Applied and Environmental Microbiology, June 2007, p. 4082–4088 Vol. 73, No. 12
0099-2240/07/$08.00+0 doi:10.1128/AEM.01820-06

DNA Microarray-Based Identification of Serogroups and Virulence Gene Patterns of *Escherichia coli* Isolates Associated with Porcine Postweaning Diarrhea and Edema Disease▽†

Weiqing Han,[1,2,3] Bin Liu,[1,2,3] Boyang Cao,[1,2,3] Lothar Beutin,[4] Ulrike Krüger,[4] Hongbo Liu,[1,2,3] Yayue Li,[1,2,3] Yanqun Liu,[1,2,3] Lu Feng,[1,2,3] and Lei Wang[1,2,3]*

TEDA School of Biological Sciences and Biotechnology, Nankai University, 23 Hongda Street, TEDA, Tianjin 300457, China[1]; Tianjin Key Laboratory of Microbial Functional Genomics, 23 Hongda Street, TEDA, Tianjin 300457, China[2]; Tianjin Research Center for Functional Genomics and Biochip, 23 Hongda Street, TEDA, Tianjin 300457, China[3]; and National Reference Laboratory for Escherichia coli, Federal Institute for Risk Assessment (BfR), Diedersdorfer Weg 1, D-12277 Berlin, Germany[4]

Received 2 August 2006/Accepted 13 April 2007

***Escherichia coli* strains causing postweaning diarrhea (PWD) and edema disease (ED) in pigs are limited to a number of serogroups, with O8, O45, O138, O139, O141, O147, O149, and O157 being the most commonly reported worldwide. In this study, a DNA microarray based on the O-antigen-specific genes of all 8 *E. coli* serogroups, as well as 11 genes encoding adhesion factors and exotoxins associated with PWD and ED, was developed for the identification of related serogroups and virulence gene patterns. The microarray method was tested against 186 *E. coli* and *Shigella* O-serogroup reference strains, 13 *E. coli* reference strains for virulence markers, 43 *E. coli* clinical isolates, and 12 strains of other bacterial species and shown to be highly specific with reproducible results. The detection sensitivity was 0.1 ng of genomic DNA or 10^3 CFU per 0.3 g of porcine feces in mock samples. Seventeen porcine feces samples from local hoggeries were examined using the microarray, and the result for one sample was verified by the conventional serotyping methods. This microarray can be readily used to screen for the presence of PWD- and ED-associated *E. coli* in porcine feces samples.**

Postweaning diarrhea (PWD) and edema disease (ED) are two of the most prevalent porcine diseases worldwide. They account for substantial economical losses and are major causes of death in weaned pigs (3, 14). PWD is primarily caused by enterotoxigenic *Escherichia coli* (ETEC), while ED is caused by verocytotoxigenic *E. coli*. Virulence factors associated with PWD and ED include adhesion factors (F4, F5, F6, F18, F41, and intimin) and exotoxins (STa, STb, LT, Stx2e, and EAST1) (2, 12, 13, 23). Pathogenic strains colonize the small intestine through different types of adhesion factors and generate one or more exotoxins responsible for the diseases (12, 16). PCR-based methods have been developed to detect these virulence factor genes in porcine *E. coli* (13, 23, 24).

E. coli causing PWD and ED in pigs is limited to a few O serogroups, including O8, O45, O138, O139, O141, O147, O149, and O157, which are most commonly reported worldwide (12, 14, 16, 22). For example, these eight serogroups accounted for 81.2% of all isolates examined and 88.9% of typeable isolates in Denmark (12). Conventional O serotyping of *E. coli* strains in clinical specimens and environmental samples is laborious and time-consuming. The presence of capsules, capsular-like fimbriae, and rough lipopolysaccharide complicates the O serotyping of porcine-pathogenic *E. coli* strains and prompted us to develop a molecular typing method.

The O antigen, which consists of 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. There are 186 O-antigen forms recognized for *E. coli* (including *Shigella*) (11). Genes involved in biosynthesis of O antigen are normally clustered on the chromosome between two housekeeping genes, *galF* and *gnd*, in *E. coli*. Some O-antigen genes, including those encoding glycosyltransferases, O-unit flippase (Wzx), and O-antigen polymerase (Wzy), are often specific for different O antigens (36) and can be used as targets in molecular typing. PCR assays based on O-antigen-specific genes of *E. coli* O45, O138, O139, and O157 have been developed by us and others (6, 35, 36).

Conventional PCR is used to amplify a single target gene but the method is laborious for detection of multiple serogroup-specific genes and virulence factor genes. Multiplex PCR can amplify multiple targets in a single reaction, but a clear length differentiation among PCR products is required, which leads to a challenge for primer design. Recently DNA microarrays have been applied to microbial detection and community analysis (4, 19, 26, 27, 31, 34, 38). The approach involves the immobilization of numerous oligonucleotide DNA probes on a solid support to which fluorescence-labeled amplified target DNA is hybridized, which was shown to be rapid, reliable, and sensitive (5, 20, 25, 37). In this study, the O-antigen gene clusters of *E. coli* O141, O147, and O149 were sequenced and analyzed. These data were combined with published data from other O-antigen gene clusters to develop a genotyping microarray.

* Corresponding author. Mailing address: TEDA School of Biological Sciences and Biotechnology, Nankai University, 23 Hongda Street, TEDA, Tianjin 300457, China. Phone: 86-22-66229588. Fax: 86-22-66229596. E-mail: wanglei@nankai.edu.cn.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 20 April 2007.

The microarray also included genes encoding adhesion factors and exotoxins of porcine *E. coli* strains. The microarray was examined for its specificity and sensitivity and applied to 17 porcine feces samples.

