

Application of RNA Polymerase β -Subunit Gene (*rpoB*) Sequences for the Molecular Differentiation of *Legionella* Species

Kwan Soo Ko,¹ Hae Kyung Lee,² Mi-Yeoun Park,² Keun-Hwa Lee,¹ Yeo-Jun Yun,¹ So-Yon Woo,¹ Hiroshi Miyamoto,³ and Yoon-Hoh Kook^{1*}

Department of Microbiology and Cancer Research Institute, Institute of Endemic Diseases, SNUMRC, Seoul National University College of Medicine, and Clinical Research Institute, Seoul National University Hospital, Seoul 110-799,¹ and Laboratory of Rickettsial and Zoonotic Disease, Department of Microbiology, Korean National Institute of Health, Seoul 122-701,² Korea, and Department of Microbiology, School of Medicine, University of Occupational and Environmental Health, Kitakyushu 807, Japan³

Received 21 December 2001/Returned for modification 4 March 2002/Accepted 23 April 2002

The nucleotide sequences of the partial *rpoB* gene were determined from 38 *Legionella* species, including 15 serogroups of *Legionella pneumophila*. These sequences were then used to infer the phylogenetic relationships among the *Legionella* species in order to establish a molecular differentiation method appropriate for them. The sequences (300 bp) and the phylogenetic tree of *rpoB* were compared to those from analyses using 16S rRNA gene and *mip* sequences. The trees inferred from these three gene sequences revealed significant differences. This sequence incongruence between the *rpoB* tree and the other trees might have originated from the high frequency of synonymous base substitutions and/or from horizontal gene transfer among the *Legionella* species. The nucleotide variation of *rpoB* enabled more evident differentiation among the *Legionella* species than was achievable by the 16S rRNA gene and even by *mip* in some cases. Two subspecies of *L. pneumophila* (*L. pneumophila* subsp. *pneumophila* and subsp. *fraseri*) were clearly distinguished by *rpoB* but not by 16S rRNA gene and *mip* analysis. One hundred and five strains isolated from patient tissues and environments in Korea and Japan could be identified by comparison of *rpoB* sequence similarity and phylogenetic trees. These results suggest that the partial sequences of *rpoB* determined in this study might be applicable to the molecular differentiation of *Legionella* species.

Legionella pneumophila, the causative agent of Legionnaires' disease and the type species of *Legionella*, was first recognized in 1977 following an epidemic of acute pneumonia in Philadelphia, Pennsylvania. Legionellosis is usually acquired in the community and accounts for 2 to 15% of all community-acquired pneumonia that requires hospitalization in the United States (27). So far, of the more than 40 *Legionella* species that have been characterized and classified by various methods (8, 30), 21 have been reported to be pathogenic to humans (20, 30). A classification scheme constructed by various methods can be efficiently used for the rapid and precise identification of clinical isolates in the diagnostic laboratory. One of these methods is the determination of molecular phylogenetic relationships or typing (18). *Legionella* species have also been previously analyzed using 16S rRNA gene (1, 6, 8) and macrophage infectivity potentiator (*mip*) sequences (19, 20). The 16S rRNA gene is by far the most widely used molecular marker to determine the phylogenetic relationships of bacteria (31). However, its utility has been questioned because of its heterogeneity (4). The *mip* gene encodes Mip, which has been reported to play a significant role in virulence (2). It has been shown that *mip* is useful for the discrimination of *Legionella* species (19, 20). However, incongruences have been found between the phylogenies elucidated from the 16S rRNA gene and *mip* sequences (8, 20). Relatively low branch-supporting bootstrap values have

restricted the comprehensive elucidation of the phylogenetic relationships within the *Legionella* species. The use of only a single gene for phylogenetic study makes it difficult to distinguish the phylogenetic relationships among species from the viewpoint of gene history (18). Therefore, the importance of comparing the sequences of several genes to evaluate a comprehensive bacterial phylogeny has been stressed (18). Consequently, in the present study we investigated the value of another gene, *rpoB*, to supplement the classification and identification scheme for *Legionella* species.

The *rpoB* gene encodes the β subunit of DNA-dependent RNA polymerase (24), and rifampin resistance is related to mutations in a particular region of *rpoB* (10). Recently, *rpoB* sequences were used as an alternative tool either for determining the phylogeny of or for identifying enteric bacteria (16), *Mycobacterium* (10), spirochetes including *Borrelia* (14, 21), and *Bartonella* (22). In this study, partial *rpoB* sequences (300 bp), containing a region that though highly conserved still has a remarkable ability for species differentiation (10, 12, 14, 22), were determined for genotypic classification of *Legionella* species and the results were compared with those results from the 16S rRNA gene and *mip*. In addition, culture isolates of *Legionella* species were identified based on *rpoB* sequences.

