AEM
Journals.ASM.org

Development of a DNA Microarray Method for Detection and Identification of All 15 Distinct O-Antigen Forms of *Legionella pneumophila*

Boyang Cao,[a,b,c,d] Fangfang Yao,[a,b,c,d] Xiangqian Liu,[a,b,c,d] Lu Feng,[a,b,c,d] Lei Wang[a,b,c,d,e]

Key Laboratory of Molecular Microbiology and Technology of the Ministry of Education, TEDA College, Nankai University, TEDA, Tianjin, People's Republic of China[a]; TEDA School of Biological Sciences and Biotechnology, Nankai University, TEDA, Tianjin, People's Republic of China[b]; Tianjin Research Center for Functional Genomics and Biochips, TEDA College, Nankai University, TEDA, Tianjin, People's Republic of China[c]; Tianjin Key Laboratory of Microbial Functional Genomics, TEDA College, Nankai University, TEDA, Tianjin, People's Republic of China[d]; Tianjin Biochip Corporation, TEDA, Tianjin, People's Republic of China[e]

***Legionella* is ubiquitous in many environments. At least 50 species and 70 serogroups of the Gram-negative bacterium have been identified. Of the 50 species, 20 are pathogenic, and *Legionella pneumophila* is responsible for the great majority (approximately 90%) of the Legionnaires' disease cases that occur. Furthermore, of the 15 *L. pneumophila* serogroups identified, O1 alone causes more than 84% of the Legionnaires' disease cases that occur worldwide. Rapid and reliable assays for the detection and identification of *L. pneumophila* in water, environmental, and clinical samples are in great demand. *L. pneumophila* bacteria are traditionally identified by their O antigens by immunological methods. We have recently developed an O serogroup-specific DNA microarray for the detection of all 15 distinct O-antigen forms of *L. pneumophila*, including serogroups O1 to O15. A total of 35 strains were used to verify the specificity of the microarray, including 15 *L. pneumophila* O-antigen standard reference strains and seven *L. pneumophila* clinical isolates as target strains, seven reference strains of other non-*pneumophila Legionella* species as closely related strains, and six non-*Legionella* bacterial species as nonrelated strains. The detection sensitivity was 1 ng of genomic DNA or 0.4 CFU/ml in water samples with filter enrichment and plate culturing. This study demonstrated that the microarray allows specific, sensitive, and reproducible detection of *L. pneumophila* serogroups. To the best of our knowledge, this is the first report of a microarray serotyping method for all 15 distinct O-antigen forms of *L. pneumophila*.**

In 1976, an outbreak of severe pneumonia among the participants of an American Legion convention in Philadelphia, PA, led to the recognition of Legionnaires' disease (1). The disease was later found to be caused by the Gram-negative bacterium *Legionella pneumophila* (*Legionella* for the legionnaires who were infected at the convention and *pneumophila* meaning "lung loving"), which belongs to the family *Legionellaceae*.

Legionella is widespread in freshwater habitats, and its ubiquity is probably due to the bacterium's capacity to survive under a wide range of environmental conditions (2). Parthuisot et al. studied the diversity and dynamics of *Legionella* species under different thermal and wastewater discharge conditions during an annual cycle, and the results suggest that *Legionella* spp. may be present at significant concentrations in many more freshwater environments than previously thought, highlighting the need for further ecological studies (3).

At least 50 *Legionella* species are known, and 70 serogroups have been identified, most notably, in *L. pneumophila* (4). The great majority (approximately 90%) of Legionnaires' disease cases that occur are caused by *L. pneumophila*, and among the at least 15 *L. pneumophila* serogroups identified, O1 alone is responsible for more than 84% of the Legionnaires' disease cases that occur worldwide (3, 5, 6). A study investigating 259 clinical and 3,128 environmental strains isolated in France in 2001 and 2002 showed that *L. pneumophila* serogroup O1 accounted for 28.2% of the environmental *Legionella* isolates, in contrast to 95.4% of the clinical isolates (7). Furthermore, the prospective surveillance of the extent of *Legionella* pollution conducted at three hot spring recreational areas in Beijing, China, in 2011 indicated that 51.9% of the spring water samples were *Legionella* positive, and their concentrations ranged from 1 to 2,218 CFU/liter. Again, *L. pneumophila* was the most frequently isolated species (98.9%), and serogroups 3 (25.3%), 6 (23.4%), 5 (19.2%), 1 (18.5%), 2 (10.2%), 8 (0.4%), 10 (0.8%), 9 (1.9%), and 12 (0.4%) were identified (8).

Of the methods currently used for the routine diagnosis of *Legionella* infection, *in vitro* culture is the most specific but its sensitivity is variable and unreliable because of the presence of a low number of viable organisms in most specimens (9). The detection of *Legionella* antigen in urine by enzyme immunoassays is very specific, and commercially available systems are able to detect *L. pneumophila* serogroup O1 but fail to detect other serogroups (10). Serological typing methods with monoclonal and multiclonal antibodies detect *L. pneumophila* only with the aid of intense preculturing (7, 8). Therefore, there is a need to establish a rapid and specific method that is at the same time more sensitive than the methods currently used. Amplification of *Legionella*-specific DNA sequences has been used to detect the bacterium in environmental water by using multilocus sequence analysis (11), the DNA gyrase subunit B gene (*gyrB*) (12, 13), the macrophage infectivity potentiator gene (*mip*) (14), and 5S rRNA (15). These technologies can typically provide much faster results but cannot distinguish the pathogen at the serogroup level.

Received 14 June 2013 **Accepted** 16 August 2013
Published ahead of print 23 August 2013
Address correspondence to Lei Wang, wanglei@nankai.edu.cn.

doi:10.1128/AEM.01957-13

The O antigen, which consists of many oligosaccharide unit (O unit) repeats, is part of the lipopolysaccharide (LPS) in the outer membrane of Gram-negative bacteria and contributes major antigenic variability to the cell surface (16). *L. pneumophila* uses the ATP-binding cassette transporter-dependent process for the synthesis and translocation of O antigens (17, 18, 19, 20, 21). In *L. pneumophila*, genes for O-antigen synthesis are normally located in a cluster that maps to the region between *hisH* and *wecA* on the chromosome (18).

Recently, PCR or PCR-based DNA microarray analysis of O-antigen-specific genes for several different serogroups of *Escherichia coli* and *Shigella* spp. have been developed (22, 23, 24, 25, 26). DNA microarrays detect thousands of specific DNA sequences simultaneously. In this study, a PCR-based DNA microarray based on the *wzx*, *wzy*, and *wecA* genes was established to detect all 15 distinct O-antigen forms (O1 to O15) of *L. pneumophila*. The microarray method described in this communication is specific, sensitive, and reliable and can be used as a better alternative to the traditional serotyping procedure, which is laborious and frequently cross-reactive.

