



# Defining the Genetic Features of O-Antigen Biosynthesis Gene Cluster and Performance of an O-Antigen Serotyping Scheme for *Escherichia albertii*

Hong Wang<sup>1†</sup>, Han Zheng<sup>2†</sup>, Qun Li<sup>1</sup>, Yanmei Xu<sup>2</sup>, Jianping Wang<sup>2</sup>, Pengcheng Du<sup>3</sup>, Xinqiong Li<sup>1</sup>, Xiang Liu<sup>1</sup>, Ling Zhang<sup>1</sup>, Nianli Zou<sup>1</sup>, Guodong Yan<sup>1</sup>, Zhengdong Zhang<sup>1</sup>, Huaiqi Jing<sup>2</sup>, Jianguo Xu<sup>2</sup> and Yanwen Xiong<sup>2\*</sup>

<sup>1</sup> Zigong Center for Disease Control and Prevention, Zigong, China, <sup>2</sup> State Key Laboratory of Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China, <sup>3</sup> Beijing Key Laboratory of Emerging Infectious Diseases, Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China

## OPEN ACCESS

### Edited by:

Leonard Peruski,  
US Centers for Disease Control and  
Prevention, United States

### Reviewed by:

Kai Zhou,  
Zhejiang University, China  
Bin Liu,  
Nankai University, China  
Lei Wang,  
Nankai University, China

### \*Correspondence:

Yanwen Xiong  
xiongyanwen@icdc.cn

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Infectious Diseases,  
a section of the journal  
Frontiers in Microbiology

Received: 11 April 2017

Accepted: 12 September 2017

Published: 26 September 2017

### Citation:

Wang H, Zheng H, Li Q, Xu Y, Wang J,  
Du P, Li X, Liu X, Zhang L, Zou N,  
Yan G, Zhang Z, Jing H, Xu J and  
Xiong Y (2017) Defining the Genetic  
Features of O-Antigen Biosynthesis  
Gene Cluster and Performance of an  
O-Antigen Serotyping Scheme for  
*Escherichia albertii*.  
*Front. Microbiol.* 8:1857.  
doi: 10.3389/fmicb.2017.01857

*Escherichia albertii* is a newly described and emerging diarrheagenic pathogen responsible for outbreaks of gastroenteritis. Serotyping plays an important role in diagnosis and epidemiological studies for pathogens of public health importance. The diversity of O-antigen biosynthesis gene clusters (O-AGCs) provides the primary basis for serotyping. However, little is known about the distribution and diversity of O-AGCs of *E. albertii* strains. Here, we presented a complete sequence set for the O-AGCs from 52 *E. albertii* strains and identified seven distinct O-AGCs. Six of these were also found in 15 genomes of *E. albertii* strains deposited in the public database. Possession of *wzy/wzx* genes in each O-AGC strongly suggest that O-antigens of *E. albertii* were synthesized by the Wzx/Wzy-dependent pathway. Furthermore, we performed an O-antigen serotyping scheme for *E. albertii* based on specific antisera against seven O-antigens and a high throughput xTAG Luminex assay to simultaneously detect seven O-AGCs. Both methods accurately identified serotypes of 64 tested *E. albertii* strains. Our data revealed the high-level diversity of O-AGCs in *E. albertii*. We also provide valuable methods to reliably identify and serotype this bacterium.

**Keywords:** *Escherichia albertii*, O-AGC, serotype, antiserum, xTAG Luminex

## INTRODUCTION

Lipopolysaccharide (LPS) molecules form the outer leaflet of the outer membrane of many Gram-negative bacteria and are essential components of the bacterial cell envelope. The O-antigen polysaccharide constitutes the exterior part of the LPS and consists of oligosaccharide repeats (O-units) containing three to six sugar residues. The O-antigen plays an important role in resistance to phagocytosis and complement-mediate lytic action (Murray et al., 2003, 2006; Duerr et al., 2009; Saldias et al., 2009). Meanwhile, the O-antigen is a major surface antigen and is responsible for serological diversity of Gram-negative bacteria which are clinically and

epidemiologically important to classify various strains. O-antigen has also provided a basis for development of vaccine against many pathogens. The genes required for O-antigen biosynthesis are clustered at a chromosomal locus, named the O-antigen biosynthesis gene cluster (O-AGC) in many bacteria. Generally, the genes in O-AGC are clustered into three major classes: sugar synthesis genes, glycosyltransferase genes, and O-unit processing genes. Polymerization of the O-units into an O-antigen is mostly mediated through two of three pathways in Gram-negative species: Wzx/Wzy-dependent pathway and ABC transporter-dependent pathway (Valvano, 2003). Synthase-dependent pathway, the third pathway, is rarely present in Gram-negative species. O-AGC is always located between the conserved *galF* (encoding UTP-glucose-1-phosphate uridylyltransferase) and *gnd* (encoding 6-phosphogluconate dehydrogenase) genes in many species of the *Enterobacteriaceae*, such as *E. coli* (Iguchi et al., 2015) and *Cronobacter sakazakii* (Mullane et al., 2008). Two genes, encoding the O antigen flippase (*wzx*) and O antigen polymerase (*wzy*), are unique in most of the O-AGCs, and have been used as targets for molecular O serotyping (DebRoy et al., 2016).

*Escherichia albertii* is a newly described and emerging diarrheagenic pathogen, which is associated with both sporadic infections and outbreaks in humans (Ooka et al., 2013; Asoshima et al., 2014; Murakami et al., 2014; Brandal et al., 2015; Inglis et al., 2015). It was initially identified as *Hafnia alvei* and later proposed as *E. albertii*, a new species within the genus *Escherichia* (Huys et al., 2003). *E. albertii* strains were often misidentified as *E. coli*, *Hafnia*, *Salmonella*, or *Yersinia ruckeri* as the lack of specific biochemical characteristics (Abbott et al., 2003). Thus, the prevalence of *E. albertii* may be underestimated owing to the lack of effective methods to discriminate *E. albertii* from other members of the *Enterobacteriaceae*. To date, little information on the *E. albertii* O-antigen is available. Only several chemical structures of the O-specific polysaccharide (OPS) of *E. albertii* were reported in previous study (Eserstam et al., 2002; Naumenko et al., 2017; Zheng et al., 2017). There is no comprehensive scheme for O-antigen classification of *E. albertii*. This study was aimed to investigate the prevalence and characteristics of O-AGCs in *E. albertii* strains and develop an O-antigen serotyping scheme and a high throughput detection assay to simultaneously detect all types of these O-AGCs.