Fifty-five reference strains of 38 *Legionella* species were used in this study (Table 1). Among the 15 serogroups of *L. pneumophila*, three strains of serogroups, i.e., 4, 5, and 15, were *L. pneumophila* subsp. *fraseri*, while the others were *L. pneumophila* subsp. *pneumophila*. One hundred and five culture isolates, of which 10 strains were isolated from pneumonia patients and the others from cooling water, were identified by

* Corresponding author. Mailing address: Department of Microbiology, Seoul National University College of Medicine, 28 Yongon-dong, Chongno-gu, Seoul 110-799, Korea. Phone: (82) 2-740-8306. Fax: (82) 2-743-0881. E-mail: yhkook@plaza.snu.ac.kr.

TABLE 1. *Legionella* strains used to determine the *rpoB* sequences in this study

Species (serogroup)	Strain no. ^b	Accession no.		
		<i>rpoB</i>	16S rRNA ^c	<i>mip</i> ^d
<i>L. pneumophila</i> (1) ^a	ATCC 33152	AF367748	M36023	S42595
<i>L. pneumophila</i> (1) ^a	ATCC 33153	AY036036		
<i>L. pneumophila</i> (1) ^a	SF9	AY036037		
<i>L. pneumophila</i> (1) ^a	ATCC 43109	AY036038		
<i>L. pneumophila</i> (2) ^a	ATCC 33154	AY036039		
<i>L. pneumophila</i> (3) ^a	ATCC 33155	AY036040		
<i>L. pneumophila</i> (4) ^a	ATCC 33156	AY036041		
<i>L. pneumophila</i> (5) ^a	ATCC 33216	AY036042		
<i>L. pneumophila</i> (6) ^a	ATCC 33215	AY036043		
<i>L. pneumophila</i> (7) ^a	ATCC 33823	AY036044		
<i>L. pneumophila</i> (8) ^a	ATCC 35096	AY036045		
<i>L. pneumophila</i> (9) ^a	ATCC 35289	AY036046		
<i>L. pneumophila</i> (10) ^a	ATCC 43283	AY036047		
<i>L. pneumophila</i> (11) ^a	ATCC 43130	AY036048		
<i>L. pneumophila</i> (12) ^a	ATCC 43290	AY036049		
<i>L. pneumophila</i> (13) ^a	ATCC 43736	AY036050		
<i>L. pneumophila</i> (14) ^a	ATCC 43709	AY036051		
<i>L. pneumophila</i> (15) ^a	ATCC 35351	AY036052		
<i>L. adelaidensis</i>	UOEH 13562	AF367721	Z49716	U91606
<i>L. anisa</i> ^a	ATCC 35292	AF367722	Z32635	AF022312
<i>L. birminghamensis</i> ^a	UOEH 11749	AF367723	Z49717	AF047743
<i>L. bozemanii</i> (1) ^a	ATCC 33217	AF367724	Z49719	U91609
<i>L. brunensis</i>	UOEH 12655	AF367725	Z32636	AF022311
<i>L. cherrii</i> ^a	UOEH 10742	AF367726	Z49720	U91635
<i>L. cincinnatiensis</i> ^a	UOEH 12201	AF367727	Z49721	AF022358
<i>L. dumoffii</i> ^a	ATCC 33279	AF367728	Z32637	AF022313
<i>L. erythra</i> (1)	ATCC 35303	AF367729	M36027	U92203
<i>L. fairfieldensis</i>	UOEH 13563	AF367730	Z49722	U92214
<i>L. feeleei</i> (1) ^a	ATCC 35072	AF367731	Z49740	U92205
<i>L. geestiana</i>	ATCC 49504	AF367732	Z49723	ND ^e
<i>L. gormanii</i> ^a	ATCC 33297	AF367733	Z42639	AF047747
<i>L. gratiana</i>	ATCC 49413	AF367734	Z49725	U92206
<i>L. hackeliae</i> (1) ^a	ATCC 35250	AF367735	M36028	U92207
<i>L. israelensis</i> ^a	ATCC 43119	AF367736	Z32640	U92208
<i>L. jamestowniensis</i>	ATCC 35298	AF367737	Z49726	AF022339
<i>L. jordani</i> ^a	HM 7000	AF367738	Z32667	U92209
<i>L. lansingensis</i> ^a	ATCC 49751	AF367739	Z49727	U92210
<i>L. londiniensis</i>	ATCC 49505	AF367740	Z49728	AF022346
<i>L. longbeachae</i> (1) ^a	ATCC 33462	AF367741	M36029	X83036
<i>L. maceachernii</i> ^a	ATCC 35300	AF367742	Z32641	AF022315
<i>L. micdadei</i> ^a	ATCC 33218	AF367743	M36032	AF023175
<i>L. moravica</i>	ATCC 43877	AF367744	Z49729	U92212
<i>L. nautarum</i>	ATCC 49506	AF367745	Z49730	U92213
<i>L. oakridgensis</i> ^a	HM 7002	AF367746	Z32643	U92214
<i>L. parisiensis</i> ^a	UOEH 11745	AF347747	Z49731	U92215
<i>L. quinlivanii</i> (1)	ATCC 43830	AF367749	Z49733	AF022347
<i>L. rubrilucens</i>	ATCC 35304	AF367750	Z32643	AF022357
<i>L. sainthelensi</i> (1) ^a	ATCC 35248	AF367751	Z49734	U92219
<i>L. santicrocuis</i>	UOEH 11746	AF367752	Z49735	U92220
<i>L. shakespearei</i>	ATCC 49655	AF367753	Z49736	U92221
<i>L. spiritensis</i> (1)	UOEH 11199	AF367754	M36030	AF047751
<i>L. steigerwaltii</i>	UOEH 11747	AF367755	Z49737	U92223
<i>L. tucsonensis</i> ^a	ATCC 49180	AF367756	Z32644	U92224
<i>L. wadsworthii</i> ^a	ATCC 33877	AF367757	Z49738	U92225
<i>L. worsleiensis</i>	ATCC 49508	AF367758	Z49739	U9222