All strains were inoculated into Luria-Bertani medium and incubated overnight at 37°C. Genomic DNA was prepared as previously described (1). O-antigen gene clusters from *E. coli* O141, O147, and O149 reference strains were sequenced using the protocol described previously (10), and sequences of 15,601 bp (12 *orf* genes), 10,252 bp (8 *orf* genes), and 8,729 bp (9 *orf* genes) were obtained between *galF* and *gnd*, respectively (see Fig. S1 in the supplemental material). All *orf* genes were assigned functions based on amino acid identities of their products to proteins of known functions and named accordingly (see Tables S1, S2, and S3 in the supplemental material). Sequences of *E. coli* O8, O45, O138, O139, and O157 O-antigen gene clusters were retrieved from GenBank.

The *wzy* gene was employed as the target gene for typing of *E. coli* O45, O138, O139, O141, O147, O149, and O157. Genes from the O-antigen gene cluster of *E. coli* O8 share high-level identity (from 88.4% to 96.3%) to those from the *Klebsiella pneumoniae* O5 *wb* gene cluster (21), which is also responsible for the synthesis of O antigen. The specific region of the glycosyltransferase gene *orf469*, which is the most heterogeneous gene, was used for typing *E. coli* O8. *E. coli* O149 and *Shigella boydii* type 1 have identical O antigens (8), and their O-antigen gene clusters were found to be nearly identical (99.8% to 100% identity for each gene set). The O-antigen structures of *E. coli* O147 and *Shigella flexneri* 6 are similar, with the only difference being the presence or absence of an O acetyl group on the side chain (7, 15), and their O-antigen gene clusters are expected to share high-level DNA identity. The invasion plasmid antigen gene (*ipaH*) has been shown to be unique to *Shigella* and enteroinvasive *E. coli* (EIEC) and can be used for the detection of *Shigella* from non-EIEC *E. coli* (17, 30, 33). None of the *E. coli* strains belonging to serogroups O147 and O149 are recognized as EIEC. Therefore, the *ipaH* gene was used as the target gene to differentiate *E. coli* O147 from *S. flexneri* 6 and *E. coli* O149 from *S. boydii* type 1. Genes encoding adhesion factors (F4, F5, F6, F18, F41, and intimin) and exotoxins (STa, STb, LT, Stx2e, and EAST1) were selected as target genes to reveal the virulence gene patterns of PWD- and ED-associated serogroups, and their sequences were retrieved from GenBank.

A total of 21 pairs of primers were designed based on each of the target genes described above, and they were divided into two groups to generate target DNAs by multiplex PCR (Table 1). Group 1 targets the O-antigen-specific genes and the *ipaH* gene, and group 2 targets the virulence genes. In addition, both groups contain the primer pair for amplifying the 16S rRNA gene as the internal positive control. The primer concentrations were optimized based on the final intensities of hybridization signals, which can be analyzed by the interpretation software developed in-house. An attempt to use the same concentration (0.2 μM) for all of the primers failed due to the detection of negative or weak fluorescence signals for some of the genes. Therefore, different primer concentrations (0.1 μM to 0.6 μM) were tested, and the best combinations are listed in Table 1. Multiplex PCR was performed with 50 to 100 ng of template DNA in a final volume of 50 μl containing the following: 1× PCR buffer; 2.5 mM $MgCl_2$; 400 μM (each) dATP, dCTP, dGTP, and dTTP; and 2.5 U *Taq* DNA polymerase. The PCR cycle was performed with the initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 50°C for 30 s, and 72°C for 1.5 min, concluding with a final elongation at 72°C for 5 min. The agarose gel images for the multiplex PCR are shown in Fig. S2 in the supplemental material. PCR products were then purified with the Microcon centrifugal filter devices kit (Millipore Corporation). To label PCR products, only reverse primers were used in a PCR, and 0.3 μl of 25 nM Cy3-dUTP was included. Ten microliters of the purified amplification products generated from the first-round multiplex PCR was added as the template. All labeled DNA was purified with the Microcon centrifugal filter devices kit and stored at −20°C in the dark until use.

Probes based on each of the target genes were designed using OligoArray 2.0 (28) (Table 2), synthesized with a 5′ amidocyanogen modifier, and added with a poly(T) spacer consisting of 15 thymine nucleotides. Probes (1 μg/μl in 50% dimethyl sulfoxide) were spotted onto an aldehyde group modified glass slide (CEL Corporation) to make a microarray, and three replicates were spotted for each probe to eliminate any possible physical defects in the glass slide. A probe based on the conserved region of the 16S rRNA gene was used as a positive control (Table 2). A probe containing 40 poly(T) oligonucleotides labeled with Cy3 at the 3′ end was used as the positional reference and printing control (see Fig. S3 in the supplemental material). For each target gene, one to four probes were used, and the layout of the array is shown in Fig. S3 in the supplemental material. To carry out the microarray assay, 10 μl of the labeled target DNA was mixed with 10 μl of preheated (50°C) hybridization buffer (25% formamide, 0.1% sodium dodecyl sulfate, 6× SSPE [1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA {pH 7.7}]) and hybridized with the array at 50°C for 15 h. The hybridized microarray was scanned with a laser beam of 532 nm using the 4100A biochip scanner (Axon Corporation) with the following parameters: photomultiplier tube gain of 600 and pixel size of 5 μm. Two files were generated, one for the images saved as .tif files and the other for the signal intensity, saved as .gpr. The signal-to-noise ratio was calculated for each spot using Bactarray Analyzer 1.0, developed in-house, with the threshold set at 3.0. A positive detection result was reported when all the probes of the given target gene generated hybridization signals above the signal-to-noise-ratio threshold.

