Molecular markers for detection of pathogenic *Escherichia coli* strains belonging to serogroups O138 and O139

Lei Wang a,b, Bin Liu a,b, Qingke Kong a,b, Hartmut Steinrück c, Gladys Krause c, Lothar Beutin c, Lu Feng a,b,*

a TEDA School of Biological Sciences and Biotechnology, Nankai University, 23# HongDa Street, TEDA, Tianjin 300457, PR China
b Tianjin Key Laboratory of Microbial Functional Genomics, TEDA College, Nankai University, 23# HongDa Street, TEDA, Tianjin 300457, PR China
c National Reference Laboratory for Escherichia coli, Federal Institute for Risk Assessment (BfR), Diedersdorfer Weg 1, D-12277 Berlin, Germany

Received 5 July 2005; received in revised form 23 September 2005; accepted 3 October 2005

Abstract

*Escherichia coli* strains belonging to O-serogroup 138 and 139 are important as disease agents in pigs causing post-weaning diarrhea and edema disease. Several types of shiga toxin-producing O138 and O139 strains were isolated from diarrheic humans and from cattle and food of bovine origin. Serotyping is the current method for detection of O138 and O139 strains but its applicability can be limited due to the presence of capsules and capsular-like bacterial surface antigens and in the case of rough LPS. To overcome these difficulties for diagnosis, we have developed a specific PCR method suitable for detection of different types of O138 and O139 strains. The O-antigen gene clusters of *E. coli* O138 and O139 type strains were sequenced, and the genes were identified on the basis of homology. By screening against 186 *E. coli* and *Shigella* type strains, two genes specific to each of *E. coli* O138 and O139 were identified, respectively, and were tested on 15 clinical and environmental isolates of those two serogroups in a double-blind test. The sensitivity of the PCR assays was determined, and the detection limits were 2 pg per µl of chromosomal DNA and 2 CFU per 10 g of water or pork samples. PCR-based detection of O-antigen specific genes of *E. coli* O138 and O139 was shown to be accurate, highly sensitive and rapid, and is suggested as a new diagnostic tool for investigations of infections and outbreaks with these strains in animals and humans and for control of food.

© 2005 Elsevier B.V. All rights reserved.

Keywords: *E. coli* O138; *E. coli* O139; PCR assay; Molecular typing

1. Introduction

Certain strains of *Escherichia coli* are agents of post-weaning diarrhea (PWD) and of edema disease (ED) in pigs. PWD and ED strains are associated with...
a few O-serogroups of *E. coli* worldwide. Among these, serogroups O138 and O139 are most frequently involved (Bell et al., 2001; Frydendahl, 2002; Gannon et al., 1988; Garabal et al., 1996; Nagy et al., 1990; Parma et al., 2000).

Representative O138 and O139 from pigs with PWD and ED often carry capsular antigens such as K81 and K82 (K12), respectively (Bertschinger and Gyles, 1994). Routine detection of such encapsulated strains is performed by slide agglutination with specific OK-antisera (Bertschinger and Gyles, 1994; Orskov and Orskov, 1978). However, serotyping can be of limited value for detection since O-rough strains with the typical virulence attributes were isolated from cases of PWD and ED in pigs (Aarestrup et al., 1996; Frydendahl, 2002). Additional investigations such as phage typing of capsules and genotyping of virulence markers are employed for diagnosis of PWD and ED strains. On the other hand, detection of typical porcine ED and PWD associated serotypes is required in investigations of outbreaks on farms with high fatality rates and important economical losses (Bertschinger and Gyles, 1994; Hampson, 1994).

Classical PWD and ED strains do not occur in humans or animals other than pigs, but other types of shiga toxin-producing O138 and O139 strains were isolated from diseased humans and from cattle and beef products (Takeda, 1999; Montenegro et al., 1990; Read et al., 1992; Sandhu et al., 1996). These strains do not carry K81 or K82/K12 capsules and are therefore not detectable with PWD and ED strain specific OK sera (Orskov and Orskov, 1978).

