The Journal of Microbiology (2012) Vol. 50, No. 1, pp. 127–136
Copyright ⓒ 2012, The Microbiological Society of Korea
DOI 10.1007/s12275-012-1243-1

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

Wang Guan[1], Ying Xu[1,2], Da-li Chen[1], Jia-nan Xu[1], Yu Tian[1], and Jian-ping Chen[1,3*]

[1]*Department of Parasitology, West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, Sichuan, P. R. China*
[2]*Department of Clinical Laboratories, the First Affiliated Hospital of Chengdu Medical College, Chengdu, 610500, Sichuan, P. R. China*
[3]*Animal Disease Prevention and Food Safety Key Laboratory of Sichuan Province, Sichuan University, Chengdu 610064, Sichuan, P. R. China*

(Received May 16, 2011 / Accepted August 31, 2011)

Legionellosis (Legionnaires' disease; LD) is a form of severe pneumonia caused by species of *Legionella* bacteria. Because inhalation of *Legionella*-contaminated aerosol is considered the major infection route, routine assessments of potential infection sources such as hot water systems, air-conditioner cooling water, and municipal fountains are of great importance. In this study, we utilized *in vitro* culture and multilocus sequence analysis (MLSA) targeting 16S rRNA, *mip*, *rpoB*, and *mip-rpoB* concatenation to isolate and identify *Legionella* spp. from 5 municipal fountains in Chengdu City, Sichuan Province, China. Our results demonstrated that 16S rRNA was useful for initial identification, as it could recognize isolates robustly at the genus level, while the genes *mip*, *rpoB*, and *mip-rpoB* concatenation could confidently discriminate *Legionella* species. Notably, the three subspecies of *L. pneumophila* could be distinguished by the analysis based on *rpoB*. The serotyping result of strain CD-1 was consistent with genetic analysis based on the concatenation of *mip* and *rpoB*. Despite regular maintenance and sanitizing methods, 4 of the 5 municipal fountains investigated in this study were positive for *Legionella* contamination. Thus, regularly scheduled monitoring of municipal fountains is urgently needed as well as vigilant disinfection. Although the application of MLSA for inspection of potential sites of infection in public areas is not standard procedure, further investigations may prove its usefulness.

Keywords: *Legionella*, MLSA, 16S rRNA, *mip*, *rpoB*, municipal fountain

Introduction

To date, there are 55 valid species in the bacterial genus *Legionella* (http://www.ncbi.nlm.nih.gov/Taxonomy, accessed July 25, 2011). Among them, approximately 20 are involved in human diseases. *Legionella* infection is mainly caused by inhalation of contaminated aerosol, leading to flu-like Pontiac fever or the more serious legionellosis, also known as Legionnaires' disease (LD). The latter is a form of severe pneumonia with a fatality rate that may reach 50% in immunocompromised patients (Fields *et al*., 2002; Yu *et al*., 2002). One species, *Legionella pneumophila*, is responsible for about 90% of LD cases. The other species are rarely pathogenic, although *Legionella longbeachae* accounts for about 30% of cases of LD in Australia and New Zealand (Helbig *et al*., 2002; Newton *et al*., 2010).

Since air-conditioner cooling water was identified as the source of contamination in the first LD outbreak in Philadelphia, 1976, numerous studies have shown that *Legionella* is ubiquitous in various water environments, residing in free-living amoebae and ciliated protozoa (Tyndall and Domingue, 1982; Anand *et al*., 1983; Fields *et al*., 1984; Rowbotham, 1986). LD outbreaks have been associated with all kinds of water resources (Hoge and Breiman, 1991; Cooper *et al*., 2008), even family hot water systems, asphalt paving machines, and whirlpool spas (Mathys *et al*., 2008; Coscolla *et al*., 2010; Euser *et al*., 2010). It was reported that a cluster of LD cases was linked to a decorative water fountain, despite standard maintenance and sanitizing methods (Palmore and Stalk, 2009). Thus, it is of great importance to assess and monitor the *Legionella*-contamination status of municipal fountain systems.

Currently there are a number of methods for detection of *Legionella* in environmental samples, such as competitive polymerase chain reaction (PCR; Declerk *et al*., 2005) and urinary antigen enzyme immunoassay (Blanco *et al*., 2008). However, accurate identification of *Legionella* species can be quite difficult due to the serological cross-reactivity among serogroups and species, biochemical inertness, and similar phenotypic identity across different species (Edelstein, 2007).

More rapid and precise identification of *Legionella* spp. is provided by sequence analysis. The 16S rRNA gene is the most utilized genetic marker to infer taxonomic relationships among species, as it is one of the most conserved genes that have been identified (Brenner, 1986). Several phylogenic studies of the genus *Legionella* based on 16S rRNA showed that most species in this genus could be distinguished correctly (Fry *et al*., 1991; Wilson *et al*., 2007). However, it has been proven that heterogeneous copies of 16S rRNA can occur in the genome, resulting in misidentifications at the species level (Clayton *et al*., 1995; Ueda *et al*., 1999), and it

*For correspondence. E-mail: jpchen007@163.com; Fax: +86-28-85501248

is so highly conserved that some species cannot be discriminated by 16S rRNA. For example, the 16S rRNAs of *L. erythra* and *L. rubrilucens* are 99.4% similar to each other (Ratcliff *et al.*, 1998; Ko *et al.*, 2002; Rubin *et al.*, 2005).

The DNA sequences of the macrophage infectivity potentiator (*mip*) and RNA polymerase subunit β (*rpoB*) genes have been proposed for analyzing phylogenic relationships and differentiating among *Legionella* species. The *mip* gene encodes an immunophilin of the FK506-binding protein class, which is involved in the ability of *L. pneumophila* to survive uptake into phagocytic cells (Cianciotto *et al.*, 1989; Hacker and Fischer, 1993; O'Connell *et al.*, 1996; Doyle *et al.*, 1998). The *mip* locus is considered an ideal candidate for a classification scheme, because it has shown greater mutational variation than 16S rRNA and is relatively stable (Ratcliff *et al.*, 1998; Sussman, 2002). However, some species and some environmental isolates could not be confidently discriminated by the *mip*-based scheme, such as *L. geestiana* and European wild strain LC4381 (Ratcliff *et al.*, 1998).

The *rpoB* gene encodes the β subunit of DNA-dependent RNA polymerase. Mutation in a particular region of *rpoB* is related to rifampin resistance (Severinov *et al.*, 1996; Kim *et al.*, 1999). Various studies have proven that *rpoB*-based analysis could effectively overcome the intrinsic limitations of the intraspecies heterogeneity of 16S rRNA (Mollet *et al.*, 1997; Dahllof *et al.*, 2000). It has also been proven that *rpoB* sequence analysis could clearly differentiate among *Legionella* species (Ko *et al.*, 2002). However, the phylogenic tree of *rpoB* showed a quite different topology from the tree of 16S rRNA and *mip*, which may be due to horizontal gene transfer between *Legionella* species (Holmes *et al.*, 1999).

Each of the loci described above has its respective advantages and disadvantages when utilized for species identification and phylogenic inference. However, the combination of the three loci improves identification and can be used in a multilocus sequence analysis (MLSA) approach (Rubin *et al.*, 2005; Gomes-Valero *et al.*, 2009). Thus, MLSA has been recommended for *Legionella* spp. identification and phylogenic analysis (Gomes-Valero *et al.*, 2009).

