<table>
<thead>
<tr>
<th>Title</th>
<th>THE NUCLEOTIDE SEQUENCE OF THE GENES, fanE AND fanF, OF Escherichia coli K99 FIMBRIAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>ONO, Eriko; LAVIN, Martin F.; NAIKI, Masaharu</td>
</tr>
<tr>
<td>Citation</td>
<td>Japanese Journal of Veterinary Research, 39(1): 1-10</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1991-05-30</td>
</tr>
<tr>
<td>DOI</td>
<td>10.14943/jjvr.39.1.1</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/3237">http://hdl.handle.net/2115/3237</a></td>
</tr>
<tr>
<td>Type</td>
<td>bulletin (article)</td>
</tr>
<tr>
<td>File Information</td>
<td>KJ00002377463.pdf</td>
</tr>
</tbody>
</table>

HOKKAIDO UNIVERSITY

Erika \textsc{ono}^{1}, Martin F. \textsc{lavin}^{2}, Masaharu \textsc{naiki}^{1,3,4}

(Accepted for Publication: January 7, 1991)

K99 fimbriae of enterotoxigenic \textit{Escherichia coli} consist of eight different subunits. A major subunit called fimbrillin forms fimbrial structure and a minor subunit called adhesin localizes at the tip of fimbriae and recognizes host receptor ganglioside. Within this eight gene cluster, \textit{fanE} and \textit{fanF} have not yet been sequenced. In this study, \textit{fanE} and \textit{fanF} genes were sequenced by analyzing several DNA fragments produced by endonuclease or exonuclease digestion. The \textit{fanE} gene encoded 227 amino acids containing 20 amino acids of signal peptide starting from GTG (valine) and showed a homology to \textit{fanA-fanB}. The \textit{fanF} gene encoded 271 amino acids containing 20 amino acids of signal peptide starting from ATG (methionine) and showed homologies to the \textit{fanD} gene, fimbrillin gene of F41, adhesin gene of P fimbriae (\textit{papG}) and adhesin gene of Type 1 fimbriae (\textit{fimH}). E and F subunits had fifteen and fourteen hydrophobic domains, respectively, which periodically appeared possibly forming a hydrophobic region.

Key words: Enterotoxigenic \textit{Escherichia coli}, K99 fimbriae, \textit{fanE} and \textit{fanF} DNA sequence, hydrophobic domain.

INTRODUCTION

Enterotoxigenic \textit{Escherichia coli} possessing K99 fimbrial antigen are frequently isolated from newborn calves, piglets, and lambs suffering from diarrhea\textsuperscript{21}. It is known that fimbriae are important as an adhesive and colonization factor of bacteria on the mucosal surface of the small intestine in the first stage of infection\textsuperscript{1,5,6}. Fimbriae have been classified by their antigenicity and subunit molecular weight. The fimbriae are composed of fimbrillin which forms main part of fimbriae and adhesin which recognizes carbohydrate receptors on the surface of host cells\textsuperscript{13,23,24}.

---

\textsuperscript{1} Department of Biochemistry, Faculty of Veterinary Medicine, Hokkaido University, Sapporo 060, Japan.
\textsuperscript{2} Queensland Institute of Medical Research, Herston, Brisbane, Queensland 4006, Australia.
\textsuperscript{3} Present address; Department of Veterinary Science, National Institute for Health, Kamio-saki 2-10-35, Shinagawa-ku, Tokyo 141, Japan.
\textsuperscript{4} To whom correspondence should be addressed.
The K99 fimbrial gene cluster named fanA to fanH encodes eight different subunits\(^{15}\). The nucleotide sequences of fanA, fanB, fanC, fanD, fanG, and fanH have already been published\(^{16,17,19}\). The functions of these subunits were analyzed and it has been established that fanC encodes fimbrillin\(^{16}\), fanA and fanB encode regulatory proteins which control expression of fimbriae\(^{17,20}\), fanD encodes platform protein\(^{19}\) and fanG and fanH encode minor subunits\(^{18}\). However, there is no available information about the gene encoding adhesin. Since it is highly possible that either fanE or fanF encodes this receptor protein, in this study, the complete gene structure of fanE and fanF were determined.

**MATERIALS AND METHODS**

*Host cells and plasmid vector*: As a host bacterial strain, JM109 or JPA101 was used, and as a vector, pUC19\(^{26}\) or BLUESCRIPT\(^{22}\) was prepared for isolation of single-strand DNA. Isolation of plasmid DNA, agarose gel electrophoresis, and transformation were performed according to conventional methods\(^{22}\). Transformants were cultured in LB media or 2YT media\(^{22}\) with 50 \(\mu\)g/ml of ampicillin.

