Journal of Medical Microbiology (2009), **58**, 000–000 **DOI** 10.1099/jmm.0.2008/000794-0

Molecular detection of all 34 distinct O-antigen forms of *Shigella*

Yayue Li,[1,2,4,5,6]† Boyang Cao,[3,4,5,6]† Bin Liu,[3,4,5,6] Dan Liu,[1]‡ Qili Gao,[7] Xia Peng,[3,4,5,6] Junli Wu,[3,4,5,6] David A. Bastin,[1] Lu Feng[3,4,5,6] and Lei Wang[1,3,4,5,6]

Correspondence
Lei Wang
wanglei@nankai.edu.cn

[1]Tianjin Biochip Corporation, 23 Hongda Street, TEDA, Tianjin 300457, PR China

[2]Tianjin University of Science and Technology, Tianjin 300457, PR China

[3]TEDA School of Biological Sciences and Biotechnology, Nankai University, 23 Hongda Street, TEDA, Tianjin 300457, PR China

[4]Tianjin Research Center for Functional Genomics and Biochips, 23 Hongda Street, TEDA, Tianjin 300457, PR China

[5]Tianjin Key Laboratory of Microbial Functional Genomics, 23 Hongda Street, TEDA, Tianjin 300457, PR China

[6]The Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, PR China

[7]Tianjin Entry-Exit Inspection and Quarantine Bureau, Tianjin 300457, PR China

Received 29 January 2008
Accepted 28 August 2008

Shigella is the cause of shigellosis or bacillary dysentery, the occurrence of which is estimated to be 165 million cases per year worldwide, resulting in 1.1 million deaths. Rapid and reliable assays for detecting and identifying *Shigella* in food, environmental and clinical samples are therefore necessary. *Shigella* species are traditionally identified by their O antigens. This study developed a DNA microarray targeting O-serotype-specific genes to detect all 34 distinct O-antigen forms of *Shigella*, including *Shigella boydii* types 1–18, *Shigella dysenteriae* types 1–13, *Shigella flexneri* types 1–5 and 6, and *Shigella sonnei*. A total of 282 strains were used to test the specificity of the microarray, including 186 *Shigella* and *Escherichia coli* representative strains, 86 *Shigella* clinical isolates and ten strains of other bacterial species that are commonly isolated from food or clinical stool specimens. The oligonucleotide probes were printed on the microarray in concentrations from 1 to 100 μM, and 10 μM proved to be the optimal probe concentration. The detection sensitivity for each serotype was 50 ng genomic DNA or 1 c.f.u. in 25 g milk powder sample following a 6 h enrichment in broth. The microarray is specific, sensitive and reproducible, and, to our knowledge, is the first report of a microarray for serotyping all O-antigen forms of *Shigella*.

INTRODUCTION

Despite the fact that *Shigella* was discovered over a century ago (Shiga, 1906), shigellosis remains an important public health challenge, especially in developing countries. Worldwide, the number of *Shigella* cases exceeds 165 million per year (Kotloff *et al.*, 1999), causing 1.1 million deaths, and more than 65 % of cases are in children under 5 years of age (Bennish, 1991; Koterski *et al.*, 2005). The infectious dose, which can be as low as 10 cells, enables the disease to be spread effectively by contaminated food or water, and also by person-to-person contact (DuPont *et al.*, 1989).

Shigella species are classified serologically by their O antigen only, because they lack H (flagellar) and K (capsular) antigens. Four species are classified on the basis of biochemical and O-antigen serological differences: *Shigella dysenteriae* (13 serotypes); *Shigella flexneri* (14 serotypes); *Shigella boydii* (18 serotypes); and *Shigella sonnei*, which is a single serotype. There are two distinct O-antigen forms for the 14 *Shigella flexneri* serotypes, types 1–5 and type 6 (Cheah *et al.*, 1991; Simmons & Romanowska, 1987), the other serotypes being phage-encoded modifications of the two O-unit structures. *Shigella sonnei*, *Shigella flexneri* and *Shigella dysenteriae* type 1 are the most

†These authors contributed equally to this work.

‡Present address: Division of Molecular Bioscience, John Curtin School of Medical Research, The Australian National University, PO Box 334, Canberra, ACT 2601, Australia.

Abbreviation: EIEC, enteroinvasive *E. coli*.

Table 1. *Shigella* clinical isolates and strains of other bacterial species

Species	Serotype	No. of strains of each source	Total no.	Microarray result
***Shigella* clinical isolates used to test the specificity of the probes (*n*=86)**				
Shigella sonnei		1*, 2†	3	*S. sonnei*
Shigella flexneri	Type 1a	1†, 1‡	2	*S. flexneri* types 1–5§
	Type 1b	1†, 1‡	2	*S. flexneri* types 1–5§
	Type 2a	1†, 1‡	2	*S. flexneri* types 1–5§
	Type 2b	1†, 1‡	2	*S. flexneri* types 1–5§
	Type 3	1†	1	*S. flexneri* types 1–5§
	Type 3a	1‡	1	*S. flexneri* types 1–5§
	Type 3b	1‡	1	*S. flexneri* types 1–5§
	Type 4a	1‡	1	*S. flexneri* types 1–5§
	Type 4b	1†, 1‡	2	*S. flexneri* types 1–5§
	Type 5	1‡	1	*S. flexneri* types 1–5§
	Type 5a	1†	1	*S. flexneri* types 1–5§
	Type 5b	1†	1	*S. flexneri* types 1–5§
	Type 6	1†, 1‡	2	*S. flexneri* type 6
	Type 6a	1‡	1	*S. flexneri* type 6
Shigella boydii	Type 1	1†, 1‡	2	*S. boydii* type 1
	Type 2	1†, 1‡, 1‖	3	*S. boydii* type 2
	Type 3	1†, 1‡	2	*S. boydii* type 3
	Type 4	1†, 1‡	2	*S. boydii* type 4
	Type 5	1‡	1	*S. boydii* type 5
	Type 6	1†, 1‡	2	*S. boydii* type 6
	Type 7	1†, 1‡	2	*S. boydii* type 7
	Type 8	1†, 1‡	2	*S. boydii* type 8
	Type 9	1†, 1‡	2	*S. boydii* type 9
	Type 10	1†, 1‡	2	*S. boydii* type 10
	Type 11	1†, 1‡	2	*S. boydii* type 11
	Type 12	1†, 1‡	2	*S. boydii* type 12
	Type 13	1†, 1‡	2	*S. boydii* type 13
	Type 14	1†, 1‡	2	*S. boydii* type 14
	Type 15	1†, 1‡	2	*S. boydii* type 15
	Type 16	1†, 1‡	2	*S. boydii* type 16
	Type 17	1†, 1‡	2	*S. boydii* type 17
	Type 18	1†, 1‡	2	*S. boydii* type 18
Shigella dysenteriae	Type 1	1†, 1‡	2	*S. dysenteriae* type 1
	Type 2	1†, 1‡	2	*S. dysenteriae* type 2
	Type 3	1†, 1‡	2	*S. dysenteriae* type 3
	Type 4	1†, 1‡	2	*S. dysenteriae* type 4
	Type 5	1†, 1‡	2	*S. dysenteriae* type 5
	Type 6	1†, 1‡	2	*S. dysenteriae* type 6
	Type 7	1†, 1‡	2	*S. dysenteriae* type 7
	Type 8	1†, 1‡	2	*S. dysenteriae* type 8
	Type 9	1†, 1‡	2	*S. dysenteriae* type 9
	Type 10	1†, 1‡, 1‖	3	*S. dysenteriae* type 10
	Type 11	1†, 1‡, 1‖	3	*S. dysenteriae* type 11
	Type 12	1†, 1‡	2	*S. dysenteriae* type 12
	Type 13	1‡	1	*S. dysenteriae* type 13
Other bacterial species used to test the specificity of the probes (*n*=10)				
Salmonella enterica	O51, O54, O59, O60	4‡	4	Negative
Staphylococcus aureus		1‖, 1¶	2	Negative
Bacillus cereus		2¶	2	Negative
Vibrio cholerae	O48, O64	2‡	2	Negative

Table 1. cont.

*ATCC, Manassas, Virginia, USA.
†Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine, Beijing, China.
‡Institute of Medical and Veterinary Science, Adelaide, Australia.
§Strains of *Shigella flexneri* types 1a, 1b, 2a, 2b, 3, 3a, 3b, 4a, 4b, 5, 5a, and 5b hybridized to the specific probes for *Shigella flexneri* types 1–5, a result in line with the fact that they have a common O-antigen gene cluster.
||National Center for Medical Culture Collection, China.
¶Institute of Microbiology, Chinese Academy of Sciences, China.

common species isolated from cases of shigellosis (Ahmed *et al.*, 1997; Bopp *et al.*, 2003).

