HHS Public Access

Author manuscript

Int J Food Microbiol. Author manuscript; available in PMC 2016 May 04.

Published in final edited form as:

Int J Food Microbiol. 2015 May 4; 200: 13–17. doi:10.1016/j.ijfoodmicro.2014.12.009.

Phenotypic and phylogenetic analysis of *Vibrio parahaemolyticus* isolates recovered from diarrhea cases in Guangdong Province, China

Baisheng Li[a,1], **Jinyan Luo**[a,b,1], **Hailing Tan**[a], **Bixia Ke**[a], **Dongmei He**[a], **Changwen Ke**[a], **John D. Klena**[c,d], and **Yonghui Zhang**[a,*]

[a]Center for Disease Control and Prevention of Guangdong Province, China

[b]Center for Disease Control and Prevention of Nanshan, Shenzhen, China

[c]Global Disease Detection Branch, Division of Global Health Protection, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, USA

[d]International Emerging Infections Program, US Centers for Disease Control and Prevention, Beijing, China

Abstract

Vibrio parahaemolyticus has emerged as a common foodborne pathogen of global concern. In this study, 108 *V. parahaemolyticus* isolates that recovered from diarrhea cases (n = 96) and seafood products (n = 12) in Guangdong Province from 2007 to 2011 were characterized by serotyping, *tdh* and *trh* toxin gene detection and multilocus sequence typing (MLST). The dominant serotypes from the cases were O3:K6, O4:K8 and O1: KUT (untyped). However, most isolates recovered from seafood products belonged to other serotypes. None of the isolates carried the *trh* gene, while the major isolates from the cases were *tdh* positive. MLST analysis revealed 31 sequence types (STs); 17 STs were unique in this study. eBURST analysis revealed four clonal complexes (CC), The majority of the isolates (n = 58, all from cases and *tdh*+) were grouped into the CC3, which included O3:K6, O4:K68 and O1:KUT isolates. The CC3 was the most prevalent clonal complex, and all of the CC3 isolates were recovered from clinical cases of geographically diverse origin. As to the CC345, which was completely constituted by O4:K8, was another important clonal complex affecting Guangdong Province. Ongoing surveillance of *V. parahaemolyticus* in diarrhea patients and seafood products remains a public health priority for Guangdong Province, China.

Keywords

Vibrio parahaemolyticus; Multilocus sequence typing; Phylogenetic analysis

[*]Corresponding author. zyh@cdcp.org.cn (Y. Zhang).

[1]These authors contributed equally to this work.

Disclosure statement

No competing financial interests exist.

1. Introduction

Vibrio parahaemolyticus is a Gram negative, genetically and antigenically diverse bacterial species, which naturally inhabits marine and estuarine environments worldwide. Consumption of raw or undercooked oysters, or other seafood products, is major risk factor for *V. parahaemolyticus*-associated diarrhea, especially in coastal areas (Vongxay et al., 2008). The thermostable direct hemolysin (TDH) and the TDH-related hemolysin (TRH) are considered as the most important virulent elements (Shirai et al., 1990). Before the pandemic serotype O3:K6 clone first appeared in Southeast Asia in 1996, *V. parahaemolyticus* was only sporadically identified as a cause of gastro-enteritis in the region and was associated with multiple serotypes (Chen et al., 2011). Since 1996, *V. parahaemolyticus* serotype O3:K6 has emerged as a pandemic clone associated with a dramatic increase in global food-borne illness (Matsumoto et al., 2000).

Guangdong has a large coastline along its southern border and many of its citizens regularly consume raw or undercooked seafood. As a result, *V. parahaemolyticus* has become a common regional pathogen (Zhou and Zhang, 2012). Surveillance for *V. parahaemolyticus* has not been routine, however, so the most prevalent circulating isolates and their associated serotypes and virulence factors, are not known. Further, the phylogenetic relationships between locally circulating O3:K6-expressing isolates and the pandemic O3:K6 clone remain to be understood.

