<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</td>
</tr>
<tr>
<td>File Information</td>
<td>KJ00002377463.pdf</td>
</tr>
<tr>
<td>HOKKAIDO UNIVERSITY Collection of Scholarly and Academic Papers: HUSCAP</td>
<td></td>
</tr>
</tbody>
</table>
THE NUCLEOTIDE SEQUENCE OF THE GENES, \( \text{fanE} \) AND \( \text{fanF} \), OF \( \text{Escherichia coli} \) K99 FIMBRIAE.

Erika ONO\(^1\), Martin F. LAVIN\(^2\), Masaharu NAIKI\(^1,3,4\)

(Accepted for Publication: January 7, 1991)

K99 fimbriae of enterotoxigenic \( \text{Escherichia coli} \) 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, \( \text{fanE} \) and \( \text{fanF} \) have not yet been sequenced. In this study, \( \text{fanE} \) and \( \text{fanF} \) genes were sequenced by analyzing several DNA fragments produced by endonuclease or exonuclease digestion. The \( \text{fanE} \) gene encoded 227 amino acids containing 20 amino acids of signal peptide starting from GTG (valine) and showed a homology to \( \text{fanA-fanB} \). The \( \text{fanF} \) gene encoded 271 amino acids containing 20 amino acids of signal peptide starting from ATG (methionine) and showed homologies to the \( \text{fanD} \) gene, fimbrillin gene of F41, adhesin gene of P fimbriae (\( \text{papG} \)) and adhesin gene of Type 1 fimbriae (\( \text{fimH} \)). E and F subunits had fifteen and fourteen hydrophobic domains, respectively, which periodically appeared possibly forming a hydrophobic region.

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

INTRODUCTION

Enterotoxigenic \( \text{Escherichia coli} \) possessing K99 fimbrial antigen are frequently isolated from newborn calves, piglets, and lambs suffering from diarrhea\(^{21}\). 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\(^{1,5,6}\). 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\(^{13,23,24}\).

\(^1\) Department of Biochemistry, Faculty of Veterinary Medicine, Hokkaido University, Sapporo 060, Japan.
\(^2\) Queensland Institute of Medical Research, Herston, Brisbane, Queensland 4006, Australia.
\(^3\) Present address; Department of Veterinary Science, National Institute for Health, Kamiosaki 2–10–35, Shinagawa-ku, Tokyo 141, Japan.
\(^4\) To whom correspondence should be addressed.
The K99 fimbrial gene cluster named *fanA* to *fanH* encodes eight different subunits\(^1\). 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 (pUC19). 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 reseach 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 (TOYOBO 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\(^\text{TM}\) 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
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\(^\text{10}\), using Genetyx software supplied from SDC Software Exproitation Co., Ltd. (Tokyo, Japan).

RESULTS

Primary structure of \textit{fanE} and \textit{fanF}: The nucleotide sequences of \textit{fanE} and \textit{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 \textit{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 \textit{HincII}, \textit{CiaI} and \textit{AvaI}. The starting codon of \textit{fanE}, GTG (valine) was overlapped by the stopping codon of \textit{fanD}. Comparing fimbrial DNA sequences, \textit{fanE} had homology to \textit{fanA–fanB} (44.5%). The \textit{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 \textit{Sau3AI} and \textit{XmnI}. The starting region of \textit{fanF} was also overlapped by the terminal region of \textit{fanE} as observed at the \textit{fanD–fanE} region. \textit{fanF} showed homologies to \textit{fanD} (45.7%), \textit{papG} (47.8%) of P fimbriae, \textit{fimH} (44.5%) of Type 1 fimbriae, and F41 fimbriae (45.1%). It is known that \textit{papG} and \textit{fimH} encode adhesin subunits of P and Type 1 fimbriae, respectively.

