Characterization of unstable pEntYN10 from enterotoxigenic *Escherichia coli* (ETEC) O169:H41

Erika Ban1,2, Yuka Yoshida1,2, Mitsuko Wakushima1, Takeaki Wajima2, Takashi Hamabata3, Naoki Ichikawa1, Hiroyuki Abe4, Yasuhiko Horiguchi5, Yukiko Hara-Kudo5, Eriko Kage-Nakadai6, Taro Yamamoto7, Takayuki Wada7,* and Yoshikazu Nishikawa1

1Department of Food and Human Health Sciences; Graduate School of Human Life Science; Osaka City University; Osaka, Japan; 2Department of Microbiology; School of Pharmacy; Tokyo University of Pharmacy and Life Sciences; Tokyo, Japan; 3Research Institute; National Center for Global Health and Medicine; Tokyo, Japan; 4Department of Molecular Bacteriology; Research Institute for Microbial Diseases; Osaka University; Osaka, Japan; 5Division of Microbiology; National Institute of Health Sciences; Tokyo, Japan; 6The OCU Advanced Research Institute for Natural Science and Technology; Osaka City University; Osaka, Japan; 7Department of International Health; Institute of Tropical Medicine; Nagasaki University; Nagasaki, Japan

*These authors equally contributed to this work.

**Keywords:** cell adhesion, *Escherichia coli*, ETEC, genetic diversity, plasmid

---

**Introduction**

Each year, enterotoxigenic *Escherichia coli* (ETEC) cause nearly one billion cases of diarrheal diseases worldwide. The pathogens often pollute foods and drinking water and which cause endemic prevalence mainly in infants in developing countries and which infect travelers from developed countries. Infection gives rise to a severe health hazards in infected areas, resulting in numerous deaths among children.1,2 Etiologically, ETEC harbors one or more cell adhesion factors to colonize on epithelial cells of intestinal surfaces of hosts (colonization factors, CFs). The CFs comprise fimbrial or afimbrial structures on surfaces of bacilli, which facilitate binding to epithelial cells.3,4 Bacilli binding to host cells proliferate on the surface and secrete heat-stable enterotoxin (ST) and/or heat-labile enterotoxin (LT), which disturb the intestinal secretory state, thereby causing watery diarrhea of infected patients.5,6 Such virulence factors of ETEC are generally encoded on plasmids, which can drive the complicated horizontal evolution of pathogenicity of the species.7

Actually, ETEC strains have been defined by possession of their enterotoxin, although their genetic diversity is broad and various serotype strains have been determined as ETEC families from symptomatic patients or foodborne disease outbreaks.3 In spite of low variety in the definitive enterotoxins, CFs of ETEC strains represent wide variation, which explains their diverse antigenicity and structural characteristics.3 In fact, ETEC isolates generally have one or more previously identified CFs. Although adhesive structural molecules of more than 25 types have been identified as CFs, the variation of CFs is not regarded as being completely identified.4 Plasmids encoding CFs and enterotoxins also represent broad diversity including numerous insertion sequences (ISs) and transposons. Host bacilli losing or curing their plasmids lose their fimbrial structure on surface and adhesive activities to host cells generally. This result implies that the virulence of ETEC is closely associated with the retention of its plasmid.5,10

*E. coli* O169:H41 is an ST-producing ETEC isolates which causes foodborne diarrheic disease.11,12 It was found first as an etiological strain of foodborne case in Japan. Subsequently, its global dissemination has been reported from various countries.13-17 The serotype exhibited variation in ribotypes and outer membrane protein binding patterns, showing its diversity from different outbreaks.18 One unique characteristic of the strain was its
atypical adherence manner to epithelial cell lines such as HEp-2, which resembles that of enteroaggregative E. coli (EAggEC). It is considered that CFs and enterotoxin of the serotype strains are also encoded in plasmids because the phenotype disappears easily after its passage in liquid medium. This disappearance suggests that its virulence plasmid is unstable in vitro and that it is potentially cured out of host infection stage smoothly.

For the specific examination of its unique cell adhesion and unstable retention from the genetic perspective in this study, the complete sequence of a plasmid, pEntYN10, purified from the serotype strain YN10 was determined. Putative CDSs were annotated. In addition, CF-encoding genes were compared with previously reported sequences to assess their potential as virulence factors.