The specificity of the DNA microarray was tested using 186 *E. coli* and *Shigella* O-serogroup reference strains, which represent all of the 186 O serogroups recognized, 13 *E. coli* reference strains for virulence markers, and 12 strains of other bacterial species (see Table S4 in the supplemental material). Multiple strains of different sources representing each of the target serogroups and virulence genes were used to test the microarray, and the numbers of strains are 6, 3, 9, 9, 6, 3, 7, and 9 for O8, O45, O138, O139, O141, O147, O149, and O157, respectively, and 10, 3, 2, 21, 3, 10, 19, 21, 11, 16, and 15 for virulence factors F4, F5, F6, F18, F41, intimin, STa, STb, LT, Stx2e, and EAST1, respectively. The sources of these strains are given in Table S4 in the supplemental material. All of the strains, belonging to the eight serogroups or carrying virulence markers, consistently hybridized to their corresponding probes,

TABLE 1. Primers and their concentrations in multiplex PCR[a]

Primer name	Target serogroup-specific gene or virulence factor (serogroup)	T_m (°C)	Sequence (5′–3′)[b]	Size (bp)	GenBank accession no. or reference	Concn in multiplex PCR (μM)
Group 1						
wl-5599	*orf469* (O8)	55.4	(F) 8124-TGTTCGTGCGATGGACC-8140	1,185	AB010150	0.6
wl-5600	*orf469* (O8)	55.1	(R) 9308-TATCCGAATGGGCCTTCT-9291	1,185	AB010150	0.6
wl-5617	*wzy* (O45)	51.8	(F) 9401-TGCTTCAATTTGGCTGTT-9418	842	AY771223	0.2
wl-5618	*wzy* (O45)	49.6	(R) 10242-CGTTGGCATTATCGTCTA-10225	842	AY771223	0.2
wl-818	*wzy* (O138)	50.3	(F) 11536-TGCCGACAACATTATCAA-11553	993	DQ109551	0.2
wl-5588	*wzy* (O138)	51.5	(R) 12528-CAAACTTTACCCGACGAA-12511	993	DQ109551	0.2
wl-5589	*wzy* (O139)	56.1	(F) 5003-ATAACGCATCCGCCAACT-5020	1,037	DQ109552	0.2
wl-5590	*wzy* (O139)	48.7	(R) 6039-CCGACTAATACGGAAACA-6022	1,037	DQ109552	0.2
wl-5591	*wzy* (O141)	47.5	(F) 8097-TGAACCTGGGTTTACATT-8114	746	This study	0.2
wl-5592	*wzy* (O141)	47.6	(R) 8842-GTACAATTATCATTGCGAGT-8823	746	This study	0.2
wl-5595	*wzy* (O147)	47.9	(F) 6208-TTTTGCTCTTATGGAACC-6225	689	This study	0.4
wl-5596	*wzy* (O147)	47.6	(R) 6896-ATAACGCCAAGTTGATTT-6879	689	This study	0.4
wl-5597	*wzy* (O149)	48.6	(F) 5755-TTTGGTGCAGATACTCAGA-5773	709	This study	0.3
wl-5598	*wzy* (O149)	48.5	(R) 6463-GAACAATAGATGCGATACAA-6444	709	This study	0.3
wl-5593	*wzy* (O157)	51.2	(F) 845-TCAGCGGCTAAGTTGATT-862	861	AF061251	0.2
wl-5594	*wzy* (O157)	52.1	(R) 1705-ATTTGCTCCCATGTCTCC-1688	861	AF061251	0.2
wl-5795	*ipaH*	54.9	(F) 963-TGACCGCCTTTCCGATA-979	585	AY206449	0.1
wl-5796	*ipaH*	52.0	(R) 1547-TTCTCCAGCATCTCATAC/TTTC-1527	585	AY206449	0.1
wl-3110	16S rRNA[c]	53.1	(F) AGAGTTTGATCC/ATGGCTCAG	919	37	0.15
wl-3111	16S rRNA[c]	54.5	(R) CCGTCAATTCCTTTGAGTTT	919	18	0.15
Group 2						
wl-5621	F4	53.4	(F) 169-TGATTTCAATGGTTCGGTC-187	770	V00292	0.2
wl-5622	F4	55.2	(R) 938-TTGCTACGTTCAGCGGAG-921	770	V00292	0.2
wl-5625	F5	47.7	(F) 235-CAGGCTGCTATTAGTGGT-252	356	M35282	0.2
wl-5626	F5	45.2	(R) 590-GCTGAAGTAGTAAATACGC-572	356	M35282	0.2
wl-5627	F6	47	(F) 378-AGTTACTGCCAGTCTATGC-396	396	M35257	0.2
wl-5628	F6	46.6	(R) 773-TTTGTATCAGGATTCCCT-756	396	M35257	0.2
wl-5623	F18	52.2	(F) 66-GCAAGGGGATGTTAAATTC-84	447	AY569976	0.4
wl-5624	F18	49.3	(R) 512-TTGTAAGTAACCGCGTAAG-494	447	AY569976	0.4
wl-5629	F41	48.7	(F) 185-TGACATTATTATTGGTGGTG-204	592	M21788	0.4
wl-5630	F41	49.9	(R) 776-GTGACTGAGGTCATCCC-760	592	M21788	0.4
wl-5633	Intimin	56.6	(F) 28-ACCCGGCACAAGCATAAG-45	929	AF530556	0.2
wl-5634	Intimin	51.2	(R) 956-TTATTGTATGACTCATGCCAG-936	929	AF530556	0.2
wl-5637	STap	52.6	(F) 286-ATGAAAAAGCTAATGTTGGC-305	193	M25607	0.6
wl-5638	STap	49	(R) 478-TACAACAAAGTTCACAGCAG-459	193	M25607	0.6
wl-5641	STb	52.7	(F) 464-AATATCGCATTTCTTCTTGC-483	204	AY028790	0.2
wl-5642	STb	54.4	(R) 667-GCATCCTTTTGCTGCAAC-650	204	AY028790	0.2
wl-5631	LT	47.4	(F) 914-CTATTACAGAACTATGTTCGG-934	291	S60731	0.6
wl-5632	LT	53.2	(R) 1204-TACTGATTGCCGCAATTG-1187	291	S60731	0.6
wl-5619	Stx2e	49.7	(F) 982-ATGAAGAAGATGTTTATAGCG-1002	264	AY368993	0.4
wl-5620	Stx2e	57.1	(R) 1245-TCAGTTAAACTTCACCTGGGC-1225	264	AY368993	0.4
wl-5635	EAST1	53.2	(F) 62-TGCCATCAACACAGTATATCC-82	109	L11241	0.2
wl-5636	EAST1	54.3	(R) 170-GCGAGTGACGGCTTTGT-154	109	L11241	0.2
wl-3110	16S rRNA[c]	53.1	(F) AGAGTTTGATCC/ATGGCTCAG	919	37	0.15
wl-3111	16S rRNA[c]	54.5	(R) CCGTCAATTCCTTTGAGTTT	919	18	0.15

[a] The multiplex PCR primer sets were divided into two groups. The first group targets the O-antigen-specific genes and *ipaH* gene, and the second targets the virulence genes.