The O antigen, which consists of many repeats of an oligosaccharide unit (O unit), is part of the lipopolysaccharide in the outer membrane of Gram-negative bacteria and contributes major antigenic variability to the cell surface. In *Shigella*, genes for O-antigen synthesis are normally located in a gene cluster, which maps between *galF* and *gnd* on the chromosome, except for *Shigella sonnei* (Lai *et al.*, 1998). The O-antigen genes generally fall into three main classes: (i) genes of the nucleotide sugar biosynthesis pathways; (ii) glycosyltransferase genes; and (iii) O-unit processing genes [O-antigen flippase gene (*wzx*) and O-antigen polymerase gene (*wzy*)]. Glycosyltransferase and O-unit processing genes are specific to a particular O antigen (Beutin *et al.*, 2005; DebRoy *et al.*, 2005; Feng *et al.*, 2004a, b; Tao *et al.*, 2004, 2005; Wang *et al.*, 2005).

Traditional identification of *Shigella* spp. in the clinical laboratory is based mainly on isolation of *Shigella* by selective culture media, followed by phenotype identification and serotyping. This process may take 3–5 days to obtain results. *Shigella* and *Escherichia coli* are closely related species and share a common molecular origin (Lan & Reeves, 2002). Enteroinvasive *E. coli* (EIEC), enterohaemorrhagic *E. coli* and enteropathogenic *E. coli* all share one or more clinical or pathogenic features with *Shigella* (Johnson, 2000), and many *E. coli* O antigens cross-react serologically with *Shigella* O antigens.

Recently, more rapid approaches for the identification of particular *Shigella* and *E. coli* O serotypes have been developed using PCR assays based on O-serotype-specific genes (DebRoy *et al.*, 2005; Feng *et al.*, 2004a, b; Fratamico *et al.*, 2003; Tao *et al.*, 2005), but most can only identify one pathogen at a time. Furthermore, it is difficult to differentiate bands of a similar size in a multiplex PCR mixture for all distinct O-antigen forms of *Shigella*, and systems such as the Taqman probe-based real-time PCR are limited by the ability to detect only four fluorophores in a single reaction (Panicker *et al.*, 2004).

By contrast, DNA microarrays allow thousands of specific DNA sequences to be detected simultaneously. Such a microarray for classification of five serotypes of *E. coli* and ten serotypes of *Shigella* was developed by us (Li *et al.*, 2006), and served as a prototype for an array of all serotypes of *Shigella* and *E. coli*. In this study, a new DNA microarray was established to detect all 34 distinct O-antigen forms of *Shigella*, including the 10 *Shigella* serotypes used in the previous study (Li *et al.*, 2006). This microarray targeting the *wzx*, *wzy* and glycosyltransferase genes is able to detect all distinct O-antigen forms, including *Shigella boydii* types 1–18, *Shigella dysenteriae* types 1–13, *Shigella flexneri* types 1–5 and 6, and *Shigella sonnei*. We present here an evaluation of the specificity, sensitivity and reliability of this microarray for the serotyping of *Shigella* isolates.

METHODS

Bacterial strains. One hundred and eighty-six representative strains of all *Shigella* and *E. coli* serotypes (Feng *et al.*, 2004b), 86 *Shigella* clinical isolates and 10 strains of *Salmonella enterica*, *Staphylococcus aureus*, *Bacillus cereus* and *Vibrio cholerae* were used in this study (Table 1). Serotypes of the 86 clinical isolates were identified using commercial antisera obtained from Chengdu Institute of Biological Products (data not shown).

Genomic DNA extraction. Genomic DNA was extracted from 1.5 ml broth culture using a DNA centrifugal filter-extraction kit (Tiangen). The DNA template used for sensitivity testing in milk powder samples was extracted as follows. Twenty-five grams of milk powder (Sanlu) was diluted in 225 ml 2-YT broth (1.6 % tryptone, 1 % yeast extract, 0.5 % NaCl, pH 7.0) and the samples were inoculated with 1 ml of 10-fold serial dilutions of cultures of *Shigella* strains. The enrichments were incubated at 37 °C for 6 h in a rotary shaker at 180 r.p.m. and DNA was extracted from 1 ml broth.

Multiplex PCR and labelling of the target genes. Multiplex PCR was used to amplify and label the target genes, and was carried out in three groups (Table 2). All groups contained a primer pair for the 16S rRNA gene (WL-1319/WL-1320; Table 2) from Li *et al.* (2006). Each multiplex PCR mixture contained 1× PCR buffer (50 mM KCl, 10 mM Tris/HCl, pH 8.3), 0.5 mM $MgCl_2$, 100 μM dNTPs, 1.5 U *Taq* DNA polymerase (TaKaRa Biotechnology), 0.047 μM each of two primers based on the 16S rRNA gene, 0.14 μM each of primers based on the target genes of the O antigen, 0.15 nM Cy3–dUTP (Amersham Biosciences) and 50–100 ng DNA template in a final reaction volume of 30 μl. The PCR amplification was performed with initial denaturation at 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 50 °C for 1 min and 72 °C for 1 min, and a final extension at 72 °C for 5 min. Two microlitres of each PCR product was run on an agarose gel to check for the amplicons and the rest was stored at –20 °C.

Oligonucleotide probe design. For each serotype, two to five probes matching the target DNA were designed by OligoArray 2.0 (Rouillard *et al.*, 2003) based on GenBank sequences and an in-house

Table 2. Multiplex PCR primers and oligonucleotide probes used in the study

Serotype	Target gene†	O-antigen gene cluster sequence	PCR group	Primer sequences (5′→3′)	Product size (bp)	Probe sequences (5′→3′)
Shigella sonnei	*wzy*	gb/AF285971	C	WL-4127: TGAGGTTTCACGTTTCTC WL-3238: AATAATCCCTAACTGAGCC	817	OA-151: GGCACTGGATTAGGTGTTGCAAATTATGTAAAGGCTA OA-614: TGATTATCGTCGAGTTGAGTTAGTATTTATTGGGGTTGAT OA-615: TGGTTCTTTTGGTGTCATTTTGCATTGGTAAAGATGAGCT
Shigella flexneri types 1–5	*wzx*	gb/AY900451	C	WL-4004: CACTTGTTGGGTATGCTGG WL-3232: CCGGCAAACAGATTAGAAA	782	OA-145: GGGCAGTGTTTCCAAGGTTAAGTAACATCAAAGACTTTAA OA-616: CGGTTGGATGACAGTAAGTAATATCATAAGTCCGGTCATG
S. flexneri type 6	*wzx*	Laboratory stock	A	WL-5161: TTAAGAGCGATCATTTC WL-5135: CCATCCAAGCGGACATT	739	OA-1731: CGTCTGAAGGTATACAACGCTTAACGATATTACCAAGTG OA-1732: CATGGAGGATTACTAAGTGCTATATATGGTTTAGTTATGGCA OA-1733: TATTTCAGCTACCATACCCTTATACTTGTTAAACCAAGTCTG OA-1739: GGTTTAGTTATGGCAAGAGTCTTATCACTTATAGTGACCT
Shigella boydii type 1	*wzx*	gb/AY630255	A	WL-3369: GGCGGTAATTTACTTGAC WL-3370: AACACATTCTGCGTAACTC	641	OA-58: GCAATTACTGATTCATTAAGATTTGGAGGCGTGTTTTCTG OA-1604: CGGTGCAAAGTTAATTCCGCTAACGATAATATGTTGTTTT OA-1605: GGCGGTAATTTACTTGACAAATCAGATATGGCTAATGTTG OA-1710: ATCAAGGTTTCTTTATATTCTTTCATGACAGAATGCCTG
S. boydii type 2	*wzx*	Laboratory stock	B	WL-2911: ATTAGGGTTTTGTCAAGCA WL-2912: CTCCGACTCCATATCTTATT	833	OA-62: TTGGCTACTTTAACGCTGCTAATACAATTAGAAATGCAGC OA-1607: AGCCCAGAGTCTAATGAATCCTATTTATCAAGCTGTTTAT OA-1608: TCAGGTTCATTACTCCTGTGCATACTTTCTCCAATAATAA
S. boydii type 3	*wzy*	Laboratory stock	B	WL-3375: GTTAGGTTTGTTTCAAGGTC WL-4782: TAGGAATGAGTATGAAAGA	1072	OA-161: ATGCCGAATGAAAATACAATGTTGTTAGGTGATGGG OA-1610: GAGTTGAAGGGAGAGTCATTTATATCATTTTCGGCAA OA-1611: TCAGTTTACAAGTCATTCGACAAATGATTTATCCAGCATG
S. boydii type 4	*wzx*	gb/AF402312	A	WL-3035: GACCTCCTTTATTGGTTGC WL-4783: TCAAAATTGATAATTGTATGC	862	OA-66: TTGGTGTTGGAACCTAAAAGTTTCGTTGTCTTTTCAACTG OA-1612: CATGCCTGTATTTAGTAAACTGTATCACCTGCATGGC OA-1613: CTGAAAGAATCGTGAAGGTGTTATCGATTACAACAACGC
S. boydii type 5	*wzx*	gb/AF402313	A	WL-3039: CATTGCTGTTTCTGGGTTT WL-3040: TACCTGTGATTGCCGTCTG	938	OA-70: TGGCACGTTGCTTTCATTATTTAGAGATACCTTATCACCA OA-71: ACATGGTGTTTTATTCTTGACATAGTATTCACGACAGCAT OA-1614: GCGATTGTATCTCCTTTTGTTTGGGAAGACACAA
S. boydii type 6	*orf**	gb/AF402314	A	WL-4036: AGAACATCAGGAAAACGC WL-4037: GCTATTGCTAATGATGCTGT	1568	OA-619: TCAATGATGAACGCCACATAGACGAAGCCC OA-1649: CCCAGATTGTTTGAATGGAATATACTCAATCACAGCATCA OA-1650: GCTGGATTAGATATTTATATACGTCATGTTTCTCACCCCA OA-1712: CGTCATGTTTCTCACCCCAGATTGTTTGAATGGAATATAC OA-1713: AGTCGATGATAGCTTTTCAACAGGTAAATCGATGAGGAA
S. boydii type 7	*wzx*	Laboratory stock	C	WL-3377: GTGTTGACTGCTGGATTTC	572	OA-76: ATTCCACTTCCATCAATCATATCGTTGATTATTCCTGCTG