**Results**

**General structure of pEntYN10**

To ascertain the genetic features related to its instability and virulence of the host strain ETEC YN10, the plasmid pEntYN10 was purified directly; its sequence was determined. Using the next-generation sequencer FLX system, 17,583,280 bp of 50,934 reads (average length = 345.22 bp, GC content = 46.65%) were obtained. From the read data, 45 contigs longer than 500 bp (N50 = 12,474 bp) were assembled. They were finally combined by gap-filling analysis into single large circular sequences of 145,082 bp. The GC content was 46.15%. On the plasmid, 182 CDSs were estimated by annotation (a circular map is presented in Figure 1; gene list on Table S1). This plasmid codes repA1 and repA2 replication genes of incFII group, which means that it belongs to the RepFII plasmid family. The F-transfer pili genes are drastically truncated in the plasmid with the result that only genes traI and traX remain. This truncation of the transfer system was structurally similar to pETEC_80 (Table 1). Although its gene components are similar to the plasmid of the family sequenced previously, pEntYN10 is marked by its large size, high percentage of ISs, small number of stability genes, and various repertories of adhesion genes, as described below.

**Plasmid stability genes on pEntYN10**

Instability of pEntYN10 in vitro has been noticed in its passage process and has been discussed in the previous paper. Actually, it was observed that the plasmid was lost during passages more drastically than a referential ETEC strain H10407 along with in vitro passages.

**Potential virulent genes encoded on pEntYN10**

An outstanding feature of pEntYN10 was that 3 CFs, CS6, CS8 (CFA/III)-like, and K88 (F4)-like, coincided in the single plasmid. These three CFs preserved their respective gene structures and low GC contents as those of other ETEC strains reported previously (Fig. 3), with unique
| Host strain | ETEC YN10 | C921b-1 | H10407 | ETEC 1392/75 | E24377A | ETEC 4266 |
|-------------|-----------|---------|--------|-------------|---------|-----------|
| Name        | pEntYN10  | pCoo    | pETEC666 | pEntH10407 | p1018 | pETEC_80 |
| Accession no. | AP014654 | NC_007635 | FN649418 | NC_013507 | NC_009786 | NC_00974 |
| Size (bp)   | 145,082   | 98,396  | 94,797  | 66,681      | 67,094 | 101,857 |
| No. of predicted CDSs | 182 | 94 | 113 | 84 | 99 | 133 |
| Stability genes | psiAB, relE | stbAB, ccdAB | stbAB, psiAB | stbAB, psiAB | stbAB, psiAB | stbAB, psiAB |
| Colonization factors | CS6, CS8, K88 (F4) | CS1 | CFA/I | — | — | — |
| No. of IS or transposase (%) | 83 (45.6) | 18 (19.1) | 56 (49.6) | 20 (23.8) | 11 (11.1) | 50 (37.6) |

*Registered plasmid includes artificial kanamycin resistant gene *aphA.*
sequence variations (Table 2). Adhesion to HEp-2 cells with an aggregative pattern was unique to YN10 and disappeared in a derivative strain lacking the plasmid after passages (Fig. 4), which suggests that the plasmid correlated to the property of adhesion to host cells.

CS6 genes have high levels of sequence identity with previously reported genes (Table 2). CS6-coding genes cssA and cssB are classifiable respectively into 3 (A-I, A-II, and A-III) and 2 subtypes (B-I and B-II).28 The CS6 variant of pEntYN10 was found to be unique but similar to A-II/B-II by amino acid sequence homology (Fig. 5).

Similarity of amino acid sequences between CS8 on pEntYN10 and previously reported ETEC strains was lower than those of CS6 (Table 2). The CS8 components reportedly form a long rod-like fimbrial structure on the cell surface of ETEC, to attach specifically to human intestinal cells.29,30 The main structural protein is CofA, which possesses structural homology with the main fimbrial component of V. cholerae, TcpA.31 The amino acid sequence of CofA of pEntYN10 was similar to that of pSH1134, a plasmid of an ETEC strain 260-1 (Fig. 6A). Recently, 5 residues in this protein were reported as highly conserved to construct the filamentous structure.31,32 These residues of CofA in pEntYN10 were also preserved (Fig. 6B), which supports the functional conservation.