[b] F, forward primer; R, reverse primer.

[c] The forward primer (37) and the reverse primer (18) target the conserved regions of the 16S rRNA gene. The amplicon binds to the 16S rRNA gene-based probe, which was used as the internal positive control.

TABLE 2. Oligonucleotide probes used in this study

Probe name	Target serogroup-specific gene or virulence factor (serogroup)	T_m (°C)[a]	Sequence (5′–3′)	GenBank accession no. or reference
OA_1937	*orf469* (O8)	79.16	9237-TCTGAAAACGGCAGAACTTTATCAAATCGACGATAAAT-9274	AB010150
OA_1938	*orf469* (O8)	79.44	9147-TGAAGCTTTGATTCAGAAAATCCTTGCTATCTCACAAG-9184	AB010150
OA_1939	*orf469* (O8)	79.04	8350-GACGATGTATACAAACCCTACTACTTACAGAAAGTTGAGT-8389	AB010150
OA_1940	*wzy* (O45)	83.16	10147-GCTTCCAGGACCTATTCCCCAGTGTTATCAGAAATCTCAT-10186	AY771223
OA_1941	*wzy* (O45)	81.23	10007-TCAATATTTGTTGTCACCAGAAGGACAATTTCTGTCGAGT-10046	AY771223
OA_1942	*wzy* (O45)	82.27	9605-GCTCATCATTTGGTGCTTTGTGATAATTCCTGATGTGGTT-9644	AY771223
OA_1943	*wzy* (O45)	82.62	9445-TGACTACTAATATTAGCCCGCTCTCAAATGAGATGGTGGT-9484	AY771223
OA_1917	*wzy* (O138)	79.63	12125-AGTTCTCGCTGTAGTTATCACCTCTGATCGATTATCTAAT-12164	DQ109551
OA_1918	*wzy* (O138)	80.14	11871-TGGTTGCATTTTGGAGTTTTCGATATAGCTACAGGTTTAT-11910	DQ109551
OA_1919	*wzy* (O138)	79.32	11700-GGAAACCCAGGTTATCTTAGTATGTCAACTGTATTGATCT-11739	DQ109551
OA_1920	*wzy* (O138)	79.33	11601-GCATGTAAGAATGCATACTCTTTATCGAGACAAATGGTTT-11640	DQ109551
OA_1921	*wzy* (O139)	79.68	5857-TGACTCGAAATTCAGTGAATTCAGTCATTTTGCAAAAGAT-5896	DQ109552
OA_1922	*wzy* (O139)	79.28	5837-AGCTCAAACATATTCAAAACTGACTCGAAATTCAGTGAAT-5876	DQ109552
OA_1923	*wzy* (O139)	79.42	5338-TGAACGAGGATTAGAAGAGCAGGTATATAGTACCATAACA-5377	DQ109552
OA_1924	*wzy* (O139)	79.63	5047-ACTATCAATGGCTTATATAGCTCACATAACTTTGGGGTTT-5086	DQ109552
OA_1925	*wzy* (O141)	79.35	8777-CGTTTGAATGGTTATACATGGCTAACAAAGGTAGATTGTT-8816	This study
OA_1926	*wzy* (O141)	79.37	8728-TCCATAGCTTACCGATTTTCGTACTATGGTTATGCTTTTA-8767	This study
OA_1927	*wzy* (O141)	79.58	8639-TCATCTGTCTGCTTTCCATTCAAGGAAGACATCTTAAATA-8678	This study
OA_1928	*wzy* (O141)	79.73	8537-TGAGTGGTTTTATTGCAACAAAATTAAAAGGATACGCTGA-8576	This study
OA_1933	*wzy* (O147)	79.3	6789-TTCTGGTTTGGGTAATGTCATCAAAATTCAAACTCCTATC-6828	This study
OA_1934	*wzy* (O147)	79.12	6704-CTGATTCTTCTTTAGGAGTTAGATGGGATCAGTTTACTGT-6743	This study
OA_1935	*wzy* (O147)	80.24	6519-AGTGGCTGGTAATTTTGCATATTTTGTTTCGATATGCTTG-6558	This study
OA_1936	*wzy* (O147)	79.24	6418-GCTCTCATTCCATTTGCCTTTATGTTGCATATAGTCATAA-6457	This study
OA_1966	*wzy* (O149)	79.75	5995-TCATTTATAGGCATTGAGTTGTTAACTAATGCTCTGAGGC-6034	This study
OA_1967	*wzy* (O149)	78.81	5881-AGTTGCCATATTAGTTGGTTAATTTGGATATGGCCAATAT-5920	This study
OA_1968	*wzy* (O149)	78.85	6035-AGGGATTCTCAATAGCAGTACTGTTATTATCGTTCTGTTT-6074	This study
OA_1929	*wzy* (O157)	79.94	1632-AGGGAATAAAGCATCAAGACTTATTTTATGGAGAACGGTT-1671	AF061251
OA_1930	*wzy* (O157)	79.05	1580-TTTCCGACACCAGAGTTAGAAAAGGAATTAAAAGCAATAA-1619	AF061251
OA_1932	*wzy* (O157)	79.7	1310-GGTTATCGTTCTGAATTGGTGTTGCTCATTCTTCAATATA-1349	AF061251
OA_1994	*wzy* (O157)	79.07	1245-ATGATGAGCATAAATTCAAACAGAGGACCATCATATTTGT-1284	AF061251
OA_1987	*ipaH*	75.13	1180-GATAATGATACCGGCGCTCTGCTCTCC-1206	AY206449
OA_1990	*ipaH*	74.28	1352-AGGAAATGCGTTTCTATGGCGTGTCG-1377	AY206449
OA_1944	Stx2e	79.18	1077-TAATGAGGATAATACCTTTACTGTGAAGGTGTCAGGAAG-1115	AY368993
OA_1945	Stx2e	79.79	1178-TGACTGTAACAATCATATCTAATACCTGCAGTTCAGGC-1215	AY368993
OA_1946	Stx2e	82.89	1112-GAAGAGAATACTGGACGAACAGATGGAATTTGCAGCCAT-1150	AY368993
OA_1959	Intimin	80.96	618-GTTACAACATTATGGAACGGCAGAGGTTAATCTGCAGAG-656	AF530556
OA_1995	Intimin	80.45	667-TTTGACGGTAGTTCACTGGACTTCTTATTACCGTTCTATG-706	AF530556
OA_1957	LT	79.8	1112-ACACATTAAGAATCACATATCTGACCGAGACCAAAATTGA-1151	S60731
OA_1958	LT	80.16	1001-GCAAAAGAGAAATGGTTATCATTACATTTAAGAGCGGCG-1039	S60731
OA_1961	EAST1	82.3	92-CATCCAGTTATGCATCGTGCATATGGTGCACAACAG-127	L11241
OA_1947	F4	81.63	202-TATCACTGCAGATGATTATCGTCAGAAATGGGAATGGA-239	V00292
OA_1948	F4	79.62	250-AGGTCTTAATGGATTTGGTAATGTATTGAATGACCTGACC-289	V00292
OA_1951	F5	80.23	456-TAATACTTCATTCACTACGGCTGAATACACTCACACTTCT-495	M35282
OA_1952	F5	80.68	441-TGGTGGTGCTAATATTAATACTTCATTCACTACGGCTGAA-480	M35282
OA_1953	F6	79.64	427-ATAGATCTTGGAGAGTTGTCTACTTCTGCTCTTAAAGCTA-466	M35257
OA_1954	F6	79.97	397-CAAGTGGATACTTCTAATCTGTCGCAAACCATAGATCTTG-436	M35257
OA_1949	F18	78.23	102-AACTACCTGTAATTTGACACCACAAATAAGTGGCACTGTA-141	AY569976
OA_1950	F18	78.62	411-ATCAAAAGCAACTACGGAAGGTTTCAAATTTACTGC-446	AY569976