To understand the phenotypic and phylogenetic characteristics of *V. parahaemolyticus* isolates, serotyping, PCR detection of the genes *tdh and trh*, encoding for TDH and TRH, respectively and multilocus sequence typing (MLST) analysis were performed on 108 isolates collected from diarrheal cases and seafood samples in Guangdong Province from 2007 to 2011. We observed that: (1) the top three serotypes were O3:K6, O4:K8 and O1:KUT; (2) 93 of 96 clinical isolates, but only 2 of 12 isolates from seafood, were *tdh* positive; none of the isolates carried *trh*; and (3) 31 different STs were identified, 17 of which were novel to this study. We believe that the results of this study will inform future efforts to detect pathogenic isolates and forecast local disease outbreaks.

2. Materials and methods

2.1. Bacterial isolates

A total of 108 *V. parahaemolyticus* isolates recovered from 2007 to 2011 were included in this study. Among these isolates, 96 were recovered from clinical cases and 12 from seafood products. Samples were pre-enriched in the Alkaline Peptone Water medium (APW) containing 3.5% NaCL at 37 °C overnight before inoculating to the thiosulfate citrate bile-salts sucrose (TCBS) agar. Those green shiny colonies from TCBS were selected for next identification using the oxidize test, triple sugar iron agar (TSI) reaction and serotyped by slide agglutination with a *V. parahaemolyticus* antiserum (Denka Seiken, Tokyo, Japan).

2.2. Ethics statement

All these 96 clinical samples were stool samples and collected by the nurse in the sentinel hospitals assigned by the foodborne disease surveillance project of Guangdong Province. All

samples used in this study were anonymized and received the approval from the Institutional Review Board of Centre for Disease Control and Prevention of Guangdong Province.

2.3. Chromosomal DNA preparation

DNA was extracted from bacterial isolates using the QIAamp DNA Mini kit according to the manufacturer's instructions (Qiagen, Inc., Shanghai, China). DNA extracts were dissolved in Tris–EDTA (10 mM Tris–HCl, 0.10 mM EDTA [pH 8.0]) buffer and stored at 4 °C. Dilutions of template DNA were made with sterile distilled water to obtain a final concentration of approximately 100 ng/mL.

2.4. Detection of toxin genes

A multiplex PCR assay was performed to determine the presence of *tdh* and *trh*. PCR was performed using a thermal cycler (Biometra, Inc., Biometra TGradient, Göttingen, Germany.) with the previously published primers (Bej et al., 1999). Target sequences were amplified after an initial denaturation at 94 °C for 3 min, with 25 cycles consisting of 94 °C for 40 s, 60 °C for 40 s, and 72 °C for 40 s and a final extension step of 72 °C for 4 min. Amplified products were separated on a 1% agarose gel, stained with 1% ethidium bromide, and imaged using a Gel Doc EQ system (Bio-Rad Inc., Hercules, CA, USA).

2.5. MLST

A seven-locus MLST system described previously (Gonzalez-Escalona et al., 2008) was used to genotype the isolates. The *dnaE* (DNA polymerase III, α subunit), *gyrB* (DNA gyrase, subunit B), *recA* (recombinase A), *dtdS* (threonine 3-dehydrogenase), *pntA* (transhydrogenase, α subunit), *pryC* (dihydro-orotase) and *tnaA* (tryptophanase) genes were analyzed as previously described. Primer sequences and conditions for PCR amplification are available at http://pubmlst.org/vparahaemolyticus/info/protocol.html. PCR products were purified using a QIAquick PCR Purification kit as described by the manufacturer (Qiagen). Cycle sequencing was carried for both DNA strands of the amplified targets using M13 universal primers and BigDye chemistry (Perkin Elmer Applied Biosystems, Foster City, CA, USA), with an initial denaturation of 96 °C for 1 min, followed by 25 cycles of denaturation at 96 °C for 10 s, primer annealing at 50 °C for 5 s, and extension at 72 °C for 4 min. Sequence data were analyzed using ABI Prism 3100 Genetic Analyzer (Perkin Elmer Applied Biosystems). Alignment of sequences was performed using DNAStar (version V7.10). Numbers for alleles and sequence types (STs) were assigned according to the PubMLST *V. parahaemolyticus* database.