*Endonuclease and the other enzymes*: AvaI, BamHI, BglII, ClaI, HincII, KpnI, PstI, XbaI, exonuclease III, S1 nuclease, and *E. coli* DNA polymerase I (Klenow fragment) (TAKARA Co. Ltd., Kyoto, Japan) were used to prepare several DNA fragments from an original plasmid pFK99 (pBR322). T4 ligase (NIPPON GENE Co. Ltd., Toyama, Japan) was used for ligation.

*Sequence procedure*: Single-strand DNA was prepared according to SEQUENASE protocol or Biochemical research products: Protocol and application committed to science worldwide (Promega protocol). 6% polyacrylamide gel containing 7M urea was used in the electrophoresis for sequencing. Samples for sequencing were prepared using SEQUENASE Ver. 2.0 kit (TOYODO Co. Ltd., Osaka, Japan). pFK99 (pBR322) which contained K99 fimbrial gene cluster (fanA to fanH) were kindly supplied by Dr. de Graaf, Vrije University, Amsterdam, the Netherlands. BamHI–BamHI fragment encoding the whole K99 gene cluster of pFK99 (pBR322)\(^{25}\) was cloned into pUC19 vector (pFK99(pUC19)) and used for subcloning (Fig. 1).

*Preparation of subclones containing fanE and fanF*: Subclones including fanE and fanF genes were prepared (Fig. 1). KpnI–AvaI fragment isolated after agarose gel electrophoresis using GENECLEAN™ Kit (FUNAKOSHI Co. Ltd., Tokyo, Japan) was treated with Klenow enzyme and inserted again into pUC19 cleaved with KpnI and HincII, to make pFK99KA. pFK99AH (AvaI–HincII, fragment) and pFK99KC (KpnI–ClaI fragment) were prepared by using each particular restriction enzyme site. pFK99 (KK3.3R) was prepared from pFK99 (pUC19) by using the KpnI site. Digestion of pFK99(KK3.3R) with exonuclease III was performed at 37°C and at 30 sec intervals, samples of partially digested DNA were taken according to Promega protocol. After cloning, DNA length was checked by agarose gel electrophoresis and 15...
The nucleotide sequence of the genes.

subclones were selected to sequence completely.

Analysis of hydrophobic character of products: The hydrophobic character of the gene products was analyzed by the method described by Kyte and Doolittle, using Genetyx software supplied from SDC Software Exproilation Co., Ltd. (Tokyo, Japan).

RESULTS

Primary structure of fanE and fanF: The nucleotide sequences of fanE and fanF are shown in Fig. 2 with the corresponding amino acid sequences. Each structural gene had −35 and −10 promoter sequences upstream from the ribosome−binding site (S. D.). The fanE product had a signal peptide of 20 amino acids and a mature molecule consisted of 207 amino acids. There are 3 restriction enzyme sites for HincII, Clal and AvaI. The starting codon of fanE, GTG (valine) was overlapped by the stopping codon of fanD. Comparing fimbrial DNA sequences, fanE had homology to fanA−fanB (44.5%). The fanF product however, had a signal peptide of 20 amino acids and a mature protein consisted of 251 amino acids. There are 2 restriction enzyme sites for Sau3AI and XmnI. The starting region of fanF was also overlapped by the terminal region of fanE as observed at the fanD−fanE region. fanF showed homologies to fanD (45.7%), papG (47.8%) of P fimbriae, fimH (44.5%) of Type 1 fimbriae, and F41 fimbriae (45.1%). It is known that papG and fimH encode adhesin subunits of P and Type 1 fimbriae, respectively.

fanE

<table>
<thead>
<tr>
<th>-35</th>
<th>-10</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAG AGA GTC TTT TTT TGA TCA TGT ACC ATA AGT ACT TCG GGA CTA AAA GAT AAT TTA TCA GAA ATT CAG GAG GTA AAC T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTG AAT AAA TTT ATT TCT ATT ATA GCT CTC TGT GTA TTC ACC TCT TAC GCC TGT ATT GCC GCT TTT ACT CTC AAT ACG CCG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Val Asn Lys Phe Ile Ser Ile Ile Ala Leu Cys Val Phe Ser Ser Tyr Ala Ala Ala Phe Thr Leu Asn Ser Thr Arg |

