ISSN 1466-2523

Animal Health Research Reviews **19**(1); 1–16
doi:10.1017/S1466252317000093

Review

Molecular serogrouping of *Escherichia coli*

Chitrita DebRoy[1]*, Pina M. Fratamico[2] and Elisabeth Roberts[1]

[1] *Department of Veterinary and Biomedical Sciences, E. coli Reference Center, The Pennsylvania State University, University Park, PA, USA*
[2] *U.S. Department of Agriculture, Eastern Regional Research Center, Agricultural Research Service, Wyndmoor, PA, USA*

Received 17 May 2017; Accepted 21 September 2017;
First published online 22 February 2018

Abstract

O-antigens present on the surface of *Escherichia coli* provide antigenic specificity for the strain and are the main components for O-serogroup designation. Serotyping using O-group-specific antisera for the identification of *E. coli* O-serogroups has been traditionally the gold-standard for distinguishing *E. coli* strains. Knowledge of the O-group is important for determining pathogenic lineage, classifying *E. coli* for epidemiological studies, for determining virulence, and for tracing outbreaks of diseases and sources of infection. However, serotyping has limitations, as the antisera generated against each specific O-group may cross-react, many strains are non-typeable, and others can autoagglutinate or be rough (lacking an O-antigen). Currently, the nucleotide sequences are available for most of the 187 designated *E. coli* O-groups. Public health and other laboratories are considering whole genome sequencing to develop genotypic methods to determine O-groups. These procedures require instrumentation and analysis that may not be accessible and may be cost-prohibitive at this time. In this review, we have identified unique gene sequences within the O-antigen gene clusters and have targeted these genes for identification of O-groups using the polymerase chain reaction. This information can be used to distinguish O-groups by developing other platforms for *E. coli* diagnostics in the future.

Keywords: serotyping, genotyping, PCR, sequencing, O-antigen gene cluster.

Introduction

Pathogenic strains of *Escherichia coli* can cause many different types of diseases in people and animals. In people, they cause diarrhea, hemorrhagic colitis, hemolytic uremic syndrome, urinary tract infections, and neonatal meningitis, among other types of illnesses. In cattle, they produce calf scours and mastitis; in pigs they are the etiologic agents for post-weaning diarrhea and edema disease; and in chickens they cause peritonitis and airsaccullitis. In horses, extra-intestinal pathogenic *E. coli* (ExPEC) cause bronchopneumonia, hemolytic uremic syndrome, and encephalopathy (DebRoy *et al.*, 2008; Dickinson *et al.*, 2008). In dogs, cats, and tigers, ExPEC strains have been reported to cause bronchopneumonia (Handt *et al.*, 2003; Sura *et al.*, 2007; Carvallo *et al.*, 2010).

*Corresponding author. E-mail: rcd3@psu.edu

The *E. coli* pathotypes have been classified based on the type of diseases they may cause. The intestinal *E. coli*, that cause diarrhea are characterized by the virulence attributes they carry, are divided into a number of major groups, including the Shiga toxin-producing *E. coli* (STEC) of which enterohemorrhagic *E. coli* (EHEC) are a subset, as well as enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC), and enteroaggregative *E. coli* (EAggEC) (Kaper *et al.*, 2004). ExPEC, strains that cause diseases outside the intestinal tract, are associated with neonatal meningitis, urinary tract infection, and septicemia in human beings, as well as pneumonia and urinary tract infection in animals (Russo and Johnson, 2003; Smith *et al.*, 2007).

E. coli strains can be distinguished by their O- and H-antigens that are located in the outer membrane of *E. coli*. Lipopolysaccharides (LPS) that are part of the outer membrane of Gram-negative bacteria are composed of three parts: lipid A,

the core region, and the O-antigen polysaccharide that is specific for each O-serogroup and is exposed on the cell surface (Stenutz *et al.*, 2006; Bos *et al.*, 2007). LPS, as a pathogenic entity, is responsible for stimulating the innate immune system in the host that recognizes a bioactive component known as the Toll-like receptor 4(TLR4), which is present on the surface of cells of the innate immune system (Poltorak *et al.*, 1998; Akira *et al.*, 2006). O-antigens are known to contribute to virulence as specific clones of some serotypes are associated with certain diseases (Achtman and Pluschke, 1986). Outbreaks of gastrointestinal illness caused by EHEC O157:H7 were first described in 1983 (Riley *et al.*, 1983; Remis *et al.*, 1984). Among ExPEC strains, O4 and O6 serogroups have been implicated in causing fetal pneumonia in dogs, cats, and tigers (Handt *et al.*, 2003; Sura *et al.*, 2007; Carvallo *et al.*, 2010). O-groups provide important pathotype information, and therefore, are considered essential for outbreak investigation, risk management, epidemiological studies, and treatment options.

Serological classification

Kaufmann (1943, 1944, 1947) classified *E. coli* into 20 O-groups on the basis of an antigenic scheme that he developed by using boiled cultures of *E. coli* for antisera production. In this method, antisera are generated against *E. coli* strains belonging to each O-group by injecting the boiled cultures into rabbits. An agglutination reaction that occurs with one of these antisera with an unknown strain (that is also boiled) represents the O-group of the unknown strain. This serotyping scheme has been used as the 'gold standard' for classification of *E. coli*. Identification of a combination of three surface antigens, O-antigen, H-flagellar antigen, and K-capsular antigen, was initially conceived for subtyping *E. coli*. However, most laboratories lack capabilities to type K-antigens, and therefore, serotyping based on O- and H-typing has been considered as the standard. In the past 70 years, *E. coli* strains have been classified into serogroups O1–O187, except for six O-groups that have not been designated (O31, O47, O67, O72, O94, and O122) (Ørskov and Ørskov, 1984; Scheutz *et al.*, 2004). Although four O-groups have been sub-grouped previously as O18ab/O18ac, O28ab/O28ac, O112ab/O112ac, and O125ab/O125ac, we recently reported O18ab/O18ac, O112ab/O112ac, and O125ab/O125ac to have identical O-AGCs (DebRoy *et al.*, 2016). Eight O groups were previously designated as OX1–OX8 by Ewing and Tatum (1956); however, OX1 is now designated as O170, OX2 as O169, OX3 as O174, OX4 as O146, OX5 as O168, OX6 as O171, and OX7 as O175 (Scheutz *et al.*, 2004). Eleven other OX groups have been informally used by many laboratories including ours. Some of the OX groups have been found to be 98–99% related to the designated O-groups. These are OX6/O168 (OX6 strains may vary, and we found OX6 to be similar to O168), OX9/O184, OX10/O159, OX19/O11, OX21/O163, OX38/O128, and OX43/O19. OX groups OX13, OX18, OX25, OX28, and OX38 have been found to be unique (DebRoy *et al.*, 2016). Currently, the Staten Serum Institut in Denmark (http://www.ssi.dk/English.aspx) has listed serogroups up to O188. Additional O-groups have been since discovered (Iguchi *et al.*, 2016) that have not yet been designated.

Traditional serotyping still remains one of the most comprehensive and simple methods (DebRoy *et al.*, 2011a, b); however, the method is not reliable as many strains cross-react with other O-group antisera, and results are equivocal if the organism is encapsulated, rough, does not carry LPS, or is autoagglutinating. In our experience of serotyping thousands of *E. coli* strains, about 25% of the strains cannot be assigned to any O-group and are designated either negative, rough, autoagglutinating, or as belonging to multiple serogroups. Similar observations have also been reported by others (Lacher *et al.*, 2014). The O-antigens are highly immunogenic, and variability among the groups provides the basis for serotyping and classification of *E. coli*. Carbohydrate analysis methods have been valuable in the identification of the structures of numerous O-antigens (MacLean *et al.*, 2006; Chafchaouni-Moussaoui *et al.*, 2011).

Biosynthesis of O-antigens

The genes within the O-antigen gene clusters (O-AGC) that are responsible for the biosynthesis of O-antigens are located on the chromosome between the two housekeeping genes, *galF* and *gnd*, and can be amplified by polymerase chain reaction (PCR) targeting of a 39-bp JUMPstart region upstream of the O-AGC and *gnd* (Batchelor *et al.*, 1992; Bastin *et al.*, 1993; Hobbs and Reeves, 1994). O-antigens are composed of repeat units of oligosaccharides (O-unit) that are sugar residues varying considerably in number of units, arrangement, and linkage of the O-units, thus making the O-antigen the most variable region of the bacterial cell. There are three major gene classes that compose the O-AGC: the nucleotide sugar synthesis genes, sugar transferase genes, and O-unit processing genes (Bronner *et al.*, 1994; Keenleyside and Whitfield, 1996; Daniels *et al.*, 1998; Linton and Higgins, 1998; Samuel and Reeves, 2003; Liu *et al.*, 2008a; Wang *et al.*, 2010). Nucleotide sugar synthesis genes encode for proteins that are responsible for sugar synthesis, specific to certain O-groups, and some of these genes are conserved among various O-AGCs. There are four genes, *rmlB* (dTDP-glucose 4, 6-dehydratase), *rmlD* (dTDP-4-dehydrorhamnose reductase), *rmlA* (glucose-1-phosphate thymidylyltransferase), and *rmlC* (dTDP-4-dehydrorhamnose 3,5-epimerase) that are involved in the biosynthesis of dTDP-L-rhamnose and that exhibit relatedness in different O-groups and often group together in the cluster (DebRoy *et al.*, 2016). In 56 of the O-AGCs, the *manB* gene encoding for phosphomannomutase and *manC* encoding for mannose-1-phosphate guanyltransferase responsible for the biosynthesis of GDP-D-mannose (Samuel and Reeves, 2003) are present. Two other genes involved in biosynthesis of UDP-L-FucNAc derived from UDP-GlcNAc, *fnlA* (UDP-glucose epimerase) and *fnlC* (UDP-N-acetylglucosamine 2-epimerase), were identified in 15 O-AGCs. VioA and VioB that carry out transamination of dTDP-6-deoxy-D-xylo-4-hexulose to dTDP-4-amino-4,6-dideoxy-D-glucose (VioN) and VioB that N-acetylates VioN to dTDP-VioNAc were found to be associated with the O-AGC

