

# Sequential Multiplex PCR Approach for Determining Capsular Serotypes of *Streptococcus pneumoniae* Isolates

Rekha Pai, Robert E. Gertz, and Bernard Beall\*

Respiratory Diseases Branch, Division of Bacterial and Mycotic Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

Received 12 August 2005/Returned for modification 20 September 2005/Accepted 13 October 2005

**Accurate serotyping is essential to monitor the changes in the seroepidemiology of *Streptococcus pneumoniae*. We devised a simple and schematic sequence-based system of seven multiplex PCRs, in a sequence order based upon Active Bacterial Core surveillance (ABCs) serotype distribution during 2002 to 2003, to reliably deduce specific pneumococcal serotypes. A total of 421 isolates from ABCs were randomly chosen to evaluate this system. Two hundred twenty-nine of the isolates (54.3%) were specifically assigned 1 of 17 serotypes by the multiplex PCR system, with the results in complete concordance with conventional serotyping. One hundred seventy-two additional isolates (40.9%) were assigned to 11 specific sets of 2 to 4 serotypes that with one exception (serotypes 6A and 6B) consisted of the single frequently occurring targeted serotype and 1 to 3 additional rare serotypes primarily within the same serogroup as the targeted serotype. Only 20 isolates (4.8%) could not be assigned specific serotypes or serotype sets, since they were either of rare serotypes not included in the assay design or were nonserotypeable. Overall, we found this system to be highly reliable, with the potential to greatly reduce our reliance upon conventional serotyping. Especially important is the capability of this system to give serotype-determining potential to any facility that lacks the expensive typing sera and expertise needed for conventional serotyping yet has the modest equipment necessary for DNA amplification and electrophoresis.**

*Streptococcus pneumoniae* is an important pathogen that causes severe life-threatening illnesses in the elderly and children. Annually, this organism accounts for ~1.2 million deaths in children due to pneumonia and meningitis, mostly in developing countries (7). In the United States alone in 2000, *S. pneumoniae* caused an estimated 17,000 cases of invasive disease in children <5 years of age, including 700 cases of meningitis (1). Contributing to virulence is a capsular polysaccharide, the immunochemistry of which helps to differentiate pneumococci into 90 distinct serotypes (9). However, many of these serotypes are rarely recovered from serious disease, and only about 15 serotypes cause the majority of invasive pneumococcal disease worldwide (7, 23). Although this increases the probability of developing effective vaccines that target the most frequent types, the distribution of serotypes can vary with age, geography, and time, posing greater challenges for vaccine development.

The seven-valent pneumococcal conjugate vaccine (PCV7; Prevnar) was licensed for use in the United States in 2000, and a significant decline in invasive pneumococcal disease caused by the vaccine-targeted serotypes has been observed in young children (11, 28). However, little is known about the long-term impact of the vaccine, with early reports indicating some replacement of vaccine serotypes (VT) by non-vaccine types (NVT), both among the nasopharyngeal colonizers and among invasive isolates (6, 11, 28; CDC, unpublished data). Such changes in the seroepidemiology of the organism need to be constantly monitored to evaluate the effect and appropriateness of newer vaccines. Increases in disease caused by previ-

ously uncommon NVT could necessitate changes in vaccine composition, emphasizing the need for continued surveillance (15, 22).

Currently, serotype distribution is monitored by culture of the organism followed by serological determination of the capsular type by the standard capsular test (9, 16). The high cost of antisera, subjectivity in interpretation, and technical expertise requirements are serious drawbacks of the system. The development of PCR-based serotyping systems has the potential to overcome some of the difficulties associated with serologic testing (3, 13). In addition, the development of PCR-based assays for direct detection of select serotypes from clinical specimens could be a valuable aid in surveillance, particularly in situations where culture is insensitive (14, 25). Production of capsule is largely controlled by capsular polysaccharide synthesis genes located at the *cps* locus, typically with the same general genetic organization and flanked by the conserved *dexB* and *aliA* genes (21). The first four genes are conserved in almost all serotypes, while the central parts of the loci contain the serotype-specific genes that serve as the basis for differentiation of pneumococci by PCR-based approaches. Though these molecular approaches to serotyping appear promising, the existing systems identify a limited number of serotypes (3, 13, 14, 24) and need to be expanded to establish a more comprehensive approach for routine surveillance. Recently, an approach relying upon amplification of the entire *cps* locus, followed by restriction fragment length polymorphism analysis, was assessed (2). Although this scheme shows promise, we prefer an approach that does not include segments of the *cps* locus that are highly conserved between serotypes, but instead relies upon short sequences specific to individual serotypes or serogroups. The sequences of the *cps* loci from all of the known 90 pneumococcal serotypes have been completed