HincII

<table>
<thead>
<tr>
<th>-20</th>
<th>-15</th>
<th>-10</th>
<th>-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAT ATA TAT GAA GGA CAG CAA AGT GTA TCT GTT AAC CAT AAC GAA AGT GAA CAT AAA TAT GAT GGT GGT GTA TGG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr Ile Tyr Asn Gly Gin Gin Ser Val Ser Val Asn Ile His Asn Gin Ser Gin His Lys Tyr Gly Gin Val Trp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Clal

<table>
<thead>
<tr>
<th>-35</th>
<th>-40</th>
<th>-45</th>
<th>-50</th>
<th>-55</th>
<th>-60</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT GAT AAT ATA GAC AAG MAT GGA GAG GTC GTA TTT TTT TCT CCA AGT CCA ATG GTG TCC AAA TTT GAC CCG AAA CAG AMG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile Asp Asp Asp Lys Asn Gly Gin Val Val Phe Ser Pro Ser Pro Met Val Phe Lys Leu Asn Pro Lys Gin Lys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
</tbody>
</table>

CiaI

<table>
<thead>
<tr>
<th>-65</th>
<th>-70</th>
<th>-75</th>
<th>-80</th>
<th>-85</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAA ATA GTG GCC ATT GTT AAT ATT AAT GAT ATT TTA GCT AAA GAC AGA GAA TCG ATT TTT TGG GTT ATT GTT CAG GAG ATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gin Ile Val Arg Ile Val Asn Ile Asn Asp Asp Leu Pro Lys Asp Arg Gin Ser Ile Phe Thr Leu Asn Val Gin Glu Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td>85</td>
</tr>
</tbody>
</table>
The nucleotide sequence of the genes.

CCT GCA GCT GCA AMG GGA GAT GGG GGT AGT CTG TCA CTG GCA ATT AAT AAT GCA GTA AAA TTA ATA TAT GCA OCA ATG

Pro Pro Ala Pro Lys Gly Asp Arg Glu Ser Leu Ser Leu Ala Ile Asn Asn Arg Val Lys Leu Ile Tyr Arg Pro Ile Ala

90 95 100 105 110 115

Ava!

CTA AAA ATT GCT GCA GAT GAG GCA GAA AAT ATT AAT AGT ATT AGC TCG GCA ATT AAT AGC AGG

Leu Lys Asn Gly Arg Asp Glu Ala Glu Asn Asn Ile Lys Leu Ile Asn Ser Gly Thr Asp Ser Cys Leu Glu Asn Thr Thr

120 125 130 135 140

CCA TAT TAC TTT GCA ATT AGT GAT GTT AAA ATT GCG AAA TCA ATT GAT TTA ATT TCT GAT GCA AAA AAT AGG GCA

Pro Tyr Tyr Phe Ala Ile Ser Asp Val Lys Ile Asn Gly Asn Ile Asp Leu Asn Ser Asp Ala Lys Asn Lys Met Gly

145 150 155 160 165

GTA TIC TCC CAA TGC AAA TGT CTG GCA AAT GTA AAT ACT AGT GGA AAT GTA ACA GCA

Val Phe Ser Pro Phe Ser Lys Val Cys Leu Gly Asn Val Asn Thr Ser Gly Asn Ile Thr Val Val Ala Phe Asn Asp Tyr

170 175 180 185 190 195

S.D.

GCC GTT GCA ACC AGC TAC ACT GTT CAA AGG AGT AAA ATG AAA ATT AAA TAAT AAT TTA TTA TTA TTA TTA TTA TTA

Gly Val Ala Thr Ser Tyr Thr Val Gin Arg Ser Lys

200 205

fanF

Met Lys Asn Lys Tyr Asn Leu Leu Phe Leu Phe Leu Leu

-20 -15 -10

TGT TAT GGA GAT GTG GGG CTG GCA GCA TGC ACA GGG AAA CTG AAA ATC TCA CCT GGT TAT AGT GGC CAT ACT TAT TCA TTT

Cys Tyr Gly Asp Val Ala Leu Ala Ala Cys Thr Gly Leu Lys Ile Ser Pro Gly Tyr Ser Gly His Thr Tyr Ser Phe

-5 5 10 15 20

GAT TGC AGT ATT CGA AAT AGT AGT ATA GCA AGA ACG TAC CGT TCT GAT TCA AAA ATT GTT GTG GAT GGC GAC CAG

Asp Ser Ser Ile Pro Asn Asn Ser Asn Ile Ala Arg Tyr Leu Val Glu Ile Ser Gly Lys Ile Cys Asp Ala Asp Gin