Table 2. cont.

Serotype	Target gene†	O-antigen gene cluster sequence	PCR group	Primer sequences (5′→3′)	Product size (bp)	Probe sequences (5′→3′)
				WL-3378: CGATATGATTGATGGAAGTG		OA-77: CCCATTGATTTTGCATCAAATAAGTGGCTTTCTTAAAGGA
						OA-575: CTTGATACTGATAATGGCGTTGAATAAAGCACTAGTTCCA
S. boydii type 8	*wzy*	Laboratory stock	A	WL-2917: TGTTTGCTTTTAGCATCCC	937	OA-82: GCATCAGGGGATTAGGTATATCAGGTTCTGTTGCC
				WL-2918: GATAATATGTCCCATCCTTT		OA-83: GCATTATTCAGCTCTGTTATAATTGCGTCGCATTTAATCC
						OA-1615: TTGCGTCGCATTTAATCCAGCATAACAATGATAATGTGTT
S. boydii type 9	*wzy*	gb/AF402315	C	WL-3045: GCGTTGGTTGGTGAAAGAG	877	OA-86: AACTGAGTTCACTTATGGTTCGAGAACCTTTACTCCATTT
				WL-4073: TTCCCACAAATCAAACCA		OA-583: GGATTTTAATACAACTGAGTTCACTTATGGTTCGAGAACC
S. boydii type 10	*wbaM*	gb/AY693427	A	WL-4036: AGAACATCAGGAAAACGC	255	OA-1649: CCCAGATTGTTTGAATGGAATATACTCAATCACAGCATCA
				WL-4037: GCTATTGCTAATGATGCTGT		OA-1650: GCTGGATTAGATATTTATATACGTCATGTTTCTCACCCCA
						OA-1712: CGTCATGTTTCTCACCCCAGATTGTTTGAATGGAATATAC
						OA-1713: AGTCGATGATAGCTTTTCAACAGGTAAATCGATGAGGAA
S. boydii type 11	*wzy*	gb/AY529126	B	WL-4784: GATATTCTTATAGAGGTGGTT	738	OA-92: CTTTTCTCTCCCATACCTGTATTAGCTGCACGAATAAGTC
				WL-2920: GTGCTGGTATAGTTATTGG		OA-1616: TTTATCTGCATGATTTTACTCAAATTAGAGCTGGCTTAGC
						OA-1617: CTTTTGTTTTCGCAAATTTCATGGTTCCTTTTCTCTCCCA
						OA-1618: ATTTCATGGTTCCTTTTCTCTCCCATACCTGTATTAGCTG
S. boydii type 12	*wzy*	Laboratory stock	A	WL-3383: GACTTCAGATACCCTCACT	1006	OA-96: TGGTGAAGGTTATGATGGTATCTGGCTCCATAATATGTTA
				WL-3384: GGTACAGCTACACCCAAT		OA-589: ATCCATGGGGAGGTTTAATTGTTCATGGTGAAGGTTATGA
						OA-1700: GGTTTAATTGTTCATGGTGAAGGTTATGATGGTATCTGGC
						OA-1701: GGGTATCAGGTTACTATACTATGTTTGTATGGCTGTTCA
S. boydii type 13	*wzy*	gb/AY369140	C	WL-61: AAAGATTGGTAGCGTCGG	847	OA-100: TTATCTTGAGTTTAGTCGAATGCAACGAGTAGTCGCG
				WL-62: TGAAGCCCTGGTAAAGTGC		OA-590: TTATTCTCATATTCGAATGTTACTTTTACCAGCGCAACTG
						OA-591: GCTGCGGGGAATAGAAATAAGAATCGTCTGTATTACTTTA
S. boydii type 14	*wzy*	Laboratory stock	A	WL-2317: CTATCAGCGGTGATCTTTAC	484	OA-165: ACTATTGCTGAGTTGGAATTATCTTTCCTCTACTGGTC
				WL-2318: TGTTATCTCCTCCTGTTCCT		OA-1619: GGAACAGGAGGAGATAACAGTTCAAATGAACGTTT
						OA-1702: TTCAGGATTAATTGATTTCTTATCAGGAACAGGAGGAGAT
S. boydii type 15	*wzy*	Laboratory stock	B	WL-2923: GAAGAGTATGGGATGGTAG	1218	OA-118: GCTTAAAGTCCTTATGGGAAGCTTTGTTCTGTATACTGTT
				WL-2924: CAACAACACTGTATAATAACCA		OA-119: GTTAAGCTTGAGAATGCATACGTGCTTGTAGTAATGATGG
						OA-1708: CGAGACGGTATAAATTGAAGTTTAGTGTGCTTAAAGTCC
						OA-1709: TTCCTTATGGGAAGCTTTGTTCTGTATACTGTTCCTTTA
S. boydii type 16	*wfaU*	gb/DQ371800	C	WL-395: CCATACGGATAATGTTGAG	698	OA-318: TTGTATGTAATCGCAAATACTCGCAGTACATGTGGA
				WL-396: TCTTTGTCTTCTCGGCTA		OA-319: CCTGAGCGATTAGAACAACTGTATCTCGAGTCAAGA
						OA-596: GACATGGGAGTGGAATGAAAACTAAAATTGCTGAGGCTTT
						OA-597: GTGAGATTGTTGATTGCTTTAGCCGAGAAGACAAAGATAT

Table 2. cont.