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

TABLE 1. Oligonucleotide primers used in this study

Primer pair	Sanger's strain no./ GenBank accession no. <sup>b</sup>	Primer sequence (5'→3')	Nucleotide position <sup>c</sup>	Product size (bp)
1-f 1-r	Z83335	CTC TAT AGA ATG GAG TAT ATA AAC TAT GGT TA CCA AAG AAA ATA CTA ACA TTA TCA CAA TAT TGG C	10469 10715	280
3-f 3-r	Z47210	ATG GTG TGA TTT CTC CTA GAT TGG AAA GTA G CTT CTC CAA TTG CTT ACC AAG TGC AAT AAC G	9179 9519	371
4-f <sup>a</sup> 4-r	AF316639	CTG TTA CTT GTT CTG GAC TCT CGA TAA TTG G GCC CAC TCC TGT TAA AAT CCT ACC CGC ATT G	9558 9957	430
5-f 5-r	AY336008	ATA CCT ACA CAA CTT CTG ATT ATG CCT TTG TG GCT CGA TAA ACA TAA TCA ATA TTT GAA AAA GTA TG	6123 6450	362
6A/B-f (biotin) <sup>d</sup> 6A/B-r	AF316640	AAT TTG TAT TTT ATT CAT GCC TAT ATC TGG TTA GCG GAG ATA ATT TAA AAT GAT GAC TA	8656 8877	250
7F-f 7F-r	554/62	CCT ACG GGA GGA TAT AAA ATT ATT TTT GAG CAA ATA CAC CAC TAT AGG CTG TTG AGA CTA AC	13356 14150	826
7C-f 7C-r	Sutcliff	CTA TCT CAG TCA TCT ATT GTT AAA GTT TAC GAC GGG A GAA CAT AGA TGT TGA GAC ATC TTT TGT AAT TTC	9438 9665	260
8-f 8-r	573/62	GAT GCC ATG AAT CAA GCA GTG GCT ATA AAT C ATC CTC GTG TAT AAT TTC AGG TAT GCC ACC	8084 8348	294
9V-f 9V-r	980/68	CTT CGT TAG TTA AAA TTC TAA ATT TTT CTA AG GTC CCA ATA CCA GTC CTT GCA ACA CAA G	12345 13070	753
10A-f 10A-r	1006/38	GGT GTA GAT TTA CCA TTA GTG TCG GCA GAC GAA TTT CTT CTT TAA GAT TCG GAT ATT TCT C	12423 13021	628
11A-f 11A-r	1813/39	GGA CAT GTT CAG GTG ATT TCC CAA TAT AGT G GAT TAT GAG TGT AAT TTA TTC CAA CTT CTC CC	11640 12071	463
12F-f 12F-r	6312	GCA ACA AAC GGC GTG AAA GTA GTT G CAA GAT GAA TAT CAC TAC CAA TAA CAA AAC	14407 14753	376
14-f 14-r	34359	CTT GGC GCA GGT GTC AGA ATT CCC TCT AC GCC AAA ATA CTG ACA AAG CTA GAA TAT AGC C	7968 8145	208
15A-f 15A-r	389/39	ATT AGT ACA GCT GCT GGA ATA TCT CTT C GAT CTA GTG AAC GTA CTA TTC CAA AC	7114 8313	436
15B/C-f 15B/C-r	7904/39	TTG GAA TTT TTT AAT TAG TGG CTT ACC TA CAT CCG CTT ATT AAT TGA AGT AAT CTG AAC C	7314 7779	496
16F-f 16F-r	Nr.34361	CTG TTC AGA TAG GCC ATT TAC AGC TTT AAA TC CAT TCC TTT TGT ATA TAG TGC TAG TTC ATC C	11521 12478	988
17F-f 17F-r	Rose (am) vac.	TTC GTG ATG ATA ATT CCA ATG ATC AAA CAA GAG GAT GTA ACA AAT TTG TAG CGA CTA AGG TCT GC	10484 11145	693
Sg18-f Sg18-r	4593/40	CTT AAT AGC TCT CAT TAT TCT TTT TTT AAG CC TTA TCT GTA AAC CAT ATC AGC ATC TGA AAC	12687 13230	573
19A-f <sup>a</sup> 19A-r	AF094575	GTT AGT CCT GTT TTA GAT TTA TTT GGT GAT GT GAG CAG TCA ATA AGA TGA GAC GAT AGT TAG	12118 12566	478
19F-f 19F-r	U09239	GTT AAG ATT GCT GAT CGA TTA ATT GAT ATC C GTA ATA TGT CTT TAG GGC GTT TAT GGC GAT AG	7828 8100	304
20-f 20-r	34365 (am)	GAG CAA GAG TTT TTC ACC TGA CAG CGA GAA G CTA AAT TCC TGT AAT TTA GCT AAA ACT CTT ATC	9567 10048	514
22F-f 22F-r	1772/40 (da)	GAG TAT AGC CAG ATT ATG GCA GTT TTA TTG TC CTC CAG CAC TTG CGC TGG AAA CAA CAG ACA AC	11055 11666	643
23F-f <sup>a</sup> 23F-r	AF057294	GTA ACA GTT GCT GTA GAG GGA ATT GGC TTT TC CAC AAC ACC TAA CAC ACG ATG GCT ATA TGA TTC	12118 12566	384

