Escherichia coli serogroup O2 and O28ac O-antigen gene cluster sequences and detection of pathogenic *E. coli* O2 and O28ac by PCR

Pina M. Fratamico, Xianghe Yan, Yanhong Liu, Chitrita DebRoy, Brian Byrne, Aine Monaghan, Seamus Fanning, and Declan Bolton

Abstract: The O-antigen gene clusters of *Escherichia coli* serogroups O2 and O28ac were sequenced, and PCR assays were developed to identify strains belonging to these 2 serogroups. Sixteen and 8 open reading frames were mapped to these loci in *E. coli* O2:H4 U 9-41 and *E. coli* O28ac:H25 96-3286, respectively. The *wzx* (O-antigen flippase) and *wzy* (O-antigen polymerase) genes in the *E. coli* O2 and O28ac O-antigen gene clusters were selected as targets for PCR assays for their identification. PCR assays targeting the *wzx* and *wzy* genes were specific for these serogroups, with one exception. *Escherichia coli* serogroup O42 strains gave positive results with *wzx* and *wzy* PCR assays targeting *E. coli* O28ac, and antiserum raised against O42 cross-reacted with serogroup O28ac strains. The O-antigen gene cluster of a strain of *E. coli* serogroup O42 was sequenced, and there were only 3 nt differences between the O-antigen gene clusters of the O28ac and O42 strains. Multiplex PCR assays targeting the O2 *wzx* gene, the stx_1, stx_2, *hly*, *eae*, and *saa* genes, and the O28ac *wzx*, *ial*, *ipaC*, and *ipaH* genes were developed for detecting Shiga toxin-producing *E. coli* O2 strains and enteroinvasive *E. coli* O28ac strains, respectively. The O2 and O28ac *wzx* and *wzy* genes can be used as diagnostic markers in PCR assays for rapid identification of these serogroups as an alternative to serotyping, and the multiplex PCR assays targeting serogroup-specific genes in combination with virulence genes can be used to identify and to detect pathogenic serogroup O2 and O28ac strains.

Key words: Escherichia coli O2, *Escherichia coli* O28ac, O-antigen, multiplex PCR detection, virulence genes.

Résumé : Les groupes géniques de l'antigène O d'*Escherichia coli* des groupes sérologiques O2 et O28ac ont été séquencés, et des essais par PCR ont été développés afin d'identifier les souches appartenant à ces deux groupes sérologiques. Seize et 8 cadres de lecture ouverts ont été respectivement cartographiés sur ces locus chez *E. coli* O2:H4 U 9-41 et *E. coli* O28ac:H25 96-3286. Les gènes *wzx* (flippase de l'antigène O) et *wzy* (polymérase de l'antigène O) des groupes géniques de l'antigène O chez *E. coli* O2 et O28ac ont été choisis comme cibles des essais par PCR en vue de leur identification. Les essais PCR ayant pour cibles *wzx* et *wzy* étaient spécifiques à ces groupes sérologiques à une exception près. Les souches *E.coli* du groupe sérologique O42 généraient des résultats positifs en PCR pour *wzx* et *wzy* ciblant *E. coli* O28ac, et l'antisérum produit contre O42 réagissait de façon croisée avec les souches du groupe sérologique O28ac. Le groupe génique de l'antigène O d'une souche d'*E. coli* du groupe sérologique O42 a été séquencé, révélant qu'il n'y avait que 3 nt de différence entre les groupes géniques de l'antigène O des souches O28ac et O42. Des essais par PCR multiplex ciblant le gène *wzx*, les gènes stx_1, stx_2, *hly, eae,* et *saa* de O2 et les gènes *wzx, ial, ipaC et ipaH* de O28ac ont été développés pour détecter les souches d'*E. coli* O2 productrice de la toxine de Shiga et les souches d'*E. coli* O28ac entéroinvasives, respectivement. Comme alternative au typage sérologique, les gènes *wzx* et *wzy* des groupes O2 et O28ac peuvent être utilisés comme marqueurs diagnostics en PCR afin d'identifier rapidement ces groupes sérologiques, et les essais par PCR multiplex ciblant les gènes spécifiques des groupes sérologiques et les gènes de virulence peuvent être utilisés pour identifier et détecter les souches pathogènes des groupes sérologiques O2 et O28ac.

Mots-clés : Escherichia coli O2, *Escherichia coli* O28ac, antigène O, détection par PCR multiplex, gènes de virulence.

[Traduit par la Rédaction]

Received 26 October 2009. Revision received 12 January 2010. Accepted 26 January 2010. Published on the NRC Research Press Web site at cjm.nrc.ca on 15 April 2010.

P.M. Fratamico,[1] X. Yan, and Y. Liu. Eastern Regional Research Center, Agriculture Research Service, US Department of Agriculture, 600 East Mermaid Lane, Wyndmoor, PA 19038, USA.
C. DebRoy. E. coli Reference Center, Department of Veterinary and Biomedical Sciences, The Pennsylvania State University, 104 Wiley Lab, Wiley Lane, University Park, PA 16802, USA.
B. Byrne, A. Monaghan, and D. Bolton. Food Safety Department, Ashtown Food Research Centre, Teagasc, Ashtown, Dublin 15, Ireland.
S. Fanning. UCD Centre for Food Safety, School of Agriculture, Food Science, and Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland.

[1]Corresponding author (e-mail: pina.fratamico@ars.usda.gov).

Introduction

The 6 major categories of human diarrheagenic *Escherichia coli* strains include (1) Shiga toxin-producing *E. coli* (STEC), of which enterohemorrhagic *E. coli* (EHEC) is a subgroup; (2) enterotoxigenic *E. coli* (ETEC); (3) enteropathogenic *E. coli* (EPEC); (4) enteroaggregative *E. coli* (EAggEC); (5) enteroinvasive *E. coli* (EIEC); and (6) diffusely adherent *E. coli* (DAEC) (Smith and Fratamico 2005). Numerous STEC (also referred to as verocytotoxin-producing *E. coli* or VTEC) serotypes have been identified and have been responsible for sporadic cases and outbreaks of hemorrhagic colitis and hemolytic uremic syndrome (Nataro and Kaper 1998; Karmali 2003). STEC that cause hemorrhagic colitis and hemolytic uremic syndrome are often referred to as EHEC (Gyles 2007). In the United States, *E. coli* O157:H7 is the most important EHEC serotype. STEC–EHEC strains possess a number of virulence genes, and the genes that encode Shiga toxins are the most critical. There are 2 types of Shiga toxins, Shiga toxin 1 (Stx1) and Shiga toxin 2 (Stx2), and a number of variants of these toxins have been described (Gyles 2007). Other virulence factors of *E. coli* O157:H7 and non-O157 STEC include the *eae* (encoding for the intimin protein), *hlyA* (a plasmid-encoded hemolysin), *saa* (STEC agglutinating adhesin) (Paton et al. 2001), *toxB* (plasmid-encoded adhesin), and *efaI* (EHEC factor for adherence) (Gyles 2007). STEC serogroup O2 strains belonging to different H types, including H^-, H1, H4, H5, H8, H25, H27, H29, H45, and others have been cultured from humans with diarrheal illness, healthy and diseased animals, and food (Reischl et al. 2002; Bettelheim 2007; Cho et al. 2007; Islam et al. 2007; http://www.microbionet.com.au/vtec1u.htm). Increasing numbers of cases and outbreaks due to non-O157 STEC infections are being reported worldwide. Improved detection methods for non-O157 STEC, including STEC O2 strains, will increase our understanding of their environmental importance and their contribution to morbidity and mortality. Furthermore, the use of antibiotics for the treatment of a STEC infection is contraindicated, since it increases the risk of developing hemolytic uremic syndrome (Wong et al. 2000); therefore, the use of rapid methods is important for diagnosis of STEC infections. *Escherichia coli* serogroup O2 strains also cause extraintestinal infections in chickens, turkeys, and other avian species, and these strains are known as avian pathogenic *E. coli* (APEC) (Dho-Moulin and Fairbrother 1999).