Serotype	Target gene†	O-antigen gene cluster sequence	PCR group	Primer sequences (5′→3′)	Product size (bp)	Probe sequences (5′→3′)
S. boydii type 17	*wzy*	Laboratory stock	B	WL-365: TACCGTTCTCGCCAATAC	1116	OA-540: TGGTCTAGATATCCGGCATTTTCATTTGAGCCATTT
				WL-4035: AAAGATAGTAAAATGGCTCAA		OA-1654: GGAGTGATAGCAGCATTAATTATACTATTTGCTTCTGTGC
S. boydii type 18	*wzx*	gb/AY948196	C	WL-955: CAGGGCACAAACTATCT	1096	OA-313: TTTGTCGTGTCTATTGGTATTGTATTCGGGCAATGG
				WL-958: TAATTCGGAATGTGCT		OA-599: TGGAGTTCCATTATCCGAGTGGTATTTTGGTAACAATAAT
						OA-600: TACTCATTAGGCTTGCTATTTACTGGGCTATTATCAGTTT
Shigella dysenteriae type 1	*wzy*	gb/L07293	B	WL-3051: TATTTCGTACCATTAAGTGC	925	OA-116: CCAAAACTGATATTTTGCCTTGGATCAACATTGGTGTATG
				WL-3052: ACAACAAATAATGCTCCTG		OA-117: TGGAGTGTTTGTATTTGGACTGTTCGCTTTATCAAGTATT
						OA-1622: GACCCTGAAGGCAAGATTATTTCAGCAATATTTGATACGT
						OA-1623: TCAAGTATTATCTTGGGTAAGTTCTCTTCAGACCCTGAAG
S. dysenteriae type 2	*wzx*	Laboratory stock	B	WL-5775: TTATTGGCCTTCTGACTT	635	OA_1999: ACCAGTATTGTTAAAATCTGTTTCTGTAGTGGCTATAGGA
				WL-5776: ACTATGAGCAGCTATTCC		OA_2000: TGGAGGTTCATATGCAGTGCTAATACAAATACCAGTATTG
						OA_2001: GTTAGGTCATTACATTGGAGGTTCATATGCAGTGCTAATA
S. dysenteriae type 3	*wzx*	Laboratory stock	B	WL-3391: GCGGCAGACTATGGCTTAT	947	OA-122: CGATGGTGTTAGCATACGAGAAAATGTCATTTTACGCATA
				WL-3392: GCCACAACAATGACCAACC		OA-1624: CTGTAATTCTATGTGCGGATGAGTTGTTAAAGATATGGCT
						OA-1625: ATACACTTACGCAATCATTAATTACAGCATTAAGGCCTCA
						OA-1626: CCAGATTATACAGTTGAATTTGTGAGGTTGGTCATTGTTG
S. dysenteriae type 4	*wzy*	Laboratory stock	C	WL-4131: TTTCTGCTTATTCATTATTG	946	OA-127: TCGCATTGCTTGGTATTAGAGCTGGTAGTGTTAATT
				WL-4125: ACTACCAGCTCTAATACC		OA-128: GCAGCTTCTCTTTACATACGTGCATTATCAGAAACGGGAA
						OA-602: ACTAGTCTAGATAAAATGGCTAGCTCAGACAATCTTTCTG
						OA-603: TGGATATTTAATTCGGCTATTCTCTATAGCTAGTGAGCCT
S. dysenteriae type 5	*wzy*	Laboratory stock	B	WL-1178: AACGAGTATGGTGGATTG	840	OA-169: GTAAGCGAGGTCAATAAAATGCTAATGACTGGTGAGT
				WL-1179: CGCATTTCGATAAATCAC		OA-170: GCTATGGTACCTATAGTTCTTATCTTTTGGGACCTCTT
						OA-1627: GCTCTAAAGCTTCGCTCATATTAATTTTGATGAGTGGTGT
						OA-1628: GGACCTCTTGATTTAGCTGTTCAATTAGCAGAAGAACTAT
S. dysenteriae type 6	*wzx*	Laboratory stock	B	WL-2933: TAGATGTGCTGCCTTTGAT	639	OA-171: ATCCGTATCAACATTGGGTATTTATTCGCTAGCTATGCA
				WL-2934: AATACTCTAGGGCCTCATG		OA-172: CTGTCCCTTGGATATGTTCTTAACGATAGATACCATGAG
						OA-1629: ATAGCATGTGTATTTTCATCAGTAGTTATGCCTCTGTCCC
S. dysenteriae type 7	*wzy*	gb/AY380835	A	WL-3389: CAATGTTGATAGGATGTTC	1079	OA-131: TCGGCCATTTAGCGTTGATATATGAAAATGCAGATGATAT
				WL-3390: CAGACTTATTATACATGCG		OA-132: TCCTTATGTTCTCTATTGCTTCTTTTCTCCATTTTGCCTG

Table 2. cont.

Serotype	Target gene†	O-antigen gene cluster sequence	PCR group	Primer sequences (5′→3′)	Product size (bp)	Probe sequences (5′→3′)
						OA-1631: GGCATAAACTATCTGTTAAATAGATTTCAACATGTCGGCC
						OA-1632: TGCTGGTTCCTTATTATGTAGTTAGTAGATTTGCTGGTAA
S. dysenteriae type 8	*wzy*	Laboratory stock	C	WL-2939: TTCCCTCTTGTTGTATTGA	1106	OA-135: GAGGGTGTGTATGGTATGATTGATTACATATTGGAGGC
				WL-2940: ACCTTTATCAATTGCCTCC		OA-136: ACTTCAGTTTCTGGGAACGATAAATTCACACGTTTGC
						OA-606: TTCCTAAATAATCATCCATTTACTGAGGGTGTGTATGGTA
						OA-607: CGCCAATTTATTCTGTGCTATTATCGAAATTCACTTCAGT
S. dysenteriae type 9	*wzy*	Laboratory stock	A	WL-2335: GGGCTTATGAAGAGTGTA	982	OA-177: CGTTATTTGGGTTGGGAAAATATCTACTTTCCGATGGTT
				WL-2334: TTCCACCGACTCCCATTT		OA-178: AGATCTGATATTGGATACTATAGACAAATGGGAGTCGGT
						OA-1634: GCGCACTTCCTTTAGTAATTATTGCCGTCCATAACATAAT
S. dysenteriae type 10	*wzy*	Laboratory stock	C	WL-2343: CGCTGTTTCTATATTAATTG	706	OA-181: ATGGCACTGATAGTAGCGATAAAACTATTTTGATCGGG
				WL-2344: AATTGAAGTGACCAGATAAC		OA-610: CGTGCATTAATTGCTATTGTCGTAATTTCTTTCATTGTGG
						OA-611: GGAACACCTTCTTGGGATTATTTTACGCAACCACTTATTA
S. dysenteriae type 11	*wzy*	Laboratory stock	B	WL-2115: AATAACCGCCGTTCTAAT	767	OA-1640: AACACAAACAAGGTGTGAGTCGTTTCATATTCTTCTT
				WL-2114: CGAAGTAATATCCGTTCATA		OA-1642: TTTATCTTGGTAGGTTTGGTACTATTAGCACCAGGCT
						OA-1716: TGATATTCATGGACGTTTATCTTGGTAGGTTTGGTAC
S. dysenteriae type 12	*wzy*	Laboratory stock	B	WL-2945: GTTATTTACCTACGCCTCA	1045	OA-143: TCAGGCCAATTATAATGATGATTTAGGTGGCATTGGTATT
				WL-2946: AGAAATACCAATGCCACCT		OA-144: GTACCCATTTATTTTATTGCGCGATTTAGAGTCCGTAGTT
						OA-1643: AATCTTAATATTCTATTGTGGATGTCTGTTGGATGTGTGC
						OA-1644: TAAGTTTATTGCTATTCCTGGGTAGTGGATTAAATGCACT
S. dysenteriae type 13	*wzx*	Laboratory stock	A	WL-2947: TTTCAGCCAATTTTACTAAT	1420	OA-1645: AGGCTCTGTAGTATTGTTCAAAACCACTACATTAGGTTAT
				WL-2324: GGAAATGTAGATAGAAGTCC		OA-1646: TTCTTTCACCATTAATATGTATCGTAACGCTTCCACTGAT
						OA-1704: GCCTGCACTCATAAATTTCTTTCTTGCAATCGTTGGAATT
						OA-1705: TGCAATCGTTGGAATTATAGGCTCTGTAGTATTGTTCAAAA
	16S rRNA gene	gi/1255975	A, B, C	WL-1319: GACGGGTGAGTAATGTCTGG	1251	OA-153: CGGGAACTCAAAGGAGACTGCCAGTGATAA‡
				WL-1320: ATCCACGATTACTAGCGATTCC		OA-381: CGGGAACTCAAAGGAGACTGCCAGTGATAAACTGGAG‡
						WL-4006: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT‡
						OA-Cy3: TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-Cy3‡

†*orf** (mutated *wbaM*), *wbaM*, *wfaU* and glycosyltransferase genes.

‡OA-153 and OA-381, positive-control probes; WL-4006, negative-control probe; OA-Cy3, positional reference and printing control probe.