In this study, we took water samples from 5 municipal fountains in Chengdu City, Sichuan Province, China to determine the status of *Legionella* contamination and isolated *Legionella* spp. from 4 of them. The 16S rRNA, *mip* and *rpoB* sequences of the *Legionella* isolates were identified. The 16S rRNA sequence was used for initial identification, while those of *mip* and *rpoB* were separately used for further analysis and then concatenated.

Materials and Methods

Sample collection

In July 2010, 50 water samples were collected from 5 municipal fountains in Chengdu City, Sichuan Province, China. All 5 municipal fountains, termed A, B, C, D, E, were located in the urban area with a high population density. Sample collection was performed strictly according to the protocol proposed by Barabee (1987). Briefly, one-liter bottles that had been autoclaved at 126°C for 25 min were used. At each fountain, 10 samples were taken. Each one-liter water sample was directly collected from the fountain stream, and then delivered to the laboratory immediately, avoiding heat and light. Each sample was filtered through a 0.22 μm nitrocellulose membrane under vacuum. Each membrane was washed in 20 ml KCl-HCl buffer (pH 2.2) for 5 min, and then 200 μl of the buffer was transferred onto a plate with glycine (3 g/L), vancomycin (0.001 g/L), polymycin B (80,000 IU/L), and cycloheximide (0.08 g/L) and incubated for 10 d at 5% CO_2, and 37°C.

Characterization of isolates

After incubation, isolates from colonies with a superficial resemblance to *Legionella* morphologically were Gram-stained, and observed under a microscope. Short, rod-like Gram-negative bacilli were subcultured onto buffered charcoal yeast extract (BCYE, specific for *Legionella*) agar and blood agar (which cannot support *Legionella* for lack of L-cysteine), and then incubated for another 3 d. Isolates able to grow on BCYE agar but not on blood agar were transferred onto BCYE agar for serotyping and DNA extraction. Serotyping was performed using a Dryspot *Legionella* species Test Kit (Oxoid, UK). Tests for β-lactamase and autofluorescence were also performed.

DNA extraction

The total DNA of each isolate was extracted using bacterial DNA extraction kits (Omega, USA). Briefly, purified isolates in culture were suspended in Tris-ethylenediaminetetra-acetic acid (TE) buffer, and then lysed in lysozyme, proteinase K, and RNase. Lysate other than DNA was removed through a series of centrifugation and resuspension steps. The final product was 150 μl of eluate containing DNA, which was stored in Eppendorf tubes at -20°C.

PCR protocol and sequence determination

PCR analysis was performed to determine the gene sequences of 16S rRNA, *mip*, and *rpoB* of the isolates. The primers have been described previously (Ratcliff *et al.*, 1998; Ko *et al.*, 2002; Kuroki *et al.*, 2007). All primers were purchased from Majorbio (China), and dissolved in nuclease-free water to give a working concentration of 10 pmole. The total volume of the PCR mix was 50 μl, consisting of 2 μl template, 2 μl of each primer, 19 μl ddH_2O, 24 μl 2× Golden Fast Reaction Mix, and 1 μl Golden Fast DNA Polymerase (Tiangen, China).

The PCR conditions for each locus were slightly different from the references. For 16S rRNA, the PCR conditions were 95°C for 5 min, followed by 35 cycles of 95°C for 15 sec, 55°C for 5 sec, and 72°C for 25 sec, and a final elongation step at 72°C for 3 min. For *mip*, the PCR conditions were 96°C for 5 min, followed by 35 cycles of 94°C for 15 sec, 58°C for 5 sec, and 72°C for 15 sec, and a final elongation step at 72°C for 5 min. For *rpoB*, the PCR conditions were 95°C for 8 min, followed by 35 cycles of 93°C for 15 sec, 54°C for 5 sec, and 72°C for 15 sec, and a final elongation step at 72°C for 10 min. PCR reactions were conducted in a Peltier Thermal Cycler (Bio-Rad, USA).

Commercial DNA purification kits (Tiangen) were used

Table 1. Strain passport and GenBank accession number of reference strains used in this study

Strain	Strain passport	GenBank accession no.		
		16S rRNA	*mip*	*rpoB*
L. adelaidensis	ATCC[a] 49625	Z49716	U91606	AF367721
L. anisa	ATCC 35292	Z32635	U91607	AF367722
L. birminghamensis	ATCC 43702	Z49717	U91608	AF367723
L. bozemanae	ATCC 33217	Z49719	U91609	AF367724
L. brunensis	ATCC 43878	Z32636	U92227	AF367725
L. cherrii	ATCC 35252	Z49720	U92235	AF367726
L. cincinnatiensis	ATCC 43753	Z49721	U92236	AF367727
L. dumoffii	ATCC 33279	Z32637	U92237	AF367728
L. erythra	ATCC 35303	Z32638	U92203	AF367729
L. fairfieldensis	ATCC 49588	Z49722	U92204	AF367730
L. feeleii	ATCC 35072	X73395	U92205	AF367731
L. geestiana	ATCC 49504	Z49723	ND[b]	AF367732
L. gormanii	ATCC 33297	Z32639	U91638	AF367733
L. gratiana	ATCC 49413	Z49725	U92206	AF367734
L. hackeliae	ATCC 35250	M36028	U92207	AF367735
L. israelensis	ATCC 43119	Z32640	U92208	AF367736
L. jamestowniensis	ATCC 35298	Z49726	U92228	AF367737
L. jordanis	ATCC 33623	Z32667	U92209	AF367738
L. lansingensis	ATCC 49751	Z49727	U92210	AF367739
L. londiniensis	ATCC 49505	Z49728	U92229	AF367740
L. longbeachae SG1[c]	ATCC 33462	AY444740	X83036	AF367741
L. longbeachae SG2	ATCC 33484	AY444741	AF000968	AF489908
L. maceachernii	ATCC 35300	X60081	U92211	AF367742
L. micdadei	ATCC 33218	AF227162	S62141	AF367743
L. moravica	ATCC 43877	Z49729	U92212	AF367744
L. nautarum	ATCC 49506	Z49730	U92213	AF367745
L. oakridgensis	ATCC 33761	X73397	U92214	AF367746
L. parisiensis	ATCC 35299	Z49731	U92215	AF367747
L. pneumophila subsp. *pneumophila* SG1	ATCC 33152	M59157	AF022334	AF367748
L. pneumophila subsp. *pneumophila* SG2	ATCC 33154	ND	AF022316	AY036039
L. pneumophila subsp. *pneumophila* SG3	ATCC 33155	ND	AF095227	AY036040
L. pneumophila subsp. *fraseri* SG4	ATCC 33156	M36025	AF022318	AY036041
L. pneumophila subsp. *fraseri* SG5	ATCC 33216	ND	AF022319	AY036042
L. pneumophila subsp. *pneumophila* SG6	ATCC 33215	ND	AF022320	AY036043
L. pneumophila subsp. *pneumophila* SG7	ATCC 33823	ND	AF022321	AY036044
L. pneumophila subsp. *pneumophila* SG8	ATCC 35096	ND	AF022322	AY036045
L. pneumophila subsp. *pneumophila* SG9	ATCC 35289	ND	AF022323	AY036046
L. pneumophila subsp. *pneumophila* SG10	ATCC 43283	ND	AF022324	AY036047
L. pneumophila subsp. *pneumophila* SG11	ATCC 43130	ND	AF022325	AY036048
L. pneumophila subsp. *pneumophila* SG12	ATCC 43290	ND	AF022326	AY036049
L. pneumophila subsp. *pneumophila* SG13	ATCC 43736	ND	AF022327	AY036050
L. pneumophila subsp. *pneumophila* SG14	ATCC 43703	ND	AF022328	AY036051
L. pneumophila subsp. *pascullei*	ATCC 33737	AF122885	ND	AJ746052
L. quinlivanii	ATCC 43830	Z49733	U92217	AF367749
L. rubrilucens	ATCC 35304	Z32643	U92218	AF367750
L. sainthelensi	ATCC 35248	Z49734	U92219	AF367751
L. santicrucis	ATCC 35301	Z49735	U92220	AF367752
L. shakespearei	ATCC 49655	Z49736	U92221	AF367753
L. spiritensis	ATCC 35249	M36030	U92222	AF367754
L. steigerwaltii	ATCC 35302	Z49737	U92223	AF367755
L. tucsonensis	ATCC 49180	Z32644	U92224	AF367756
L. wadsworthii	ATCC 33877	Z49738	U92225	AF367757
L. worsleiensis	ATCC 49508	Z49739	U92226	AF367758
Coxiella burnetii	ATCC VR-615	AE016828	U14170	U86688

