Serotyping of *Streptococcus pneumoniae* Isolates from Nasopharyngeal Samples: Use of an Algorithm Combining Microbiologic, Serologic, and Sequential Multiplex PCR Techniques

Karen Miernyk,1,2* Carolynn DeByle,2 Marcella Harker-Jones,2 Kimberlee Boyd Hummel,2 Thomas Hennessy,2 Jay Wenger,2 and Karen Rudolph2

*Alaska Native Tribal Health Consortium1 and Arctic Investigations Program, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Anchorage, Alaska*

Received 25 March 2011/Returned for modification 15 May 2011/Accepted 12 July 2011

We evaluated nasopharyngeal carriage of *Streptococcus pneumoniae* (pneumococci) in nine Alaskan communities and used an algorithm combining microbiologic, serologic, and sequential multiplex PCR (MP-PCR) techniques to serotype the isolates. After microbiological identification as pneumococci, isolates (n = 1,135) were serotyped using latex agglutination and Quellung tests (LA/Q) as well as a series of six sequential MP-PCR assays. Results from the two methods agreed for 94% (1,064/1,135) of samples. Eighty-six percent (61/71) of the discordant results were resolved. Discordant results occurred because (i) the MP-PCR gel was misread (31/61 [51%]), (ii) the LA/Q agglutination was misinterpreted (13/61 [21%]), (iii) two serotypes or sets of serotypes were identified by MP-PCR and only one of the two was identified by LA/Q (9/61 [15%]), (iv) different serotypes or sets of serotypes were identified by LA/Q and MP-PCR and both were correct (7/61 [11%]), and (v) the capsular polysaccharide locus (cps) did not amplify during the initial MP-PCR but was present upon retesting (1/61 [2%]). Overall, isolation of pneumococci followed by MP-PCR quickly and accurately identified pneumococcal serotypes in >97% of samples and made available isolates for additional tests such as antimicrobial susceptibility. Misinterpretation of the MP-PCR gel was identified as the main source of discordance. Increasing the number of MP-PCRs from six to seven and reducing the number of serotypes in each reaction may reduce this error. This method may be of use to laboratories characterizing large numbers of *S. pneumoniae* samples, especially when antimicrobial susceptibility data are needed.

*Streptococcus pneumoniae* (pneumococci) is a major contributor to morbidity and mortality worldwide, causing diseases such as acute otitis media, sinusitis, pneumonia, and meningitis. Additionally, it is part of the upper respiratory tract flora of healthy persons, especially children. Although colonization with *S. pneumoniae* is often asymptomatic, it can progress to serious respiratory disease or invasive pneumococcal disease (IPD), the development of which requires preceeding colonization (3, 9, 24).

Studies of pneumococcal colonization began in Alaska in an effort to understand how antimicrobial use affects the carriage of nonsusceptible pneumococci (5, 11, 12). These studies continued after introduction of the heptavalent pneumococcal conjugate vaccine (PCV7) to monitor carriage of PCV7 serotypes. Data from these studies show that PCV7 use resulted in declining rates of colonization by PCV7 serotypes and increased rates of colonization by non-PCV7 serotypes (10, 13). Increasing rates of non-PCV7 IPD in Alaska (23) and the introduction of a 13-valent pneumococcal conjugate vaccine (PCV13) have led us to continue studies of pneumococcal carriage to better understand transmission, serotype replacement, and the impact of conjugate vaccine.

Historically, serotyping of pneumococci has been accomplished by culture of the organism followed by serological identification of the capsular type using latex agglutination and Quellung tests (LA/Q) (2). This method is time-consuming and subject to interpretation; however, it does provide purified pneumococcal isolates for antimicrobial susceptibility testing and genetic characterization. Recently, some laboratories have begun using multiplex PCR (MP-PCR) assays to identify and serotype pneumococci collected from the nasopharynx (1, 7, 18). This method reduces the cost and time required to identify pneumococcal serotypes but does not provide isolates for further microbiological and molecular testing; additionally, it cannot presently identify all 92 known capsular types. We have developed a pneumococcal serotyping algorithm (Fig. 1) that includes microbiological isolation of pneumococci from nasopharyngeal swabs (NPS), followed by MP-PCR of the isolates and subsequent LA/Q for some isolates to resolve a serotype when two or more serotypes cannot be distinguished from each other using our MP-PCR assays. Here we present data to support use of the algorithm and highlight the contributions of MP-PCR and LA/Q by comparing initial results obtained by MP-PCR with LA/Q results.