MATERIALS AND METHODS

Bacterial strains. The *L. pneumophila* strains used for whole-genome sequencing by the Solexa technology (27) are standard strains NCTC 11230, NCTC 11232, DSM 7514, ATCC 33216, NCTC 11406, ATCC 33823, NCTC 11985, NCTC 11986, NCTC 12000, NCTC 12179, NCTC 12180, NCTC 12181, NCTC 12174, and ATCC 35251 of serogroups O2 to O15, respectively. The bacterial strains used for microarray assays in this study are listed in Table 1. Fifteen *L. pneumophila* O standard reference strains, seven reference strains of other non-*pneumophila Legionella* species, six non-*Legionella* bacterial species, and seven *L. pneumophila* clinical isolates were used in this study (Table 1).

Genomic DNA extraction. All of the *Legionella* strains were cultured on buffered charcoal yeast extract (BCYE) agar (Hope Biotechnology Co., Ltd., Qingdao, China) in a 5% CO_2 incubator at 37°C for 48 h. The other strains were cultivated in Luria-Bertani medium (Oxoid Ltd., Basingstoke, Hampshire, England) at 37°C overnight with shaking. Genomic DNA was prepared with a DNA extraction kit (Tiangen, Beijing, China).

Target genes and oligonucleotide primer design. Sequence alignment and comparison were performed with the ClustalW program. Phylogenetic trees were constructed by the neighbor-joining method and plotted by the Molecular Evolutionary Genetics Analysis (MEGA) 3.1 software package (http://www.megasoftware.net). Bootstrap analysis was carried out on the basis of 1,000 replicates. The target genes used to design primers and probes were *wzm* for *L. pneumophila* serogroups O4, O9, O10, O12, O13, and O15; *wzt* for *L. pneumophila* serogroups O1, O2, O3, O5, O6, O7, O8, and O14; and *wecA* for *L. pneumophila* serogroup O11. All of the primer pairs were designed with the Primer Premier 5.0 software (Premier Boost International). The primer sequences and concentrations used for the multiplex PCR are listed in Table 2.

Multiplex PCR and labeling of target genes. The multiplex PCR was carried out with two groups; the first group consisted of *L. pneumophila* serogroups O1, O2, O3, O4, O5, O6, O7, and O9, and the second group consisted of *L. pneumophila* serogroups O8, O10, O11, O12, O13, O14, and O15. Amplification was performed with 50 μl of a reaction mixture consisting of 100 ng of DNA, 1× PCR buffer (50 mM KCl, 2.5 mM $MgCl_2$, 10 mM Tris-HCl [pH 8.3]); 100 μM deoxynucleoside triphosphates, 2.5 U of *Taq* DNA polymerase (TaKaRa Biotechnology [Dalian] Co. Ltd.), and each primer at the concentration shown in Table 2. The reaction parameters were 94°C for 5 min; 35 cycles of 94°C for 30 s, 57°C for 1 min, and 72°C for 1 min; and a final extension at 72°C for 5 min. An aliquot of the PCR product (2 μl) was subjected to agarose gel electrophoresis to examine the amplified DNA (Fig. 1). To label PCR products, 10 μM Cy3-dUTP

TABLE 1 Bacterial strains used in this study

Application and bacterium	Serogroup	No. of strains from each source	Total no.
Targets for probe specificity testing			
Legionella pneumophila	O1	1,[a] 1[c]	2
Legionella pneumophila	O2	1[b]	1
Legionella pneumophila	O3	1[b]	1
Legionella pneumophila	O4	1[d]	1
Legionella pneumophila	O5	1[e]	1
Legionella pneumophila	O6	1[b]	1
Legionella pneumophila	O7	1,[e] 1[c]	2
Legionella pneumophila	O8	1[b]	1
Legionella pneumophila	O9	1[b]	1
Legionella pneumophila	O10	1[b]	1
Legionella pneumophila	O11	1[b]	1
Legionella pneumophila	O12	1[b]	1
Legionella pneumophila	O13	1[b]	1
Legionella pneumophila	O14	1[b]	1
Legionella pneumophila	O15	1[e]	1
Other species used for probe specificity testing			
Aeromonas hydrophila		1[h]	1
Legionella anisa		1[d]	1
Legionella bozemanii		1[e]	1
Legionella gormanii		1[e]	1
Legionella longbeachae		1[d]	1
Legionella micdadei		1[b]	1
Legionella steigerwaltii		1[b]	1
Legionella waltersii		1[b]	1
Pseudomonas aeruginosa		1[g]	1
Salmonella		1[f]	1
Shigella dysenteriae		1[g]	1
Staphylococcus aureus		1[g]	1
Yersinia enterocolitica		1[d]	1
Blind testing ($n = 17$)			
Legionella pneumophila	O1	1[a]	1
Legionella pneumophila	O5	1[e]	1
Legionella pneumophila	O6	1[b]	1
Legionella pneumophila	O13	1[b]	1
Aeromonas hydrophila		1[h]	1
Legionella anisa		1[d]	1
Legionella bozemanii		1[e]	1
Legionella gormanii		1[e]	1
Legionella longbeachae		1[d]	1
Legionella micdadei		1[b]	1
Legionella steigerwaltii		1[b]	1
Legionella waltersii		1[b]	1
Pseudomonas aeruginosa		1[g]	1
Salmonella		1[f]	1
Shigella dysenteriae		1[g]	1
Staphylococcus aureus		1[g]	1
Yersinia enterocolitica		1[d]	1

[a] Czech Collection of Microorganisms, Masaryk University, Brno, Czech Republic.
[b] National Collection of Type Cultures, Central Public Health Laboratory, London, United Kingdom.
[c] Center for Disease Control and Prevention of Nanshan District, Shenzhen, China.
[d] German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.
[e] American Type Culture Collection, Manassas, VA.
[f] School of Molecular and Microbial Biosciences, University of Sydney, Sydney, Australia.
[g] National Center for Medical Culture Collections, Beijing, China.
[h] Academy of Military Medical Sciences, Beijing, China.