The K88 gene cluster also encodes fimbrial structure components.33 It is particularly interesting that the gene cluster of pEntYN10 had 2 paralogous copies of faeG, which encodes the major component of the K88 fimbrial structure (Fig. 3). Phylogenetic relation of FaeG sequences exhibited that these 2 paralogous genes, faeG1 and faeG2, were separated from other E. coli strains and they are more similar to that of Salmonella enterica serovar Infantis (Fig. 7). These genes were topologically close with long branches, which implies their deep genetic divergence.

**Discussion**

Recent developments in genomic analysis using next generation sequencers have yielded an extremely large number of draft sequences from widely diverse species. They were registered in databases, expanding the plentiful data resources. Such enormous amounts of information have driven a paradigm shift in various research fields, especially those of comparative genomics, genetic diversity, and molecular evolution.34 In spite of the powerful strategy, it is noteworthy that the current state of sequence techniques providing massive amounts of short reads has hindered the scrutiny of the structural arrangements of genomes, and has presented an obstacle to the specific examination of the genetic nature of virulent plasmids harboring complicated structures and active dynamics because of their many mobile elements. For instance, a virulent plasmid pCoo has an atypical structure integrated with 2 backbones of replication units (RepFIIA and RepI).21 Although such a unique structure can provide insights into the evolutionary dynamics of pathogenicity of ETEC, it is difficult to ascertain that structure from a simple accumulation of short read sequences. This study revealed the complete sequence of a plasmid of an ETEC strain, pEntYN10, which might help to estimate not only phenotypic features of the strain but also its evolutionary background.

A curious feature of pEntYN10 is its small number of stability-related genes (Table 1). The instability was concordantly verified in vitro, in contrast to stable retention of well-known ETEC strain H10407 plasmids (Fig. 2). Instability in vitro might provide benefits of an infectious habitat for the host bacilli. It is noteworthy that instability of pEntYN10 was observed only in vitro culture media. Indeed, plasmid-positive O169:H41 strains were actually isolated from clinical feces,18 perhaps because of the higher stability of the plasmid in vivo. Plasmid stability in vivo should be elucidated. Some putative stability-related genes found in pEntYN10 such as a relE family gene or resolvase genes might be a clue to approach it.

In addition to instability, results showed that the tra region of pEntYN10 is largely truncated, implying that the plasmid lacks a notorious strategy, the F-transfer system, to transfer itself horizontally to other bacilli because of a lack of an essential gene: traA. Similar truncation was reported in pETEC948 and pETEC_80 (Table 1), although that truncation was probably compensated by other coexisting plasmids in their host strains. O169 clinical strains reportedly possess no other coexisting plasmids,18 therefore, it remains to be seen whether the plasmid has the potential to be transferred horizontally in natural environments. Strong attachment to epithelial cells provided by the plasmid might enhance the survival fitness of host bacilli to compensate for the defects sufficiently as an independent replicon.

In our study, cell attachment pattern of ETEC strain YN10 in vitro was extraordinarily unique (Fig. 4). Functional

---

**Figure 2.** Stability of pEntYN10 during passages was verified during 7 d by ΔΔCt comparison between genes on plasmid (est/chromosome; gyrB) using realtime PCR. The Y-axis represents the difference of Ct value (ΔCt) between est and gyrB of respective strains. Asterisks denote statistically significant difference (*P < 0.001) by Student’s t-test between ETEC strain O169:H41 (YN10) and positive control (H10407).
redundancy of as many as 3 CFs on pEntYN10 may reinforce its adhesion to host cells in infectious stages more widely and tightly, which can enlarge their habitats as a pathogen. The redundancy may also affect its epidemiological features and pathogenicity of ETEC O169: H41. Each role of CFs that might contribute to pathogenicity and virulence should be scrutinized one by one using recombinant plasmids to uncover it.