## MATERIALS AND METHODS

### Bacterial Strains and Genomic DNA Preparation

Fifty-two strains were selected in the current study: type strain LMG20976 (Huys et al., 2003); one strain from the stool of a diarrheal patient resident in Shanghai in 2013; and 50 strains isolated from multiple sources in Zigong city of Sichuan province between 2014 and 2015 (Table 1). Thirty of these were also used in our previous study (Wang H. et al., 2016). An additional 12 strains were isolated from Luzhou city of Sichuan province in 2016 and used in an agglutination test and the

development of the high throughput xTAG Luminex detection assay (Table 1). Strains were cultured on Luria-Bertani (LB) plates (Oxoid, UK) and genomic DNA was extracted using the Wizard Genomic DNA Purification kit (Promega, Madison, MI, USA). Seven housekeeping genes were used for multi-locus sequence typing (MLST) analysis according to the *E. coli* MLST website (<http://mlst.warwick.ac.uk/mlst/dbs/Ecoli>). *E. coli* O3 and O181 antisera were purchased from Statens Serum Institut (SSI, Copenhagen, Denmark) for the agglutination test. All strains were verified to be *E. albertii* based on the combination of 16S rDNA sequencing, diagnostic multiplex PCR, and MLST analysis as described in our previous study (Wang H. et al., 2016).

### Sequencing and Bioinformatics Analysis

Forty-two strains were chosen for whole genome sequencing (WGS) based on the MLST results. For each strain, a library was constructed (500–2,000 bp) and then sequenced on an Illumina HiSeq 4,000 system (Illumina, San Diego, CA, USA) to produce 150 bp paired-end reads, which were then assembled into scaffolds using the program SOAP *de novo* (Release 1.04, <http://soap.genomics.org.cn/soapdenovo.html>). Open reading frames (ORFs) were identified and annotated using the Artemis program ([www.sanger.ac.uk](http://www.sanger.ac.uk)) and homology searches against several databases including GenBank ([www.ncbi.nlm.nih.gov/GenBank](http://www.ncbi.nlm.nih.gov/GenBank)), the Clusters of Orthologous Groups (COG; [www.ncbi.nlm.nih.gov/COG/](http://www.ncbi.nlm.nih.gov/COG/)), and Pfam ([pfam.sanger.ac.uk](http://pfam.sanger.ac.uk)) protein motif databases (Altschul et al., 1997; Tatusov et al., 2001; Bateman et al., 2002). Each O-AGC between the *galF* and *gnd* genes was extracted from the draft genome sequence. Based on these O-AGC sequences, another 10 O-AGCs were sequenced by primer walking PCR. The TMHMM (v2.0) analysis program (<http://www.cbs.dtu.dk/services/TMHMM/>) was used to identify potential transmembrane segments from the amino acid sequences. The Artemis comparison tool (ACT) (Carver et al., 2005) was used to visualize the data.

### Preparation of Specific Antisera

Based on the typing result of strains SP140089, SP150020, SP140724, D140513, T150248, T150072, and ZG141049 were initially used as standard antigen strains to produce antisera. Three New Zealand white rabbits (female, 1.5 to 2 kg body weight) were immunized intravenously with heat-killed (100°C, 2 h) cells four times with the same doses ( $2.5 \times 10^{10}$  CFU) for each strain. The second immunization booster was performed 14 days after the first immunization. The third and fourth immunization boosters were performed 5 days and 10 days after the second immunization, respectively. Serum was obtained 5 days after the last immunization. Prepared serum was used to test all *E. albertii* strains in this study by slide agglutination and the strain was heat-killed in 100°C for an hour before the agglutination test. Visual agglutination apparent within 20 s was recorded as a positive result. The antiserum that agglutinated all corresponding serotype strains but did not agglutinate any other strains from other serotype groups was referred to as specific antiserum for the corresponding serotype.