TABLE 2 Primers used in multiplex PCR assays and their concentrations

Primer	Target gene	T_m (°C)	Orientation[a]	Sequence (5′–3′)	Size (bp)	GenBank accession no.	Primer concn (μM) for: Multiplex PCR	Primer concn (μM) for: Labeling
Group A								
wl-52601	O1/O7 *wzt*	56.3	F	TGGGRTCAATGATAGATACTCC	817	Lab data	0.20	
wl-52603		56.3	R	CCAATSCAACTGWGGTCTWATC			0.20	0.20
wl-41032	O2/O3 *wzt*	49.4	F	TTCAGAAATCCTCTGGAAG	908	Lab data	0.60	
wl-41033		49.9	R	ATTTGCTTGGAGAACCTTA			0.60	0.60
wl-41037	O4 *wzm*	51.1	F	CAACTCCGGATTGGTAAA	243	Lab data	0.12	
wl-41038		54.9	R	TTCAAATCGCGGTACCTG			0.12	0.12
wl-41388	O5 *wzt*	49.4	F	ATAATAAAGCAAGCCTTGAT	293	Lab data	0.12	
wl-41389		47.7	R	TTCTGGATGAAAACCAGTC			0.12	0.12
wl-22688	O6 *wzt*	53.7	F	TAAAGATATTGTAGAGAGCCAGC	517	Lab data	0.12	
wl-22689		52.7	R	CATAGAGAGATAACCCTCACATT			0.12	0.12
wl-52835	O9 *wzm*	47.1	F	ACAGATGGTTTGCCTTAC	420	Lab data	0.40	
wl-41040		52.2	R	TTCATACCAAAACCGCAG			0.40	0.40
Group B								
wl-47909	O8/O14 *wzt*	45.5	F	CAATACGAGATTAAAAGAAA	343	Lab data	0.20	
wl-47910		40.3	R	CTTTGGTCTTAATAAGCCATC			0.20	0.20
wl-41046	O10 *wzm*	46.7	F	ACACTTTTAGGCTTTGGT	645	Lab data	0.40	
wl-41047		49.1	R	CCCAAGCATAAAAACAATA			0.40	0.40
wl-50709	O11 *wecA*	55.0	F	TTGAATTCATTATTTCTTTTCG	228	Lab data	0.40	
wl-50711		53.7	R	ATGAATAATAAACTAATTAACTGA			0.40	0.40
wl-41050	O12/O15 *wzm*	50.2	F	GGGGATATTCAACCGTTA	751	Lab data	0.40	
wl-41051		45.1	R	TAAATCCTATTACAAATATAGC			0.40	0.40
wl-41044	O13 *wzm*	50.0	F	TTGTCATTTGTGCCACAG	297	Lab data	0.40	
wl-41045		47.3	R	GCCAATTACCCTTTAAAC			0.40	0.40
Positive controls								
Wl-3110	16S rRNA		F	TGTACACACCGCCCGTC	500–1,000	AB553285.1	0.08	
Wl-3111			R	GGTACTTAGATGTTTCAGTTC		AY987650.1	0.08	0.13

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

(Amersham Biosciences UK Ltd., Little Chalfont, United Kingdom) and each reverse primer at the concentration shown in Table 2 were included in a PCR mixture. Twelve microliters of the amplification products generated from the above multiplex PCR was added as the template to 30 μl of the PCR mixture. The thermal cycling conditions were the same as for the multiplex PCR. All labeled DNA was stored at −20°C in the dark.

Oligonucleotide probe design. For each type of pathogen, one to four probes were designed by OligoArray 2.0 on the basis of the sequences in GenBank. One probe based on the 16S rRNA gene was designed as the positive control. A probe containing 40 poly(T) oligonucleotides was used as the negative control. A probe containing 40 poly(T) oligonucleotides and labeled with 3′-Cy3 was used as the positional reference and printing control. Each probe comprised a modified 5′ amino acid sequence followed by a spacer of 10 to 15 poly(T)s and a stretch of specific sequence (synthesized by AuGCT Biotechnology Corporation, Beijing, China). All of the oligonucleotide probes used are listed in Table 3.

DNA array preparation and hybridization. The probes were dissolved in 50% dimethyl sulfoxide (DMSO) at a final concentration of 1 μg/μl and printed onto aldehyde group-modified glass slides (CapitalBio Corporation, Beijing, China) with SpotArray 72 (Perkin-Elmer Corporation). Each probe was spotted in triplicate to eliminate irregular data arising from physical defects in the glass slides. Printed slides were dried

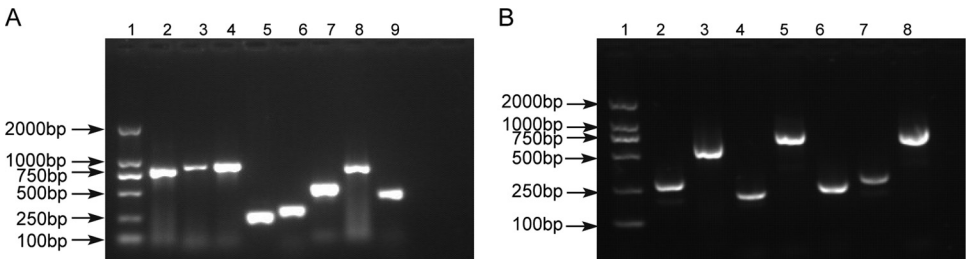

FIG 1 Agarose gel electrophoresis of multiplex PCR products. (A) Lanes: 1, molecular size standards (DL2000 marker); 2, *L. pneumophila* serogroup O1; 3, *L. pneumophila* serogroup O2; 4, *L. pneumophila* serogroup O3; 5, *L. pneumophila* serogroup O4; 6, *L. pneumophila* serogroup O5; 7, *L. pneumophila* serogroup O6; 8, *L. pneumophila* serogroup O7; 9, *L. pneumophila* serogroup O9. (B) Lanes: 1, molecular size standards (DL2000 marker); 2, *L. pneumophila* serogroup O8; 3, *L. pneumophila* serogroup O10; 4, *L. pneumophila* serogroup O11; 5, *L. pneumophila* serogroup O12; 6, *L. pneumophila* serogroup O13; 7, *L. pneumophila* serogroup O14; 8, *L. pneumophila* serogroup O15.