EIEC and *Shigella* spp. cause an invasive, dysenteric form of diarrheal illness known as bacillary dysentery, and outbreaks have been linked to unpasteurized milk, beef, and cheese. EIEC strains elaborate a cytotoxin and an enterotoxin and possess a ca. 220 kb virulence plasmid, which carries genes that encode for a type III secretion system along with genes involved in entry into epithelial cells and movement of the bacteria from cell to cell. These include the *ial* (invasion associated locus), *ipaC* (invasion plasmid antigen C), and *ipaH* (invasion plasmid antigen H) genes (Fasano et al. 1990; Parsot 2005). These genes are located on a large virulence plasmid in EIEC strains, and PCR assays targeting them facilitate the differentiation of EIEC from other pathogroups of *E. coli*. EIEC serogroup O28ac strains are primarily associated with diarrhea in children (Pal et al. 1997). Moreover, STEC O28ac strains have also been identified. A stx_2-carrying O28ac:H25 strain caused bloody diarrhea in a patient in Montana (Jelacic et al. 2003), and STEC O28ac:H21 was isolated from cattle in India (Pal et al. 1997, 1999).

The O-antigen contains many repeats of an oligosaccharide unit (O-unit) and is located on the outer membrane of gram-negative bacteria, where it contributes to the antigenic variability of the cell surface (Reeves and Wang 2002). Traditionally, *E. coli* serotyping to distinguish the different *E. coli* O serogroups is performed by agglutination reactions using antisera raised in rabbits against the different O-antigen standard reference strains. However, antisera used for serotyping can only be generated by specialized laboratories with animal facilities, and traditional serotyping is laborious, time consuming, and often generates equivocal results due to cross-reactions between different serogroups. Thus strains are frequently mistyped or mutant rough strains arise that are nontypeable. More definitive *E. coli* typing data could be achieved using rapid and specific molecular-based techniques, such as PCR.

The genes involved in the biosynthesis of O-antigens in *E. coli* are located in the *rfb*-encoding O-antigen gene cluster flanked by the *galF* and *gnd* genes in the *E. coli* chromosome. A number of *E. coli* O-antigen gene clusters have been characterized (Wang et al. 2001; DebRoy et al. 2005; Fratamico et al. 2003, 2005). Several genes in these clusters, in particular the *wzx* (O-antigen flippase) and *wzy* (O-antigen polymerase) genes, show relatively low similarity among different *E. coli* serogroups, and PCR primers targeting these loci have been developed, thereby providing for serogroup-specific PCR assays (Wang et al. 2001; DebRoy et al. 2004, 2005; Fratamico et al. 2003, 2005).

Not all *E. coli* strains belonging to serogroups O2 and O28ac carry the virulence genes characteristic of STEC–EHEC or EIEC, respectively; therefore, PCR assays targeting serogroup-specific genes and virulence genes specific for O2 and O28ac can be used to identify these 2 serogroups, as well as to distinguish pathogenic from nonpathogenic strains. Thus, the objective of the current study was to sequence the O-antigen gene clusters of *E. coli* serogroups O2 and O28ac and identify specific genes that can be used as diagnostic markers for these serogroups. As *E. coli* O2 and O28ac strains are human pathogens, multiplex PCR assays targeting virulence genes and genes in the O-antigen gene clusters of the 2 serogroups were developed. These assays could potentially be employed for detection and for clinical diagnosis of diseases caused by these *E. coli* serogroups.

Materials and methods

Bacterial strains and growth conditions

Escherichia coli strains used for sequencing were *E. coli* O2:H4 strain U 9-41 (Ørskov et al. 1977), *E. coli* O28ac:H25 strain 96-3286 (Centers for Disease Control and Prevention (CDC), Atlanta, Ga.), and *E. coli* O42:H37 strain P11a (Ørskov et al. 1977).[2] Bacteria used to test for the specificity of PCR included the following: 21 STEC O2 strains, 85 O2 field isolates,

[2] The mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

Table 1. Oligonucleotide primers used in this study.

Target gene	Primer name	Primer sequence (5′ to 3′)	Size of PCR product (bp)	Reference
wzx (O2)	O2wzxF	F: TGGCCTTGTTCGATATACTGCGGA	819	This study
	O2wzxR	R: TCACGAGCTGAGCGAAACTGTTCA		
wzy (O2)	O2wzyF	F: TGCAACTCATTGGTCTGCTTTGCC	351	This study
	O2wzyR	R: CGGAAAGCCATAACAGGTAGAGAG		
wzx (O28ac)	O28acwzxF	F: ACCAGAGCAAGGACGATTTGTCA	554	This study
	O28acwzxR	R: CAACTTTAACTTTCCCAAGCGCGG		
wzy (O28ac)	O28acwzyF	F: GGTAATACACTTGCTGTGGTGGGT	218	This study
	O28acwzyR	R: ATGATTGACCATCCCAGGCCGTAT		
stx_1	vt1-F	F: ATAAATCGCCATTCGTTGACTAC	180	Paton and Paton 1998
	vt1-R	R: AGAACGCCCACTGAGATCATC		
stx_2	vt2-F	F: GGCACTGTCTGAAACTGCTCC	255	Paton and Paton 1998
	vt2-R	R: TCGCCAGTTATCTGACATTCTG		
hlyA	hlyA-R	F: GCATCATCAAGCGTACGTTCC	534	Paton and Paton 1998
	hlyA-F	R: AATGAGCCAAGCTGGTTAAGCT		
eae	eaeA-F	F: GACCCGGCACAAGCATAAGC	384	Paton and Paton 1998
	eaeA-R	R: CCACCTGCAGCAACAAGAGG		
saa	SAADF	F: CGTGATGAACAGGCTATTGC	119	Paton and Paton 2002
	SAADR	R: ATGGACATGCCTGTGGCAAC		
ipaH	EI-1	F: GCTGGAAAAACTCAGTGCCT	424	Tornieporth et al. 1995
	EI-2	R: CCAGTCCGTAAATTCATTCT		
ipaC	ipaC-I	F: ATCATTGCTCGCCTTACTGAC	862	Boudeau et al. 1999
	ipaC-II	R: GCAATCTGACTGGCTGCCG		
ial	ial-F	F: CTGGATGGTATGGTGAGG	320	Thong et al. 2005
	ial-R	R: GGAGGCCAACAATTATTTCC		

