Molecular Analysis of the *Enterobacter sakazakii* O-Antigen Gene Locus[▽]

N. Mullane,¹ P. O'Gaora,² J. E. Nally,³ C. Iversen,⁴ P. Whyte,¹ P. G. Wall,⁵ and S. Fanning^{1*}

Centres for Food Safety and Food-borne Zoonomics, UCD Veterinary Sciences Centre,¹ School of Agriculture, Food Science & Veterinary Medicine,³ Conway Institute of Biomolecular & Biomedical Research,² and School of Public Health & Population Sciences,⁵ University College Dublin, Belfield, Dublin 4, Ireland, and Institute for Food Safety & Hygiene, Vetsuisse Faculty, University of Zurich, CH-8057 Zurich, Switzerland⁴

Received 11 October 2007/Accepted 16 April 2008

Nucleotide polymorphism associated with the O-antigen-encoding locus, *rfb*, in *Enterobacter sakazakii* was determined by PCR-restriction fragment length polymorphism analysis. Based on the analysis of these DNA profiles, 12 unique banding patterns were detected among a collection of 62 strains from diverse origins. Two common profiles were identified and were designated serotypes O:1 and O:2. DNA sequencing of the 12,500-bp region flanked by *galF* and *gnd* identified 11 open reading frames, all with the same transcriptional direction. Analysis of the proximal region of both sequences demonstrated remarkable heterogeneity. A PCR assay targeting genes specific for the two prominent serotypes was developed and applied for the identification of these strains recovered from food, environmental, and clinical samples.

Enterobacter sakazakii is an opportunistic pathogen that has been associated with food-borne illness in neonates (12, 24, 25). Powdered infant formula has been implicated as a source of intrinsic *E. sakazakii*, but extrinsic contamination at reconstitution and improper handling have also been linked with cases of infection (4, 25). Organisms identified as *E. sakazakii* have been shown to represent multiple species, and an alternative classification of *E. sakazakii* as five species within a novel genus, “*Cronobacter*”, has been proposed (20).

Lipopolysaccharide (LPS) is a major component of the outer membrane of gram-negative bacteria. This is composed of three parts, a complex lipid, called lipid A (consisting of sugars and fatty acids), that anchors the structure to the outer membrane, a conserved core oligosaccharide, and variable polysaccharide side chains (O antigen) that extend from the latter core. The O antigen is a major surface antigen present in gram-negative bacteria, and it is responsible for serological diversity. In these bacteria, LPS contains many oligosaccharide units (individual O units consisting of between 3 and 6 sugars) with typically between 10 and 30 repeats (23, 42, 47). Genes involved in O-antigen synthesis and downstream assembly map to the *rfb* locus located between the *galF* and *gnd* genes in many *Enterobacteriaceae*. The O-antigen locus varies in size for each serotype depending on the sugar composition and complexity of the antigen structure (34). O-antigen serotypes emerge as a consequence of the gene content rather than sequence variation at this locus (44). Usually three gene types map to the O-antigen gene cluster, and these include the following: (i) genes that code for enzymes involved in the synthesis of sugars forming the O subunit, (ii) genes that encode glycosyltransferases, involved in the assembly of sugar substituents in the O subunit, and (iii) those genes that code for the transporter

(*wzx*) and polymerase (*wzy*) proteins necessary for processing and assembly of the O antigen from the O subunit.

The O-antigen gene cluster of *E. sakazakii* has not been previously characterized. This article describes the molecular characterization of the *rfb* locus in *E. sakazakii*. Two major serotypes were identified.

MATERIALS AND METHODS

Bacterial strains. A summary of the *E. sakazakii* strains used in this study is presented in Table 1 with nomenclature according to the proposed alternative classification as “*Cronobacter*” species (20). The collection contained 62 isolates of clinical, food, or environmental origin. All isolates were confirmed to be *E. sakazakii* by real-time PCR using a primer set and probe targeting the *dnaG* gene on the macromolecular synthesis operon (35). All isolates in this study were distinguished by pulsed-field gel electrophoresis, and some were further characterized by 16S rRNA sequencing (data not shown).

DNA isolation. Total DNA was prepared using the Wizard Genomic DNA purification kit (Promega, Madison, WI). DNA concentrations were determined spectrophotometrically. The integrity of the purified template DNA was assessed by conventional agarose gel (1% [wt/vol]) electrophoresis, and DNA preparations were stored at 4°C.

O-antigen gene cluster amplification. The O-antigen gene cluster was identified from the sequence information of *Enterobacter sakazakii* type strain ATCC BAA-894 provided by Washington University School of Medicine (http://genome.wustl.edu/pub/organism/Microbes/Enteric_Bacteria/Enterobacter_sakazakii/assembly/Enterobacter_sakazakii-4.0/). Primer pairs Esrfb-F2 (5'-TAC CCA CTC CTC CAA GAA CG-3') and Esrfb-R2 (5'-TTT CGC CGT AGT TCA GAT CC-3'), complementary to *galF* and *gnd*, respectively, were designed to amplify the O-antigen gene cluster (MWG Biotech AG, Ebersberg, Germany). A long-range PCR was performed using the Bio-X-ACT Long DNA polymerase kit (Bioline, London, England). PCR amplification was performed in 50-μl volumes containing 1× OptiBuffer, containing 4 mM MgCl₂, 2 mM deoxynucleoside triphosphates, 0.5 μM of each primer, 4 U Bio-X-ACT Long DNA polymerase, 100 ng template DNA, and PCR-grade water (Invitrogen, CA). The PCR cycles consisted of initial denaturation at 94°C for 2 min and thirty-five cycles of 94°C for 10 s, 52°C for 30 s, and 68°C for 15 min. A final elongation step at 72°C for 10 min was used. Amplified products were visualized following electrophoresis through a conventional 1.5% (wt/vol) agarose gel stained with ethidium bromide (10 mg/ml) in a 0.5× Tris-borate-EDTA (TBE) buffer. Gels were visualized and photographed using the Gel Doc 2000 system (Bio-Rad, Hercules, CA). Amplicon sizes were estimated using Hyperladder IV (Bioline, London, England) and a 1-kb ladder (New England Biolabs, Herts, United Kingdom).

Restriction analysis (O-antigen restriction fragment length polymorphism [RFLP]). Ten microliters of O-antigen PCR product was digested with 12.5 U MboII (New England Biolabs, Herts, United Kingdom) in the supplied buffer.

* Corresponding author. Mailing address: Centres for Food Safety & Food-borne Zoonomics, UCD Veterinary Sciences Centre, University College Dublin, Belfield, Dublin 4, Ireland. Phone: (353-1) 716 6082. Fax: (353-1) 716 6091. E-mail: sfanning@ucd.ie.

[▽] Published ahead of print on 25 April 2008.