TABLE 3 Oligonucleotide probes used in this study

Probe name	Target gene or virulence factor	T_m (°C)[a]	Sequence (5′–3′)	GenBank accession no.
OA-3383	O1 *wzt*	67.8	CAGTCATGGATATCACTCGCGACTATCTC	CP000675.2
OA-3386		67.6	GCGAAATATAAATCGGAACAGGTTTGG	
OA-3761	O2 *wzt*	63.9	GTTAGCAGTTGGAGATCAGGATTTTCA	Lab data
OA-3762		61.2	GCCATGATATGAGTGCTATTGAATCGATTTG	
OA-3762	O3 *wzt*	61.2	GCCATGATATGAGTGCTATTGAATCGATTTG	Lab data
OA-3758		66.1	TATTAGCAGTAGGAGATCAAGATTTTCAAA	
OA-3702	O4 *wzm*	65.8	CAACTCCGGATTGGTAAATAAAATTTATTTT	Lab data
OA-3703		71.4	CGGTTTAATTATAATTTGCGCCACCATTTATG	
OA-3377	O5 *wzt*	69.9	CAAGCCTTGATGATGAGAATACTCAAAAATCC	Lab data
OA-3705		67.0	GAAACAGAAGAGTTCCAGAGTTTGTGTAAGG	
OA-3706		74.8	CCAAATATCACCTGGCGAACGCATAGGTTTAA	
OA-3707		66.7	ATCCAGAATAACTACTCCGACTTATGGTGAA	
OA-3373	O6 *wzt*	67.0	TAGATTCAATTGCAAGATCTCATGCCC	Lab data
OA-3374		63.8	ACAAGGTTCTTATTCATTCTCTTAAGCTTTT	
OA-3770	O7 *wzt*	66.0	CTGTGTAGAGCTTAATTGTGTGAGATTATTGG	Lab data
OA-3792		66.0	AAATGCAAATTCAGTCAGTAAAAGTTCAGTG	
OA-3793		69.0	CCATAGTCCAATGCGTTTTTTAAACTGTGTAG	
OA-3680	O8 *wzt*	64.2	TACAGATTTAGAAAATAATGGACCAGCAAT	Lab data
OA-3776		60.4	TTCACTTGTCATTACAGATTTAGAAAATAAT	
OA-3359	O9 *wzm*	67.8	CCTACAGATCAGTAGAGGGTAACGCCG	Lab data
OA-3360		65.7	TCCCTACAGATCAGTAGAGGGTAACGC	
OA-3799		68.2	GCTTGGTGGGACATATACCCTATCTGCT	
OA-3800		61.6	CCTTTGTGCCATTTTATATCATAAACTT	
OA-3801		64.4	TTTTCAGTTATTCGTGCTATTTCCTTTTT	
OA-3802		76.0	GTTTCAGCTTCACCACATTGGCTGCGGTTT	
OA-3366	O10 *wzm*	64.8	CCTTGTTGGATTTGGTTTAATAATAATTTGT	Lab data
OA-3367		64.3	TTTAGGATTAATGTTGGCTCCTTTTAGTACT	
OA-3771	O11 *wecA*	64.1	TATTATTTATTTTCTTACACCGCATTATTGG	Lab data
OA-3772		59.0	TCATCTCATATTGATTCAGTTAATTAGTTT	
OA-3773		59.4	CATATTGATTCAGTTAATTAGTTTATTATTCAT	
OA-3774		67.3	GATTATAAATCACGAAAAATGGTTTAGCCCT	
OA-3775		64.1	TTGAATTCATTATTTCTTTTCGATTCTACAG	
OA-3369	O12 *wzm*	63.4	CAGTTATAGAGATGCACTAAATGGTCGAT	Lab data
OA-3370		65.5	TTCACCTCATTGGTTAAGATTCTGGTATG	
OA-3361	O13 *wzm*	62.2	TGCCACAGTTTATGATAAACTTAATTATCA	Lab data
OA-3362		60.0	CCTACTATTTGTTATTCAAGCAATGTTT	
OA-3363		63.1	TTTGTACCACTGGCATTACAATTTTATTAT	
OA-3680	O14 *wzt*	64.2	TACAGATTTAGAAAATAATGGACCAGCAAT	Lab data
OA-3776		60.4	TTCACTTGTCATTACAGATTTAGAAAATAAT	
OA-3369	O15 *wzm*	63.4	CAGTTATAGAGATGCACTAAATGGTCGAT	Lab data
OA-3370		65.5	TTCACCTCATTGGTTAAGATTCTGGTATG	
OA-1993	16S rRNA gene	71.9	1380-TTGTACACACCGCCCGTCACACCAT-1404[b]	X80725
WL-4006			TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT[c]	
Cy3			TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-Cy3[d]	

[a] T_m was predicted using Primer Premier 5.0 software.
[b] The 16S rRNA gene-based probe was used as the positive control.
[c] The probe containing 40 poly(T) oligonucleotides was used as the negative control.
[d] The probe labeled with 3′-Cy3 was used as the positional reference and printing control.

and stored at room temperature in the dark. Before use, the slides were scanned at 532 nm for spotting quality control. Each glass slide consisted of eight individual arrays framed by a 20-μl Geneframe (CapitalBio Corporation, Beijing, China), which consisted of individual reaction chambers. A schematic diagram of the probe positions on the microarray is shown in Fig. 2. Hybridization was performed according to the following procedure. All 30 μl of the labeled PCR product was heated for approximately 2 h at 65°C until dry and diluted in 20 μl of hybridization buffer (30% formamide, 0.5% SDS, 6× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 5× Denhardt's solution). The mixture was then applied to a hybridization chamber and incubated at 43°C for 12 h in a water bath. Subsequently, the slide was washed sequentially in solution A (1× SSC, 0.2% SDS) for 3 min, in solution B (0.2× SSC) for 3 min, and in solution C (95% alcohol) for 1.5 min. The slide was dried under a gentle air stream before it was scanned.

Data acquisition and automated analysis. The slide was scanned with 532-nm laser beams with a 4100A biochip scanner (Axon Corporation), a photomultiplier tube gain of 600, and a pixel size of 5 μm. The image files were saved as .tif files, and the signal intensity was saved as .gpr files. The signal-to-noise ratio was calculated for each spot with a Bactarray Analyzer 1.0 developed in house with the threshold set at 3.0. A detection result was recorded as positive only when all of the hybridization signals generated by the probes of the target genes were above the signal-to-noise ratio threshold.

Cy3	OA-3383	OA-3386	OA-3702	OA-3703
OA-1993	OA-3761	OA-3762	OA-3758	50%DMSO
50%DMSO	OA-3377	OA-3705	OA-3706	OA-3707
Cy3	OA-3373	OA-3374	50%DMSO	OA-3770
OA-1993	OA-3771	OA-3772	OA-3773	OA-3774
WL-4006	OA-3369	OA-3370	OA-3366	OA-3367
Cy3	OA-3361	OA-3362	OA-3363	50%DMSO
50%DMSO	OA-3680	OA-3776	50%DMSO	50%DMSO
50%DMSO	OA-3792	OA-3793	OA-3359	OA-3360
50%DMSO	OA-3799	OA-3800	OA-3801	OA-3802

FIG 2 Probe positions on the slide. OA-1993 is the positive-control probe based on the 16S rRNA gene. WL-4006 is the negative-control probe. Cy3 is the positional reference and printing control probe. The blank is 50% DMSO. The rest are the specific probes for the target strains.

Testing of mock samples. BCYE medium was used for proliferation. Pure cultures of *L. pneumophila* serogroups O1, O10, and O12 were serially diluted to 10^1 to 10^6 CFU/ml, and 1 ml of the diluent was mixed with 100 ml of fresh tap water from the laboratory and vacuum filtered with a 0.22-μm membrane. The membrane was treated with 500 μl of diluted HCl (pH 3.0) for 1 min, placed face down on BCYE agar plates, and incubated in a 5% CO_2 incubator at 37°C for 3 to 5 days. The genomic DNA was then extracted from the cells for microarray hybridization.