In addition, 3 CFs encoded in pEntYN10 possessed unreported variations. The most unique features were found in K88-like fimbrial components. pEntYN10 encodes 2 copies of structural gene homologs tandemly: faeG1 and faeG2 (Fig. 3 (C)). Such duplication and their unique amino acid sequences might contribute to adhesion to human cells, although the adhesion targets of K88 fimbriae of ETEC have been regarded as restricted to porcine. The two paralogous genes have mutually diverged sequences. They are more similar to orthologous faeG of S. Infantis isolated from human patients than those of ETEC isolates (Fig. 7). That fact implies that the K88-like genes of the plasmid might enable O169:H41 strains to attach not to porcine cells but to human cells. However, gene structures and protein sequences of CS6 and CS8 on pEntYN10 were highly conserved, which suggests their functional conservation, although the parental strain YN1 was not reacted with the anti-CS8 MAb. How the unique variations found in the CFs respectively affect the affinity to the infected cells, host range, and virulence should be verified. Infection experiments using animal models might also provide more credible insights in their roles, although careful attention must be devoted to the difficulty of examining the possibility of its broad host range driven by the variety of CFs.

In conclusion, the complete sequence of a plasmid of ETEC O169:H41 showed its unique components, implying association with its phenotypic characteristics such as instability in vitro and atypical adhesion pattern to host cells. The CF variety might shed light on the evolutionary strategy and genetic diversity of ETEC pathogenicity.

Figure 3. Linear comparison of adhesion related gene clusters between pEntYN10 and other plasmids of ETEC strains reported previously. The gene clusters were retrieved from the GenBank database: (A) CS6 [accession ID:U04846]; (B) CFA/III [accession ID:AB049751]; (C) K88 gene clusters for comparison were retrieved from the complete sequence of plasmid pUMNK88 [accession ID:CP002730]. The length scales are shown in respective bars. Genes coding fimbrial structure proteins are shown as greyed. They were analyzed in detail for this study.

Materials and Methods

Bacterial strain

An ETEC YN10 was a derivative of the original strain ETEC O169:H41 YN1, isolated from a diarrheic patient during an outbreak foodborne disease of Osaka City, Japan in 1991. Our previous study revealed that the ETEC O169:H41 strains did not agglutinate human or bovine red blood cells in a mannose-resistant manner. The strains showed positive reaction in dot-blot tests using anti-CS6 antibodies. Although CS6 is a structural protein of CFA/IV, no fimbrial structure was identified.18

Instability test of plasmid pEntYN10

YN10 and a control strain H10407 were passaged for 7 d in 10 ml of TSB broth at 37°C. At each passage, 1 μl of each
To compare the adhesion patterns of respective strains quantitatively, the adhesion tests were operated 5 times and the number of

DNA sequencing

To purify the plasmid pEntYN10, the host strain YN10 was grown in Luria-Bertani broth at 37°C with vigorous aeration. The plasmid was purified by agarose extraction for separation from chromosomal DNA (NucleoBond BAC100; Macherey Nagel, GmbH and Co. KG, Germany), from cultured bacilli within at least 3 times passage from the original isolation.

The draft sequences of the plasmid were obtained using a genome sequencer (FLX System; 454 Life Sciences, Roche Applied Science, Branford, CT). Linkage of each contig estimated by assembly software GS De Novo Assembler (ver. 2.5.3) was verified using the gap filling method with PCR and direct sequencing of the amplified products.

Annotation

To check coding sequences (CDSs) on the plasmid, an automated annotation system, Microbial Genome Annotation Pipeline (MiGAP), was subjected to the complete sequence of the plasmid at first (http://www.migap.org/index.php/en). Subsequently, all annotated CDSs were curated further by search of the homologous protein sequences in GenBank using BLASTP. To avoid systematic sequencing errors by 454 FLX sequencer, putative frameshift errors caused by more than 5 nucleotide regions were checked using Sanger sequencing. The entire sequence and annotation information of the plasmid was registered at DDBJ [accession no: AP014654].

Tissue culture adhesion tests

Adhesion tests of YN10 and other E. coli strains ETEC3023, ETEC4266 (isolated from diarrheal patients admitted to the Infectious Diseases Hospital, Kolkata, India) and ETEC4266 (isolated from diarrheal patients admitted to the Infectious Diseases Hospital, Kolkata, India) and a laboratory strain HB101 for negative control to HEp-2 cells were conducted by infection with cultured bacilli (MOI = 1.0 × 10) on monolayers of HEp-2 cells in 0.5 mL of Basal Eagle’s medium as a previous study. Plasmid-cured YN10 was cloned from parental YN10 passaged during 10 d in TSB. After 3 h incubation at 37°C, the monolayers were washed 3 times with PBS(-). Then 0.5 mL of the medium was added. After more 3 h incubation, the monolayers were washed again 3 times as in the previous step. The cells were fixed with absolute methanol and were stained with Giemsa.