https://doi.org/10.1017/S1466252317000093 Published online by Cambridge University Press

for O39, O49, and O116 (DebRoy *et al.*, 2016). Sugar transferase genes encoding for glycosyl transferase proteins, transfer various precursor sugars to form oligosaccharides on the carrier lipid undecaprenyl phosphate (UndP), situated in the inner membrane facing the cytoplasm (Reeves, 1994). Highly specific and varied glycosyl transferases are associated with each of the O-AGCs. The O-antigen processing proteins, Wzx (O-antigen flippase), and Wzy (O-antigen polymerase) are involved in translocation of the O-units across the membrane and in polymerization of the sugar units, respectively. The O-unit is synthesized by sequential transfer of sugars and other constituents to the first sugar, which is then translocated and flipped across the membrane by the protein Wzx to the periplasmic face. These are further polymerized by Wzy (Mulford and Osborn, 1983; McGrath and Osborn, 1991; Reeves and Wang, 2002; Liu *et al.*, 2008a). The number of O-units in the final O-antigen is regulated by Wzz. The Wzx and Wzy proteins are both hydrophobic and have about 12 predicted transmembrane regions, and a long cytoplasmic region is predicted between the transmembrane domains in Wzy. O-groups O8, O9, O52, O60, O89, O92, O95, O97, O99, O101, and O162 are ABC transporter dependent for O-antigen processing and carry *wzt* and *wzm* genes (DebRoy *et al.*, 2016). The transportation processing is brought about by the ABC transporter protein, Wzm, which is responsible for export, and Wzt is the ATP-binding component.

Relatedness among O-AGCs

The nucleotide sequences of all designated 187 O-AGCs have been recently reported (Iguchi *et al.*, 2015a; DebRoy *et al.*, 2016). Out of these O-groups, 13 sets exhibited 98–99.9% identical nucleotide sequences. These are O2/O50, O13/O129/O135, O17/O44/O73/O77/O106, O18ab/O18ac, O42/O28ac, O46/O134, O62/O68, O90/O127, O101/O162, O107/O117, O118/O151, O123/O186, O124/O164, and O125ab/O125ac (Table 1) (DebRoy *et al.*, 2016). O-groups that carry identical nucleotide sequences may or may not cross-react in serological reactions. There were three nucleotide differences between the O-AGCs of O118 and O151, and there was an insertion sequence in O62 located between the *rmlA* and *rmlC* genes, but otherwise, the O-AGCs of O62 and O68 were similar. However, these differences may be responsible for the absence of serological cross-reactions with these serogroups (Liu *et al.*, 2008b, 2015). There are other O-groups that have identical nucleotide sequences that do not exhibit any serological cross-reactions. Further studies on the expression of genes of the O-antigens may shed light on production of the O-antigen and the immune response to these antigens.

Molecular genotyping of O-groups

Molecular serogrouping of O-antigens entails determining the O-group by molecular methods. Several molecular methods have been used to distinguish different O-groups, including restriction fragment length polymorphism (RFLP) (Coimbra, *et al.*, 2000; Beutin *et al.*, 2005a, b; Moreno *et al.*, 2006; Abbadi and Strockbine, 2007), high throughput real-time PCR (Bugarel *et al.*, 2010; Delannoy *et al.*, 2013; Tseng *et al.*, 2014), multiplex-PCR (Fratamico and DebRoy, 2010; Botkin *et al.*, 2012; Doumith *et al.*, 2012; Fratamico *et al.*, 2014; Iguchi *et al.*, 2015b), microarray (Hegde *et al.*, 2013; Lacher *et al.*, 2014; Patel *et al.*, 2016), flow cytometry (Liu and Fratamico, 2006; Hegde *et al.*, 2012; NeoSEEK™ (PCR-Mass spectroscopy) (Stromberg *et al.*, 2015), and targeted amplicon sequencing for single-nucleotide-polymorphism genotyping (Ison *et al.*, 2016). Most of these methods have been applied to distinguish pathotypes belonging to certain O-groups. A multiplex PCR method using 162 PCR primer pairs for 141 O-groups was developed by Iguchi *et al.* (2015a, b). Although currently a whole genome sequencing (WGS) approach to determine O-groups is trending (Norman *et al.*, 2015), we have developed PCR for each of the O-groups for laboratories that are still not ready for WGS because of cost or lack of analysis tools.

Target genes for detection of O-groups

The O-antigen gene clusters are generally highly heterogeneous with considerable nucleotide differences among the clusters. However, the predominant genes in the clusters are those involved in nucleotide sugar synthesis, and encoding for synthetases, transferases, epimerases, and other enzymes. Although the genes encoding specific proteins may have similarities in functions among the O-groups, the DNA sequences for some may be quite variable. In most of the clusters, two genes, encoding the O-antigen flippase (*wzx*) and the O-antigen polymerase (*wzy*) responsible for O-antigen translocation across the membrane and polymerization, are relatively unique, and therefore have been used as targets in genetic-based methods for serogroup identification. For developing PCR tests for detecting specific O-groups, primers targeting unique genes (usually *wzx* and *wzy*) in the O-AGCs were designed. In 11 O-AGCs, O8, O9, O52, O60, O89, O92, O95, O97, O99, O101, and O162, are ATP-binding Cassette (ABC) transporter integral membrane proteins that are dependent on *wzt* and *wzm* genes that assist in the O-antigen processing and transportation process. The O-AGCs for O14 and O57 could not be mapped (Iguchi *et al.*, 2015a, b; DebRoy *et al.*, 2016). O14 strains are rough and cannot be serotyped (Ørskov *et al.*, 1977), and an O-AGC could not be identified in O57. Therefore, primers for these two O-groups were not designed. We have put together all the primers that have been developed by our group and others, and provided a practical method for detecting the O-serogroup of *E. coli* with PCR (Table 2).

PCR primers, methodology, and specificity

E. coli strains belonging to the targeted serogroup were grown overnight on tryptic soy agar (TSA) plates at 37°C. A single colony was picked and resuspended in 150 µl of distilled water and heated at 100°C for 10 min. The suspension was centrifuged at

https://doi.org/10.1017/S1466252317000093 Published online by Cambridge University Press

Table 1. *Escherichia coli* O-AGCs that are 98–99% identical in nucleotide sequence

Set #	Computational analysis	Serological cross-reaction	Positive PCR reaction with target gene of the O-group listed
1	O2/O50	No	Yes
2	O13/O129/O135	Yes	Yes
4	O17/O44/O73/O77/O106	Yes, but O44 is distinct	Yes
5	O42/O28ac	Insufficient data	Yes
6	O46/O134	No	Yes
7	O62/O68	O68 cross-reacts with O62 but not vice versa	Yes
8	O90/O127	O90 cross-reacts with O127 but not vice versa	Yes
9	O101/O162	O101 cross-react with O162 but not vice versa	Yes
10	O107/O117	Yes	Yes
11	O118/O151	No	Yes
12	O123/O186	Yes	Yes
13	O124/O164	Occasional cross-reaction	Yes

12,000 ×***g***, and the supernatant containing genomic DNA was used for the PCR reaction. Primers were designed for the genes listed in Table 2 for each serogroup. They were synthesized from Integrated DNA Technologies, (Coralville, IA) and reconstituted to 100 µM concentration using nuclease-free water from the manufacturer. PCR was conducted in 96-well microtiter plates. Each well contained 300 nM of primer pairs and Denville ChoiceTaq Blue working mix (22 µl) (Denville Scientific, MA, USA). Template DNA (3 µl) was added to the mix, and the PCR was carried out using an Eppendorf Mastercycler Pro (Eppendorf, Germany) consisting of initial denaturation at 94°C for 2 min, followed by 25 cycles of 94°C for 10 s and using annealing temperatures listed in Table 2 for each O-group, followed by extension at 72°C for 15 s and the final extension at 72°C for 1 min. Reaction mixtures without template DNA and without primers served as negative controls. The amplification products were subjected to electrophoresis in 1% agarose gels at 200 V for 1 h. The gels were stained with ethidium bromide, and PCR products were visualized under UV light. Positive samples were identified based on the presence of bands of the expected sizes and compared with results using DNA from control reference strains belonging to the targeted serogroup.

PCR assays were developed using the reference O-standard strains, routinely used for serotyping, that were obtained from the World Health Organization Collaborating Center for Reference and Research on *Escherichia* and *Klebsiella* based at the Statens Serum Institut. The reference strains belong to serogroups O1 to O187 except for O31, O47, O67, O72, O94, and O122 serogroups that are not designated (Ørskov *et al.*, 1977). The specificity of the PCR assays for each serogroup was tested by using field strains or clinical isolates (see Table 2) belonging to the targeted serogroup and obtained from the *E. coli* collection at the *E. coli* Reference Center. The percentage of the strains that exhibited positive results by PCR was noted, and the ones that exhibited negative results were re-serotyped to make sure that the serogroup was correctly listed in the database. In addition, randomly selected field strains ($n = 50$) belonging to serogroups other than the targeted serogroup and 21 other bacterial species, including *Citrobacter freundii*, *Enterobacter cloacae*, *Enterococcus aerogenes*, *E. faecalis*, *Hafnia alvei*, *Klebsiella pneumoniae, Lactococcus lactis*, *Listeria monocytogenes*, *Proteus vulgaris*, *Pseudomonas aeruginosa*, *Salmonella enterica subsp.* Anatum, Arizona, Choleraesuis, and Typhimurium, *Serratia marcescens*, *Shigella boydii*, *Shigella flexneri*, *Staphylococcus aureus*, *Streptococcus pyrogenes*, *and Yersinia enterocolitica* were tested for specificity.