Testing of air conditioner condensed-water samples. Eight hundred milliliters of filter-enriched air conditioner condensed-water sample was plated onto a GVPC agar plate (Hope Bio-Technology Co., Ltd., Qingdao, China), and after incubation in a 5% CO_2 incubator at 37°C for 48 h, the culture was collected and genomic DNA was extracted for use in the multiplex PCR.

Nucleotide sequence accession numbers. The DNA sequences of the *wzm* and *wzt* genes of *L. pneumophila* serogroups O2, O3, O4, O5, O6, O7, O8, O9, O10, O12, O13, O14, and O15 have been deposited in GenBank under accession numbers KF536969 to KF536994.

RESULTS

Phylogenetic analysis of *wzm* and *wzt* genes. The *wzm* and *wzt* gene sequences of 14 of the 15 *L. pneumophila* serogroups, with the exception of serogroup O11, were determined by Solexa sequencing technology. Sequencing failed to detect the *wzm* and *wzt* genes in serogroup O11. Two phylogenetic trees of 14 *L. pneumophila* serogroup strains based on the *wzm* and *wzt* gene sequences were constructed with *E. coli* IAI1 as the outgroup reference (Fig. 3A and B). In both cases, *L. pneumophila* serogroups O1 and O7 formed a small group and the other 12 serogroups clustered together as a large group. For the *wzm* tree, the large group was divided into two equal-size subgroups, with *L. pneumophila* serogroups O2, O3, and O9 grouping together with *L. pneumophila* serogroups O6, O12, and O15 in one subgroup and *L. pneumophila* serogroups O5 and O13 grouping together with *L. pneumophila* serogroups O10, O4, O8, and O14 in the other subgroup. In the *wzt* tree, *L. pneumophila* serogroups O6, O12, and O15 separated from the smaller subgroup of *L. pneumophila* serogroups O2, O3, and O9 and instead grouped together with *L. pneumophila* serogroups O8, O14, O4, O10, O5, and O13 to form a bigger subgroup.

Optimization of PCRs. A multiplex PCR was used to amplify and label the target genes in two groups. Group A targeted serogroups O1, O2, O3, O4, O5, O6, O7, and O9, and group B targeted serogroups O8, O10, O11, O12, O13, O14, and O15. The *wzt* genes

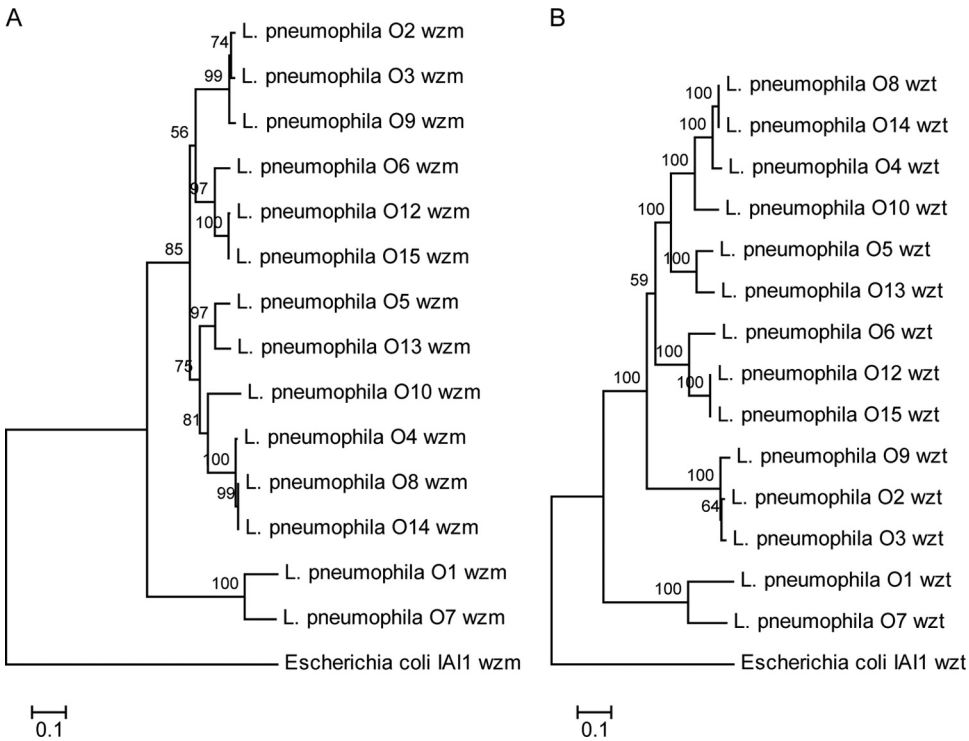

FIG 3 Unrooted phylogenetic trees constructed by the neighbor-joining method on the basis of the *wzm* and *wzt* genes. Bootstrap values were based on 1,000 replications, and only values greater than 50% are shown. (A) Unrooted *wzm* gene phylogenetic tree of 14 *L. pneumophila* serogroups constructed by the neighbor-joining method. (B) Unrooted *wzt* gene phylogenetic tree of 14 *L. pneumophila* serogroups constructed by the neighbor-joining method.