To compare the adhesion patterns of respective strains quantitatively, the adhesion tests were operated 5 times and the number of

Table 2. Predicted CDSs on pEntYN10 related to cell adhesion

| Cluster | Gene | Predicted function | Putative ortholog (accession no.) | Identity (no. of AA, %) | Reference |
|---------|------|--------------------|----------------------------------|------------------------|-----------|
| CS6 | cssD | Outer membrane usher | AAC45096 | 776/819 (94.7) | Wolf et al. 26 |
|     | cssC | Periplasmic chaperone | AAC45095 | 224/232 (96.6) | |
|     | cssB | Pilus subunit | AAC45094 | 159/167 (95.2) | |
|     | cssA | Pilus subunit | AAC45093 | 141/154 (91.6) | |
| CS8 (CFA/III) | cofP | Peptide-processing of prepilin | BAB62907 | 169/274 (61.7) | Taniguchi et al. 24 |
|     | cofJ | Unknown | BAB62906 | 191/363 (52.6) | |
|     | cofI | Type II secretion system | BAB62905 | 293/341 (85.9) | |
|     | cofH | Type II secretion system | BAB62904 | 397/510 (77.8) | |
|     | cofG | Pilus biosynthesis | BAB62903 | 93/162 (57.4) | |
|     | cofF | Pilus biosynthesis | BAB62902 | 159/275 (54.5) | |
|     | cofE | Pilus transmembrane anchor | BAB62901 | 99/186 (53.2) | |
|     | cofD | Pilus biosynthesis | BAB62900 | 412/485 (84.9) | |
|     | cofC | Pilus biosynthesis | BAB62899 | 86/160 (54.1) | |
|     | cofB | Pilus biosynthesis | BAB62898 | 336/527 (63.8) | |
|     | cofA | Pilus biosynthesis | BAB62897 | 175/239 (73.2) | |
|     | cofT | Unknown | BAB62896 | 106/147 (72.1) | |
|     | cofS | Transcriptional regulator | BAB62895 | 150/265 (64.1) | |
|     | cofR | Transcriptional regulator | BAB62894 | 66/110 (60.0) | |
|     | faeG2 | Fimbrial major subunit | YP_006131675.1 | 73/265 (27.5) | Shipard et al. 27 |
|     | faeG1 | Fimbrial major subunit | YP_006131674.1 | 113/258 (43.8) | |
|     | faeJ | Minor fimbrial subunit | YP_006131673.1 | 66/162 (41.3) | |
|     | faeI | Minor fimbrial subunit | YP_006131672.1 | 115/308 (37.3) | |
|     | faeH | Unknown | YP_006131671.1 | 129/291 (44.3) | |
|     | faeF | Minor fimbrial subunit | YP_006131670.1 | 152/309 (49.2) | |
|     | faeE | Minor fimbrial subunit | YP_006131669.1 | 66/126 (52.6) | |
|     | faeD | Outer membrane usher | YP_006131668.1 | 150/265 (64.1) | |
|     | faeC | Minor fimbrial subunit | YP_006131667.1 | 105/211 (49.8) | |
Figure 4. Light micrograph of HEp-2 cells infected by ETEC strains (A). Scale bars, 10 μm. The number of adhered bacteria was counted quantitatively for each strain (B). Asterisks denote statistically significant difference (*$P < 0.001$) by Student's t-test between ETEC strain O169:H41 (YN10) and other strains.
bacteria adhering to 60 epithelial cells was counted and its average was calculated.

**Pairwise alignment**

Pairwise alignment of amino acid sequences of orthologous CS genes were conducted using the EMBOSS needle program at the website of EMBL-EBI (http://www.ebi.ac.uk/Tools/psa/).