OA-62	OA-62	OA-1607	OA-1607	OA-1608	OA-1608	OA-Cy3	OA-Cy3	OA-58	OA-58	OA-1604	OA-1604	OA-1710	OA-1710	OA-1605	OA-1605
OA-161	OA-161	OA-1610	OA-1610	OA-1611	OA-1611	OA-381	OA-381	OA-619	OA-619	OA-1649	OA-1649	OA-1650	OA-1650	OA-1712	OA-1712
OA-66	OA-66	OA-1612	OA-1612	OA-1613	OA-1613	WL-4006	WL-4006	OA-86	OA-86	OA-583	OA-583	OA-540	OA-540	OA-1654	OA-1654
OA-70	OA-70	OA-71	OA-71	OA-1614	OA-1614	OA-153	OA-153	OA-92	OA-92	OA-1616	OA-1616	OA-1617	OA-1617	OA-1618	OA-1618
OA-76	OA-76	OA-77	OA-77	OA-575	OA-575	50%DMSO	50%DMSO	OA-96	OA-96	OA-589	OA-589	OA-1700	OA-1700	OA-1701	OA-1701
OA-82	OA-82	OA-83	OA-83	OA-1615	OA-1615	OA-381	OA-381	OA-118	OA-118	OA-119	OA-119	OA-1708	OA-1708	OA-1709	OA-1709
OA-100	OA-100	OA-590	OA-590	OA-591	OA-591	50%DMSO	50%DMSO	OA-318	OA-318	OA-319	OA-319	OA-596	OA-596	OA-597	OA-597
OA-165	OA-165	OA-1619	OA-1619	OA-1702	OA-1702	OA-153	OA-153	OA-1731	OA-1731	OA-1732	OA-1732	OA-1733	OA-1733	OA-1739	OA-1739
OA-313	OA-313	OA-599	OA-599	OA-600	OA-600	50%DMSO	50%DMSO	OA-116	OA-116	OA-117	OA-117	OA-1622	OA-1622	OA-1623	OA-1623
OA-145	OA-145	OA-616	OA-616	50%DMSO	50%DMSO	OA-381	OA-381	OA-1713	OA-1713	OA_1999	OA_1999	OA_2000	OA_2000	OA_2001	OA_2001
OA-171	OA-171	OA-172	OA-172	OA-1629	OA-1629	WL-4006	WL-4006	OA-122	OA-122	OA-1624	OA-1624	OA-1625	OA-1625	OA-1626	OA-1626
OA-177	OA-177	OA-178	OA-178	OA-1634	OA-1634	OA-153	OA-153	OA-127	OA-127	OA-128	OA-128	OA-602	OA-602	OA-603	OA-603
OA-181	OA-181	OA-610	OA-610	OA-611	OA-611	50%DMSO	50%DMSO	OA-169	OA-169	OA-170	OA-170	OA-1627	OA-1627	OA-1628	OA-1628
OA-1640	OA-1640	OA-1642	OA-1642	OA-1716	OA-1716	OA-381	OA-381	OA-131	OA-131	OA-132	OA-132	OA-1631	OA-1631	OA-1632	OA-1632
OA-143	OA-143	OA-144	OA-144	OA-1643	OA-1643	OA-1644	OA-1644	OA-135	OA-135	OA-136	OA-136	OA-606	OA-606	OA-607	OA-607
OA-151	OA-151	OA-614	OA-614	OA-615	OA-615	OA-Cy3	OA-Cy3	OA-1645	OA-1645	OA-1646	OA-1646	OA-1704	OA-1704	OA-1705	OA-1705

Fig. 1. Probe positions on the slide. OA-153 and OA-381 are the positive-control probes based on the 16S rRNA gene of *Shigella*. WL-4006 is the negative-control probe. OA-Cy3 is the positional reference and printing control probe.

database of all O-antigen gene clusters of *Shigella* and *E. coli*. For positive controls, two probes based on the 16S rRNA genes of *Shigella* (Li *et al.*, 2006) were used to detect DNA of *Shigella* or closely related organisms. A probe containing poly$(T)_{40}$ oligonucleotides was used as the negative control. A probe labelled at its 3′-end with Cy3 was used as the positional reference and printing control. Each probe was 5′-amino-modified and a poly$(T)_{10}$ tail was added at the 5′-end. All probes are listed in Table 2.

Microarray preparation. Microarrays were spotted on aldehyde-group-modified glass slides (CEL) using SpotArray72 (Perkin-Elmer). Probes were dissolved in 50 % DMSO and spotted in duplicate on the slide. The slides were dried for 24 h at room temperature, scanned at 532 nm to check the quality of spotting and stored at room temperature in the dark. Each slide consisted of six microarrays framed with a 16 μl Geneframe (Beijing Capital Biochip), which constituted individual reaction chambers. One of the six microarrays was tested with the positive-control standard of *Shigella dysenteriae* type 8 genomic DNA (100 ng μl^{-1}) to ensure that the reagents were effective. Another was tested with the negative-control standard of sterile deionized water to show that the reagents were uncontaminated. The other four were used to analyse samples. A schematic diagram of the probe positions on the microarray is shown in Fig. 1.

Hybridization process. Fifteen microlitres of labelled PCR product was baked for about 1.5 h at 65 °C until dry, diluted in 16 μl hybridization buffer (25 % formamide, 0.1 % SDS, 6 × SSPE). After heating at 98 °C for 5 min, the hybridization mixture was applied to the slide. Hybridization was carried out in a chamber at 40 °C for 16 h in a temperature-controlled water bath in the dark. After hybridization, the Geneframes were removed and the slide was washed with 1 × SSC and 0.1 % SDS for 3 min, followed by a wash with 0.05 × SSC for 3 min and finally with 95 % ethanol for 1.5 min. The slide was dried under a gentle air stream before it was scanned.

Data acquisition and automated analysis. Microarray slides were scanned at 532 nm using a GenePix Personal 4100A microarray scanner (Axon Instruments) with the following parameters: photomultiplier tube gain, 650; pixel size, 5 μm. Two files were generated with GenePix Pro 6.0, one for the images saved as TIF and the other for the signal intensity saved as GPR. The signal-to-noise ratio was calculated for each spot using the Bactarray Analyser 1.0, developed in house, with the threshold set at 3.0. For each O-antigen form, two to five probes were used, and each probe was printed in duplicate to eliminate possible physical defects in the glass slide. A serotype was confirmed and reported when the following conditions were met: (i) the positive standard, the negative standard, the two positive-control probes, the negative-control probe and the printing-control probe all provided the expected signals; (ii) more than half of all of the probes of a particular serotype generated positive signals above the signal-to-noise ratio threshold.

RESULTS AND DISCUSSION

Target gene selection and probe design

The different O-antigen forms are almost entirely due to genetic variations in their respective O-antigen gene clusters (*wzx*, *wzy* and glycosyltransferase genes). We utilized O-serotype-specific genes used in a previous study to identify 10 serotypes of *Shigella*, including *Shigella sonnei*, *Shigella flexneri* type 2a, *Shigella boydii* types 7, 9, 13, 16 and 18, and *Shigella dysenteriae* types 4, 8 and 10 (Li *et al.*, 2006), for the strains studied here, and identified specific gene cluster sequences for *Shigella boydii* types 1, 4, 5, 6, 10 and 11, and *Shigella dysenteriae* types 1 and 7 obtained from GenBank (accession nos AY630255, AF402312, AF402313, AF402314, AY693427, AY529126, L07293 and AY380835, respectively). The sequences of the other 16 serotypes were extracted from our laboratory data. For each serotype, a specific gene was chosen as the target gene for the downstream amplification and labelling (Table 2), and two to five probes were designed for each target gene.

Multiplex PCR to label the target genes

Multiplex PCR was used to amplify and label the target genes in three groups. Group A targeted *Shigella boydii* types 1, 4, 5, 6, 8, 10, 12 and 14, *Shigella dysenteriae* types 7, 9 and 13, and *Shigella flexneri* type 6; group B targeted *Shigella boydii* types 2, 3, 11, 15 and 17, and *Shigella dysenteriae* types 1, 2, 3, 5, 6, 11 and 12; group C targeted *Shigella boydii* types 7, 9, 13, 16 and 18, *Shigella dysenteriae* types 4, 8 and 10, *Shigella flexneri* types 1–5 and *Shigella*

