right and L_0 (6%), L_1 (8%), and L_3 (84%) on the left (Table 3). It is apparent from Table 3 that there are unobserved combinations of restriction sites allowing the prediction of bands and, therefore, pattern types not yet observed.

When all available data were included, these fragment-size and probability matrices served as an extended DNA-based definition of a species that include observed and unobserved fragment sizes in observed and unobserved combinations with a computable expected frequency of occurrence.

Table 3. Matrix representation of fragments derived from the *Eco*RI recognition sites flanking the G rRNA region

"Left" *Eco*RI site	Parameter	"Right" *Eco*RI sites R0	R1	R2	R3
L0	Size, kbp	3.17	3.44	4.38	5.13
	Observed, %	4.80	0	0	1.27
	Expected, %	3.26	0.03	0.03	2.74
L1	Size, kbp	3.79	4.07	5.00	5.76
	Observed, %	6.00	0.15	0	1.80
	Expected, %	4.27	0.04	0.04	3.59
L2	Size, kbp	4.51	4.79	5.72	6.48*
	Observed, %	0.37	0	0	0.22
	Expected, %	0.32	—	—	0.27
L3	Size, kbp	6.19	6.47*	7.40	8.16
	Observed, %	41.98	0.30	0.30	41.30
	Expected, %	45.09	0.38	0.44	37.92
L4	Size, kbp	7.69	7.97	8.90	9.66
	Observed, %	0.07	0	0	0.07
	Expected, %	0.08	—	—	0.07
L5	Size, kbp	9.12	9.39	10.33	11.09
	Observed, %	0.37	0	0	0.45
	Expected, %	0.44	—	—	0.37
L6	Size, kbp	10.42	10.69	11.63	12.38
	Observed, %	0.15	0	0.22	0.07
	Expected, %	0.24	—	—	0.20

The fragment sizes in this table and in Fig. 1 were computed from the sizes in Table 1 using a maximum-likelihood algorithm. The following conventions are used in this table: A dash (—) indicates a finite percentage less than 0.01%. A zero indicates that the corresponding combination has never been observed.

*These two fragments of similar size derived from different restriction sites were assigned a proportional part of the observed frequency (0.5%) of the 6.5-kbp fragment (G region in Table 1).

DISCUSSION

Strain Variation. Observed strain variation, as a function of *Eco*RI sites within and immediately surrounding the rRNA operons, was 50 pattern types. Derived from the observed pattern data, the probability of a next strain being of a new type is not >1 in 1346 + 1. Analysis of the observed sizes of the individual fragments, independent of the patterns, permitted the observations shown in Table 1. From this analysis and the arrangement of the data into matrices, prediction of unobserved combinations—patterns with one member from each polymorphic set—and estimation of frequencies of occurrence became possible. Comparison of the observed and expected frequencies indicated sampling error in the strains of certain pattern types, such as dd 0653, ATCC 19115 (E 9.2, G 8.1 in Table 2). When the unobserved combinations of bands, the predicted patterns, were added to the species description, the classification incorporated the total strain variation into 4080 types. The probability of observation of a new pattern type, including the most likely new fragment polymorphism (in the G region), was then computed at 1 in 3447. This classification also established a system of nomenclature for the types: for example, dd 1067 could be reported as E 5.2, G 3.2 (Table 2). Other region fragment sizes would be reported only if they differ from the most frequently observed values.

Predicted Bands. The species description was extended beyond unobserved combinations of observed fragment-size polymorphisms to include unobserved combinations of unobserved polymorphic fragment sizes. The observed polymorphisms of the E and F regions (Table 1) can be accounted for by mutations on just one side of the hybridizable region. The diversity of the G region is more easily explained by mutations on both sides of the hybridizable region (Tables 1 and 3). The model developed to predict sizes and frequencies of occurrence was demonstrated by its retrospective placement of all the observed G fragments. Also, retrospective visual inspection of the G bands in several types revealed G-size variants in the six strains included in G 11.1 and warranted size recomputation. In the outcome, three strains had G-fragment sizes of 11.10 ± 0.08 kbp, two strains had sizes of 10.68 ± 0.07 kbp, and one strain had a G fragment size of 11.56 kbp (fragments 53, 61, and 62, in Fig. 1). Although a conservative approach was taken considering the experimental error and the patterns were originally merged into a single pattern type (4), review of closely spaced bands prompted by Table 3, such as in the

patterns of the G 11.1 group, suggested the model's validity. Table 3 predicts 28 G-fragment sizes that could be used to re-compute the matrices and would account for 6720 types. The probability of observation of a new pattern type including the most likely new *Eco*RI site, in the G region, was computed as 1 in 4106 [the probability of the most likely new restriction site (0.838 in 1346 + 1) times the highest observed frequencies of the remaining fragments], which is less than the probability of observing a new E-fragment polymorphism, 1 in 3660 [the probability of a new E fragment (1 in 1346 + 1) times the highest observed frequencies of the remaining fragments].

***L. monocytogenes* rRNA Genes.** Band A appeared to have higher intensity relative to B or C, suggesting multiple A fragments (analog data not shown). All fragments, except G, appeared to have a highly stable *Eco*RI recognition site on one side of the fragment, most likely within the operon. The number of observed G-fragment sizes suggested the combinatorial of four to seven *Eco*RI sites on each side of the hybridizable region. One side would then have a number of *Eco*RI sites comparable to the number observed for the E and F regions, whereas the other side would show the effects of a higher mutation rate. The size of the smallest observed G fragment was ≈3 kbp, which is less than the size of a complete operon, and the expected second labeled fragment, the complement to this smallest size, could not be observed. When the smallest fragment sizes in each set were considered, the pairing of one fragment (from A or B) with one fragment (from C, D, E, F, or H) could create five genomic regions approximately the size of an rRNA operon. In this arrangement, the G region was left without a complementary fragment; however, its homology to rRNA was supported by data produced using transcribed rRNA from *E. coli* as probe. This result would indicate that the *L. monocytogenes* genome has five complete rRNA operons, in agreement with refs. 7 and 10, and that the G region is the sixth rRNA homologous region reported in ref. 8. Perhaps the genomic region surrounding the smallest G fragment lost the constraint of functionality and became rapidly evolving, similar to a pseudo gene (9) as has been frequently observed in eukaryotes, and ultimately not sufficiently homologous with rRNA sequences to be detected.

Through the recursive process described, a newly observed fragment can lead to a newly described *Eco*RI site, which can predict a new set of bands, which can predict a new set of patterns. The resulting species description is a theoretical projection to a total of 6720 types based on 50 observed pattern types. The expected frequency of occurrence of the least likely predicted pattern type is 1 in 10^{18} strains. This description can never rule out the possibility of additional pattern types based on newly observed fragment size data created by *Eco*RI sites not previously observed. The most likely new pattern type, which would include a new E polymorphism, has an expected frequency of occurrence not greater than 1 in 3660 strains. The theoretical extension of the species classification increases the probability of recognizing an unknown, with a previously unseen pattern type, as belonging to one of the established, named taxa.

1. Seeliger, H. P. R. & Jones, D. (1986) in *Bergey's Manual of Systematic Bacteriology*, eds. Sneath, P. H. A., Mair, H. S., Sharpe, M. E. & Holt, J. G. (Williams & Wilkins, Baltimore), Vol. 2, pp. 1235–1245.
2. Webster, J. A. (1983) Eur. Patent Appl. 82,305,061.2.
3. Grimont, F. & Grimont, P. A. D. (1986) *Ann. Inst. Pasteur/ Microbiol.* **137,** 165–175.
4. Bruce, J. L., Hubner, R. J., Cole, E. M., McDowell, C. I. & Webster, J. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5229–5233.
5. Webster, J. A., Bannerman, T. L., Hubner, R. J., Ballard, D. N., Cole, E. M., Bruce, J. L., Fiedler, F., Schubert, K. & Kloos, W. E. (1994) *Int. J. Syst. Bacteriol.* **44,** 454–460.
6. Southern, E. M. (1979) *Anal. Biochem.* **100,** 319–323.
7. Carriere, C., Allardet-Servant, A., Bourg, G., Audurier, A. & Ramuz, M. (1991) *J. Clin. Microbiol.* **29,** 1351–1355.
8. Michel, E. & Cossart, P. (1992) *J. Bacteriol.* **174,** 7098–7103.
9. Miyata, T., Yasunaga, T. & Nishida, T. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 7328–7332.
10. Thompson, D. E., Balsdon, J. T., Cai, J. & Collins, M. D. (1992) *FEMS Microbiol. Lett.* **96,** 219–224.