Detection of the O-antigen for all serogroups

PCR assays for most of the O-groups were between 80 and 100% specific for the targeted O-type. Although primers designed for O19, O20, and O97 exhibited positive results in PCR for the reference strains, reflecting that the PCR is designed correctly, the specificities among the clinical isolates belonging to these O-groups were very low. All the clinical isolates or field strains that exhibited negative results by PCR were re-serotyped and found to belong to the targeted O-group sometimes, and sometimes they exhibited a different O-group. The serotyping results, therefore, may be flawed. For example, we tested 44 isolates that were all serotyped as O20; however, none of them were positive by PCR targeting *wzx* or *wzy* (GenBank no. 778793.1). Sequencing O-AGCs of 44 O20 strains showed that they belonged to other serogroups such as O8, O48, O86, O111, O118, O131, O160, O178, and O-non-typeable (Dr D. W. Lacher, personal communication). Therefore, O20 antigens are not always specific for the O20 genotype. Although Iguchi *et al.* (2015a, b) reported primer sequences for detecting O20, these primers also did not show any positive results for O20 clinical isolates. Other O-groups that exhibited relatively low specificity among the field strains were O32 (62%), O34 (53%), O64 (21%), O93 (30%), and O152 (58%). The percentage of strains that exhibit a positive response by PCR may improve in these O-groups if more clinical/field isolates are tested, because the serotyping results for these O-groups are not always accurate. Thus, when we compare classical serotyping to molecular serotyping results, disparities are observed. Clinical isolates of certain O-groups, may not exhibit positive results with PCR of the targeted O-group. Further investigation as to why they are not specific for the O-type designated by serotyping needs to be determined for each individual isolate. This can be due to mutations, insertions,

https://doi.org/10.1017/S1466252317000093 Published online by Cambridge University Press

Table 2. Primer sequences for the detection of O-groups by PCR

O-group	Gene		Primers	Temp (C)	Amplicon bp	Specificity for field isolates (same O-type)		GenBank	Reference
						n = pos	% pos		
O1	*wzx*	F	GTGAGCAAAAGTGAAATAAGGAACG	55°	1098			GU299791	Li *et al.* (2010)
		R	CGCTGATACGAATACCATCCTAC						
O2/O50	*wzx*	F	TGGCCTTGTTCGATATACTGCGGA	56°	819			EU549863	Fratamico *et al.* (2010)
		R	TCACGAGCTGAGCGAAACTGTTCA						
O3	*wzx*	F	TGCGTGTCCATTCTTTCTGC	55°	376			EU694097	Ren *et al.* (2008)
		R	GAATCACCTGCAACCCCTGA						
O4	*wzx*	F	TGAAACAGCAGTGCGTGCATTCTC	59°	832	100	100	AY568960	This study
		R	GATTGCCGTGCCAATTATGCCGAA						
O5	*wzx*	F	CATACTGGCTCGCTGCCTCATT	58°	650	43	70	KP710588	This study
		R	GCCAAACAGCGGCGACATGTTA						
O6	*wzy*	F	GGATGACGATGTGATTTTGGCTAAC	55°	783	10	100	AJ426423	Grozdanov *et al.* (2002)
		R	TCTGGGTTTGCTGTGTATGAGGC						
O7	*wzx*	F	CTATCAAAATACCTCTGCTGGAATC	50°	610	10	100	AF125322	Li *et al.* (2010)
		R	TGGCTTCGAGATTAAACCTATTCCT						
O8	orf469	F	CCAGAGGCATAATCAGAAATAACAG	50°	448	10	100	AB811598	Li *et al.* (2010)
		R	GCAGAGTTAGTCAACAAAAGGTCAG						
O9	*wzm*	F	TGTTCCGGGGTGTTAACCTG	57°	197	10	100	D43637	This study
		R	TTGCCCGTGACGATCAGAAA						
O10	*wzx*	F	CAAGGACCGTTCCGTAGCAT	55°	850	10	100	KJ755557	This study
		R	AACACGCAATCCCTTGACGA						
O11/OX19	*wzx*	F	ACGGATTTAACTACTCAGCGCCCAA	57°	281	10	100	HQ388393	Li *et al.* (2011)
		R	GCGTGCGGAAAACATACTTATCGCT						
O12	*wzx*	F	GCCAGAGGTGGCACCAATTA	55°	236	13	77	KJ755558	This study
		R	CCCCTAGCGCACCTAACTTT						
O13/O129/O135	*wzx*	F	CGGGAGAGCAGTGTTTCCAA	55°	364	11	100	EU296422	This study
		R	TATTGCGGCACCCAGTAACC						
O15	*wzy*	F	TCTTGTTAGAGTCATTGGTGTATCG	55°	183	12	100	AY647261	Li *et al.* (2010)
		R	ATAAAACGAGCAAGCACCACACC						
O16	*wzx*	F	ACGTCGATGACCTTGCTTGT	59°	809	10	100	AB811601	This study
		R	AAACTCAACAAACCCGCAGC						
O17/O44/O73/ O77/O106	*wzx*	F	GCAACGTGTACGATCTTCGC	55°	693	10 (O77)	100	DQ000313	This study
		R	AGCAGCACCTTTAACACCGT						
O18ab/O18ac	*wzy*	F	GTTCGGTGGTTGGATTACAGTTAG	55°	551	10	100	GU299793	Li *et al.* (2010)
		R	CTACTATCATCCTCACTGACCACG						
O19	*wzy*	F	ATAAGCGCGAGCTTAGCTCTT		389			AB811604	Iguchi *et al.* (2015b)
		R	CACAACACGGCGCTAAGTAAA						
O20	wzx	F	GCCAAATCAAAATGGCTACCGA	57°	245	44	15	KJ778793.1	This study
		R	GTTGCAGCAAGGTTTGCTCT						
O21	*wzx*	F	CATGGTGTTTGGCGATTCTGT	59°	385	10	100	EU694098	This study
		R	GGGAGCAACAAACCGGAAAG						
O22	*wzx*	F	TGTCGCCACTACTTTCCGCGTTTA	58°	458	100	100	DQ851855	This study
		R	AGCCCATGACATTACTACGGCACT						

Table 2. (*Cont.*)

O-group	Gene		Primers	Temp (C)	Amplicon bp	Specificity for field isolates (same O-type)		GenBank	Reference
						n = pos	% pos		
O23	*wzy*	F	ACTCATTACTCTCTTTAAACACGGA	55°	487	14	93	KJ755561	This study
		R	TGCATAACTCTTCAGGCGAT						
O24	*wzx*	F	TGATTGTTTGGTAAGAC		345			DQ220292	Cheng *et al.* (2006)
		R	AGCCATGCTTTTA						
O25	*wzy*	F	AGAGATCCGTCTTTTATTTGTTCGC	55°	230			GU299796	Li *et al.* (2010)
		R	GTTCTGGATACCTAACGCAATACCC						
O26	*wzx*	F	GCGCTGCAATTGCTTATGTA	52°	151	100	100	AF529080	DebRoy *et al.* (2004)
		R	TTTCCCCGCAATTTATTCAG						
O27	*wzx*	F	TACATTTGGCGTCGCAACAG	55°	534	10	100	GU014555	This study
		R	CCGTTTAGCTTTGCTGCTTGT						
O28ac/O42	*wzx*	F	ACCAGAGCAAGGACGATTTGTCA	58°	554			DQ462205	Fratamico *et al.* (2010)
		R	CAACTTTAACTTTCCCAAGCGCGG						
O29	*wzx*	F	GATTATGGGCGGATGTCCCT	55°	625	10	100	EU294173	This study
		R	ATATACCAGCCTGCCCCTCT						
O30	*wzx*	F	CACGTTAAATGCCGGCTCTG	55°	221	9	100	KJ755563	This study
		R	CCAAACGTGCTTTTTGAGGCT						
O32	*wzx*	F	TACAGCGTTATTGGGAGGCG	55°	311	16	62	EU296410	This study
		R	AGCGTGATGCTACAGATGTGA						
O33	*wzx*	F	GGACCTCTGTTTGTGGCTGT	55°	413	11	100	KJ755564	This study
		R	GGTGTGCCTATCGCATAACC						
O34	*wzx*	F	TGGAATGTCATGGCGTGTAA	55°	477	19	53	KJ778803	This study
		R	AGGGCCGCAAATAAACCCTA						
O35	*wzx*	F	TGCCTCTTTTTGTTTCCCAGA	55°	534	10	100	FJ940774	This study
		R	ACTGAGCGAGACTCAGTCCTA						
O36	*wzx*	F	ATCTCGGACTTTCGCGAGC	55°	620	12	83	AB811613	This study
		R	TGAAATTGCATCCGCAGGAG						
O37	*wzx*	F	ACGGCACTGTCAAAACTGAT	55°	282	10	100	KJ755554	This study
		R	CTGCGCCATACCCCCATATT						
O38	*wzy*	F	TGGGTATATTTCAGGTGGCAATGA	57°	608	10	100	KP710589	This study
		R	TCGAAAGTGCTGGGAAAGGTAA						
O39	*wzx*	F	TGGGTGCTGATAACAAATACGGA	57°	131	10	100	AB811616.1	This study
		R	GCTTGTGCAATAGGTTGGGC						
O40	*wzx*	F	GGGGGCTTGGTTAGGAACTG	55°	777	12	92	EU296417	This study
		R	ACACCAAGACTATGCCAAACAG						
O41	*wzx*	F	TACGCGTTTTCGTCCAACTG	57	200	10	100	AB811617.1	This study
		R	ACCAAAGCGTAGCAAGTGGT						
O42	*wzx*	F	CCCAGTCAGCCAGGCATTAT	55°	254	11	100	FJ539194	This study
		R	CAAGCGCGGAAAATACAGGG						
O43	*wzx*	F	TTTTGTGGTACTGCGCTTGC	57°	273	10	100	KJ778789.1	This study
		R	TGCACCAGCAAAAAGCCTTC						