2.6. Assignment to clonal complexes

The program eBURST v3.0 was used to identify the different clonal complexes (http://eburst.mlst.net). The most restrictive group definition was used to define the clonal complexes, i.e., at least six of the seven alleles had to be identical to be included in the same group or clonal complex (Feil et al., 2004). The statistical confidences for the ancestral types were assessed using 1000 bootstrap resamplings. Two different STs are considered single-locus variant (SLV) when they differ from each other at a single locus. Double-locus variants (DLVs) refer to two STs differing in only two of the seven loci.

2.7. Test for recombination

The START (version 2.0) software package was used to calculate the “standardized” index of association (IAS) (Haubold and Hudson, 2000). This statistical method tests for the null hypothesis of linkage equilibrium; i.e., if IAS = 0, then alleles are independently distributed at all loci analyzed (alleles are in linkage equilibrium) and recombination has occurred frequently. The ratio between the numbers of synonymous (dS) and nonsynonymous (dN) substitutions was calculated by the method of Nei–Gojobori implemented in START. This measures the type of selection occurring at each locus. The hypothesis tested was for neutrality (dN = dS); if dN/dS < 1, then nonsynonymous, sites are under selective constraint or purifying pressure (negative selection); dN/dS > 1 indicates positive selection, and dN/dS = 1 indicates neutrality.

3. Results

Fifteen serotypes were identified among the 108 isolates, dominated by O3:K6 (n = 52; 48%), O4:K8 (n = 20; 18%) and O1:KUT (n = 12; 11%). Only four isolates (two O4:K8 and two O1:KUT) from seafood products expressed one of these serotypes. Ninety-three of the clinical but only two of the seafood product isolates were *tdh* positive. However, none of the isolates carried the *trh* gene (Table 1).

The numbers of alleles observed for each MLST locus were distributed as follows: 16 (*dnaE*), 19 (*gyrB*), 20 (*recA*), 17 (*dtdS*), 14 (*pntA*), 19 (*pyrC*) and 18 (*tnaA*). The number of polymorphic sites observed varied per locus from 17 (*pntA*) to 50 (*recA*). The nucleotide diversity ranged from 0.010 to 0.025, with the highest degree of diversity and percentage of polymorphic sites observed for *dtdS* (0.025, 6.77%) and *recA* (0.022, 6.86%). The ratio of nonsynonymous to synonymous substitutions (dN/dS) was lower than 1 for each locus analyzed using a selection test for neutrality (Table 2).

Thirty-one different STs were identified and 17 STs were novel to this study. ST-3 was the most abundant type (53 out of 108 isolates); isolates with this ST were found in three different serotypes (O3:K6, n = 47, O1:KUT, n = 5 and O4:K68, n = 1). Eight additional STs were represented by more than one isolate including ST-189 (n = 8; all O4:K8), ST-265 (n = 6; all O4:K8), ST-345 (n = 6; all O4:K8), ST-199 (n = 4; all O1:KUT), ST-337 (n = 3; all O4:K34), ST-8 (n = 2; O1:KUT and O1:K56), ST-492 (n = 2; both O3:K6) and ST-787 (n = 2; both O4:K68).

eBURST analysis revealed four clonal complexes (CC3, CC345, CC120 and CC8), and 20 singletons (Fig. 1). CC3 consisted of 58 isolates from five STs [ST-3, founder; ST-431, SLV (*gyrB*); ST-435, SLV (*recA*); ST-661, SLV (*pntA*); and ST-787, SLV (*pyrC*)] and three serovars (O3:K6, O1: KUT and O4:K68). The ST-3 O3:K6 pandemic clone is located within CC3. CC8 consists of two STs (ST-8 and ST-783) differing at the *pyrC* locus. CC120 consists of two STs (ST-120 and ST-447) differing at the *dtdS* and *pyrC* loci, respectively. The CC345 clonal complex consists of three STs [ST-189, founder, ST-265 SLV (*recA*) and ST-345, SLV (*recA*)]. The singletons did not belong to any of the clonal complexes or groups identified in this study.