**Construction of phylogenetic trees**

Phylogenetic trees based on amino acid sequences were constructed using Neighbor-Joining method with a Poisson correction model and a 500 replicate bootstrap analysis using MEGA 6.06. Bootstrap consensus trees were calculated. Then the average bootstrap values were included on the trees.
Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Funding
This study was partially supported by a grant from the Ministry of Health, Labor, and Welfare of Japan (H24-Syokuhin-Ippan-010).

References
1. Qadri F, Svennerholm AM, Faruque AS, Sack RB. Enterotoxigenic Escherichia coli in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. Clin Microbiol Rev 2005; 18:465–83; PMID:16020685; http://dx.doi.org/10.1128/CMR.18.3.465-483.2005
2. Wetteras C, Elding V. Prevalence of enterotoxigenic Escherichia coli-associated diarrhoea and carrier state in the developing world. J Health Popul Nutr 2004; 22:370–82; PMID:15663170; http://www.jstor.org/stable/23499155
3. Levine MM. Escherichia coli that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohaemorrhagic, and enteroadherent. J Infect Dis 1987; 155:377–89; PMID:3543152; http://dx.doi.org/10.1093/infdis/155.3.377
4. Gaazra W, Svennerholm AM. Colonization factors of human enterotoxigenic Escherichia coli (ETEC). Trends Microbiol 1996; 4:444–52; PMID:8950814; http://dx.doi.org/10.1016/0966-842X(96)10068-8
5. Nataro JP, Kaper JB. Diarrheagenic Escherichia coli. Clin Microbiol Rev 1998; 11:442–201; PMID:9457432
6. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, Steinsland H. Molecular mechanisms of enterotoxigenic Escherichia coli infection. Microbes Infect 2010; 12:89–98; PMID:19883790; http://dx.doi.org/10.1016/j.micinf.2009.10.002
7. Johnson TJ, Nolan LK. Pathogenomics of the virulence plasmids of Escherichia coli. Microbiol Mol Biol Rev 2009; 73:750–74; PMID:19946140; http://dx.doi.org/10.1128/MMBR.00115-09
8. Sahl JW, Steinsland H, Redman JC, Angiuoli SV, Nataro JP, Sommerfelt H, Rasko DA. A comparative genomic analysis of diverse clonal types of enterotoxigenic Escherichia coli reveals pathovar-specific conservation. Infect Immun 2011; 79:950–60; PMID:21078854; http://dx.doi.org/10.1128/IAI.00932-10
9. Sahl JW, Steinsland H, Redman JC, Angiuoli SV, Nataro JP, Sommerfelt H, Rasko DA. A comparative genomic analysis of diverse clonal types of enterotoxigenic Escherichia coli reveals pathovar-specific conservation. Infect Immun 2011; 79:950–60; PMID:21078854; http://dx.doi.org/10.1128/IAI.00932-10
10. Schifferli DM, Beachey EH, Taylor RK. The 987P fimbrial gene cluster of enterotoxigenic Escherichia coli is plasmid encoded. Infect Immun 1990; 58:149–56; PMID:1967167
11. Casey TA, Moon HW. Genetic characterization and virulence of enterotoxigenic Escherichia coli mutants which have lost virulence genes in vivo. Infect Immun 1990; 58:4156–8; PMID:2254037
12. Anda K, Iraya T, Aoki A, Saito A, Masaki H, Tokunaga Y. An outbreak of food poisoning caused by enterotoxigenic Escherichia coli O169:H41. Jpn J Food Microbiol 1993; 10:77–81; PMID:24270144; http://dx.doi.org/10.7883/yoken.66.530
13. Nishikawa Y, Hanakawa M, Ogawawara J, Moyer NP, Kimura T. Heat-stable enterotoxin-producing Escherichia coli O169:H41 in Japan. Emerg Infect Dis 1995; 1:61; PMID:8903162; http://dx.doi.org/10.3201/eid0102.950206
14. Beatty ME, Bopp CA, Wells JG, Greene KD, Puhit ND, Minna ED. Enterotoxin-producing Escherichia coli O169:H41, United States. Emerg Infect Dis 2004; 10:518–21; PMID:15109427; http://dx.doi.org/10.3201/eid1003.030268
15. Cho SH, Kim J, Oh KH, Hu JK, Seo J, Oh SS, Han MJ, Choi YH, Youn SK, Chung GT, et al. Outbreak of enterotoxigenic Escherichia coli O169 enteritis in schoolchildren associated with consumption of kimchi, Republic of Korea, 2012. Epidemiol Infect 2014; 142:616–23; PMID:24806632; http://dx.doi.org/10.1017/S0950268813001477