**MATERIALS AND METHODS**

**Study participants and bacterial isolates.** During February to May 2009, we evaluated pneumococcal carriage in nine Alaskan communities by collecting nasopharyngeal samples using flocked swabs (Copan Diagnostics, Inc., Murrieta, CA). In Anchorage, AK, an urban community of >300,000 people, we recruited...
children 3 to 59 months of age receiving care at two pediatric clinics. In the eight rural communities (median population: 635), we recruited community residents of all ages. The study was approved by the Centers for Disease Control and Prevention (CDC) and the Alaska Area Institutional Review Boards as well as the Alaska Native Tribal Health Consortium and all appropriate tribal health boards. All participants or their legal guardians provided informed consent to participate.

After collection, NPS were immediately placed into 1.0 ml skim milk-tryptone-glucose-glycerol (STGG) medium (19) and frozen in a laboratory biocooler. The inoculated STGG remained frozen throughout transport to the Arctic Investigations Program laboratory in Anchorage, AK, where they were stored at −80°C until testing. For testing, frozen vials of STGG media containing the NPS were thawed and vortexed intermittently for 10 to 15 s and 50 μl was inoculated onto selective media (Trypticase soy agar with 5% sheep blood containing gentamicin). The inoculated plates were incubated overnight at 37°C with 5% CO₂. Colonies suspected of being alpha-hemolytic were subcultured onto Trypticase soy agar with 5% sheep blood (BAP) and incubated overnight at 37°C with 5% CO₂. Colonies that appeared morphologically different were subcultured onto separate BAP. Pneumococci were confirmed using susceptibility to optochin and bile solubility (22) and frozen in sheep blood at −80°C.

**DNA extraction.** Pneumococcal isolates were retrieved from storage, subcultured on BAP, and incubated overnight at 37°C with 5% CO₂. A 1-μl loop was
used to transfer bacterial cells to 100 μl of nuclease-free water. The suspensions were vortexed, heated at 100°C for 10 min, and centrifuged for 5 min at 13,000 rpm. The supernatant containing the crude DNA was removed and stored at -30°C until use.

**Serotyping.** Pneumococcal isolates were retrieved from storage, partially thawed, and inoculated onto BAP. After overnight incubation at 37°C with 5% CO₂, isolates were serotyped by LA/Q as previously described (2). Briefly, latex agglutination (LA) was used for screening the samples. This procedure uses antibody bound to polystyrene latex particles that agglutinates in the presence of homologous antigen. The confirmatory Quellung (Q) reaction involves the interaction between pneumococcal capsular polysaccharide and its homologous antibody. Microscopic swelling of the capsule creates a change in the refractive index of the bacteria and causes the bacteria to agglutinate. Capsular typing sera were prepared by Statens Serum Institut (Copenhagen, Denmark).

Capsular serotypes were subsequently determined from DNA extracted from the purified pneumococcal colonies using six sequential MP-PCR assays (Fig. 1) similar to the method initially described by Pai et al. (20). Laboratorians performing MP-PCR were blind to the LA/Q results. The MP-PCR method used in this study deviated from that of Pai et al. in that a primer set specific for the cps locus was included only in reaction 1. Samples negative for the cps locus in reaction 1 were classified as nontypeable (NT) or serotype 38 and were not tested again in subsequent reactions. Additionally, a single positive-control mixture was created for each of the six MP-PCRs. These control mixtures included DNA from all of the serotypes included in that reaction (Fig. 2). The positive-control mixtures were run alongside unknown samples in lanes 1 and 23 of each row of a 4 by 24 (96-well) gel and used as molecular weight markers to identify a serotype or set of serotypes in the unknown samples.

To minimize the number of MP-PCR assays performed, primers were grouped together so that the most common serotypes would be detected in the first three reactions. To design our algorithm, we used serotype prevalence data from a 2008 survey of pneumococcal carriage in the same nine communities surveyed for this study. When developing the MP-PCR algorithm, we also attempted to ensure that PCR products were easily separated on the gels, so each reaction includes serotypes whose products are greater than 100 bp apart. Based on the serotype distribution from the 2008 carriage survey, we expected this algorithm to identify approximately 47% of isolates after the initial reaction, 87% of isolates after the third reaction, and 96% of isolates after all six reactions.