O45	*wzx*	F	CCGGGTTTCGATTTGTGAAGGTTG	55°	527	55	100	AY771223	DebRoy *et al.* (2005)
		R	CACAACAGCCACTACTAGGCAGAA						
O46/O134	*wzx*	F	TGATGCGGTCGGCATATTCA	59°	236	10	100	AB811621	This study
		R	CTCGAAACCGGAGCCATACA						
O48	*wzx*	F	AGTCTGTATGGCTTGGCTGT	55°	238	10	100	KJ710508.1	This study
		R	CGCCTTCCCATTCTTCACCT						
O49	*wzx*	F	TGGTTGCTACGTTGGTTGGA	55°	429	11	100	AB811623	This study
		R	TCCAAACAATTCCCGTGCCT						
O51	*wzx*	F	TTGTAAGTCCATTGGGGGCA	55°	243	10	100	AB812020	This study
		R	CAGCGCTCCTCGCTACTAAA						
O52	*wbrX*	F	AACGCGGCCTTGATGAATTTAGCC	55°	280	10	100	AY528413	This study
		R	TCTTCAGTGAATCGGCCTGGAAGA						
O53	*wzy*	F	CGGGATGTTGGCATTTGCAG	57°	539	10	100	EU289392.1	This study
		R	AGGAGTGCCAAGGTGAAACT						
O54	*wzx*	F	ATTAGCGCTGCGAAATCTGC	55°	635	10	100	AB812085	This study
		R	CGGTTCATCCGGGGGAAAAT						
O55	*wzx*	F	AATGGAACATTGCAACAGCA	55°	150	50	100	AF461121	DebRoy *et al.* (2005)
		R	TTGCTTTTCTGGAATCCACA						
O56	*wzx*	F	TGTCGGTGTTGAAGTCTAT	43°	692			DQ220293	Cheng *et al.* (2006)
		R	GCTAAAATTAATTTTATT						
O58	*wzx*	F	TGCTTGTTTGTCATTTGGGACA	55°	753	10	100	EU294175	This study
		R	GCTGCCGACAACCCTATACA						
O59	*wzx*	F	CAATCACAATTGCAAGCCTCT	55°	878	13	92	AY654590	This study
		R	AGAATCTGACATAGCTGCCG						
O60	*wzm*	F	AAGTATGCTAGGTGCGGCAT	57°	250	3	100	AB812022	This study
		R	AGGCAAACAGATTCTTGGGA						
O61	*wzx*	F	TGCTAAGTTCGTGGTGTCGTT	55°	234	10	100	GU220362	This study
		R	ACGATGAACGAGCAGCTTTA						
O62/O68	*rmlA/C*	F	CTACACTGATGTTAGCGGGTATT		1969 (O62)	2/6	100	JX501334	Liu *et al.* (2015)
		R	CCGCTTCAAATTCAGGACAATAA		1172 (O68)			KJ534585	
O63	*wzx*	F	ATAGCCAATGGTGGAAAGCGGCTA	56°	410	19	100	EU549862	This study
		R	GCCCTTGGTAATTGGGCCGATAAA						
O64	*wzy*	F	TCGCAACTTCACTCGCCATA	55°	301	19	21	AB812025	This study
		R	ATTCTCTCGCTGCTGATGGT						
O65	*wzx*	F	ACTGTGGTGTGGCGATTAAGA	55°	462	11	100	KP710592	This study
		R	CACATCCCCCACTAGTTGCC						
O66	*wzx*	F	TTGCGATGGCAGGAATAA	50°	469			DQ069297	Cheng *et al.* (2007)
		R	GCTCCAAGGCTAGTGAA						
O69	*wzx*	F	CCGACGCATAGCAGAGATGA	55°	516	12	67	KJ778804	This study
		R	TACAGGTGCAAGAACTCCCG						
O70	*wzx*	F	CAGTTCTAATGTCGGGCATTGT	55°	599	10	100	FN995094	This study
		R	CATGGGAAACGCATAAAGCCA						
O71	*wzx*	F	GCATTATTAGCCACTTCAA	55°	377	10	100	GU445927	Hu *et al.* (2010)
		R	AGCCGTATCATTTAGAGCAGA						
O74	*wzx*	F	CGCTTAGGTAAGCCTCAGATATT	55°	219	11	100	KJ778807	This study
		R	GCTTAAGGTTTGCATGGGTTAG						
O75	*wzy*	F	GAGATATACATGGGGAGGTAGGCT	55°	511	10	100	GU299795	Li *et al.* (2010)
		R	ACCCGATAATCATATTCTTCCCAAC						

Table 2. (*Cont.*)

O-group	Gene		Primers	Temp (C)	Amplicon bp	Specificity for field isolates (same O-type)		GenBank	Reference
						n = pos	% pos		
O76	*wzx*	F	CATATGCAGATTGAAGGTAG	57°	533	10	100	AB812031	This study
		R	GAAAGCCATAAAGTGCC						
O78	*wzx*	F	GGTATGGGTTTGGTGGTA	57°	992	10	100	FJ940775	Liu *et al.* (2010)
		R	AGAATCACAACTCTCGGCA						
O79	*wzx*	F	GCTATAGTGCACCAGGATTGT	55°	266	8	100	KJ778790	This study
		R	GATTGCCGTCTGCCCTAAT						
O80	*wzx*	F	GAAGCATGGCTTCTAGGGGG	59°	406	10	100	AB812032	This study
		R	TGACAAAGGTAGCCACGGAA						
O81	*wzx*	F	ATGGCACAGTTGCTGGAAGT	57°	671	14	93	KJ778811	This study
		R	TTCTGTCCCGCTCGCATAAT						
O82	*wzx*	F	ATTTAACCAGGGTGCTCGGG	59°	521	10	100	AB812034	This study
		R	ACTCGGCTTGCAGAAACTCA						
O83	*wzx*	F	GTACACCAGGCAAACCTCGAAAG	55°	362	14	71	KJ778808	Li *et al.* (2010)
		R	TTCTGTAAGCTAATGAATAGGCACC						
O84	*wzx*	F	TCAGCGTTCCAAGAAGCACT	59°	621	10	100	KJ778809	This study
		R	TGGTGTGCACTTATACATCCGA						
O85	*wzx*	F	ACTGCAGGTGAGTGGTGATT	55°	346	10	100	KJ778791	This study
		R	TGCTAAGTACAGCGCAACCT						
O86	*wzx*	F	GTATTCATTGCAGGCGTGGGCATT	59°	273	94	100	AY670704	Feng *et al.* (2005)
		R	CCCAAACGCCACTACACCGTAAA						
O87	*wzx*	F	CCCAGAGTAAGTGCATTGTATGA	55°	350	13	100	EU294177	This study
		R	GCAGCACCATCTGCTGAATA						
O88	*wzx*	F	TGTTCTCTCCCTTCGTTGGC	57°	355	10	100	KJ778812	This study
		R	CCAGCGAGCTCCTGAAATGA						
O89/O101/O162	methyl transferase	F	CGTTGCTTGGTGAAAGTGTATT	55°	418	12	100	KJ755555	This study
		R	CCGCAGTAGAAGCTACAGTTAG						
O90/O127	*wzx*	F	GTTCGCAGATTATGCCTCGC	55°	678	10	100	AB812054	This study
		R	CCCCTTGTGATGATAAAGCTGG						
O91	*wzx*	F	TTGCATCTGGCGCAATAAACACGG	59°	616	60	100	AY035396	Fratamico *et al.* (2009a)
		R	ACACCATCCCAAATACCTGCTTGC						
O92	*wzm*	F	GGATCACAGTAAGCACAGCG	57°	397	9	100	AB812040	This study
		R	AACCGAGTGCAAAGGCATGA						
O93	*wzx*	F	ACTCACGCTAAGCTACAAAGAA	55°	761	4	30	AB812041	This study
		R	TAGCCACCGAATCCAAACC						
O95	*wzm*	F	AATGCTCCTTATGCTGCGGT	57°	486	4	100	KJ755556	This study
		R	GCCTAAGATAGCCCAGCGAG						
O96	*wzx*	F	TTGCTGCACTTATCCAGGGG	55°	377	10	100	KJ778788.1	This study
		R	ACCTCCAACATAGCGTAGCG						
O97	*wzt*	F	AGGCAGATCGTCCACAGTCA		184			KJ778810	Iguchi *et al.* (2015b)
		R	ACAGGATAAATGCCAGCCAA						