Figure 7. Phylogenetic tree based on amino acid sequences of FaeG. Numbers on nodes of the phylogenetic tree correspond to bootstrap values.
15. Devaia RA, Jones TF, Ward J, Stafford L, Hardin H, Bopp C, Beatty M, Mintz E, Schaffner W. Endemically acquired foodborne outbreak of enterotxin-producing *Escherichia coli* serotype O169:H41. *Am J Med* 2006; 119:168-70; PMID:16443428; http://dx.doi.org/10.1016/j.amjmed.2005.07.063

16. Hamada K, Tsuji H, Shimada K. Outbreaks of heat stable enterotoxin-producing *Escherichia coli* O169:H41 in the Kinki district in Japan: epidemiological analysis by pulsed-field gel electrophoresis. *Jpn J Infect Dis* 1999; 52:165-7; PMID:10592898

17. Harada T, Inoh K, Yamaguchi Y, Hirai Y, Kanki M, Kawanou K, Sato T, Taguchi M, Kumeda Y. A foodborne outbreak of gastrointestinal illness caused by enterotoxigenic *Escherichia coli* serotype O169:H41 in Osaka, Japan. *Jpn J Infect Dis* 2013; 66:530-3; PMID:24270144; http://dx.doi.org/10.7883yoken.66.530

18. Nishikawa Y, Helander A, Ogasawara J, Moyer NP, Harada T, Itoh K, Yamaguchi Y, Hirai Y, Kanki M, Kawabata T, et al. The pCoo plasmid of enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* 1997; 179:6789-95; PMID:9366107; http://dx.doi.org/10.1128/JB.179.17.6789-6795.1997

19. Crossman LC, Chaudhuri RR, Beaton SA, Wells TJ, Desvaux M, Cunningham AF, Petry NK, Mahon V, Brinkley C, Hohman JL, et al. A communal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* 2010; 192:5822-31; PMID:20802035; http://dx.doi.org/10.1128/JB.00710-10

20. Froehlich B, Parkhill J, Sanders M, Quail MA, Scott JR, et al. The Escherichia coli genome: a mosaic cointegrate. *J Bacteriol* 2005; 187:6509-16; PMID:16157844; http://dx.doi.org/10.1128/JB.187.18.6509-6516.2005

21. Osuji S, Shimizu T, Ohtani K, Ichinose Y, Arimitsu H, Suzuki K, Kato M, Tsuji T. Nucleotide sequence analysis of the enterotoxigenic *Escherichia coli* Ent plasmid. *DNA Res* 2009; 16:299-309; PMID:19765799; http://dx.doi.org/10.1093/dnare/dsp015

22. Wajima T, Sabu S, Kano S, Ramamurthy T, Chatterjee NS, Hamabata T. Entero pathogenic *Escherichia coli* clinical isolate. *Plasmid* 2013; 70:345-52; PMID:23933556; http://dx.doi.org/10.1016/j.plasmid.2013.07.006

23. Taniguchi T, Akeda Y, Haba A, Yasuda Y, Yamamoto K, Honda T, Tsuikioubu K. Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. *Infect Immun* 2001; 69:5864-73; PMID:11500465; http://dx.doi.org/10.1128/IAI.69.9.5864-5873.2001

24. Verdonck F, Cos E, Schepers E, Imberechts H, Joensuu J, Geurtsen W. Preserved region in the sequence of the (K88) fimbrial adhesin FaeG suggest a donor strand mechanism in F assembly. *Vet Microbiol* 2004; 102:215-26; PMID:15327796; http://dx.doi.org/10.1016/j.vetmic.2004.06.002

25. Wolff MK, de Haan LA, Causeli FJ, Willshaw GA, Warren R, Boedeker EC, Gaastra W. The C6 colonization factor of human enterotoxigenic *Escherichia coli* contains two heterologous major subunits. *FEMS Microbiol Lett* 1997; 148:35-42; PMID:9066107; http://dx.doi.org/10.1128/FEMSL.1997.148.1.35-42.1997