TABLE 1. Members of the *Enterobacteriaceae* used in this study

Species	Strain	Origin	Serotype
<i>Enterobacter sakazakii</i> ("Cronobacter" spp.)			
<i>C. sakazakii</i>	CFS06	Acid casein powder	1
<i>C. sakazakii</i>	CFS10	Acid casein powder	2
<i>C. sakazakii</i>	CFS104	Milk powder	1
<i>C. sakazakii</i>	CFS122	Environmental	1
<i>C. malonaticus</i> ^{*a}	CFS129	Environmental	ND ^b
<i>C. sakazakii</i>	CFS131	Milk powder	1
<i>C. sakazakii</i>	CFS136	Environmental	2
<i>C. sakazakii</i>	CFS149	Environmental	2
<i>C. sakazakii</i>	CFS153	Environmental	1
<i>C. sakazakii</i>	CFS164	Environmental	1
<i>C. sakazakii</i>	CFS173	Milk powder	ND
<i>C. sakazakii</i>	CFS175	Environmental	2
<i>C. sakazakii</i>	CFS176	Environmental	2
<i>C. dublinensis</i> [*]	CFS237	Milk powder	ND
<i>C. sakazakii</i>	CFS1001	Environmental	1
<i>C. sakazakii</i>	44	Environmental	1
<i>C. sakazakii</i>	71	Environmental	2
<i>C. sakazakii</i>	79	Environmental	2
<i>C. sakazakii</i>	80	Environmental	2
<i>C. sakazakii</i>	82	Environmental	ND
<i>C. sakazakii</i>	88	Environmental	1
<i>C. sakazakii</i>	90	Environmental	1
<i>C. sakazakii</i>	93	Environmental	1
<i>C. sakazakii</i>	102	Environmental	ND
<i>C. sakazakii</i>	109	Environmental	2
<i>C. sakazakii</i>	228N	Milk powder	2
<i>C. sakazakii</i>	305N	Milk powder	1
<i>C. sakazakii</i>	343	Environmental	2
<i>C. sakazakii</i>	344	Environmental	ND
<i>C. sakazakii</i>	366	Environmental	ND
<i>C. sakazakii</i>	E604	Clinical	2
<i>C. sakazakii</i>	E632	Infant food	1
<i>C. sakazakii</i>	E770	Milk powder	1
<i>C. sakazakii</i>	E775	Milk powder	2
<i>C. sakazakii</i>	E784	Neonate	1
<i>C. sakazakii</i>	E787	Clinical	2
<i>C. sakazakii</i>	E824 (CDC 0996-77)	Human spinal (fluid)	1
<i>C. malonaticus</i> [*]	E825 (CDC 1058-77)	Human (breast abscess)	ND
<i>C. sakazakii</i>	E826 (CDC 1059-77)	Unknown	ND
<i>C. sakazakii</i>	E827 (CDC 0407-77)	Human (septum)	1
<i>C. sakazakii</i>	E828 (CDC 3128-77)	Human (septum)	1
<i>C. malonaticus</i> [*]	E829 (CDC 1716-77)	Human (Blood)	ND
<i>C. sakazakii</i>	E830 (CDC 9369-75)	Unknown	2
<i>C. sakazakii</i>	E837	Clinical	2
<i>C. sakazakii</i>	E839	Neonate	ND
<i>C. sakazakii</i>	E844	Neonate	2
<i>C. sakazakii</i>	E846	Neonate	2
<i>C. turicensis</i> [*]	E866	Neonate	ND
<i>C. sakazakii</i>	E891	Clinical	1
<i>C. sakazakii</i>	E892	Clinical	2
<i>C. sakazakii</i>	E899 (CDC 0289-81)	Clinical	2
<i>C. sakazakii</i>	E900 (CDC 0322-78)	Clinical	2
<i>C. sakazakii</i>	E901 (CDC 1123-79)	Clinical	2
<i>C. sakazakii</i>	ATCC 12868	Unknown	2
<i>C. sakazakii</i>	NCTC 8155 (CDC 9081-75)	Milk powder	2
<i>C. sakazakii</i>	NCTC 9238	Human (abdominal puss)	2
<i>C. sakazakii</i>	NCTC 29004	Unknown	2
<i>C. genomospecies</i> 1 [*]	NCTC 9529	Water	ND
<i>C. sakazakii</i>	NCTC 11467 (ATCC 29544)	Human (throat)	1
<i>C. sakazakii</i>	ATCC BAA 893	Milk powder	1
<i>C. sakazakii</i>	ATCC BAA 894	Neonate	1
<i>C. muytjensii</i> [*]	ATCC 51329	Unknown	ND
Non- <i>Enterobacter sakazakii</i> ("Cronobacter" spp.)			
<i>Escherichia coli</i> (EPEC)		ND	ND
<i>Escherichia coli</i> (ETEC)		ND	ND

Continued on following page

TABLE 1—Continued

Species	Strain	Origin	Serotype
<i>Escherichia coli</i>	K12 NC10538:06	ND	ND
<i>Salmonella enterica</i> serovar Anatum		ND	ND
<i>Salmonella enterica</i> serovar Agona	NCTC11377	ND	ND
<i>Salmonella enterica</i> serovar Typhimurium	LT2 NCTC12416	ND	ND
<i>Salmonella enterica</i> serovar Enterica PT4	NCTC13349	ND	ND
<i>Citrobacter youngae</i>		ND	ND
<i>Citrobacter sedlakii</i>		ND	ND
<i>Citrobacter braakii</i>		ND	ND
<i>Citrobacter koseri</i>		ND	ND
<i>Citrobacter freundii</i>	ATCC 8090	ND	ND
<i>Citrobacter koseri</i>	NCTC10768	ND	ND
<i>Klebsiella oxytoca</i>	ATCC 43086	ND	ND
<i>Klebsiella pneumoniae</i>	ATCC 13883	ND	ND
<i>Yersinia enterocolitica</i>		ND	ND
<i>Enterobacter helveticus</i>		ND	ND
<i>Escherichia hermannii</i>		ND	ND
<i>Enterobacter hormaechei</i>		ND	ND
<i>Enterobacter cancerogenes</i>		ND	ND
<i>Escherichia vulneris</i>		ND	ND
<i>Enterobacter cloacae</i>	NCTC11933	ND	ND
<i>Enterobacter aerogenes</i>	NCTC10006	ND	ND
<i>Enterobacter gergoviae</i>	NCTC11434	ND	ND
<i>Enterobacter agglomerans</i>	NCTC9381	ND	ND

^a *, proposed species names for taxonomic change according to Iversen et al. (20).

^b ND, not determined.

The mixture was incubated at 37°C for 2 h, followed by a final denaturation step of 70°C for 10 min to denature the endonuclease as it remains attached to DNA after cleavage (6). Failure to denature the enzyme would result in disturbed electrophoretic migration of DNA. The total volume was visualized by electrophoresis through a conventional 1.5% (wt/vol) agarose gel in a 1.0× TBE buffer at 100 V for 3 h. Fragment sizes were estimated using a 100-bp DNA ladder (New England Biolabs, Herts, United Kingdom) and DNA molecular weight marker XVI (Roche Applied Science, Indianapolis, IN). Following electrophoresis, gels were stained for 25 min in 400 ml deionized water containing 40 µl ethidium bromide (10 mg/ml) and then destained with two 20-min washes with 500 ml distilled water. Gels were visualized and photographed using the Gel Doc 2000 system (Bio-Rad, Hercules, CA). DNA fingerprints were stored as TIFFs (tagged image format files) and imported into the BioNumerics software (Applied Maths, Belgium), where dendrograms were created using the Dice coefficient using the unweighted pair group method with arithmetic mean. Band position tolerance and optimization values of 1.5% were used for all analyses.

The corresponding complete O-antigen locus from all isolates was digested at least twice from two separate DNA preparations in order to verify the reproducibility of the restriction digest.

Sequence analysis. The O-antigen gene clusters of strains NCTC 11467 and NCTC 8155 were sequenced by double-stranded sequencing and extended by primer walking (GATC-Biotech, Contance, Germany). The Blockmaker program was used to search conserved motifs (19). The NCBI BLAST service was used to search sequence databases for sequences with homology to the open reading frames (ORFs) that were identified by ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) (2). Sequence alignment and comparisons were performed using the ClustalW program (<http://www.ebi.ac.uk/cluster/>) (40). The dense alignment surface method, described by Cserzo et al. (8), was used to identify potential transmembrane segments (<http://www.sbc.su.se/~miklos/DAS>).