Continued on following page

TABLE 2—*Continued*

Probe name	Target serogroup-specific gene or virulence factor (serogroup)	T_m (°C)[a]	Sequence (5′–3′)	GenBank accession no. or reference
OA_1955	F41	79.18	649-AGAATATAGCCTATAAATGGAATGGACTCTCAAAAGCTGA-688	M21788
OA_1956	F41	79.41	401-GGCATCTTATGATGGTAGTGTTATTACACCTAGTTTCACT-440	M21788
OA_1962	STap	79.74	322-TTATCTTTCCCCTCTTTTAGTCAGTCAACTGAATCACTTG-361	M25607
OA_1963	STap	77.69	361-GACTCTTCAAAAGAGAAAATTACATTAGAGACTAAAAAGTG-401	M25607
OA_1965	STb	79.11	499-TTCTATTGCTACAAATGCCTATGCATCTACACAATCAAA-537	AY028790
OA_1996	STb	78.1	592-AGGTTTTTTAGGGGTTAGAGATGGTACTGCTGGAGCATG-630	AY028790
OA_1997[b]	16S rRNA	72.4	GTGCCAGCAGCCGCGGTAATACG	X80725

[a] T_m was predicted using OligoArray 2.0 software.
[b] The 16S rRNA gene-based probe, whose location in the 16S rRNA gene of *E. coli* is 506 to 528, was used as the positive control.

indicating that the primers/probes can well represent the corresponding serogroups or virulence genes. The hybridization patterns observed for 15 reference strains are shown in (Fig. 1). All *Shigella* strains tested in this study yielded signals with *ipaH* gene-targeted probes, while all *E. coli* strains tested negative. None of the *E. coli* strains belonging to other serogroups and strains of other bacterial species, which are likely to be present in the small intestine, bound to the serogroup-specific probes on the microarray.

A double-blind test was carried out with 43 clinical isolates (see Table S4 in the supplemental material) which have been characterized for their virulence genes by conventional PCR (see Table S5 in the supplemental material) and for their O and H serotypes with specific antisera at the Federal Institute for Risk Assessment (BfR) in Berlin, Germany. A complete agreement with the results of conventional methods was obtained, and the hybridization patterns for representative clinical isolates are shown in Fig. S4 in the supplemental material. These results showed that the microarray assay was specific and reliable.

The microarray was applied for examination of 17 porcine feces samples (0.3 g) obtained from asymptomatic adult pigs from 4 local hoggeries. After 10 h of enrichment in LB medium at 37°C, 16 out of the 17 samples gave positive signals to at least 1 serogroup, and only sample 17 gave a negative signal (see Table S6 in the supplemental material). In order to confirm the results, the culture of sample 15, which gave positive signals for O8, O45, and O147 strains, was selected and plated to isolate single colonies. A total of 126 single bacterial colonies were picked and screened by PCR using O8-, O45-, and O147-specific primers, respectively. One each of O8- and O45-positive colonies were found when 30 colonies were screened, and their identities were confirmed by using the Biolog Microstation system (Biolog, Inc.) and the conventional O serotyping methods with commercial antisera (Institute of Veterinary Drug Control, Beijing, China) (data not shown). We did not detect any O147 strains after all 126 colonies were tested, and it is possible that the corresponding colonies were missed, which is a common problem for conventional antiserum serotyping. This indicates that the microarray is more sensitive than conventional serotyping methods. The O8 and O45 isolates were further screened for their virulence gene patterns using the same microarray, and the O45 strain was found to harbor an exotoxin EAST1-encoding gene. It has been reported that EAST1 contributes to the virulence of bacterial strains; however, the significance of this gene as a pathogenicity factor is not conclusive (32). The results for these 17 porcine feces samples showed that the microarray could identify PWD- and ED-related *E. coli* strains in the presence of other microorganisms and is highly specific.