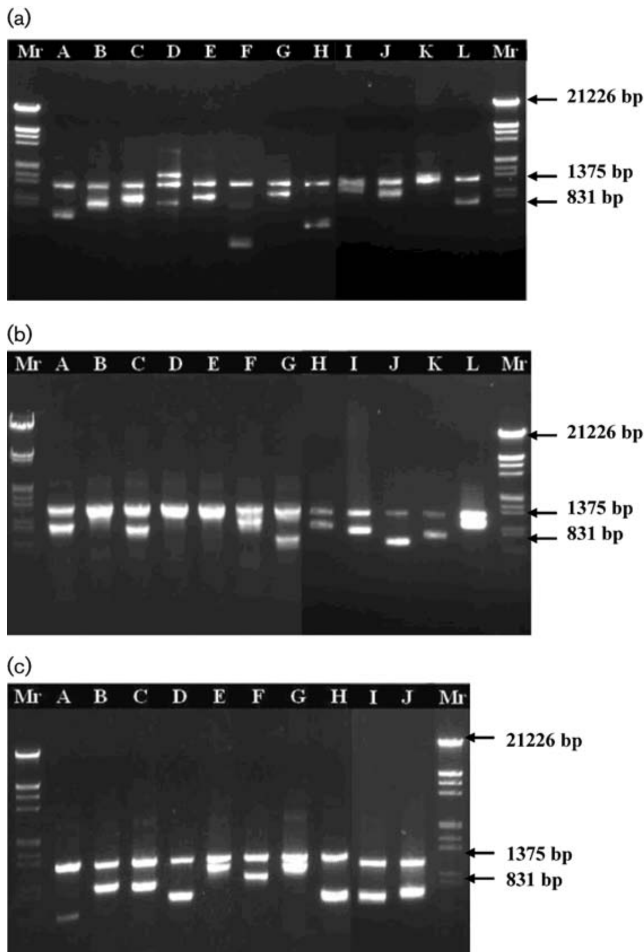

Fig. 2. Agarose gel electrophoresis of multiplex PCR products. (a) Group A. Lanes: Mr, molecular mass standards (lambda DNA/*Eco*RI+*Hin*dIII marker; Sangon); A–H, *Shigella boydii* types 1, 4, 5, 6, 8, 10, 12 and 14; I–K, *Shigella dysenteriae* types 7, 9 and 13; L, *Shigella flexneri* type 6. (b) Group B. Lanes: Mr, molecular mass standards; A–E, *Shigella boydii* types 2, 3, 11, 15 and 17; F–L, *Shigella dysenteriae* types 1, 2, 3, 5, 6, 11, and 12. (c) Group C. Lanes Mr, molecular mass standards; A–E, *Shigella boydii* types 7, 9, 13, 16 and 18; F–H, *Shigella dysenteriae* types 4, 8 and 10; I, *Shigella flexneri* type 2a; J, *Shigella sonnei*. Two bands were generated from the PCRs: one from the 16S rRNA gene (1.2 kb) and the other from the gene specific to the individual serotype.

sonnei. For each of the representative strains of all distinct O-antigen forms of *Shigella*, two bands were generated by PCR, one for the 16S rRNA gene and one for the O-serotype-specific gene (Fig. 2). The PCR products ranged from 255 to 1568 bp (Table 2). The dye Cy3–dUTP was integrated into the PCR amplicons.

Optimal probe concentration

Probes OA-1731, OA-1732, OA-1733 and OA-1739 targeting *Shigella flexneri* type 6 and probes OA-58, OA-1604, OA-1605 and OA-1710 targeting *Shigella boydii* type 1 were selected randomly and printed on slides at concentrations of 1, 3, 10, 30, 50, 80 and 100 μM. The slide was then processed to hybridize with labelled PCR fragments of the corresponding *wzx* genes. The fluorescence intensity increased steadily in the range of 1–10 μM probe and reached peak intensity at 10 μM; the signal level decreased as the probe concentration approached 100 μM (Fig. 3). At high concentrations, steric hindrance by an excess of probe molecules on the microarray surface prevents incoming target DNA molecules from hybridization, thereby reducing the fluorescence signal (Schena, 2003). Therefore, 10 μM was used as the optimal probe concentration in this study.

Specificity of oligonucleotide probes

The DNA microarray was tested on all strains of the panel of O-serotype strains of *Shigella* and *E. coli* in addition to those listed in Table 1. Each of the representative strains of all 34 distinct O-antigen forms of *Shigella* consistently hybridized to their corresponding probes. The results for selected strains are shown in Fig. 4 on arrays 1–5. Representative strains of *Shigella flexneri* types 1–5 hybridized specifically with their corresponding probes (Table 1), a result in line with the fact that they have a common basic O-antigen structure and a common O-antigen gene cluster (Cheah *et al.*, 1991).

Many *E. coli* O antigens cross-react serologically with *Shigella* O antigens. Recently, we elucidated the relationships between the structures and genetics of the O antigens of *Shigella* and *E. coli* (Liu *et al.*, 2008) and showed that 21 *Shigella* type strains share identical or closely related O-antigen gene clusters with *E. coli*. As expected, 23 representative strains of *E. coli* O serotypes (Table 3) gave the same hybridization results as the 21 corresponding *Shigella* serotypes (Table 3). All of the 86 clinical isolates of *Shigella* were successfully identified. No strains of other bacterial species (Table 1) hybridized with the serotype-specific probes (Fig. 4, arrays 6 and 7). *E. coli*, *Salmonella enterica* and *V. cholerae* have a close relationship with *Shigella*, and the 16S rRNA gene primers can work with their genomic DNA, so both the positional reference control and positive-control probes gave positive results. In contrast, *B. cereus* and *Staphylococcus aureus* have a distant relationship with *Shigella*; hence only the positional reference control probe gave a positive result.

Differentiation between groups of *Shigella* and *E. coli* with identical or closely related O antigen

As with traditional serotyping, the O-serotype-specific gene microarray was unable to differentiate the 21 O antigens that are identical or closely related in some *Shigella* and *E. coli* groups (Table 3). *Shigella* strains are non-motile; the *E. coli* inactive group, which includes EIEC, is very similar to *Shigella* in biochemical and physiological properties to the extent that Lan & Reeves (2002) suggested that the latter

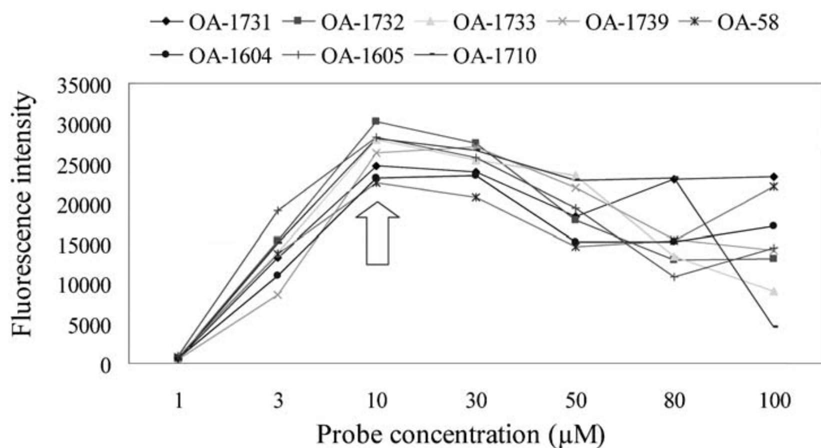

Fig. 3. Fluorescence intensity of different probe concentrations. Scanned fluorescence intensities were plotted as a function of probe concentration. The probe concentration that gave the strongest fluorescent signal (arrow) defined the optimal probe density.

should be considered a member of the extremely diverse species *E. coli*. Other shared properties are the specificity of the invasion plasmid antigen gene (*ipaH*) for *Shigella* and EIEC (Venkatesan *et al.*, 1989) and the absence of lysine decarboxylase activity, present in 90 % of *E. coli* strains, among strains of *Shigella* and EIEC strains (Maurelli *et al.*, 1998), which has been used in the diagnostic laboratory to differentiate *E. coli* from *Shigella*. Nevertheless, most EIEC strains contain the *cadA* gene encoding lysine decarboxylase, whilst *Shigella* strains lack it (Casalino *et al.*, 2003; Tao *et al.*, 2005).

In order to differentiate O antigen identical or closely related in *Shigella* and *E. coli*, primers were designed for