Serotype O:1- and O:2-specific PCR. Total DNA was prepared using the Wizard Genomic DNA purification kit (Promega, Madison, WI) and quantified by spectrophotometry (NanoDrop, Wilmington, DE). The O:1 serotype-specific PCR used the forward primer ESO:1F (5'-CAC GTT CGC CCT GCA AAA AT-3') and a reverse primer, ESO:1R (5'-GCA AGC GGC CAG ACT GGA TA-3'), designed from sequence information within *wehC* to generate a 341-bp amplicon. For O:2 serotype PCR, a forward primer, ESO:2F (5'-TCC TGC ATT TGT GGA TTT TGC-3'), and a reverse primer, ESO:2R (5'-AAC GCA TTG CGC TTG AGA AA-3'), were designed from sequence information within *wehI* to generate a 329-bp amplicon. These genes were chosen based on sequence analysis that revealed that these targets were highly conserved. PCR amplifica-

tion was performed with 50-µl volumes containing 1× PCR buffer containing 1.5 mM MgCl₂ (New England Biolabs, Herts, United Kingdom), 100 mM deoxynucleoside triphosphates, 0.1 µM of each primer (MWG Biotech AG, Ebersberg, Germany), 2.5 U *Taq* DNA polymerase (New England Biolabs, Herts, United Kingdom), 100 ng template DNA, and PCR-grade water (Invitrogen, CA). Thermal PCR conditions for O:1 were 95°C for 1 min, followed by 35 cycles of 94°C for 1 min, 56°C for 1 min (annealing), and 72°C for 1 min with a final extension of 72°C for 5 min, while O:2 PCR was annealed at 51°C for 1 min. Amplified products were analyzed by electrophoresis through a 1% (wt/vol) agarose gel at 80 V for 60 min in 1× TBE. Gels were visualized as per restriction analysis.

The specificity of each PCR was assessed using the 62 study isolates and 25 non-*E. sakazakii* *Enterobacteriaceae* strains (Table 1).

Analysis of *E. sakazakii* LPS. (i) **Cell lysis and proteinase K digestion.** Preparation of lysed cells and proteinase K digestion were performed with 12 strains as described previously by Pantophlet et al., (26) with minor modifications. Briefly, overnight cultures were subcultured after 18 h into fresh prewarmed tryptone soy broth at 37°C. Cultures were grown to late exponential phase to an optical density at 610 nm of 0.8 to give 1 × 10⁸ cells per ml. Incubation times were strain specific and ranged from 2 to 2.5 h. Following three washing and centrifugation steps with phosphate-buffered saline, the bacterial pellets were solubilized in sample buffer A (200 µl, 62.5 mM Tris-HCl [pH 8.0], 2% [wt/vol] sodium dodecyl sulfate [SDS], and 10% [vol/vol] glycerol) and were stored at -20°C. For proteinase K digestion, each sample was heated to 100°C for 10 min, followed by addition of proteinase K (750 µg in 30 µl sample buffer A [Roche, Mannheim, Germany]) and overnight digestion at 60°C. The digested product could be stored at -20°C if required.

(ii) **SDS-PAGE and silver staining.** Preparations of proteinase K digests were subjected to SDS-polyacrylamide gel electrophoresis (PAGE) with a 4% (wt/vol) stacking and a 15% (wt/vol) separating gel. Twenty microliters of each whole-cell lysate was mixed with 20 µl sample loading buffer (62.5 mM Tris-HCl [pH 8.0], 5% 2-mercaptoethanol, 2% [wt/vol] SDS, 0.01% [wt/vol] bromophenol blue, and 10% [vol/vol] glycerol). Five micrograms LPS from *Salmonella enterica* serotype Minnesota was obtained from Sigma-Aldrich and was used as a control. A rainbow high-range molecular weight marker was used as the size standard (Amersham, Buckinghamshire, England). Following electrophoresis, LPS was visualized by silver staining (43).

Nucleotide sequence accession numbers. The sequences of the O-antigen gene clusters of *E. sakazakii* strains NCTC 11467 and NCTC 8155 were deposited in the GenBank database under accession numbers EU076545 (serotype O:1) and EU076545 (serotype O:2), respectively.

FIG. 1. Restriction length profiles of *rfb*-encoding locus of *E. sakazakii* following MboII digestion. Lane 1, ATCC BAA-894; lane 2, E824; lane 3, E825; lane 4, NCTC 11467; lane 5, NCTC 8155; lane 6, 336; lane 7, 344; lane 8, 109; lane 9, CFS06; lane 10, E787; lane 11, 82; lane 12, 102; lane M1, 100-bp DNA ladder (New England Biolabs, Hertfordshire, England); lane M2, DNA molecular weight marker XVI (Roche, Mannheim, Germany). Black and gray boxes indicate profiles of strains corresponding to O:1 and O:2 serotypes, respectively.

RESULTS

O-antigen gene cluster amplification and MboII restriction analysis. Long PCR products, with a size range of between 11 and 12.5 kbp, were detected in all 62 *E. sakazakii* strains tested (data not shown), using primers targeting the flanking *galF* and *gnd* genes. Each amplicon was digested with MboII. Reproducible restriction patterns were obtained in each case, with DNA fragment sizes ranging from 250 through 2,000 bp (Fig. 1). A number of recurring DNA profiles were evident, containing from 9 to 13 restriction fragments (Fig. 1). A 400-bp MboII DNA fragment was observed in ATCC BAA-894 (Fig. 1, lane 1) that was not present in other similar O:1 RFLP profiles, including those of E824, NCTC11467, and CFS06 (Fig. 1). An *in silico*-generated MboII restriction map of the 12-kbp region, based on the results from sequencing, failed to identify the corresponding 400-bp DNA fragment. This fragment could be derived from a number of smaller, completely digested DNA fragments within this mix. Alternatively, a polymorphic site in ATCC BAA-894 may have resulted in the generation of a fragment of this size in this isolate. This DNA fragment was not considered when assigning the serotype.

Using fragment analysis software, all profiles were grouped into 12 clusters, consisting of two large cluster groups comprised of 22 and 26 strains. Each cluster group potentially represented a distinct serotype based on PCR-RFLP analysis. The two dominant profiles were selected for complete characterization and designated serotypes O:1 and O:2, respectively (Fig. 2). The remaining 14 strains were represented by nine other groups composed of between one and three strains. No serotype number was assigned to these strains on this occasion.

Sequence analysis of O-antigen gene clusters. The corresponding amplicons representing both serotypes derived from *E. sakazakii* NCTC 11467 (designated serotype O:1) and NCTC 8155 (designated serotype O:2) were sequenced. Nucleotide sequences of 12,188 bp and 12,521 bp, respectively, were obtained. In addition to *galF* and *gnd* (designated *orf1* and *orf13*, respectively), 11 individual ORFs were identified. All were orientated in the same transcriptional direction from *galF* to *gnd* (Fig. 3). As in other O-antigen clusters, all genes had a low G+C content (11, 15, 27, 30) ranging from 50.3% to 28.9% (Table 2). These 11 ORFs were assigned functions based on comparisons with the most significant homologues in the current databases and are summarized in Table 2. Gene names for both serotype-specific amplicons were assigned according to the Bacterial Polysaccharide Gene Nomenclature scheme (http://www.mmb.usyd.edu.au/BPGD/big_paper.pdf) (31).