Continued on following page

TABLE 1—Continued

Primer pair	Sanger's strain no./ GenBank accession no. <sup>b</sup>	Primer sequence (5'→3')	Nucleotide position <sup>c</sup>	Product size (bp)
31-f 31-r	Nr.34374 (am)	GGA AGT TTT CAA GGA TAT GAT AGT GGT GGT GC CCG AAT AAT ATA TTC AAT ATA TTC CTA CTC	9144 9815	701
33F-f 33F-r	AY163221	GAA GGC AAT CAA TGT GAT TGT GTC GCG CTT CAA AAT GAA GAT TAT AGT ACC CTT CTA C	181 488	338
34-f 34-r	676/74	GCT TTT GTA AGA GGA GAT TAT TTT CAC CCA AC CAA TCC GAC TAA GTC TTC AGT AAA AAA CTT TAC	7350 7725	408
35B-f 35B-r	4356/39 (da)	GAT AAG TCT GTT GTG GAG ACT TAA AAA GAA TG CTT TCC AGA TAA TTA CAG GTA TTC CTG AAG CAA G	10556 11199	677
35F-f 35F-r	361/39 (da)	GAA CAT AGT CGC TAT TGT ATT TTA TTT AAA GCA A GAC TAG GAG CAT TAT TCC TAG AGC GAG TAA ACC	7374 7858	517
38-f 38-r	9687/39	CGT TCT TTT ATC TCA CTG TAT AGT ATC TTT ATG ATG TTT GAA TTA AAG CTA ACG TAA CAA TCC	13848 14392	574
<i>cpsA</i> -f <i>cpsA</i> -r	AF057294	GCA GTA CAG CAG TTT GTT GGA CTG ACC GAA TAT TTT CAT TAT CAG TCC CAG TC	2314 2473	160

<sup>a</sup> Primers previously published as in reference 3.

<sup>b</sup> Primers were designed using the sequence either from the GenBank (accession numbers provided) or from [www.sangers.ac.uk](http://www.sangers.ac.uk) (strain numbers provided).

<sup>c</sup> Start position of each primer.

<sup>d</sup> The biotinylated forward primer is required for subsequent resolution of serotypes 6A and 6B by pyrosequencing.

recently ([http://www.sanger.ac.uk/Projects/S\\_pneumoniae/CPS/](http://www.sanger.ac.uk/Projects/S_pneumoniae/CPS/)), providing an opportunity to develop a simple sequence-based scheme for identifying most commonly occurring serotypes.