54 non-O2 field isolates, 4 *E. coli* O28ac:H$^-$ strains (provided by Dr. James Nataro, University of Maryland School of Medicine, Baltimore, Md., and described by Fasano et al. 1990), 1 *E. coli* O28ac field isolate, 35 serogroup O28 *E. coli* strains, 50 non-O28ac field isolates, and 10 O42 field isolates. The specificity of the PCR assays was also tested against 167 *E. coli* standard reference strains used for serotyping at the *E. coli* Reference Center at The Pennsylvania State University, University Park, Penn. The *E. coli* standard strains used in this study belonged to serogroups O1 through O175, but excluded O14, O31, O47, O67, O72, O93, O94, and O122, since these serogroups have been deleted from the classification (Ørskov et al. 1977). In addition, strains representative of other bacterial genera (n = 21) were tested, and these included one strain each of *Bacillus cereus*, *Citrobacter freundii*, *Enterobacter cloacae*, *Enterococcus aerogenes*, *Enterococcus faecalis*, *Hafnia alvei*, *Klebsiella pneumonia*, *Listeria monocytogenes*, *Pseudomonas aeroginosa*, *Proteus vulgaris*, *Salmonella enterica* sv. Anatum, *Salmonella enterica* sv. Arizona, *Salmonella enterica* sv. Choleraesuis, *Salmonella enterica* sv. Enteritidis, *Salmonella enterica* sv. Typhimurium, *Serratia marcescens*, *Shigella boydii*, *Staphylococcus aureus*, *Vibrio cholerae*, and *Yersinia enterocolitica*. All bacteria were grown in Luria–Bertani broth or on Luria–Bertani agar plates at 37 °C.

Determination and analysis of *E. coli* O2, O28ac, and O42 O-antigen gene cluster sequences

Genomic DNA was isolated using the DNeasy Tissue Kit (Qiagen, Inc., Valencia, Calif.) according to the manufacturer's instructions. Long PCR assays were performed to amplify the O-antigen gene clusters using the Expand Long Template PCR System (Roche Applied Science, Mannheim, Germany) and JUMPSTART (named for "just upstream of many polysaccharide-associated gene" STARTs) and GND (6-phosphogluconate dehydrogenase gene) primers that flank the *E. coli* O-antigen gene clusters essentially as previously described (Fratamico et al. 2003) and using 5 μL of template DNA (ca. 500 ng) and primers at 0.5 μmol/L. In addition, the primer annealing temperature was set at 64 °C. The sequence of the JUMPSTART sense primer was 5′-ATTGGTAGCTGTAAGCCAAGGGCGGTAGCGT-3′, and the antisense GND primer sequence was 5′-CACTGCCATACCGACGACGCCGATCTGTTGCTTGG-3′ (Invitrogen Life Technologies, Inc., Carlsbad, Calif.). Long PCR amplicons were purified using the DNA Extraction Kit (Fermentas Inc., Glen Burnie, Md.) and then subjected to partial digestion with *Nla*III and *Sau*3AI restriction enzymes to generate enriched DNA shotgun libraries by ligating these partially digested PCR products (>1 kb) with *Sph*I and *Bam*HI-digested pUC19, respectively. DNA sequencing was performed using a 3730 DNA Analyzer (Applied Biosystems, Foster City, Calif.) and the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Since the 5′- and 3′-end sequences of the long PCR products were not covered by the DNA shotgun library, the 2 ends were sequenced by using the purified long PCR product and the JUMPSTART and GND primers. Gene sequence data were assembled by using Sequencher software (Gene Codes Corporation, Ann Arbor, Mich.), and the open reading frames (ORFs) were located using Artemis software (Rutherford et al. 2000). A cutoff value of 100 amino acids was used for annotating ORFs. Putative gene function was ascertained by BlastP analysis.

Specificity of the PCR assays

Bacterial DNA used as templates for PCR assays was isolated as previously described (DebRoy et al. 2005). PCR primers (Table 1) were designed from the *E. coli* O2 and O28ac *wzx* and *wzy* genes using the Primer3 software program (http://frodo.wi.mit.edu/primer3/). Multiplex PCR reactions targeting the *wzx* and *wzy* genes of *E. coli* O2 and O28ac serogroups were performed and analyzed as described previously (DebRoy et al. 2005) using 3 mmol/L of $MgCl_2$ in a RapidCycler (Idaho Technology, Inc., Salt Lake City, Utah) and the rapid cycle DNA amplification method. For *E. coli* O2, cycling conditions consisted of denaturation at 94 °C, annealing at 56 °C, and extension at 72 °C for 55 s. For multiplex PCR assays with simultaneous amplification of the *E. coli* O28ac *wzx* and *wzy* genes, the cycling conditions consisted of denaturation at 94 °C, annealing at 58 °C, and extension at 72 °C for 25 s.

Multiplex PCR assays for detection of STEC O2 and EIEC O28ac strains

Multiplex PCR assays were developed for the detection of STEC–EHEC serogroup O2 strains, targeting the O2 *wzx*, stx_1, stx_2, *hly*, *eae*, and *saa* genes, and for EIEC O28ac strains, targeting the O28ac *wzx*, *ial*, *ipaC*, and *ipaH* genes. DNA was extracted from isolated colonies using the PrepMan Ultra reagent (Applied Biosystems) following the manufacturer's instructions. The multiplex PCR assays were performed using the QIAGEN Multiplex PCR Kit (QIAGEN, Valencia, Calif.), following the universal multiplex cycling protocol provided by the manufacturer and using an annealing temperature of 57 °C in a GeneAmp PCR System 9600 thermal cycler (Applied Biosystems). A 5 μL sample of template DNA was used in a 50 μL total reaction volume containing 0.25 μmol/L of the primers (Table 1). *Escherichia coli* strains from the culture collection at the Eastern Regional Research Center, Microbial Food Safety Research Unit, which carried each of the target virulence genes, were used as positive controls for the PCR assays. The PCR products were resolved by conventional agarose (1.5%) gel electrophoresis containing 0.5 μg/mL of ethidium bromide and then visualized under UV light and photographed.

Detection of *E. coli* O2 and O28ac in dog fecal samples

Escherichia coli O2:H27 1.2740 (*E. coli* Reference Center strain collection) and O28ac:H^- CVD/E1-1 (James Nataro, University of Maryland) strains were cultured overnight in Luria–Bertani broth and then serially diluted in 0.1% peptone. Samples of dog feces (2 g) were each inoculated with 2 mL of the diluted cultures containing ca. 10 colony-forming units of each strain in a Seward Stomacher filter bag (Cole Parmer, Vernon Hills, Ill.). Then 100 mL of modified TSB (mTSB) medium (Possé et al. 2008) was added, followed by enrichment as described by Possé and co-workers (2008). However, 10 mg/L of cefsulidin was added to the mTSB instead of 8 mg/L of novobiocin. A 1 mL sample of the enrichment was centrifuged for 2 min at 16 000*g*, the supernatant was discarded, and the pellet was washed once with 1 mL of sterile water. The pellet was resuspended in 200 μL of PrepMan Ultra, and the DNA was extracted as described previously. A 1 μL sample of the extracted DNA was used in the multiplex PCR assays as described previously, and the PCR products were analyzed by agarose gel electrophoresis.