To test the sensitivity of the microarray, serial dilution of genomic DNA (0.01 to 100 ng) was done for each of the eight target reference strains. The DNA dilutions were mixed with 500 ng of background DNA extracted from sample 17, which was not found to contain any target serogroup strains by microarray analysis, using the QIAamp DNA stool minikit (QIAGEN Inc.), and used as templates for multiplex PCR. Positive signals could be obtained from as little as 0.1 ng of genomic DNA. For porcine feces mock samples, serial dilutions from each of the pure cultures of the eight strains were mixed with approximately 0.3 g of porcine feces from sample 17. The spiked samples were cultured in 30 ml of LB medium at 37°C for 6 h, and 50 μl of genomic DNA was isolated with the QIAamp DNA stool minikit, out of which 3 μl was used as the template in the microarray assay. Cells from all eight strains were detected at levels as low as 10^3 CFU per 0.3 g of porcine feces.

It has been widely accepted that serotyping can be used as an epidemiological marker for pathogenicity of *E. coli* (2, 29). Conventionally, antigenic analysis of different serogroups in *E. coli* is performed by carrying out laborious agglutination reactions using antisera raised against the O reference strains. In addition, cross-reactions among different serogroups often occur, giving equivocal results; O serotyping fails if O-rough strains (which do not produce an O antigen as a result of mutations in one or more of the multiple genes controlling O-antigen synthesis and polymerization) are present, which was found not infrequently with isolates from PWD and ED (35). PCR-based tests amplifying specific genes in the *E. coli* O-antigen gene clusters have been found to be serogroup specific (6, 35). In this study, through sequencing of three *E. coli* O-antigen gene clusters which have not been identified before, we obtained specific genes for all of the eight major *E. coli* serogroups associated with PWD and ED.

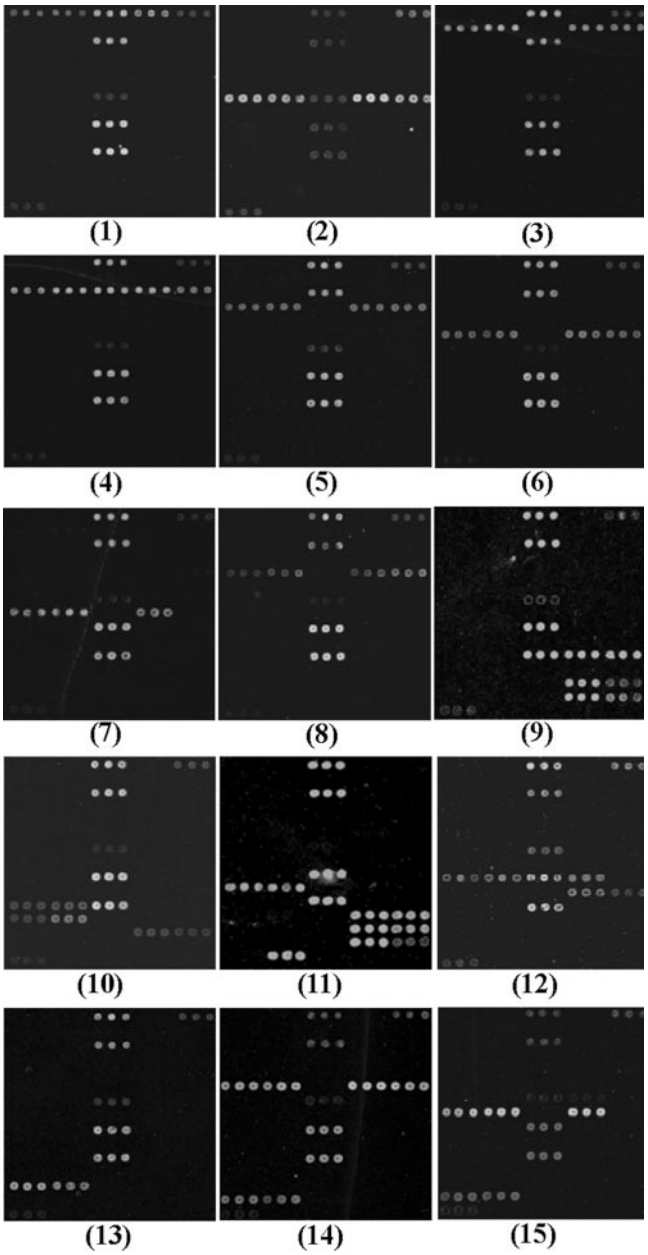

FIG. 1. Hybridization patterns for different *E. coli* strains. (1) *E. coli* O8. (2) *E. coli* O45. (3) *E. coli* O138. (4) *E. coli* O139. (5) *E. coli* O141. (6) *E. coli* O147. (7) *E. coli* O149. (8) *E. coli* O157. (9) *E. coli* strain harboring virulence factors STap, STb, and F6. (10) *E. coli* strain harboring virulence factors STap, F5, and F41. (11) *E. coli* strain harboring virulence factors LT, EAST1, STap, STb, and F4. (12) *E. coli* strain harboring virulence factors Stx2e and F18. (13) *E. coli* strain harboring virulence factor intimin. (14) *S. flexneri* 6. (15) *S. boydii* type 1.

Nevertheless, serotyping only indirectly identifies porcine-pathogenic strains. Therefore, 11 virulence factor genes associated with PWD and ED were also targeted in our microarray to give further indications of possible pathogenicity of the targeted serogroups. *E. coli* pathogenicity is a complex, multifactorial mechanism involving a large number of virulence factors which vary according to the pathotype, and a single component might not be sufficient to transform an *E. coli* strain into a pathogenic one but could only play a role in combination with other virulence determinants. The inclusion of virulence gene probes in the microarray is useful for virulence assessment of *E. coli* strains. Because several virulence traits are located on transferable genetic elements (9), it is likely that new pathogenic serogroups may emerge. The distribution and frequencies of pathogenic serogroups can vary considerably from one geographic region to another and over time in a given region. Characterization for the presence of virulence factor genes will detect all of the potential pathogenic strains, even if they appear in serogroups other than the currently recognized pathogenic ones.