\begin{verbatim}
GAG AGA GTC TTT TTT GGA TCA TGT ACC ATA ACT TGG GGA GTA AAA GAT ATT TTA TCA GAA ATT CAG GAG GTA AAC T
\end{verbatim}

\begin{verbatim}
GTG ATT AAA TTT ATT TCT ATT ATA GCT TCT TGT GTA TTA ACC TCT TAC TAC GCT ATT GCC GCT TTT ACT CTC AAT AGT AGC C
\end{verbatim}

\begin{verbatim}
Val Asn Lys Phe Ile Ser Ile Ile Ala Leu Cys Val Phe Ser Ser Thr Ala Ala Ala Phe Thr Leu Asn Ser Thr Arg Val Asn Ala Ala Phe Thr Leu Asn Ser Thr Arg
\end{verbatim}

\begin{verbatim}
HincII
\end{verbatim}

\begin{verbatim}
TAT ATT GAA GGA GCA GAT GTA TCT GTT AAC ATC CAT AAC GAA AGT GAA CAT AAA ATT GCT GGT GGT TCT TAC TAC GCT TTT TAC TAC
\end{verbatim}

\begin{verbatim}
Tyr Ile Tyr Asn Gly Gln Gln Ser Val Ser Val Ser Val Ser Val Ser Ile His Asn Glu Ser Glu His Lys Tyr Gly Gly Gln Val Trp
\end{verbatim}

\begin{verbatim}
HincII
\end{verbatim}

\begin{verbatim}
ATT GAT ATT ATA GAA GAG AAT GGA GAC GGT GTA TTT TCT TCT CCA AGT GCA ATG GTG TAC AAA TGC AAC CGC AAA CAG AGC
\end{verbatim}

\begin{verbatim}
Ile Asp Asp Ile Asp Asp Lys Asn Gly Gln Gln Asp Val Val Phe Ser Ser Ser Met Val Phe Lys Leu Asn Pro Lys Gly Lys
\end{verbatim}

\begin{verbatim}
CiaI
\end{verbatim}

\begin{verbatim}
CAA ATA GTC GCC ATT GTT AAT AAT GAT ATT AAT AAT GAT AAT GAA GCA AGA GAA TGG ATT TCT TCT TTC ATT AAT GTT GAC GAG ATC
\end{verbatim}

\begin{verbatim}
Gln Ile Val Arg Ile Val Asn Ile Asn Asp Asp Asp Lys Pro Lys Arg Asp Gly Ser Ile Phe Trp Leu Asn Val Glu Ile
\end{verbatim}
The nucleotide sequence of the genes.