O98	*wzx*	F	GGGCGTATTGAGGTTCTTGT	55°	675	14	86	DQ180602	This study
		R	CAACTGAAATGGAGCAGCAAAT						
O99	*wzx*	F	CGCAACCAACAGAAGGTACTA	55°	619	10	100	FJ940773	This study
		R	CCTGGCCGCGTATGATAAA						
O100	*wzx*	F	TTCCTCAGTCGCACTTTGGG	57°	593			KJ778805	Iguchi *et al.* (2015b)
		R	ACACCGACTGCCAATCTTGA						
O101/O162	*wzt*	F	TGTGGGTGTTGAATTTCCCA	57°	661	10	100	KJ778806	This study
		R	GCCCTCTAACCACAGGACTT						
O102	*wzx*	F	CGTGAATTGGGAGCAGAAAGG	55°	200	10	100	AB812047	This study
		R	AGAGCTGCACCAACTAAACCC						
O103	*wzx*	F	TTGGAGCGTTAACTGGACCT	57°	321	59	100	AY532664	Fratamico *et al.* (2005)
		R	GCTCCCGAGCACGTATAAG						
O104	*wzx*	F	TTCGGTCAAAGCAGATATCGCAGG	55°	451	14	100	AF361371	This study
		R	CAGTTGCAATAGCTGCGCCTAAAG						
O105	wzx	F	CCTTGGATGCCTAAAAGAAGCA	59°	514	4	100	EU294171	This study
		R	CATGGCGACCAAAAGGCAAA						
O107/O117	*wzx*	F	GCAATTGGCGCACAGAAACA	55°	343	12	100	EU694095	This study
		R	AAACCAAAGCAGTTTTTGGGAA						
O108	*wzy*	F	AGCTTCCCTGTCTACGGTTGA		647			KP710597	Iguchi *et al.* (2015b)
		R	CCATCCCATCACCAAATTGA						
O109	*wzx*	F	TCTCTCTCGACATACCCGCGCTT	59°	204			HM485572	Iguchi *et al.* (2015b)
		R	ACCGTAGCCCAAAGAGCCACA						
O110	*wzx*	F	AGGATTTGGCGAAAATGGCT	55°	419	16	75	AB812049.1	This study
		R	CGGAGACCAACCCCGATAAT						
O111/OX21	*wzx*	F	TGTTTCTTCGATGTTGCGAG	59°	438			AF078736	DebRoy *et al.* (2011a, b)
		R	GCAAGGGACATAAGAAGCCA						
O112ab	*wzy*	F	CGGGTTAACAGCCCATTTTT		241			EU296413	Iguchi *et al.* (2015b)
		R	CAGCCCCCATTTACCAGTAAT						
O112ac	*wzx*	F	CTGTCCTTTTGCGCGAATTA		1180			EU296045	Iguchi *et al.* (2015b)
		R	AAATCCCAGAGCAAGGGTAGA						
O113	*wzx*	F	GGGTTAGATGGAGCGCTATTGAGA	55°	771	50	100	AF172324	DebRoy *et al.* (2004)
		R	AGGTCACCCTCTGAATTATGGCAG						
O114	*wzx*	F	CAGGTTTAAGTTGGGTA	55°	603	10	100	AY573377	Feng *et al.* (2004)
		R	AAGAAGAAAGTCTGGGTA						
O115	*wzx*	F	TTGGCTGGGGCATATTTCGT	55°	243	10	100	GU068041	This study
		R	GTGTTGCAACCGAAGCCAAT						
O116	*wzx*	F	CTTTTGCGTTGTGCCTCGAA	55°	355	14	85	AB812051	This study
		R	ATACTGCCCTACGTTTGCGG						
O118/O151	*wzx*	F	GTGGGAGTCTGAATCAAGTTGCGA	60°	344	69	100	DQ990684	Liu *et al.* (2008b)
		R	AGCAACCTTACCCAATCCTAAGGG						
O119	*wzx*	F	CTGGGGCAATCTGCTTTCCT	55°	421	13	100	GQ499368	This study
		R	CCCAAGGTATTGTTGCCCCT						
O120	*wzx*	F	CTGGTTTTGTTGTTGCATTGCT	55°	535	11	100	AB812052	This study
		R	CCAGTTGGTGCCAACCAAAG						
O121	*wzx*	F	AGGCGCTGTTTGGTCTCTTA	55°	310	50	100	AY208937	Fratamico *et al.* (2003)
		R	TCGCTACCGCTAATGATTCC						
O123/O186	*wzx*	F	ACAATTAGGGCCTGGTGCAT	55°	619	11	100	DQ676934	This study
		R	TGTGCTAGCGCTAAAGGACT						

Table 2. (*Cont.*)

O-group	Gene		Primers	Temp (C)	Amplicon bp	Specificity for field isolates (same O-type)		GenBank	Reference
						n = pos	% pos		
O124/O164	*wzx*	F	GGCGGCAGACTATGGCTTAT	55°	214	10	100	AB972420	This study
		R	CCCCCTTGCCAAAAGAAGAC						
O125ab	*wzx*	F	TGAGTATGCTGGGTTGTGGG	55°	423	10	100	AB812053	This study
		R	TGAGCATTTTCGACGTAGCAT						
O126	*wzx*	F	TTAGCTCTCGTAGAGGCTGGTGTT	56°	925	77	100	DQ465248	Liu *et al.* (2007)
		R	ATGTCATTCCTGGGACGCGAATGT						
O128/OX38	*wzx*	F	GCCATTACGACGTTGATGACT	55°	768	10	100	AY217096	This study
		R	TGCAACCCCAATAGCAAAAGC						
O129	*wzx*	F	GATGTAGGCCTGACTCGCTC	55°	516	11	100	AB972421	This study
		R	TGTCATCCAACCGCCAAAGA						
O130	*wzx*	F	AGGAGGGCTAATTGCATCCG	55°	277	10	100	EU296421	This study
		R	TGGCATCAATGCTTGGTTGA						
O131	*wzx*	F	TAGGTGCTAGTGAGGCTGGG	55°	291	10	100	KJ755544	This study
		R	CCTGCCCTAAACAAAGCCAG						
O132	*wzx*	F	GGGAAGGAGCAAATGGGGTT	55°	369	11	100	AB812056	This study
		R	CCCCTGAACTGCTTGCCTAA						
O133	*wzy*	F	TCTGCGTTATGGCAACTGTCA		1017			KJ710509	Iguchi *et al.* (2015b)
		R	CACTCGCAAACGTCTCACATT						
O136	*wzx*	F	ATGGATCAAGAAGGGGGTGG	57°	551	16	75	KJ755546	This study
		R	AAATCGCCATGCCTATGCCA						
O137	*wzx*	F	GTTTGGCTAGGGAATGGGGG	55°	580	10	100	AB972423	This study
		R	GAATTGCGACTGTTGCCGTC						
O138	*wzx*	F	GCAGCAATGCCTGCTGTTTT	55°	696	9	100	DQ109551	This study
		R	AGCGTATGCAACCCCAATGA						
O139	*wzx*	F	ATAATGCAGCCCTTGGTCTC	55°	400	10	100	DQ109552	This study
		R	CACAGTCCAGAGACCCCGTA						
O140	*wzx*	F	TTGCATTGATACTTGCGGCG	55°	183	10	100	KJ755552	This study
		R	ACGCGGCTATCCAAACATCT						
O141	*wzy*	F	TGAACCTGGGTTTACATT	50°	746			DQ868765	Han *et al.* (2007)
		R	GTACAATTATCATTGCGAGT						
O142	*wzx*	F	ATTGTTCCCGGTGCTTTTGC	55°	186	10	100	KJ755549	This study
		R	TGCAAATTCAGGTCCCATCCA						
O143	*wzx*	F	AGCATCTGCTCTGGTTGGAC	55°	233	10	100	EU294164	This study
		R	CGGATACTGTCACAAATCGCA						
O144	*wzx*	F	TCGCTATCAATCCTCAACTGGT	55°	359	9	100	KJ755550	This study
		R	CAACTGTGGCTCATGCCAAT						
O145	*wzx*	F	ACTGGGATTGGACGTGGATA	60°	222	70	100	AY647260	Fratamico *et al.* (2009b)
		R	AGGCAAGCTTTGGAAATGAA						
O146	*wzx*	F	AGGGTGACCATCAACACACTTGGA	56°	640	98	100	DQ465249	Liu *et al.* (2007)
		R	AGTTCAATACTGTCGCAGCTCCTC						

O147	*wzx*	F	AATCGGCTTTGGAAGCTGGA	55°	519	10	100	DQ868766	DebRoy *et al.* (2010)
		R	ATTGCGGCTCCAACAATTCC						
O148	*wzx*	F	TATTACCATGCCCCAAGCCC	57°	438	10	100	DQ167407	This study
		R	GCTGGTGACTATGTTGTGGC						
O149	*wzy*	F	TTTGGTGCAGATACTCAGA		709			DQ091854	Han *et al.* (2007)
		R	GAACAATAGATGCGATACAA						
O150	*wzx*	F	TGCGTTCACTTGCTGGTTTG	55°	313	10	100	EU294168	This study
		R	ATGAGTGCAGGCACTTGAA						
O152	*wzx*	F	AAGAGGCGCGGTTCAATCTT	55°	387	19	58	EU294170	This study
		R	AACGTACCAAAAAGCCCGAA						
O153	*wzx*	F	TCGTCCTATGCGCAGTATTCA	55°	799	9	100	KJ755551	This study
		R	GGGGCAACTACTACGGACAC						
O154	*wzx*	F	CCGACACAGTTAGGTGCGTA	55°	453	10	100	AB812064	This study
		R	ACCCATCAAAACGGCAAACG						
O155	*wzx*	F	TCCTGGTGAATTTGGCTTGC	55°	542	9	100	AY657020	This study
		R	TTGGCCAAGACGAGCGTAAT						
O156	*wzx*	F	GACGAGTCGGGATGGTGTTT	59°	928	9	88	KJ755559	This study
		R	ACCCCGTGAAAGATGGACAC						
O157	*wzx*	F	TCGAGGTACCTGAATCTTTCCTTCTGT		894	100	100	AF061251	Wang and Reeves (1998)
		R	ACCAGTCTTGGTGCTGCTCTGACA						
O158	*wzx*	F	CGTCCGAAGCAAAGATGA		624			GU068044	Wang *et al.* (2010)
		R	ATACTGAATATGCCCGTGA						
O159	*wzy*	F	TGTGTATGTTAGGCGGGGTAA		298			EU294176	Iguchi *et al.* (2015b)
		R	AGTCGGTTCCATTTGTTGCA						
O160	*wzx*	F	CTTGGTTGTTTCAGGGGCTTG	55°	512	10	100	KJ755560	This study
		R	TCCACCTAAACATCCCAGAATCG						
O161	*wzx*	F	TATGTTGGCGGATATTCGGT		349			GU220361	Iguchi *et al.* (2015b)
		R	AGGCAACGGATGGAATTGAT						
O163	*wzx*	F	TGTCATCCAAGGGGGTTCAG	59°	606	47	100	AB812068	This study
		R	TTGGGATTCTTGCTGCTGCT						
O165	*wzx*	F	AACTGTTTATCCGAAGTGGTAG	59°	630	10	100	GU068045	This study
		R	CACGCTTTAACGCATACAG						
O166	*wzy*	F	TTCATAGCTGGCCTCCTTGTT		462			GU299794	Iguchi *et al.* (2015b)
		R	TCTATTCGCCGAATCCTTTCT						
O167	*wzy*	F	TCAGGGGCAATTACAATCCTT		403			EU296408	Iguchi *et al.* (2015b)
		R	TCGCGCATAGAATAGCATGTC						
O168/OX6	*wzx*	F	TGTCGACTTTGGGAAATGTGG	59°	336	10	100	EU296403	This study
		R	CTGCAGAGGCCAATTCAGGT						
O169	*wzx*	F	TGCGGCATATTATCCAATGGT	57°	188	10	100	KJ778796	This study
		R	CTGCACCTGTGACACCATGA						
O170	*wzx*	F	CAAGACGTATGGGTTGCTGC	57°	226	10	100	KJ778797	This study
		R	TTGGACCCGAGACAAATCCG						
O171	*wzx*	F	TGCTCAAGTGGCATGCAGAT	55°	281	12	100	AB812071.1	This study
		R	TGCAACCTGATATCCAGCAGT						
O172	*wzx*	F	TGTTGATGGTTAAAATA		704			AY545992	Guo *et al.* (2004)
		R	CATAATAAGGACATGAC						
O173	*wzx*	F	AGATTAGCCTTCCAGGGGCT	59°	403	11	100	GU068046	This study
		R	TTAATGCAGCGCCGTAACCA						