[a] ATCC, American Type Culture Collection, Manassas, VA, USA.
[b] ND, No data in the GenBank.
[c] SG, serogroup

to resolve the PCR products on 2% agarose gel stained with ethidium bromide, according to the manufacturer's protocol. DNA sequencing was performed with ABI BigDye Terminator chemistry on an ABI 3730 automated sequencer. The determined sequences were then submitted to GenBank.

Sequence alignment and analysis

The sequences of 53 reference *Legionella* strains were retrieved from GenBank. *Coxiella burnetii*, the pathogen of Q fever, served as an out-group. GenBank accession numbers are listed in Table 1. All sequences were aligned by CLUSTAL X (Thompson *et al*., 1997) with default gap penalties. MEGA 4.0 (Tamura *et al*., 2007) was used for minor adjustments to the aligned matrix and concatenation of the aligned matrices of *mip* and *rpoB*. MEGA 4.0 was also used to calculate the p-distance within each aligned matrix. Compositional heterogeneity was evaluated using chi-square ($\chi 2$) tests implemented in PAUP*4.0b10 (Swofford, 2002)

Bayesian phylogenic analyses

To conduct Bayesian analysis, the Akaike Information Criterion (Akaike, 1974), following its recent approbation (Posada and Buckley, 1998, 2004), was applied using Modeltest 3.7 to choose the best-fit model of nucleotide substitution. MrBayes v3.2 was used to conduct the Bayesian analysis (Ronquist and Huelsenbeck, 2003). Posterior probability (PP) distribution was estimated by allowing four incrementally heated Markov chains to proceed four million times, with samples taken from every 200 generations. To ensure our analyses were not restricted from the global optimum, analyses were repeated beginning with a different starting tree (Huelsenbeck *et al*., 2002). The first one million generations were discarded

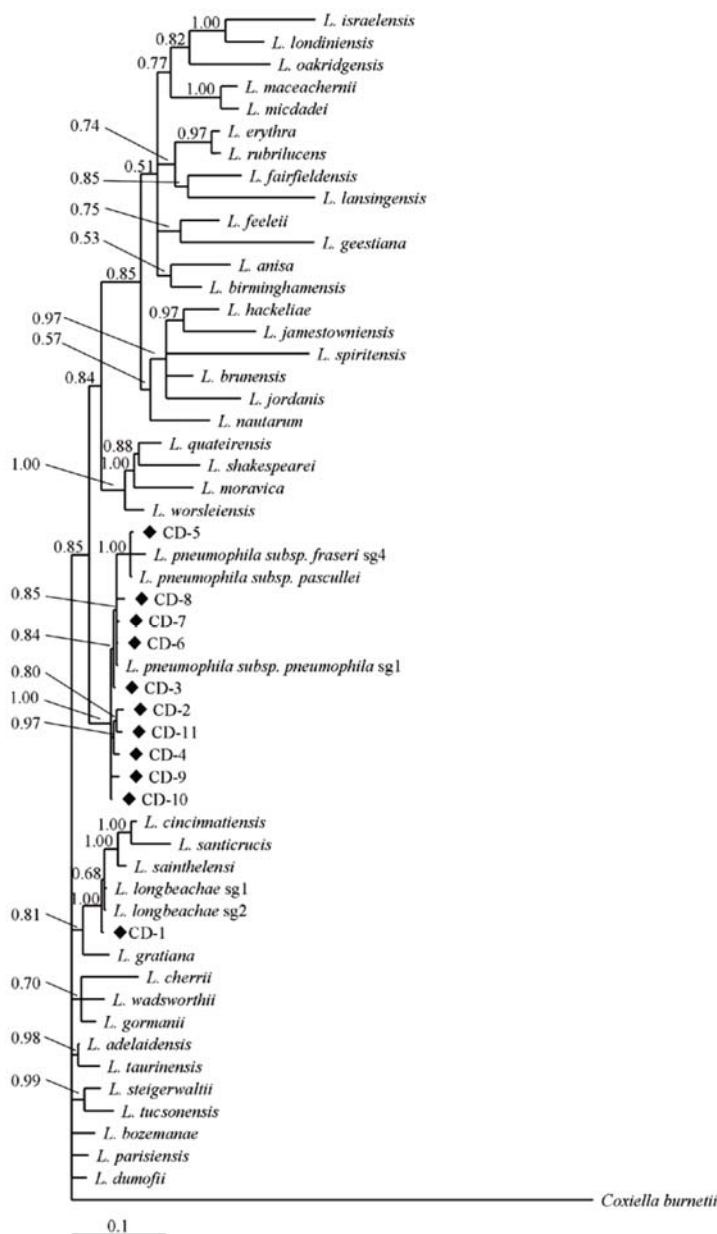

Fig. 1. The 50% majority-rule consensus tree from Bayesian inference of the 16S rRNA dataset using MrBayes v.3.2. Bayesian posterior probabilities are shown at nodes. The taxa tagged with (◆) indicate the isolates of this study.

in case this chain reached a stationary distribution, and the remaining samples from the independent runs were pooled to obtain the final approximation of the posterior distribution of trees. The posterior distribution was summarized as a 50% majority-rule consensus to form a robust phylogeny. The results of Bayesian analyses were accessed with Treeview v1.6.6 (Page, 1996).

Results

Isolate characterization

Ten days after incubation, the colonies of 11 strains (2, 4, 3, and 2 from fountains B, C, D, and E, respectively) morphologically similar to *Legionella* and with strict L-cysteine dependence were isolated. These were termed CD-1 to CD-11. All isolates were positive for β-lactamase but negative for autofluorescence. Serotyping tests showed that strain CD-1 reacted strongly with *L. longbeachae* serogroup 1, and strains CD-2, -3, -4, -6, -7, and -8 reacted strongly with *L. pneumophila* serogroup 1. CD-5 reacted strongly with serogroup 5. CD-9, -10, and -11 reacted strongly with *L. pneumophila* serogroups 3, 2, and 10, respectively. However, both CD-9 and CD-10 reacted relatively weakly with *L. pneumophila* serogroup 1.

