JOURNAL OF BACTERIOLOGY, Mar. 1987, p. 1352–1357
0021-9193/87/031352-06$02.00/0
Vol. 169, No. 3

Evolutionary Origin of Pathogenic Determinants in Enterotoxigenic *Escherichia coli* and *Vibrio cholerae* O1

TATSUO YAMAMOTO,[1]* TAKASHI GOJOBORI,[2] AND TAKESHI YOKOTA[1]

Department of Bacteriology, School of Medicine, Juntendo University, Bunkyo-ku, Tokyo, Japan,[1] and National Institute of Genetics, Mishima, Shizuoka, Japan[2]

Received 14 October 1986/Accepted 12 December 1986

Three families of the evolutionarily related pathogenic determinants in enterotoxigenic *Escherichia coli* and *Vibrio cholerae* O1, a family of cholera enterotoxin (CT) and heat-labile enterotoxin (LT) including CT, LTh, and LTp, a family of heat-stable enterotoxin I (STI) including STIa and STIb, and a family of K88 enteroadhesion fimbriae including K88ab, K88ac, and K88ad were analyzed for synonymous (silent) nucleotide substitutions by using the gene nucleotide sequences of earlier reports and the LTp gene nucleotide sequence presented in this paper. The data suggested that the divergences between LT and CT and between STIa and STIb occurred in the remote past, whereas those betwen LTh and LTp and between members of the K88 family occurred very recently. We concluded that the LT gene is a foreign gene that has been acquired by *E. coli* to form an enteropathogen. This provides evolutionary evidence of species-to-species transfer of pathogenic determinants in procaryotes.

Enterotoxigenic *Escherichia coli* causes choleralike diarrhea in humans and food animals. It colonizes the intestinal epithelium with fimbrial adhesins and produces enterotoxins such as heat-labile toxin (LT) and heat-stable toxin I (STI) (46). Human-colonizing enterotoxigenic *E. coli* produces fimbriae named (e.g.) CFA, and swine-colonizing enterotoxigenic *E. coli* produces fimbriae named (e.g.) K88 (11, 30). The K88 family includes K88ab, K88ac, and K88ad having different epitopes. LT has structural and functional features in common with cholera enterotoxin (CT) produced from *Vibrio cholerae* O1, a causative agent of cholera, and is subdivided into two (LTp and LTh) having different epitopes; LTp is found only in porcine isolates of enterotoxigenic *E. coli*, whereas LTh is found only in human isolates (6, 13, 14, 18, 45, 55). STI also has two members, one (STIa) being produced from both porcine and human isolates of enterotoxigenic *E. coli* and another (STIb) being produced only from human isolates (1, 3, 42, 48, 52). The virulence genes for LT, STI, and K88 of enterotoxigenic *E. coli* are carried by plasmids, whereas the gene for CT of *V. cholerae* O1 is on the chromosome (38). Nucleotide sequence knowledge has been accumulated in the three multigene families: the CT-LT gene family (8, 33, 34, 39, 49, 60, 61), the STI gene family (42, 48), and the K88 gene family (10, 12). However, little is understood of their evolutionary relatedness. In this paper, we describe the corrected, complete nucleotide sequence of the LTp subunit A (LTp-A) gene and, by analyzing synonymous nucleotide substitutions of the gene sequences, we discuss the divergence in evolution of the bacterial pathogenic determinants.

The genes for LTp-A and LTp-B, originating from an enterotoxin plasmid (ENT plasmid), P307 (47), and an ENT-R plasmid, pCG86 (17) which codes for antibiotic resistance as well, were sequenced by the chemical method of Maxam and Gilbert (36) as previously described (60). In these experiments, recombinant plasmids EWD299 (carrying the LTp genes of ENT plasmid P307 [9, 47]) and pPMC2 (carrying the LTp genes of ENT-R plasmid pCG86 [4]) were used as DNA sources. The LTp-A genes on P307 and pCG86 were identical in sequence (Fig. 1). The LTp-B genes on P307 (8, 32) and pCG86 (Fig. 1) were also identical in sequence. It was also clear (Fig. 1) that the differences between the three subunit A gene sequences of LTp, LTh, and CT are the result of a series of single base changes. The present sequence of the LTp-A gene (Fig. 1) disagrees with earlier sequence data (49) made on the LTp-A gene of P307, with which significant deletion and insertion of nucleotides had been pointed out (34, 39, 60). The amino acid sequence of the LTp-A gene product (precursor to LTp-A), deduced from the present nucleotide sequence, is also shown in Fig. 1. When an amino acid sequence comparison was made between the LTp-A1 and LTh-A1 fragments (ADP-ribosyltransferase; a proteolytic cleavage product of subunit A which plays a key role in the toxin function [position 1 to 192 in Fig. 1 A gene region]), LTp-A1 differed from LTh-A1 by only one amino acid at position 4 (Arg↔Lys). In addition, in terms of a local hydrophilicity profile (19) and a predicted secondary structure (5), LTp-A1 was identical to LTh-A1 which had been previously described (58, 60).

