The Journal of Microbiology, February 2007, p. 69-74
Vol. 45, No. 1

NOTE

Genetic Characterization of the *Escherichia coli* O66 Antigen and Functional Identification of its *wzy* Gene

Jiansong Cheng[1,2], Bin Liu[1,2], David A. Bastin[3], Weiqing Han[1,2], Lei Wang[1,2,3], and Lu Feng[1,2,*]

[1]*TEDA School of Biological Sciences and Biotechnology, Nankai University, 23# Hong Da Street, TEDA, Tianjin 300457, P. R. China*
[2]*Tianjin Key Laboratory of Microbial Functional Genomics, TEDA College, Nankai University, 23# Hong Da Street, TEDA, Tianjin 300457, P. R. China*
[3]*Tianjin Biochip Technology Corporation, 23 HongDa Street, TEDA, Tianjin 300457, P. R. China*

(Received April 14, 2006 / Accepted January 4, 2007)

***Escherichia coli* is a clonal species, and occurs as both commensal and pathogenic strains, which are normally classified on the basis of their O, H, and K antigens. The O-antigen (O-specific polysaccharide), which consists of a series of oligosaccharide (O-unit) repeats, contributes major antigenic variability to the cell surface. The O-antigen gene cluster of *E. coli* O66 was sequenced in this study. The genes putatively responsible for the biosynthesis of dTDP-6-deoxy-L-talose and GDP-mannose, as well as those responsible for the transfer of sugars and for O-unit processing were identified based on their homology. The function of the *wzy* gene was confirmed by the results of a mutation test. Genes specific for *E. coli* O66 were identified via PCR screening against representatives of 186 *E. coli* and *Shigella* O type strains. The comparison of intergenic sequences located between *galF* and the O-antigen gene cluster in a range of *E. coli* and *Shigella* showed that this region may perform an important function in the homologous recombination of the O-antigen gene clusters.**

Keywords: *Escherichia coli* O-antigen gene cluster, PCR typing, molecular evolution, O-antigen polymerase gene (*wzy*)

Lipopolysaccharide (LPS), a crucial component of the outer membranes of gram-negative bacteria, is comprised of three distinct regions: lipid A, an oligosaccharide core, and an O-specific polysaccharide (O-antigen). The O-antigen is constructed from many repeats of an oligosaccharide unit (O-unit), which normally harbors two to eight residues of a broad range of sugars. O-units are synthesized via the sequential transfer of a sugar phosphate and sugars from respective nucleotide sugars to the carrier lipid, undecaprenyl phosphate (UndP). O-units are then flipped across the membrane and polymerized to form polysaccharide chains, which are subsequently transferred to the independently synthesized core-lipid A, resulting in the formation of LPS (Reeves, 1994). Due to variations in the types of sugars present, their arrangements within the O-unit, and the linkage between O-units, the O-antigen is one of the most variable of the cell constituents. The surface O-antigen is subjected to intense selection by the host immune system and by bacteriophages, which may account to some degree for the maintenance of many different O-antigen forms within species including *E. coli*. Currently, more than 180 O serotypes have been detected in the *E. coli* scheme (Stenutz *et al.*, 2006).

Characteristically, genes specific for O-antigen synthesis tend to be located in a gene cluster, which maps between *galF* and *gnd* in *E. coli* and *Salmonella*. Previous studies have indicated that genetic variation in the gene clusters can cause major differences among the diverse forms of O-antigens. The inter- and intra-species lateral transfer of O-antigen genes is crucial in the distribution of polymorphic forms (Wang *et al.*, 2001; Reeves and Wang, 2002).

Genes associated with O-antigen synthesis are generally classified into three classes: (1) nucleotide sugar precursor synthesis genes; (2) genes for the transfer of sugars for the construction of the O-unit; (3) O-unit processing genes encoding for flippase (Wzx) and polymerase (Wzy). *wzx* and especially *wzy* genes always evidence high levels of specificity to individual O antigen forms and are difficult to identify, and thus harbor potential for use in PCR assays for the rapid detection of relevant strains.

