

Use of Pyrosequencing To Differentiate *Streptococcus pneumoniae* Serotypes 6A and 6B

Rekha Pai,¹ Josef Limor,² and Bernard Beall^{1*}

Bacterial and Mycotic Diseases, Respiratory Diseases Branch,¹ and Scientific Resources Program, Biotechnology Core Facility Branch,² Centers for Disease Control and Prevention, Atlanta, Georgia

Received 29 April 2005/Returned for modification 30 May 2005/Accepted 8 June 2005

Accurate serotyping of *Streptococcus pneumoniae* remains important to monitor the changes in seroepidemiology of the organism over time. Though several PCR-based systems have been developed for this purpose, the cross-reactivity within serogroups often limits discrimination between types. All serogroup 6 isolates can be identified using a multiplex PCR system; however, due to the high sequence homology between the *cps-6B* and *cps-6A* loci, serotypes 6A and 6B cannot be differentiated by this method. We describe the use of pyrosequencing to reliably differentiate between serotypes 6A and 6B using a previously described single nucleotide polymorphism at codon 195 of the *cps* locus *wciP* gene. We observed complete concordance between capsular serotyping results and *wciP* pyrosequencing among 210 isolates examined, indicating that pyrosequencing is a rapid and accurate technique for deducing serotypes 6A and 6B.

Streptococcus pneumoniae is a major cause of morbidity and mortality worldwide, leading to ~1.2 million deaths among young children, mostly in developing countries (6). Immunity to pneumococcal infection is largely type specific with 90 distinct immunological types being described (4). However, not all known serotypes cause serious infections; ~15 serotypes cause the majority of invasive pneumococcal disease worldwide (3). Although this increases the probability of developing effective vaccines targeting the most frequent types, the distribution of serotypes can vary with age, geography, and time, posing greater challenges for vaccine development. Therefore, monitoring the changes in seroepidemiology of this organism is important, especially to evaluate the effect of newer vaccines.

Conventionally, serotyping is performed using the standard capsular reaction test (12), although the high cost of antisera and the technical expertise required are limitations of the system. The development of PCR-based serotyping systems has overcome some of the difficulties associated with serologic testing (1, 9), although most of the PCR systems described are limited to identifying serogroups that include cross-reacting serotypes (e.g., they can identify serotypes 6A and 6B only as serogroup 6), leading to incomplete serotype identification. The availability of the sequences of the capsular biosynthetic loci (*cps*) from all 90 pneumococcal serotypes at the website www.sanger.ac.uk and development of newer techniques such as pyrosequencing for accurate determination of single nucleotide polymorphisms (SNPs) may resolve this issue. Pyrosequencing is a real-time DNA sequencing technique for generating short reads rapidly and inexpensively (13). The method employs coupled enzymatic reactions to detect inorganic pyrophosphate (PP_i) released as a result of nucleotide incorporation by DNA polymerase. The released PP_i is converted to ATP by ATP sulfurylase, which provides the energy for luciferase to oxidize luciferin and generate light that is detected by a charge-coupled device camera and is seen as a peak in the pyrogram (7, 8, 13).

Recently, the difference between serotypes 6A and 6B has been putatively correlated with a single nonsynonymous substitution in the putative rhamnosyl transferase gene (*wciP*) (11). We utilized this information to test a pyrosequencing assay to reliably differentiate types 6A and 6B among isolates initially deduced as serogroup 6 from a PCR assay.

Primers targeting a *wciP* gene segment of types 6A and 6B, which encompasses the putative key polymorphism at codon 195 (see GenBank accession AF246897), were designed using the pyrosequencing assay design software (Biotage AB). *wciP*-f (5' AATTTGTATTTTATTCATGCCTATATCTGG) (biotinylated at 5' end) and *wciP*-r (5' TTAGCGGAGATAATTTAAATGATGACTA) were used to amplify a 250-bp product. PCR was done in 25- μ l volumes with 1 \times PCR buffer (Promega Inc., WI), 2.5 mM of MgCl₂, and 200 μ M of each deoxynucleoside triphosphate (New England Biolabs, MA) with 250 nM of each primer and 2.0 U of *Taq* DNA polymerase (Promega Inc., WI). Thermal cycling was performed with the following conditions: 94°C for 4 min and 30 amplification cycles of 94°C for 45 s, 55°C for 45 s, and 65°C for 2 min. Products were detected on 2% NuSieve agarose gels (Cambrex Bio Science, Inc., Rockland, ME).