In this study, the microarray assay was applied to 254 strains, including 186 *E. coli* and *Shigella* O serogroup reference strains, 13 *E. coli* reference strains for virulence markers, 43 *E. coli* clinical isolates, and 12 strains of other bacterial species. The microarray was also applied to 17 porcine feces samples. All results showed that the microarray assay is specific and reliable. The applicable samples of the microarray fall into two categories: bacterial isolates and porcine feces. For bacterial isolates, the microarray can be utilized as a specific and reliable alternative to the conventional O serotyping and PCR-based detection of virulence genes to characterize both serogroups and virulence gene patterns simultaneously.

When the microarray assay is directly applied to porcine feces samples, it can be used as an efficient preliminary screening method for surveillance of pig farms or for an epidemiological investigation in a certain region. The first step is to detect the presence of potential pathogenic O serogroups directly with porcine feces samples using the microarray; once the pathogenic serogroups are found, further plating can be carried out to isolate the targeted serogroup colonies; finally, the isolated strains can then be examined by the same microarray for the identification of virulence gene patterns. Although the microarray can be applied to detect both pathogenic serogroups and virulence genes directly with porcine feces samples at the same time, the serogroup-specific and virulence genes might be from different host strains.

The diseases of PWD and ED commonly occur in piglets within the first 14 days after weaning (12). Adult pigs carrying the PWD- and ED-associated serogroup strains with related virulence genes may not show any symptoms but can serve as an infection source for piglets, leading to an outbreak of PWD and ED in a large pig farm due to rapid transmission of pathogens, so preliminary screening for the presence of PWD- and ED-related *E. coli* strains using this microarray is important when starting to breed or buy piglets. This is the first DNA microarray for comprehensive detection of serogroups and pathotypes of *E. coli* associated with PWD and ED and is promising as a new diagnostic tool for investigations of sporadic infections and outbreaks and for surveillance of serogroup/pathotype distribution of these strains in different geographic locations.

Nucleotide sequence accession numbers. The DNA sequences of the *E. coli* O141, O147, and O149 O-antigen gene clusters have been deposited in GenBank under the accession numbers DQ868765, DQ868766, and DQ868764.

This work was supported by the NSFC Key Program (30530010) and funds from the Science and Technology Committee of Tianjin City (05YFGZG04700) to L.W. and L.F.