With the exception of the five listed in Table 1, primer pairs used in this study were previously described by Pai et al. (20) (cps locus, serotypes 3, 4, 7A/7F, 7B/7C/40, 10A, 15A, 16F, 17F, 19A, 19F, 20, 23F, 31, 34, 33A/33F/37, 35B, and 35F/47F, and serogroups 6, 11, 18, and 22). Dias et al. (8) (serotypes 9L/9N and 14), and the CDC's *Streptococcus* laboratory (http://www.cdc.gov/nicidod/biotech/strep/pcr.htm) (serotypes 8 and 23A). Primer pairs listed in Table 1 were selected using Primer3 (21) software from sequences obtained from the National Center for Biotechnology Information (NCBI; accessed October 2008). Primer pairs were tested for potential cross-reactivity with nonhomologous serotypes and serogroups using the Basic Local Alignment Search Tool software from NCBI. They were also tested for specificity using DNA extracted from all serotypes included in our algorithm. All primers were synthesized at the Biotechnology Core Facility Branch, CDC.

MP-PCRs were performed in 25-μl reaction volumes that contained 2× multiplex PCR master mix (Qiagen Inc., Valencia, CA), 0.2 μM each primer, and 3 μl crude DNA extract. Reaction conditions were as follows: 95°C for 15 min followed by 25 cycles of 94°C for 30 s, 60°C for 90 s, and 72°C for 90 s and a final extension of 72°C for 10 min. The amplified products were analyzed on a 4 by 24 (96-well) 3% Tris-acetate-EDTA gels at 90 V for 37 min (Bio-Rad, Hercules, CA).

At the completion of all serotype testing, MP-PCR results were compared with those obtained by LA/Q. We attempted to resolve discordant results by performing singleplex PCR and additional LA/Q on the samples in discordance. So as to not introduce bias into the interpretation of the LA/Q results, the laboratorians doing the confirmatory LA/Q testing were blind to the original LA/Q and MP-PCR results. Because interpretation of singleplex PCR results is not subjective, confirmation by singleplex PCR was not performed blindly. Singleplex PCR was performed using primer pairs specific for each of the two serotypes or sets of serotypes that were discordant between MP-PCR and LA/Q. Samples that had the same original results by both methods after subsequent testing were reported as “not resolved.”

**RESULTS**

NPS were collected from 3,895 people (Fig. 1). Of these, 32% (1,242/3,895) carried pneumococci, resulting in 1,359

---

**TABLE 1. New oligonucleotide primers used in sequential MP-PCR assays for serotyping *S. pneumoniae* isolates from nasopharyngeal samples obtained in Alaska in 2009**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Product size (bp)</th>
<th>Primer sequence (5'-3')</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-f</td>
<td>942</td>
<td>GCA GAA AAA GAT GGG GAC AA</td>
<td>wzx</td>
</tr>
<tr>
<td>12-r</td>
<td>TAA AAT GCC CGC GAT TAA AG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15B/15C-f</td>
<td>350</td>
<td>ACC AGT TAC GGA GTA AGA TAT GCA G</td>
<td>wzy</td>
</tr>
<tr>
<td>15B/15C-r</td>
<td></td>
<td>CGT GTC ACA TTA CTA CCT GTT CTT G</td>
<td></td>
</tr>
<tr>
<td>21-f</td>
<td>831</td>
<td>TTC TTA AAA ATT ACG CCT ATA ATC TCT CTT</td>
<td></td>
</tr>
<tr>
<td>21-r</td>
<td>GTT ACA TTT TCT TCA CAG ACT TAT AAT CAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23B-f</td>
<td>226</td>
<td>GAC AGA GTA TGG TTT TTA TGA TTT GGT</td>
<td>wzx</td>
</tr>
<tr>
<td>23B-r</td>
<td>TTG AGA AAT TAA CTC CGA TAA AGA AAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29-f</td>
<td>217</td>
<td>CTA CGG CAA AGT TGG GAG TT</td>
<td></td>
</tr>
<tr>
<td>29-r</td>
<td>AAG CGA GAA GTA TCT GTA TGT CCA</td>
<td></td>
<td>wzy</td>
</tr>
</tbody>
</table>
samples. Both LA/Q and MP-PCR testing was completed on 84% (1,135/1,359) of the samples. The two serotyping methods showed complete concordance for 94% (1,064/1,135) of the samples, representing 27 different serotypes or sets of serotypes, including NT samples (Table 2). The serotype was determined in the first two MP-PCRs for 79% (840/1,064) of the samples with complete concordance. For 59% (633/1,064) of the concordant samples, the MP-PCR assays narrowed the results to 1 of 13 sets of serotypes (including NT samples); Quellung testing was required to further characterize these samples.