Nucleotide sequence accession numbers

The DNA sequences of the O-antigen gene cluster of O2:H4 U 9-41, *E. coli* O28ac:H25 96-3286, and *E. coli* O42:H37 P11a were deposited in GenBank under the accession No. EU549863, DQ462205, and FJ539194, respectively.

Results and discussion

The genes that encode proteins within the *E. coli* O-antigen gene clusters consist of 3 categories: nucleotide sugar biosynthesis, glycosyl transferase, and the O-antigen processing genes (Samuel and Reeves 2003). Nucleotide sugar biosynthesis gene products are involved in the biosynthesis of the nucleotide sugar precursors in the O-antigen, which occurs in the cytoplasm. Glycosyl transferases usually have narrow substrate specificity and are responsible for transferring the various precursor sugars to form an oligosaccharide on a carrier lipid, undecaprenyl phosphate (UndP), located on the cytoplasmic side of the inner membrane. O-antigen processing proteins include a flippase (Wzx) and the O-antigen polymerase (Wzy). Wzx flips the O-unit across the inner membrane. After the UndP-linked O-unit has been translocated across the cytoplasmic membrane, the O-units are enzymatically bonded by Wzy through the formation of a classical glycosidic linkage. Although both Wzx and Wzy are membrane proteins and usually have low similarity among different bacteria, the action of Wzx is not very specific, while the Wzy activity is substrate specific (Samuel and Reeves 2003).

DNA sequences of 8579 and 14 063 bp were obtained from the *E. coli* O28ac and O2 O-antigen gene clusters, respectively. The O-antigen gene cluster of *E. coli* O2:H4 U 9-41 contained 16 ORFs (Table 2), and that of *E. coli* O28ac:H25 96-3286 contained 8 ORFs (Table 3), all oriented in the same transcriptional direction from *galF* to *gnd*. The deduced amino acid sequences from these ORFs were used to search the NCBI database to aid in their annotation and to determine their possible functions. Gene names were assigned on the basis of the Bacterial Polysaccharide Gene Nomenclature system described by Reeves et al. (1996) (http://www.microbio.usyd.edu.au/BPGD/big-paper.pdf).

Sequence analysis of the *E. coli* O2 O-antigen gene cluster

Genes involved in the sugar biosynthetic pathway

The O2 O-antigen gene cluster shown in Table 2 has 4 ORFs that encode for the structural genes *rmlA*, *rmlB*, *rmlC*, and *rmlD*. These genes are found in O-antigen gene clusters of other bacteria and encode for enzymes in the dTDP-L-rhamnose pathway, which make rhamnose from glucose-1-phosphate. *rmlA–D* encode for the enzymes glucose-1-phosphate thymidylyltransferase, dTDP-glucose 4,6 dehydratase, dTDP-4-dehydrorhamnose 3,5-epimerase, and dTDP-6-deoxy-D-glucose-3,5 epimerase, respectively, and the ORFs showed relatively high similarity to *rmlABCD* in other bacteria. As in other O-antigen gene clusters, the order of the genes in *E. coli* O2 was *rmlBDAC* (Samuel and

Table 2. Open reading frames (ORFs) in the O-antigen gene cluster of *Escherichia coli* serogroup O2.

ORF	Proposed gene name	Location	No. of amino acids	Putative function	Most significant homology (GenBank accession No.)	% identity, % similarity
1	*rmlB*	66–1151	361	dTDP-Glucose 4,6-dehydratase	dTDP-Glucose 4,6-dehydratase (*Escherichia coli* IAI39) (YP_002406996.1)	98, 99
2	*rmlD*	1151–2050	299	dTDP-6-Deoxy-D-glucose 3,5-epimerase	dTDP-6-Deoxy-L-mannose-dehydrogenase (*Escherichia coli*) (ACA24825.1)	97, 98
3	*rmlA*	2109–2984	291	Glucose-1-phosphate thymidylyltransferase	Glucose-1-phosphate thymidylyltransferase (*Escherichia coli* IAI39) (YP_002406998.1)	99, 99
4	*ftdA*	2999–3412	137	NDP-hexose isomerase	NDP-hexose isomerase (*Yersinia intermedia* ATCC 29909) (ZP_04635116.1)	67, 80
5	*ftdC*	3399–3866	155	WxcM-like protein	Hypothetical protein PROVRETT_01740 (*Providencia rettgeri* DSM 1131) (ZP_03638653.1)	71, 86
6	*ftdB*	3859–4965	368	Aminotransferase	WblQ protein (*Photorhabdus luminescens* subsp. *laumondii* TTO1) (NP_931971.1)	65, 80
7	*wzx*	4965–6227	420	O-antigen flippase	Polysaccharide biosynthesis protein (*Pectobacterium carotovorum* subsp. *carotovorum* PC1) (YP_003016888.1)	50, 70
8	*wegP*	6244–7290	348	Glycosyl transferase	Hypothetical protein FIC_01940 (*Flavobacteriaceae bacterium* 3519–10) (YP_003096444.1)	29, 50
9	*rmlC*	7315–7872	185	dTDP-4-dehydrorhamnose 3,5-epimerase	RmlC (*Escherichia coli*) (ACA24796.1)	70, 84
10	*wzy*	7941–8903	320	O-antigen polymerase	Hypothetical protein Gura_3055 (*Geobacter uraniireducens* Rf4) (YP_001231799.1)	26, 44
11	*wegQ*	8896–9783	295	Glycosyl transferase	Glycosyl transferase, putative, gt2D (*Cellvibrio japonicus* Ueda107) (reflYP_001983904.1)	31, 54
12	*wegR*	9814 – 10 890	358	Glycosyl transferase	Glycosyl transferase, group 1 (*Shewanella frigidimarina* NCIMB 400) (YP_751504.1)	57, 75
13	*yhhI*	11 070 – 11 666	198	Transposase	Putative transposase (*Escherichia coli* UMN026) (YP_002411411.1)	83, 89
14	ORF14	11 251 – 11 562	103	H repeat-associated protein	H repeat-associated protein (*Escherichia coli*) (ACD75787.1)	86, 87
15	ORF15	11 853 – 12 206	117	Transposase	Transposase (*Shigella dysenteriae* 1012) (ZP_03065072.1)	85, 89
16	*wegW*	12 275 – 13 903	542	Sulfatase	Putative transmembrane sulfatase protein (*Stenotrophomonas maltophilia* K279a) (YP_001970541.1)	39, 56

Reeves 2003). However, O-antigen gene clusters may carry only *rmlAB* or *rmlABC*, and these genes might be involved in pathways for the synthesis of other sugars (Samuel and Reeves 2003; DebRoy et al. 2005; Fratamico et al. 2005). For example, synthesis of dTDP-L-talose requires *rmlABC* (Samuel and Reeves 2003), and *rmlAB* together with *fdtA*, *fdtB*, and *fdtC* are involved in the synthesis of dTDP-3-acetamido-3,6-dideoxy-alpha-D-galactose (Feng et al. 2004*b*).