\[
\begin{align*}
\text{CCT GCA GCT GCA AMG GGA GAT GGG GGT AGT CTG TCA CTG GCA ATT AAT GAA GTA AAA TTA ATA TAT GCA CCA ATT GCT} \\
\text{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} \\
\text{90} & \quad 95 & \quad 100 & \quad 110 & \quad 115 \\
\text{Ava} \\
\text{CTA AAT ATG GGT GCA GAT GAG GCA GAA AAT ATT AAT CTG ATG AGC TCG GCC AGG GAT TCT TGC CTT GAA ATT ACA AGG} \\
\text{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} \\
\text{120} & \quad 125 & \quad 130 & \quad 135 & \quad 140 \\
\text{CCA TAT TAC TTT GCA ATT ATG GAT GTT AAA ATT GGC AAA TCA ATT GAT TTA ATT TCT GAT GCA AAA AAT AGG GCA} \\
\text{Pro Tyr Tyr Phe Ala Ile Ser Asp Val Lys Ile Asn Gly Lys Ser Ile Asn Leu Asn Ser Asp Ala Lys Asn Lys Met Gly} \\
\text{145} & \quad 150 & \quad 155 & \quad 160 & \quad 165 \\
\text{GTA TIC TCC CTC TCG AAA GTG TCT CTG GGA AAT GTA AAT ACT AGT GGA MC ATC ACG GTA ACA GCA} \\
\text{Val Phe Ser Pro Phe Ser Lys Cys Leu Gly Asn Val Asn Thr Ser Gly Asn Ile Thr Val Thr Ala Phe Asn Asp Tyr} \\
\text{170} & \quad 175 & \quad 180 & \quad 185 & \quad 190 & \quad 195 \\
\text{GTC ATT GCA ACC AGC TAC ACT GTT CAA AGG ACT AAA TA ATG AAA MAT AAA TAT TAT TTA TTA TTA TTA TTA CTG} \\
\text{S.D.} \\
\text{Gly Val Ala Thr Ser Tyr Thr Val Gin Arg Ser Lys ++ +} \\
\text{200} & \quad 205 & \quad 210 \\
\text{TGT TAT GGA GAT GGT GGG CTG GCA GCA TGG ACA GGA AAA CTG AAA ATC TCA CCT GGT TAT AGT GGC CAT ACT TAT TCA TTT} \\
\text{Cys Tyr Gly Asp Val Ala Leu Ala Ala Cys Thr Gly Lys Leu Lys Ile Ser Pro Gly Tyr Ser Gly His Thr Tyr Ser Phe} \\
\text{220} & \quad 225 & \quad 230 & \quad 235 & \quad 240 & \quad 245 \\
\text{GAT TOC AGT ATT CCA AAT AGT AAT GTA GCA ACA TAC CTG GTC GTA ATT TCT GAG AAA ATT GTT TGT GAT GCG CAG} \\
\text{Asp Ser Ser Ile Pro Asn Asn Ser Ile Ala Asp Tyr Gly Lys Ser Gly Asp Asp Ile Thr Val Gin Leu Ile Ser Gly Lys Ile Val Cys Asp Ala Asp Gin} \\
\text{255} & \quad 260 & \quad 265 & \quad 270 & \quad 275 \\
\text{TCA GGC TGG GAT GGT AAA GGT TAT GCT CAA TTA CAT CTAT TCA TCA GGT GCC TTA TCT GTA AGT GTC AGT GAT GGG} \\
\text{Ser Gly Trp Asp Gly Lys Arg Tyr Ala Gin Leu His Leu Tyr Ser Ser Gly Asp Ala Leu Cys Gin Ser Val Ser Gly Asp Gin} \\
\text{300} & \quad 305 & \quad 310 & \quad 315 & \quad 320 & \quad 325 \\
\text{TIT AAA TAT AGG TCA TTG GGG CTG TCA TGG CTG TTT CCG AAC GTA AGC GAT CAC TGT GCA GCC GCA AAA ATA} \\
\text{ATT ACA TTY AGG TCA TAT GTG TCG CTT GCA TCA AGT CCT GCC AGC AAA ACA TAC CAT CAC TGT GCA GCC GCA AAA ATA} \\
\text{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 Glu Ile} \\
\text{80} & \quad 85 & \quad 90 & \quad 95 & \quad 100 & \quad 105 & \quad 110 & \quad 115 & \quad 120 & \quad 125
ONO, E. et al.

**Fig. 2.** Primary sequence of *fanE* and *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 *fanE* which periodically appeared.

**DISCUSSION**

Various kinds of fimbrial genes were cloned from *Escherichia coli* or other bacterial cells. Among them, fimbrial genes of *E. coli* K2, 3, 4, 11, 12, 14) and Type 17, 8, 9) were well
The nucleotide sequence of the genes.

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 fanC encodes fimbrillin of K99 antigen (16.5 KDa)\(^{16}\). The A subunit (11.0 KDa)\(^{17}\), B subunit (10.8 KDa)\(^{17}\), D subunit (84.5 KDa)\(^{19}\), G subunit (16.9 KDa)\(^{18}\), and H subunit (16.3 KDa)\(^{18}\) 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{fanE} and \textit{fanF} were sequenced for the first time. The starting codon of \textit{fanE} 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{fanD}, \textit{fanE}, and \textit{fanF} overlapped each other in the structural gene regions. A similar overlapping was reported in \textit{fanG} and \textit{fanH} of K99 fimbriae\(^{18}\). In addition, the \textit{fanF} had homologies to the genes encoding adhesin of P and Type 1 fimbriae. In contrast, \textit{fanE} showed a homology only to the \textit{fanA}–\textit{fanB} connecting region. Thus, the E subunit seems a very unique subunit within the K99 fimbriae. Regarding the hydrophobic character of \textit{fanE} and \textit{fanF}, fifteen and fourteen peaks were detected, respectively. In addition, the peaks of \textit{fanE} 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{fanF} had some homology to the adhesin gene of other fimbriae. Further analysis using mutants lacking \textit{fanE} and \textit{fanF} is necessary to determine the function of these subunits. Such experiments are now in progress.

REFERENCES