^a *Legionella* species pathogenic to human.

^b ATCC, American Type Culture Collection; UOEH, University of Occupational and Environmental Health; HM, Hiroshi Miyamoto.

^c 16S rRNA gene sequences used by Hookey et al. (8).

^d *mip* sequences used by Ratcliff et al. (19, 20).

^e ND, not determined (19, 20).

rpoB sequence analysis. These isolates were grown on buffered charcoal yeast extract agar and identified on the basis of cysteine requirement, autofluorescence, and biochemical tests such as those for gelatinase, urease, catalase, oxidase, peroxidase, and β -lactamase activities; hippurate hydrolysis; nitrate

reduction; and browning of tyrosine-supplemented agar (8). They were provided by Mi-Yeoun Park at the Korean National Institute of Health and H. Miyamoto at the University of Occupational and Environmental Health for blinded identification by *rpoB* sequence analysis.

DNAs were extracted using the bead beater-phenol extraction method (10) and were used as a template for PCR. A primer pair, RL1 (5'-GAT GAT ATC GAT CAY CTD GG-3') and RL2 (5'-TTC VGG CGT TTC AAT NGG AC-3'), designed from the consensus sequences of *Escherichia coli* (GenBank accession no. V00339), *Coxiella burnetii* (U86688), and *L. pneumophila* (AF087812), was used to amplify a portion of *rpoB* DNA (369 bp) containing the Rif^r region (10, 17). Template DNA (ca. 50 ng) and 20 pmol of each primer were added to a PCR mixture tube (AccuPower PCR PreMix; Bioneer, Daejeon, Korea) containing 1 U of *Taq* DNA polymerase, each deoxynucleoside triphosphate at a concentration of 250 μ M, 10 mM Tris-HCl (pH 8.3), 10 mM KCl, 1.5 mM MgCl₂, and gel loading dye (14). The final volume was adjusted to 20 μ l with distilled water. The reaction mixture was then subjected to 30 cycles for amplification. Each cycle consisted of 30 s at 95°C for denaturation, 30 s at 55°C for annealing, and 1 min at 72°C for extension, followed by final extension at 72°C for 5 min (model 9700 Thermocycler; Perkin-Elmer Cetus). Amplified PCR products were purified for sequencing using a QIAEX II gel extraction kit (Qiagen, Hilden, Germany).

Sequences of the purified PCR products were directly determined with forward and reverse primers using an Applied Biosystems automated sequencer (model 377) and a BigDye Terminator Cycle Sequencing kit (Perkin-Elmer Applied Biosystems, Warrington, United Kingdom). For the sequencing reaction, 30 ng of purified PCR products, 2.5 pmol of each primer, and 4 μ l of BigDye Terminator RR mix (part no. 4303153; Perkin-Elmer Applied Biosystems) were mixed and adjusted with distilled water to a final volume of 10 μ l. The reaction was run with 5% (vol/vol) dimethyl sulfoxide for 30 cycles of 15 s at 95°C, 5 s at 50°C, and 4 min at 60°C. To confirm the identity of the reference strains used in this study, we also amplified and sequenced about 400 bp from the 3' end of the 16S rRNA gene (8). In addition, *mip* sequence analysis (402 bp) for 79 Korean isolates was performed to compare with the result of *rpoB* sequence analysis. For the amplification and sequencing of *mip* DNA, a new primer set, ML1 (5'-GAT AAG TTG TCT TAT AGC ATT GG-3') and ML2 (5'-TCT GTC CAT CCT GGG ATA ACT TG-3'), was used.

The partial *rpoB* sequences (300 bp) were aligned using the multiple alignment program CLUSTAL X (29) and the sequences determined were submitted to GenBank (Table 1). Phylogenetic trees of *Legionella* were constructed by the neighbor-joining (NJ) method and the parsimony method in the PAUP program (28). In the NJ method, pairwise distances were calculated using the maximum likelihood option, and in parsimony analysis, heuristic searches were conducted with the option of tree bisection reconstruction branch swapping. *C. burnetii*, the etiologic agent of Q fever, was used as an outgroup. Branch supporting values were evaluated with 1,000 bootstrap replications. The 16S rRNA gene (1,396 bp) and *mip* (525 bp) sequences of *Legionella* were retrieved from the GenBank database and used to infer the phylogenetic relationships by the same method. Incongruence length difference (ILD) tests (partition homogeneity test in the PAUP program) (3) were conducted to determine whether *rpoB*, the 16S rRNA gene, and *mip* data sets were coalescent together or not. The degree of incongruence between trees was assessed by comparing the log likelihood values (11) of three NJ trees with

those of the phylogenetic topologies obtained from the other genes (7). All analyses were performed using the PAUP program (28).