The difficulties for O-serotyping which may be caused by the presence of capsules, capsule-like fimbriae and rough LPS prompted us to develop a PCR-based detection method specific for O138 and O139 antigen synthetic genes. In *E. coli*, the genes for O-antigen synthesis are generally clustered on the chromosome between *galF* and *gnd*, and those encoding glycosyltransferases; O-unit flippase (*Wzx*) and O-antigen polymerase (*Wzy*) are often specific to individual O-antigen gene clusters. In recent years, PCR assays based on these gene clusters have been established in a number of *E. coli* serotypes, such as O15, O26, O86, O91, O113, O104, O111, O114, O121, O157 and O172 (Beutin et al., 2005; Feng et al., 2005, 2004b; Wang and Reeves, 1998; Guo et al., 2004), and were shown to be accurate, sensitive and rapid.

In the present study, the O-antigen gene clusters of *E. coli* O138 and O139 type strains were sequenced, and the genes were identified on the basis of homology. By screening with all the 186 *E. coli* (including *Shigella*) type strains, four genes specific for *E. coli* O138 and O139 were identified, respectively. PCR methods based on O-antigen-specific genes were established for the identification and detection of strains belonging to these two pathogenic serogroups.

2. Materials and methods

2.1. Bacterial strains and plasmids

The *E. coli* typing strains G1242 (O138:K81:H-) (CDC62-57 (Orskov et al., 1977)) and G1208 (O139:K12(82):H1) (CDC63-57 (Orskov et al., 1977)), G1198 (O139:K82:H1), G1446 (O139:K12:H1?) and G1656 (O139:H56), were kindly provided by P. R. Reeves and Yuliy Ratiner. The strains G1120 (*E. coli* O138:K81,88ac:H19) and G1122 (O139:K82:H1) were obtained from The National Institute for the Control of Pharmaceutical and Biological Products, Beijing, PR China. The clinical strains listed in Table 1 and the reference strains for adhesins F4 (K88), F5 (K99), F6 (987P) and F18 (F107) fimbriae were from the collection of the Federal Institute for Risk Assessment (BfR) in Berlin, Germany. Other *E. coli* and *Shigella* type strains used in this study were as previously described (Feng et al., 2004b).

2.2. Serotyping and detection of virulence markers

Serotyping of O (lipopolysaccharide)- and H-antigens and detection of flagellar (*fliC*)-specific sequences was performed as described elsewhere (Beutin et al., 2005). Detection of O138:K81 and O139:K12(82) and of adhesins F4 (K88), F5 (K99), F6 (987P) and F18 (F107) was performed by slide-agglutination with OK sera produced at the Federal Institute for Risk Assessment (BfR), Berlin, Germany. All strains listed in Table 1 were tested for Vero cytotoxicity and for production of hemolysins as previously described (Beutin et al., 2005). The following virulence genes were investigated by PCR with the primers and PCR-conditions as described.
before (Beutin et al., 2005): shiga toxins (stx-genes), intimin (eae-gene), EHEC-hemolysin (ehxA-gene), alpha-hemolysin (hlyA-gene), heat-labile enterotoxin (LT-I), heat stable enterotoxins (STIh and STIp). Subtyping of stx-genes was performed as described previously (Beutin et al., 2004). The presence of the stx_{2e} gene was confirmed by a specific PCR using primers BB1 and BB2 as previously described (Beutin et al., 1995).

### 2.3. Construction of DNaseI shot gun bank

Chromosomal DNA was prepared as previously described (Bastin and Reeves, 1995). Primers #1523