A promoter region sequence of the LTp operon was also determined (Fig. 1). This sequence was similar to the corresponding sequence of the LTh operon but markedly different from that of the CT operon (Fig. 1). Tandem repetitions of the unique sequence TTTTGAT located 77 base pairs upstream of the start of the CT-A gene, which play a role in the expression of the CT operon (39), were not found in the promoter region of the LTp or LTh operon (Fig. 1).

Using the sequence data in Fig. 1, as well as data from earlier reports, we estimated the number of synonymous (silent) nucleotide substitutions (40) between the homologous genes in the virulence gene families of CT-LT, STI, and K88 in an attempt to determine their evolutionary relationships (Table 1). Synonymous substitutions were the most significant between the LT and CT genes and between the

* Corresponding author.

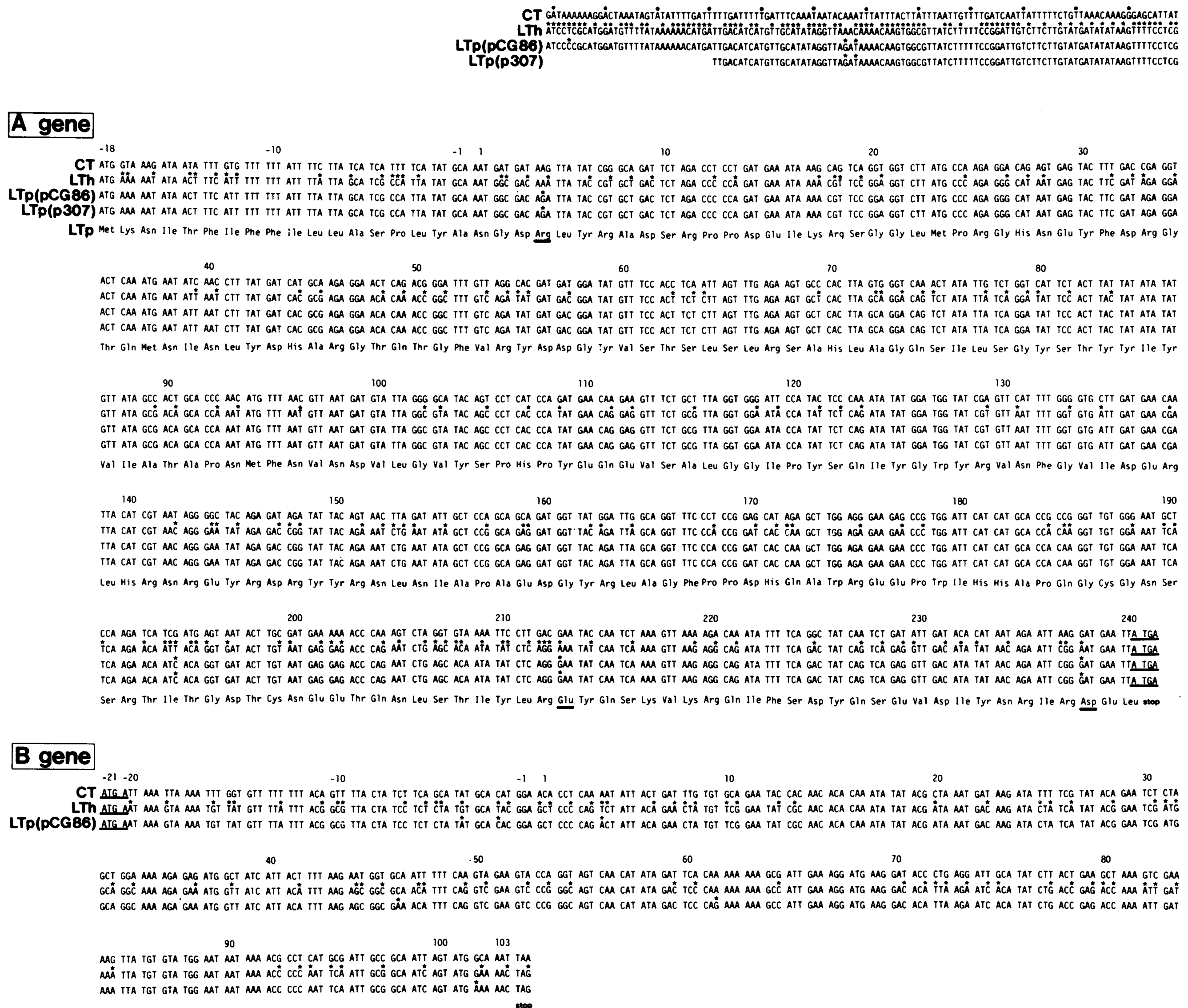

FIG. 1. DNA sequences of the LTp-A and LTp-B genes compared with those of the LTh and CT genes. The sequence data of the LTh and CT genes are from references 60 and 61 and references 33 and 39, respectively. The two underlined nucleotides (ATGA) of the A and B genes overlap. The LTp and CT gene nucleotides that differ from the nucleotides of the LTh genes are marked with asterisks. In the A gene sequences, the numbers −1 to −18 and 1 to 240 represent the coding sequence regions for a signal peptide (18 amino acids) and for subunit A (240 amino acids), respectively. The previously reported codon sequence, GAT, at position 189 of the LTh-A gene (60) was a misreading, and the corrected sequence, AAT (position 189), is indicated. The predicted amino acid sequence of the LTp-A gene product is shown below the LTp nucleotide sequences; underlined amino acids differ from those of LTh (60). In the B gene sequences, the numbers −1 to −21 and 1 to 103 represent the coding regions for a signal peptide and for subunit B, respectively. Nucleotide sequences upstream from the A gene are shown at the top.