In this study, the O-antigen gene cluster of *E. coli* O66 was sequenced. The genes responsible for the synthesis of the O-antigen were identified based on homology. Via a deletion test, the function of its *wzy* gene was confirmed. Via the screening of 186 *E. coli* and *Shigella* type strains, genes specific to *E. coli* O66 were identified. We also evaluated the non-encoding region located between the *galF* and the initial gene of the-O-antigen gene cluster, which may mediate the homologous recombination by which O serotypes are formed.

The *E. coli* O66 type strain was acquired from the Institute

* To whom correspondence should be addressed.
(Tel) 86-22-6622-9592; (Fax) 86-22-6622-9596
(E-mail) fenglu63@nankai.edu.cn

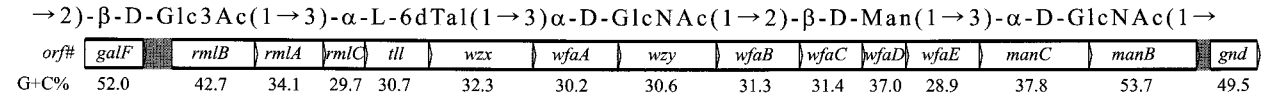

Fig. 1. O-antigen structure and O-antigen gene cluster of *E. coli* O66. GlcNAc, N-acetylglucosamine; Tal, talose; Man, mannose.

Table 1. Characteristics of the ORFs in the *E. coli* O66 O-antigen gene cluster

Gene name	Position of gene	Conserved domain(s)	Similar protein(s), strain(s) (GenBank accession no.)	% Identical / % Similar (No. of aa overlap)	Putative function of protein
rmlB	1137..2213	NAD-dependent epimerase/ dehydratase family (PF01370) *E* value=$1.3\times e^{-213}$	RmlB, *E. coli* O91(AAK60448)	97/98(358)	dTDP-D-glucose 4,6-dehydratase
rmlA	2210..3073	Nucleotidyl transferase (PF00483) *E* value=$3.7\times e^{-108}$	RmlA, *A. hydrophila* (AAM74475)	76/87(287)	Glucose-1-phosphate thymidylyl transferase
rmlC	3070..3621	dTDP-4-dehydrorhamnose 3,5-epimerase (PF00908) *E* value=$4.2\times e^{-93}$	RmlC, *A. hydrophila* (AAM22546)	71/82(183)	dTDP-6-deoxy-D-glucose-3,5 epimerase
tll	3611..4456	NAD dependent epimerase/ dehydratase family (PF01370) *E* value=$9.1\times e^{-6}$	UrdR, *S. fradiae* (AAF72551)	30/52(150)	dTDP-6-deoxy-L-lyxo-h exulose reductase
wzx	4453..5790	Polysaccharide biosynthesis protein (PF01943) *E* value=$1.9\times e^{-3}$	Wzx, *Y. pseudotuberculosis* O:1b (CAB63295)	31/55(437)	O-antigen flippase
wfaA	5783..6907	Glycosyl transferases group 1 (PF00534) E value=$5.2\times e^{-2}$	Putative glycosyl transferase, *S. boydii* 5 (AAL27331)	21/44(349)	Glycosyltransferase
wzy	6897..8159				O-antigen polymerase
wfaB	8159..9235	Glycosyl transferases group 1 PF00534 *E* value=$1.9\times e^{-32}$	Putative glycosyl transferase, *S. boydii* 6 (AAL27338)	31/52(357)	Glycosyltransferase
wfaC	9232..10047	Glycosyl transferase (PF00535) *E* value=$3.3\times e^{-35}$	Glycosyl transferase, *Shewanella baltica* OS155 (ZP_00581321)	36/54(275)	Glycosyltransferase
wfaD	10044..10592	Bacterial transferase hexapeptide (PF00132) *E* value=$4.6\times e^{-13}$	PatA, *P. mirabilis* (AAK38353)	44/61(172)	O-acetyl transferase
wfaE	10558..11553		CpsH, *Streptococcus iniae* (AAN63793)	35/57(326)	Glycosyltransferase
manC	11557..12987	Mannose-6-phosphate isomerase (PF01050) *E* value=$5.6\times e^{-120}$	ManC, *E. coli* CFT073 (AAN81013)	60/78(473)	Mannose-1-phosphate guanylyltransferase
manB	12974..14368	Phosphoglucomutase/phosphomannomutase (PF02879) *E* value=$5.9\times e^{-59}$	ManB, *E. coli* CFT073 (AAN81012)	97/98(451)	Phosphomannomutase