The Active Bacterial Core surveillance (ABCs), part of the Centers for Disease Control and Prevention's (CDC) Emerging Infections Program, has conducted active, population-based, laboratory-based surveillance for invasive pneumococcal disease since 1995 (4). The surveillance program currently includes nine sites (as of 2003) with over 25,000,000 persons under surveillance (<http://www.cdc.gov/ncidod/dbmd/abc/>). Over 3,000 invasive ABCs pneumococcal isolates are currently received annually at the CDC *Streptococcus* Laboratory for serotype determination. We used serotype distribution data from the 2002-to-2003 ABCs to develop a simple, cost-effective, multiplex PCR-based approach for deducing serotypes within large sets of isolates.

#### MATERIALS AND METHODS

**Bacterial isolates.** Sixty-five clinical isolates of pneumococci representing different serotypes and serogroups were used as a control set to test primer specificity and to determine cross-reactivity. The serotypes included 1, 2, 3, 4, 5, 8, 13, 14, 20, 21, 31, 34, 37, 38, 39, 40, 44, 46, 6A, 6B, 7A, 7B, 7C, 7F, 9A, 9N, 9L, 9V, 10A, 10F, 11A, 11D, 11F, 12A, 12B, 12F, 15A, 15B, 15C, 15F, 16A, 16F, 17F, 18A, 18B, 18C, 18F, 19A, 19F, 19C, 22F, 22A, 23A, 23B, 23F, 25F, 28A, 24F, 33F, 33A, 35B, 35F, 35A, 35C, and 47F. An additional 5 to 10 isolates for each of the 29 serotypes targeted in this study were used to additionally assess the primer sets. Finally, a total of 421 additional pneumococcal isolates obtained from children and adults during the ABC surveillance period 2002 to 2003 were chosen for a blinded study to evaluate the assay. Each pneumococcal isolate was confirmed using standard microbiological tests, including colony morphology, optochin susceptibility, and bile solubility (26).

**Serotyping.** Preliminary serotyping (to detect pools) was performed by the latex agglutination test with capsular typing sera prepared at the CDC. All serotypes were confirmed using the Quellung reaction.

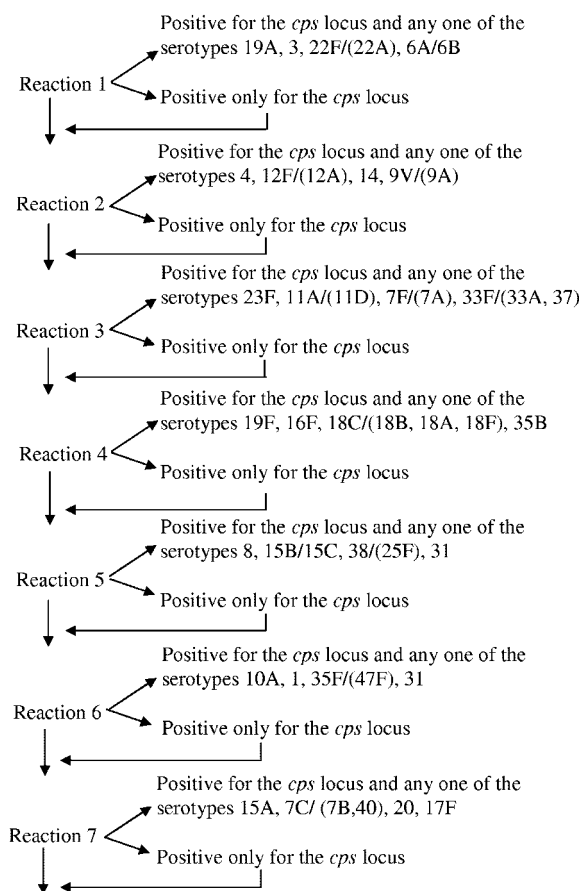
**Molecular capsular typing.** (i) **Oligonucleotide primers.** Twenty-nine primer pairs were designed to target serotypes 1, 3, 4, 5, 6A/B, 7F, 7C, 8, 9V, 10A, 11A, 12F, 14, 15A, 15B/C, 16F, 17F, 18, 19A, 19F, 20, 22F, 23F, 31, 33, 34, 35B, 35F,