Similarities between the partial *rpoB* DNA sequences (300 bp) of *Legionella* species were lower than 95% in all cases except for *L. jamestownensis*-*L. londiniensis* (98.7%). The sequence divergence in *rpoB* was 3.5 times greater than that for the 16S rRNA gene, indicating the greater variance of *rpoB*. The *mip* sequences were more divergent than those of *rpoB* by a factor of 1.5. However, the relationship between the pairwise dissimilarities of *rpoB* and other genes was not linear, which suggested that the degree of sequence divergences was not proportional to the time needed for speciation. In contrast, the pairwise distances between the *mip* and 16S rRNA genes exhibited a close linear correlation (data not shown). The homoplasy index of the *rpoB* sequences (0.779), calculated by parsimony analysis using the PAUP program, was higher than those of the 16S rRNA gene and *mip* sequences (0.558 and 0.693, respectively).

The phylogeny inferred from *rpoB* sequences has long terminal branches with the exception of *L. jamestownensis* and *L. londiniensis*. However, bootstrap values supporting each branch in the *rpoB* phylogeny were relatively lower (Fig. 1) than those of the 16S rRNA gene and *mip* phylogenies (Fig. 2). Although *rpoB*, 16S rRNA gene, and *mip* phylogenies demonstrated partially similar relationships, they represented different topologies in many respects (Fig. 1 and 2). According to the results of ILD tests (3) of PAUP (28), three of the gene data sets were incongruent with each other ($P < 0.001$), indicating that three gene sequences could not be combined. Log likelihood tests (11) also indicated that the tree topology of one gene differed significantly from those of the other two genes ($P < 0.0001$).

The *rpoB* phylogeny indicated that *L. pneumophila* was closely related to *L. gormanii* (Fig. 1), a blue-white autofluorescent species that differs from *L. pneumophila* in many respects. In the *mip* tree (Fig. 2B), however, *L. pneumophila* was placed at a basal position of the B-BW and B groups (see Fig. 2 legend) and the clade of *L. moravica*, *L. worsleiensis*, and *L. shakespearei*. The bootstrap value supporting their relationships was relatively high, 91%. Sequence dissimilarities of *rpoB* between two subspecies of *L. pneumophila* (subsp. *pneumophila* and subsp. *fraseri*) ranged from 13.2 to 12.4%. These values exceeded those observed in other species, such as *L. steigerwaltii* and *L. cherrii* (10.4%) and *L. anisa* and *L. parisiensis* (9.2%).

There are eight blue-white autofluorescent species in the genus *Legionella*: *L. anisa*, *L. parisiensis*, *L. bozemanii*, *L. cherrii*, *L. steigerwaltii*, *L. gormanii*, *L. dumoffii*, and *L. tucsonensis* (8, 20). All of these species belong to ubiquinone group B (13, 20). In the *mip* tree (Fig. 2B), the clade formation of these eight blue-white autofluorescent species with *L. wadsworthii* was robustly supported by bootstrap analysis (97%), but there was no equivalent single clade in the *rpoB* (Fig. 1) and 16S rRNA gene phylogenies (Fig. 2A).

The other species belonging to ubiquinone group B, *L. cinncinnatiensis*, *L. santicrucis*, *L. longbeachae*, *L. sainthelensi*, and *L. gratiana* (13, 20), all grouped well into a single clade in both the *rpoB* (Fig. 1) and *mip* (Fig. 2B) phylogenies. While their relationships were supported only moderately by the *rpoB* phy-

