Use of Partial 16S rRNA Gene Sequencing for Identification of *Legionella pneumophila* and Non-*pneumophila* *Legionella* spp.[▽]

D. A. Wilson,¹ U. Reischl,² G. S. Hall,¹ and G. W. Procop^{3*}

The Cleveland Clinic, Cleveland, Ohio¹; University of Regensburg, Regensburg, Germany²; and
The University of Miami, Miami, Florida³

Received 26 July 2006/Returned for modification 11 September 2006/Accepted 6 November 2006

We examined 49 *Legionella* species, 26 *L. pneumophila* and 23 non-*pneumophila* *Legionella* spp., using partial 16S rRNA gene sequencing. This approach accurately identified all the *L. pneumophila* isolates, characterized all non-*pneumophila* *Legionella* isolates as such within this genus, and classified most (20/23; 87%) of the non-*pneumophila* *Legionella* isolates to the species level.

Although *Legionella pneumophila* is the most frequent cause of legionellosis, non-*pneumophila* *Legionella* species may also cause serious or fatal disease (1, 2). Non-*pneumophila* *Legionella* species in respiratory specimens, however, are not detected by the *Legionella pneumophila* direct fluorescent antigen test. Similarly, the *Legionella* urinary antigen test will be negative for urine specimens from patients with infections caused by these bacteria. However, non-*pneumophila* *Legionella* species, like *L. pneumophila*, may be cultured on charcoal-yeast extract (CYE) agar. Unfortunately, fluorescent antibody stains are not commercially available for confirmation/identification of non-*pneumophila* *Legionella* isolates. Isolates suspected to represent non-*pneumophila* *Legionella* species must be sent to state public health laboratories for confirmation. The identification of these bacteria by public health facilities is not timely enough for clinical purposes in our experience.

It has been demonstrated that broad-range PCR targeting an rRNA gene may be used to detect the *Legionella* genus (3, 4). Therefore, we examined the ability of 16S rRNA gene sequencing as a means by which to identify *Legionella* species and compared these results to those obtained by traditional serotyping. We also wanted to identify GenBank (NCBI) entries that correlated with the serotypes of these isolates. Therefore, we examined a collection of DNA extracts from 49 *Legionella* species (26 *L. pneumophila* and 23 non-*pneumophila* *Legionella* spp.) by partial 16S rRNA gene sequencing using the OpenGene System (Bayer, Berkeley, CA). The resulting sequences were submitted as BLAST queries to GenBank (NCBI). The species level identification as achieved by DNA sequencing was then compared with the identification achieved by routine serotyping.

We examined DNA extracts of *Legionella* species from The Cleveland Clinic, Cleveland, OH, and the University of Regensburg, Regensburg, Germany. Nucleic acid extracts were obtained from isolates of *Legionella* spp. recovered from clinical specimens received at The Cleveland Clinic. The species of these isolates were determined by serologic studies performed at the Ohio State Department of Health. DNA extracts were

obtained using a QIAGEN DNA mini kit (QIAGEN, Inc., Valencia, CA) or a Roche MagNA Pure LC instrument (Roche Diagnostics Corporation, Indianapolis, IN), according to the manufacturer's guidelines. DNA extracts of the *Legionella* spp. were received from the Institute of Medical Microbiology and Hygiene, University of Regensburg, Regensburg, Germany. The serotyping of the isolates from the University of Regensburg was performed in Germany. The nucleic acid extracts from these *Legionella* spp. were prepared using the Roche MagNA Pure LC instrument (Roche Diagnostics Corporation, Indianapolis, IN). The non-*pneumophila* *Legionella* species tested consisted of *Legionella bozemanii* (three isolates), *L. longbeachae* (four isolates), *L. micdadei* (five isolates), *L. jordanis* (two isolates), *L. feeleyi* (two isolates), and one isolate each of *L. oakridgensis*, *L. israelensis*, *L. wadsworthii*, *L. tucsonensis*, *L. sainthelensis*, *L. birminghamiensis*, and *L. gormanii*.