O-antigen processing genes

ORF7 showed 70% similarity to a polysaccharide biosynthesis protein in *Pectobacterium carotovorum* (Table 2) and also some similarity to a number of Wzx proteins in different microorganisms. ORF7 contained 12 transmembrane segments shown by HMMTOP analysis (Tusnády et al. 2001), which is typical of Wzx proteins. Therefore, ORF7 was proposed to encode a flippase and was named *wzx*. ORF10 shared 44% similarity to a hypothetical protein Gura_3055 in *Geobacter uraniireducens* Rf4, and HMMTOP analysis showed that it contained 9 transmembrane helices. Therefore, ORF10 was proposed to encode the O-antigen polymerase and was named *wzy*.

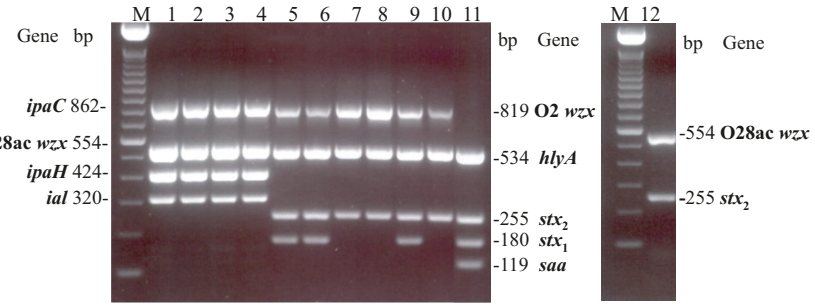
Transferase and other genes

ORF3 (glucose-1-phosphate thymidylyltransferase) is involved in the synthesis of dTDP-rhamnose. ORF8, ORF11,

Table 3. Open reading frames (ORFs) in the O-antigen gene cluster of *Escherichia coli* serogroup O28ac.

ORF	Proposed gene name	Location	No. of amino acids	Putative function	Most significant homology (GenBank accession No.)	% identity, % similarity
1	*wegS*	50–445	131	Glycerol-3-phosphate cytidyltransferase	WffW (*Shigella dysenteriae*) (ACA24835.1)	81, 92
2	*wzx*	456–1685	409	O-antigen flippase	Polysaccharide biosynthesis protein (*Actinobacillus succinogenes* 130Z) (YP_001344120.1)	35, 55
3	*wbeS*	1682–2860	392	Glycerol phosphotransferase	Teichoic acid biosynthesis protein RodC-related protein (*Methanothermobacter thermautotrophicus* strain Delta H) (NP_275508.1)	27, 48
4	*wzy*	2857–4236	459	O-antigen polymerase	Hypothetical protein PA2G_02559 (*Pseudomonas aeruginosa* 2192) (ZP_04935163.1)	30, 50
5	*wbeX*	4229–5239	186	Glycosyl transferase	Hypothetical protein plu3015 (*Photorhabdus luminescens* subsp. *laumondii* TTO1) (NP_930247.1)	35, 54
6	*wbeY*	5236–6153	305	Glycosyl transferase	Glycosyl transferase family protein (*Pectobacterium carotovorum* subsp. *brasiliensis*) (PBR1692)	52, 68
7	*wbeZ*	6150–7220	356	Acetyltransferase	Acetyltransferase (*Aliivibrio salmonicida* LFI1238) (YP_002265404.1)	41, 60
8	*glf*	7217–8362	381	UDP-galactopyranose mutase	Glf (*Shigella boydii*) (gb|ACD37076.1)	65, 82

Fig. 1. Agarose gel showing the results of multiplex PCR assays for the detection of enteroinvasive *Escherichia coli* serogroup O28ac, targeting the O28ac *wzx* gene (554 bp) and the *ial* (320), *ipaH* (424), and *ipaC* (862) genes, and for detection of STEC serogroup O2, targeting the O2 *wzx* gene (819 bp) and the stx_1 (180), stx_2 (255), *hly* (534), *eae* (384), and *saa* (119) genes. Lane M, molecular weight standards, 100 bp DNA ladder; lane 1, *E. coli* O28ac:H$^-$ CVD/EI-1; lane 2, *E. coli* O28ac:H$^-$ CVD/EI-2; lane 3, *E. coli* O28ac:H$^-$ CVD/EI-4; lane 4, *E. coli* O28ac:H$^-$ CVD/EI-5; lane 5, *E. coli* O2:H27 1.2740; lane 6, *E. coli* O2:H5 1.2741; lane 7, *E. coli* O2:H27 SJ5; lane 8, *E. coli* O2:H27 SJ6; lane 9, *E. coli* O2:H27 E18; lane 10 *E. coli* O2:H$^-$ E51; lane 11, *E. coli* O74 1.2803; lane 12, multiplex PCR assay using DNA from *E. coli* O28ac:H25 96-3286 and targeting the Shiga toxin-producing *E. coli* (STEC) virulence genes used for the STEC O2 multiplex PCR assays described previously.



and ORF12 (glycosyl transferases) encode for putative sugar transferases. ORF6 showed homology to genes that encode for glutamine amidotransferases. In *E. coli* O145, glutamine amidotransferase (WbuY) was proposed to be involved in the formation of an ammonia tunnel, which transfers ammonia to another protein (Feng et al. 2005). Since there are 3 glycosyl transferases, the structure of the O2 O-antigen gene cluster is likely a tetrasaccharide or pentasaccharide containing L-rhamnose; however, the structure and sugar composition of the O-unit in the O2 serotype strain has yet to be determined. The gene cluster also contains genes that encode for transposases (ORF13, ORF15) and a sulfatase (ORF16). The ORF14 showed similarity to an H repeat-associated protein. In *E. coli* K12, 5 *Rhs* loci associated with H-rpt sequences are dispersed within the genome and provide homology for RecA-dependent rearrangements (Zhao et al. 1993). Xiang et al. (1994) found that the *Salmonella* serogroup D2 O-antigen gene cluster was a hybrid of those of serogroup D1 and E1, and the D1 and E1 sequences were separated by an H-rpt element associated with *Rhs* loci. Thus, the serogroup D2 O-antigen gene cluster was produced through recombination events with serogroups D1 and E1. The presence of transposases (ORF13 and ORF15) and H repeat-associated protein (ORF14) between the genes encoding for the glycosyl transferase and sulfatase indicate the likelihood that in the case of the *E. coli* O2 O-antigen gene cluster, the sulfatase gene (ORF16, Table 2) might have been acquired through a DNA recombination event.

Sequence analysis of the *E. coli* O28ac O-antigen gene cluster

Sugar biosynthetic pathway genes

The structure of the *E. coli* O28ac O-antigen polysaccharide has been elucidated and contains *N*-acetylglucosamine, acetylgalactofuranose, and glycerol phosphate (Rundlöf et al. 1996). Genes that are responsible for *N*-acetylglucosamine and glycerol synthesis are regarded as housekeeping genes and were not present in the O28ac O-antigen gene cluster. ORF8 showed 82% similarity with UDP-galactopyranose mutase (Table 3), which catalyzes the reversible interconversion of UDP-galactopyranose and UDP-galactofuranose reaction (Nassau et al. 1996). The UDP-galactopyranose encoded by *glf* has been shown experimentally to be involved in galactofuranose biosynthesis (Nassau et al. 1996). Since the *E. coli* O28ac O-unit contains acetylgalactofuranose, it would be expected that the gene encoding UDP-galactopyranose mutase would be present in the O-antigen gene cluster of *E. coli* O28ac. The UDP-galactopyranose mutase gene was also present in the *E. coli* O16 (Marolda et al. 2004) and O52 (Feng et al. 2004*a*) O-antigen gene clusters. Therefore, ORF8 was identified and named *glf* based on its high homology with Glf.