of Medical and Veterinary Science, in Adelaide, Australia. The construction of the DNaseI shotgun bank, sequencing, and data analysis were conducted using previously described methods (Feng *et al.,* 2004). For *E. coli* O66, a sequence of 15,137 bases encompassing 13 open reading frames (ORFs) from *galF* to *gnd*, was obtained. All of the ORFs evidence the same transcriptional direction from *galF* to *gnd* (Fig. 1). The putative genes were designated in accordance with the bacterial polysaccharide gene nomenclature (BPGN) system (www.microbio.usyd.edu.au/BPGD/default. htm). Functions of each gene within the O-antigen gene cluster were predicted via searches of available databases, and are summarized in Table 1. The DNA sequence of the *E. coli* O66 O-antigen gene cluster was deposited in the Genbank database under

the accession number DQ069297.

The *E. coli* O66 antigen harbors two particular sugars, a mannose and a 6-deoxy-L-talose (Fig. 1) (Jann *et al.,* 1995). GDP-mannose is synthesized from fructose-6-phosphate by the products of *manA*, *manB,* and *manC*. *manA* is located elsewhere in *E. coli* (Jensen and Reeves, 2001), whereas *manB* and *manC* are in the O-antigen gene cluster. *orf12* and *orf13* were determined to share a high level of identity with many *manB* and *manC* genes, and are therefore designated *manB* and *manC*, respectively. dTDP-6-deoxy-L-talose differs from dTDP-L-rhamose in terms of the stereochemistry of its C-4 carbon. In *Actinobacillus actinomycetemcomitans*, the pathway for dTDP-6-deoxy-L-talose biosynthesis has been characterized, and involves the products of *rmlA*, *rmlB*, *rmlC,* and *tll*. Tll and RmlD (dTDP-L-rhamnose synthase) catalyze the reduction of the same substrate, dTDP-6-deoxy-L-lyxo-4-hexulose, and the products can be identified on the basis of the stereospecificity of the reductase activity (Nakano *et al.,* 2000). Orf 1, Orf 2, and Orf 3, respectively, were determined to share a high level of identity with a variety of reported RmlB, RmlA, and RmlC forms. Orf4 shares 52% similarity with UrdR, which is an NDP-hexose 4-ketoreductase required for the formation of NDP-D-olivose in *Streptomyces fradiae* (Hoffmeister *et al.,* 2000). Therefore, *orf1, orf2, orf3, and orf4* have been theorized to synthesize dTDP-6-deoxy-L-talose, and were designated *rmlB*, *rmlA*, *rmlC,* and *tll,* respectively.

The only two ORFs encoding for the predicted membrane proteins are *orf5* and *orf7*. Orf5 harbors 10 predicted transmembrane segments. It also shares 55% and 53% similarity with the O-unit flippases of *Yersinia pseudotuberculosis* type O:1 b and *Y. pestis* KIM, respectively. Whereas Orf7 shares no similarity with any of the known proteins in the Genpept database, it harbors 11 predicted transmembrane segments with a large periplasmic loop of 74 amino acid residues, a topology characteristic of Wzy proteins (Daniels *et al.,* 1998). Therefore, *orf5* and *orf7* were identified as putative O-unit flippase and O-antigen polymerase genes, respectively.

The product of the *wecA* gene, which is located outside of the O antigen gene cluster, is known to be responsible for the transfer of GlcNAc-1-phosphate or GalNAc-1-phosphate to the carrier lipid, UndP, thereby initiating O-unit synthesis in the *E. coli* strains whose antigens utilize GlcNAc (N-acetylglucosamine) or GalNAc (N-acetylgalactosamine) as the first sugar in the O unit (Alexander and Valvano, 1994). Thus, we expected four sugar transferase genes within the O-antigen gene cluster on the basis of the *E. coli* O66 antigen structure. After a search of databases including the Genbank, COG, and Pfam databases, *orf6*, *orf8*, *orf9*, and *orf11* were implicated as putative transferase genes and designated *wfaA*, *wfaB*, *wfaC*, and *wfaE*, respectively.

The O-antigen of *E. coli* O66 is modified by an acetyl group on glucose. Orf10 was determined to share 44% sequence identity with PatA, a putative O-antigen acetylase in *Proteus mirabilis*. We propose that *orf10* is responsible for the O acetylation of glucose in *E. coli* O66 antigen, and have designated it *wfaD*.

We identified the *wzx* gene of the *E. coli* O66 O-antigen gene cluster on the basis of its sequence similarity and hydrophobic profiles. The Wzx protein of *E. coli* O66 was grouped with other known or putative Wzx proteins using the BlockMaker program. The consensus sequences of motifs as shown by BlockMaker analysis were used to run the program PSI-BLAST for a search of the Genpept database. In this analysis, the input Wzx proteins and many other distantly related Wzx proteins were retrieved after three iterations (*E* value= $1\times e^{-04}$). The analysis further verified the designation of the *wzx* gene.

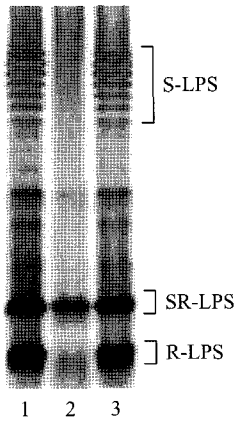

Fig. 2. Mutation analysis of *E. coli* O66 *wzy* gene. Membrane extracts were run on SDS-PAGE gels and stained by silver staining. Lanes: 1, G1074 (*E. coli* O66 type strain); 2, H1249 (G1074 missing the *wzy* gene); 3, H1250 (H1249 with plasmid pLW1072 carrying *E. coli* O66 *wzy* gene).

Owing to the low observed levels of sequence similarity, we did not obtain the expected results via the method mentioned above when identifying the *wzy* gene. Thus, we confirmed the identity of the *wzy* gene of *E. coli* O66 via a mutation test. In *E. coli* and *Shigella*, *wzy* mutation results in a particular semi-rough LPS (SR-LPS) phenotype of only one O-unit, attached to core-lipid A (Daniels *et al.,* 1998). The *wzy* gene of *E. coli* O66 was supplanted by a chloramphenicol acetyltransferase (CAT) gene, using the RED recombination system of phage lambda (Datsenko and Wanner, 2000; Yu *et al.,* 2000). The CAT gene was then PCR-amplified from the pKK232-8 plasmid (Pharmacia) using the following primers: wl-1066; 5'-GCACCAATAATATTATCGTT-GTACT ATATTTCAACAACAAAGGAGCTAAGGAAGCTAAAAT GG-3' and wl-1067; 5'-CCCAAAAGAAATCTTAGATAGCA AAAGAAATACTCGCCAGTAAAAAAATTACGCCCCG-3' which bound to the 5' and 3' ends of the gene, with each primer harboring 40 bp based on the *E. coli* O66 DNA flanking the *wzy* gene. The PCR products were then transformed into the *E. coli* O66 strain harboring pKD20, and chloramphenicol-resistant transformants were selected following the induction of the RED genes, in accordance with the protocol developed by Datsenko and Wanner (Datsenko and Wanner, 2000). PCR using primers specific for the CAT and *wzy* genes was conducted in order to confirm the replacement.

The mutant strain H1249 generated only a lipid-A/core and one oligosaccharide unit of the LPS, whereas the wild-type strain generated normal LPS (Fig. 2). The *wzy* gene was PCR-amplified using the primers wl_4898; 5'-TAATCCA TGGAAAATAAGATATC-ATTCTATAT-3' and wl_4899; 5'-A

Table 2. PCR specificity test with *E.coli* O66 genes

Gene	Forward primers	Reverse primers	Length of the PCR fragment (bp)	Annealing temperature (°C) of PCR
wzx	wl-914, 5'-TTGCGATGGCAGGAATAA-3'	wl-915, 5'-GCTCCAACGGCTAGTGAA-3'	469	50
	wl-916, 5'-TGGATGTGGGTTAGTTT-3'	wl-917, 5'-CTGATGTAATTCTGGGTA-3'	345	50
wzy	wl-918, 5'-CGAGCAAATTAAATCCAC-3'	wl-919, 5'-TCAACACTAAACGAAACG-3'	301	50
	wl-920, 5'-ATTATTAGGTGATTTTCGAACT-3'	wl-921, 5'-CAAGTGGATTTAATTTGCTCGC-3'	290	50