Since multiplex PCR-based serotyping may become a viable alternative to conventional serotyping (1, 9), we incorporated primers *wciP*-f and *wciP*-r into a multiplex PCR also containing primers targeting *cps* loci of three of the most predominant serotypes seen in the Centers for Disease Control and Prevention's (CDC's) Active Bacterial Core surveillance (see <http://www.cdc.gov/ncidod/dbmd/abcs/>) during 2002 and 2003 (types 19A, 3, and 22F). The primer sequences, product sizes, and concentrations used for the reaction are listed in Table 1. All 250-bp PCR products obtained using the multiplex reaction were identified as serogroup 6 and were stored at 4°C until further use in pyrosequencing.

Pyrosequencing was performed with the Pyro Gold kit (Biotage AB, Uppsala, Sweden), and primer *wciP*-s (5' CCA

* Corresponding author. Mailing address: CDC Respiratory Diseases Branch, Mailstop C02, 1600 Clifton Rd., NE, Atlanta, GA 30333. Phone: (404) 639-1237. Fax: (404) 639-4215. E-mail: BBEALL@CDC.GOV.

cartridge was then loaded with deoxynucleoside triphosphates, enzyme, and substrate supplied with the kit, and the cartridge and the plate were placed in the instrument (Biotage AB) for analysis. Since each nucleotide addition takes 1 minute for completion, only 26 min was required to complete all 96 reactions.

Two hundred ten serogroup 6 isolates obtained through Active Bacterial Core surveillance were included for analysis (2). These included sterile-site isolates from both adults and children obtained during 2002 and 2003. In the first phase of this study, 175 isolates were amplified in PCRs containing *wciP*-f and *wciP*-r and subsequently pyrosequenced with *wciP*-s. Of the 175 products tested by pyrosequencing, 86 were identified as serotype 6B (Fig. 1) and 88 were serotype 6A (Fig. 1), correlating perfectly with conventional serotyping results. Though the genetic diversity among the isolates in this set was not determined, the strains of serotypes 6A and 6B have been previously found to be genetically diverse (5), and the assay can therefore be used to reliably differentiate serotypes even among strains with high genetic diversity. Only one isolate among our isolate set was consistently identified as serotype 6A by pyrosequencing and not 6B as originally recorded by serotyping. Repeat serotyping revealed that the original results were in error and that the isolate was actually type 6A.

In the second phase of this work, primers *wciP*-f and *wciP*-r were also incorporated as part of a multiplex PCR system used in part to identify serogroup 6 isolates, which also contained primers to identify serotypes 19A, 3, and 22F (CDC unpublished data), and 35 strains of serogroup 6 identified by the multiplex PCR were further subjected to pyrosequencing. There was no cross-reactivity between the primers in the multiplex PCR, and the specificity of each primer was checked against pneumococci of 59 different serogroups/types during initial standardization. Since the pyrosequencing includes a cleanup procedure before the sequencing reaction, the high primer concentration in the multiplex reaction mixture did not interfere with the pyrosequencing. All of the 35 isolates identified as serogroup 6 through multiplex PCR were differentiated as 6A or 6B by pyrosequencing, achieving an additional level of differentiation that we cannot currently achieve by conventional PCR. In addition to the SNP at residue 195, another nonsynonymous substitution that differentiated all 6A and 6B isolates tested was observed at codon 192, corresponding to an alanine-to-serine alteration (Fig. 1). It must be noted, however, that codon 192 was not entirely reliable in differentiating serotype 6A and 6B isolates (11).

Accurate determination of pneumococcal serotypes is a continuing priority, especially since the introduction of the heptavalent pneumococcal conjugate vaccine (PCV7). Monitoring of serotypes subsequent to introduction of the highly effective PCV7 is necessary to record the decline in vaccine types while simultaneously detecting possible serotype replacements with nonvaccine types. Though capsular testing has long been the preferred method for serotyping, the subjectivity in interpretation and the technical expertise required make alternate approaches attractive (1). The development of multiplex PCR systems that are rapid, reliable, and cost-effective has the capacity to make serotyping more affordable in resource-limited settings.