Sequence analysis

After DNA extraction, the amplicons of 16S rRNA, *mip*,

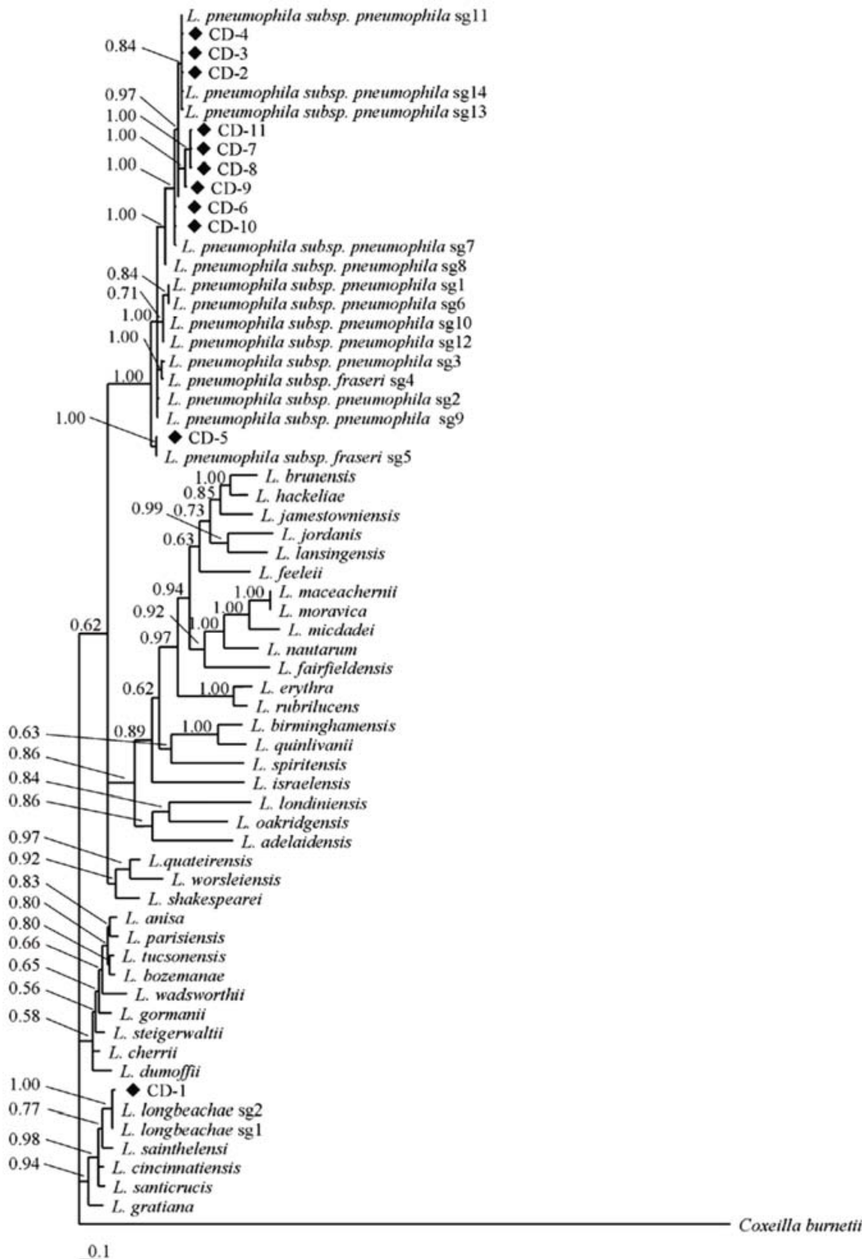

Fig. 2. The 50% majority-rule consensus tree from Bayesian inference of the *mip* dataset using MrBayes v.3.2. Bayesian posterior probabilities are shown at nodes. The taxa tagged with (◆) indicate the isolates of this study.

and *rpoB* of the 11 isolates were obtained. The length of sequences ranged from 1,366 to 1,432 bp for 16S rRNA, 585 to 619 bp for *mip*, and 303 to 334 bp for *rpoB*. All of the sequences were deposited in GenBank, with accession numbers HQ645020-HQ645030 for 16S rRNA, HQ645031-HQ645041 for *mip*, and HQ645042-HQ645052 for *rpoB*. The sequences were aligned by CLUSTAL X with sequences of the type strain retrieved from GenBank. A base stationarity test showed insignificant differences among all taxa in base composition bias: 16S rRNA, χ^2=23.18, df=159, P=1.00; *mip*, χ^2=115.46, df=189, P=0.99; *rpoB*, χ^2=90.72, df=192, P=1.00; and *mip-rpoB*, χ^2=147.01, df=186, P=0.98. Prior to the Bayesian analysis, the most adequate model of evolution, the nucleotide substitution models GTR+I+G for 16S rRNA, *mip* and *rpoB*, and TVM+I+G for *mip-rpoB*, were selected by Modeltest 3.7.

Phylogenic relationships

For the aligned matrix of 16S rRNA, the overall mean p-distance was 0.056. In the phylogenic tree inferred from this matrix (Fig. 1), strain CD-1 formed a clade with *L. longbeachae* serogroup (SG)1 and SG2, *L. santhilensi*, *L. santicrucis*, and *L. cincinnatiensis*, supported by the PP 1.00. The p-distance was 0.001 to *L. longbeachae* SG1 and SG2, 0.010 to *L. santhilensi*, 0.031 to *L. santicrucis*, and 0.014 to *L. cincinnatiensis*. The other 10 isolates formed a group with *L. pneumophila* supported by the PP 1.00; overall mean p-distance of the group was 0.006.

In the phylogenic tree inferred from the 16S rRNA matrix (Fig. 1), within the clade of *L. pneumophila*, strain CD-5 was in the clade of *L.* ssp. *pascullei* and *L.* ssp. *fraseri* supported by the PP 1.00. The p-distance was 0.006 to *L.* ssp. *fraseri*, 0.001 to *L.* ssp. *pascullei*, but 0.007 to *L.* ssp. *pneu-*