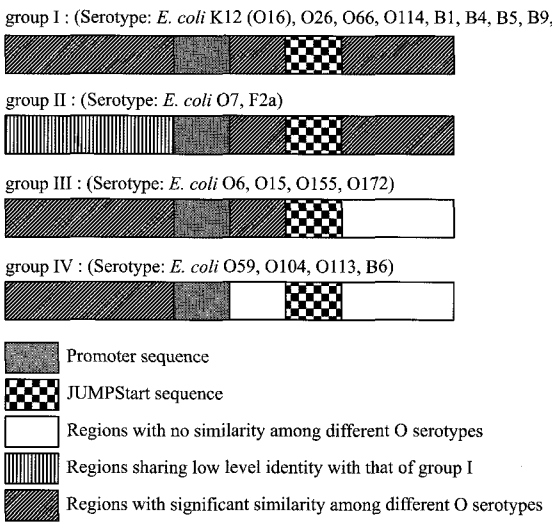

Fig. 3. Patterns of intergenic region between *galF* and O-antigen gene cluster. Group I sequence is used as reference. B, *Shigella boydii* D, *Shigella dysenteriae* F, *Shigella flexneri*.

TAGGATCCATACTTTTCATTATTTTAGC-3', and the resultant PCR product was cloned into pTRC99A to construct the pLW1072 plasmid. The mutation of the semi-rough LPS (SR-LPS) phenotype was trans-complemented by the pLW1072 plasmid, which harbors the corresponding *wzy* gene, to the smooth phenotype (Fig. 2). Therefore, we are confident of the designation of the *wzy* genes.

The *wzx* and *wzy* from different O-antigen gene clusters evidence low-level sequence identity and are specific to individual forms of the O-antigen. Primers were designed on the basis of the *wzx* and *wzy* sequences of *E. coli* O66 (Table 2). Two pairs of primers for each gene were employed in the screening of DNA pools consisting of 186 O type strains of *E. coli* and *Shigella* (Feng *et al.*, 2004). The quality of the chromosomal DNA was assessed via the PCR amplification of the *mdh* gene. All primer pairs generated no bands in any of the O type strains of *E. coli* and *Shigella*, with the exception of the pool containing *E. coli* O66. Therefore, the primer pairs utilized in this study can be thought to be specific to *E. coli* O66, and may be employed in the development of a PCR assay for the identification of this type.

Large non-encoding regions have been detected between the *galF* gene and the O-antigen gene cluster of *E. coli* and *Shigella*. This region harbors two *cis*-regulated elements that are vital for the expression of the O-antigen gene cluster. One of these is the JUMPStart sequence, a 39 bp highly conserved sequence which has been detected in a variety of different polysaccharide clusters (Hobbs and Reeves, 1994; Leeds and Welch, 1997). The other is the promoter for the O-antigen gene cluster, which is located between the JUMPStart sequence and the *galF* gene. Promoters of O-antigen gene clusters have been detected in *Salmonella* LT2 and *E. coli* O7, both of which harbor the -35 and -10 sequences "TTGTTT" and "GACAAT", and the base G five bases downstream of the -10 element is the initial transcription site (Jiang *et al.*, 1991; Marolda and Valvano, 1998). Using the ClustalW program (Thompson *et al.*, 1994), we compared the putative promoter sequences and the JUMPStart sequences of the O-antigen gene clusters in 25 *Shigella* and *E. coli* O serogroups, 22 of which were from published data, and those of *E. coli* O59 and O155 were from separate studies (Guo *et al.*, 2005). The results indicated that they were highly conserved. The consensus sequences of TTGTT TccagagcggattggtaaGACAATtagcG and cagtgc(t/a)ctggtagctgt (a/t)aagccaggggcggtagcgtg were shown for the promoter sequences and the JUMPStart sequences, respectively. An 8 bp *ops* subsequence (ggcggtag) within JUMPStart has been determined to contribute to an increase in the transcription of distal genes of the *E. coli* hemolysin gene cluster (Bailey *et al.*, 1996; Leeds and Welch, 1997). This highly-conserved motif is located at the 3' end of the JUMPStart sequences, and the only variation observed within this sequence is in *E. coli* O15, in which an A is substituted for the first G.

The *rml* gene set has been observed in many O-antigen gene clusters in *E. coli* and *Shigella*. Four genes grouped together for dTDP-rhamnose biosynthesis are generally detected in the following order: *rmlB, rmlD, rmlA,* and *rmlC,* and are located at the 5' end of the O-antigen gene clusters in *E. coli* and *Shigella* (Wang *et al.*, 2001). The *RmlA*, *rmlB,* and *rmlC* genes may also be involved in the synthesis pathways of other nucleotide sugars. We compared the intergenic regions between the *galF* gene and the O-antigen gene cluster among 25 different O serogroups, 12 of which harbored the *rmlB gene* at the 5' end of the gene cluster, and 20 of which can be classified into 4 groups (Fig. 3). It was determined that the *galF-rmlB* intergenic regions (371-372 bp in length) share a high level of identity (between 93% and 99%) among 10 serogroups [*E. coli* K12 (O16), O26, O66, O114, *S. boydii* O1, O4, O5, O9, O11, and *S. dysenteriae* O1], with the exceptions of *E. coli* O7 and *S. flexneri* 2a. We also determined that the intergenic regions in many other O serogroups without the *rmlB* gene set at the 5' end of the O-antigen gene cluster harbor segments which share greater than 90% identity with the corresponding portions

of the *galF-rmlB* intergenic regions. The *galF-rmlB* intergenic regions, with the exceptions of those in *E. coli* O7 and *S. flexneri* 2a, are referred to herein as group I. In group I, the non-encoding sequence located between *galF* and the first O-antigen gene evidences a length of 371 to 372 bp, the promoter region is from positions 160 to 209, and the JUMPStart sequence stretches from positions 259 to 297. In *E. coli* O7 and *S. flexneri* 2a, the region between the *galF* gene and the promoter share 69 to 73% identity with those of group I, and the region between the promoter sequence and the *rmlB* genes share 86 to 90% identity with those of group I. The intergenic regions located upstream of the O-antigen gene cluster in *E. coli* O7 and *S. flexneri* 2a are referred to herein as group II. The genes of the *rml* gene set in *E. coli* O7 and *S. flexneri* 2a also diverge significantly from those of the group I members. Four serogroups, including *E. coli* O6, O15, O155, and O172 are grouped into group III, in which the regions located between the *galF* gene and the JUMPStart sequence share 91% to 99% identity with those of group I. However, the region located between the JUMPStart sequence and the first O-antigen gene in each of the group III strains shares a low degree of identity and is of a different length with those of group I, as well as between each other. *E. coli* O59, O104, O113 and *S. boydii* O6 were grouped into group IV, in which the region between the *galF* gene and the promoter sequence was found to share 95% to 99% identity with those of group I, and the region between the promoter sequence and the first O-antigen gene shares no significant homology. The *galF*-O-antigen gene cluster intergenic regions of *E. coli* O55, O111, O157, *S. dysenteriae* O7, and *S. boydii* O13 diverge substantially, and thus cannot be grouped. The pattern of variation in the intergenic region has been suggested to be induced by recombination events in which the *galF*-O-antigen gene cluster intergenic region, in addition to the O-antigen genes and flanking genes can be reshuffled after an O-antigen gene cluster arrives in *Shigella* and *E. coli*. This indicates that the different forms of the intergenic region have a similar nature, and relate to different groups of *E. coli*. Rhamnose is a common sugar, which is widely distributed throughout the O-antigens of Gram-negative bacteria (Li and Reeves, 2000). The presence of members in groups III and IV indicates that the O-antigen gene clusters containing the *rml* gene set at the 5' end may represent an ancient form in *E. coli* and *Shigella*. Additionally, two highly conserved regions: the promoter sequences and JUMPStart sequences, are putative homologous recombination sites of the O-antigen gene clusters.

Acknowledgements

This work was supported by NSFC General Programs (30270029 and 30530010) to LW and LF.

References

Alexander, D.C. and M.A. Valvano. 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing N-acetylglucosamine. *J. Bacteriol.* 176, 7079-7084.