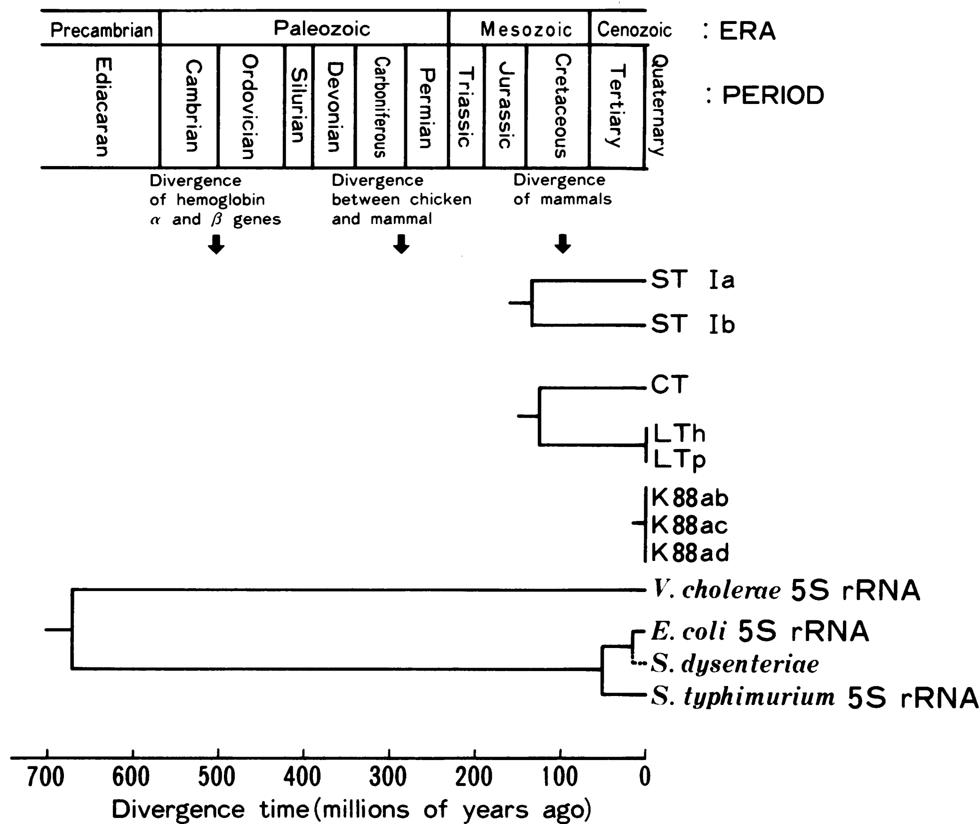

FIG. 2. Divergence in evolution of the pathogenic determinants and the 5S rRNA sequences of the bacterial pathogens causing diarrheal diseases. The evolutionary distance between two bacterial species was calculated based on the nucleotide substitutions of 5S rRNA sequences as described by Kimura and Ohta (29). The proportion of different nucleotides between the two 5S rRNA sequences compared (*E. coli* 5S rRNA and *V. cholerae* O1 5S rRNA [35]) was corrected for multiple substitutions as described in the footnote to Table 1. Evolutionary time was estimated by using the corrected number (K') obtained and $(1.8 \pm 0.5) \times 10^{-10}$ per nucleotide site per year as the evolutionary rate of nucleotide substitution for 5S rRNA sequences (20, 21). The divergence between *E. coli* 5S rRNA and *Salmonella typhimurium* 5S rRNA and that between *E. coli* ribosomal proteins and *Shigella dysenteriae* ribosomal proteins (drawn as a dotted line in the figure) are from references 20 and 23 and reference 22, respectively.

STIa and STIb genes, markedly low between the LTp and LTh genes, and also very low or even rare among the K88ab, K88ac, and K88ad genes.

Since the evolutionary rate of synonymous substitution for different genes is known to be roughly uniform (41), we estimated the divergence time for each pair of the homologous virulence genes (Fig. 2). The STIa and STIb genes, as well as the CT and LT genes, seemed to have diverged in the remote past, ~130 million years ago. In contrast, the divergence between the LTp and LTh genes seemed to have been very recent, ~0.9 million years ago. During the Cretaceous period, to which time the divergence between the LT and CT genes corresponds, mammalian forms with a placenta seem to have arisen (57). Moreover, the European wild pig, *Sus scrofa*, seems to have appeared in the earlier Pleistocene (the Quaternary period), ~0.011 to 3 million years ago (7); this period of time roughly overlaps the predicted time of divergence of the LTp and LTh genes. From this viewpoint, we speculated that the divergence of the toxin genes may reflect the divergence of animal hosts of the toxin-producing bacterial parasites. Occurrence of divergence among members of the