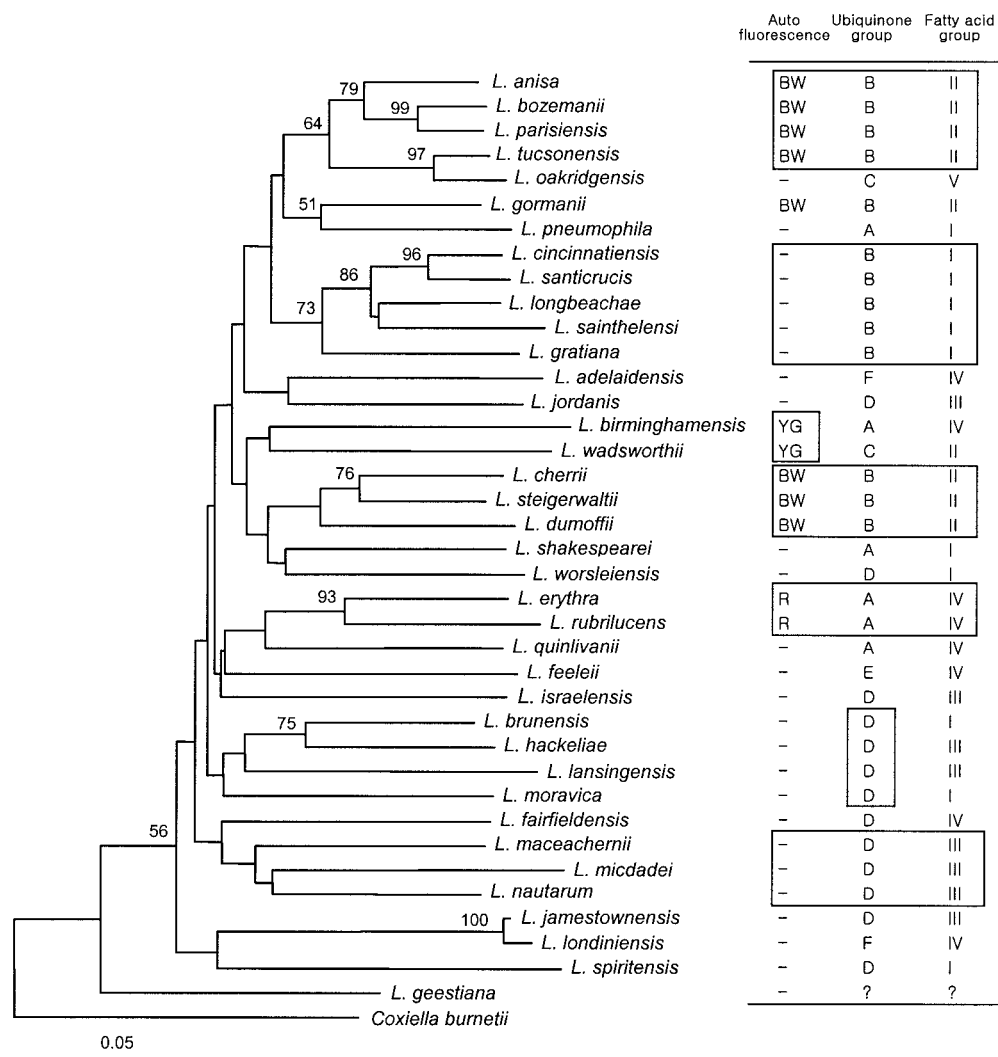


FIG. 1. Phylogenetic relationships of *Legionella* species inferred from partial *rpoB* DNA sequences. This tree was constructed by the NJ method in PAUP (28). The species *C. burnetii* was used as an outgroup. The bootstrap values presented at corresponding branches were evaluated from 1,000 replications. Values below 50% are not indicated. The autofluorescence (BW, blue-white; YG, yellow-green; R, red), ubiquinone (A to F), and fatty acid (I to IV) groups of each *Legionella* species (8, 9, 13, 20, 30) are indicated at the right. Biochemical traits identical to phylogenetic grouping are represented in boxes. The scale bar represents 5 substitutions per 100 nucleotides.

logeny (bootstrap value of 73%), they were fully supported by the *mip* phylogeny. In the 16S rRNA gene tree (Fig. 2A), *L. gratiana* displayed a close relationship with one of the subgroups in the blue-white autofluorescence group, though with very low bootstrap support. The other four species formed a clade that was fully supported (bootstrap value of 100%) in the 16S rRNA gene tree (Fig. 2A).

In the *rpoB* tree, *L. maceachernii*, *L. micdadei*, and *L. nautarum* formed one clade (Fig. 1), which was identical in the *mip* phylogeny (Fig. 2B). These three species belong to the same ubiquinone group (D) and fatty acid group III (9, 13, 20). While their relationships were poorly supported by a very low bootstrap value in the *rpoB*, the support was much stronger as determined by the *mip* phylogeny. However, in the 16S rRNA gene phylogeny, rather than clustering with *L. maceachernii* and *L. micdadei*, *L. nautarum* formed a distinct group along with *L. israelensis* (Fig. 2A).

The red autofluorescent species, *L. erythra* and *L. rubrilucens*, formed a distinct clade in all analyses (R group in Fig. 2), showing a particularly high similarity in terms of 16S rRNA gene sequences (99.4%), though the sequence similarities in *rpoB* and *mip* were not as high (88.7 and 89.0%, respectively). The yellow-green autofluorescent species, *L. birminghamensis* and *L. wadsworthii* (30), were clustered into one group in the *rpoB* phylogeny (Fig. 1) but not in the other two trees. These two species with yellow-green autofluorescence showed a sequence similarity of 80.7%.

L. geestiana, which contains unique isoprenoid quinone (Q-14) (8, 13) and fatty acid compositions (5), was placed at a basal position with respect to all *Legionella* species in both *rpoB* and 16S rRNA gene phylogenies (Fig. 1 and 2A). For *rpoB* sequences, the sequence similarities between *L. geestiana* and all other species were below 80%.