26. Shepard SM, Danzeisen JL, Isaaceon RE, Seemann T, Koeze NS. PCR-based identification of common colonization factors of *Escherichia coli*. *J Infect Dis* 1998; 176:157-67; PMID:9490802; http://dx.doi.org/10.1086/313433

27. Hidaoka A, Hosokawa K, Tashita K, Fujishara S, Ogasawara J, Hase A, Haru-Kado Y, Nishikawa Y. Multiple real-time PCR for exhaustive detection of diarrhoeagenic *Escherichia coli*. *J Appl Microbiol* 2009; 106:410-20; PMID:19020309; http://dx.doi.org/10.1111/j.1365-2672.2008.04003.x

28. Taylor DE, Rooker M, Keelan M, Ng LK, Martin I, Perna NT, Burland NT, Blatmar FR. Genomic variability of O islands encoding tellurite resistance in enterohemorrhagic *Escherichia coli* O157:H7 isolates. *J Bacteriol* 2002; 184:4609-8; PMID:12169592; http://dx.doi.org/10.1128/JB.184.17.4609-4608.2002

29. Sugawara H, Ohyama A, Mori H, Kurokawa K. *Escherichia coli* genome annotation pipeline (MiGAP) for diverse users. *The 20th International Conference on Genome Informatics* (GIW2009). Yokohama, Japan, 2009.

30. Hidaoka A, Bhowmick R, Nandy RK, Ramamurthy T, Chatterjee NS. PCR-based identification of common colonization factors antigens of enterotoxigenic *Escherichia coli*. *J Clin Microbiol* 2007; 45:3068-71; PMID:17596357; http://dx.doi.org/10.1128/JCM.00666-07

31. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013; 30:2725-9; PMID:24132122; http://dx.doi.org/10.1093/molbev/msr197

32. Li J, Lim MS, Li S, Brock M, Piqee ME, Woods VL, Jr, Craig L. Virbiov cholerae toxin-coupled pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry. *Structure* 2008; 16:137–48; PMID:18184591; http://dx.doi.org/10.1016/j.str.2007.10.027

33. Mol O, Oudega B. Molecular and structural aspects of fimbrial biosynthesis and assembly in *Escherichia coli*. *FEBS Microbiol Rev* 1996; 19:25-52; PMID:8716546; http://dx.doi.org/10.1111/j.1574-6976.1996.d00252.x

34. Schuster SC. Next-generation sequencing transforms today’s biology. *Nat Methods* 2008; 5:16-8; PMID:18165802; http://dx.doi.org/10.1038/nmeth1156

35. Blanco J, Blanco M, Gutab M, Gunal J, Gonzalez EA. Enterotoxins, colonization factors and serotypes of enterotoxigenic *Escherichia coli* from humans and animals. *Microbiologia* 1991; 5:7-73; PMID:1684712

36. Arviv G, Tibia K, Steck N, Salmon-Divit M, Corduas A, Rahav G, Graul GA, Gal-Mor O. A unique megaplasmid contributes to stress tolerance and pathogenicity of an emergent *Salmonella enterica* serovar *Infantis* strain. *Environ Microbiol* 2014; 16:997–9; PMID:24320043; http://dx.doi.org/10.1111/1462-2920.12351

37. Sugawara H, Ohyama A, Mori H, Kurokawa K. *Escherichia coli* genome annotation pipeline (MiGAP) for diverse users. *The 20th International Conference on Genome Informatics* (GIW2009). Yokohama, Japan, 2009.

38. Hidaoka A, Bhowmick R, Nandy RK, Ramamurthy T, Chatterjee NS. PCR-based identification of common colonization factors antigens of enterotoxigenic *Escherichia coli*. *J Clin Microbiol* 2007; 45:3068-71; PMID:17596357; http://dx.doi.org/10.1128/JCM.00666-07

39. Sugawara H, Ohyama A, Mori H, Kurokawa K. *Escherichia coli* genome annotation pipeline (MiGAP) for diverse users. *The 20th International Conference on Genome Informatics* (GIW2009). Yokohama, Japan, 2009.