K88 gene family may overlap or be much more recent than divergence of the LTp and LTh genes (Table 1; Fig. 2). It is believed that *V. cholerae* O1 produces diseases only in humans (humans are natural reservoir hosts of *V. cholerae* O1). Therefore, it is extremely interesting that CT, a major virulence factor of cholera, was already present before the appearance of humans.

The time of divergence of the *E. coli* and *V. cholerae* O1 5S rRNAs was estimated by using the reported sequences (35) as described by Kimura and Ohta (29); the number of nucleotide substitutions (K' [Fig. 2, legend]) was 0.2421 ± 0.0519. The estimated divergence time, ~670 million years ago, was apparently much earlier than that between the LT and CT genes (Fig. 2).

We therefore concluded that the LT gene is a foreign gene which was acquired by *E. coli* to form an enteropathogen in the recent evolutionary past compared with the origin of the two bacterial species *E. coli* and *V. cholerae*. It presumably originated in the *V. cholerae* CT gene ancestor. This conclusion is consistent with the finding that, with respect to codon usage, the correlation coefficient between the LTh genes (A and B) and the *E. coli* gene mixture (Table 2; 16) was only 0.0646 (this correlation may be obtained with a probability of 0.62 by chance), whereas the correlation coefficient between the LTh genes (A and B) and the CT genes (A and B) was 0.6313 (this correlation may be obtained with a probability of less than 10^{-6} by chance alone; see Table 2 for each subunit gene). In agreement with this, the frequency of usage of *E. coli* optimal codons (24, 25) was very low (Table 2) compared with that of other *E. coli* genes (24, 25), indicating that, as far as codon usage is concerned,

TABLE 1. The number of synonymous substitutions between evolutionarily related virulence genes[a]

Gene family	Homologous virulence genes	Synonymous substitution number (K)
CT-LT	CT-A, LTh-A	1.1604 ± 0.1728
	CT-A, LTp-A	1.1578 ± 0.1721
	LTh-A, LTp-A	0.0055 ± 0.0055
	CT-B, LTh-B	1.3246 ± 0.3116
	CT-B, LTp-B	1.4243 ± 0.3539
	LTh-B, LTp-B	0.0122 ± 0.0122
STI	STIa, STIb	1.3504 ± 0.4437
K88	K88ab, K88ad	0.039 ± 0.015
	K88ab, K88ad	0.000 ± 0.000
	K88ab, K88ac	0.000 ± 0.000
	K88ac, K88ad	0.009 ± 0.007