4. Discussion

In this study, 15 different serotypes of *V. parahaemolyticus* were identified among the isolates recovered. As a traditionally used method for the characterization of *V. parahaemolyticus*, serotyping failed to reveal the depth of genetic and clonal relationships among isolates. It is known that different bacterial serotypes can share common molecular types, such as the MLST profiles (Chao et al., 2009). In this study, several isolates belonging to genetically distant STs shared the same serotype, and isolates with the same STs displayed different serotypes, which suggested that MLST better resolves the phylogenetic relatedness of *V. parahaemolyticus* isolates than serotyping alone; this resolution is essential for investigation of bacterial population structures, and can provide valuable information about the emergence of new epidemic clones.

In this study, 20 singletons were identified among the 31 STs using eBURST. These singletons are currently not part of any recognized clonal complex. Seventeen STs previously unidentified were observed in this study; this result was confirmed by the curator of PubMLST database. New STs were recovered from the patient and related seafood products. The newly discovered ST-787, belonging to CC3 [SLV (pyrC)], suggesting that the housekeeping genes within isolates of CC3 are continuing to evolve. Previous study suggested the *recA* gene was highly diverse (Yu et al., 2011). A recombination test performed as part of this study also showed that the *recA* and *dtdS* loci displayed higher nucleotide diversity than the other five loci. The “standardized” index of association (IAS) of the entire isolate collection was 0.8787, indicating that the alleles were in linkage disequilibrium or were nonrandomly distributed (Table 2). It was consisted with previous study that a nonrandom distribution of alleles in the *V. parahaemolyticus* population is general, although recombination may also be occurring within different subpopulations (Gonzalez-Escalona et al., 2008).

The CC3 in our study was dominated by isolates expressing O3:K6 or its serovariants (O1:KUT and O4:K68) (Ansaruzzaman et al., 2005; Nair et al., 2007). A previous study had showed that the CC3 consists of isolates belonging to a pandemic clonal complex with the lineage originating from ST-3 (Gonzalez-Escalona et al., 2008). Since 1996, this clonal complex has been identified as a dominant serotype from clinical cases of diarrhea reported from various Southeast Asian countries, including India, Japan, Thailand, Bangladesh, Taiwan, and Vietnam, as well as from the United States (Chowdhury et al., 2004). In our study, all the CC3 isolates were recovered from clinical cases of geographically diverse origin. Also, more than half isolates from the cases were formed CC3, which indicated that the CC3 isolates were the most prevalent among clinical cases locally. With respect to CC345, it was completely constituted by O4:K8 isolates consisting of three STs (ST-189, ST-345, and ST-265). In this study, there was no apparent clonal relationship found between the O4:K8 and O3:K6 clonal complex. This suggests that CC345, represented by isolates of O4:K8 serotype, is another important clonal complex affecting Guangdong Province. Additional surveillance and characterization of these isolates, to provide information about the genetic characteristics, pathogenicity and spread of this emerging group, is critical for understanding diarrheal illness associated with *V. parahaemolyticus* in Guangdong.

Acknowledgments

This work was funded by the Mega-projects of Science and Technology Research of China (No. 2012ZX10004-213) and the Guangdong Province Medical Research Foundation (No. A2014090).

References

Ansaruzzaman M, Lucas M, Deen JL, Bhuiyan NA, Wang XY, Safa A, Sultana M, Chowdhury A, Nair GB, Sack DA, Von Seidlein L, Puri MK, Ali M, Chaignat CL, Clemens JD, Barreto A. Pandemic serovars (O3:K6 and O4:K68) of *Vibrio parahaemolyticus* associated with diarrhea in Mozambique: spread of the pandemic into the African continent. J Clin Microbiol. 2005; 43:2559–2562. [PubMed: 15956363]

Bej AK, Patterson DP, Brasher CW, Vickery MC, Jones DD, Kaysner CA. Detection of total and hemolysin-producing *Vibrio parahaemolyticus* in shellfish using multiplex PCR amplification of *tl*, *tdh* and *trh*. J Microbiol Methods. 1999; 36:215–225. [PubMed: 10379807]

Chao GX, Jiao XN, Zhou XH, Yang ZQ, Huang JL, Pan ZM, Zhou LP, Qian XQ. Serodiversity, pandemic O3:K6 clone, molecular typing, and antibiotic susceptibility of foodborne and clinical *Vibrio parahaemolyticus* isolates in Jiangsu, China. Foodborne Pathog Dis. 2009; 6(8):1021–1028. [PubMed: 19630509]

Chen Y, Stine OC, Badger JH, Gil AI, Nair GB, Nishibuchi M, Fouts DE. Comparative genomic analysis of *Vibrio parahaemolyticus*: serotype conversion and virulence. BMC Genomics. 2011; 12:294. [PubMed: 21645368]

Chowdhury NR, Stine OC, Morris JG, Nair GB. Assessment of evolution of pandemic *Vibrio parahaemolyticus* by multilocus sequence typing. J Clin Microbiol. 2004; 42(3):1280–1282. [PubMed: 15004094]

Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol. 2004; 186:1518–1530. [PubMed: 14973027]

Gonzalez-Escalona N, Martinez-Urtaza J, Romero J, Espejo RT, Jaykus LA, DePaola A. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. J Bacteriol. 2008; 190:2831–2840. [PubMed: 18281404]

Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Linkage analysis Bioinformatics. 2000; 16:847–848. [PubMed: 11108709]

Matsumoto C, Okuda J, Ishibashi M, Iwanaga M, Garg P, Rammamurthy T, Wong HC, Depaola A, Kim YB, Albert MJ, Nishibuchi M. Pandemic spread of an O3:K6 clone of *Vibrio parahaemolyticus* and emergence of related strains evidenced by arbitrarily primed PCR and toxRS sequence analyses. J Clin Microbiol. 2000; 38:578–585. [PubMed: 10655349]

Nair GB, Ramamurthy T, Bhattacharya SK, Dutta B, Takeda Y, Sack DA. Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. Clin Microbiol Rev. 2007; 20:39–48. [PubMed: 17223622]

Shirai H, Ito H, Hirayama T, Nakamoto Y, Nakabayashi N, Kumagai K, Takeda Y, Nishibuchi M. Molecular epidemiologic evidence for association of thermostable direct hemolysin (TDH) and TDH-related hemolysin of *Vibrio parahaemolyticus* with gastroenteritis. Infect Immun. 1990; 58:3568–3573. [PubMed: 2228229]

Vongxay K, Pan Z, Zhang X, Wang S, Cheng S, Mei L, Xu C, Fang W. Occurrence of pandemic clones of *Vibrio parahaemolyticus* isolates from seafood and clinical samples in a Chinese coastal province. Foodborne Pathog Dis. 2008; 5:127–134. [PubMed: 18370608]

Yu Y, Hu W, Wu B, Zhang P, Chen J, Wang S, Fang W. *Vibrio parahaemolyticus* isolates from southeastern Chinese coast are genetically diverse with circulation of clonal complex 3 strains since 2002. Foodborne Pathog Dis. 2011; 8:1169–1176. [PubMed: 21883006]

Zhou H, Zhang J. Surveillance of other infectious diarrheal diseases in China. Dis Surv. 2012; 27:184–187. in chinese.