<table>
<thead>
<tr>
<th>Strain Source, Illness, origin (reference)</th>
<th>Serotype\textsuperscript{b}</th>
<th>Virulence markers\textsuperscript{c}</th>
<th>Hly-genotype\textsuperscript{e}</th>
<th>stx-genotype\textsuperscript{f}</th>
<th>LT-I\textsuperscript{g}</th>
<th>STIp\textsuperscript{h}</th>
</tr>
</thead>
<tbody>
<tr>
<td>00296 Piglet, PWD, New Zealand</td>
<td>O138:K81:H19 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00297 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O138:K81:H19 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00298 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O138:K81:H19 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00299 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O138:K81:H19 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00667 Piglet, ED, Germany (this work, 1989)</td>
<td>O138:K81:H14 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00668 Piglet, ED, Germany (this work, 1989)</td>
<td>O138:K81:H14 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00728 Piglet, ED, Germany (this work, 1989)</td>
<td>O138:K81:H14 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00853 Piglet, ED, Germany (this work, 1989)</td>
<td>O138:K81:H14 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00649 Cattle, AS, Germany (Montenegro et al., 1990)</td>
<td>O139:H19 E-hly</td>
<td>ehxA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00295 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O139:K12(K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00301 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O139:K12 (K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00302 Piglet, PWD, New Zealand (Hampson et al., 1988)</td>
<td>O139:K12 (K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>00868 Piglet, ED, Germany (this work, 1989)</td>
<td>O139:K12(K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>01104 Piglet, ED, Germany (Wittig and Fabricius, 1992)</td>
<td>O139:K12(K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>03373 Piglet, ED, Germany (this work, 1993)</td>
<td>O139:K12(K82):H1 Alpha</td>
<td>hlyA</td>
<td>stx_{2e}</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Year of publication, in case of unpublished data: year of isolation.
\textsuperscript{b} Nomenclature of K antigens according to Orskov et al. (1977).
\textsuperscript{c} All strains were negative for genes encoding intimin (eae-gene) and the human type heat-stable enterotoxin (STIh), as well as for adhesins F4 (K88), F5 (K99), F6 (987P) and F18 (F107) by slide-agglutination with specific OK-antisera.
\textsuperscript{d} Hemolytic phenotype detected on washed sheep blood-agar plates as described (Beutin et al., 2005). Alpha, alpha hemolytic; E-hly, enterohemolytic.
\textsuperscript{e} PCR detection of hlyA (alpha-hlyA gene), ehxA (EHEC-hlyA gene).
\textsuperscript{f} The presence of stx-genes corresponded to the results for Vero cell toxicity (data not shown).
\textsuperscript{g} PCR detection of heat-labile enterotoxin LT-I. –, negative; +, positive.
\textsuperscript{h} PCR detection of porcine heat stable enterotoxin (STIp). –, negative; +, positive.
and #1524 (Feng et al., 2004b) based on the galF and gnd genes, respectively, were used to amplify E. coli O138 and O139 O-antigen gene cluster, using the Expand Long Template PCR system from Roche. The PCR cycles used were as follows: denaturation at 94 °C for 10 s, annealing at 60 °C for 30 s and extension at 68 °C for 15 min. The PCR products were digested with DNaseI and the resulting DNA fragments were cloned into pGEM-T easy to produce a bank using the method described previously (Wang and Reeves, 1998).

2.4. Sequencing and analysis

Sequencing was carried out using an ABI 3730 automated DNA sequencer from Applied Biosystems. Sequence data were assembled using the Staden Package (Staden, 1996). The program Artemis (Rutherford et al., 2000) was used for annotation. The program BlockMaker (Henikoff et al., 1995) was used to search conserved motifs. Blast and PSI-BLAST were used for searching databases including GenBank, COG and Pfam protein motif databases. The TMHMM v2.0 analysis program (http://www.cbs.dtu.dk/services/TMHMM-2.0/) was used to identify potential transmembrane segments from the deduced amino acid sequence.

2.5. Specificity assay and sensitivity test by O-serogroup specific PCR

Chromosomal DNA prepared from each of 186 representative strains of all E. coli O-serotypes, was examined for its quality by PCR amplification of the mdh gene using primers described previously (Wang and Reeves, 1998). A total of 13 pools of DNA were made, each containing DNA from 12 to 19 strains (Feng et al., 2004b). E. coli O138:K81:H- (G1242), O139:K12(82):H1 (G1208), O139:K81,88ac:H19 (G1120), O139:K82:H1 (G1122), O139:K82:H1 (G1198), O139:K12:H? (G1446) and O139:H56 (G1656), were added to pool 13, respectively. Pools were screened using primers based on specific genes of E. coli O138 and O139 (Table 2). The PCR cycles used were as follows: denaturation at 94 °C for 2 min; annealing for 30 s (in different Tm) (Table 2), and extension at 72 °C for 1 min, 30 cycles. And a final extension step at 72 °C for 5 min. PCR was carried out in a total volume of 30 μl, of which 10 μl was run on an agarose gel to check for amplified DNA. Template DNA from 24 clinical E. coli isolates including 8 E. coli O138, 7 E. coli O139 (Table 1) and 9 E. coli strains belonging to other O-serogroups (data not shown) were prepared as described (Guo et al., 2004) and screened with the primers listed in Table 2 in a double-blind test.

<table>
<thead>
<tr>
<th>Serogroup</th>
<th>Gene</th>
<th>Base positions</th>
<th>Forward primer/reverse primer</th>
<th>Annealing temp. (°C) of PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>O138</td>
<td>wzx</td>
<td>4632–5909</td>
<td>w1-814 (5388–5407) 5’T-TTTTATACAGCACCACCTCAGCAGT-3′/w1-815 (5827–5847) 3’T-TGAACATCTTTAACGACACGTTAGTGTG-5’S′</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>w1-816 (4690–4708) 5’T-TAGCCACATACCTGCTTGG-3’S′/w1-817 (5393–5411) 3’T-TATGTTTGGTGTAGTGTGTCAGAAGT-5’S′</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>wzy</td>
<td>11469–12680</td>
<td>w1-818 (11536–11554) 5’T-TGCGCGACACCGAGCATTTAAGT-3’S′/w1-819 (12023–12041) 3’T-TTGTCCGAGTTAGTGTGTCAGAAGT-5’S′</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>w1-820 (12024–12042) 5’T-AGAACGACACGGTGGTGGTGGTG-3’S′/w1-821 (12502–12520) 3’T-AGGACCTAACGAGCATTTAAGT-5’S′</td>
<td>55</td>
</tr>
<tr>
<td>O139</td>
<td>wzx</td>
<td>7016–8236</td>
<td>w1-620 (7219–7236) 5’T-AGAAGACATCACCGCGGAGTGTG-3’S′/w1-621 (7708–7725) 3’T-TGGACCGACCGAGCATTTAAGT-5’S′</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>w1-622 (7712–7729) 5’T-GGTGCAGTGGTTATGTTAATCAGAAGT-3’S′/w1-623 (8100–8117) 3’T-TATGGGACAAAGAGCATTTAAGT-5’S′</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>wzy</td>
<td>4922–6160</td>
<td>w1-616 (5150–5167) 5’T-AGGATTTAATGAGTGTGCT-3’S′/w1-617 (5393–5410) 3’T-CCACGGAGTAGGATGTTAATCAGAAGT-5’S′</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>w1-618 (5343–5362) 5’T-AGGATTTAATGAGTGTGCT-3’S′/w1-619 (6046–6063) 3’T-ATTGCGGTAGATACAAAA-5’S’</td>
<td>55</td>
</tr>
</tbody>
</table>
The sensitivity of the PCR assays was tested with tenfold serial dilutions of chromosomal DNA of *E. coli* O138 or O139 with the primer pairs designed (Table 2). Primer pairs wl-814/wl-815 and wl-818/wl-819 from *E. coli* O138, and wl-620/wl-621 and wl-616/wl-617 from *E. coli* O139 were also used to screen for *E. coli* strains of these two O-serogroups in water and pork samples using the method as described previously (Guo et al., 2004).

### 2.6. Nucleotide sequence accession number

The DNA sequence of the *E. coli* O138 and O139 O-antigen gene cluster has been deposited in GenBank under the accession numbers DQ109551 and DQ109552.

### 3. Results and discussion

#### 3.1. Sequence analysis of the *E. coli* O138 O-antigen gene cluster

A sequence of 14,139 bases from *galF* to *gnd* was obtained from *E. coli* O138, and 11 open reading frames (*orfs*) in addition to *galF* and *gnd* were identified (Fig. 1). Putative genes were assigned functions based on their similarity to *orfs* from the available databases (Table 3).