O-antigen processing genes

ORF2 showed 55% similarity to the polysaccharide biosynthesis protein that was involved in polysaccharide biosynthesis in *Actinobacillus succinogenes* 130Z. In addition, ORF2 shows some similarity to a number of Wzx proteins in different microorganisms and contained 10 transmembrane helices as shown by HMMTOP analysis (Tusnády et al. 2001). Therefore, ORF2 was proposed to encode a flippase and was named *wzx*. ORF4 shared 50% similarity to a hypothetical protein PA2G_02559 in *Pseudomonas aeruginosa* 2192. ORF4 also showed similarity to secreted polysaccharide polymerases in *Bacillus cereus*. HMMTOP analysis showed that ORF3 contained 15 transmembrane helices; therefore, ORF4 was proposed to encode a polymerase and was named *wzy*.

Transferase genes

The *E. coli* O28ac O-unit is a tetrasaccharide, and there are 2 glycosyltransferases in the O-antigen gene cluster. ORF5 and ORF6, named *wbeX* and *wbeY*, respectively, both show homologies to glycosyltransferases. ORF1, named *wegS*, showed very high similarity (92%) to the WffW protein in *Shigella dysenteriae* that might have glycerol-3-phosphate cytidyltransferase activity, indicating that ORF1 could be a glycerol-3-phosphate cytidyltransferase. ORF3, named *wbeS*, showed 48% similarity to a teichoic acid biosynthesis protein that might have glycerophosphotransferase activity, indicating that ORF3 could be a glycerophosphotransferase. Since the polysaccharide of *E. coli* O28ac contains glycerol phosphate, we proposed that ORF1 and ORF3 might be responsible for the transfer of a phosphate to glycerol to form glycerol phosphate. ORF7 (*wbeZ*) showed 60% similarity to the acyltransferase in *Aliibrio salmonicida*. Since the *E. coli* O28 polysaccharide repeat unit contains acetylgalactofuranose, ORF7 could be responsible for the addition of the acetyl group to galactofuranose to form acetylgalactofuranose.

Specificity of the PCR assays and sequencing of the *E. coli* O42 O-antigen gene cluster

The PCR assays targeting the *E. coli* O2 *wzx* and *wzy* genes showed 100% specificity when tested using DNA from the *E. coli* and non-*E. coli* strains listed previously. There were no nonspecific PCR products detected. *Escherichia coli* and *Shigella* species often display similar antigenic components, complicating their differentiation and reliable identification. Cheasty and Rowe (1983) investigated the antigenic relationships between several EIEC serogroups, including O28ac and *Shigella* species. Cross-reactions were found between *E. coli* O112ac, O124, O136, O152, and O164 and *Shigella* species, indicating that they had similar O-antigens. However, there was no cross-reactivity of *E. coli* O28ac with *Shigella* species using specific antisera. In the current study, the PCR assays targeting the *E. coli* O28ac *wzx* and *wzy* genes were specific when tested using DNA from the *E. coli* and non-*E. coli* strains listed previously; however, 10 *E. coli* O42 field isolates and the *E. coli* O42 P11a standard strain produced PCR products using primers targeting the *E. coli* O28ac *wzx* and *wzy* genes. Therefore, the O-antigen gene cluster of *E. coli* O42:H37 P11a was sequenced to determine the similarity of the O28ac and O42 O-antigen gene clusters. When both complete sequences were compared, only 3 point mutations in the *wbeX* (ORF5) and *wbeY* (ORF6) genes were observed. All these mutations resulted in amino acid substitutions in the corresponding ORFs. An adenine to cytosine transversion in the *wbeX* gene (at nucleotide 5158) resulted in a Lys (DQ462205) to Asn (FJ539194) substitution, whereas a guanidine to adenine transition within the *wbeY* gene (nucleotide 5561) resulted in a Ser (DQ462205) to Asn (FJ539194) substitution. A thymine to cytosine transition in the *wbeY* gene (nucleotide 5996) resulted in a change of amino acid from Leu (DQ462205) to Ser (FJ539194). Thus *E. coli* O28ac and O42 likely represent one serogroup, since the sequences of the O-antigen gene clusters were very similar, and there was cross-reactivity in the antisera raised against O28ac and O42. However, structural analyses comparing the O-antigens of these 2 strains are warranted.

Multiplex PCR assays for detection of STEC O2 and EIEC O28ac strains

The multiplex PCR assay for detection of STEC O2 targeted the O2 *wzx* gene (819 bp product), and the STEC *saa* (119 bp), stx_1 (180 bp), stx_2 (255 bp), *eae* (384 bp), and *hlyA* (534 bp) virulence genes. For EIEC O28ac, the target genes were the O28ac *wzx* gene (554 bp) and the virulence genes, *ial* (320 bp), *ipaH* (424 bp), and *ipaC* (862 bp). All 4 O28ac strains were isolated from patients with diarrhea, and all were positive for the O28ac *wzx*, *ial*, *ipaH*, and *ipaC* genes (Fig. 1). Interestingly, the strain used for sequencing the O28ac O-antigen gene cluster in this study, clinical isolate *E. coli* O28ac:H25 96-3286 obtained from the CDC, harbored the O28ac *wzx* and *wxy* genes; however, it lacked the EIEC *ial*, *ipaH*, and *ipaC* genes. Instead, this strain carried the stx_2 gene, but not the *hlyA*, *eae*, *saa*, and stx_1 genes associated with STEC (Fig. 1, lane 12). Therefore, STEC

O28ac strains that cause human illness can be found, and it is possible that these strains might carry genes that encode for virulence factors other than *hlyA*, *eae*, and *saa*.

None of the STEC O2 strains examined in this study carried the *eae* gene that encodes for the intimin protein involved in the formation of attaching and effacing lesions or the *saa* gene that encodes for the STEC autoagglutinating adhesin (Fig. 1). Two of the strains were clinical isolates obtained from the CDC (O2:H27 SJ5 and O2:H27 SJ6). Therefore, the *eae* and *saa* genes might not be essential for causing diarrhea, and these strains might carry other virulence genes. All the O2 strains were positive for the O2 *wzx* gene and the *hlyA* genes, and they carried stx_2 or stx_1 and stx_2. A STEC serogroup O74 strain that was tested carried the *saa* gene, in addition to stx_1, stx_2, and *hlyA* (Fig. 1, lane 11).

Escherichia coli O2:H27 1.2740 and O28ac:H$^-$ CVD/E1-1 were detected in dog feces (2 g) inoculated with 10 colony-forming units after 24 h of enrichment in mTSB using the multiplex PCR assays. PCR products for the O2 *wzx* gene, stx_1, stx_2, and *hlyA* genes in samples inoculated with *E. coli* O2:H27 and PCR products for the O28ac *wzx* gene, *ial*, *ipaH*, and *ipaC* genes in samples inoculated with *E. coli* O28ac: H$^-$ were obtained (data not shown).