O:1 O-antigen gene cluster. Genes *rmlA* and *rmlB* (*orf2* and *orf3*) (Table 2; Fig. 3a) encode a glucose-1-phosphate thymidyl-transferase and a dTDP-D-glucose 4,6-dehydratase, respectively, catalyzing the conversion of glucose-1-phosphate to a dTDP-6-deoxy-D-xylo-4-hexulose intermediate that represents a branch point for the downstream biosynthetic pathways (33). The corresponding proteins encoded by these ORFs showed 87 and 81% identity to other known RmlB and RmlA homologues, respectively. The biosynthetic pathway and their corresponding functions have been well characterized for many gram-negative and -positive bacteria (17, 18, 28).

rml-encoding genes function in the biosynthesis of nucleotide-activated dTDP-L-rhamnose and are usually located toward

FIG. 2. Dendrogram generated using BioNumerics software program at a similarity of 80% of MboII restriction digests. Dotted lines indicate uncertain bands. Proposed new species names are included and are marked by asterisks (20). The black and gray boxes in this figure indicate profiles of strains corresponding to serotypes O:1 and O:2, respectively, in the dendrogram (see Fig. 1).

a) O:1 antigen locus

b) O:2 antigen locus

FIG. 3. Genetic organization of the NCTC 11467 (O:1) (a) or NCTC 8155 (O:2) (b) O-antigen gene clusters. Arrows labeled with the corresponding gene names represent individual ORFs and the direction of transcription.

the 5' end of the O-antigen gene cluster in the order *rmlBDAC* (45). However, the four genes are not always clustered together (38, 41), as with *E. sakazakii* serotype O:1 in this study. BLASTp analysis of Orf4, consisting of 133 amino acids, showed homology to WbtA, a putative enzyme previously iden-

tified in *Escherichia coli* (Table 2). The N-terminal part of this ORF protein encoded a transacetylase activity, while the C terminus was predicted to encode an isomerase function and was designated FdtA. The latter is probably responsible for the conversion of dTDP-6-deoxy-D-xyllohex-4-ulose into dTDP-6-

TABLE 2. Putative genes in *E. sakazakii* O:1 or O:2 O-antigen gene cluster

ORF	Location in sequence (bp)	G+C content (%)	Size (aa ^a)	Similar protein	Source (accession no.)	% Identical/% similar (no. of aa)	Putative function
O:1 cluster							
<i>orf1</i>	1–891	53.4	296	GalF	<i>Salmonella enterica</i> (AE017220)	87/95 (295)	UDP-glucose pyrophosphorylase
<i>orf2</i>	1282–2361	50.3	282	RfbB	<i>Shigella flexnerii</i> (NP_707936)	87/93 (273)	dTDP-D-glucose 4,6 dehydratase
<i>orf3</i>	2358–3230	40.9	290	RmlA	<i>Escherichia coli</i> (AAT77170)	81/90 (287)	dTDP-glucose pyrophosphorylase
<i>orf4</i>	3232–3633	35.8	133	WbtA	<i>Escherichia coli</i> (ABK27315)	55/73 (131)	Putative isomerase
<i>orf5</i>	3626–4078	31.8	150	FdtC	<i>Escherichia coli</i> (AAT77171)	49/58 (147)	Acetyltransferase
<i>orf6</i>	4079–5014	34.4	311		<i>Xanthomonas axonopodis</i> pv. Citri (NP_642021)	45/62 (306)	Hypothetical protein
<i>orf7</i>	5011–6114	36.7	367	FdtB	<i>Escherichia coli</i> (AAT77174)	58/73 (365)	Aminotransferase
<i>orf8</i>	6111–7373	37.3	420	Wzx	<i>Escherichia coli</i> (AAT77175)	46/67 (413)	O-antigen flippase
<i>orf9</i>	7373–8317	32.5	314	WbtE	<i>Escherichia coli</i> (AAS73170)	28/50 (254)	Putative glycosyltransferase
<i>orf10</i>	8356–9636	31.9	426	Wzy	<i>Escherichia coli</i> (AAZ85713)	19/43 (272)	O-antigen polymerase
<i>orf11</i>	9623–10501	36.7	292	WbgQ	<i>Shigella boydii</i> (AAL27350)	32/51 (281)	Putative glycosyltransferase
<i>orf12</i>	10491–11255	34.0	254		<i>Shewanella baltica</i> (ZP_01435692)	55/68 (243)	Glycosyltransferase
<i>orf13</i>	11469–12875	54.3	468	Gnd	<i>Escherichia coli</i> (ZP_00926655)	93/97 (469)	6-Phosphogluconate dehydrogenase
O:2 cluster							
<i>orf1</i>	1–594	55.1	197	GalF	<i>Escherichia coli</i> (AAS99158)	83/90 (197)	UDP-glucose pyrophosphorylase
<i>orf2</i>	986–2068	50.3	360	RmlB	<i>Escherichia coli</i> (ABI98981)	84/91 (360)	dTDP-glucose 4,6-dehydratase
<i>orf3</i>	2065–2967	49.5	300	RmlD	<i>Shigella boydii</i> (AAY23334)	80/88 (298)	dTDP-4-dehydrorhamnose reductase
<i>orf4</i>	3017–3895	46.6	292	RmlA	<i>Salmonella enterica</i> (P55254)	91/96 (292)	dTDP-glucose synthase
<i>orf5</i>	3899–4441	35.9	180	RfbC	<i>Shigella boydii</i> (YP_407363)	81/90 (180)	dTDP-4-dehydrorhamnose 3,5-epimerase
<i>orf6</i>	4488–5744	30.2	418	RfbX	<i>Escherichia coli</i> (NP_416541)	27/52 (406)	Putative O-antigen transporter
<i>orf7</i>	5741–6718	33.2	325		<i>Escherichia coli</i> (ZP_01699967)	22/42 (280)	Glycosyltransferase
<i>orf8</i>	6730–7485	29.4	251	WbsV	<i>Shigella boydii</i> (AAS98030)	30/48 (241)	Glycosyltransferase
<i>orf9</i>	7475–8473	31.0	332	Wzy	<i>Shigella boydii</i> (AAS98031)	27/43 (317)	O-antigen polymerase
<i>orf10</i>	8493–9332	28.9	279	WbsW	<i>Shigella boydii</i> (AAS98032)	30/47 (268)	Glycosyltransferase
<i>orf11</i>	9334–10341	31.9	335	WbgM	<i>Escherichia coli</i> (YP_541302)	30/46 (110)	Putative galactosyltransferase
<i>orf12</i>	10331–11095	29.9	254	WbbL	<i>Escherichia coli</i> (P36667)	40/58 (257)	Putative lipopolysaccharide biosynthesis glycosyltransferase
<i>orf13</i>	11462–12868	54.4	468	Gnd	<i>Escherichia coli</i> (ZP_00926655)	93/97 (469)	6-Phosphogluconate dehydrogenase

^a aa, amino acids.

deoxy-D-xylohex-3-ulose and showed good homologies with the putative FdtA isomerase from a number of organisms (*Photobacterium luminescens* subsp. *Laumondii* [NP_931970] and *E. coli* [ABK27315]). This gene product is thought to be essential for the biosynthesis of dTDP-3-acetamido-3,6-dideoxy- α -D-galactose (dTDP-D-Fucp3NAc), known alternatively as 3-acetamidofucose.

orf5 was determined to encode an acetyltransferase activity and is an important protein family involved in O-antigen biosynthesis (36). This ORF protein exhibited 49% identity to an acetyltransferase found in *E. coli* that is involved in the metabolism of nucleotide-activated sugar precursors. BLAST searches revealed highest homologies to genes similar to the FdtC acetyltransferase (Table 2).

Orf7 was identified as a putative aminotransferase belonging to the DegT/DnrJ/EryC1/StrS aminotransferase family. This ORF protein displayed 58% identity to FdtB, an aminotransferase found in *E. coli*. Members of this family are probably all pyridoxal-phosphate-dependent enzymes with a variety of molecular functions, including regulation of cell wall biogenesis. BLAST searches revealed homologies to genes similar to the FdtB aminotransferase. In this study, *orf7* was designated *fdtB* (Table 2).

Based on previous characterization of the enzymes of the biosynthetic pathway of D-Fucp3NAc in *Aneurinibacillus thermoaerophilus*, *orf4*, *orf5*, and *orf7* were designated *fdtA*, *fdtC*, and *fdtB*, respectively (28). dTDP-activated D-Fucp3NAc, the end product, acts as a precursor for the assembly of structural polysaccharides in gram-positive and -negative organisms.