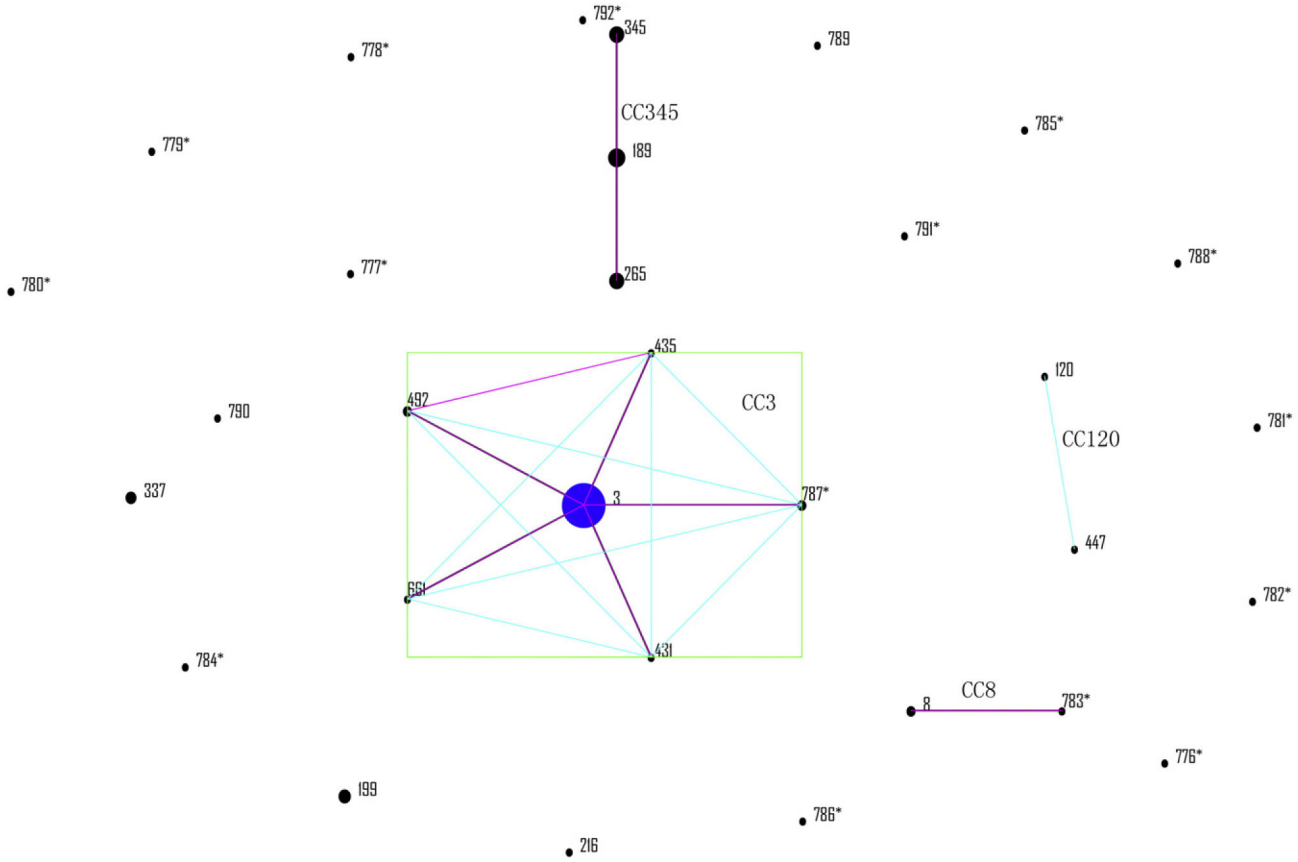

Fig. 1.

V. parahaemolyticus "population snapshot" of 108 strains obtained using eBURST v3. Four groups were defined using stringent criteria (5/7 shared alleles). Among those four groups, the clonal complex CC3 was dominant, and their predicted clonal ancestor-ST-3 is shown in blue. Another three clonal complexes CC345 and CC 8 with SLVs of each other are shown connected by pink lines. CC3 and CC120 with DLV STs are shown connected by chartreuse lines. The sizes of the circles are relative to the numbers of strains in the ST.

Int J Food Microbiol. Author manuscript; available in PMC 2016 May 04.

Table 1

Serotypes, sequence types, allele profiles, toxin genes and sources of the *V. parahaemolyticus* strains analyzed.