The decontaminated, broad-range partial 16S rRNA gene amplification used for the amplification of the *Legionella* isolates has been used for many years in our laboratory for sequence-based identification of difficult-to-identify bacteria. The decontaminated PCR assay was performed with 14.5 µl of DNA extract (supernatant) in a total volume of 25.75 µl. The decontaminated PCR master mixture contained 2.5 µl of 10× reaction buffer II (without MgCl₂) (final concentration, 1×), 1.5 µl of 25 mM MgCl₂ (final concentration, 1.5 mM), 2.0 µl of 2.5 mM of each deoxynucleoside triphosphate (final concentration, 200 µM), 2.5 µl of 100% dimethyl sulfoxide (final concentration, 10%), 1.0 µl of a forward primer (5 µM) (final concentration, 0.2 µM), 1.0 µl of a reverse primer (5 µM) (final concentration, 0.2 µM), 0.25 µl of AmpliTaq polymerase (5 U/µl) (final concentration, 1.25 U), and 0.5 µl of ALU I (10 U/µl) (final concentration, 5 U). The master mixture was incubated at 37°C for 90 min to activate the enzyme ALU. This was followed by incubation at 65°C for 30 min to deactivate the enzyme.

The sequence of the forward primer was 5'-GTTAAGTCC CGCAACGA-3', which corresponded with *Staphylococcus aureus* GenBank entry N315 (NC-002745), positions 1100 to 1116. The sequence of the reverse primer was 5'-AGGAGGT GATCCAGCC-3', which corresponded with *S. aureus* GenBank entry N315 (NC-002745), positions 1551 to 1536. Computer-assisted comparison of these primers with those for *Legionella*

* Corresponding author. Mailing address: University of Miami Miller School of Medicine, Department of Pathology (R-5), 1611 NW 112th Avenue, Holtz Center, 2090, Miami, FL 33136. Phone: (305) 585-5068. Fax: (305) 326-9363. E-mail: gprocop@med.miami.edu.

[▽] Published ahead of print on 15 November 2006.

pneumophila subsp. *pneumophila* strain Philadelphia 1 (complete genome; GenBank accession number NC_002942) demonstrated complete homology in the reverse primer and a single nucleotide mismatch in the forward primer. This mismatch, however, was 6 nucleotides from the 3' end of the primer. The amplicon length calculated from this comparison was 420 nucleotides.

The PCR was performed on a model 9600 GeneAmp PCR system (Applied Biosystems). Denaturation was accomplished at 94°C for 2 min. Thirty-five cycles of PCR that consisted of 94°C for 30 s, 50°C for 30 s, and 72°C for 60 s were performed. A final extension phase at 72°C for 5 min was performed, followed by a 4°C hold.

The DNA sequencing was performed with an Open-Genie Long-Read DNA sequencer (Bayer HealthCare) according to the protocol provided by manufacturer. A Cy5/Cy5.5 dye primer cycle sequence kit (Bayer HealthCare) was used for the sequencing PCR; dideoxynucleotides are included in the kit. The forward and reverse amplification primers were used as sequencing primers. The forward primer was labeled at the 5' end with Cy5.5, whereas the reverse primer was labeled with Cy5.

The master mixture for the sequencing PCR had a 22.0-μl total volume and consisted of 2.5 μl of sequencing buffer, 4.0 μl of 100% dimethyl sulfoxide, 5.5 μl of a Cy5.5-labeled primer (0.20 μM), 2.2 μl of a Cy5.0-labeled primer (0.20 μM), 4.0 μl of diluted enzyme, 1.8 μl distilled H₂O, and 2.0 μl of PCR product for sequencing. For primers, a 20 μM stock was diluted 1:100 for a concentration of 0.20 μM. Each reaction mixture consisted of a 3-μl aliquot each of ddA, ddC, ddG, and ddT and a 5-μl aliquot of the master mixture.

The thermocycling for the sequencing reactions was performed with a model 9600 GeneAmp PCR system (Applied Biosystems). It began with an initial denaturation at 94°C for 2 min followed by 35 cycles of PCR that consisted of 94°C for 20 s, 50°C for 45 s, and 70°C for 60 s. A final extension was performed at 72°C for 2 min, followed by a hold at 4°C.

Amplicons were generated for all *Legionella* isolates tested, as determined by gel electrophoresis, as were DNA sequences. For the majority of the isolates tested, the reverse sequencing primer produced a higher-quality sequence than the forward primer. The combination of the sequences generated by this bidirectional sequencing was used to resolve the identities of any individual nucleotides in the sequence that were in question.

All of the *L. pneumophila* strains were accurately identified to the species level by DNA sequencing. Of the 26 *L. pneumophila* isolates sequenced, all matched best with GenBank entries for *L. pneumophila*, with >95% matches; most had 98 or 99% matches. The percent matches for the *L. pneumophila* isolates were 96% (2 isolates), 97% (1 isolate), 98% (11 isolates), 99% (11 isolates), and 100% (1 isolate). The lengths of the amplicons sequenced for the *L. pneumophila* isolates ranged from 213 to 398 nucleotides. For reference, matching 394 nucleotides of an amplicon with a length of 398 nucleotides would result in a 99% match.