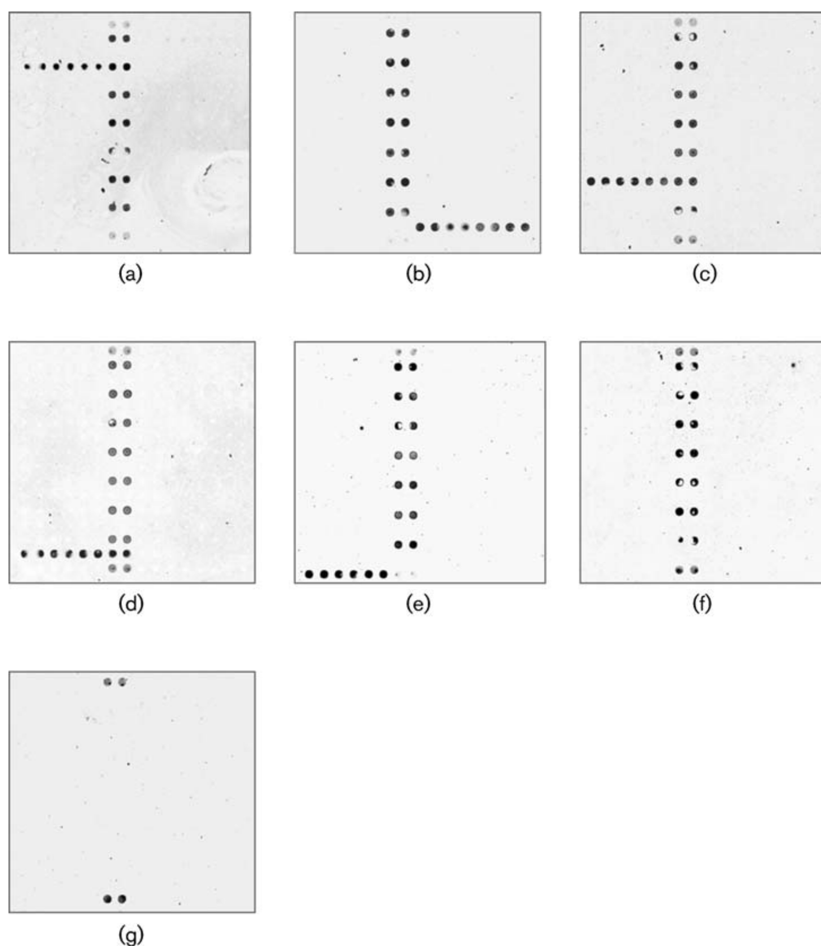

Fig. 4. Microarray hybridization patterns of different pathogens. (a) *Shigella boydii* type 5; (b) *Shigella dysenteriae* type 8; (c) *Shigella dysenteriae* type 9; (d) *Shigella dysenteriae* type 12; (e) *Shigella sonnei*; (f) *E. coli*, *Salmonella enterica* and *V. cholerae*; (g) *B. cereus* and *Staphylococcus aureus*. The two lines of dots located in the middle of the array are positive-control probes; the two dots in the first and last lines are positional reference control probes and the others are serotype-specific probes. The positional reference control probes are always positive. The positive-control probes may or may not give a positive result depending on whether the test bacteria are closely related to *Shigella*.

Table 3. Differentiation between groups of O antigen that are identical or closely related in *Shigella* and *E. coli*

Group	Strains	Result by:		
		DNA microarray	Amplification of *ipaH*	Amplification of *cadA*
1	*Shigella boydii* type 1	*S. boydii* type 1	+	−
	E. coli O149, ETEC*		−	+
2	*S. boydii* type 3	*S. boydii* type 3	+	−
	E. coli O167, EIEC		−	+
3	*S. boydii* type 4	*S. boydii* type 4	+	−
	E. coli O53		−	+
4	*S. boydii* type 5	*S. boydii* type 5	+	−
	E. coli O79		−	+
5	*S. boydii* type 8	*S. boydii* type 8	+	−
	E. coli O143, EIEC		+	+
6	*S. boydii* type 11	*S. boydii* type 11	+	+
	E. coli O105		−	+
7	*S. boydii* type 14	*S. boydii* type 14	+	−
	E. coli O32		−	+
8	*S. boydii* type 15	*S. boydii* type 15	+	−
	E. coli O112ab, EIEC		+	+
9	*Shigella dysenteriae* type 1	*S. dysenteriae* type 1	+	+
	E. coli O148, ETEC*		−	+
10	*S. dysenteriae* type 2	*S. dysenteriae* type 2	+	−
	E. coli O112ac, EIEC		+	+
11	*S. dysenteriae* type 3	*S. dysenteriae* type 3	+	−
	E. coli O124, EIEC		−	+
12	*S. dysenteriae* type 4	*S. dysenteriae* type 4	+	−
	E. coli O168		−	+
13	*S. dysenteriae* type 5	*S. dysenteriae* type 5	+	−
	E. coli O58		−	+
14	*S. dysenteriae* type 6	*S. dysenteriae* type 6	+	−
	E. coli O130		−	+
15	*S. dysenteriae* type 7	*S. dysenteriae* type 7	+	−
	E. coli O121		−	+
16	*S. dysenteriae* type 9	*S. dysenteriae* type 9	+	−
	E. coli O40		−	+
17	*S. dysenteriae* type 11	*S. dysenteriae* type 11	+	−
	E. coli O29		−	+
18	*S. dysenteriae* type 12	*S. dysenteriae* type 12	+	−
	E. coli O152, EIEC		+	+
19	*S. dysenteriae* type 13	*S. dysenteriae* type 13	+	−
	E. coli O150		−	+
20	*Shigella flexneri* type 2a	*S. flexneri* types 1–5	+	−
	E. coli O13		−	+
	E. coli O129		−	+
	E. coli O135, EIEC		−	−
21	*S. flexneri* type 6	*S. flexneri* type 6	+	−
	E. coli O147		−	+

*ETEC, enterotoxigenic *E. coli*, which produces heat-labile enterotoxin, heat-stable enterotoxin, or both.

ipaH (*ipaH*-F, 5′-TGACCGCCTTTCCGATA-3′, and *ipaH*-R, 5′-TTCTCCAGCATCTCATA-3′) and *cadA* (*cadA*-F, 5′-TTCAAAAACATCGATAACGA-3′, and *cadA*-R, 5′-ACGGTATGCACCGTGAAT-3′). With the primers for *ipaH*, all 21 *Shigella* strains and four out of seven EIEC strains (Table 3) gave PCR products of the expected size (584 bp), whilst other *E. coli* strains failed to generate an amplicon. Moreover, bands of the predicted size (669 bp) were observed with *Shigella boydii* type 11, *Shigella dysenteriae* type 1 and 22/23 tested *E. coli* strains with the primer pair for *cadA*. The combined results of amplification of these genes

indicated that all 21 groups of O antigen identical or closely related in *Shigella* and *E. coli* could be differentiated. Therefore, when the microarray result is positive for one of the 21 *Shigella* cross-reaction serotypes listed in Table 3, further PCR assays based on *ipaH* and *cadA* are recommended.

Sensitivity of detection with genomic DNA

Serial dilutions of genomic DNA of *Shigella* strains G1222 (*Shigella dysenteriae* type 7, multiplex PCR group A), G1191 (*Shigella boydii* type 11, multiplex PCR group B) and G1221 (*Shigella dysenteriae* type 8, multiplex PCR group C) to give 1000, 100, 50, 10 and 1 ng DNA were used to test the sensitivity of the microarray. Strong positive signals were generated at 10 ng or above (not shown), and 50 ng was chosen as the standard DNA amount for the study.

Sensitivity of detection with milk powder

DNA extracts from 10-fold serial dilutions of *Shigella* strains G1222, G1191 and G1221 in the range of 10^1–10^6 c.f.u. ml^{-1} were labelled and hybridized to the microarray. All specific probes correctly detected 10^4 c.f.u. ml^{-1} or greater. In addition, all *Shigella* strains were detected following 6 h enrichment with an initial cell count of 1 c.f.u. in 25 g milk powder. This level of sensitivity is in compliance with the food standards of the Codex Alimentarius Commission (CAC) (Principles for the Establishment and Application of Microbiological Criteria for Foods; CAC/GL 21-1997).

Double-blind test

A double-blind test was performed to verify the reliability and specificity of the microarray. A total of 36 *Shigella* strains and nine strains of other bacterial species were used to hybridize to the microarray without disclosure of their identity during testing. The results for the *Shigella* strains, which included one strain from each of *Shigella boydii* types 1–18, *Shigella dysenteriae* types 1–13, *Shigella flexneri* types 1a, 2a, 3 and 6, and *Shigella sonnei*, were consistent with traditional phenotyping and serotyping methods. As before, none of the control species reacted in the microarray.

The microarray has many advantages over traditional bacterial culture and serotyping methods: (i) it is rapid and allows a high throughput of isolates, with processing time within a working day, whilst traditional identification of *Shigella* in the clinical laboratory may take 3–5 days; (ii) it is able to detect O-rough strains, which are due to mutations in the O-antigen gene cluster; and (ii) cross-reactions are uncommon. However, as with traditional serotyping, the microarray has its limits in distinguishing *Shigella* and *E. coli* strains that share the same O-antigen structures. This resulted in 23 *E. coli* O serotypes giving the same hybridization results as the 21 corresponding *Shigella* types; however, PCR of the *ipaH* and *cadA* genes was able to distinguish between the two groups of strains.

With the exception of *Shigella flexneri* type 6, all of the other *Shigella flexneri* serotypes share an identical O-antigen basic chain structure, and the same variations in *E. coli* or *Salmonella enterica* would not be given separate status. These variations result from phage modifications within the genome. As a result, the specific probes OA-145 and OA-616 (Table 2), based on the *wzx* gene of *Shigella flexneri* type 2a, can hybridize with all *Shigella flexneri* type 1–5 strains to detect them as a group. Further resolution within this group still requires traditional serotyping.

Shigellosis is a global human health problem and no licensed vaccines are currently available to protect against infection (NIAID, 2003). A rapid and comprehensive means of detecting all *Shigella* serotypes is important for the microbiological monitoring of food to protect consumer health. The DNA microarray method described here is specific and sensitive, and offers a promising tool for wide applications in food safety, basic microbiological research and epidemiological surveillance.