(i) **Sugar biosynthesis pathway genes.** It is known that the O-unit of *E. coli* O138 consists of four sugar residues: two L-Rhamnose (L-Rha), a 2-acetamido-2-deoxy-D-glucose (D-GlcNAc), and a 2-acetamido-2-deoxy-D-galacturonic acid (D-GalNAcA) (Linnerborg et al., 1997) (Fig. 2). Since genes for the synthesis of common sugars including GlcNAc are located outside O-antigen gene cluster, only genes involved in the synthesis of dTDP-L-Rha and UDP-D-GalNAcA were expected in the *E. coli* O138 O-antigen gene cluster.

Orf1, Orf2, Orf3 and Orf4 shared between 84 and 98% identity to many known RmlB, D, A and C, respectively. *Rml* genes are responsible for the synthesis of dTDP-L-Rhamnose, and usually found together at the 5′ end of the O-antigen gene clusters in the order of *rmlBDAC*. The *rml* gene set has been well characterized both genetically and biochemically. Therefore, orf1, orf2, orf3 and orf4 were identified as *rmlB*, *rmlD*, *rmlA* and *rmlC*, respectively, and named accordingly.

Orf8 and Orf9 shared 58 shared 58% and 66% identity to *WbpO* and *WbpP* of *Pseudomonas aeruginosa* O6, respectively. *WbpO* and *WbpP* are responsible for the synthesis of UDP-GalNAcA and their function have been experimentally proved (Ishiyama et al., 2004; Creuzenet et al., 2000). Orf8 and Orf9 also shared 61% and 49% identity to Gna and Gne of *S. dysenteriae* type 7, respectively. *Gna* and *gne* have been proposed to share the same function as *wbpO* and *wbpP*, respectively (Feng et al., 2004a). Since GalNAcA is also present in *E. coli* O138, orf8 and orf9 were proposed to encode enzymes responsible for the synthesis of UDP-GalNAcA, and named *gna* and *gne*, respectively.

![E. coli O138 O-antigen gene cluster](image)

Fig. 1. The O-antigen gene clusters of *E. coli* O138 and O139. All the genes are transcribed in a direction from *galF* to *gnd*.
O-unit processing genes. Both Wzx and Wzy are typical inner membrane proteins. The only two orfs encoding predicted membrane proteins are orf5 and orf11. Orf5 has 11 predicted transmembrane segments, and belong to the RfbX (wzx) family (Pfam01943, E-value = 4.1 x 10^-45). Members of this family are membrane proteins involved in the export of O-antigen and teichoic acid. Orf5 also showed 69 and 47% similarity to the Wzx proteins of S. flexneri 2a and E. coli O121. When the three proteins were analyzed using the BLOCKMAKER program, six conserved motifs (31, 30, 13, 16, 5, 46 amino acid residues) were revealed. By using the consensus sequence of the conserved motifs to run the program PSI-BLAST to search the Genpept database.

(ii) O-unit processing genes. Both Wzx and Wzy are typical inner membrane proteins. The only two orfs encoding predicted membrane proteins are orf5 and orf11. Orf5 has 11 predicted transmembrane segments, and belong to the RfbX (wzx) family (Pfam01943, E-value = 4.1 x 10^-45). Members of this family are membrane proteins involved in the export of O-antigen and teichoic acid. Orf5 also showed 69 and 47% similarity to the Wzx proteins of S. flexneri 2a and E. coli O121. When the three proteins were analyzed using the BLOCKMAKER program, six conserved motifs (31, 30, 13, 16, 5, 46 amino acid residues) were revealed. By using the consensus sequence of the conserved motifs to run the program PSI-BLAST to search the Genpept database.

\[
\begin{align*}
\text{E. coli O138 O-antigen structure} \\
\beta-D-Glc(\rightarrow 3) \\
\rightarrow 3)-\alpha-L-Rha(1 \rightarrow 4)-\alpha-D-GalNAc(1 \rightarrow 3)-\beta-D-GlcNAc(1)
\end{align*}
\]

\[
\text{E. coli O139 O-antigen structure} \\
\rightarrow 2)-\alpha-L-Rha(1 \rightarrow 3)-\alpha-L-Rha(1 \rightarrow 4)-\alpha-D-GalNAcA(1 \rightarrow 3)-\beta-D-GlcNAc(1)
\]