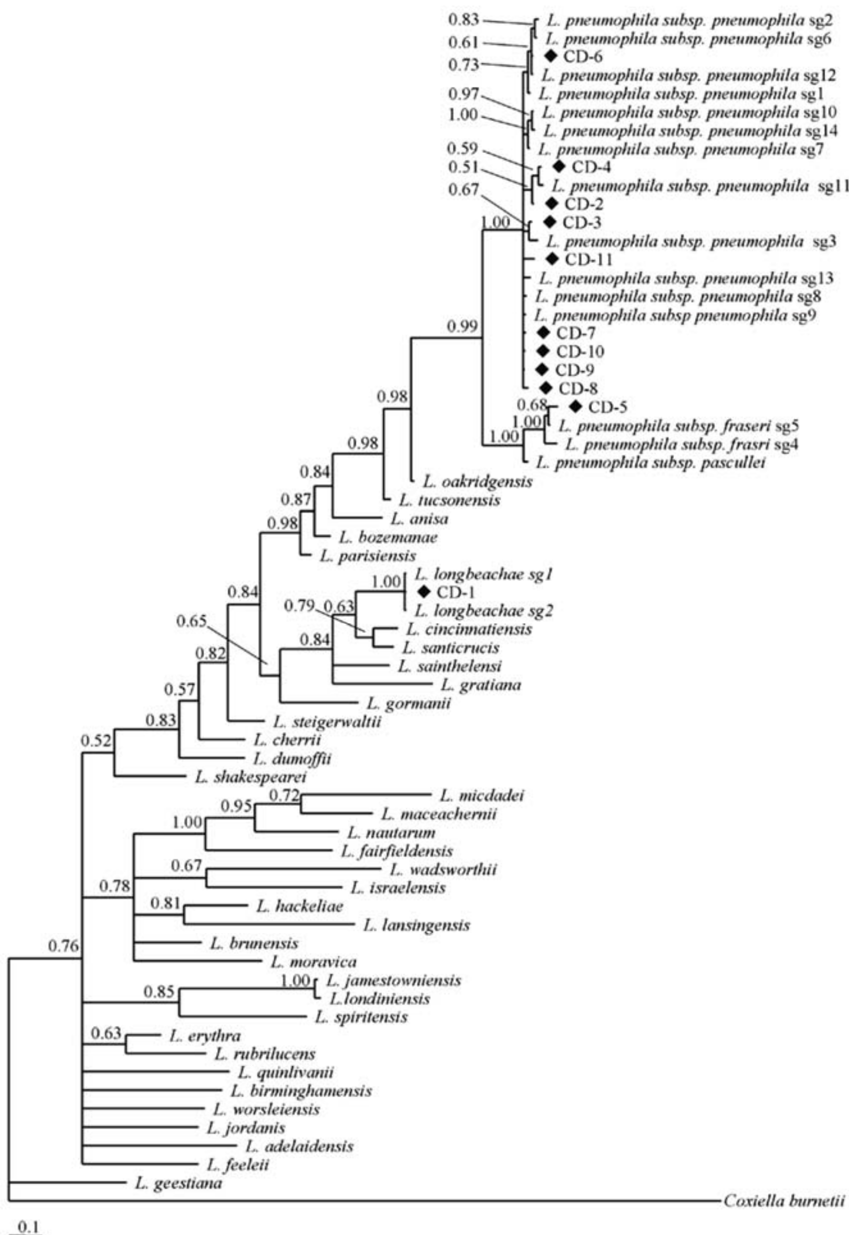

Fig. 3. The 50% majority-rule consensus tree from Bayesian inference of the *rpoB* dataset using MrBayes v.3.2. Bayesian posterior probabilities are shown at nodes. The taxa tagged with (◆) indicate the isolates of this study.

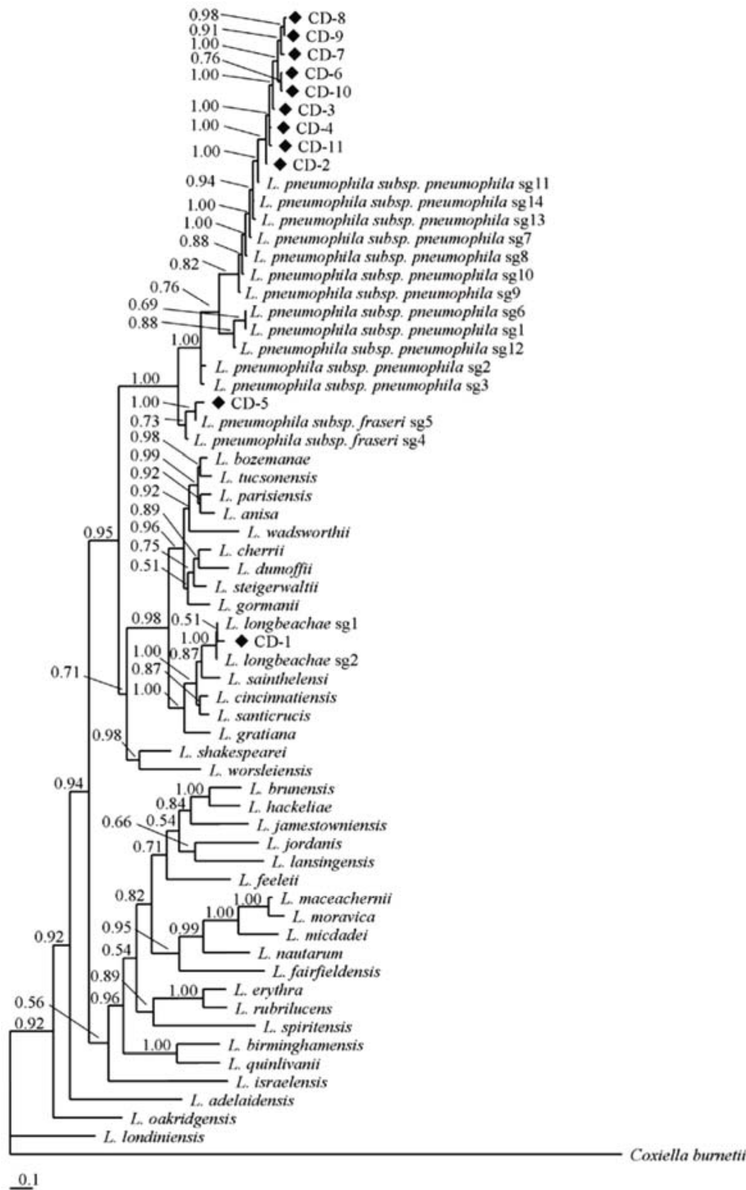

Fig. 4. The 50% majority-rule consensus tree from Bayesian inference of the *mip-rpoB* combined dataset using MrBayes v.3.2. Bayesian posterior probabilities are shown at nodes. The taxa tagged with (◆) indicate the isolates of this study.

mophila. However, the p-distances of the other 9 isolates to *L.* ssp. *pneumophila* ranged from 0.001 (CD-6) to 0.004 (CD-2), to *L.* ssp. *fraseri* ranged from 0.009 (CD-3) to 0.013 (CD-2), and to L. ssp. *pascullei* ranged from 0.007 (CD-3) to 0.011 (CD-2).

For the aligned matrix of *mip*, the overall mean p-distance was 0.361. In the phylogenic tree inferred from this matrix (Fig. 2), strain CD-1 formed a branch within the clade of *L. longbeachae* SG1 (p-distance: 0.012) and SG2 (p-distance: 0.015) supported by the PP 1.00. The other 10 isolates formed a group with *L. pneumophila* supported by the PP 1.00; the overall mean p-distance of this group was 0.022.

In the phylogenic tree inferred from the *mip* matrix (Fig. 2), in the clade of *L. pneumophila*, two distinct clades were formed. Strain CD-5 was in the clade of *L.* ssp. *fraseri* SG5 (p-distance: 0.002) supported by the PP 1.00. However, the p-distance of CD-5 to the other type strains of different serogroups ranged from 0.042 (*L.* ssp. *pneumophila* SG1) to 0.058 (*L.* ssp. *pneumophila* SG14). The other 9 isolates were in the clade of *L.* ssp. *pneumophila* supported by the PP 1.00. Notably, the type strain *L.* ssp. *fraseri* SG4 formed a group with *L.* ssp. *pneumophila* SG3 (p-distance: 0.005) supported by the PP 1.00, but not with *L.* ssp. *fraseri* SG5 (p-distance: 0.048).

For the aligned matrix of *rpoB*, overall mean p-distance was 0.239. In the phylogenic tree inferred from this matrix (Fig. 3), strain CD-1 formed a branch within the clade of *L. longbeachae* SG1 and SG2 supported by the PP 1.00 (p-distance: 0.001). The other 10 isolates formed a group with *L. pneumophila* supported by the PP 0.99; the overall mean p-distance of this group was 0.059.