Species identification of the 105 strains was accomplished

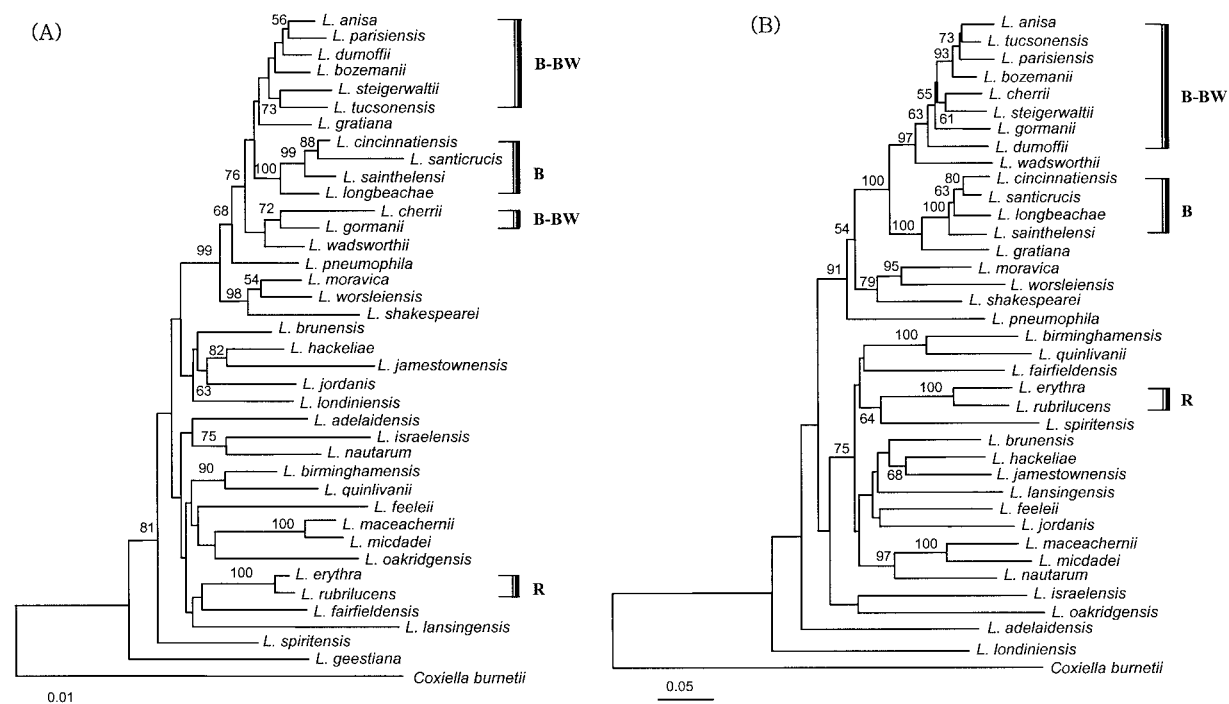


FIG. 2. Gene trees based on 16S rRNA gene sequences (A) and *mip* sequences (B). These were constructed using the same method as described above for Fig. 1 for the *rpoB* sequences. The *mip* sequence of *L. geestiana* was not available in GenBank. B, ubiquinone group B; R, red autofluorescence group; and B-BW, ubiquinone group B and blue-white autofluorescence group. The scale bars represent the number of inferred substitutions per site.

using *rpoB* sequence analysis of reference strains to measure the similarities and thereby infer the phylogenetic reconstruction (data not shown). Of these, 98 isolates belonged to *L. pneumophila* and exhibited 99.3 to 100% *rpoB* sequence similarities with the reference strains. The others were identified as other *Legionella* species (two *L. bozemanii*, two *L. dumoffii*, two *L. feeleii*, and one *L. micdadei* species), which showed 98.3 to 100% sequence similarities with each reference strain. It was interesting that 87 strains could be identified as *L. pneumophila* subsp. *pneumophila* and 11 as *L. pneumophila* subsp. *fraseri*. These strains of *L. pneumophila* subsp. *pneumophila* and subsp. *fraseri* showed 97.0 to 100% and 98.3 to 100% sequence similarities within each subspecies, respectively. The identified strains that belong to each subspecies showed 87.3 to 89.7% sequence similarity. However, no correlation was observed between the *rpoB* sequence and those of serogroups. For example, *rpoB* sequences of four reference strains belonging to serogroup 1 of *L. pneumophila* were not identical. On the other hand, *mip* sequences of 97 *L. pneumophila* strains, including 18 reference strains and 79 isolates, showed 93.4 to 100% similarities. However, unlike in *rpoB* analysis, no clear differentiation between two subspecies of *L. pneumophila* was observed in *mip* sequence analysis.