The protein encoded by *orf8* was determined to be a hydrophobic protein, with 12 predicted transmembrane spanning segments, characteristic of the common O-antigen flippase protein (27). Members of this family are proteins involved in the export of O antigen and teichoic acid. This gene product shared 46% identity to several Wzx proteins in *Escherichia coli*. When this ORF was analyzed using the Blockmaker program, 10 conserved motifs were revealed (consisting of 31, 25, 24, 42, 37, 41, 16, 45, 40, and 50 residues). The consensus sequences of these motifs were searched using the PSI-BLAST program against the GenPept database, and other distantly related Wzx proteins were retrieved following three iterations. *orf8* was therefore proposed to be the O-antigen transporter gene, *wzx*.

The protein encoded by *orf10* is another hydrophobic protein, with 11 predicted transmembrane segments, a characteristic feature of O-antigen polymerase genes (33). The *wzy* gene also shared 43% similarity to Wzy proteins of *E. coli* and *Salmonella enterica*. When this ORF protein was analyzed using the Blockmaker program, four conserved motifs were revealed (with 37, 33, 32, and 23 amino acids, respectively). Searching PSI-BLAST using these motifs identified other distantly related Wzy proteins after one iteration.

Searches of the current databases using *orf6* did not permit precise function-related assignment of gene names on this occasion. Only one corresponding match was identified, against a hypothetical protein in *Xanthomonas axonopodis*. It was not possible to confidently assign a function to this ORF in *E. sakazakii*. Therefore, it was designated *wehA*, encoding a hypothetical protein (Table 2).

orf9, *orf11*, and *orf12* are members of the glycosyltransferase group 2 family, which transfer a wide range of sugar groups

(http://www.cazy.org/fam/acc_GT.html). Orf9 exhibited 28% identity to *E. coli* WbtE, while Orf11 exhibited 19% identity to WbgQ, a putative glycosyltransferase of *Shigella boydii*, while Orf12 showed 55% identity to a glycosyltransferase of *Shewanella baltica*. Therefore, *orf9*, *orf11*, and *orf12* were proposed to encode putative glycosyltransferases and were designated *wehB*, *wehC*, and *wehD*.

All of the O:1 antigen genes are tightly linked, and some overlapping reading frames occur. The largest of these were identified between *orf10* and *orf11* (consisting of 14 nucleotides, wherein a start-stop overlap was identified) and *orf11* and *orf12* (11-nucleotide overlap), which would indicate functional interaction between transcriptional units.

The O-antigen gene cluster of *Enterobacter sakazakii* NCTC 11467 was 99.1% similar at the nucleotide sequence level to the same region from ATCC BAA-894 (data not shown).

O:2 O-antigen gene cluster. A complete L-rhamnose biosynthetic *rml* operon was identified in *E. sakazakii* NCTC 8155 comprising *orf2* through *orf5* (Fig. 3b; Table 2). This operon is widely distributed in many gram-negative bacteria, and the four-step biosynthetic pathway produces dTDP-L-rhamnose from glucose 1-phosphate, the initial substrate (33). In this study the gene order *rmlBDAC* was assigned based on comparisons with other similar sequences and the high level of identity (ranging between 80 and 91%) with other *Enterobacteriaceae* (Table 2).

Analysis of the *orf6* coding sequence identified a gene product with 12 predicted transmembrane segments, similar to O-antigen transporters (14, 27). This ORF protein displayed 27% identity to RfbX proteins of *E. coli*. When Orf6 was analyzed using the Blockmaker program, seven conserved motifs were revealed (of 30, 44, 12, 34, 24, 19, and 37 amino acids, respectively). These consensus sequences were analyzed using PSI-BLAST. Other distantly related Wzx proteins were retrieved after three iterations. *orf6* was therefore proposed to be an O-antigen transporter gene and was named *wzx* accordingly.

Orf7, -8, and -10 were similar to members of the group 1 glycosyltransferases (Table 2b). In the former, 22% identity was observed compared to a functionally similar protein in *E. coli*. In the case of Orf8 and Orf10, both exhibited 30% identity to the WbsV and WbsW glycosyltransferases found in *Shigella boydii*. All three putative glycosyltransferase genes were designated *wehE*, *wehF*, and *wehG*, respectively (Fig. 3b).

The protein encoded by *orf9* had 11 predicted transmembrane segments, typical of O-antigen polymerases (10, 45). When compared to similar Wzy proteins in *Shigella boydii*, Orf9 displayed 27% similarity. Five conserved motifs were revealed with Blockmaker (of 14, 10, 38, 13, and 22 residues). These consensus sequences were similar to those of other O-antigen polymerase genes in other organisms, including *Shigella boydii* (AAS98031) and *Salmonella enterica* (YP_217085). Based on these findings, *orf9* was designated *wzy*.

Orf11 also belongs to the group 1 family of glycosyltransferases. This ORF protein also exhibited 30% identity to WbgM, a putative galactosyltransferase responsible for catalyzing the formation of the α (1-3)-galactosyl linkage to GalNAc. *orf11* encodes a putative glycosyltransferase and is named *wehH*.

Orf12 is a member of the group 2 family of glycosyltransferases with 40% identity to the putative glycosyltransferase WbbL of *E. coli*. The ORF protein also displayed 42% identity

FIG. 4. A 1.5% agarose gel image of specific-serotype O:1 (top panel) or serotype O:2 (bottom panel) amplicons, generated by PCR. Lane 1, CFS131; lane 2, E824; lane 3, CFS06; lane 4, 44; lane 5, 93; lane 6, E891; lane 7, E784; lane 8, CFS153; lane 9, 88; lane 10, 305N; lane 11, ATCC BAA 893; lane 12, ATCC BAA 894; lane 13, E899; lane 14, 343; lane 15, 80; lane 16, E775; lane 17, 71; lane 18, E846; lane 19, E830; lane 20, 228N; lane 21, CFS136; lane 22, CFS149; lane 23, no-template control; lane M, 100-bp DNA ladder (New England Biolabs, Hertfordshire, England). Dashed-line box indicates serotype O:1 strains, while continuous-line box indicates serotype O:2 strains.

to WcvF of *Vibrio vulnificus*, which is a characterized rhamnosyltransferase for sugar linkage that has been shown to be essential for capsule expression (37). This ORF is proposed to encode a glycosyltransferase and is designated *wehI*.

As was evident in the corresponding O:1 antigen cluster, some of the genes in the O:2 locus are also linked through overlapping ORFs. The largest of these were between *orf8* and *orf9* (11 nucleotides) and *orf11* and *orf12* (11 nucleotides).

PCR assay to detect *E. sakazakii* O:1 and O:2 serotypes. Two glycosyltransferase genes, *wehC* (in serotype O:1) and *wehI* (in serotype O:2), located toward the 3' end of each cluster, were selected and compared against the current database entries. In silico analysis showed that both loci were unique and may be specific for O:1 and O:2 serotypes of *E. sakazakii*. To test this hypothesis, primer pairs were designed based on these genes and used to amplify internal regions of each ORF. The 62 study isolates, including both serotypes, together with 25 non-*E. sakazakii* members of the *Enterobacteriaceae*, were selected to determine the utility of these gene targets to reliably detect the two serotypes of *E. sakazakii*.

Amplicons were obtained from *E. sakazakii* strains corresponding to the two serotypes only. PCR products of 341 and 329 bp (Fig. 4), corresponding to serotypes O:1 and O:2 and represented by the strains NCTC11467 (O:1) and NCTC8155 (O:2) (Table 1; Fig. 2), were obtained on each occasion. No PCR products were amplified from any of the non-O:1, non-O:2 *E. sakazakii* strains, nor from other members of the *Enterobacteriaceae* (data not shown).