All of the 23 non-*pneumophila* *Legionella* isolates were correctly categorized as non-*pneumophila* *Legionella* by DNA sequencing. The 20 isolates of non-*pneumophila* *Legionella* that were correctly characterized by partial 16S rRNA gene sequencing had percent matches of 90% (one isolate), 96% (three isolates), 97% (six isolates), 98% (eight isolates), and

99% (two isolates). The amplicon lengths for the non-*pneumophila* *Legionella* isolates sequence ranged from 187 to 385 nucleotides. Three of the 23 non-*pneumophila* *Legionella* isolates, *L. gormanii* (1 isolate), *L. feelei* (1 isolate), and *L. bozemanii* (1 isolate) were identified by DNA sequencing as species different from those designated by serotyping. Of the three isolates whose identities could not be confirmed by DNA sequencing, one of the three tested isolates of *L. bozemanii* had a 98% match for two GenBank entries for *L. longbeachae* and one entry for *L. bozemanii* (354-nucleotide amplicon). One *L. feelei* isolate could not be confirmed; it had a 98% match with a GenBank entry for *L. donaldsonii*, but the second- and third-best matches were for *L. feelei* (344-nucleotide amplicon). The *L. gormanii* isolate could not be confirmed by this method; it had a 97% match with *L. cherrii* (336-nucleotide amplicon). The second-, third-, and fourth-best matches for this isolate were *L. bozemanii*, *L. steigerwaltii*, and *L. gormanii*, respectively.

We conclude, given these findings, that partial 16S rRNA gene sequencing is an acceptable alternative for the identification *Legionella* isolates to the genus level and for the differentiation of *L. pneumophila* from non-*pneumophila* *Legionella* species. This technology has the capability of accurately identifying *L. pneumophila* and also accurately identifies many of the non-*pneumophila* *Legionella* species.

The sequence-based identification of any microorganism is reliant on the integrity of the sequences within the database. This study identifies numerous GenBank (NCBI) entries that correlated with the species level identification achieved by traditional serotyping. Of the *L. pneumophila* isolates tested, GenBank entry AE017354 matched 21 isolates, whereas 2 matched best with accession number AJ496383 and 3 matched best with accession number AF129523. Of the non-*pneumophila* *Legionella* spp. whose sequence-based identification matched their serologic species determination, *L. bozemanii* matched with accession number Z49718, *L. longbeachae* with accession number AY444741, *L. micdadei* with accession number AF227162, *L. jordanis* with accession number X73396, *L. oakridgenesis* with accession number X73397, *L. israelensis* with accession numbers Z32640 and X73408, *L. wadsworthii* with accession number X73401, *L. tucsonensis* with accession number Z32644, *L. sainthelensis* with accession number X73399, *L. birminghamensis* with accession number Z49717, and *L. feelei* with accession number X73406. This study corroborates the validity of the GenBank sequence entries for the aforementioned *Legionella* spp. We believe that these sequences could be used with confidence for the sequence-based identification of the corresponding *Legionella* spp., if a high degree of complementarity is achieved (i.e., a high percent match).

REFERENCES

- Pedro-Botet, M. L., and M. Sabria. 2005. Legionellosis. *Semin. Respir. Crit. Care Med.* 26:625–634.
- Taylor, T. H., and M. A. Albrecht. 1995. *Legionella bozemanii* cavitary pneumonia poorly responsive to erythromycin: case report and review. *Clin. Infect. Dis.* 20:329–334.
- Reischl, U., H.-J. Linde, N. Lehn, O. Landt, K. Barratt, and N. Wellinghausen. 2002. Direct detection and differentiation of *Legionella* spp. and *Legionella pneumophila* in clinical specimens by dual-color real-time PCR and melting curve analysis. *J. Clin. Microbiol.* 40:3814–3817.
- Hayden, R. T., J. R. Uhl, X. Qian, M. K. Hopkins, M. C. Aubry, A. H. Limper, R. V. Lloyd, and F. R. Cockerill. 2001. Direct detection of *Legionella* species from bronchoalveolar lavage and open lung biopsy specimens: comparison of LightCycler PCR, in situ hybridization, direct fluorescence antigen detection, and culture. *J. Clin. Microbiol.* 39:2618–2626.