Serotypes	No. of strains	ST$^{\psi}$	Allele profiles							Toxin genes		Sources	Clonal complex
			dnaE	*gyrB*	*recA*	*dtdS*	*pntA*	*pyrC*	*tnaA*	*tdh*	*trh*		
O3:K6	47	3	3	4	19	4	29	4	22	+	−	Case	CC3
O3:K6	2	492	3	4	189	4	29	4	22	+	−	Case	
O3:K6	1	431	3	225	19	4	29	4	22	+	−	Case	CC3
O3:K6	1	435	3	4	31	4	29	4	22	+	−	Case	CC3
O3:K6	1	661	3	4	19	4	26	4	22	+	−	Case	CC3
O4:K8	7	189	11	48	3	48	26	48	26	+	−	Case	CC345
O4:K8	1	189	11	48	3	48	26	48	26	+	−	Food	CC345
O4:K8	6	265	11	48	107	48	26	48	26	+	−	Case	CC345
O4:K8	5	345	11	48	19	48	26	48	26	+	−	Case	CC345
O4:K8	1	345	11	48	19	48	26	48	26	+	−	Food	CC345
O1:KUT	5	3	3	4	19	4	29	4	22	+	−	Case	CC3
O1:KUT	4	199	22	28	17	13	8	19	14	+	−	Case	
O1:KUT	1	8	28	4	82	88	63	69	1	+	−	Case	CC8
O1:KUT	1	447	60	108	86	13	18	62	51	−	−	Food	CC120
O1:KUT	1	790*	51	73	46	13	18	249	24	−	−	Food	
O4:K68	1	3	3	4	19	4	29	4	22	+	−	Case	CC3
O4:K68	2	787*	3	4	19	4	29	48	22	+	−	Case	CC3
O1:K56	1	8	28	4	82	88	63	69	1	+	−	Case	CC8
O1:K56	1	783*	28	4	82	88	63	140	1	+	−	Case	CC8
O1:K38	1	777*	51	104	178	146	28	45	12	+	−	Case	
O2:K3	1	216	98	135	112	107	77	97	26	+	−	Case	
O2:K3	1	786*	98	4	112	107	77	4	23	−	−	Food	
O2:K3	1	788*	98	4	30	32	77	48	82	+	−	Case	
O2:K28	1	791*	5	41	217	198	26	48	48	+	−	Case	
O2:K28	1	782*	35	104	26	13	100	5	47	−	−	Food	
O3:K29	1	120	60	108	86	98	18	45	51	+	−	Case	CC120

Serotypes	No. of strains	ST$^{\psi}$	Allele profiles							Toxin genes		Sources	Clonal complex
			dnaE	*gyrB*	*recA*	*dtdS*	*pntA*	*pyrC*	*tnaA*	*tdh*	*trh*		
O4:K9	1	789*	14	30	141	78	4	35	13	+	−	Case	
O4:K34	2	337	47	139	53	19	50	143	26	−	−	Case	
O4:K34	1	337	47	139	53	19	50	143	26	−	−	Food	
O4:K42	1	785*	35	138	101	116	80	4	57	−	−	Food	
O4:KUT	1	776*	49	314	38	219	26	203	26	+	−	Case	
O4:KUT	1	792*	230	167	136	224	50	177	17	−	−	Food	
O5:K17	1	778*	28	23	31	198	4	80	179	−	−	Food	
O5:K17	1	781*	31	132	73	13	4	145	1	−	−	Case	
O5:KUT	1	779*	100	122	31	69	47	202	99	−	−	Food	
O5:KUT	1	780*	44	297	73	33	53	101	33	−	−	Food	
O5:KUT	1	784*	35	43	38	21	31	37	37	+	−	Case	

CC3: clonal complex 3; CC8: clonal complex 8; CC345: clonal complex 345; and CC120: clonal complex 120.

$^{\psi}$ST, sequence type.

*the ST types were first defined in this study.

Table 2

Summary of loci statistics included in the MLST.

Chromosome and locus	Fragment size (bp)	No. of alleles	GC content (%)	Nucleotide diversity (SE)	No. of Polymorphic site (%)	d_N/d_S ratio
I						
dnaE	557	16	48.57	0.011 (0.001)	25 (4.49)	0.0281
gyrB	592	19	47.78	0.015 (0.001)	40 (6.76)	0.0000
recA	729	20	45.59	0.022 (0.002)	50 (6.86)	0.0065
II						
dtdS	458	17	50.27	0.025 (0.002)	31 (6.77)	0.0000
pntA	430	14	43.88	0.010 (0.001)	17 (3.95)	0.0103
pyrC	493	19	48.52	0.011 (0.001)	26 (5.27)	0.0565
tnaA	423	18	48.47	0.013 (0.001)	23 (5.44)	0.0053