Pyrosequencing assays might generally allow distinguishing serotypes within strains that can currently only be serogrouped using PCR-based techniques. It is possible that similar assays can be developed to deduce the serotypes within other cross-reacting serogroups (e.g., serogroups 12, 18, 22, and 7). The assay described here rapidly and reliably detected the key polymorphism believed to be the genetic basis of the structural difference between the serotype 6A and 6B capsules (10). Though the initial setup costs for pyrosequencing are considerable, the cost per test is low at less than \$1 (United States). With pyrosequencing finding greater application for diagnosis of infectious and noninfectious diseases (10, 14), the cost may drop with time, making it more affordable. This is the first report on the use of pyrosequencing for deducing serotypes of *S. pneumoniae* and highlights the possibility of developing multiplex pyrosequencing systems for rapid and reliable determination of pneumococcal serotypes.

Rekha Pai was an International Emerging Infectious Diseases Fellow sponsored by the American Public Health laboratories and CDC.

We sincerely thank the Active Bacterial Core surveillance investigators for providing the clinical isolates used in the study. We thank the CDC National Vaccine Program Office and the CDC Antimicrobial Resistance Working Group for the necessary resources for this work. We are grateful to Robert E. Gertz for technical assistance.

REFERENCES

1. Brito, D. A., M. Ramirez, and H. De Lencastre. 2003. Serotyping *Streptococcus pneumoniae* by multiplex PCR. *J. Clin. Microbiol.* **41**:2378–2384.
2. Flannery, B., S. Schrag, N. M. Bennett, R. Lynfield, L. H. Harrison, A. Reingold, P. R. Cieslak, J. Hadler, M. M. Farley, R. R. Facklam, E. R. Zell, C. G. Whitney, et al. 2004. Impact of childhood vaccination on racial disparities in invasive *Streptococcus pneumoniae* infections. *JAMA* **291**:2197–2203.
3. Hausdorff, W. P., J. Bryant, C. Kloek, P. R. Paradiso, and G. R. Siber. 2000. The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulations and use, part II. *Clin. Infect. Dis.* **30**:122–140.
4. Henrichsen, J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **33**:2759–2762.
5. Gertz, R. E., Jr., M. C. McEllistrem, D. J. Boxrud, Z. Li, V. Sakota, T. A. Thompson, R. R. Facklam, J. M. Besser, L. H. Harrison, C. G. Whitney, and B. Beall. 2003. Clonal distribution of invasive pneumococcal isolates from children and selected adults in the United States prior to 7-valent conjugate vaccine introduction. *J. Clin. Microbiol.* **41**:4194–4216.
6. Global Programme for Vaccines and Immunization. 1996. Report of the meeting of Scientific Group of Experts (SAGE) of the Children's Vaccine Initiative and the Global Programme for Vaccines and Immunization. World Health Organization, Geneva, Switzerland.
7. Kobayashi, N., T. W. Bauer, D. Togawa, I. H. Lieberman, H. Sakai, T. Fujishiro, M. J. Tuohy, and G. W. Procop. 2005. A molecular Gram stain using broad range PCR and pyrosequencing technology: a potentially useful tool for diagnosing orthopaedic infections. *Diagn. Mol. Pathol.* **14**:83–89.
8. Langae, T., and M. Ronaghi. 2005. Genetic variation analyses by pyrosequencing. *Mutat. Res.* **573**:96–102.
9. Lawrence, E. R., D. B. Griffiths, S. A. Martin, R. C. George, and L. M. Hall. 2003. Evaluation of semiautomated multiplex PCR assay for determination of *Streptococcus pneumoniae* serotypes and serogroups. *J. Clin. Microbiol.* **41**:601–607.
10. Lee, S. S., K. M. Kim, H. Thi-Le, S. S. Yea, I. J. Cha, and J. G. Shin. 2005. Genetic polymorphism of CYP2C9 in a Vietnamese Kinh population. *Ther. Drug Monit.* **27**:208–210.
11. Mavroidi, A., D. Godoy, D. M. Aanensen, D. A. Robinson, S. K. Hollingshead, and B. G. Spratt. 2004. Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J. Bacteriol.* **186**:8181–8192.
12. Obaro, S., and R. Adegbola. 2002. The pneumococcus: carriage, disease and conjugate vaccines. *J. Med. Microbiol.* **51**:98–104.
13. Ronaghi, M., and E. Elahi. 2002. Pyrosequencing for microbial typing. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **782**:67–72.
14. Tuohy, M. J., G. S. Hall, M. Sholtis, and G. W. Procop. 2005. Pyrosequencing as a tool for the identification of common isolates of *Mycobacterium* sp. *Diagn. Microbiol. Infect. Dis.* **51**:245–250.