Bailey, M.J., C. Hughes, and V. Koronakis. 1996. Increased distal gene transcription by the elongation factor RfaH, a specialized homologue of NusG. *Mol. Microbiol.* 22, 729-737.

Daniels, C., C. Vindurampulle, and R. Morona. 1998. Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). *Mol. Microbiol.* 28, 1211-1222.

Datsenko, K.A. and B.L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 97, 6640-6645.

Guo, H., Q. Kong, J. Cheng, L. Wang, and L. Feng. 2005. Characterization of the *Escherichia coli* O59 and O155 O-antigen gene clusters: the atypical *wzx* genes are evolutionary related. *FEMS Microbiol. Lett.* 248, 153-161.

Feng, L., S.N. Senchenkova, J. Yang, A.S. Shashkov, J. Tao, H. Guo, G. Zhao, Y.A. Knirel, P. Reeves, and L. Wang. 2004. Structural and Genetic Characterization of the *Shigella boydii* Type 13 O Antigen. *J. Bacteriol.* 182, 383-392.

Hobbs, M. and P.R. Reeves. 1994. The JUMPstart sequence: a 39 bp element common to several polysaccharide gene clusters. *Mol. Microbiol.* 12, 855-856.

Hoffmeister, D., K. Ichinose, S. Domann, B. Faust, A. Trefzer, G. Drager, A. Kirschning, C. Fischer, E. Kunzel, D. Bearden, J. Rohr, and A. Bechthold. 2000. The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. *Chem. Biol.* 7, 821-831.

Jann, B., A. Shashkov, V. Torgov, H. Kochanowski, G. Seltmann, and K. Jann. 1995. NMR investigation of the 6-deoxy-L-talose-containing O45, O45-related (O45rel), and O66 polysaccharides of *Escherichia coli*. *Carbohydr. Res.* 278, 155-165.

Jensen, S.O. and P.R. Reeves. 2001. Molecular evolution of the GDP-mannose pathway genes (*manB* and *manC*) in *Salmonella enterica*. *Microbiology* 147, 599-610.

Jiang, X.M., B. Neal, F. Santiago, S.J. Lee, L.K. Romana, and P.R. Reeves. 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). *Mol. Microbiol.* 5, 695-713.

Leeds, J.A. and R.A. Welch. 1997 Enhancing transcription through the *Escherichia coli* hemolysin operon, *hlyCABD*: RfaH and upstream JUMPStart DNA sequences function together via a postinitiation mechanism. *J. Bacteriol.* 179, 3519-3527.

Li, Q. and P.R. Reeves. 2000. Genetic variation of dTDP-L-rhamnose pathway genes in *Salmonella enterica*. *Microbiology.* 146, 2291-2307.

Marolda, C.L. and M.A. Valvano. 1998. Promoter region of the *Escherichia coli* O7-specific lipopolysaccharide gene cluster: structural and functional characterization of an upstream untranslated mRNA sequence. *J. Bacteriol.* 180, 3070-3079.

Nakano, Y., N. Suzuki, Y. Yoshida, T. Nezu, Y. Yamashita, and T. Koga. 2000. Thymidine diphosphate-6-deoxy-L-lyxo-4-hexulose reductase synthesizing dTDP-6-deoxy-L-talose from Actinobacillus actinomycetemcomitans. *J. Biol. Chem.* 275, 6806-6812.

Reeves, P.P. and L. Wang. 2002 Genomic organization of LPS-specific loci. *Curr. Top. Microbiol. Immunol.* 264, 109-135.

Reeves, P.R. 1994. Biosynthesis and assembly of lipopolysaccharide, p. 281-314. *In* A. Neuberger and L.L.M. van Deenen (ed.), Bacterial cell wall-1994. Elsevier Science Publishers, Amsterdam, Netherlands.

Stenutz, R., A. Weintraub, and G. Widmalm. 2006. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol. Rev.* 30, 382-403.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W-Impoving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specifc gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.

Wang, L., W. Qu, and P.R. Reeves. 2001. Sequence analysis of four *Shigella boydii* O-antigen loci: implication for *Escherichia coli* and *Shigella* relationships. *Infect. Immun.* 69, 6923-6930.

Yu, D., H.M. Ellis, E.C. Lee, N.A. Jenkins, N.G. Copeland, and D.L. Court. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 91, 5978-5983.