[a] The nucleotide sequence data of the LTp, LTh, and CT genes are those in Fig. 1. The sequence data of the STIa and STIb genes (each 216 base pairs long) are from references 48 and 42, respectively. See reference 12 for the first K88ab-K88ad comparison listed and reference 10 for the remaining K88 pairs; the K88ab and K88ad gene sequences are 855 base pairs long, and the reported K88ac gene sequence is 849 base pairs long. The number of synonymous (silent) nucleotide substitutions was calculated by the method of Miyata and Yasunaga (40) as follows. First, the proportion of synonymous substitution (P) was obtained by the formula P = number of synonymous substitutions observed (n_s)/number of synonymous sites in the gene (N_s). Then, we corrected P for multiple substitutions by the formula (26, 28) K (average number of synonymous substitutions per site) = $-3/4 \ln (1-4/3P)$, assuming that the rates of nucleotide substitutions among nucleotides are equal; we also used other methods (15, 27, 51), in which the rates of nucleotide substitutions among nucleotides were unequal, but the values of evolutionary distances did not change much. Finally, we estimated divergence time (T) from the formula $T = K/2v$, where v is the evolutionary rate of synonymous substitution. Here, we used $(5.1 \pm 0.3) \times 10^{-9}$ per nucleotide site per year as v (41). The data obtained are shown in Fig. 2.

the LT gene is not fully efficient in the *E. coli* translational system. These data on the LT and CT genes provide strong support for a species-to-species transfer event in the evolution of bacterial pathogenic determinants. It has been shown that the LTp gene can be expressed in *V. cholerae* O1 (43). The concept of species-to-species transfer of genes in evolution has been described as a cross-species gene transfer (50), a horizontal gene transfer (2), or a nondivergence theory of evolution (56).

Similarly, based on the codon usage data summarized in Table 2, we concluded that the STI (a or b) gene, as well as the STII gene (a virulence gene of porcine isolates) of enterotoxigenic *E. coli*, is a foreign gene which was acquired in the past. The STI gene may have originated in bacterial species such as *Yersinia enterocolitica* or *V. cholerae* non-O1 which produce an STI-like enterotoxin (53, 54); the gene sequences have not been determined. More interestingly, it has been demonstrated that STI has extensive amino acid sequence homology with conotoxins of the marine snail *Conus geographus* (42).

In marked contrast, K88 genes could be more efficient in the *E. coli* translational system, because the frequency of usage of *E. coli* optimal codons, as well as the correlation coefficient of codon usage, compared with the *E. coli* gene mixture was extremely high (Table 2). It was concluded that the K88 genes are evolutionarily closer to *E. coli* than are the enterotoxin genes.

The ENT plasmid P307 and the ENT-R plasmid pCG86 have been shown to share a common enterotoxin-coding region as evidenced by heteroduplex analysis by electron microscope (37). The present study showed that the LTp genes on P307 and pCG86 are identical in sequence (there are no single base changes), suggesting that occurrence of

TABLE 2. Summary of genetic characteristics of the virulence genes

Gene family and gene	G+C content (mol%)	Frequency of usage of *E. coli* optimal codons[a]	Correlation coefficient (probability by chance) of codon usage when compared with:		Source (reference) of gene sequences
			E. coli gene mixture[b]	Other genes	
CT-LT					
LTh A	38.2	0.38[c]	0.034 (0.790)	0.724 ($<10^{-5}$)[d]	60; this study
LTp A	38.6	0.38	0.046 (0.724)		This study
CT A	38.5	0.41[c]	0.020 (0.880)		39
LTh B	37.1	0.44[c]	0.209 (0.105)	0.701 ($<10^{-5}$)[e]	61
LTp B	36.8	0.44	0.219 (0.090)		8, 32
CT B	32.8	0.40[c]	0.178 (0.170)		33
STI					
STIa	31.0	0.50[f]	0.131 (0.315)	0.809 ($<10^{-5}$)[g]	48
STIb	31.9	0.50	0.105 (0.421)		42
STII					
STII	36.2	0.46	0.223 (0.085)	0.489 ($<10^{-4}$)[g]	31, 44
K88					
K88ab	45.1	0.59	0.348 (0.006)		12
K88ab	45.1	0.61	0.355 (0.005)		10
K88ac	45.1	0.59	0.334 (0.008)		10
K88ad	46.4	0.57	0.379 (0.003)		12
K88ad	44.6	0.60	0.373 (0.003)		10