Fig. 2. Structure of the E. coli O138 (Linnerborg et al., 1997) and O139 (Marie et al., 1998) O-antigens.
database, only distantly related Wzx proteins were retrieved after three iterations except for Wzx proteins of *S. flexneri 2a* and *E. coli* O121. Orf5 was proposed to be an O-unit flippase gene (wzx) and named accordingly. Orf11 has 11 predicted transmembrane segments and a large periplasmic loop of 64 amino acid residues, which is a characteristic topology for O-antigen polymerases (Daniels et al., 1998). Orf11 also showed 68 and 49% similarity to Wzy proteins of *P. aeruginosa* O13 and *P. aeruginosa* O15, respectively. The three proteins were analyzed using the BLOCK-MAKER program and six conserved motifs (12, 20, 16, 17, 54, 26 amino acid residues) were revealed. By searching the Genpept database, orf11 was proposed as an O-unit polymerase gene (wzy) and named accordingly.

(iii) Genes encoding sugar transferases. The *wecA* gene, which is outside the O-antigen gene cluster, encodes the enzyme for transferring GlcNAc-1-phosphate or GalNAc-1-phosphate, either of which may serve as the first sugar in the O-unit, to a lipid carrier (Und-P) to initiate O-unit synthesis. Therefore, three more glycosyltransferase genes were expected from the O138 O-antigen gene cluster. Orf10 showed 48% identity and 67% similarity to WbnI of *S. dysenteriae* type 7, which was predicted to be an glycosyltransferase responsible for the linkage α-ᴅ-GalNAcA(1 → 3)ᴅ-GlcNAc (Feng et al., 2004a). The same linkage is also present in the *E. coli* O138 O-antigen. Therefore, orf10 was proposed to encode a glycosyltransferase responsible for the linkage α-ᴅ-GalNAcA(1 → 3)ᴅ-GlcNAc in *E. coli* O138, and was named wfaH. Orf6 and Orf7 belong to the Glycos_trans family (Pfam00536, E-value = 2 × 10^-23) and Glycos_transf family (Pfam00534, E-value = 5.3 × 10^-6), respectively. Orf7 also showed 44% identity to glycosyltransferase of *P. aeruginosa* O14. Orf6 and orf7 were proposed to encode glycosyltransferases, and named as wfaF and wfaG, respectively.

3.2. Sequence analysis of the *E. coli* O139 O-antigen gene cluster

A sequence of 12,507 bases from *galF* to *gnd* was obtained from *E. coli* O139, and 10 orfs in addition to *galF* and *gnd* were identified (Fig. 1). Putative genes were assigned functions based on their similarity to orfs from the available databases (Table 4).

Table 4
Characteristics of the ORFs located in the *Escherichia coli* O139 O-antigen gene cluster