and 38. A primer pair (primers *cpsA*-f and *cpsA*-r) was also included as an internal control targeting the *cpsA* locus found in all pneumococci (17). To design primers, all available relevant *cps* gene sequences were aligned (MegAlign) and primers were selected using the PrimerSelect program of DNASTAR version 5. The primers were designed to target the following genes (serotypes are indicated in parentheses): *cap1H* (1), *cap3C* (3), *wzy* (4, 33F, and 18), *wxy* (5), *wcip* (6), *cap8G* (8), *cps9VL* (9V), *cps14H* (14), *cps19aK* (19A), *cps19J* (19F), and *cps23fG* (23F). Although functional annotations were not available for the rest of the serotypes for which primers were designed, many of these sequences showed homology to other pneumococcal *cps* operon genes (e.g., *cps9VI* with 11A and 9V; *cps6bS* with 17F, 6A, and 6B; *cps4H* with serotypes 4 and 20; *cps19fG* with 19F and 35F; *cps19cS* with 22F and 19C; and *cps6bI* with 34, 6A, and 6B). Every primer pair was compared to all other *cps* sequences both at [http://www.sanger.ac.uk/Projects/S\\_pneumoniae/CPS/](http://www.sanger.ac.uk/Projects/S_pneumoniae/CPS/) and at [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) to determine the specificity and cross-reactivity with all other pneumococci. All primers were synthesized at the Biotechnology Core Facility Branch, CDC. The primer designations, sequences, and product sizes are listed in Table 1.

(ii) **Multiplex PCR scheme.** The primers were grouped into seven multiplex reactions as shown in Fig. 1, except the serotype 5 primer set was not included in any of these reactions. These primers were grouped together based on serotype distributions among invasive pneumococci recovered in the United States (7; CDC, unpublished data). Each reaction was designed to include four primer pairs targeting serotype-specific regions of four different serotypes and also included an internal positive control targeting all known pneumococcal *cps* operons (Fig. 1). In addition, the compatibility between primers included in two entirely different reactions was also assessed by modifying the multiplex PCRs 1 and 2, to include two PCRs that could detect the most prevalent types in other geographic regions (7). In the modified reactions, reaction 1 contained primers for serotypes 19A, 1, 6A/6B, and 19F while reaction 2 could detect serotypes 14, 5, 7F/7A (the rare serotype 7A was also detected by 7F primers), and 23F. These two reactions would detect the eight most predominant serotypes in regions of Asia and Africa (7).

(iii) **DNA extraction.** Pneumococcal isolates were retrieved from storage by subculture on blood agar plates (tryptic soy agar base supplement with 5% sheep blood) and incubated overnight at 37°C in 5% CO<sub>2</sub>. Bacterial cells were suspended in 250 μl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), and the turbidity was adjusted to that of MacFarland standard 1. The suspension was heated at 100°C for 5 min (13) and immediately frozen at -20°C for 5 min. These lysates were stored at -20°C until further use.

(iv) **PCRs.** The PCRs were performed in 25-μl volumes, with each reaction mixture containing the following: 1× PCR buffer (20 mM Tris-HCl, pH 8.0, 100



Conventional serotyping to deduce remaining serotypes (or if nontypeable)

FIG. 1. Schematic approach to molecular capsular typing, indicating the seven reactions and the serotypes detected in each reaction.

mM KCl, 0.1 mM EDTA, 1 M dithiothreitol, 0.5% Tween 20, 0.5% Nonidet P-40; Promega Inc., Madison, Wis.), 200  $\mu$ M of each deoxynucleoside triphosphate (New England Biolabs, Beverly, MA), 2.5 mM of  $MgCl_2$ , 2.0 U of *Taq* DNA polymerase (Promega Inc.), and primers with concentrations as specified in Table 2. Crude extract (2.5  $\mu$ l) was used as the DNA template for each PCR. Thermal cycling was performed in Perkin-Elmer GeneAmp PCR system 2700 (Applied Biosystems) under the following conditions: 94°C for 4 min followed by 30 amplification cycles of 94°C for 45 s, 54°C for 45 s, and 65°C for 2 min 30 s.

(v) **PCR product detection.** The PCR products were analyzed by gel electrophoresis on 2% NuSieve agarose gels (Cambrex Bio Science, Inc., Rockland, ME) in 1 $\times$  TAE buffer (40 mM Tris, 20 mM of glacial acetic acid, 1 mM EDTA, pH 8.0) at 120 V for 45 min. Gels were stained with ethidium bromide (0.5  $\mu$ g/ml), and gel images were recorded. The sizes of the PCR products were determined by comparison with the molecular size standard (50-bp ladder; Novagen, Inc.).