Table 2. (*Cont.*)

O-group	Gene		Primers	Temp (C)	Amplicon bp	Specificity for field isolates (same O-type)		GenBank	Reference
						n = pos	% pos		
O174	*wzx*	F	CGCTTGGAATTGGGCCTGTA	59°	514	10	100	DQ008592	This study
		R	TCCAGCTGAGTAAAGAGACTTACAA						
O175	*wzx*	F	TGTTTAATTGTTCCTCCGCTCA	55°	343	11	100	KJ739597	This study
		R	GCAGCCCAACCAAACTTAGC						
O176	*wzy*	F	GCGAGCATATACGTTCCGTT	55°	404	2	100	KJ778798	This study
		R	GCAAGGCTTTGCTCATCGTA						
O177	*wzx*	F	TCGGTGTTTGAAGGGGAAG	59°	749	10	90	DQ008593	Beutin *et al.* (2005b)
		R	GTCCATGCATATGCCGTTC						
O178	*wzx*	F	CTGTCCGTAGTGAGGTTGGC	57°	495	12	91	KJ778799	This study
		R	ACCTCCAGATCGGTCCTTAATC						
O179	*wzx*	F	CAAAGCACACTTGATGGCGT	57°	541	10	90	KJ778800	This study
		R	CCTCCACCCCAGCATAACTG						
O180	*wzx*	F	GCACGATAGCTACAGGTCGG	59°	589	7	100	JQ751058	This study
		R	CCAGCAGCTGTAGGGGAAAA						
O181	*wzx*	F	TGGAGTAACGAAATACACCGCT	55°	595	11	82	KJ778801	This study
		R	AGATTGCCAATAACCAGAAGCA						
O182	*wzx*	F	ATTACCTTGGCCCTGCTTCC	55°	630	1	100	KJ778802	This study
		R	ACTGCATAACCAATGTAGCCGA						
O183	*wzx*	F	GATACCAGTACGGCCTTGGG	59°	691	6	100	AB627352	This study
		R	TGGGGCAACATCTGTTCACT						
O184/OX9	*wzx*	F	CCGTGGTGGAGAGCTTCATT	59°	845	2	100	AB812080	This study
		R	GCACCACGAGACAACATTGG						
O185	*wzx*	F	GCTGCATTTGCCATGCTCTT	55°	765	5	100	AB812081	This study
		R	ACTTAACGAAGCGAGCCCAA						
O187	*wzx*	F	ATTCACAGCGGTAGGAACCC	55°	792	2	100	KJ739600	This study
		R	TTTGCGCCCTGAACACCATA						

pos, positive by PCR; bp, base pairs.

or deletions in the targeted gene, which can be further analyzed to determine the actual O-group. On the other hand, serotyping may not be always accurate, and strains that showed negative results by PCR may actually belong to other serogroups. It has been reported (Plainvert *et al.*, 2007) that a pathogenic clone of *E. coli* that causes meningitis belonging to serogroup O45 has two O45 polysaccharides, representing two different antigens. The O-AGC for this *E. coli* O45 clone is genetically related to the O45 reference strain, although they show low DNA sequence homology of the orthologous genes. Both the meningitis strain and the reference strain react to O45 antisera serologically.

When bacteria other than *E. coli* were used for testing specificity, most of the PCR assays for O-groups were highly specific for *E. coli* except for O13 strains, in which PCR assays targeting the O13 *wzx* gene reacted positively with *S. flexneri*. O147 strains with PCR assays targeting the O147 *wzx* gene exhibited positive results with *S. boydii*. It is known that the O-antigens of *E. coli* and *Shigella* species are structurally closely related (Liu *et al.*, 2010; Knirel *et al.*, 2016). The O-antigens of O13, O129, and O135 are almost identical and the structure is shared by the *S. flexneri* O-antigen group. While *S. flexneri* O-antigen F2a reacts with *E. coli* O13 serologically, F4b and F5 react with O135 and O129, respectively (Ewing, 1986; Morona *et al.*, 1995). *Shigella* O-antigens, F1–5 have close structural similarity and have the same basic O-unit structure, and the distinguishing features are due to phage-encoded post-polymerization modifications in O acetylation and/or glucosylation (Liu *et al.*, 2008a).

Future directions

While serotyping has been a reliable and low-cost method for determining the *E. coli* O-groups, it has many disadvantages. It is now imperative that a platform be developed for detecting O-types based on molecular methods that are fast, reliable, and cost-effective. Although the PCR is a comparatively easy and low-cost method, developing other formats such as microarray, for high throughput processing could substantially reduce the labor and time required. WGS is now being considered by many public health laboratories for routine diagnostics as the cost of DNA sequencing is declining (Joensen *et al.*, 2015; Kwong *et al.*, 2015). Joensen *et al.* (2014) developed the web tool known as SerotypeFinder for identifying *E. coli* serotypes by analysis of the WGS data. Another method developed by Ingle *et al.* (2016) was used to rapidly and accurately detect serotypes of *E. coli* directly from WGS reads using short read data. Their approach bypasses the need for *de novo* assembly of the genome by directly screening short read data against the curated database of alleles linked to known *E. coli* O- and H-groups (the EcOH database), using the software package SRST2. Applying this approach, they explored the genetic diversity of >1500 *E. coli* genomes from multiple sources and found serotype diversity, revealing 38 novel O-AGCs. The O-antigen region is subjected to diversity due to selection pressure from the host immune system, as well as due to insertion of phages. Therefore, recombination events may occur in O-AGCs, resulting in O-antigens from related strains having different O-types and from different lineages having similar O-types. As more WGS data are becoming available, diversity in the O-AGC region will be realized. As different approaches for typing *E. coli* serogroups are explored, it is hoped that in the near future, an easy, rapid and reliable molecular-based method based on the sequences of the O-AGCs will become available for routine testing of *E. coli* serogroups.

References

Abbadi SH and Strockbine NA (2007). Identification of *Escherichia coli* flagellar types by restriction of the amplified *fliC* gene. *Egyptian Journal of Medical Microbiology* **16**: 225–233.

Achtman M and Pluschke G (1986). Clonal analysis of descent and virulence among selected *Escherichia coli*. *Annual Review of Microbiology* **40**: 185–210.

Akira S, Uematsu S and Takeuchi O (2006). Pathogen recognition and innate immunity. *Cell* **124**: 783–801.

Bastin DA, Stevenson G, Brown AH and Reeves PR (1993). Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. *Molecular Microbiology* **7**: 725–734.

Batchelor RA, Alifano P, Biffali E, Hull SI and Hull RA (1992). Nucleotide sequences of the genes regulating O-polysaccharide antigen chain length (*rol*) from *Escherichia coli* and *Salmonella typhimurium*: protein homology and functional complementation. *Journal of Bacteriology* **174**: 5228–5236.

Beutin L, Strauch E, Zimmermann S, Kaulfuss S, Schaudinn C, Männel A and Gelderblom HR (2005a). Genetical and functional investigation of *fliC* genes encoding flagellar serotype H4 in wildtype strains of *Escherichia coli* and in a laboratory *E. coli K*-12 strain expressing flagellar antigen type H48. *BMC Microbiology* **5**: 4.

Beutin L, Kong Q, Feng L, Wang Q, Krause G, Leomil L, Jin Q and Wang L (2005b). Development of PCR assays targeting the genes involved in synthesis and assembly of the new *Escherichia coli* O174 and O177 O antigens. *Journal of Clinical Microbiology* **43**: 5143–5149.

Bos MP, Robert V and Tommassen J (2007). Biogenesis of Gram-negative bacterial outer membrane. *Annual Review of Microbiology* **61**: 191–214.

Botkin DJ, Galli L, Sankarapani V, Soler M, Rivas M and Torres AG (2012). Development of a multiplex PR assay for detection of Shiga toxin-producing *Escherichia coli*, enterohemorrhagic *E. coli*, and enteropathogenic *E. coli* strains. *Frontiers in Cellular and Infection Microbiology* **2**: 8.

Bronner D, Clarke BR and Whitfield C (1994). Identification of an ATP-binding cassette transport system required for translocation of lipopolysachharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumonia* serotype O1. *Molecular Microbiology* **14**: 505–519.

Bugarel M, Beutin L, Martin A, Gill A and Fach P (2010). Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with hemorrhagic colitis and hemolytic uremic syndrome in humans. *International Journal of Food Microbiology* **142**: 318–329.

Carvallo F, DebRoy C, Baeza E, Hinckley L and Gilbert K (2010). Extraintestinal pathogenic *Escherichia coli* induced necrotizing pneumonia in a white tiger cub. *Journal of Veterinary Diagnostic Investigation* **22**: 136–140.

Chafchaouni-Moussaoui I, Novikov A, Bhrada F, Perry MB, Filali-Maltouf A and Caroff M (2011). A new rapid and microscale hydrolysis, using triethylamine citrate, for lipopolysaccharide characterization by mass spectrometry. *Rapid Communication in Mass Spectrometry* **14**: 2043–2048.

https://doi.org/10.1017/S1466252317000093 Published online by Cambridge University Press

Cheng J, Wang Q, Wang W, Wang Y, Wang L and Feng L (2006). Characterization of *E. Coli* O24 and O56 O antigen gene clusters reveals a complex evolutionary history of the O24 gene cluster. *Current Microbiology* **53**: 470–476.

Cheng J, Liu B, Bastin DA, Han W, Wang L and Feng L (2007). Genetic characterization of the *Escherichia coli* O66 antigen and functional identification of its *wzy* gene. *Journal of Microbiology* **45**: 69–74.

Coimbra RS, Grimont F, Lenormand P, Burguière P, Beutin L and Grimont PAD (2000). Identification of *Escherichia coli* O-serogroups by restriction of the amplified O-antigen gene cluster (rfp-RFLP). *Research in Microbiology* **151**: 639–654.