When these PCRs were tested using mixtures of O:1 and O:2 DNA in combination with DNA purified from all of the *E. sakazakii* strains and other members of *Enterobacteriaceae*, in this study (Table 1) only the specific target amplicons for O:1 and O:2 were observed each time (data not shown). These data suggest that the assay is capable of detecting both serotypes.

Analysis of *E. sakazakii* LPS. The LPS profiles of 12 selected strains were determined. Proteinase K-prepared LPS followed by SDS-PAGE and visualization by silver staining was

used to analyze the LPS profile in each case. Various numbers of bands and intensities were obtained (Fig. 5a and b). An intensely stained low-molecular-mass band that migrated with a mobility similar to that of LPS from *Salmonella enterica* serotype Minnesota was common to all *E. sakazakii* strains and was predicted to be the lipid A core (located toward the end of each gel). O-antigen sugar repeats representing smooth LPS were visible for several strains. Some of these strains (e.g., NCTC 8155 in Fig. 5b, lane 3) showed a typical smooth (S) LPS profile with separation of bands of high molecular weight, consistent with an LPS structure containing an O-chain polysaccharide composed of a single, uniform, repeating glycoside residue.

All strains belonging to serotype O:2 exhibited a similar O-antigen chain length repeat distribution, with the exception of ATCC 12868 (Fig. 5b, lane 2) and E892 (an R-form strain). The former displayed bands that migrated faster and some that stained less intensely, while E892 appeared to lack O-antigen repeats and was identified accordingly as having an R-form LPS (Fig. 5b, lanes 2 and 5, respectively).

Consistencies in S-form and O-antigen chain length repeat distribution were observed between serotype O:1 strains ATCC BAA-894 and 90, which displayed identical S-form LPS (Fig. 5a, lanes 2 and 6). In contrast, CFS1001 showed an R-form LPS (Fig. 5a, lane 4).

O-antigen sugar repeat similarities were identified for each serotype, such as strains CFS10 (O:2) and 90 (O:1), which, although of different intensities, shared three bands ranging from 26 to 31 kDa (Fig. 5a, lanes 5 and 6). This suggested the conservation of polymer units between the two serogroups.

DISCUSSION

O-antigen gene clusters characterized previously for *E. coli*, *Salmonella*, and other members of the *Enterobacteriaceae* are located between the *galF* and *gnd* genes (33). Based on data presented in this article, the genetic organization of the corresponding locus in the opportunistic neonatal pathogen *E. sakazakii* was similar to that found in related members of the *Enterobacteriaceae*. The O-antigen-encoding locus of 62 *E. sakazakii* strains was analyzed initially by PCR-RFLP. Based on the DNA restriction patterns obtained, two major groups within the species *sakazakii* were identified when the banding profiles were compared (Fig. 2). These were designated serotypes O:1 and O:2, respectively, and are represented by the strains *E. sakazakii* NCTC 11467 (or ATCC 29544) and NCTC 8155 (or CDC 9081-75). The corresponding *rfb*-encoding regions were completely characterized. Both loci are the first of the *E. sakazakii* serotypes to be defined. An additional 9 potential serotypes, represented by 14 strains, were identified, and these were distinct from previously characterized serotypes O:1 and O:2 based on their corresponding PCR-RFLP profiles (Fig. 2).

Analysis of the gene content and sequence similarities revealed two distinct clusters containing 11 unique ORFs. Typically all of these genes had a low G+C content (<40%), a feature found previously in other O-antigen gene clusters (11, 15, 27, 30). Based on the currently available data from the draft *E. sakazakii* genome sequence (http://genome.wustl.edu/pub/organism/Microbes/Enteric_Bacteria/Enterobacter_sakazakii

FIG. 5. Silver-stained SDS-PAGE of purified LPS samples. The relative positions of the O-antigen side chain and lipid A core are indicated by the brackets. (a) Lane M, rainbow molecular weight marker (high range; RPN 800) (Amersham Life Science, Buckinghamshire, England); lane 1, *S. enterica* serotype Minnesota LPS (5 μ g); lane 2, ATCC BAA-894 (O:1); lane 3, E901 (O:2); lane 4, CFS1001 (O:1); lane 5, CFS10 (O:2); lane 6, 90; lane 7, CFS149 (O:1); lane 8, CFS173 (neither O:1 nor O:2); lane 9, undigested E901. Asterisks indicate O-antigen bands of similar mobilities. (b) Lane M, molecular marker, lane 1, *S. enterica* serotype Minnesota LPS (5 μ g); lane 2, ATCC 12868 (O:2); lane 3, NCTC 8155 (O:2); lane 4, E844 (O:2); lane 5, E892 (O:2); lane 6, 228N (O:2); lane 7, undigested ATCC 12868. Asterisks indicate O-antigen bands of different mobilities in digested NCTC 12868.

/assembly/Enterobacter sakazakii-4.0/), this divergence in G+C content at both loci suggested that O:1 and O:2 antigen DNA may have originated from a species other than *E. sakazakii*. The locus was probably acquired relatively recently by lateral gene transfer (LGT). These genetic loci are thought to be subject to strong frequency-dependent selection linked to the avoidance of host defense systems and survival in the environment (29, 46). Previous studies have shown recombination between O-antigen gene clusters of *E. coli* (5), between those of *E. coli* and *Klebsiella* (39), and between those of *E. coli* and *Shigella* (13). Identical O-antigen gene clusters among different species have also been reported, and these are likely to originate from a common ancestor (33, 46). In contrast, recent reports suggested that LGT is insertion sequence mediated (9). In relation to *E. sakazakii*, it is not possible at this time to conclude whether or not this locus was acquired through LGT. Interestingly, our results show that strains CFS129 and E839 group into the same serotype (Table 1; Fig. 2). These strains are now reclassified as "*Cronobacter malonaticus*" and "*Cronobacter sakazakii*," respectively, and these data suggest that the same serotype may be represented by more than one species.

Of the 62 *E. sakazakii* isolates studied, 77% (representing 48 isolates) were subdivided into one of the two O-antigen locus types based on the MboII restriction digest patterns obtained. Furthermore, it was not possible to attribute any relationship to a given serotype and its corresponding biotype (data not shown). *Enterobacter sakazakii* is a taxonomically diverse group of organisms represented by at least five DNA:DNA hybridization groups (20). According to this recently proposed classification scheme, strains belonging to serotypes O:1 and O:2 all belong to the same DNA:DNA hybridization group, proposed as "*Cronobacter sakazakii* comb. nov.," while none of the other "*Cronobacter*" species belong to either serogroup (Fig. 2). Moreover, DNA sequence comparisons using the Artemis tool (32) confirmed the distinct genetic composition associated with serotypes O:1 and O:2 (data not shown). Each O-antigen locus contained genes specific for sugar synthesis, located toward the proximal end of the locus, along with genes encoding the O-unit flippase (*wzx*) and O-antigen polymerase (*wzy*).

Closer analysis of the gene content of both O-antigen clusters identified genes responsible for dTDP-sugar biosynthetic pathways. L-rhamnose and 3-acetamidofucose, both 6-deoxyhexose sugars, are synthesized from glucose-1-phosphate as dTDP sugars in NCTC11467 (O:1) and NCTC8155 (O:2), respectively. Similar genes were previously found in many bacteria (33) and in *Aneurinibacillus thermoaerophilus* for 3-acetamidofucose (28).

O-antigen processing genes and related glycosyltransferases often display low levels of primary structure similarity (33). This is particularly evident for glycosyltransferases. Indeed, the precise function of only a limited number of putative glycosyltransferases has been elucidated. Based on our sequence analysis, we proposed functions for the three glycosyltransferases from O:1 and the five glycosyltransferases from O:2. The O-antigen processing genes, *wzx* and *wzy*, encode typical inner membrane proteins, with *Wzx* containing 12 transmembrane segments and *Wzy* proteins containing 11 transmembrane segments (27, 45). These characteristics, along with the identi-

cation of conserved motifs, support their identification in *E. sakazakii* serotypes O:1 and O:2.

A number of acetyltransferases were also identified in the O-antigen region of NCTC 11467 which are associated with the modification of the O-subunit structure. In many cases acylation defines the serotype, and this has been shown to have an effect on the host-microbe interaction for bacteria (22, 36).

orf10 from NCTC 11467 was not homologous to any nucleotide or amino acid sequence when subjected to BLAST analysis. Therefore, no function could be assigned to this ORF based on sequence homology at this time.

This study also identified two specific *E. sakazakii* O:1 and O:2 O-antigen regions, which form the basis for the development of molecular detection assays for the identification of both serotypes. Previous studies have reported the successful use of genes within the O-antigen gene cluster as molecular markers for the identification of specific bacterial serotypes (3, 14, 15). In our analysis, we carried out extensive BLAST searches of the current databases and determined that *wehC* and *wehI* were unique to *E. sakazakii* and that both loci were also highly polymorphic. On this basis, these genes were selected as serotype-specific markers for detection of *E. sakazakii* serotypes O:1 and O:2, respectively. The PCR assays reported here may represent a convenient way to monitor the occurrence of the *E. sakazakii* O:1 and O:2 serotypes among clinical, environmental, and food isolates.

Bacteria can on occasion switch between smooth (S) and rough (R) forms. These forms arise following mutations in one or more genes controlling synthesis and polymerization (7). Amplification and subsequent RFLP analysis provided no clues as to whether an isolate exhibited an S or R form. As an example, the clinical strain E892 produced a DNA amplicon of similar size along with a MboII fragment pattern identical to those of other *E. sakazakii* O:2 serotypes; however, this strain had an R-form LPS, as determined by SDS-PAGE (Fig. 5b, lane 5). Interestingly, an ORF located 228 bp downstream of *gnd* in ATCC BAA-894 was identified which shared 83% similarity to *wzzB*, the O-antigen chain length-determining gene of *Shigella flexneri* (data not shown). Franco et al. (16) showed that O-antigen chain heterogeneity in *E. coli* was a result of amino acid substitutions in the *Wzz* protein. This could explain differences in chain length seen in *E. sakazakii* and/or that some LPS profiles appeared rough due to bands staining at a low intensity.

The loss of O-antigen subunits corresponded to the appearance of R-form LPS and increased cell surface hydrophobicity. Additionally, the growth temperature can influence the organization of LPS, and this could explain the appearance of R-form LPS in some strains. al-Hendy et al. (1) showed a decrease in O-antigen expression in *Yersinia enterocolitica* at 37°C compared to that in strains cultured at 25°C, while no differences were shown in an *E. coli* clone grown at these temperatures. These observations could suggest the presence of a thermally induced regulatory gene located outside the O-antigen locus in the former, while a similar gene was absent in *E. coli*. The presence of such a locus controlling the expression of the O-antigen region could explain the appearance of different forms of O antigen within the same *E. sakazakii* serotype. The most successful of bacteria are those that can modulate their surface according to their environment. *Entero-*

bacter sakazakii is an environmental bacterium (21) whose phenotypic plasticity would serve to aid adaptation and competitive fitness.

In conclusion, we have described the first characterization of two *rfb*-encoding loci defining two serotypes, O:1 and O:2, in *E. sakazakii*. These data should serve to further our understanding of this emerging neonatal pathogen. Serotype-specific genes were identified and were used to develop a PCR assay. Based on these primer pairs, these gene targets could be used to identify *E. sakazakii* serotype O:1 and O:2 strains from environmental, food, and clinical sources, contributing to the protection of neonates.

ACKNOWLEDGMENTS

We thank the Bacterial Polysaccharide Gene Database and in particular Peter Reeves and Monica Cunneen for their advice on nomenclature and allocation of new gene names. We thank Avril Monahan for technical assistance. We thank Denise Drudy and Maria Karczmarczyk, who provided helpful comments on the manuscript.

We acknowledge the financial support provided through the Irish government's Food Institutional Research Measure (FIRM), grant no. 05/R&D/D/363.