(vi) **Pyrosequencing assay.** All isolates identified as serogroup 6 by multiplex PCR were further differentiated as types 6A or 6B by a pyrosequencing assay designed to detect the single nucleotide polymorphism at codon 195 of the *cps* locus *wciP* gene as described previously (19). Pyrosequencing is a real-time DNA sequencing technique for generating short reads rapidly and inexpensively (24).

## RESULTS

**Optimization of PCR.** All serotype-specific primers were first tested with individual isolates of the targeted serotypes. To ensure that the primers would detect all strains within the given serotype, an additional 5 to 10 different clinical isolates within each targeted serotype were amplified with its corre-

sponding primer sets. While the annealing temperature in these individual reactions was at 60°C, the thermal cycling conditions were gradually altered when the primers were combined in multiplex reactions, with 2°C decrements in the annealing temperature to allow optimal annealing for all primers in the mixture (8). The final annealing temperature suitable for all reactions was 54°C. Similarly the optimal extension time and temperature were also determined. The final reactions yielded products ranging from 250 bp to 988 bp as shown in Table 1 and Fig. 2.

**Primer specificity.** Serotype or serogroup specificities of the specific primer pairs were assessed by amplifying each strain in the control set. Of the 29 primer pairs designed to target specific serotypes, 18 were completely specific for the targeted serotype and included primer pairs used to detect serotypes 1, 3, 4, 5, 8, 10A, 14, 15A, 15B/C, 16F, 17F, 19A, 19F, 20, 23F, 31, 34, and 35B. (We consider 15B and 15C as one serotype since they interconvert.) The serotype 6B primer set was also cross-reactive with the common serotype 6A *cps* operon sequence. Ten additional primer pairs were not completely serotype specific but were limited to an additional 1 to 3 rare serotypes, most within the same serogroup (Fig. 1 and Table 3). These included primer pairs for 22F, 12F, 9V, 11A, 7F, 33A, 18C, 38, 35F, and 7C.

TABLE 2. Primer concentrations in seven multiplex PCR<sup>a</sup>

Reaction	Primers	Primer concn ( $\mu$ M)
1	19A-f, 19A-r	1
	3-f, 3-r	1.5
	22F-f, 22-r	1.5
	6A/B-f (biotin), 6A/B-r	0.5
2	4-f, 4-r	1.5
	14-f, 14-r	1.0
	12F-f, 12F-r	1.5
	9V-f, 9V-r	1.5
3	23F-f, 23F-r	1.5
	7F-f, 7F-r	2.0
	11A-f, 11A-r	1.0
	33F-f, 33F-r	1.0
4	19F-f, 19F-r	1.5
	16F-f, 16F-r	2.0
	sg18-f, sg18-r	1.25
	35B-f, 35B-r	1.0
5	8-f, 8-r	1.5
	15B/C-f, 15B/C-r	1.5
	38-f, 38-r	1.5
	31-f, 31-r	2.0
6	1-f, 1-r	1.5
	10A-f, 10-r	1.5
	35F-f, 35F-r	1.5
	34-f, 34-r	1.5
7	20-f, 20-r	1.5
	7C-f, 7C-r	1.5
	17F-f, 17F-r	1.5
	15A-f, 15A-r	1.5

<sup>a</sup> The universal capsular primers *cpsA*-f and *cpsA*-r were included in all reactions at 0.5  $\mu$ M each.