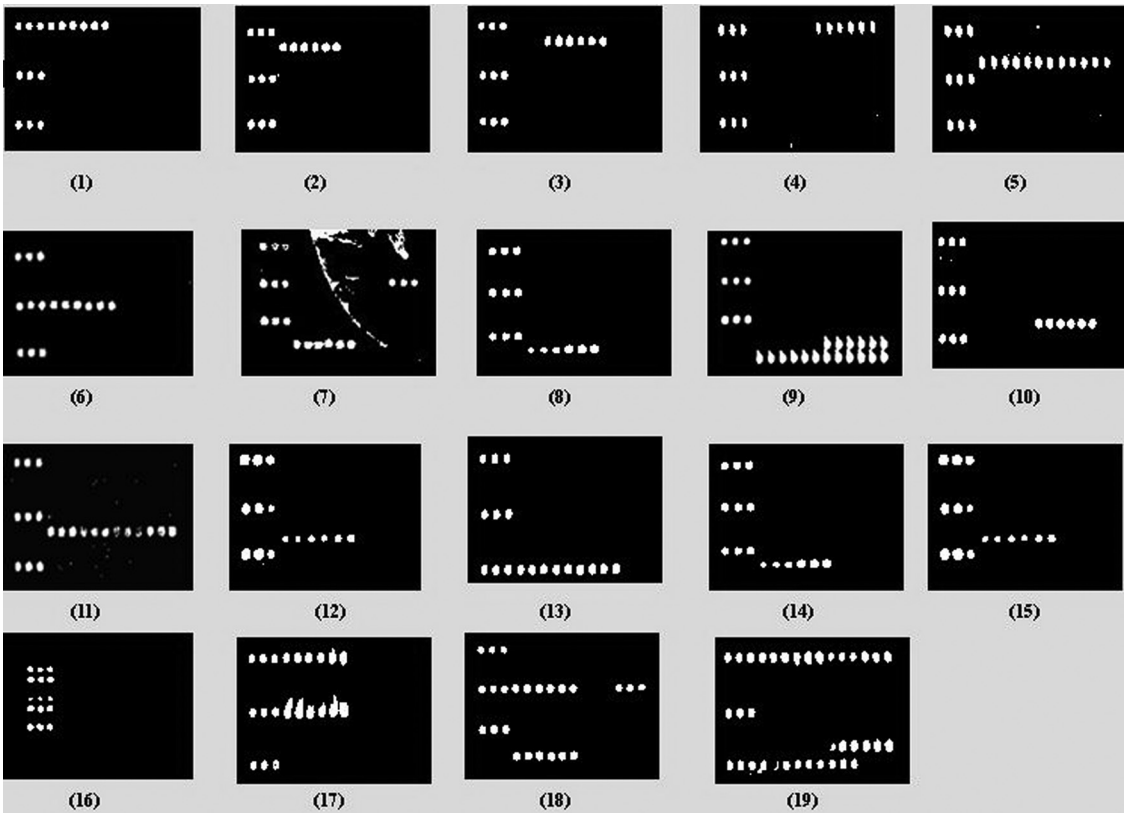

FIG 4 Microarray differentiation of pathogens. Panels: 1, *L. pneumophila* serogroup O1; 2, *L. pneumophila* serogroup O2; 3, *L. pneumophila* serogroup O3; 4, *L. pneumophila* serogroup O4; 5, *L. pneumophila* serogroup O5; 6, *L. pneumophila* serogroup O6; 7, *L. pneumophila* serogroup O7; 8, *L. pneumophila* serogroup O8; 9, *L. pneumophila* serogroup O9; 10, *L. pneumophila* serogroup O10; 11, *L. pneumophila* serogroup O11; 12, *L. pneumophila* serogroup O12; 13, *L. pneumophila* serogroup O13; 14, *L. pneumophila* serogroup O14; 15, *L. pneumophila* serogroup O15; 16, non-*pneumophila Legionella* species; 17, *L. pneumophila* serogroups O1 and O6; 18, *L. pneumophila* serogroups O6 and O7; 19, *L. pneumophila* serogroups O1, O4, O10, and O13.

of *L. pneumophila* serogroups O2 and O3 are very similar (97.5% identical); therefore, the same primer pair, wl-41032/wl-41033, was used to amplify the target *wzt* genes of both of these serogroups. Since the *wzm* and *wzt* genes of serogroups O8 and O14 are the same, primer pair wl-47909/wl-47910 was used to amplify *wzm* of *L. pneumophila* serogroups O8 and O14. Also, since the *wzm* and *wzt* genes of serogroups O12 and O15 are 100% identical, primer pair wl-41050/wl-41051 was used to amplify *wzt* of serogroups O12 and O15. The *wecA* gene was used to detect serogroup O11. The multiplex PCR was used to streamline the test and maintain the specificity of individual amplicons. Initially, all of the primers were used at 0.2 μM. However, several pathogens failed to generate the expected hybridization signals under the conditions used. Consequently, primer concentrations of 0.12 to 0.60 μM were tested to establish the optimal concentration (Table 2). The amplicons of the 11 groups of target pathogens were amplified under optimized conditions, and the PCR product lengths varied from 481 to 1,000 bp (Fig. 1A and B).

Probe specificity. The DNA microarray was tested by using 35 strains, including 15 *L. pneumophila* O standard reference strains, seven reference strains of other non-*pneumophila Legionella* species, six non-*Legionella* bacterial species, and seven *L. pneumophila* clinical isolates (Table 1). A total of 38 probes were used in the microarray, including 35 probes for specific genes, 1 positive control, 1 negative control, and 1 positional reference and printing control (Table 3). The probe positions on the slide are shown in Fig. 2. All of the representative strains belonging to the 15 groups consistently hybridized to the corresponding probes with 100% specificity. In addition, in order to distinguish *L. pneumophila* serogroup O2 from serogroup O3, specific probe OA-3761 was designed for *L. pneumophila* serogroup O2, specific probe OA-3758 was designed for *L. pneumophila* serogroup O3, and probe OA-3762 was designed for both serogroups. The hybridization results are shown in Fig. 4, panels 1 to 16.

Sensitivity of detection with genomic DNA. Serial 10-fold dilutions of 10, 1.0, 0.1, and 0.01 ng of genomic DNA of *L. pneumophila* serogroup O1 from the first PCR group and serogroup O12 from the second PCR group were used as the templates in a multiplex PCR to test its sensitivity. Positive signals were generated by using 1 ng of DNA for both groups A and B. Therefore, the sensitivity of detection with genomic DNA was set at 1 ng of DNA.

Simultaneous detection of multiple pathogens. Genomic DNAs from two pathogens of *L. pneumophila* serogroups O1 and O6, *L. pneumophila* serogroups O6 and O7, and four pathogens of *L. pneumophila* serogroups O1, O4, O10, and O13 were mixed and used as templates to test the specificity of the microarray assay. Data demonstrated that the probes were able to hybridize with and detect multiple pathogens in the sample (Fig. 4, panels 17 to 19).

Blind test. A blind test was performed in order to verify the

reliability and specificity of the microarray method. A total of 17 environmental and clinical isolates (Table 1) were used for hybridization to the microarray without disclosure of their identities during testing. The test samples included *L. pneumophila* serogroup O1, *L. pneumophila* serogroup O5, *L. pneumophila* serogroup O6, *L. pneumophila* serogroup O13, *L. anisa*, *L. bozemanii*, *L. gormanii*, *L. longbeachae*, *L. micdadei*, *L. steigerwaltii*, *L. waltersii*, *Aeromonas hydrophila*, *Pseudomonas aeruginosa*, *Salmonella*, *Shigella dysenteriae*, *Staphylococcus aureus*, and *Yersinia enterocolitica*. The detection results are consistent with those of conventional methods.

Testing of mock samples. *L. pneumophila* serogroups O1, O10, and O12 were detected at levels of 20 CFU/100 ml, 90 CFU/100 ml, and 11 CFU/100 ml, respectively. On average, as few as 40 CFU/100 ml or 0.4 CFU/ml of *L. pneumophila* can be detected by filter enrichment and culturing.

Testing of environmental isolates and confirmation by sequencing. A total of five *L. pneumophila* isolates, G2759, G2761, G2762, G2763, and G2765, from air conditioner condensed-water samples were obtained from the Center for Disease Control and Prevention of Nanshan District, Shenzhen, China. The microarray profile identified them as belonging to *L. pneumophila* serogroups O1, O2, O3, O7, and O5, respectively. The identities of these samples were further validated by amplification and sequencing of the *wzt* gene of each sample (data not shown).

Testing of real water samples and confirmation by sequencing. Seven batches of condensed-water samples from air conditioners collected, enriched, and provided by the Center for Disease Control and Prevention, Shanghai, China, were analyzed by the microarray method. The microarray profile showed that three of the seven samples were *L. pneumophila* serogroup O1 (Fig. 4, panel 1). The remaining four samples tested yielded positive-control probe signals only. Further validation was done by amplification and sequencing of the *wzt* gene of each sample. The overall results obtained by the microarray method produced 100% accuracy, indicating that the microarray method has the practical ability to detect and differentiate *L. pneumophila* serogroups.

DISCUSSION

In this report, we describe a microarray method for the determination of *L. pneumophila* O serogroups on the basis of the O serogroup-specific genes. This is the first report of the comprehensive detection and identification of all of the 15 serogroups of *L. pneumophila*. In comparison with other molecular-analysis-based detection methods, such as PCR and DNA sequencing, the advantages of the present method include (i) high throughput, (ii) high specificity, and (iii) greater efficiency. PCR is rapid and easy to process, but it cannot differentiate amplicons of similar sizes from different serogroups. DNA sequencing is very accurate, but it requires more expensive equipment, costs more, and takes a lot of time and hands-on work for library preparation and sequencing (28, 29).

The O antigen contains many repeats of an O unit and constitutes part of the LPS present in the outer membrane of Gram-negative bacteria (30). The O antigen exhibits variation in terms of the types of sugars, the arrangement of the sugars within the O unit, and different types of linkages within and between O units (26). This makes LPS one of the most variable cell constituents and provides a basis for the serotyping of many Gram-negative bacteria (31, 32, 33). So far, only *L. pneumophila* serogroup O1 LPS gene clusters have been determined. The *L. pneumophila* serogroup O1 LPS gene locus includes genes that are involved in LPS core oligosaccharide biosynthesis (*rmlA* to *rmlD*, rhamnosyltransferases, acetyltransferase), and LPS O-chain biosynthesis and translocation (*neuC*, *neuB*, *neuA*, *wecA*, *wzt*, and *wzm*) (18). In this study, *wzt* and *wzm* were used as the target genes for 14 of the 15 *L. pneumophila* serogroups, with the exception of serogroup O11, for which *wecA* was used as the target gene instead.

The resolving power of the microarray assay is comparable to that of traditional immunological methods; for example, this method is able to differentiate closely related serogroups O1 and O7 and closely related serogroups O2 and O3 by the use of serogroup-specific probes. The current version of the microarray failed to differentiate serogroup O8 from O14 and serogroup O12 from O15 because the target genes used are the same. Future addition of other target genes in the O-antigen gene cluster or in other regions of the genome should be able to differentiate these serogroups.

A two-step multiplex PCR was used for amplification and labeling. In the first step, the target genes were amplified with the forward and reverse primers, and in the second step, the single-stranded DNA was labeled with the reverse primers. The two-step PCR not only enhanced amplification efficiency but also generated single-stranded PCR products for hybridization. As little as 1 ng of DNA or 0.4 CFU/ml before filtering and culturing enrichment can be detected by the microarray. In the case of water samples obtained from the environment, the bacterial concentration, especially in bottled water, can be as low as $<$0.1 CFU/ml, which falls below the threshold of detection by immune- or molecular-analysis-based assays. To overcome this problem, a two-step filtering and culturing process was used here to enrich the target bacteria. In the first step, vacuum filtering was used to collect all of the bacteria in the samples, and in the second step, the filter membrane was treated with diluted HCl and plated on the selective culture medium BCYE. The acid treatment used is based on the acid resistance feature of *Legionella* spp. By this method, *L. pneumophila* could be detected in as little as 2 or 3 days, 9 to 10 days earlier than by the traditional detection and identification method (ISO 11731:1998).

The DNA microarray method described here is specific and sensitive and offers a promising tool for wide applications in basic microbiological research and epidemiological surveillance. The method described here will be further verified with various clinical samples of bronchoalveolar lavage fluid and sputum and other samples, as well as environmental water samples, in the next stage of our study. The microarray has many advantages over traditional bacterial culture and serotyping methods; (i) it allows high sample throughput, (ii) it is able to detect O rough strains raised as a consequence of mutations in the O-antigen gene cluster, and (iii) cross-reactions are less common.

In conclusion, this study presents a new multiplex PCR-based microarray assay for the comprehensive detection and identification of all 15 *L. pneumophila* O serogroups. This new method provides an accurate and reliable approach by which to differentiate *L. pneumophila* isolates at the serogroup level, contributes significantly to large-scale epidemiology studies, and could be used to monitor local, regional, and national trends in human legionellosis.

ACKNOWLEDGMENTS

This work was supported by grants from the National Key Program for Infectious Diseases of China (2013ZX10004216-001-001), the National 863 Program of China (2012AA020103, 2011AA100901-2), the National 973 Program of China (2011CB504900), the National Natural Science Foundation of China (NSFC) Program (81171524, 31270144, 31170094, 31030002, and 31270003), and the Tianjin Research Program of Application Foundation and Advanced Technology (12JCYBJC15100 and 10JCYBJC10000).