**TABLE 1** | *E. albertii* strains used in this study.

Strain	Source	Geography	Isolation time	Serotype	MLST sequence type (ST)	Accession number of O-AGCs	References
LMG20976	Faeces of diarrhoeal child	Dhaka, Bangladesh	90's	O7	ST383	KY574555	Wang H. et al., 2016
SP140128	Mutton	Zigong city, Sichuan province	2014	O1.1	ST4479	KY574559	Wang H. et al., 2016
SP140150	Chicken intestines	Zigong city, Sichuan province	2014	O1.2	ST4479	KY574562	Wang H. et al., 2016
SP140152	Chicken intestines	Zigong city, Sichuan province	2014	O2	ST3762	KY574563	Wang H. et al., 2016
SP140089	Chicken meat	Zigong city, Sichuan province	2014	O1.3	ST4479	KY574558	Wang H. et al., 2016
SP140149	Chicken intestines	Zigong city, Sichuan province	2014	O1.3	ST4479	KY574561	Wang H. et al., 2016
SP140148	Chicken intestines	Zigong city, Sichuan province	2014	O3	ST4633	KY574560	Wang H. et al., 2016
SP140047	Chicken intestines	Zigong city, Sichuan province	2014	O1.4	ST4479	KY574556	Wang H. et al., 2016
SP140084	Chicken meat	Zigong city, Sichuan province	2014	O3	ST4633	KY574557	Wang H. et al., 2016
SP140602	Chicken intestines	Zigong city, Sichuan province	2014	O4	ST4638	KY574564	Wang H. et al., 2016
SP140610	Chicken intestines	Zigong city, Sichuan province	2014	O3	ST4633	KY574565	Wang H. et al., 2016
SP140618	Chicken intestines	Zigong city, Sichuan province	2014	O2	ST3762	KY574566	Wang H. et al., 2016
SP140619	Chicken intestines	Zigong city, Sichuan province	2014	O5	ST1996	KY574567	Wang H. et al., 2016
SP140637	Chicken intestines	Zigong city, Sichuan province	2014	O4	ST4634	KY574568	Wang H. et al., 2016
SP140638	Chicken intestines	Zigong city, Sichuan province	2014	O4	ST4596	KY574569	Wang H. et al., 2016
SP140645	Chicken intestines	Zigong city, Sichuan province	2014	O4	ST4636	KY574570	Wang H. et al., 2016
SP140674	Chicken intestines	Zigong city, Sichuan province	2014	O6	ST4637	KY574571	Wang H. et al., 2016
SP140692	Duck meat	Zigong city, Sichuan province	2014	O2	ST3762	KY574572	Wang H. et al., 2016
SP140701	Chicken meat	Zigong city, Sichuan province	2014	O4	ST4638	KY574573	Wang H. et al., 2016
SP140724	Chicken intestines	Zigong city, Sichuan province	2014	O3	ST4633	KY574574	Wang H. et al., 2016
SP140733	Duck intestines	Zigong city, Sichuan province	2014	O4	ST4634	KY574575	Wang H. et al., 2016
SP140748	Duck intestines	Zigong city, Sichuan province	2014	O1.5	ST4479	KY596022	Wang H. et al., 2016
SP140749	Duck intestines	Zigong city, Sichuan province	2014	O1.6	ST4479	KY574576	Wang H. et al., 2016
SP140753	Duck intestines	Zigong city, Sichuan province	2014	O1.7	ST4479	KY574577	Wang H. et al., 2016
SP140754	Duck intestines	Zigong city, Sichuan province	2014	O1.4	ST4479	KY574578	Wang H. et al., 2016
SP140771	Chicken intestines	Zigong city, Sichuan province	2014	O3	ST4633	KY574579	Wang H. et al., 2016
SP140791	Pork meat	Zigong city, Sichuan province	2014	O5	ST4619	KY574580	Wang H. et al., 2016
SP140807	Chicken intestines	Zigong city, Sichuan province	2014	O4	ST4636	KY574581	Wang H. et al., 2016
SP140813	Chicken intestines	Zigong city, Sichuan province	2014	O3	ST4633	KY574582	Wang H. et al., 2016
SP140837	Duck intestines	Zigong city, Sichuan province	2014	O2	ST4639	KY574583	Wang H. et al., 2016
SP140839	Duck meat	Zigong city, Sichuan province	2014	O2	ST4639	KY574584	Wang H. et al., 2016
ZG141049	Diarrhea patient	Zigong city, Sichuan province	2014	O7	ST4947	KY574602	This study
ZG140067	Healthy human	Zigong city, Sichuan province	2014	O7	ST4947	KY574601	This study
D140513	Faeces of egret	Zigong city, Sichuan province	2014	O4	ST4634	KY574553	This study
SP150020	Duck intestines	Zigong city, Sichuan province	2015	O2	ST3762	KY574585	This study
SP150021	Duck intestines	Zigong city, Sichuan province	2015	O2	ST3762	KY574586	This study
SP150027	Duck intestines	Zigong city, Sichuan province	2015	O2	ST3762	KY574587	This study
SP150036	Duck intestines	Zigong city, Sichuan province	2015	O5	ST1996	KY574588	This study
SP150104	Duck intestines	Zigong city, Sichuan province	2015	O1.1	ST4479	KY574589	This study
T150072	Faeces of healthy butcher	Zigong city, Sichuan province	2015	O6	ST4619	KY574598	This study
T150248	Faeces of healthy butcher	Zigong city, Sichuan province	2015	O5	ST4619	KY574599	This study
T150298	Faeces of healthy butcher	Zigong city, Sichuan province	2015	O5	ST4637	KY574600	This study
SH13EC413	Faeces of diarrhoeal patient	Shanghai	2013	O6	ST4488	KY596024	This study
SP150175	Duck intestines	Zigong city, Sichuan province	2015	O1.1	ST4606	KY574590	This study
SP150183	Duck intestines	Zigong city, Sichuan province	2015	O1.1	ST4606	KY574591	This study

(Continued)

TABLE 1 | Continued

Strain	Source	Geography	Isolation time	Serotype	MLST sequence type (ST)	Accession number of O-AGCs	References
SP150185	Duck intestines	Zigong city, Sichuan province	2015	O6	ST4637	KY596023	This study
SP150193	Chicken intestines	Zigong city, Sichuan province	2015	O1.1	ST4479	KY574592	This study
SP150242	Duck intestines	Zigong city, Sichuan province	2015	O1.8	ST4606	KY574593	This study
SP150249	Duck intestines	Zigong city, Sichuan province	2015	O1.9	ST4606	KY574594	This study
SP150253	Duck intestines	Zigong city, Sichuan province	2015	O1.9	ST4606	KY574595	This study
SP150265	Duck intestines	Zigong city, Sichuan province	2015	O1.10	ST4606	KY574596	This study
SP150270	Duck intestines	Zigong city, Sichuan province	2015	O3	ST4633	KY574597	This study
LX16057	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX160195	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX160135	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4606		This study
LX16061	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX16053	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX160190	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX160154	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4606		This study
LX16058	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4479		This study
LX160100	Chicken intestines	Luzhou city, Sichuan province	2016	O1	ST4606		This study
LX16059	Chicken intestines	Luzhou city, Sichuan province	2016	O2	New		This study
LX16054	Chicken intestines	Luzhou city, Sichuan province	2016	O2	ST3762		This study
LX160162	Chicken intestines	Luzhou city, Sichuan province	2016	O4	ST4638		This study

## Development of the High Throughput xTAG Luminex Detection Assay

MagPlex -xTAG Microspheres (superparamagnetic beads in 6.5 microns diameter) precoupled with a 24- base oligonucleotide “anti-TAG” sequence were used in the assay. Sequences and working concentration of the primers used to amplify *E. albertii* specific gene *lysP* (Hyma et al., 2005) and the serotype-specific *wzy* gene were listed in Table 2. The primers were designed based on the principles described in previous study (Bai et al., 2015). Briefly, the lengths of the primers were between 22 and 26 oligonucleotides, their melting temperatures were between 49 and 52°C, and the amplification size ranged between 100 and 500 base pairs. In order to conjugate with MagPlex -xTAG Microspheres, a corresponding 24- base oligonucleotide “TAG” sequence was added at the 5' terminus of each upstream primer. To facilitate the interaction between MagPlex -xTAG Microspheres and amplified productions, a spacer was made by incorporating a 12-carbon amine containing group between “TAG” sequence and primer. Each downstream primer was biotinylated at the 5' terminus. The products were amplified using cycling parameters at 94°C for 5 min; 30 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 30 s, followed by a final elongation step at 72°C for 10 min. The threshold of the detection limit was determined by using serially diluted DNA from a representative strain of each serotype. Pure genomic templates from 113 stocked strains used in our previous study (Wang Y. et al., 2016) were also used to determine the specificity of the system in the study. Two independent experiments were performed to establish the sensitivity and specificity of the system.

## Ethics Statement

This study was reviewed and approved by the ethics committee of the National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention. The rights and the welfare of the rabbits used in the study were adequately protected. All necessary steps were taken to minimize suffering and distress to the rabbits in these studies.

## RESULTS

### Grouping and General Features of the O-AGCs

On the basis of sequences and genetic structures of the entire O-AGC regions, the O-AGCs from 52 strains were placed into seven groups (denoted O1–O7) where O1 ( $n = 17$ ) was the most prevalent, followed by O4 ( $n = 9$ ), O2 ( $n = 8$ ), O3 ( $n = 6$ ), O5 ( $n = 5$ ), O6 ( $n = 4$ ), and O7 ( $n = 3$ ). In our previous studies, we dissected the chemical structures of the O-specific polysaccharides (OPSs) of *E. albertii* O1, O3, O4, O6, and O7 (Naumenko et al., 2017; Zheng et al., 2017). The predicted gene functions of O-AGCs were consistent with their OPS structures. The data indicated that O-AGCs extracted are responsible for the O-antigen synthesis of *E. albertii*.

All O-AGCs carried the *wzx* and *wzy* genes. The DNA sequence identities of the *wzy* and *wzx* genes were >99.9% within the same serotype group and <8% among different serotype groups. Genes coding for enzymes involved in the synthesis of sugars forming the O subunit and glycosyltransferases were found in each serotype group.

Significant differences among the seven groups were also observed where the size ranged from 7.2 kb (O5, including 7



**TABLE 2** | Serotype-specific primers used in this study.

Serotype	No. of beads coupled with "anti-TAG" sequences	Sequence (5'–3')	Working concentration ( $\mu$ M)	Sensitivity (pg)	PCR product size (bp)
O1	39	<u>ACAAATATCTAACTACTATCACAA</u> – <b>12C</b> –TCCAGTCTTCTTTCGGAAATTT *AAGTTCATGCGTGGAAAAATAC	0.3	0.5	107
O2	42	<u>CACTACACATTTATCATAACAAAT</u> – <b>12C</b> –ATAGCGGGGTATTGGATTTAC *TACAACCGACAAGAAGAAACAA	0.3	1	232
O3	43	<u>AACTTCTCTCTCTATTCTTATT</u> – <b>12C</b> –ATCTTCACGCTCTTTTACTGA *TATAACCTGCAATTACCGAAG	0.3	1	275
O4	44	<u>TCATCACTTTCTTTACTTTACATT</u> – <b>12C</b> –TTACTGCGTTGATGAAAGTTTG *CGCAATAACGGTAAACAAAGAA	0.3	0.5	102
O5	45	<u>TACACAATATTCATCATACTAAC</u> – <b>12C</b> –GCGGGGATTATTACTTTTAGGT *CTCCATATCGCAGGTCAAATA	0.3	10	271
O6	46	<u>TTAAACAATCTACTATTCAATCAC</u> – <b>12C</b> –GTGGGTGAAAGTAAGGTCAATA *TCTGAAAATGGGATGAATGACA	0.3	1	212
O7	47	<u>TCTCTTTAAACACATTCAACAATA</u> – <b>12C</b> –AGATATAACGTCGGCATTGATT *ATAGCAACCCAACCCACATAAAA	0.3	10	250
All ( <i>lysP</i> )	48	<u>AATCAACACACAATAACATTCATA</u> – <b>12C</b> –GGGCGCTGCTTTCATATATTCTT *TCCAGATCCAACCGGGAGTATCAGGA	0.1	1	252

*xTAG* sequences and the 12-carbon amine containing group are indicated by underline and bold text, respectively. \*Indicates reverse primer is biotinylated at the 5' terminus.

genes) to 16.4 kb (O4, including 16 genes) and the G+C content of seven O-AGCs ranged from 29.2% (O5) to 38.9% (O4). Simple insertions of transposase genes were found in *E. albertii* O1, O6, and O7 without any gene disruption (**Figure 1**). Genetic heterogeneity was only found within *E. albertii* O1-AGC. Within the six other O-AGC groups, we observed high sequence conservation (>99% DNA sequence identity).

Within *E. albertii* O1-AGC, all strains harbored a transposase gene in the 5' region. The tandem repeat number of "CTCTG" in the transposase gene was different between the strains (eight types of tandem repeat were found from 11 copies to 73 copies). Meanwhile, two types of transposase gene were found in central regions of *E. albertii* O1-AGC. Nineteen *E. albertii* O1 strains were assigned into 10 subtypes based on the variable sequence and organization of transposase genes, named *E. albertii* O1.1-AGC to O1.10-AGC. *E. albertii* O1.1-AGC ( $n = 5$ ) was the dominant subtype (**Table 1**).

## MLST of *E. albertii*

Thirteen sequence types (STs) were found in 52 strains. Eleven of these were reported previously (Wang H. et al., 2016). *E. albertii* O1, O2, O4, O5, and O7 contained multiple STs (**Table 1**). All strains with identical STs were clustered into the same serotype group.

## Comparison of O-AGCs between *E. albertii* and Other Species

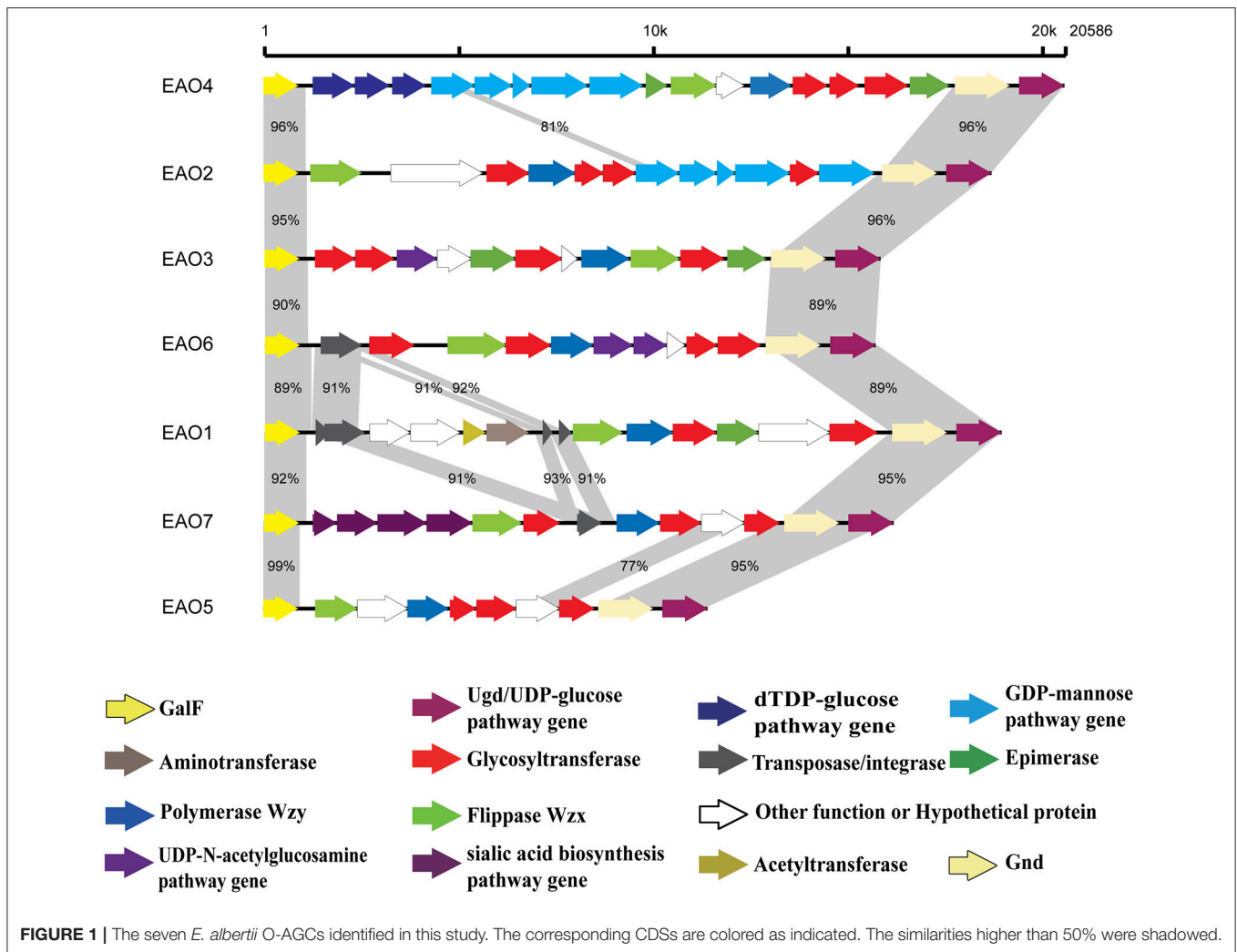
In our previous studies, we have found that the O-antigen of *E. albertii* O1, O3, O4, O6, and O7 is structurally and genetically related to the O-antigens of other species (Naumenko et al., 2017; Zheng et al., 2017). In the study, we compared the O-AGCs of *E. albertii* O2 and O5 to those of other species.

*orf8* to *orf13* in O-AGC of *E. albertii* O2 showed 91% similarity to genes (*orf10* to *orf15*) in O-AGC of *E. coli* O156 (GenBank accession no. AB812065). Five from them (*orf8-11* and *orf13*) are homologs of *gmd*, *fcl*, *gmm*, *manB* and *manC* genes. *manB* and *manC* genes were involved in the synthesis of GDP-d-Man from Fru-6-P. *gmd* and *fcl* genes are responsible for the synthesis of GDP-l-Fuc from GDP-d-Man. The protein encoded by *gmm* gene, which catalyzes hydrolysis of GDP-d-Man to yield GDP and d-Man, has been suggested to participate in the regulation of cell wall biosynthesis by influencing the concentration of GDP-d-Man in the cell wall (Perepelov et al., 2015; Duan et al., 2016). *orf12* of *E. albertii* O2 is homologs of *orf14* (glycosyltransferase) of *E. coli* O156. *orf14* of *E. coli* O156 named *wefY* is responsible for the formation of  $\alpha$ -l-Fucp-(1 $\rightarrow$ 3)-d-GlcpNAc linkage (Duan et al., 2016). It is reasonable to propose that *E. albertii* O2 has the similar linkage.

Two ORFs of *E. albertii* O5 (*orf6-7*) showed 80% and 76% identities in amino acid to the *glf* and *wfeT* genes of *Shigella dysenteriae* type 3 O-antigen gene cluster (GenBank: EU296415), respectively. *E. albertii* O5 may share the OPS main chain synthesized by *glf* and *wfeT* genes with *S. dysenteriae* type 3 (Liu et al., 2008).

## Distribution of Seven O-AGCs in Published *E. albertii* Sequences

Four complete genome sequences and 25 draft sequences of *E. albertii* strains have been reported recently (Fiedoruk et al., 2014; Ooka et al., 2015). Meanwhile, two additional genome draft sequences (BBMY00000000 and NZ\_CH991859.1) were available in the NCBI database. Analysis of their O-AGCs was performed and six types of O-AGCs reported in the present study were also found in 15 previously published genomes. Contrary to Chinese



strains, *E. albertii* O7 ( $n = 4$ ) was the most prevalent amongst these, followed by O2 ( $n = 3$ ), O1 ( $n = 2$ ), O3 ( $n = 2$ ), O4 ( $n = 2$ ), and O5 ( $n = 2$ ) (Table S1). Another 16 published genomes harbored novel O-AGCs, which were placed into 13 groups. All 16 O-AGCs possessed similar characteristics to those of *E. albertii* O1-7 in that: (i) located in a fixed region between *galF* and *gnd*; (ii) carried the *wzx* and *wzy* genes. It is noteworthy that *wzy* was serotype-specific gene of 13 O-AGCs.

The size ranged from 8.1 kb (CB9791, including 8 genes) to 16.4 kb (NIAH\_Bird\_23, including 14 genes) and the G+C content ranged from 31.5% (K7394) to 39.1% (CB9791). Thirteen of them were found homologies in O-AGCs of *E. coli*, including O41, O49, O58, O65, O115, O128, O130, O152, O182, and O184 (Table S1). Three of them were not found homologies in any O-AGCs of other species.

### Establishment of *E. albertii* O Serogroups

Seven antisera were ultimately selected for the current *E. albertii* serotyping scheme, and all 52 tested *E. albertii* strains were clearly assigned to one of these seven serotypes. The agglutination results

of 52 strains were completely consistent with their O-AGCs grouping results. All antisera were specific for their homologous strains. In general, homologous titers were high, varying from 640 to 1280 (Table 3). An additional 12 strains isolated in 2016 were also typed using the seven antisera. Nine of these were typed as *E. albertii* O1, while the others were typed as *E. albertii* O2 (2) and *E. albertii* O4 (1), respectively (Table 1).

### Development and Evaluation of a High Throughput xTAG Luminex Assay to Simultaneously Detect Seven O-AGCs

Specific detection was based on the unique sequence of *wzy* for each O-AGC (Table 2). The *wzy* gene was amplified in a multiplex PCR format. The detection limit for the seven O-AGCs varied from 0.5 to 10 pg of purified DNA per reaction.

The performance of this system was tested on 64 strains used in the study. Cross- and non-specific amplification between sequences was not observed. All 64 strains were correctly designated to corresponding serotype groups, which were completely consistent with their agglutination test results.

**TABLE 3** | Agglutination of *E. albertii* antisera\*.

Antiserum type	Agglutination titer to serotype strains						
	O1 (SP140089)	O2 (SP150020)	O3 (SP140724)	O4 (D140513)	O5 (T150248)	O6 (T150072)	O7 (ZG141049)
O1	1,280						
O2		1,280					
O3			1,280			<20	
O4				1,280			
O5			<20		1,280		
O6						1,280	
O7						<20	640

\*Agglutination titers lower than 10 are not shown.

## DISCUSSION

Serotyping remains the “gold standard” for identifying and monitoring organisms. The chemical composition and structure of the O antigen show high levels of variation even within a single species revealing it to be serologically diverse. Prior to the present study, little was known about the distribution and diversity of O-AGCs in *E. albertii*. Combination of chemical structures of the O-specific polysaccharides (OPSs) of *E. albertii* (Naumenko et al., 2017; Zheng et al., 2017) and sequence analysis, serotyping scheme in the present study, we defined the characteristics of O-AGCs in *E. albertii*: (i) Similar to *E. coli*, all O-AGCs of *E. albertii* were located in a fixed region of the genome between *galF* and *gnd*. (ii) High diversity among different O-AGCs group was observed. (iii) O-antigens of *E. albertii* were synthesized by the Wzx/Wzy-dependent pathway. The Wzx/Wzy-dependent assembly pathway is conserved in a wide range of both Gram-negative and Gram-positive bacteria, and is encoded in dedicated gene clusters. Within 185 well identified O-AGCs in *E. coli*, 174 of these were synthesized by the Wzx/Wzy pathway (Iguchi et al., 2015). Within 90 serotypes of *Streptococcus pneumoniae*, 88 of these were synthesized by the Wzx/Wzy pathway (Bentley et al., 2006).

In this study, we found and named seven serotypes as *E. albertii* O1–O7 in Chinese strains. Amongst these, the O1 serotype comprised approximately 40% (26/64) of all the strains, which was the most dominant serotype. Even though *E. albertii* O1 was the dominant serotype identified in this study, it was not found in strains from diarrheal patients in the current and previous studies (Fiedoruk et al., 2014; Ooka et al., 2015) (Table 1 and Table S1). Further studies are needed to understand the relationship between serotype and pathogenic potential. Different from Chinese strains, 19 types of O-AGCs were found in 31 public genomes of *E. albertii*, which were composed of 6 serotypes reported in the present study and 13 different O-AGCs. This may suggest that host specificity and ecological environment may contribute to the serotype diversity of strains between China and other countries.

To date, 20 O-AGCs of *E. albertii* were identified. It is relatively lower compared to 185 O serogroups of *E. coli* (Iguchi et al., 2015), 54 serotypes of *Shigella* spp. (Muthuirulandi Sethuvel et al., 2017) and more than 200 serotypes of *Vibrio cholerae* (Stine

and Morris, 2014). We cannot rule out the possibility that more serotypes may be found with additional testing of *E. albertii* strains.

In this study, we performed an O-antigen serotyping scheme for *E. albertii* based on specific antisera against seven O-antigens, but the conventional serotyping method using the agglutination test with serotype-specific antisera is laborious, time-consuming and expensive. High-throughput molecular serotyping methods allow for simultaneous detection of multiple nucleic acid sequences in a single reaction, and can greatly reduce the time, cost, and work. These technologies have become attractive alternatives to conventional serotyping methods. mPCR coupled to Luminex xTAG technology-based detection provides a clear and attractive approach for multiplex analysis. The low conservation between *wzy* genes of different serotypes means *wzy* gene is an excellent molecular marker for molecular serotyping. In present study, a high throughput xTAG Luminex assay using unique sequence of *wzy* for each serotype to simultaneously detect seven O-AGCs was developed. All the tested strains were accurately typed into seven O-AGCs which were completely consistent with their seroagglutination results. The detection system can be completed in 40 min post-PCR amplification. The limitation of the system is that only seven serotypes revealed in the study can be detected. Thirteen O-AGCs present in public genome of *E. albertii* (Table S1) were not added in the system for lacking strains to evaluate the system. However, the system has great potential to increase the multiplicity in a single reaction.

Comparing to whole genome, the pretty low G+C content of 20 O-AGCs suggest that they may have originated from a different species. *E. albertii* has recently been recognized as a close relative of *E. coli* (Ooka et al., 2015). It is noteworthy that 12 O-AGCs of *E. albertii* were found homologies in O-AGCs of *E. coli*. Additionally, many genes present in O-AGCs of *E. albertii* were also widely distributed in O-AGCs of *E. coli*. Moreover, all strains of *E. albertii* O3 and O6 were agglutinated with *E. coli* O181 and O3 serum, respectively. It is noteworthy that identical O-AGCs among different species have also been reported in previous studies (Sugiyama et al., 1997; Cheng et al., 2006; Feng et al., 2007). Meanwhile, the *E. albertii* strains with identical O-antigens were isolated from diverse sources and belonged to different sequence types. The finding suggested the O-AGCs

can also readily spread among *E. albertii* strains, even among *Enterobacteriaceae*. Further studies are needed to understand the ability of this organism to spread and cause disease.

In conclusion, our data revealed the highly genetic diversity of O-AGCs in *E. albertii* and that *E. albertii* O1 was the dominant serotype. Our study provided valuable serotyping methods for the epidemiological study of this newly emerging enteric pathogen.

## AUTHOR CONTRIBUTIONS

HZ, HJ, JX, and YX designed the project; HW, QL, XL, LZ, NZ, GY, and ZZ carried out the sampling work; HW, HZ, YXu, JW, PD, and XqL carried out the experiments and generated data; HZ and YX analyzed data and drafted the manuscript. All authors have read and approved the final version of the manuscript.

## REFERENCES

- Abbott, S. L., O'Connor, J., Robin, T., Zimmer, B. L., and Janda, J. M. (2003). Biochemical properties of a newly described *Escherichia* species, *Escherichia albertii*. *J. Clin. Microbiol.* 41, 4852–4854. doi: 10.1128/JCM.41.10.4852-4854.2003
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Asoshima, N., Matsuda, M., Shigemura, K., Honda, M., Yoshida, H., Hiwaki, H., et al. (2014). Identification of *Escherichia albertii* as a causative agent of a food-borne outbreak occurred in 2003. *Jpn. J. Infect. Dis.* 67, 139–140. doi: 10.7883/yoken.67.139
- Bai, X., Liu, Z., Ji, S., Gottschalk, M., Zheng, H., and Xu, J. (2015). Simultaneous detection of 33 *Streptococcus suis* serotypes using the luminex xTAG(R) assay. *J. Microbiol. Methods* 117, 95–99. doi: 10.1016/j.mimet.2015.07.018
- Bateman, A., Birney, E., Cerruti, L., Durbin, R., Ewinger, L., Eddy, S. R., et al. (2002). The Pfam protein families database. *Nucleic Acids Res.* 30, 276–280. doi: 10.1093/nar/30.1.276
- Bentley, S. D., Aanensen, D. M., Mavroidi, A., Saunders, D., Rabinowitsch, E., Collins, M., et al. (2006). Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet.* 2:e31. doi: 10.1371/journal.pgen.0020031
- Brandal, L. T., Tunstø, H. S., Ranheim, T. E., Lobersli, I., Lange, H., and Wester, A. L. (2015). Shiga toxin 2a in *Escherichia albertii*. *J. Clin. Microbiol.* 53, 1454–1455. doi: 10.1128/JCM.03378-14
- Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis Comparison Tool. *Bioinformatics* 21, 3422–3423. doi: 10.1093/bioinformatics/bti553
- 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. *Curr. Microbiol.* 53, 470–476. doi: 10.1007/s00284-006-0032-7
- DeRoy, C., Fratamico, P. M., Yan, X., Baranzoni, G., Liu, Y., Needleman, D. S., et al. (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
- Duan, Z., Senchenkova, S. N., Guo, X., Perepelov, A. V., Shashkov, A. S., Liu, B., et al. (2016). Structure and gene cluster of the O-antigen of *Escherichia coli* O156 containing a pyruvic acid acetal. *Carbohydr. Res.* 430, 24–28. doi: 10.1016/j.carres.2016.04.025
- Duerr, C. U., Zenk, S. F., Chassin, C., Pott, J., Gutle, D., Hensel, M., et al. (2009). O-antigen delays lipopolysaccharide recognition and impairs antibacterial host defense in murine intestinal epithelial cells. *PLoS Pathog.* 5:e1000567. doi: 10.1371/journal.ppat.1000567

## ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (81371762, 81290345), the State Key Laboratory of Infectious Disease Prevention and Control (2015SKLID504, 2016SKLID309), the National Basic Research Priorities Program of China (2015CB554201), the National Key Research and Development Plan (2016YFC1201903, 2016YFC1202700), and the Health and Family Planning Commission of Sichuan Province (150259).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01857/full#supplementary-material>

- Eserstam, R., Rajaguru, T. P., Jansson, P. E., Weintraub, A., and Albert, M. J. (2002). The structure of the O-chain of the lipopolysaccharide of a prototypal diarrheagenic strain of *Hafnia alvei* that has characteristics of a new species under the genus *Escherichia*. *Eur. J. Biochem.* 269, 3289–3295. doi: 10.1046/j.1432-1033.2002.03009.x
- Feng, L., Perepelov, A. V., Zhao, G., Shevelev, S. D., Wang, Q., Senchenkova, S. N., et al. (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. doi: 10.1099/mic.0.2006/001107-0
- Fiedoruk, K., Daniluk, T., Swiecicka, I., Murawska, E., Sciepek, M., and Leszczynska, K. (2014). First complete genome sequence of *Escherichia albertii* strain KF1, a new potential human enteric pathogen. *Genome Announc.* 2:e00004-14. doi: 10.1128/genomeA.00004-14
- Huys, G., Cnockaert, M., Janda, J. M., and Swings, J. (2003). *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int. J. Syst. Evol. Microbiol.* 53, 807–810. doi: 10.1099/ijs.0.02475-0
- Hyma, K. E., Lacher, D. W., Nelson, A. M., Bumbaugh, A. C., Janda, J. M., Strockbine, N. A., et al. (2005). Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. *J. Bacteriol.* 187, 619–628. doi: 10.1128/JB.187.2.619-628.2005
- Iguchi, A., Iyoda, S., Kikuchi, T., Ogura, Y., Katsura, K., Ohnishi, M., et al. (2015). A complete view of the genetic diversity of the *Escherichia coli* O-antigen biosynthesis gene cluster. *DNA Res.* 22, 101–107. doi: 10.1093/dnares/dsu043
- Inglis, T. J., Merritt, A. J., Bzdyl, N., Lansley, S., and Urosevic, M. N. (2015). First bacteraemic human infection with *Escherichia albertii*. *New Microbes New Infect.* 8, 171–173. doi: 10.1016/j.nmni.2015.07.003
- Liu, B., Knirel, Y. A., Feng, L., Perepelov, A. V., Senchenkova, S. N., Wang, Q., et al. (2008). Structure and genetics of Shigella O antigens. *FEMS Microbiol. Rev.* 32, 627–653. doi: 10.1111/j.1574-6976.2008.00114.x
- Mullane, N., O'gaora, P., Nally, J. E., Iversen, C., Whyte, P., Wall, P. G., et al. (2008). Molecular analysis of the *Enterobacter sakazakii* O-antigen gene locus. *Appl. Environ. Microbiol.* 74, 3783–3794. doi: 10.1128/AEM.02302-07
- Murakami, K., Etoh, Y., Tanaka, E., Ichihara, S., Horikawa, K., Kawano, K., et al. (2014). Shiga toxin 2f-producing *Escherichia albertii* from a symptomatic human. *Jpn. J. Infect. Dis.* 67, 204–208. doi: 10.7883/yoken.67.204
- Murray, G. L., Attridge, S. R., and Morona, R. (2003). Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol. Microbiol.* 47, 1395–1406. doi: 10.1046/j.1365-2958.2003.03383.x
- Murray, G. L., Attridge, S. R., and Morona, R. (2006). Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. *J. Bacteriol.* 188, 2735–2739. doi: 10.1128/JB.188.7.2735-2739.2006



- Muthuirulandi Sethuvel, D. P., Devanga Ragupathi, N. K., Anandan, S., and Veeraraghavan, B. (2017). Update on: shigella new serogroups/serotypes and their antimicrobial resistance. *Lett. Appl. Microbiol.* 64, 8–18. doi: 10.1111/lam.12690
- Naumenko, O. I., Zheng, H., Senchenkova, S. N., Wang, H., Li, Q., Shashkov, A. S., et al. (2017). Structures and gene clusters of the O-antigens of *Escherichia albertii* O3, O4, O6, and O7. *Carbohydr. Res.* 449, 17–22. doi: 10.1016/j.carres.2017.06.008
- Ooka, T., Ogura, Y., Katsura, K., Seto, K., Kobayashi, H., Kawano, K., et al. (2015). Defining the genome features of *Escherichia albertii*, an emerging enteropathogen closely related to *Escherichia coli*. *Genome Biol. Evol.* 7, 3170–3179. doi: 10.1093/gbe/evv211
- Ooka, T., Tokuoka, E., Furukawa, M., Nagamura, T., Ogura, Y., Arisawa, K., et al. (2013). Human gastroenteritis outbreak associated with *Escherichia albertii*, Japan. *Emerg. Infect. Dis.* 19, 144–146. doi: 10.3201/eid1901.120646
- Perepelov, A. V., Guo, X., Filatov, A. V., Liu, B., and Knirel, Y. A. (2015). Structure and gene cluster of the O-antigen of *Escherichia coli* O43. *Carbohydr. Res.* 416, 32–36. doi: 10.1016/j.carres.2015.08.008
- Saldias, M. S., Ortega, X., and Valvano, M. A. (2009). *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *J. Med. Microbiol.* 58, 1542–1548. doi: 10.1099/jmm.0.013235-0
- Stine, O. C., and Morris, J. G. Jr. (2014). Circulation and transmission of clones of *Vibrio cholerae* during cholera outbreaks. *Curr. Top. Microbiol. Immunol.* 379, 181–193. doi: 10.1007/82\_2013\_360
- Sugiyama, T., Kido, N., Kato, Y., Koide, N., Yoshida, T., and Yokochi, T. (1997). Evolutionary relationship among rfb gene clusters synthesizing mannose homopolymer as O-specific polysaccharides in *Escherichia coli* and *Klebsiella*. *Gene* 198, 111–113. doi: 10.1016/S0378-1119(97)00300-4
- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22–28. doi: 10.1093/nar/29.1.22
- Valvano, M. A. (2003). Export of O-specific lipopolysaccharide. *Front. Biosci.* 8, s452–s471. doi: 10.2741/1079
- Wang, H., Li, Q., Bai, X., Xu, Y., Zhao, A., Sun, H., et al. (2016). Prevalence of eae-positive, lactose non-fermenting *Escherichia albertii* from retail raw meat in China. *Epidemiol. Infect.* 144, 45–52. doi: 10.1017/S0950268815001120
- Wang, Y., Wang, Y., Zhang, L., Li, M., Luo, L., Liu, D., et al. (2016). Endonuclease restriction-mediated real-time polymerase chain reaction: a novel technique for rapid, sensitive and quantitative detection of nucleic-acid sequence. *Front. Microbiol.* 7:1104. doi: 10.3389/fmicb.2016.01104
- Zheng, H., Shashkov, A. S., Xiong, Y., Naumenko, O. I., Wang, H., Senchenkova, S. N., et al. (2017). Structure and gene cluster of the O-antigen of *Escherichia albertii* O1 resembling the O-antigen of *Pseudomonas aeruginosa* O5. *Carbohydr. Res.* 446–447, 28–31. doi: 10.1016/j.carres.2017.04.024

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wang, Zheng, Li, Xu, Wang, Du, Li, Liu, Zhang, Zou, Yan, Zhang, Jing, Xu and Xiong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.