25 30 35 40 45

TCA GGC TGG GAT GGT AAA GGT TAT GCT CAA TTA CAT CTT TAT TCA CCT GGT GCC TTA TCT GAT AGC ACT GTC AGT GAT GGG

Ser Gly Trp Asp Gly Lys Arg Tyr Ala Gin Leu His Leu Tyr Ser Ser Gly Ala Leu Cys Gin Ser Val Ser Gly Asp Gly

50 55 60 65 70 75

ATT ACA TTY AGG TCA GTA GAT GTC GCA TCT TCC CAA TTT GCC AGT TCA ATC GCA GCA GCC CAA ATA

Ile Thr Phe Arg Ser Asn Val Ser Gly Leu Ser Trp Arg Phe Pro Asn Gly Ile Pro Tyr His Cys Ala Ala Gly Gin Ile

80 85 90 95 100

AAT CAT GCC GGT ATA AAT TGG GAT AGA AAT GGT AAA GGT ACC TGG ATT GCT GAT ATA GCA CAT GAA ATA TTT TTA

Asn Leu Gly Gly Ile Lys Tyr Ala Asp Arg Asn Gly Lys Val Thr Trp Asn Pro Gly Leu Arg His Glu Ile Phe Leu

105 110 115 120 125
Fig. 2. Primary sequence of \textit{fanE} and \textit{fanF} genes. An arrow shows a position cleaved by digestion with signal peptidase I. –35 and –10 indicate promoter sequences. S. D. means ribosome-binding site. Stopping codon is indicated by ***.

**Characterization of E and F subunits**: According to computer analysis, the E subunit had partial homologies to the B, C, and D subunits of K99 fimbriae in its amino acid Sequence. The F subunit was partially homologous to the G adhesin subunit of P fimbriae, C and H subunits of K99 fimbriae. The hydrophobic character of these proteins are shown in Fig. 3. Fifteen and fourteen hydrophobic domains appeared in the E and F subunits, respectively. Particularly notable are the hydrophobic peaks of \textit{fanE} which periodically appeared.

**DISCUSSION**

Various kinds of fimbrial genes were cloned from \textit{Escherichia coli} or other bacterial cells. Among them, fimbrial genes of \textit{p2,3,4,11,12,14} and Type \textit{17,8,9} were well
The nucleotide sequence of the genes.

**FanE**

![Graph showing hydrophobic character of FanE](image)

**FanF**

![Graph showing hydrophobic character of FanF](image)

Fig. 3. Hydrophobic character of E and F subunits. Signal peptide is removed at a position indicated by an arrow. Hydrophobic peaks are indicated by boxes.
analyzed. It is known that the fimbrial gene and amino acid sequence of the products have some homologies.

Roosendaal et al. have reported on the K99 gene cluster and indicated that \textit{fan}C encodes fimbrillin of K99 antigen (16.5 KDa)\textsuperscript{16}. The A subunit (11.0 KDa)\textsuperscript{17}, B subunit (10.8 KDa)\textsuperscript{17}, D subunit (84.5 KDa)\textsuperscript{19}, G subunit (16.9 KDa)\textsuperscript{18}, and H subunit (16.3 KDa)\textsuperscript{18} have been reported to be homologous in their nucleotide sequences. In the analysis of hydrophobicity, the periodic appearance of hydrophobic domain was detected.

In this study, \textit{fan}E and \textit{fan}F were sequenced for the first time. The starting codon of \textit{fan}E was GTG, coding for valine. This starting codon is unique among the fimbrial genes of which nucleotide sequences have been reported. It was found that \textit{fan}D, \textit{fan}E, and \textit{fan}F overlapped each other in the structural gene regions. A similar overlapping was reported in \textit{fan}G and \textit{fan}H of K99 fimbriae\textsuperscript{18}). In addition, the \textit{fan}F had homologies to the genes encoding adhesin of P and Type 1 fimbriae. In contrast, \textit{fan}E showed a homology only to the \textit{fan}A-\textit{fan}B connecting region. Thus, the E subunit seems a very unique subunit within the K99 fimbriae. Regarding the hydrophobic character of \textit{fan}E and \textit{fan}F, fifteen and fourteen peaks were detected, respectively. In addition, the peaks of \textit{fan}E appeared periodically, indicating that those formed the hydrophobic part of the E subunit.

At present it is not known which gene encodes K99 adhesin. However, it was found that the \textit{fan}F had some homology to the adhesin gene of other fimbriae. Further analysis using mutants lacking \textit{fan}E and \textit{fan}F is necessary to determine the function of these subunits. Such experiments are now in progress.

REFERENCES


The nucleotide sequence of the genes.