Recently, it was demonstrated that the *rpoB*-based approach to microbial community analysis or identification could practically overcome the inherent limitations of 16S rRNA gene intraspecies heterogeneity (4). In this study, we have shown that *rpoB* sequences are as useful as 16S rRNA gene and *mip* sequences in accessing and evaluating the relationships among the *Legionella* species. Such a simple genetic analysis promises

to provide a practical scheme for the classification of *Legionella* species and the identification of culture isolates in the diagnostic or reference clinical laboratory.

In contrast to analysis of other genes, *rpoB* analysis has several advantages. In spite of the sequence's shortness (300 bp), *rpoB* sequence analysis was able to clearly differentiate among the *Legionella* species. One such example was the clear differentiation of *L. erythra* and *L. rubrilucens* by *rpoB* analysis. Because of their high similarity in 16S rRNA gene sequences (99.4%) and their DNA hybridization test results, the latter representing a value of just below 70% (23), they had long been considered subspecies of the same species (8). Another example was the differentiation of blue-white autofluorescent species. While their 16S rRNA gene sequences indicated similarities of at least 97.8% (8), their *rpoB* sequences represented the most significant one, was the differentiation of subspecies. *L. pneumophila* includes two subspecies (*L. pneumophila* subsp. *pneumophila* and subsp. *fraseri*), which featured 99.2% similarity (8) with 16S rRNA gene sequence and were not differentiated by *mip* analysis (19). However, *rpoB* analysis was able to clearly distinguish them (12). Coupled with the high nucleotide similarity within each subspecies (97.0 to 100% and 98.3 to 100%), the significant difference of nucleotides between the two subspecies (87.3 to 89.7%) indicates that *rpoB* can be a better marker in differentiation of *L. pneumophila* than can *mip*, which could not discriminate two subspecies of *L. pneumophila*. These results confirm the usefulness of *rpoB* sequence analysis in population genetic and epidemiological studies of *L. pneumophila*, including molecular typing.

The *rpoB* tree showed quite a different topology from those of the 16S rRNA gene and *mip* sequences (Fig. 1 and 2). The ILD and log likelihood tests suggested that phylogenetic relationships inferred from the three genes were statistically significantly different. The lack of congruence between the *rpoB*, 16S rRNA gene, and *mip* trees did not result only from frequent synonymous base substitutions. Otherwise, the observed incongruence among the three gene trees is most likely the result of horizontal gene transfer between *Legionella* species (7, 15, 25, 26). It is known that such horizontal gene transfer disrupts the treelike branching pattern, thus complicating the phylogenetic relationships of species (7).

Due to these discrepancies among the three gene trees, none was able to exactly explain the phylogenetic relationships within the *Legionella* species and the evolution of this species. This suggests that the use of several markers, such as *rpoB* in combination with 16S rRNA gene or *mip*, may be necessary for the reliable identification and phylogenetic study of *Legionella*. Such an approach will reduce the risk of error in molecular typing or identification.

K. S. Ko and H. K. Lee equally contributed as joint first authors.

We thank H. S. Jung (School of Biological Sciences, Seoul National University, Seoul, Korea) for his help with the phylogenetic analysis and critical reading of the manuscript. We also thank B. S. Field, R. F. Benson, and E. Brown (National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Ga.) and M. J. Kim (College of Medicine, Korea University, Seoul, Korea) for providing reference strains and clinical isolates.

This work was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health and Welfare, Seoul, Republic of Korea (01-PJ10-PG6-01GM03-0002), and in part by the BK21 project for Medicine, Dentistry, and Pharmacy.