<table>
<thead>
<tr>
<th>Orf no.</th>
<th>Gene name</th>
<th>G+C content (%)</th>
<th>Similar protein(s), strain(s) (GenBank accession no.)</th>
<th>% Identical aa/% similar aa (total no. of aa)</th>
<th>Putative function of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rmlB</td>
<td>42.8</td>
<td>RmlB, <em>Shigella flexneri</em> 2a str. 301 (AAN43643)</td>
<td>95/97(361)</td>
<td>dTDP-glucose 4,6-dehydratase</td>
</tr>
<tr>
<td>2</td>
<td>rmlD</td>
<td>46.1</td>
<td>RmlD, <em>Shigella flexneri</em> 2a str. 301 (AAN43642)</td>
<td>97/98(299)</td>
<td>dTDP-6-deoxy-D-glucose-3,5 epimerase</td>
</tr>
<tr>
<td>3</td>
<td>rmlA</td>
<td>43.9</td>
<td>RmlA, <em>Shigella flexneri</em> 2a str. 301 (AAN43641)</td>
<td>97/99(292)</td>
<td>Glucose-1-phosphate thymidylyltransferase</td>
</tr>
<tr>
<td>4</td>
<td>rmlC</td>
<td>38.6</td>
<td>RmlC, <em>Shigella flexneri</em> 2a str. 301 (AAN43640)</td>
<td>86/91(171)</td>
<td>dTDP-4-dehydrodhamnosyltransferase</td>
</tr>
<tr>
<td>5</td>
<td>wzy</td>
<td>29.4</td>
<td>Wzy, <em>Streptococcus pneumoniae</em> (AAK20689)</td>
<td>21/45(353)</td>
<td>O-antigen polymerase</td>
</tr>
<tr>
<td>6</td>
<td>wfaI</td>
<td>31.0</td>
<td>Rhamnosyltransferase, <em>Pseudomonas aeruginosa</em> (AA04519)</td>
<td>34/57(297)</td>
<td>Glycosyltransferase</td>
</tr>
<tr>
<td>7</td>
<td>wzx</td>
<td>30.2</td>
<td>Wzx, <em>Versinia enterocolitica</em> (type 0:8) (AAC60766)</td>
<td>23/46(384)</td>
<td>O-antigen flippase</td>
</tr>
<tr>
<td>8</td>
<td>wfaJ</td>
<td>33.0</td>
<td>WbpR, <em>Pseudomonas aeruginosa</em> (AAF24001)</td>
<td>36/53(343)</td>
<td>Glycosyltransferase</td>
</tr>
<tr>
<td>9</td>
<td>wfaK</td>
<td>33.4</td>
<td>WbpH, <em>Escherichia coli</em> (AAO39702)</td>
<td>25/45(366)</td>
<td>Glycosyltransferase</td>
</tr>
<tr>
<td>10</td>
<td>wfaL</td>
<td>29.0</td>
<td>Rhamnosyl transferase, <em>Vibrio vulnificus</em> (AAM34821)</td>
<td>47/63(227)</td>
<td>Glycosyltransferase</td>
</tr>
</tbody>
</table>
(i) Sugar biosynthesis pathway genes. The O-unit of *E. coli* O139 consists of seven sugar residues: four L-Rhamnose, a d-GlcNAc, a d-Glc and a d-GalA (Marie et al., 1998) (Fig. 2). While d-GlcNAc, d-Glc and d-GalA are common sugars, only genes for the synthesis of Rhamnose are expected in the O139 O-antigen gene cluster.

Orfs 1–4 were identified as rmlB, rmlD, rmlA and rmlC, respectively, based on their high level identity to other known *rml* genes (identity between 86 and 97%) from other *E. coli* and *Shigella* strains and named accordingly.

(ii) O-unit processing genes. Both Orf7 and Orf5 were found to have 11 predicted transmembrane segments, and Orf5 also have a large periplasmic loop of 84 amino acid residues. Using the same analysis as in *E. coli* O138, orf7 and orf5 were proposed as *wza* and *wzy* gene, respectively, and named accordingly.

(iii) Genes encoding sugar transferases. Orf6, 8, 9 and 10 showed different levels of similarity to putative glycosyltransferases (Table 4). Orf6 and 10 were also related to glycosyltransferase family 2 described by Wiggins (Wiggins and Munro, 1998), and Orf8 and 9 belong to the glycosyltransferase family 1. Orf8 showed 36% identity to WbpR of *Pseudomonas aeruginosa* O5, which is a glycosyltransferase responsible for transferring α-d-GalA via (1→2) linkage. Therefore, Orf8 is a glycosyltransferase for α-d-GalA(1→2)-L-Rha linkage. Orf10 showed 34% identity and 58% similarity to WbbL of *E. coli* O16 (K-12), which is a rhamnosyltransferase for sugar linkage 1-Rha-α-(1→3)-d-GlcNAc. The same function was proposed for Orf10 as the same linkage also presents in O139. Orf6, 8, 9 and 10 were proposed to encode glycosyltransferases and named wfaI, wfaJ, wfaK and wfaL, respectively. Except for *wecA*, six transferase genes were needed for the synthesis of O139 O-antigen. However, there were only four glycosyltransferase genes found in the O139 O-antigen gene cluster.