REFERENCES

1. al-Hendy, A., P. Toivanen, and M. Skurnik. 1991. The effect of growth temperature on the biosynthesis of *Yersinia enterocolitica* O:3 lipopolysaccharide: temperature regulates the transcription of the *rfb* but not of the *rfa* region. *Microb. Pathog.* 10:81–86.
2. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
3. Bogdanovich, T., E. Carniel, H. Fukushima, and M. Skurnik. 2003. Use of O-antigen gene cluster-specific PCRs for the identification and O-genotyping of *Yersinia pseudotuberculosis* and *Yersinia pestis*. *J. Clin. Microbiol.* 41:5103–5112.
4. Bowen, A. B., and C. R. Braden. 2006. Invasive *Enterobacter sakazakii* disease in infants. *Emerg. Infect. Dis.* 12:1185–1189.
5. Cheng, J., Q. Wang, W. Wang, Y. Wang, L. Wang, and L. Feng. 2006. Characterization of *E. coli* O24 and O56 O antigen gene clusters reveals a complex evolutionary history of the O24 gene cluster. *Curr. Microbiol.* 53:470–476.
6. Coimbra, R. S., F. Grimont, and P. A. Grimont. 1999. Identification of *Shigella* serotypes by restriction of amplified O-antigen gene cluster. *Res. Microbiol.* 150:543–553.
7. Coimbra, R. S., F. Grimont, P. Lenormand, P. Burguiere, L. Beutin, and P. A. Grimont. 2000. Identification of *Escherichia coli* O-serogroups by restriction of the amplified O-antigen gene cluster (*rfb*-RFLP). *Res. Microbiol.* 151:639–654.
8. Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.* 10:673–676.
9. Cunneen, M. M., and P. R. Reeves. 2007. The *Yersinia kristensenii* O11 O-antigen gene cluster was acquired by lateral gene transfer and incorporated at a novel chromosomal locus. *Mol. Biol. Evol.* 24:1355–1365.
10. 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.
11. DebRoy, C., P. M. Fratamico, E. Roberts, M. A. Davis, and Y. Liu. 2005. Development of PCR assays targeting genes in O-antigen gene clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. *Appl. Environ. Microbiol.* 71:4919–4924.
12. Drudy, D., N. R. Mullane, T. Quinn, P. G. Wall, and S. Fanning. 2006. *Enterobacter sakazakii*: an emerging pathogen in powdered infant formula. *Clin. Infect. Dis.* 42:996–1002.
13. Feng, L., A. V. Perepelov, G. Zhao, S. D. Shevelev, Q. Wang, S. N. Senchenkova, A. S. Shashkov, Y. Geng, P. R. Reeves, Y. A. Knirel, and L. Wang. 2007. Structural and genetic evidence that the *Escherichia coli* O148 O antigen is the precursor of the *Shigella dysenteriae* type 1 O antigen and identification of a glucosyltransferase gene. *Microbiology* 153:139–147.
14. Feng, L., S. N. Senchenkova, J. Tao, A. S. Shashkov, B. Liu, S. D. Shevelev, P. R. Reeves, J. Xu, Y. A. Knirel, and L. Wang. 2005. Structural and genetic characterization of enterohemorrhagic *Escherichia coli* O145 O antigen and development of an O145 serogroup-specific PCR assay. *J. Bacteriol.* 187:758–764.
15. Fitzgerald, C., L. Gheesling, M. Collins, and P. I. Fields. 2006. Sequence analysis of the *rfb* loci, encoding proteins involved in the biosynthesis of the *Salmonella enterica* O17 and O18 antigens: serogroup-specific identification by PCR. *Appl. Environ. Microbiol.* 72:7949–7953.
16. Franco, A. V., D. Liu, and P. R. Reeves. 1996. A Wzz (Cld) protein determines the chain length of K lipopolysaccharide in *Escherichia coli* O8 and O9 strains. *J. Bacteriol.* 178:1903–1907.
17. Giraud, M. F., F. M. Gordon, C. Whitfield, P. Messner, S. A. McMahon, and J. H. Naismith. 1999. Purification, crystallization and preliminary structural studies of dTDP-6-deoxy-D-xylo-4-hexulose 3,5-epimerase (RmlC), the third enzyme of the dTDP-L-rhamnose synthesis pathway, from *Salmonella enterica* serovar Typhimurium. *Acta Crystallogr.* 55:706–708.
18. Giraud, M. F., H. J. McMiken, G. A. Leonard, P. Messner, C. Whitfield, and J. H. Naismith. 1999. Overexpression, purification, crystallization and preliminary structural study of dTDP-6-deoxy-L-xylo-4-hexulose reductase (RmlD), the fourth enzyme of the dTDP-L-rhamnose synthesis pathway, from *Salmonella enterica* serovar Typhimurium. *Acta Crystallogr.* 55:2043–2046.
19. Henikoff, S., J. G. Henikoff, W. J. Alford, and S. Pietrovski. 1995. Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* 163:GC17–GC26.
20. Iversen, C., A. Lehner, N. Mullane, E. Bidlas, I. Cleenwerck, J. Marugg, S. Fanning, R. Stephan, and H. Joosten. 2007. The taxonomy of *Enterobacter sakazakii*: proposal of a new genus *Cronobacter* gen. nov. and descriptions of *Cronobacter sakazakii* comb. nov., *Cronobacter sakazakii* subsp. *sakazakii*, comb. nov., *Cronobacter sakazakii* subsp. *malonaticus* subsp. nov., *Cronobacter turicensis* sp. nov., *Cronobacter mytjensii* sp. nov., *Cronobacter dublinensis* sp. nov. and *Cronobacter* genomospecies 1. *BMC Evol. Biol.* 7:64.
21. Kandhai, M. C., M. W. Reij, K. van Puyvelde, O. Guillaume-Gentil, R. R. Beumer, and M. van Schothorst. 2004. A new protocol for the detection of *Enterobacter sakazakii* applied to environmental samples. *J. Food Prot.* 67:1267–1270.
22. Lerouge, I., C. Verreth, J. Michiels, R. W. Carlson, A. Datta, M. Y. Gao, and J. Vanderleyden. 2003. Three genes encoding for putative methyl- and acetyl-transferases map adjacent to the *wzm* and *wzt* genes and are essential for O-antigen biosynthesis in *Rhizobium elii* CE3. *Mol. Plant-Microbe Interact.* 16:1085–1093.
23. Marolda, C. L., P. Lahiry, E. Vines, S. Saldias, and M. A. Valvano. 2006. Micromethods for the characterization of lipid A-core and O-antigen lipopolysaccharide. *Methods Mol. Biol.* 347:237–252.
24. Mullane, N. R., R. Drudy, P. Whyte, M. O'Mahony, A. G. M. Scannell, P. G. Wall, and S. Fanning. 2006. *Enterobacter sakazakii*: biological properties and significance in dried infant milk formula (IMF) powder. *Int. J. Dairy Technol.* 59:102–111.
25. Mullane, N. R., C. Iversen, B. Healy, C. Walsh, P. Whyte, P. G. Wall, T. Quinn, and S. Fanning. 2007. *Enterobacter sakazakii* an emerging bacterial pathogen with implications for infant health. *Minerva Pediatr.* 59:137–148.
26. Pantophlet, R., L. Brade, L. Dijkshoorn, and H. Brade. 1998. Specificity of rabbit antisera against lipopolysaccharide of *Acinetobacter*. *J. Clin. Microbiol.* 36:1245–1250.
27. Perelle, S., F. Dilasser, J. Grout, and P. Fach. 2002. Identification of the O-antigen biosynthesis genes of *Escherichia coli* O91 and development of a O91 PCR serotyping test. *J. Appl. Microbiol.* 93:758–764.
28. Pfostl, A., A. Hofinger, P. Kosma, and P. Messner. 2003. Biosynthesis of dTDP-3-acetamido-3,6-dideoxy-alpha-D-galactose in *Aneurinibacillus thermoaerophilus* L420-91T. *J. Biol. Chem.* 278:26410–26417.
29. Reeves, P. 1993. Evolution of *Salmonella* O antigen variation by interspecific gene transfer on a large scale. *Trends Genet.* 9:17–22.
30. Reeves, P. P., and L. Wang. 2002. Genomic organization of LPS-specific loci. *Curr. Top. Microbiol. Immunol.* 264:109–135.
31. Reeves, P. R., M. Hobbs, M. A. Valvano, M. Skurnik, C. Whitfield, D. Coplin, N. Kido, J. Klena, D. Maskell, C. R. Raetz, and P. D. Rick. 1996. Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol.* 4:495–503.
32. Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945.
33. Samuel, G., and P. Reeves. 2003. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr. Res.* 338:2503–2519.
34. Schnaitman, C. A., and J. D. Klena. 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol. Rev.* 57:655–682.
35. Seo, K. H., and R. E. Brackett. 2005. Rapid, specific detection of *Enterobacter sakazakii* in infant formula using a real-time PCR assay. *J. Food Prot.* 68:59–63.
36. Slauch, J. M., A. A. Lee, M. J. Mahan, and J. J. Mekalanos. 1996. Molecular characterization of the *oafA* locus responsible for acetylation of *Salmonella typhimurium* O-antigen: OafA is a member of a family of integral membrane trans-acylases. *J. Bacteriol.* 178:5904–5909.
37. Smith, A. B., and R. J. Siebeling. 2003. Identification of genetic loci required for capsular expression in *Vibrio vulnificus*. *Infect. Immun.* 71:1091–1097.
38. Stenutz, R., A. Weintraub, and G. Widmalm. 2006. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol. Rev.* 30:382–403.
39. Sugiyama, T., N. Kido, Y. Kato, N. Koide, T. Yoshida, and T. Yokochi. 1997. Evolutionary relationship among *rfb* gene clusters synthesizing mannose

- homopolymer as O-specific polysaccharides in *Escherichia coli* and *Klebsiella*. *Gene* **198**:111–113.
40. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
41. **Trefzer, A., J. A. Salas, and A. Bechthold.** 1999. Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat. Prod. Rep.* **16**:283–299.
42. **Trent, M. S., C. M. Stead, A. X. Tran, and J. V. Hankins.** 2006. Diversity of endotoxin and its impact on pathogenesis. *J. Endotoxin Res.* **12**:205–223.
43. **Tsai, C. M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal. Biochem.* **119**:115–119.
44. **Verma, N. K., N. B. Quigley, and P. R. Reeves.** 1988. O-antigen variation in *Salmonella* spp.: *rfb* gene clusters of three strains. *J. Bacteriol.* **170**:103–107.
45. **Wang, L., B. Liu, Q. Kong, H. Steinruck, G. Krause, L. Beutin, and L. Feng.** 2005. Molecular markers for detection of pathogenic *Escherichia coli* strains belonging to serogroups O 138 and O 139. *Vet. Microbiol.* **111**:181–190.
46. **Wang, W., A. V. Perepelov, L. Feng, S. D. Shevelev, Q. Wang, S. N. Senchenkova, W. Han, Y. Li, A. S. Shashkov, Y. A. Knirel, P. R. Reeves, and L. Wang.** 2007. A group of *Escherichia coli* and *Salmonella enterica* O antigens sharing a common backbone structure. *Microbiology* **153**:2159–2167.
47. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* **75**:39–68.