ACKNOWLEDGEMENTS

This work was supported by China National 863 Program (2006AA020703, 2006AA06Z409), China NSFC General Program (20536040) and NSFC Key Program (30530010).

REFERENCES

Ahmed, F., Clemens, J. D., Rao, M. R., Ansaruzzaman, M. & Haque, E. (1997). Epidemiology of shigellosis among children exposed to cases of *Shigella* dysentery: a multivariate assessment. *Am J Trop Med Hyg* **56**, 258–264.

Bennish, M. L. (1991). Potentially lethal complications of shigellosis. *Rev Infect Dis* **13** (*Suppl. 4*), S319–S324.

Beutin, L., Tao, J., Feng, L., Krause, G., Zimmermann, S., Gleier, K., Xia, Q. & Wang, L. (2005). Sequence analysis of the *Escherichia coli* O15 antigen gene cluster and development of a PCR assay for rapid detection of intestinal and extraintestinal pathogenic *E. coli* O15 strains. *J Clin Microbiol* **43**, 703–710.

Bopp, C. A., Brenner, F. W., Fields, P. I., Wells, J. G. & Strockbine, N. A. (2003). *Escherichia*, *Shigella*, and *Salmonella*. In *Manual of Clinical Microbiology*, 8th edn, pp. 654–671. Edited by P. R. Murray, E. J. Baron, J. H. Jorgensen, M. A. Pfaller & R. H. Yolken. Washington, DC: American Society for Microbiology.

Casalino, M., Latella, M. C., Prosseda, G. & Colonna, B. (2003). CadC is the preferential target of a convergent evolution driving enteroinvasive *Escherichia coli* toward a lysine decarboxylase-defective phenotype. *Infect Immun* **71**, 5472–5479.

Cheah, K. Y., Beger, D. W. & Manning, P. A. (1991). Molecular cloning and genetic analysis of the *rfb* region from *Shigella flexneri* type 6 in *Escherichia coli* K-12. *FEMS Microbiol Lett* **83**, 213–218.Medline

DebRoy, C., Fratamico, P. M., Roberts, E., Davis, M. A. & Liu, Y. (2005). Development of PCR assays targeting genes in O-antigen gene

clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. *Appl Environ Microbiol* **71**, 4919–4924.

DuPont, H. L., Levine, M. M., Hornick, R. B. & Formal, S. B. (1989). Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* **159**, 1126–1128.

Feng, L., Tao, J., Guo, H., Xu, J., Li, Y., Rezwan, F., Reeves, P. & Wang, L. (2004a). Structure of the *Shigella dysenteriae* 7 O antigen gene cluster and identification of its antigen specific genes. *Microb Pathog* **36**, 109–115.

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

Fratamico, P. M., Briggs, C. E., Needle, D., Chen, C. Y. & DebRoy, C. (2003). Sequence of the *Escherichia coli* O121 O-antigen gene cluster and detection of enterohemorrhagic *E. coli* O121 by PCR amplification of the *wzx* and *wzy* genes. *J Clin Microbiol* **41**, 3379–3383.

Johnson, J. R. (2000). *Shigella* and *Escherichia coli* at the crossroads: machiavellian masqueraders or taxonomic treachery? *J Med Microbiol* **49**, 583–585.

Koterski, J. F., Nahvi, M., Venkatesan, M. M. & Haimovich, B. (2005). Virulent *Shigella flexneri* causes damage to mitochondria and triggers necrosis in infected human monocyte-derived macrophages. *Infect Immun* **73**, 504–513.

Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K. & Levine, M. M. (1999). Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* **77**, 651–666.

Lai, V., Wang, L. & Reeves, P. R. (1998). *Escherichia coli* clone Sonnei (*Shigella sonnei*) had a chromosomal O-antigen gene cluster prior to gaining its current plasmid-borne O-antigen genes. *J Bacteriol* **180**, 2983–2986.

Lan, R. & Reeves, P. R. (2002). *Escherichia coli* in disguise: molecular origins of *Shigella*. *Microbes Infect* **4**, 1125–1132.

Li, Y., Liu, D., Cao, B., Han, W., Liu, Y., Liu, F., Guo, X., Bastin, D. A., Feng, L. & Wang, L. (2006). Development of a serotype-specific DNA microarray for identification of some *Shigella* and pathogenic *Escherichia coli* strains. *J Clin Microbiol* **44**, 4376–4383.

Liu, B., Knirel, Y. A., Feng, L., Perepelov, A. V., Senchenkova, S. N., Wang, Q., Reeves, P. & Wang, L. (2008). Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* **32**, 627–653.

Maurelli, A. T., Fernandez, R. E., Bloch, C. A., Rode, C. K. & Fasano, A. (1998). "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**, 3943–3948.

NIAID (2003). *NIAID Biodefense Research Agenda for Category B and C Priority Pathogens*. NIH Publication No. 03-5315.

Panicker, G., Call, D. R., Krug, M. J. & Bej, A. K. (2004). Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. *Appl Environ Microbiol* **70**, 7436–7444.

Rouillard, J. M., Zuker, M. & Gulari, E. (2003). OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res* **31**, 3057–3062.

Schena, M. (2003). Microarray surfaces. In *Microarray Analysis*, pp. 95–120. New York: Wiley.

Shiga, K. (1906). Observations on the epidemiology of dysentery in Japan. *Philipp J Sci* **1**, 485.

Simmons, D. A. R. & Romanowska, E. (1987). Structure and biology of *Shigella flexneri* O antigens. *J Med Microbiol* **23**, 289–302.

Tao, J., Feng, L., Guo, H., Li, Y. & Wang, L. (2004). The O-antigen gene cluster of *Shigella boydii* O11 and functional identification of its *wzy* gene. *FEMS Microbiol Lett* **234**, 125–132.

Tao, J., Wang, L., Liu, D., Li, Y., Bastin, D. A., Geng, Y. & Feng, L. (2005). Molecular analysis of *Shigella boydii* O1 O-antigen gene cluster and its PCR typing. *Can J Microbiol* **51**, 387–392.

Venkatesan, M. M., Buysse, J. M. & Kopecko, D. J. (1989). Use of *Shigella flexneri ipaC* and *ipaH* gene sequences for the general identification of *Shigella* spp. and enteroinvasive *Escherichia coli*. *J Clin Microbiol* **27**, 2687–2691.

Wang, L., Liu, B., Kong, Q., Steinruck, H., Krause, G., Beutin, L. & Feng, L. (2005). Molecular markers for detection of pathogenic *Escherichia coli* strains belonging to serogroups O 138 and O 139. *Vet Microbiol* **111**, 181–190.

JMM Journal of Medical Microbiology

Offprint Order Form

PAPER jmm2008/000794 **Please quote this number in any correspondence**

Authors Y. Li and others **Date** ____________

I would like 25 free offprints, plus [ ] additional offprints, giving a **total of** [ ] **offprints**

Dispatch address for offprints (BLOCK CAPITALS please)

____

____

____

____

Please complete this form **even if you do not want extra offprints.** Do not delay returning your proofs by waiting for a purchase order for your offprints: the offprint order form can be sent separately.

Please pay by credit card or cheque with your order if possible. Alternatively, we can invoice you. All remittances should be made payable to **'Society for General Microbiology'** and crossed **'A/C Payee only'**.

Tick one

☐ Charge my credit card account (give card details below)
☐ I enclose a cheque/draft payable to Society for General Microbiology
☐ Purchase order enclosed

Return this form to: JMM Editorial Office, Marlborough House, Basingstoke Road, Spencers Wood, Reading RG7 1AG, UK.

CHARGES FOR ADDITIONAL OFFPRINTS

Copies	25	50	75	100	125	150	175	200	Per 25 extra
No. of pages									
1-2	£23	£40	£58	£76	£92	£110	£128	£145	£23
3-4	£35	£58	£81	£104	£128	£150	£173	£191	£29
5-8	£46	£76	£104	£133	£162	£191	£219	£249	£35
9-16	£58	£92	£128	£162	£196	£231	£267	£301	£40
17-24	£70	£110	£151	£191	£231	£272	£312	£353	£46
each 8pp extra	£18	£23	£29	£35	£40	£46	£53	£58	

OFFICE USE ONLY
Issue:
Vol/part:
Page nos:
Extent:
Price:
Invoice: IR/

PAYMENT BY CREDIT CARD *(Note: we cannot accept American Express)*

Please charge the sum of £__________ to my credit card account.

My Mastercard/Visa number is (*circle appropriate card; no others acceptable*):

[| | | | | | | | | | | | | | |] Expiry date [| | |] Security Number [| |]

Signature: ____________________ Date: ____________

Cardholder's name and address*: ______________________________

____

**Address to which your credit card statement is sent. Your offprints will be sent to the address shown at the top of the form.*

May 2006