Conclusions

Previously, we developed a model microarray-based *E. coli* typing assay for rapid identification of different serogroups of *E. coli* in a single platform (Liu and Fratamico 2006). Sequences in the O-antigen gene clusters of several *E. coli* serogroups were spotted onto these microarrays, and specific signals were generated for each serogroup tested using labeled long PCR products or genomic DNA. Sequence information of additional O-antigen gene clusters, including those from *E. coli* O2 and O28ac, will help us to broaden our development of these diagnostic DNA microarrays for the accurate identification of *E. coli* isolates. We are also investigating other types of multiplex detection systems that can be used to identify the *E. coli* O serogroup, H type, and virulence genes simultaneously on a single diagnostic platform, which will facilitate identification and detection of pathogenic strains. The *wzx* and *wzy* genes of *E. coli* serogroups O2 and O28ac can be used as diagnostic markers in PCR assays for rapid identification of these serogroups as an alternative to sertoyping or to confirm serotyping results. Additionally, the multiplex PCR assays targeting the O2 and O28ac *wzx* and *wzy* genes in combination with STEC–EHEC and EIEC virulence genes can potentially be used to identify and detect pathogenic strains of these serogroups in food, fecal, and environmental samples.

Acknowledgements

The technical assistance of Dr. David Needleman, Dr. James Smith, Ms. Sue Lawler, Ms. Lori Bagi, and Ms. Elisabeth Roberts is gratefully acknowledged.

References

Bettelheim, K.A. 2007. The non-O157 Shiga-toxigenic (verocytotoxigenic) *Escherichia coli*; under-rated pathogens. Crit. Rev. Microbiol. **33**(1): 67–87. doi:10.1080/10408410601172172. PMID:17453930.

Boudeau, J., Glasser, A.-L., Masseret, E., Joly, B., and Darfeuille-Michaud, A. 1999. Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. Infect. Immun. **67**(9): 4499–4509. PMID:10456892.

Cheasty, T., and Rowe, B. 1983. Antigenic relationships between the enteroinvasive *Escherichia coli* O antigens O28ac, O112ac, O124, O136, O143, O144, O152, and O164 and *Shigella* O antigens. J. Clin. Microbiol. **17**(4): 681–684. PMID:6189854.

Cho, S., Fossler, C.P., Diez-Gonzalez, F., Wells, S.J., Hedberg, C.W., Kaneene, J.B., et al. 2007. Antimicrobial susceptibility of Shiga toxin-producing *Escherichia coli* isolated from organic dairy farms, conventional dairy farms, and county fairs in Minnesota. Foodborne Pathog. Dis. **4**(2): 178–186. doi:10.1089/fpd.2006.0074. PMID:17600485.

DebRoy, C., Roberts, E., Kundrat, J., Davis, M.A., Briggs, C.E., and Fratamico, P.M. 2004. Detection of *Escherichia coli* serogroups O26 and O113 by PCR amplification of the *wzx* and *wzy* genes. Appl. Environ. Microbiol. **70**(3): 1830–1832. doi:10.1128/AEM.70.3.1830-1832.2004. PMID:15006811.

DebRoy, C., Fratamico, P.M., Roberts, E., Davis, M.A., and Liu, Y. 2005. Development of PCR assays targeting genes in O-antigen gene clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. Appl. Environ. Microbiol. **71**(8): 4919–4924. doi:10.1128/AEM.71.8.4919-4924.2005. PMID:16085897.

Dho-Moulin, M., and Fairbrother, J.M. 1999. Avian pathogenic *Escherichia coli* (APEC). Vet. Res. **30**(2–3): 299–316. PMID:10367360.

Fasano, A., Kay, B.A., Russell, R.G., Maneval, D.R., Jr., and Levin, M.M. 1990. Enterotoxin and cytotoxin production by enteroinvasive *Escherichia coli*. Infect. Immun. **58**(11): 3717–3723. PMID:2228242.

Feng, L., Senchenkova, S.N., Yang, J., Shashkov, A.S., Tao, J., Guo, H., et al. 2004*a*. Synthesis of the heteropolysaccharide O antigen of *Escherichia coli* O52 requires an ABC transporter: structural and genetic evidence. J. Bacteriol. **186**(14): 4510–4519. doi:10.1128/JB.186.14.4510-4519.2004. PMID:15231783.

Feng, L., Wang, W., Tao, J., Guo, H., Krause, G., Beutin, L., and Wang, L. 2004*b*. Identification of *Escherichia coli* O114 O-antigen gene cluster and development of an O114 serogroup-specific PCR assay. J. Clin. Microbiol. **42**(8): 3799–3804. doi:10.1128/JCM.42.8.3799-3804.2004. PMID:15297533.

Feng, L., Senchenkova, S.N., Tao, J., Shashkov, A.S., Liu, B., Shevelev, S.D., et al. 2005. Structural and genetic characterization of enterohemorrhagic *Escherichia coli* O145 O antigen and development of an O145 serogroup-specific PCR assay. J. Bacteriol. **187**(2): 758–764. doi:10.1128/JB.187.2.758-764.2005. PMID:15629947.

Fratamico, P.M., Briggs, C.E., Needle, D., Chen, C.Y., and DebRoy, C. 2003. Sequence of the *Escherichia coli* O121 O-antigen gene cluster and detection of enterohemorrhagic *E. coli* O121 by PCR amplification of the *wzx* and *wzy* genes. J. Clin. Microbiol. **41**(7): 3379–3383. doi:10.1128/JCM.41.7.3379-3383.2003. PMID:12843098.

Fratamico, P.M., DebRoy, C., Strobaugh, T.P., Jr., and Chen, C.Y. 2005. DNA sequence of the *Escherichia coli* O103 O antigen gene cluster and detection of enterohemorrhagic *E. coli* O103 by PCR amplification of the *wzx* and *wzy* genes. Can. J. Microbiol. **51**(6): 515–522. doi:10.1139/w05-049. PMID:16121232.

Gyles, C.L. 2007. Shiga toxin-producing *Escherichia coli*: an overview. J. Anim. Sci. **85**(13 Suppl.): E45–E62. doi:10.2527/jas.2006-508. PMID:17085726.

Islam, M.A., Heuvelink, A.E., de Boer, E., Sturm, P.D., Beumer, R.R., Zwietering, M.H., et al. 2007. Shiga toxin-producing *Escherichia coli* isolated from patients with diarrhoea in Bangladesh. J.

Med. Microbiol. **56**(3): 380–385. doi:10.1099/jmm.0.46916-0. PMID:17314370.

Jelacic, J.K., Damrow, T., Chen, G.S., Jelacic, S., Bielaszewska, M., Ciol, M., et al. 2003. Shiga toxin-producing *Escherichia coli* in Montana: bacterial genotypes and clinical profiles. J. Infect. Dis. **188**(5): 719–729. doi:10.1086/376999. PMID: 12934188.