In the phylogenic tree inferred from the *rpoB* matrix (Fig. 3), two distinct clades were formed within the clade *L. pneumophila*. Strain CD-5 was in the clade of *L.* ssp. *fraseri* supported by the PP 1.00, and formed a group with *L.* ssp. *fraseri* SG5 supported by the PP 0.68. The p-distance was

0.039 to L. ssp. *fraseri* SG4 and 0.015 to L. ssp. *fraseri* SG5. However, the p-distance of CD-5 to *L.* ssp. *pascullei* was 0.062, and to other type strains of *L.* ssp. *pneumophila* ranged from 0.153 (*L.* ssp. *pneumophila* SG8) to 0.182 (*L.* ssp. *pneumophila* SG2). The other 9 isolates formed a group with *L.* ssp. *pneumophila* supported by the PP 1.00; the overall mean p-distance of this group was 0.023.

For the combined matrix of *mip* and *rpoB*, the overall mean p-distance was 0.319. In the phylogenic tree inferred from this matrix (Fig. 4), strain CD-1 formed a branch within the clade of *L. longbeachae* supported by the PP 1.00. Notably, it formed a group with *L. longbeachae* SG1 supported by the PP 0.51. The p-distance was 0.023 to *L. longbeachae* SG1 and 0.025 to *L. longbeachae* SG2.

In the phylogenic tree inferred from the *mip* and *rpoB* combined matrix (Fig. 4), the other 10 isolates formed a group with *L. pneumophila* supported by the PP 1.00; the overall mean p-distance of this group was 0.039. Within the clade of *L. pneumophila*, two distinct clades were formed. Strain CD-5 was in the clade of *L.* ssp. *fraseri* supported by the PP 0.73, and formed a group with *L.* ssp. *fraseri* SG5 supported by the PP 1.00. The p-distance was 0.070 to *L.* ssp. *fraseri* SG4 and 0.027 to L. ssp. *fraseri* SG5. To other type strains of *L.* ssp. *pneumophila*, the p-distance ranged from 0.104 (*L.* ssp. *pneumophila* SG1) to 0.117 (*L.* ssp. *pneumophila* SG14). The other 9 isolates formed a group with *L.* ssp. *pneumophila* supported by the PP 1.00. The overall mean p-distance of this group was 0.024.

Discussion

Since the first outbreak of LD in 1976, artificial aquatic environments have been regarded as the most important source of infection. In 2009, it was reported to the European Surveillance Scheme for Travel-Associated Legionnaires' Disease that 49 environmental samples, taken from travelers' accommodations with clusters of LD cases, were found positive for *Legionella* (Joseph *et al*., 2010). The existence of *Legionella* in municipal fountains has been the subject of various studies, including a report of a cluster of LD cases linked to a decorative water fountain despite standard maintenance and sanitizing methods (Palmore and Stock, 2009). This highlights the potential threat of an LD outbreak originating from these sources.

In this study, we examined 50 water samples from 5 municipal fountains in Chengdu, China. An acidic buffer treatment (KCl-HCl, pH 2.2) and the selective culture media GVPC were combined to isolate *Legionella* from the samples. Eleven bacterial strains with similar morphology to *Legionella* were isolated. All of them were positive for β-lactamase activity, and negative for autofluorescence.

GVPC agar is a selective medium for environmental *Legionella* isolation (Bartie *et al*., 2003; Luck *et al*., 2004). However, there have been reports that some other species with similar colonial morphology and Gram-stain morphology to species of *Legionella* are able to grow on GVPC agar (Qu *et al*., 2009), compromising the correct isolation and identification of *Legionella*.

Serotyping is a classic method for strain identification. However, its reliability is also seriously hampered by crossactivity among serogroups, and even among different species (Bangsborg *et al*., 1991; Boswell, 1996). In this study, we also encountered the problem of cross-reactivity, as the isolated strains CD-8 and CD-9 both reacted with two different serogroups.

Because of the potential problems associated with the questionable selectivity of GVPC agar and cross-reactivity among serogroups, to ensure accurate identification of the 11 isolated strains we utilized MLSA, based on 16S rRNA, *mip*, *rpoB*, and the concatenation of *mip* and *rpoB*. The gene sequences of 16S rRNA, *mip*, and *rpoB* previously proved to be useful in molecular differentiation of *Legionella* spp. However, prior to the current study *mip* and *rpoB* had not been combined to differentiate within the genus *Legionella*.

Chi-square (χ2) tests of the four aligned matrices showed that there was no obvious compositional heterogeneity but sufficient variable sites, suggesting that they were suitable for discrimination and phylogenic analysis. Strain CD-1 was not confidently discriminated in the phylogenic tree inferred from the 16S rRNA matrix, although the p-distance was 0.001 to *L. longbeachae* and ranged from 0.010 (*L. santhilensi*) to 0.085 (*L.* geestiana) to the other *Legionella* species. However, CD-1 was clearly identified as *L. longbeachae* in the phylogenic trees inferred from *mip*, *rpoB*, and *mip*-*rpoB* matrices, each supported by the PP 1.00.

It is well known that genetic analyses rarely parallel serological analyses. Interestingly, CD-1 formed a clade with *L. longbeachae* SG1, supported by the PP 0.51 in the phylogenic tree inferred from the combined matrix of *mip* and *rpoB*, while it could not be distinguished from *L. longbeachae* SG1 or *L. longbeachae* SG2 in the phylogenic tree inferred from *mip* or *rpoB*. Although the clade was not robustly supported, further investigation is needed, because it may be useful in identifying the *L. longbeachae* strains with poor seroagglutination reactivity or cross-reactivity (Steele *et al*., 1990).

The other 10 isolates confidently formed a group with *L. pneumophila* in all the four phylogenic trees, supported by the PP 0.99 for the *rpoB* dataset, and 1.00 for the other three datasets. Based on Brenner's study (1988), there were 3 subspecies within *L. pneumophila*. Subspecies *pneumophila* was in serogroups from 1 through 14, while *L.* ssp. *pascullei* was in serogroups 1, 4, and 5, and *L.* ssp. *fraseri* was only in serogroup 5.

In the phylogenic tree inferred from the 16S rRNA matrix, strain CD-5, recognized as a strain of serogroup 5 by serotyping, formed a clade with *L.* ssp. *pascullei* and *L.* ssp. *fraseri* (PP: 1.00), but could not be discriminated. However, CD-5 could be clearly differentiated between the two subspecies in the phylogenic tree inferred from the *rpoB* matrix (PP: 1.00). Although CD-5 formed a clade with the type strain *L.* ssp. *fraseri* SG5 (PP: 1.00) in the phylogenic tree inferred from the *mip* matrix, the discriminability of *mip* within the species *L. pneumophila* was not as reliable as *rpoB*. This was because the type strain *L.* ssp. *fraseri* SG4 grouped with *L.* ssp. *pneumophila* SG3 but not with *L.* ssp. *fraseri* SG5, which was consistent with Ratcliff's study (1998).

In the phylogenic tree inferred from the combined matrix of *rpoB* and *mip*, CD-5 formed a clade with *L.* ssp. *fraseri* SG5 (PP: 1.00). However, as *L.* ssp. *pascullei* has been rarely reported since it was recognized, we were unable to find any reliable *mip* sequence of *L.* ssp. *pascullei*, and thus the three subspecies of *L. pneumophila* could not be confidently differentiated based on *mip* and *mip-rpoB* concatenation. The other 9 isolates were grouped with *L.* ssp. *pneumophila* in the phylogenic tree inferred from the *mip*, *rpoB*, and *mip-rpoB* matrices (PP: 1.00). However, the sequence analyses of *L.* ssp. *pneumophila* were uncorrelated with serological analyses, even in the combined matrix of *mip* and *rpoB*.