Daniels C, Vindurampulle C and Morona R (1998). Overexpression and topology of *Shigella flexneri* O antigen polymerase (Rfc/Wzy). *Molecular Microbiology* **28**: 1211–1222.

DebRoy C, Roberts E, Kundrat J, Davis MA, Briggs CE and Fratamico PM (2004). Detection of *Escherichia coli* serogroups O26 and O113 by PCR amplification of *wzx* and *wzy* genes. *Applied and Environmental Microbiology* **70**: 1830–1832.

DebRoy C, Fratamico PM, Roberts E, Davis MA 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. *Applied and Environmental Microbiology* **71**: 4919–4924.

DebRoy C, Roberts E, Jayarao BM and Brooks JW (2008). Bronchopneumonia associated with extraintestinal pathogenic *Escherichia coli* in a horse. *Journal of Veterinary Diagnostic Investigations* **20**: 661–664.

DebRoy C, Roberts E, Davis MA and Bumbaugh AC (2010). Detection of non-serotypable Shiga toxin-producing *Escherichia coli* strains derived from pigs by multiplex PCR assay. *Foodborne Pathogens and Disease* **7**: 1407–1414.

DebRoy C, Roberts E and Fratamico PM (2011a). Detection of O antigens in *Escherichia coli*. *Animal Health Research Review* **12**: 169–185.

DebRoy C, Roberts E, Valadez AM, Dudley EG and Cutter CN (2011b). Detection of Shiga toxin producing *Escherichia coli* O26, O45, O103, O111, O113, O121, O145, and O157 serogroups by multiplex PCR of the *wzx* gene of the O-antigen gene cluster. *Foodborne Pathogens and Disease* **8**: 651–652.

DebRoy C, Fratamico PM, Yan X, Baranzoni GM, Liu Y, Needleman DS, Tebbs R, O'Connell CD, Allred A, Swimley M, Mwangi M, Kapur V, Raygoza Garay JA, Roberts EL and Katani R (2016). Comparison of O-antigen gene clusters of all O-serogroups of *Escherichia coli* and proposal for adopting a new nomenclature for O-typing. *PLoS ONE* **11**: e0147434. doi: 10.1371/journal.pone.0147434.

Delannoy S, Beutin L and Fach P (2013). Discrimination of enterohemorrhagic *Escherichia coli* (EHEC) from non-EHEC strains based on detection of various combinations of type III effector genes. *Journal of Clinical Microbiology* **51**: 3257–3262.

Dickinson CE, Gould DH, Davidson AH, Avery PR, Legare ME, Hyatt DR and DebRoy C (2008). Hemolytic-uremic syndrome in a postpartum mare concurrent with encephalopathy in the neonatal foal. *Journal of Veterinary Diagnostic Investigation* **20**: 239–242.

Doumith M, Day MJ, Hope R, Wain J and Woodford N (2012). Improved multiplex PCR strategy for rapid assignment of the four major *Escherichia coli* phylogenetic groups. *Journal of Clinical Microbiology* **50**: 3108–3110.

Ewing WH (1986). The genus shigella. In: Edwards PR and Ewing WH (eds) *Edwards and Ewing's Identification of Enterobacteriaceae*, 4th edn. New York: Elsevier Science Publishing Co., Inc, pp. 135–172.

Ewing WH and Tatum HW (1956). US Communicable Disease Center. Studies on the serology of the *Escherichia coli* group. [Available online at http://www.worldcat.org/title/studies-on-the-serology-of-the-escherichia-coli-group/oclc/654668490].

Feng L, Wang W, Tao J, Guo H, Krause G, Beutin L and Wang L (2004). Identification of *Escherichia coli* O114 O-antigen gene cluster and development of an O114 serogroup-specific PCR assay. *Journal of Clinical Microbiology* **42**: 3799–3804.

Feng L, Han W, Wang Q, Bastin DA and Wang L (2005). Characterization of *Escherichia coli* O86 O-antigen gene cluster and identification of O86-specific genes. *Veterinary Microbiology* **106**: 241–248.

Fratamico PM and DebRoy C (2010). Detection of *Escherichia coli* O157: H7 in food using real-time multiplex PCR assays targeting the stx1, stx2, wzyO157, and the fliCh7 or *eae* genes. *Food Analytical Methods* **3**: 330–337.

Fratamico PM, Briggs CE, Needle D, Chen CY 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. *Journal of Clinical Microbiology* **41**: 3379–3383.

Fratamico PM, DebRoy C, Strobaugh TP Jr and Chen CY (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. *Canadian Journal of Microbiology* **51**: 515–522.

Fratamico PM, DebRoy C and Liu Y (2009a). The DNA sequence of the *Escherichia coli* O22 O antigen gene cluster and detection of pathogenic strains belonging to *E. coli* serogroups O22 and O91 by multiplex PCR assays targeting virulence genes and genes in the respective O-antigen gene clusters. *Food Analytical Methods* **2**: 169–179.

Fratamico PM, DebRoy C, Miyamoto T and Liu Y (2009b). Detection of entero-hemorrhagic *Escherichia coli* O145 in food by PCR targeting genes in the *E. coli* O145 O antigen gene cluster and the Shiga toxins 1 and 2 genes. *Foodborne Pathogen and Disease* **6**: 605–611.

Fratamico PM, Yan X, Liu Y, DebRoy C, Byrne B, Monaghan A, Fanning S and Bolton D (2010). *Escherichia coli* serogroup O2 and O28ac O-antigen gene cluster sequences and detection of pathogenic *E. coli* O2 and O28ac by PCR. *Canadian Journal of Microbiology* **56**: 308–316.

Fratamico PM, Wasilenko JL, Garman B, DeMarco DR, Varkey S, Jensen M, Rhoden K and Tice G (2014). Evaluation of a multiplex PCR real-time PCR method for detecting Shiga toxin-producing *Escherichia coli* in beef and comparison to the U.S. Department of Agriculture Food Safety and Inspection Service Microbiology Laboratory Guidebook Method. *Journal of Food Protection* **77**: 180–188.

Grozdanov L, Zahringer U, Blum-Oehler G, Brade L, Henne A, Knirel YA, Schombel U, Schulze J, Sonnenborn U, Gottschalk G, Hacker J, Rietschel ET and Dobrindt U (2002). A single nucleotide exchange in the wzy gene is responsible for the semirough O6 lipopolysaccharide phenotype and serum sensitivity of *Escherichia coli* strain Nissle 1917. *Journal of Bacteriology* **184**: 5912–5925.

Guo H, Feng L, Tao J, Zhang C and Wang L (2004). Identification of *Escherichia coli* O172 O-antigen gene cluster and development of a serogroup-specific PCR assay. *Journal of Applied Microbiology* **97**: 181–190.

Han W, Liu B, Cao B, Beutin L, Kruger U, Liu H, Li Y, Liu Y, Feng L and Wang L (2007). DNA microarray-based identification of serogroups and virulence gene patterns of *Escherichia coli* isolates associated with porcine postweaning diarrhea and edema disease. *Applied and Environmental Microbiology* **73**: 4082–4088.

Handt LK, Stoffregen DA, Prescott JS, Pouch WJ, Ngai DTW, Anderson CA, Gatto NT, DebRoy C, Fairbrother JM, Motzel SL and Klein HJ (2003). Clinical and microbiologic characterization of hemorrhagic pneumonia due to extraintestinal pathogenic *E. coli* in four young dogs. *Comparative Medicine* **53**: 663–670.

Hobbs M and Reeves PR (1994). The JUMPstart sequence: a 39 bp element common to several polysaccharide gene clusters. *Molecular Microbiology* **12**: 855–856.

Hegde NV, Jayarao BM and DebRoy C (2012). Rapid detection of the top six non-O157 Shiga toxin-producing *Escherichia coli* O groups in ground beef by flow cytometry. *Journal of Clinical Microbiology* **50**: 2137–2139.

Hegde NV, Praul C, Gehring A, Fratamico PM and DebRoy C (2013). Rapid identification of the six clinically relevant O serogroups of Shiga toxin-producing *Escherichia coli* by antibody microarray. *Journal of Microbiological Methods* **93**: 273–276.

Hu B, Perepelov AV, Liu B, Shevelev SD, Guo D, Senchenkova SN, Shashkov AS, Feng L, Knirel YA and Wang L (2010). Structural and genetic evidence for the close relationship between *Escherichia coli* O71 and *Salmonella enterica* O28 O-antigens. *FEMS Immunology and Medical Microbiology* **59**: 161–169.

Iguchi A, Iyoda S, Kikuchi T, Ogura Y, Katsura K, Ohnishi M, Hayashi T and Thomson NR (2015a). A complete view of the genetic diversity of the *Escherichia coli* O-antigen biosynthesis gene cluster. *DNA Research* **22**: 101–107.

Iguchi A, Iyoda S, Seto K, Morita-Ishihara T, Sheutz F, Ohnishi M and Pathogenic *E. coli* Working Group in Japan (2015b). *Escherichia coli* O-genotyping PCR: a comprehensive and practical platform for molecular O serogrouping. *Journal of Clinical Microbiology* **53**: 2427–2432.

Iguchi A, Iyoda S, Seto K, Nishii H, Ohnishi M, Mekata H, Ogura Y and Hayashi T (2016). Six novel O genotypes from Shiga toxin-producing *Escherichia coli*. *Frontiers in Microbiology* **7**: 765.

Ingle DJ, Valcanis M, Kuzevski A, Tauschek M, Inouye M, Stinear T, Levine MM, Robins-Browne RM and Holt KE (2016). *In silico* serotyping of *E. coli* from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. *Microbial Genomics* **2**: e000064. doi: 10.1099/mgen.0.000064.

Ison SA, Delannoy S, Bugarel M, Nagaraja TG, Renter DG, den Bakker HC, Nightingale KK, Fach P, Loneragan GH (2016). Targeted amplicon sequencing for single-nucleotide-polymorphism genotyping of attaching and effacing *Escherichia coli* O26:H11 cattle strains via a high-throughput library preparation technique. *Applied and Environmental Microbiology* **82**: 640–649.

Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM and Aarestrup FM (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *Journal of Clinical Microbiology* **52**: 1501–1510.

Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM and Scheutz F (2015). Rapid and easy in silico serotyping of *Escherichia coli* using whole genome sequencing (WGS) data. *Journal of Clinical Microbiology* **53**: 2410–2426.

Kaper JB, Nataro JP and Mobley HL (2004). Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* **2**: 123–140.

Kaufmann F (1943). Ueber neue tthermolabile Ko¨rperantigen der Colibakterien. *Acta Pathologica et Microbiologica Scandinavica* **20**: 21–44.

Kaufmann F (1944). Zur Serologie der Coli-Gruppe. *Acta Pathologica et Microbiologica Scandinavica* **21**: 20–45.

Kaufmann F (1947). The serology of *E. coli* group. *Journal of Immunology* **57**: 71–100.

Keenleyside WJ and Whitfield C (1996). A novel pathway for O polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. *Journal of Biological Chemistry* **271**: 28581–28592.

Knirel YA, Qian C, Shashkov AS, Sizova OV, Zdorovenko EL, Naumenko OI, Senchenkova SN, Perepelov AV and Liu B (2016). Structural relationships between genetically closely related O-antigens of *Escherichia coli* and *Shigella* spp. *Biochemistry (Moscow)* **81**: 600–608.

Kwong JC, McCallum N, Sintchenko V and Howden BP (2015). Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**: 199–210.

Lacher DW, Gangiredla J, Jackson SA, Elkins CA and Feng PC (2014). Novel microarray design for molecular serotyping of Shiga toxin-producing *Escherichia coli* strains isolated from fresh produce. *Applied and Environmental Microbiology* **80**: 4677–4682.

Li D, Liu B, Chen M, Guo D, Guo X, Liu F, Feng L and Wang L (2010). A multiplex PCR method to detect 14 *Escherichia coli* serogroups associated with urinary tract infections. *Journal of Microbiological Methods* **82**: 71–77.

Li Y, Perepelov AV, Guo D, Shevelev SD, Senchenkova SN, Shahskov AS, Liu B, Wang L, Knirel YA (2011). Structural and genetic relationships of two pairs of closely related O-antigens of Escherichia coli and *Salmonella enterica*: *E. coli* O11/*S. Enterica* O16 and *E. coli* O21/*S. Enterica* O38. *FEMS Immunology and Medical Microbiology* **61**: 258–268.

Linton KJ and Higgins CF (1998). The *Escherichia coli* ATP binding cassette (ABC) proteins. *Molecular Microbiology* **28**: 5–13.

Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L (2008a). Structure and genetics of *Shigella* O antigens. *FEMS Microbiology Reviews* **32**: 627–653.

Liu B, Wu F, Li D, Beutin L, Chen M, Cao B and Wang L (2010). Development of a serogroup-specific DNA microarray for identification of *Escherichia coli* strains associated with bovine septicemia and diarrhea. *Veterinary Microbiology* **142**: 373–378.

Liu Y and Fratamico PM (2006). *Escherichia coli* O antigen typing using DNA microarrays. *Molecular and Cellular Probes* **20**: 239–244.

Liu Y, DebRoy C and Fratamico PM (2007). Sequencing and analysis of the *Escherichia coli* serogroup O117, O126, and O146 O-antigen gene clusters and development of PCR assays targeting serogroup O117-, O126-, and O146-specific DNA sequences. *Molecular and Cellular Probes* **21**: 295–302.

Liu Y, Fratamico PM, DebRoy C, Bumbaugh AC and Allen JW (2008b). DNA sequencing and identification of serogroup-specific genes in the *Escherichia coli* O118 O antigen gene cluster and demonstration of antigenic diversity but only minor variation in DNA sequence of the O antigen clusters of *E. coli* O118 and O151. *Foodborne Pathogens and Disease* **5**: 449–457.

Liu Y, Yan X, DebRoy C, Fratamico PM, Needleman DS, Li R, Wang W, Losada L, Brinka L, Radune D, Toro M, Hegde N and Meng J (2015). *Escherichia coli* O-antigen gene clusters of serogroups O62, O68, O131, O140, O142, and O163: DNA sequences and similarity between O62 and O68, and PCR-based serogrouping. *Biosensors* **5**: 51–68.

MacLean LL, Webb AC and Perry MB (2006). Structural elucidation of the O-antigenic polysaccharide from enterohemorrhagic (EHEC) *Escherichia coli* O48:H21. *Carbohydrate Research* **341**: 2543–2549.

McGrath BC and Osborn MJ (1991). Localization of the terminal steps of O antigen synthesis in *Salmonella typhimurium*. *Journal of Bacteriology* **173**: 649–654.

Moreno ACR, Guth BEC and Martinez MB (2006). Can the *fliC* PCR-restriction fragment length polymorphism technique replace classic serotyping methods for characterizing the H antigen of enterotoxigenic *Escherichia coli* strains? *Journal of Clinical Microbiology* **44**: 1453–1458.

Morona R, Van Den Bosch L and Manning P (1995). Molecular, genetic, and topological characterization of O antigen chain regulation in *Shigella flexneri*. *Journal of Bacteriology* **177**: 1059–1068.

Mulford CA and Osborn MJ (1983). An intermediate step in translocation of lipopolysaccharide to outer membrane of *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences USA* **80**: 1159–1163.

Norman KN, Clawson ML, Strockbine NA, Mandrell RE, Johnson R, Ziebell K, Zhao S, Fratamico PM, Stones R, Allard MW and Bono JL (2015). Comparison of whole genome sequences from human and non-human *Escherichia coli* O26 strains. *Frontiers in Cellular and Infection Microbiology* **5**: 21.

Ørskov F and Ørskov I (1984). Serotyping of *Escherichia coli*. *Methods in Microbiology* **14**: 43–112.

Ørskov I, Ørskov F, Jann B and Jann K (1977). Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. *Bacteriological Reviews* **41**: 667–710.

Patel IR, Gangiredla J, Lacher DW, Mammel MK, Jackson SA, Lampel KA and Elkins CA (2016). FDA *Escherichia coli* identification (FDA-ECID) microarray: a pangenome molecular toolbox for serotyping, virulence profiling, molecular epidemiology, and phylogeny. *Applied and Environmental Microbiology* **82**: 3384–3394.

Plainvert C, Bidet P, Peigne C, Barbe V, Médigue C, Denamur E, Bingen E and Bonacorsi S (2007). A new O-antigen gene cluster has a key role in the virulence of the *Escherichia coli* meningitis clone O45:K1:H7. *Journal of Bacteriology* **189**: 8528–8536.

Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B and Beutler B (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**: 2085–2088.

Reeves PR (1994). Biosynthesis and assembly of lipopolysaccharides. In Ghuysen J-M and Hakenbeck R (eds) *New Comprehensive Biochemistry, Bacterial Cell Wall*. Elsevier Science Publisher, Amsterdam, Chap. 13, pp. 281–317.

Reeves PR and Wang L (2002). Genomic organization of LPS-specific loci. *Current Topics in Microbiological Immunology* **264**: 109–135.

Remis RS, MacDonald KL, Riley LW, Puhr ND, Wells JG, Davis BR, Blake PA and Cohen ML (1984). Sporadic cases of hemorrhagic colitis associated with *Escherichia coli* O157:H7. *Annals of Internal Medicine* **101**: 624–626.

Ren Y, Liu B, Cheng J, Liu F, Feng L and Wang L (2008). Characterization of *Escherichia coli* O3 and O21 O antigen gene clusters and development of serogroup-specific PCR assays. *Journal of Microbiological Methods* **75**: 329–334.

Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, Hebert RJ, Olcott ES, Johnson LM, Hargrett NT, Blake PA and Cohen ML (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New England Journal of Medicine* **308**: 681–685.

Russo TA and Johnson JR (2003). Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection* **5**: 449–456.

Samuel G and Reeves P (2003). Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydrate Research* **338**: 2503–2519.

Scheutz F, Cheasty T, Woodward D and Smith HR (2004). Designation of O174 and O175 to temporary O groups OX3 and OX7, and six new *E. coli* O groups that include verocytotoxin-producing *E. coli* (VTEC): O176, O177, O178, O179, O180 and O181. *APMIS* **112**: 569–584.

Smith JL, Fratamico PM and Gunther N (2007). Extraintestinal pathogenic *Escherichia coli* (ExPEC). *Foodborne Pathogens and Disease* **4**: 134–163.

Stenutz R, Weintraub A and Widmalm G (2006). The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiology Review* **30**: 382–403.

Stromberg ZR, Baumann NW, Lewis GL, Sevart NJ, Cernicchiaro N, Renter DG, Marx DB, Phebus RK, Moxley RA (2015). Prevalence of enterohemorrhagic *Escherichia coli* O26, O45, O103, O111, O121, O145, and O157 on hides and preintervention carcass surfaces of feedlot cattle at harvest. *Foodborne Pathogens and Disease* **12**: 631–638.

Sura R, Van Kruiningen HJ, DebRoy C, Hinckley LS, Greenberg KJ, Gordon Z and French RA (2007). Extraintestinal pathogenic *E. coli* (ExPEC) induced acute necrotizing pneumonia in cats. *Zoonoses and Public Health* **54**: 307–313.

Tseng M, Fratamico PM, Bagi L, Delannoy S, Fach P, Manning SD and Funk JA (2014). Diverse virulence gene content of Shiga toxin-producing *Escherichia coli* from finishing swine. *Applied and Environmental Microbiology* **80**: 6395–6402.

Wang L and Reeves PR (1998). Organization of *Escherichia coli* O157 O antigen gene cluster and identification of its specific genes. *Infection and Immunity* **66**: 3545–3551.

Wang Q, Ruan X, Wei D, Hu Z, Wu L, Yu T, Feng L and Wang L (2010). Development of a serogroup-specific multiplex PCR assay to detect a set of *Escherichia coli* serogroups based on the identification of their O-antigen gene clusters. *Molecular and Cellular Probes* **24**: 286–290.