It is well-accepted that the addition of glucosyl residues to the O-antigen in *E. coli*, *Shigella* and *Salmonella* was due to the modification genes located outside the O-antigen gene cluster (Allison and Verma, 2000). The same was proposed for the branch linkage β-d-Glc(1→3)-L-Rha in O139. The ability of a single glycosyltransferase to transfer the same sugar successively to what appear to be different substrates has been reported previously. As Orf10 was a proposed rhamnosyltransferase and there are two L-Rha-α-(1→3)-linkages present, it is suggested that Orf10 is responsible for the two sugar linkages L-Rha-α-(1→3)-d-GlcNAc and L-Rha-α-(1→3)-L-Rha in O139.

3.3. Identification of *E. coli* O138 and O139-serogroup specific genes

Two pairs of primers based on each of *wzx* and *wzy* genes from *E. coli* O138 and O139 were designed (Table 2), and used to screen DNA pools containing representatives of the 186 known O-antigen forms of *E. coli* and *Shigella* strains. With either primer pairs used, only the band of correct size was observed in the pool containing *E. coli* O138 or O139 strains depending on the primers used. There were no PCR products were detected in other pools.

A double-blind test was performed on *E. coli* clinical isolates including eight *E. coli* O138 and seven *E. coli* O139 strains (Table 1) and nine *E. coli* strains of other O-serogroups (data not shown). All of the *E. coli* O138 and O139 strains gave the expected PCR product corresponding to each primer pair, and no PCR products were obtained from other strains. Therefore, all of the primer pairs are specific to *E. coli* O138 and O139, respectively, and can be used for the development of a PCR assay for the identification and detection of *E. coli* O138 and O139 strains.

3.4. Sensitivity of the O-serogroup specific PCR assays

The primer pairs described in Table 2 were used to amplify serially diluted chromosomal DNA prepared from the *E. coli* reference strains for G1242 (O138) and G1208 (O139). Positive PCR results were obtained from as little as 2 pg per μl of DNA for either of the type strains.

Specific primer pairs w1-814/wl-815 and w1-818/ w1-819 from *E. coli* O138, and w1-620/wl-621 and w1- 616/wl-617 (O139) were used for detection of serogroup O138 and O139 reference strains from
spiked samples of pork and water. As few as $10^3$ CFU/g were detected in the samples when examined directly and 2 CFU/10 g in samples when the samples were further incubated at 37 °C for 12 h. 

*E. coli* O138 and O139 strains are worldwide occurring important pathogens in pigs, causing PWD and ED. Shiga toxin-producing (STEC) strains belonging to these O-groups were also identified as pathogens in humans (Takeda, 1999) and were isolated from cattle and food of bovine origin (Montenegro et al., 1990; Read et al., 1992; Sandhu et al., 1996). The classical method of serotyping for diagnosis of these pathogens is restricted to the classical porcine pathogenic types but the OK sera are not suitable for detection of other types of O138 and O139 strains which were found in cattle, bovine products and infected humans. O-serotyping fails if O-rough strains are present which was found not un frecuently observed with isolates from PWD and ED (Aarestrup et al., 1996; Frydendahl, 2002). The specific PCR that we have developed for detection of *E. coli* O138 and O139 strains overcomes the problems caused by serological detection and is faster and more accurate than conventional serotyping. The *E. coli* O138 and O139 specific PCR assay could be extremely useful tool in the epidemiologic investigation of piglet edema disease (ED) and post-weaning diarrhea (PWD) outbreaks and in detection of such *E. coli* serogroup strains from other sources including infected humans.

**Acknowledgments**

This work was supported by the Chinese National Science Fund for Distinguished Young Scholars (30125001), NSFC General Program (30270029, 30370339, 30370023), the 863 Program (2002AA2Z2051), and funding from the Science and Technology Committee of Tianjin City (013181711) to LF and LW. The authors thank Karl Bettelheim, Melbourne, Australia, for providing clinical isolates of *E. coli* from New Zealand.

**References**