Based on MLSA targeting 16S rRNA, *mip*, *rpoB*, and *mip-rpoB* concatenation, all isolates were clearly identified. CD-1 was recognized as a strain of *L. longbeachae*, while the other 10 isolates were *L. pneumophila*. At the subspecies level, strain CD-5 was identified as *L.* ssp. *fraseri* while the other 9 strains were *L.* ssp. *pneumophila*. Thus, strains of *Legionella* were isolated from 4 of the 5 municipal fountains investigated in this study, indicating that these 4 fountains were contaminated with *Legionella*, and disinfection and long-term surveillance of these municipal fountains are urgently needed. To our knowledge, this is the first study which has focused on the *Legionella*-contaminated status of municipal fountains in China.

In conclusion, we combined *in vitro* culture and MLSA targeting 16S rRNA, *mip* and *rpoB* to isolate and identify *Legionella* spp. from 4 of 5 municipal fountains in Chengdu, China. A total of 11 strains were isolated. One strain was identified as *L. longbeachae*, while the other 10 strains were *L. pneumophila*. Of the three loci used in this study, 16S rRNA was useful for initial identification as it could recognize isolates robustly at the genus level, while *mip*, *rpoB*, and the *mip-rpoB* concatenation could confidently discriminate *Legionella* species. Notably, the three subspecies of *L. pneumophila* could be distinguished by the analysis based on *rpoB*. It is interesting that the serotyping result of CD-1 was consistent with genetic analysis based on concatenation of *mip* and *rpoB*, while it could not be identified to the serogroup level by *mip* or *rpoB* alone. None of the loci used in this study could identify *L.* ssp. *pneumophila* strains to the serogroup level. The municipal fountains that were found to be *Legionella*-positive in this study may be a potential threat to human health, as all are located in densely populated areas of this city. Thus, disinfection and long-term surveillance of these municipal fountains is urgently needed.

Acknowledgements

We are grateful to our colleagues, Lin Liao and Jinxian Chen, for their kind help in sampling and serotyping. We would like to thank Dr. Xian-guang Guo (Chengdu Institute of Biology, Chinese Academy of Sciences) and two anonymous reviewers for their helpful comments that improved this manuscript. This study is supported by grants from the National Natural Sciences Foundation of China (No. 81001345, No. 30800094 and No. 81171607) and Doctoral Program of Higher Education Research Fund (No. 2006060092).

References

Akaike, H. 1974. A new look at the statistical model identification. IEEE. *Trans. Automatic. Control.* **19**, 716–723.

Anand, C.M., Skinner, A.R., Malic, A., and Kurtz, J.B. 1983. Interaction of *L. pneumophila* and a free living amoeba (*Acanthamoeba palestinensis*). *J. Hyg. Camb.* **91**, 167–178.

Bangsborg, J.M., Shand, G., Pearlman, E., and Hoiby, N. 1991. Cross-reactive *Legionella* antigens and the antibody response during infection. *APMIS* **99**, 854–865.

Barabee, J.M., Gowman, G.W., Martin, W.T., Fields, B.S., and Morrill, W.E. 1987. Protocol for sampling environmental sites for legionellae. *Appl. Environ. Microbiol.* **53**, 1454–1458.

Bartie, C., Vente, S.N., and Nel, L.H. 2003. Identification methods for *Legionella* from environmental samples. *Water Res.* **37**, 1362–1370.

Blanco, S., Prat, C., Sanchez, M.D., Ferrer, D., Pellicer, T., Haba, L., Latorre, I., Vilaplana, C., Ausina, V., and Dominguez, J. 2008. Evaluation of a *Legionella* urinary antigen enzyme immunoessay for rapid detection of *Legionella pneumophila* in water samples. *Int. J. Hyg. Envir. Health* **211**, 168–171.

Boswell, T.C. 1996. Serological cross reaction between *Legionella* and campylobacterin the rapid microagglutination test. *J. Clin. Pathol.* **49**, 584–586.

Brenner, D.J. 1986. Classification of Legionellaceae: current status and remaining questions. *Isr. J. Med. Sci.* **22**, 620–632.

Brenner, D.J., Steigerwalt, A.G., Epple, P., Bibb, F.W., Mckinney, R.M., Starnes, R.W., Colville, J.M., Selander, R.M., and Edelstein, P.H. 1988. *Legionella pneumophila* serogroup Lansing 3 isolated from a patient with fatal pneumonia, and description of *L. pneumophila* subsp. *pneumophila* subsp. nov., *L. pneumophila* subsp. *fraseri* subsp. nov., and *L. pneumophila* subsp. *pascullei* subsp. nov. *J. Clin. Microbiol.* **26**, 1695–1703.

Cianciotto, N.P., Eisenstein, B.I., Mody, C.H., Toews, G.B., and Engleberg, N.C. 1989. A *Legionella pneumophila* gene encoding a species-specific surface protein potentiates initiation of intracellular infection. *Infect. Immun.* **57**, 1255–1262.

Clayton, R.A., Sutton, G., Hinkle, P.S., Bult, C., and Fields, C. 1995. Intra specific variation in small-subunit rRNA sequences in GenBank, why single sequences may not adequately represent prokaryotic taxa. *Int. J. Syst. Bacteriol.* **45**, 595–599.

Cooper, I.R., White, J., Mahenthiralingam, E., and Hanlon, G.W. 2008. Long-term persistence of a single *Legionella pneumophila* strain possessing the *mip* gene in a municipal shower despite repeated cycles of chlorination. *J. Hosp. Infect.* **70**, 154–159.

Coscolla, M., Fenolla, J., Escribano, I., and Gonzalez-Candelas, F. 2010. Legionellosis outbreak associated with asphalt paving machine, Spain, 2009. *Emerg. Infect. Dis.* **16**, 1381–1387.

Dahllof, I., Baillie, H., and Kjelleberg, S. 2000. RpoB-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl. Environ. Microbiol.* **66**, 3376–3380.

Declerk, P., Behets, J., Lammertyn, E., Labeau, I., Anne, J., and Ollevier, F. 2005. Detection and quantification of *Legionella pneumophila* in water samples using competitive PCR. *Can. J. Microbiol.* **52**, 584–590.

Doyle, R.M., Steele, T.W., McLennan, A.M., Parkinson, I.H., Manning, P.A., and Heuzenroeder, M.W. 1998. Sequence analysis of the *mip* gene of the soilborne pathogen *Legionella longbeachae*. *Infect. Immun.* **66**, 1492–1499.

Edelstein, P.H. 2007. Legionella. *In* Murray, P.R., Baron, E.J., Jorgensen, J.H., Landry, M.L., and Pfalle, M.A. (eds.), Manual of clinical microbiology, pp. 835–849. 9th ed. ASM Press, Washington, D.C., USA.