Karmali, M.A. 2003. The medical significance of Shiga toxin-producing *Escherichia coli* infections: an overview. Methods Mol. Med. **73**: 1–7. PMID:12375423.

Liu, Y., and Fratamico, P. 2006. *Escherichia coli* O antigen typing using DNA microarrays. Mol. Cell. Probes, **20**(3-4): 239–244. doi:10.1016/j.mcp.2006.01.001. PMID:16537102.

Marolda, C.L., Vicarioli, J., and Valvano, M.A. 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. Microbiology, **150**(12): 4095–4105. doi:10.1099/mic.0.27456-0. PMID:15583162.

Nassau, P.M., Martin, S.L., Brown, R.E., Weston, A., Monsey, D., McNeil, M.R., and Duncan, K. 1996. Galactofuranose biosynthesis in *Escherichia coli* K-12: identification and cloning of UDP-galactopyranose mutase. J. Bacteriol. **178**(4): 1047–1052. PMID:8576037.

Nataro, J.P., and Kaper, J.B. 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11**(1): 142–201. PMID:9457432.

Ørskov, I., Ørskov, F., Jann, B., and Jann, K. 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol. Rev. **41**(3): 667–710. PMID:334154.

Pal, T., Al-Sweih, N.A., Herpay, M., and Chugh, T.D. 1997. Identification of enteroinvasive *Escherichia coli* and *Shigella* strains in pediatric patients by an IpaC-specific enzyme-linked immunosorbent assay. J. Clin. Microbiol. **35**(7): 1757–1760. PMID: 9196187.

Pal, A., Ghosh, S., Ramamurthy, T., Yamasaki, S., Tsukamoto, T., Bhattacharya, S.K., et al. 1999. Shiga-toxin producing *Escherichia coli* from healthy cattle in a semi-urban community in Calcutta, India. Indian J. Med. Res. **110**: 83–85. PMID:10612908.

Parsot, C. 2005. *Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors. FEMS Microbiol. Lett. **252**(1): 11–18. doi:10.1016/j.femsle.2005.08.046. PMID:16182469.

Paton, A.W., and Paton, J.C. 1998. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for *stx*$_1$, *stx*$_2$, *eaeA*, enterohemorrhagic *E. coli hlyA, rfb*O111, and *rfb*O157. J. Clin. Microbiol. **36**(2): 598–602. PMID:9466788.

Paton, A.W., and Paton, J.C. 2002. Direct detection and characterization of Shiga toxigenic *Escherichia coli* by multiplex PCR for *stx*$_1$, *stx*$_2$, *eae, ehxA*, and *saa*. J. Clin. Microbiol. **40**(1): 271–274. doi:10.1128/JCM.40.1.271-274.2002. PMID:11773130.

Paton, A.W., Srimanote, P., Woodrow, M.C., and Paton, J.C. 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. Infect. Immun. **69**(11): 6999–7009. doi:10.1128/IAI.69.11.6999-7009.2001. PMID:11598075.

Possé, B., De Zutter, L., Heyndrickx, M., and Herman, L. 2008. Quantitative isolation efficiency of O26, O103, O111, O145 and O157 STEC serotypes from artificially contaminated food and cattle faeces samples using a new isolation protocol. J. Appl. Microbiol. **105**(1): 227–235. doi:10.1111/j.1365-2672.2008.03739.x. PMID:18284481.

Reeves, P.P., and Wang, L. 2002. Genomic organization of LPS-specific loci. Curr. Top. Microbiol. Immunol. **264**(1): 109–135. PMID:12014174.

Reeves, P.R., Hobbs, M., Valvano, M.A., Skurnik, M., Whitfield, C., Coplin, D., et al. 1996. Bacterial polysaccharide synthesis and gene nomenclature. Trends Microbiol. **4**(12): 495–503. doi:10.1016/S0966-842X(97)82912-5. PMID:9004408.

Reischl, U., Youssef, M.T., Kilwinski, J., Lehn, N., Zhang, W.L., Karch, H., and Strockbine, N.A. 2002. Real-time fluorescence PCR assays for detection and characterization of Shiga toxin, intimin, and enterohemolysin genes from Shiga toxin-producing *Escherichia coli*. J. Clin. Microbiol. **40**(7): 2555–2565. doi:10.1128/JCM.40.7.2555-2565.2002. PMID:12089277.

Rundlöf, T., Weintraub, A., and Widmalm, G. 1996. Structural studies of the enteroinvasive *Escherichia coli* (EIEC) O28 O-antigenic polysaccharide. Carbohydr. Res. **291**: 127–139. PMID:8864227.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. 2000. Artemis: sequence visualization and annotation. Bioinformatics, **16**(10): 944–945. doi:10.1093/bioinformatics/16.10.944. PMID:11120685.

Samuel, G., and Reeves, P. 2003. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr. Res. **338**(23): 2503–2519. doi:10.1016/j.carres.2003.07.009. PMID:14670712.

Smith, J.E., and Fratamico, P.M. 2005. Diarrhea-inducing *Escherichia coli*. *In* Foodborne pathogens: microbiology and moleclar biology. Caister Academic Press, Norfolk, UK. pp. 357–382.

Thong, K.L., Hoe, S.L., Puthucheary, S.D., and Yasin, R.M. 2005. Detection of virulence genes in Malaysian *Shigella* species by multiplex PCR assay. BMC Infect. Dis. **5**(1): 8. doi:10.1186/1471-2334-5-8. PMID:15707504.

Tornieporth, N.G., John, J., Salgado, K., de Jesus, P., Latham, E., Melo, M.C.N., et al. 1995. Differentiation of pathogenic *Escherichia coli* strains in Brazilian children by PCR. J. Clin. Microbiol. **33**(5): 1371–1374. PMID:7615758.

Tusnády, G.E., and Simon, I. 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics, **17**(9): 849–850. doi:10.1093/bioinformatics/17.9.849. PMID:11590105.

Wang, L., Briggs, C.E., Rothemund, D., Fratamico, P., Luchansky, J.B., and Reeves, P.R. 2001. Sequence of the *E. coli* O104 antigen gene cluster and identification of O104 specific genes. Gene, **270**(1–2): 231–236. doi:10.1016/S0378-1119(01)00471-1. PMID:11404020.

Wong, C.S., Jelacic, S., Habeeb, R.L., Watkins, S.L., and Tarr, P.I. 2000. The risk of the hemolytic–uremic syndrome after antibiotic treatment of *Escherichia coli* O157:H7 infections. N. Engl. J. Med. **342**(26): 1930–1936. doi:10.1056/NEJM200006293422601. PMID:10874060.

Xiang, S.H., Hobbs, M., and Reeves, P.R. 1994. Molecular analysis of the *rfb* gene cluster of a group D2 *Salmonella enterica* strain: evidence for its origin from an insertion sequence-mediated recombination event between group E and D1 strains. J. Bacteriol. **176**(14): 4357–4365. PMID:8021222.

Zhao, S., Sandt, C.H., Feulner, G., Vlazny, D.A., Gray, J.A., and Hill, C.W. 1993. *Rhs* elements of *Escherichia coli* K-12: complex composites of shared and unique components that have different evolutionary histories. J. Bacteriol. **175**(10): 2799–2808. PMID:8387990.