REFERENCES

1. **Bastin, D. A., and P. R. Reeves.** 1995. Sequence and analysis of the O antigen gene (*rfb*) cluster of *Escherichia coli* O111. Gene **164:**17–23.
2. **Bertschinger, H. U.** 1999. Postweaning *E. coli* diarrhea and edema disease. *In* B. E. Straw, S. D'Allaire, W. L. Mengeling, and D. J. Taylor (ed.), Diseases of swine. Iowa State University Press, Ames.
3. **Bertschinger, H. U., and C. L. Gyles.** 1994. Oedema disease of pigs, p. 171–191, *Escherichia coli* in domestic animals and humans, vol. 8. CAB International, Wallingford, United Kingdom.
4. **Bruant, G., C. Maynard, S. Bekal, I. Gaucher, L. Masson, R. Brousseau, and J. Harel.** 2006. Development and validation of an oligonucleotide microarray for detection of multiple virulence and antimicrobial resistance genes in *Escherichia coli*. Appl. Environ. Microbiol. **72:**3780–3784.
5. **Chizhikov, V., A. Rasooly, K. Chumakov, and D. D. Levy.** 2001. Microarray analysis of microbial virulence factors. Appl. Environ. Microbiol. **67:**3258–3263.
6. **DebRoy, C., P. M. Fratamico, E. Roberts, M. A. Davis, and Y. Liu.** 2005. Development of PCR assays targeting genes in O-antigen gene clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. Appl. Environ. Microbiol. **71:**4919–4924.
7. **Dmitriev, B. A., Y. A. Knirel, O. K. Sheremet, A. A. Shashkov, N. K. Kochetkov, and I. L. Hofman.** 1979. Somatic antigens of *Shigella*. The structure of the specific polysaccharide of *Shigella newcastle* (*Sh. flexneri* type 6) lipopolysaccharide. Eur. J. Biochem. **98:**309–316.
8. **Ewing, W. H.** 1986. Edwards and Ewing's identification of the Enterobacteriaceae, 4th ed. Elsevier Science Publishers, Amsterdam, The Netherlands.
9. **Fairbrother, J. M.** 1999. Neonatal *E. coli* diarrhea. *In* B. E. Straw, S. D'Allaire, W. L. Mengeling, and D. J. Taylor (ed.), Diseases of Swine. Iowa State University Press, Ames.
10. **Feng, L., S. N. Senchenkova, J. Yang, A. S. Shashkov, J. Tao, H. Guo, G. Zhao, Y. A. Knirel, P. Reeves, and L. Wang.** 2004. Structural and genetic characterization of the *Shigella boydii* type 13 O antigen. J. Bacteriol. **182:**383–392.
11. **Feng, L., W. Wang, J. Tao, H. Guo, G. Krause, L. Beutin, and L. Wang.** 2004. Identification of *Escherichia coli* O114 O-antigen gene cluster and development of an O114 serogroup-specific PCR assay. J. Clin. Microbiol. **42:**3799–3804.
12. **Frydendahl, K.** 2002. Prevalence of serogroups and virulence genes in *Escherichia coli* associated with postweaning diarrhoea and edema disease in pigs and a comparison of diagnostic approaches. Vet. Microbiol. **85:**169–182.
13. **Frydendahl, K., H. Imberechts, and S. Lehmann.** 2001. Automated 5′ nuclease assay for detection of virulence factors in porcine *Escherichia coli*. Mol. Cell. Probes **15:**151–160.
14. **Hampson, D. J.** 1994. Postweaning *Escherichia coli* diarrhoea in pigs, p. 171–191. *In* C. L. Gyles (ed.), *Escherichia coli* in domestic animals and humans, vol. 8. CAB International, Wallingford, United Kingdom.
15. **Hygge Blakeman, K., A. Weintraub, and G. Widmalm.** 1998. Structural determination of the O-antigenic polysaccharide from the enterotoxigenic *Escherichia coli* O147. Eur. J. Biochem. **251:**534–537.
16. **Imberechts, H., H. De Greve, and P. Lintermans.** 1992. The pathogenesis of edema disease in pigs. Vet. Microbiol. **31:**221–233.
17. **Kong, R. Y., S. K. Lee, T. W. Law, S. H. Law, and R. S. Wu.** 2002. Rapid detection of six types of bacterial pathogens in marine waters by multiplex PCR. Water Res. **36:**2802–2812.
18. **Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace.** 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl. Acad. Sci. USA **82:**6955–6959.
19. **Li, Y., D. Liu, B. Cao, W. Han, Y. Liu, F. Liu, X. Guo, D. A. Bastin, L. Feng, and L. Wang.** 2006. Development of a serotype-specific DNA microarray for identification of some *Shigella* and pathogenic *Escherichia coli* strains. J. Clin. Microbiol. **44:**4376–4383.
20. **Maynard, C., F. Berthiaume, K. Lemarchand, J. Harel, P. Payment, P. Bayardelle, L. Masson, and R. Brousseau.** 2005. Waterborne pathogen detection by use of oligonucleotide-based microarrays. Appl. Environ. Microbiol. **71:**8548–8557.
21. **Merino, S., M. Altarriba, L. Izquierdo, M. M. Nogueras, M. Regue, and J. M. Tomas.** 2000. Cloning and sequencing of the *Klebsiella pneumoniae* O5 *wb* gene cluster and its role in pathogenesis. Infect. Immun. **68:**2435–2440.
22. **Nagy, B., and P. Z. Fekete.** 1999. Enterotoxigenic *Escherichia coli* (ETEC) in farm animals. Vet. Rec. **30:**259–284.
23. **Ojeniyi, B., P. Ahrens, and A. Meyling.** 1994. Detection of fimbrial and toxin genes in *Escherichia coli* and their prevalence in piglets with diarrhoea. The application of colony hybridization assay, polymerase chain reaction and phenotypic assays. Zentbl. Vet. B **41:**49–59.
24. **Osek, J.** 1998. Application of polymerase chain reaction for determination of toxins in *Escherichia coli* strains isolated from pigs with diarrhea. Acta Microbiol. Pol. **47:**409–413.
25. **Panicker, G., D. R. Call, M. J. Krug, and A. K. Bej.** 2004. Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. Appl. Environ. Microbiol. **70:**7436–7444.
26. **Peplies, J., C. Lachmund, F. O. Glockner, and W. Manz.** 2006. A DNA microarray platform based on direct detection of rRNA for characterization of freshwater sediment-related prokaryotic communities. Appl. Environ. Microbiol. **72:**4829–4838.
27. **Rhee, S. K., X. Liu, L. Wu, S. C. Chong, X. Wan, and J. Zhou.** 2004. Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. Appl. Environ. Microbiol. **70:**4303–4317.
28. **Rouillard, J. M., M. Zuker, and E. Gulari.** 2003. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. **31:**3057–3062.
29. **Scheutz, F., T. Cheasty, D. Woodward, and H. R. Smith.** 2004. Designation of O174 and O175 to temporary O groups OX3 and OX7, and six new *E. coli* O groups that include verocytotoxin-producing *E. coli* (VTEC): O176, O177, O178, O179, O180 and O181. APMIS **112:**569–584.
30. **Theron, J., D. Morar, M. Du Preez, V. S. Brozel, and S. N. Venter.** 2001. A sensitive seminested PCR method for the detection of *Shigella* in spiked environmental water samples. Water Res. **35:**869–874.
31. **Tiquia, S. M., L. Wu, S. C. Chong, S. Passovets, D. Xu, Y. Xu, and J. Zhou.** 2004. Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples. BioTechniques **36:**664–670, 672, 674–675.
32. **Veilleux, S., and J. D. Dubreuil.** 2006. Presence of *Escherichia coli* carrying the EAST1 toxin gene in farm animals. Vet. Res. **37:**3–13.
33. **Venkatesan, M. M., J. M. Buysse, and A. B. Hartman.** 1991. Sequence variation in two *ipaH* genes of *Shigella flexneri* 5 and homology to the LRG-like family of proteins. Mol. Microbiol. **5:**2435–2445.
34. **Vora, G. J., C. E. Meador, M. M. Bird, C. A. Bopp, J. D. Andreadis, and D. A. Stenger.** 2005. Microarray-based detection of genetic heterogeneity, antimicrobial resistance, and the viable but nonculturable state in human pathogenic *Vibrio spp*. Proc. Natl. Acad. Sci. USA **102:**19109–19114.
35. **Wang, L., B. Liu, Q. Kong, H. Steinruck, G. Krause, L. Beutin, and L. Feng.** 2005. Molecular markers for detection of pathogenic *Escherichia coli* strains belonging to serogroups O 138 and O 139. Vet. Microbiol. **111:**181–190.
36. **Wang, L., and P. R. Reeves.** 1998. Organization of *Escherichia coli* O157 O antigen gene cluster and identification of its specific genes. Infect. Immun. **66:**3545–3551.
37. **Warsen, A. E., M. J. Krug, S. LaFrentz, D. R. Stanek, F. J. Loge, and D. R. Call.** 2004. Simultaneous discrimination between 15 fish pathogens by using 16S ribosomal DNA PCR and DNA microarrays. Appl. Environ. Microbiol. **70:**4216–4221.
38. **Wu, L., X. Liu, C. W. Schadt, and J. Zhou.** 2006. Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl. Environ. Microbiol. **72:**4931–4941.