Euser, S.M., Pelgrim, M., and Den Boer, J.W. 2010. Legionnaires' disease and Pontiac fever after using a private outdoor whirl-

pool spa. *Scand. J. Infect. Dis.* Epub. **42**, 910–916.
Fields, B.S., Benson, R.F., and Besser, R.E. 2002. *Legionella* and Legionnaires' disease, 25 years of investigation. *Clin. Microbiol. Rev.* **15**, 506–526.
Fields, B.S., Shotts, E.B., Feeley, J.C., Gorman, G.W., and Martin, W.T. 1984. Proliferation of *Legionella pneumophila* as an intracellular parasite of the ciliated protozoan *Tetrahymena pyriformis*. *Appl. Environ. Microbiol.* **47**, 467–471.
Fry, N.K., Warwick, S., Saunders, N.A., and Embley, T.M. 1991. The use of 16S ribosomal RNA analyses to investigate the phylogeny of the family Legionellaceae. *J. Gen. Microbiol.* **137**, 215–222.
Gomez-Valero, L., Rusniok, C., and Buchrieser, C. 2009. Legionella pneumophila, population genetics, phylogeny and genomics. *Infect. Genet. Evol.* **9**, 727–739.
Hacker, J. and Fischer, G. 1993. Immunophilins, structure-function relationship and possible role in microbial pathogenicity. *Mol. Microbiol.* **10**, 445–456.
Helbig, J.H., Bernander, S., Castellani Pastoris, M., Etienne, J., Gaia, V., Lauwers, S., Lindsay, D., Lück, P.C., Marques, T., Mentula, S., and *et al.* 2002. Pan-European study on culture-proven Legionnaires' disease, distribution of *Legionella pneumophila* serogroups and monoclonal subgroups. *Eur. J. Clin. Microbiol. Infect. Dis.* **21**, 710–716.
Hoge, C.W. and Breiman, R.F. 1991. Advances in the epidemiology and control of *Legionella. Epidemiol. Rev.* **13**, 329–335.
Holmes, E.C., Urwin, R., and Maiden, M.C.J. 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**, 741–749.
Huelsenbeck, J.P., Larget, B., Miller, R.E., and Ronquist, F. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst. Biol.* **51**, 673–688.
Joseph, C.A., Ricketts, K.D., Yadav, R., and Patel, S., on behalf of the European Working Group for Legionella Infections. 2010. Travel-associated Legionnaires' disease in Europe in 2009. *Euro. Surveill.* **15**, 19683.
Kim, B.J., Lee, S.H., Lyu, M.A., Kim, S.J., Bai, G.H., Kim, S.S., Chae, G.T., Kim, E.C., Cha, C.Y., and Kook, Y.H. 1999. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *J. Clin. Microbiol.* **37**, 1714–1720.
Ko, K.S., Lee, H.K., Park, M.Y., Lee, K.H., Yun, Y.J., Woo, S.Y., Miyamoto, H., and Kook, Y.H. 2002. Application of RNA polymerase beta-subunit gene (*rpoB*) sequences for the molecular differentiation of *Legionella* species. *J. Clin. Microbiol.* **40**, 2653–2658.
Kuroki, H., Miyamoto, H., Fukuda, K., Iihara, H., Kawamura, Y., Ogawa, M., Wang, Y., Ezaki, T., and Taniguchi, H. 2007. *Legionella impletisoli* sp. nov. and *Legionella yabuuchiae* sp. nov., isolated from soils contaminated with industrial wastes in Japan. *Syst. Appl. Microbiol.* **30**, 273–279.
Luck, P.C., Igel, L., Helbig, J.H., Kuhlisch, E., and Jaztzwauk, L. 2004. Comparison of commercially available media for the recovery of *Legionella species*. *Int. J. Hyg. Environ. Health* **207**, 589–593.
Mathys, W., Stanke, J., Harmuth, M., and Jungle-Mathys, E. 2008. Occurence of *Legionella* in hot water system of single-family residences in suburbs of two German cities with special reference to solar and district heating. *Int. J. Hyg. Environ. Health* **211**, 179–185.
Mollet, C., Drancourt, M., and Raoult, D. 1997. *RpoB* sequence analysis as a novel basis for bacterial identfication. *Mol. Microbiol.* **26**, 1005–1011.
Newton, H.J., Ang, D.K., van Driel, I.R., and Hartland, E.L. 2010. Molecular pathogenesis of infections caused by *Legionella pneumophila*. *Clin. Microbiol. Rev.* **23**, 274–298.
O'Connell, W.A., Dhand, L., and Cianciotto, N.P. 1996. Infection of macrophage-like cells by *Legionella* species that have not been associated with disease. *Infect. Immun.* **64**, 4381–4384.
Page, R.D.M. 1996. TREEVIEW, an application to display phylogenetic trees on personal computers. *CABIOS* **12**, 357–358.
Palmore, T.N. and Stock, F. 2009. A cluster of nosocomial Legionnaires' disease linked to a contaminated hospital decorative water fountain. *Infect. Control. Hosp. Epidemiol.* **30**, 764–768.
Posada, D. and Buckley, T.R. 1998. Modeltest, testing the model of DNA substitution. *Bioinfomatics* **14**, 817–818.
Posada, D. and Buckley, T.R. 2004. Model selection and model averaging in phylogenetics advantages of Akaike information criterion and Bayesian approach. *Syst. Biol.* **53**, 793–808.
Qu, P., Deng, X., Zhang, J., Chen, J., Zhang, J., Zhang, Q., Xiao, Y., and Chen, S. 2009. Identification and characterization of the *Francisella* spp. strain 08HL01032 isolated in air condition systems. *Acta. Microbiologica. Sinica.* **49**, 1003–1010.
Ratcliff, R.M., Lanser, J.A., Manning, P.A., and Heuzenroeder, M.W. 1998. Sequence-based classification scheme for the genus *Legionella* targeting the *mip* gene. *J. Clin. Microbiol.* **36**, 1560–1567.
Ronquist, F. and Huelsenbeck, J.P. 2003. Mrbayes 3, Bayesian phylogenic inference under mixed models. *Bioinformatics* **19**, 1572–2574.
Rowbotham, T.J. 1986. Current views on the relationships between amoebae, legionellae and man. *Isr. J. Med. Sci.* **22**, 678–689.
Rubin, C.J., Thollesson, M., Kirsebom, L.A., and Herrmann, B. 2005. Phylogenetic relationships and species differentiation of 39 *Legionella* species by sequence determination of the RNase P RNA gene *rnpB*. *J. Syst. Bacteriol.* **55**, 2039–2049.
Severinov, K., Mustaev, A., Kukarin, A., Muzzin, O., Bass, I., Darst, S.A., and Goldfarb, A. 1996. Structural modules of the large subunits of RNA polymerase. *J. Biol. Chem.* **271**, 27969–27974.
Steele, T.W., Lanser, J., and Sangster, N. 1990. Isolation of *Legionella longbeachae* serogroup1 from potting mixes. *Appl. Environ. Microbiol.* **56**, 49–53.
Sussman, M. 2002. Molecular medical microbiology. Academic Press, Inc., London, England.
Swofford, D.L. 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer, Sunderland, MA, USA.
Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4, Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Bio. Evol.* **24**, 1596–1599.
Thompson, J.D., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The Clustal X Windows interface, flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
Tyndall, R.L. and Domingue, E.L. 1982. Cocultivation of *Legionella pneumophila* and free-living amoebae. *Appl. Environ. Microbiol.* **44**, 954–959.
Ueda, K., Seki, T., Kudo, T., Yoshida, T., and Kataoka, M. 1999. Two distinct mechanisms cause heterogeneity of 16S rRNA. *J. Bacteriol.* **181**, 78–82.
Wilson, D.A., Reischl, U., Hall, G.S., and Procop, G.W. 2007. Use of partial 16S rRNA gene sequencing for identification of *Legionella pneumophila* and non-*pneumophila Legionella* spp. *J. Clin. Microbiol.* **45**, 257–258.
Yu, V.L., Plouffe, J.F., Pastoris, M.C., Stout, J.E., Schousboe, M., Widmer, A., Summersgill, J., File, T., Heath, C.M., Paterson, D.L., and *et al.* 2002. Distribution of *Legionella* species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis, an international collaborative survey. *J. Infect. Dis.* **186**, 127–128.