There were 71 discordant results, and of these, 61 (86%) were resolved using the described methods. There were five reasons for discordance: (i) the MP-PCR gel was misread, and the LA/Q result was confirmed (31/61 [51%]) (Table 3); (ii) LA/Q was misinterpreted, and the MP-PCR result was confirmed (13/61 [21%]) (data not shown); (iii) two serotypes or sets of serotypes were identified by MP-PCR, and only one of the two was identified by LA/Q (9/61 [15%]) (Table 4); (iv) different serotypes or sets of serotypes were identified by LA/Q and MP-PCR, and both were subsequently confirmed as correct (7/61 [11%]) (Table 4); and (v) the capsular polysaccharide locus (cps) did not amplify during MP-PCR 1, so the sample was reported as NT/serotype 38, but the cps locus and serotype-specific gene (serotype 21) were subsequently shown to be present (1/61 [2%]).

For 10/1,135 samples (0.9%), the correct serotype could not be resolved with the methods used in this study. The majority (7/10) of the unresolved results were isolates identified as serotype 29 by LA/Q and serotype 35B by MP-PCR. Additionally, two isolates were identified twice as NT by LA/Q but were shown by MP-PCR to contain the cps locus as well as a serotype-specific gene (serotypes 8 and 33A/33F/37), and one isolate was identified as serotype 19F by LA/Q and serotype 19A by MP-PCR. Excluding samples “not resolved,” MP-PCR accurately identified the serotype or set of serotypes for >97% (1,093/1,125) of the samples compared with standard serotyping methods.

**DISCUSSION**

The introduction of effective pneumococcal conjugate vaccines has stimulated interest in understanding aspects of nasopharyngeal carriage among populations receiving the vaccine, since carriage precedes invasive disease. However, using traditional LA/Q for large nasopharyngeal carriage studies is time-consuming, so we sought to use a combination of LA/Q and MP-PCR methods to improve efficiency. Here we have described an algorithm in which we initially isolate pneumococci from NPS stored in STGG medium and then perform MP-PCR supplemented with LA/Q to determine the serotype or set of serotypes.
of the isolated organisms. This strategy resulted in accurate serotyping data for >97% of the isolates compared with standard serotyping methods, simplified the laboratory work by decreasing the number of samples that require serotyping by LA/Q, and provided purified pneumococci for additional tests.

The successful use of genetic approaches to serotype pneumococci began by focusing on a few serotypes of epidemiological importance (4, 16). As the genetic sequences of more serotypes became available, accurate MP-PCR algorithms to identify an increasing number of serotypes were published (15, 20). Eventually researchers began manipulating these algorithms to identify prominent serotypes causing IPD in different regions of the world (8, 14, 17) and using the algorithms to serotype pneumococci from nasopharyngeal samples (1, 7, 18). The currently published reports describing MP-PCR methods to serotype pneumococci collected from the nasopharynx describe DNA templates prepared directly from STGG medium containing the NPS. These methods allow for identification of pneumococcal serotypes but do not provide purified isolates for antimicrobial susceptibility testing or genetic characterization, which are important parts of a comprehensive program for monitoring carriage of *S. pneumoniae*. The algorithm used in this study overcomes this limitation, as the initial step is to isolate and freeze pneumococcal colonies from the STGG medium. These colonies are then available for additional testing, if necessary.

The main disadvantage of using the published genetic approaches for serotyping pneumococci is that the assays currently cannot identify all known capsular types. Many serotypes are resolved to a set of genetically related serotypes (i.e., 6A/6B/6C, 7A/7F, and 15B/15C) rather than an individual serotype. The sequential MP-PCR used in this study can identify 31 different serotypes or sets of serotypes, including NT isolates. However, by initially isolating and freezing the organisms, all known serotypes can be distinguished because samples not identified by MP-PCR or identified as one of a set of serotypes can be removed from storage and grown on BAP and the serotype can be determined using conventional methods. Even though 59% of samples in our study required further LA/Q after completion of the MP-PCR portion of the algorithm, the additional work was simplified with the knowledge of the set of serotypes to which the samples belong.

The high degree of complete concordance (94%) between MP-PCR and LA/Q in our study confirms that MP-PCR can accurately identify serotypes of *S. pneumoniae* collected from the nasopharynx. Excluding those “not resolved,” MP-PCR correctly identified >97% of samples in our study. Additionally, MP-PCR accurately determined the serotype of 13 samples misclassified by LA/Q and identified a second serotype not found by LA/Q in nine samples.