[a] See references 24 and 25.
[b] Genes for *lacI*, *lacY*, *trpA*, *recA*, lipoprotein, and a part of RNA polymerase (16).
[c] Taken from reference 59.
[d] CT-A.
[e] CT-B.
[f] Taken from reference 24.
[g] STIb.

the ENT-R plasmid (from the ENT plasmid) is a very recent event—apparently in the more recent evolutionary past compared with the divergence of the LTp and LTh genes and probably since clinical use of antibiotics by humans, which must have acted as a very strong selective agent.

We thank Werner K. Maas for the LTp coding plasmids and helpful suggestions; Takaaki Tamura for partial LTp gene sequencing; Etsuko Nakamura for synonymous substitution analysis in part; Hiroshi Hori, Takeru Akazawa, and Nobuo Tamiya for stimulating discussions; and Yoshifumi Takeda for continuous support and interest.

This work was supported by a grant-in-aid for special project research and a grant-in-aid for scientific research from the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

1. **Aimoto, S., T. Takao, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani.** 1982. Amino-acid sequence of a heat-stable enterotoxin produced by human enterotoxigenic *Escherichia coli*. Eur. J. Biochem. **129**:257–263.
2. **Busslinger, M., S. Rusconi, and M. L. Birnstiel.** 1982. An unusual evolutionary behaviour of a sea urchin histone gene cluster. EMBO J. **1**:27–33.
3. **Chan, S.-K., and R. A. Giannella.** 1981. Amino acid sequence of heat-stable enterotoxin produced by *Escherichia coli* pathogenic for man. J. Biol. Chem. **256**:7744–7746.
4. **Chen, T.-M., A. J. Mazaitis, and W. K. Maas.** 1985. Construction of a conjugative plasmid with potential use in vaccines against heat-labile enterotoxin. Infect. Immun. **47**:5–10.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
6. **Clements, J. D., D. C. Flint, and F. A. Klipstein.** 1982. Immunological and physicochemical characterization of heat-labile enterotoxins isolated from two strains of *Escherichia coli*. Infect. Immun. **38**:806–809.
7. **Cornwall, I. W.** 1968. Prehistoric animals and their hunters. Faber & Faber, London.
8. **Dallas, W. S., and S. Falkow.** 1980. Amino acid sequence homology between cholera toxin and *Escherichia coli* heat-labile toxin. Nature (London) **288**:499–501.
9. **Dallas, W. S., D. M. Gill, and S. Falkow.** 1979. Cistrons encoding *Escherichia coli* heat-labile toxin. J. Bacteriol. **139**:850–858.
10. **Dykes, C. W., I. J. Halliday, M. J. Read, A. N. Hobden, and S. Harford.** 1985. Nucleotide sequences of four variants of the K88 gene of porcine enterotoxigenic *Escherichia coli*. Infect. Immun. **50**:279–283.
11. **Evans, D. G., R. P. Silver, D. J. Evans, Jr., D. G. Chase, and S. L. Gorbach.** 1975. Plasmid-controlled colonization factor associated with virulence in *Escherichia coli* enterotoxigenic for humans. Infect. Immun. **12**:656–667.
12. **Gaastra, W., P. Klemm, and F. K. de Graaf.** 1983. The nucleotide sequence of the K88ad protein subunit of porcine enterotoxigenic *Escherichia coli*. FEMS Microbiol. Lett. **18**:177–183.
13. **Geary, S. J., B. A. Marchlewicz, and R. A. Finkelstein.** 1982. Comparison of heat-labile enterotoxins from porcine and human strains of *Escherichia coli*. Infect. Immun. **36**:215–220.
14. **Gill, D. M., J. D. Clements, D. C. Robertson, and R. A. Finkelstein.** 1981. Subunit number and arrangement in *Escherichia coli* heat-labile enterotoxin. Infect. Immun. **33**:677–682.
15. **Gojobori, T., K. Ishii, and M. Nei.** 1982. Estimation of average number of nucleotide substitutions when the rate of substitution varies with nucleotide. J. Mol. Evol. **18**:414–423.
16. **Greene, P. J., M. Gupta, H. W. Boyer, W. E. Brown, and J. M. Rosenberg.** 1981. Sequence analysis of the DNA encoding the *Eco* RI endonuclease and methylase. J. Biol. Chem. **256**: 2143–2153.
17. **Gyles, C. L., S. Palchaudhuri, and W. K. Maas.** 1977. Naturally occurring plasmid carrying genes for enterotoxin production and drug resistance. Science **198**:198–199.
18. **Honda, T., T. Tsuji, Y. Takeda, and T. Miwatani.** 1981. Immunological nonidentity of heat-labile enterotoxins from human and porcine enterotoxigenic *Escherichia coli*. Infect. Immun. **34**:337–340.
19. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3824–3828.
20. **Hori, H.** 1976. Molecular evolution of 5S RNA. Mol. Gen. Genet. **145**:119–123.
21. **Hori, H., K. Higo, and S. Osawa.** 1977. Molecular evolution of ribosomal components, p. 240–260. *In* M. Kimura (ed.), Molecular evolution and polymorphism: Proceedings of the Second Taniguchi International Symposium on Biophysics. National Institute of Genetics, Mishima, Japan.
22. **Hori, H., and S. Osawa.** 1978. Evolution of ribosomal proteins in *Enterobacteriaceae*. J. Bacteriol. **133**:1089–1095.
23. **Hori, H., and S. Osawa.** 1979. Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. Proc. Natl. Acad. Sci. USA **76**:381–385.
24. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J. Mol. Biol. **151**:389–409.
25. **Ikemura, T., and H. Ozeki.** 1983. Codon usage and transfer RNA contents: organism-specific codon-choice patterns in reference to the isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. **47**:1087–1097.
26. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism III. Academic Press, Inc., New York.
27. **Kimura, M.** 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. USA **78**:454–458.
28. **Kimura, M., and T. Ohta.** 1972. On the stochastic model for estimation of mutational distance between homologous proteins. J. Mol. Evol. **2**:87–90.
29. **Kimura, M. and T. Ohta.** 1973. Eukaryotes-prokaryotes divergence estimated by 5S ribosomal RNA sequences. Nature (London) New Biol. **243**:199–200.
30. **Klemm, P.** 1985. Fimbrial adhesins of *Escherichia coli*. Rev. Infect. Dis. **7**:321–340.
31. **Lee, C. H., S. L. Moseley, H. W. Moon, S. C. Whipp, C. L. Gyles, and M. So.** 1983. Characterization of the gene encoding heat-stable toxin II and preliminary molecular epidemiological studies of enterotoxigenic *Escherichia coli* heat-stable toxin II producers. Infect. Immun. **42**:264–268.
32. **Leong, J., A. C. Vinal, and W. S. Dallas.** 1985. Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons from *Escherichia coli* of human and porcine origin. Infect. Immun. **48**:73–77.
33. **Lockman, H., and J. B. Kaper.** 1983. Nucleotide sequence analysis of the A2 and B subunits of *Vibrio cholerae* enterotoxin. J. Biol. Chem. **258**:13722–13726.
34. **Lockman, H. A., J. E. Galen, and J. B. Kaper.** 1984. *Vibrio cholerae* enterotoxin genes: nucleotide sequence analysis of DNA encoding ADP-ribosyltransferase. J. Bacteriol. **159**:1086–1089.
35. **MacDonell, M. T., and R. R. Colwell.** 1984. Identical 5S rRNA nucleotide sequence of *Vibrio cholerae* strains representing temporal, geographical, and ecological diversity. Appl. Environ. Microbiol. **48**:119–121.
36. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
37. **Mazaitis, A. J., R. Maas, and W. K. Maas.** 1981. Structure of a naturally occurring plasmid with genes for enterotoxin production and drug resistance. J. Bacteriol. **145**:97–105.
38. **Mekalanos, J. J.** 1983. Duplication and amplification of toxin genes in Vibrio cholerae. Cell **35**:253–263.
39. **Mekalanos, J. J., D. J. Swartz, G. D. N. Pearson, N. Harford, F. Groyne, and M. de Wilde.** 1983. Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development. Nature (London) **306**:551–557.
40. **Miyata, T., and T. Yasunaga.** 1980. Molecular evolution of

mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. J. Mol. Evol. **16**:23–36.

41. **Miyata, T., T. Yasunaga, and T. Nishida.** 1980. Nucleotide sequence divergence and functional constraint in mRNA evolution. Proc. Natl. Acad. Sci. USA **77**:7328–7332.
42. **Moseley, S. L., J. W. Hardy, M. I. Huq, P. Echeverria, and S. Falkow.** 1983. Isolation and nucleotide sequence determination of a gene encoding a heat-stable enterotoxin of *Escherichia coli*. Infect. Immun. **39**:1167–1174.
43. **Neill, R. J., B. E. Ivins, and R. K. Holmes.** 1983. Synthesis and secretion of the plasmid-coded heat-labile enterotoxin of *Escherichia coli* in *Vibrio cholerae*. Science **221**:289–291.
44. **Picken, R. N., A. J. Mazaitis, W. K. Maas, M. Rey, and H. Heyneker.** 1983. Nucleotide sequence of the gene for heat-stable enterotoxin II of *Escherichia coli*. Infect. Immun. **42**:269–275.
45. **Robertson, D. C., J. L. McDonel, and F. Dorner.** 1985. *E. coli* heat-labile enterotoxin. Pharmacol. Ther. **28**:303–339.
46. **Sack, R. B.** 1980. Enterotoxigenic *Escherichia coli*: identification and characterization. J. Infect. Dis. **142**:279–286.
47. **So, M., W. S. Dallas, and S. Falkow.** 1978. Characterization of an *Escherichia coli* plasmid encoding for synthesis of heat-labile toxin: molecular cloning of the toxin determinant. Infect. Immun. **21**:405–411.
48. **So, M., and B. J. McCarthy.** 1980. Nucleotide sequence of the bacterial transposon Tn1681 encoding a heat-stable (ST) toxin and its identification in enterotoxigenic *Escherichia coli* strains. Proc. Natl. Acad. Sci. USA **77**:4011–4015.
49. **Spicer, E. K., and J. A. Noble.** 1982. *Escherichia coli* heat-labile enterotoxin: nucleotide sequence of the A subunit gene. J. Biol. Chem. **257**:5716–5721.
50. **Syvanen, M.** 1986. Cross-species gene transfer: a major factor in evolution? Trends Genet. **2**:63–66.
51. **Takahata, N., and M. Kimura.** 1981. A model of evolutionary base substitutions and its application with special reference to rapid change of pseudogenes. Genetics **98**:641–657.
52. **Takao, T., T. Hitouji, S. Aimoto, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani.** 1983. Amino acid sequence of a heat-stable enterotoxin isolated from enterotoxigenic *Escherichia coli* strain 18D. FEBS Lett. **152**:1–5.
53. **Takao, T., Y. Shimonishi, M. Kobayashi, O. Nishimura, M. Arita, T. Takeda, T. Honda, and T. Miwatani.** 1985. Amino acid sequence of heat-stable enterotoxin produced by *Vibrio cholerae* non-O1. FEBS Lett. **193**:250–254.
54. **Takao, T., N. Tominaga, Y. Shimonishi, S. Hara, T. Inoue, and A. Miyama.** 1984. Primary structure of heat-stable enterotoxin produced by *Yersinia enterocolitica*. Biochem. Biophys. Res. Commun. **125**:845–851.
55. **Takeda, Y., T. Honda, H. Sima, T. Tsuji, and T. Miwatani.** 1983. Analysis of antigenic determinants in cholera enterotoxin and heat-labile enterotoxins from human and porcine enterotoxigenic *Escherichia coli*. Infect. Immun. **41**:50–53.
56. **Tamiya, N., and T. Yagi.** 1985. Non-divergence theory of evolution: sequence comparison of some proteins from snakes and bacteria. J. Biochem. **98**:289–303.
57. **Valentine, J. W.** 1978. The evolution of multicellular plants and animals. Sci. Am. **239**:104–117.
58. **Yamamoto, T., T. Nakazawa, T. Miyata, A. Kaji, and T. Yokota.** 1984. Evolution and structure of two ADP-ribosylation enterotoxins, *Escherichia coli* heat-labile toxin and cholera toxin. FEBS Lett. **169**:241–246.
59. **Yamamoto, T., A. Suyama, N. Mori, T. Yokota, and A. Wada.** 1985. Gene expression in the polycistronic operons of *Escherichia coli* heat-labile toxin: a new model of translational control. FEBS Lett. **181**:377–380.
60. **Yamamoto, T., T. Tamura, and T. Yokota.** 1984. Primary structure of heat-labile enterotoxin produced by *Escherichia coli* pathogenic for humans. J. Biol. Chem. **259**:5037–5044.
61. **Yamamoto, T., and T. Yokota.** 1983. Sequence of heat-labile enterotoxin of *Escherichia coli* pathogenic for humans. J. Bacteriol. **155**:728–733.