REFERENCES

- Birtles, R. J., T. J. Rowbotham, D. Raoult, and T. G. Harrison. 1996. Phylogenetic diversity of intra-amoebal legionellae as revealed by 16S rRNA gene sequence comparison. *Microbiology* **142**:3525–3530.
- Cianciotto, N. P., B. I. Eisenstein, C. H. Mody, G. B. Toews, and N. C. Engleberg. 1989. A *Legionella pneumophila* gene encoding a species-specific surface protein potentiates initiation of intracellular infection. *Infect. Immun.* **57**:1255–1262.
- Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* **14**:733–740.
- Dahlöf, L., H. Baillie, and S. Kjelleberg. 2000. *rpoB*-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl. Environ. Microbiol.* **66**:3376–3380.
- Diogo, A., A. Verissimo, M. F. Nobre, and M. S. Da Costa. 1999. Usefulness of fatty acid composition for differentiation of *Legionella* species. *J. Clin. Microbiol.* **37**:2248–2254.
- Fry, N. K., S. Warwick, N. A. Saunders, and T. M. Embley. 1991. The use of 16S ribosomal RNA analyses to investigate the phylogeny of the family *Legionellaceae*. *J. Gen. Microbiol.* **137**:215–222.
- Holmes, E. C., R. Urwin, and M. C. J. Maiden. 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**:741–749.
- Hookey, J. V., N. A. Saunders, N. K. Fry, R. J. Birtles, and T. G. Harrison. 1996. Phylogeny of *Legionellaceae* based on small-subunit ribosomal DNA sequences and proposal of *Legionella lytica* comb. nov. for *Legionella*-like amoebal pathogens. *Int. J. Syst. Bacteriol.* **46**:526–531.
- Jantzen, E., A. Sonesson, T. Tangen, and J. Eng. 1993. Hydroxy-fatty acid profiles of *Legionella* species: diagnostic usefulness assessed by principal component analysis. *J. Clin. Microbiol.* **31**:1413–1419.
- Kim, B.-J., S.-H. Lee, M.-A. Lyu, S.-J. Kim, G.-H. Bai, S.-S. Kim, G.-T. Chae, E.-C. Kim, C.-Y. Cha, and Y.-H. Kook. 1999. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *J. Clin. Microbiol.* **37**:1714–1720.
- Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**:170–179.
- Ko, K. S., H. K. Lee, M.-Y. Park, M.-S. Park, K.-H. Lee, S.-Y. Woo, Y.-J. Yun, and Y.-H. Kook. 2002. Population genetic structure of *Legionella pneumophila* inferred from RNA polymerase gene (*rpoB*) and *DotA* gene (*dotA*) sequences. *J. Bacteriol.* **184**:2123–2130.
- Lambert, M. A., and C. W. Moss. 1989. Cellular fatty acid composition and isoprenoid quinone contents of 23 *Legionella* species. *J. Clin. Microbiol.* **27**:465–473.
- Lee, S.-H., B.-J. Kim, J.-H. Kim, K.-H. Park, S.-J. Kim, and Y.-Y. Kook. 2000. Differentiation of *Borrelia burgdorferi* sensu lato on the basis of RNA polymerase gene (*rpoB*) sequences. *J. Clin. Microbiol.* **38**:2557–2562.
- Mintz, C. S. 1999. Gene transfer in *Legionella pneumophila*. *Microbes Infect.* **1**:1203–1209.
- Mollet, C. M. Drancourt, and D. Raoult. 1997. *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* **26**:1005–1011.
- Nielsen, K., P. Hindersson, N. Høiby, and J. M. Bangsbo. 2000. Sequencing of the *rpoB* gene in *Legionella pneumophila* and characterization of mutations associated with rifampin resistance in the *Legionellaceae*. *Antimicrob. Agents Chemother.* **44**:2679–2683.
- Palys, T., L. K. Nakamura, and F. M. Cohan. 1997. Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int. J. Syst. Bacteriol.* **47**:1145–1156.
- Ratcliff, R. M., J. A. Lanser, P. A. Manning, and M. W. Heuzenroeder. 1998. Sequence-based classification scheme for the genus *Legionella* targeting the *mip* gene. *J. Clin. Microbiol.* **36**:1560–1567.
- Ratcliff, R. M., S. C. Donnellan, J. A. Lanser, P. A. Manning, and M. W. Heuzenroeder. 1997. Interspecies sequence differences in the Mip protein from the genus *Legionella*: implication for function and evolutionary relatedness. *Mol. Microbiol.* **25**:1149–1158.
- Renesto, P., K. Lorvellec-Guillon, M. Drancourt, and D. Raoult. 2000. *rpoB* gene analysis as a novel strategy for identification of spirochetes from the genera *Borrelia*, *Treponema*, and *Leptospira*. *J. Clin. Microbiol.* **38**:2200–2203.
- Renesto, P., J. Gouvenet, M. Drancourt, V. Roux, and D. Raoult. 2001. Use of *rpoB* gene analysis for detection and identification of *Bartonella* species. *J. Clin. Microbiol.* **39**:430–437.
- Saunders, N. A., N. Doshi, and T. G. Harrison. 1992. A second serogroup of *Legionella erythra* serologically indistinguishable from *Legionella rubrilucens*. *J. Appl. Bacteriol.* **72**:262–265.
- Severinov, K., A. Mustaev, A. Kukarin, O. Muzzin, I. Bass, S. A. Darst, and A. Goldfarb. 1996. Structural modules of the large subunits of RNA polymerase. *J. Biol. Chem.* **271**:27969–27974.
- Smith, N. H., E. C. Homles, G. M. Donovan, G. A. Carpenter, and B. G. Spratt. 1999. Networks and groups within the genus *Neisseria*: analysis of *argF*, *recA*, *rho*, and 16S rRNA sequences from human *Neisseria* species. *Mol. Biol. Evol.* **16**:773–783.
- Stone, B. J., and Y. A. Kwaik. 1999. Natural competence for DNA transformation by *Legionella pneumophila* and its association with expression of type IV pili. *J. Bacteriol.* **181**:1395–1402.
- Swanson, M. S., and B. K. Hammer. 2000. *Legionella pneumophila* pathogenesis: a fateful journey from amoebae to macrophages. *Annu. Rev. Microbiol.* **54**:567–613.
- Swofford, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer Associates, Sunderland, Mass.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**:4876–4882.
- Winn, W. C. 1999. *Legionella*, p. 572–582. In P. R. Murray, E. J. Baron, M. A. Tenover, and R. H. Tenover, (ed.), *Manual of clinical microbiology*, 7th ed. American Society for Microbiology, Washington, D.C.
- Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.