REFERENCES

1. **Fraser DW, Tsai TR, Orenstein W, Parkin WE, Beecham HJ, Sharrar RG, Harris J, Mallison GF, Martin SM, McDade JE, Shepard CC, Brachman PS.** 1977. Legionnaires' disease: description of an epidemic of pneumonia. N. Engl. J. Med. **297:**1189–1197.
2. **Carvalho FR, Nastasi FR, Gamba RC, Foronda AS, Pellizari VH.** 2008. Occurrence and diversity of Legionellaceae in polar lakes of the Antarctic peninsula. Curr. Microbiol. **57:**294–300.
3. **Parthuisot N, West NJ, Lebaron P, Baudart J.** 2010. High diversity and abundance of *Legionella* spp. in a pristine river and impact of seasonal and anthropogenic effects. Appl. Environ. Microbiol. **76:**8201–8210.
4. **Fields BS, Benson RF, Besser RE.** 2002. *Legionella* and Legionnaires' disease: 25 years of investigation. Clin. Microbiol. Rev. **15:**506–526.
5. **Marston BJ, Plouffe JF, File TM, Jr, Hackman BA, Salstrom SJ, Lipman HB, Kolczak MS, Breiman RF.** 1997. Incidence of community-acquired pneumonia requiring hospitalization. Results of a population-based active surveillance study in Ohio. The Community-Based Pneumonia Incidence Study Group. Arch. Intern. Med. **157:**1709–1718.
6. **Muder RR, Yu VL.** 2002. Infection due to *Legionella* species other than *L. pneumophila*. Clin. Infect. Dis. **35:**990–998.
7. **Doleans A, Aurell H, Reyrolle M, Lina G, Freney J, Vandenesch F, Etienne J, Jarraud S.** 2004. Clinical and environmental distributions of *Legionella* strains in France are different. J. Clin. Microbiol. **42:**458–460.
8. **Qin T, Yan G, Ren H, Zhou H, Wang H, Xu Y, Zhao M, Guan H, Li M, Shao Z.** 2013. High prevalence, genetic diversity and intracellular growth ability of *Legionella* in hot spring environments. PLoS One **8:**e59018. doi: 10.1371/journal.pone.0059018.
9. **Maiwald M, Helbig JH, Luck PC.** 1998. Laboratory methods for the diagnosis of *Legionella* infections. J. Microbiol. Methods **33:**59–79.
10. **Murdoch DR.** 2003. Diagnosis of *Legionella* infection. Clin. Infect. Dis. **36:**64–69.
11. **Guan W, Xu Y, Chen DL, Xu JN, Tian Y, Chen JP.** 2012. Application of multilocus sequence analysis (MLSA) for accurate identification of *Legionella* spp. isolated from municipal fountains in Chengdu, China, based on 16S rRNA, *mip*, and *rpoB* genes. J. Microbiol. **50:**127–136.
12. **Zhou G, Cao B, Dou Y, Liu Y, Feng L, Wang L.** 2011. PCR methods for the rapid detection and identification of four pathogenic *Legionella* spp. and two *Legionella pneumophila* subspecies based on the gene amplification of *gyrB*. Appl. Microbiol. Biotechnol. **91:**777–787.
13. **Zhou G, Wen S, Liu Y, Li R, Zhong X, Feng L, Wang L, Cao B.** 2011. Development of a DNA microarray for detection and identification of *Legionella pneumophila* and ten other pathogens in drinking water. Int. J. Food Microbiol. **145:**293–300.
14. **Su HP, Tung SK, Tseng LR, Tsai WC, Chung TC, Chang TC.** 2009. Identification of *Legionella* species by use of an oligonucleotide array. J. Clin. Microbiol. **47:**1386–1392.
15. **Cunningham SA, Sloan LM, Uhl JA.** 2009. Validation of a real-time PCR assay for the detection of *Legionella* species in respiratory samples. Abstracts of the Annual Meeting of the Association for Molecular Pathology, General Meeting, Kissimmee, FL, 19 to 22 November 2009.
16. **Guo H, Yi W, Song JK, Wang PG.** 2008. Current understanding on biosynthesis of microbial polysaccharides. Curr. Top. Med. Chem. **8:**141–151.
17. **Knirel YA, Rietschel ET, Marre R, Zahringer U.** 1994. The structure of the O-specific chain of *Legionella pneumophila* serogroup 1 lipopolysaccharide. Eur. J. Biochem. **221:**239–245.
18. **Lüneberg E, Zetzmann N, Alber D, Knirel YA, Kooistra O, Zähringer U, Frosch M.** 2000. Cloning and functional characterization of a 30 kb gene locus required for lipopolysaccharide biosynthesis in *Legionella pneumophila*. Int. J. Med. Microbiol. **290:**37–49.
19. **Bronner D, Clarke BR, Whitfield C.** 1994. Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. Mol. Microbiol. **14:**505–519.
20. **Guo D, Bowden MG, Pershad R, Kaplan HB.** 1996. The *Myxococcus xanthus rfbABC* operon encodes an ATP-binding cassette transporter homolog required for O-antigen biosynthesis and multicellular development. J. Bacteriol. **178:**1631–1639.
21. **Kelly RF, Whitfield C.** 1996. Clonally diverse *rfb* gene clusters are involved in expression of a family of related D-galactan O-antigens in *Klebsiella* species. J. Bacteriol. **178:**5205–5214.
22. **Li Y, Cao B, Liu B, Liu D, Gao Q, Peng X, Wu J, Bastin DA, Feng L, Wang L.** 2009. Molecular detection of all 34 distinct O-antigen forms of *Shigella*. J. Med. Microbiol. **58:**69–81.
23. **Li Y, Liu D, Cao B, Han W, Liu Y, Liu F, Guo X, Bastin DA, 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.
24. **Ren Y, Liu B, Cheng J, Liu F, Feng L, Wang L.** 2008. Characterization of *Escherichia coli* O3 and O21 O antigen gene clusters and development of serogroup-specific PCR assays. J. Microbiol. Methods **75:**329–334.
25. **Wang L, Liu B, Kong Q, Steinrück H, Krause G, Beutin L, Feng L.** 2005. Molecular markers for detection of pathogenic *Escherichia coli* strains belonging to serogroups O138 and O139. Vet. Microbiol. **111:**181–190.
26. **Wang L, Reeves PR.** 1998. Organization of *Escherichia coli* O157 O antigen gene cluster and identification of its specific genes. Infect. Immun. **66:**3545–3551.
27. **Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, et al.** 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456:**53–59.
28. **Shallom SJ, Tae H, Sarmento L, Preston D, McIver L, Franck C, Dickerman A, Adams LG, Garner HR.** 2012. Comparison of genome diversity of *Brucella* spp. field isolates using Universal Bio-signature Detection Array and whole genome sequencing reveals limitations of current diagnostic methods. Gene **509:**142–148.
29. **Guo D, Liu B, Liu F, Cao B, Chen M, Hao X, Feng L, Wang L.** 2013. Development of a DNA microarray for molecular identification of all 46 *Salmonella* O serogroups. Appl. Environ. Microbiol. **79:**3392–3399.
30. **Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L.** 2008. Structure and genetics of *Shigella* O antigens. FEMS Microbiol. Rev. **32:**627–653.
31. **Kim S, Frye JG, Hu J, Fedorka-Cray PJ, Gautom R, Boyle DS.** 2006. Multiplex PCR-based method for identification of common clinical serotypes of *Salmonella enterica* subsp. *enterica*. J. Clin. Microbiol. **44:**3608–3615.
32. **Sun Y, Wang M, Liu H, Wang J, He X, Zeng J, Guo X, Li K, Cao B, Wang L.** 2011. Development of an O-antigen serotyping scheme for *Cronobacter sakazakii*. Appl. Environ. Microbiol. **77:**2209–2214.
33. **Sun Y, Wang M, Wang Q, Cao B, He X, Li K, Feng L, Wang L.** 2012. Genetic analysis of the *Cronobacter sakazakii* O4 to O7 O-antigen gene clusters and development of a PCR assay for identification of all *C. sakazakii* O serotypes. Appl. Environ. Microbiol. **78:**3966–3974.