Of the incorrect MP-PCR results, all but one was the result of misreading the MP-PCR gel; the bands were too close together to be resolved accurately. One of the variables to consider when testing >1,300 samples is how to organize the work to optimize time and resources. With this method, we created positive-control mixtures for each reaction, used them as molecular weight markers, and ran the reactions on 4 by 24 (96-well) gels. Organizing the work in this manner had the advantage of optimizing the number of samples tested at any one time; however, it sometimes made it difficult to determine the correct sizes of the unknown bands. Fortunately, the sequential MP-PCR algorithm is flexible and can be easily updated to reflect organizational changes or serotype fluctuations (8, 17, 20). As part of work done subsequent to this study, we redesigned our MP-PCR algorithm to attempt to make it easier to determine the sizes of the bands in our unknown samples. We increased the number of reactions from six to seven, which decreased the number of serotypes or sets of serotypes in each reaction from five to four. We specifically focused on those serotypes and sets of serotypes listed in Table 3 (serotypes 3, 10A, 15A/15F, 19A, 21, and 35F/47 and serogroups 9, 11, 18, and 22) which were difficult to identify in this study. For those serotypes we redesigned some of the primer sets so the product sizes were different from those in this study. We also moved most of the serotypes into different reaction mixtures from where they were previously, thus separating the two serotypes that were difficult to discriminate from each other. These changes have led to better band resolution and easier gel interpretation. Data from the updated algorithm have shown a 0.3% (3/975) error caused by misreading the gel. This compares to the 2.7% (31/1,135) error found in this study.

Pai et al. first described three samples that were not amplified by their internal control (cps locus) primer set (20). All three of these were NT by traditional serotyping methods. Although NT organisms are not often isolated from cases of IPD, they represent approximately 10% of isolates found in the nasopharynx (6). In our study they accounted for 12% of all samples, making them more common than any serotype except for serotype 19A. It is important, therefore, to identify these organisms in the first MP-PCR so resources are not spent on the additional testing of samples without a capsule. Based upon the report by Pai et al. (20), as well as work done in our laboratory previous to this study, we identified all samples not amplified by the primers specific for the cps locus as suspected NT, and they were removed from subsequent reaction mixtures. All samples containing the gene for the cps locus were thought to be serotypeable and continued through all six MP-PCR assays or until their serotypes were determined. With one exception, this method correctly identified every sample as either NT or serotypeable. The single exception was a sample that was initially not amplified by the primers specific for the cps locus but was serotype 21 by LA/Q. Subsequent singleplex PCR analysis confirmed that the sample did contain the cps locus and the serotype 21-specific gene. It is likely that the DNA template was inadvertently left out of the reaction tube during the initial MP-PCR.

It is important to note that data generated by this method could be slightly different from data generated by LA/Q and will reinforce the serotype frequency distribution used to develop the algorithm. In this study there were seven samples (<1%) where LA/Q identified one capsular type, most likely the more abundant serotype in the sample, and MP-PCR identified another, the serotype or set of serotypes tested earliest in the sequential MP-PCR algorithm. Confirmatory testing showed that both serotypes were present in these samples. In future testing where LA/Q is not completed on every sample, only the serotype that is tested for earliest in the MP-PCR algorithm will be identified. This is acceptable for our large, population-based studies but should be considered by groups doing vaccine studies on an individual level.
This method is also not suitable for detecting multiple colonizing serotypes unless they have different morphologies upon the initial subculture. Samples were not exhaustively searched for multiple serotypes using MP-PCR. If multiple serotypes are present, the more common serotype (the one tested for earliest in the MP-PCR algorithm) will most likely be the one identified.

In March 2010, the U.S. Food and Drug Administration licensed PCV13, and it replaced PCV7 in the Alaska childhood vaccine schedule in April 2010. Continued evaluation of serotype distribution and antimicrobial susceptibility of pneumococci carried in the nasopharynx will be critical for evaluating the impact of this new vaccine. We have described here a combined microbiologic, MP-PCR, and serologic algorithm to test for earliest in the MP-PCR algorithm) will most likely be the one identified.

Callicarriage in Alaska as a part of our program to evaluate the licensed PCV13, and it replaced PCV7 in the Alaska childhood vaccine schedule in April 2010. Continued evaluation of serotype distribution and antimicrobial susceptibility of pneumococci carried in the nasopharynx will be critical for evaluating the impact of this new vaccine. We have described here a combined microbiologic, MP-PCR, and serologic algorithm to be the one identified.

To plan we use this algorithm to monitor pneumococcal carriage in Alaska as a part of our program to evaluate the effectiveness of the new PCV13 vaccine, and it could be considered by all laboratories interested in characterizing large numbers of pneumococci either from the nasopharynx or from episodes of IPD.

ACKNOWLEDGMENTS

We thank the entire staff of the CDC’s Arctic Investigations Program for their contribution to the data, particularly Alisa Reasonover, Julie Morris, and Carolyn Zanis, who assisted with the microbiology work.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES