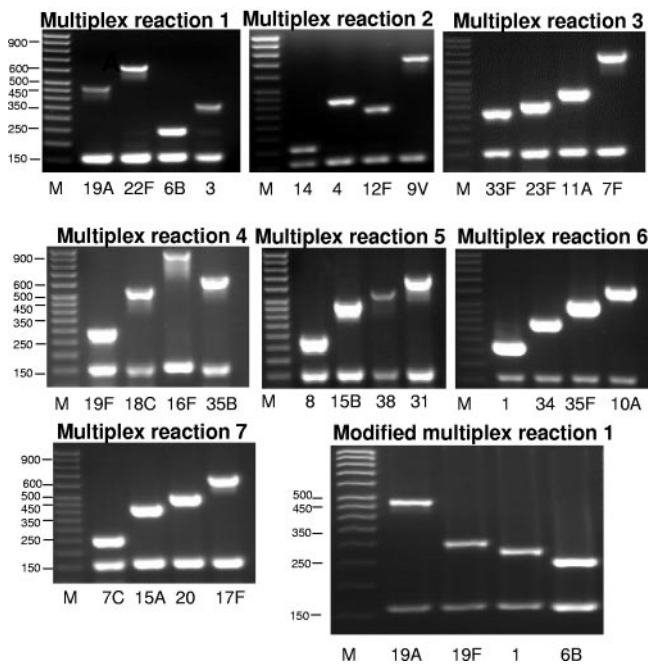


FIG. 2. Representative multiplex reactions. The serotypes of four strains tested for each of eight different multiplex formulations are shown below their respective lanes.

**Deduction of pneumococcal serotypes and serogroups using multiplex PCR.** Four hundred twenty-one random invasive pneumococcal isolates obtained during 2002 to 2003 were tested using the multiplex PCR system illustrated in Fig. 1. All results were in complete concordance with this scheme. For 54.3% (229/421) of the isolates, 1 of 17 individual serotypes was predicted, which was in complete concordance with conventional serotyping results (Table 3). These included eight isolates, representing five serotypes deduced by PCR, where the PCR results did not agree with conventional serotyping results (data not shown). When the conventional serotyping was repeated, it was clear that the original conventional test was in error since retesting matched PCR method results.

For 40.9% (172/421) of the isolates, the multiplex scheme narrowed the results to 1 of 11 serotype sets that included a major serotype plus 1 to 3 minor serotypes that were primarily within the same serogroup. For all of these except serogroup 6, only one serotype of the set is relatively common among ABCs, with the additional serotypes (in parentheses in Table 3) rare among current ABCs pneumococcal isolates (accounting for only 0 to 0.2% of isolates among year 2003 isolates;  $n = 2,798$ ). If potential cross-reactions were disregarded, the original targeted serotypes would have been correctly predicted among 95.2% of the isolates.

The 20 (4.7%) isolates for which serotypes could not be deduced by the system included 10 serotypes not included in the multiplex reaction scheme (Table 3). Three isolates that were repeatedly not amplified by specific primer sets as well as the internal control corresponded to two nontypeable isolates and a single type 9V isolate from original conventional serotyping. All three of these isolates were subsequently retested by conventional serotyping and found to be nonserotypeable.

Since the first three reactions in the sequential multiplex scheme are designed to cover the predominant serotypes within the United States, we attempted to make the system expeditious by performing PCR for the first three reactions simultaneously on all isolates. This allowed us to obtain results for 62.5% (263/421) of the isolates within the first three reactions and would theoretically identify ~68% of the ABCs isolates obtained during 2002 to 2003.

**Pyrosequencing.** The multiplex assay identified 35 of the 421 isolates as serogroup 6. Therefore, 14  $\mu$ l of the amplified product obtained through reaction 1 was used for the pyrosequencing assay as described previously. Pyrosequencing accurately identified the serotype of these isolates as serotypes 6A ( $n = 19$ ) and 6B ( $n = 16$ ).

**Modification of primer combinations.** While the seven reactions described above were designed to cover most of the predominant serotypes reported in the United States, reactions 1 and 2 were altered based on the prevalence data in other geographic regions (Asia and Africa) to include primers targeting serotypes 1, 5, 6A/B, 7F, 14, 19A, 19F, and 23F. These altered reactions (reaction 1 contained primers for serotypes 19A, 1, 6, and 19F and reaction 2 included serotypes 14, 5, 7F, and 23F) were a combination of serotypes different

TABLE 3. Serotyping results deduced by sequential multiplex PCR for 421 randomly chosen invasive pneumococcal isolates

Serotype result(s) from multiplex PCR <sup>a</sup>	Actual conventional serotyping result(s) (no. of positive isolates)	% of total ( $n = 421$ )
1	1 (13)	3.1
3	3 (24)	5.7
4	4 (23)	5.5
6A/6B, 35 isolates	6A (19), 6B (16)	8.3
7F/(7A)	7F (20), 7A (2)	5.2
7C/(7B)/(40)	7C (8), 7B (1)	2.1
8	8 (8)	1.9
9V/(9A)	9V (15)	3.6
14	14 (20)	4.8
20	20 (6)	1.4
31	31 (8)	1.9
34	34 (7)	1.6
38/(25F)	38 (8)	1.9
10A	10A (9)	2.1
11A/(11D)	11A (10)	2.3
12F/(12A)/(44)/(46)	12F (21)	4.9
15A	15A (9)	2.1
15B/C	15B/C (15)	3.2
16F	16F (9)	2.1
17F	17F (10)	2.3
18C/(18B)/(18F)/(18A), 3 isolates	18C (8), 18B (3), 18F (1), 18A (1)	3.1
19A	19A (34)	8.1
19F	19F (10)	2.3
22F/(22A)	22F (19), 22A (1)	5.0
23F	23F (16)	3.8
33F/(33A)/(37)	33F (10)	2.4
35B	35B (8)	1.9
35F/(47F)	35F (9)	2.1
Nontypeable	9N (4), 10F (1), 13 (2), 21 (2), 23A (1), 23B (3), 28A (1), 28F (1), 29 (1), 41A (1), NT (3)	4.8

<sup>a</sup> Serotypes in parentheses occur in 0 to 0.2% of ABCs isolates (according to year 2003 data).

from that in the original sequential PCR scheme. A 4% gel was used to detect these products (example shown in Fig. 2).

## DISCUSSION

Accurate serotype determination of *Streptococcus pneumoniae* is of critical importance as vaccine development presently relies on serotype prevalence data (7). Monitoring serotype prevalence patterns is also very important after the introduction of conjugate vaccines (7- to 11-valent), since the possibility of replacement of the vaccine types by nonvaccine types exists (15, 21, 27). For example, the recent increase in serotype 19A isolates obtained from children and adults in the postvaccine period (Jan 2003 to July 2004) highlights the need for constant serotype surveillance to determine the changing epidemiology of the organism (20). This has also led to renewed interest in developing accurate and efficient systems for pneumococcal capsular typing. Though the existence of 90 different capsular types makes it difficult to develop a typing scheme based on genetic approaches, the recent availability of sequence information of all serotypes ([http://www.sanger.ac.uk/Projects/S\\_pneumoniae/CPS/](http://www.sanger.ac.uk/Projects/S_pneumoniae/CPS/)) increases the prospect of achieving this objective. Also, not all of these 90 capsular types cause serious infections, and therefore development of a capsular typing scheme targeting most serotypes frequently associated with serious disease is a promising approach. We have attempted to develop a schematic approach to serotype the isolates obtained through our surveillance program, using 28 primer pairs divided among seven multiplex PCRs. Each multiplex reaction is designed to sequentially include the most frequently occurring serotypes, with seven reactions progressively covering about 88% of the isolates obtained each year through ABCs. The multiplex assay described here is supplemented by a pyrosequencing assay for resolving the highly homologous serotypes 6A and 6B (19). We also demonstrate the potential of altering the combinations of serotypes in different multiplex PCRs to more efficiently detect predominant serotypes in regions outside the United States.

Our study indicates that *cps* genes can be targeted to efficiently and reliably determine pneumococcal serotypes. The primers *cpsA-f* and *cpsA-r* were included in all reactions as an internal control targeting the *cpsA* gene, a highly conserved gene present in all pneumococci (18). This locus was amplified in all the strains characterized, except in four strains which were also nontypeable by conventional serotyping (three nontypeable and one type 9V rendered nontypeable on repeat capsular testing). The absence of amplification among such rough strains has been reported previously (3) and is probably due to the absence of the *cps* locus or *cps* sequence alterations. Seventeen of the primers in this sequential multiplex PCR system were found to be highly specifically associated with no cross-reactivity. These primers could reliably detect all isolates tested of corresponding types. This observation is particularly important as the primer designs for many serotypes have been based on the sequence information obtained from a single isolate. (All serotype sequences available at [www.sangers.ac.uk](http://www.sangers.ac.uk) are based upon one isolate of the corresponding serotype.) We did not find evidence suggesting that the genetic variability between different strains of the same serotype may affect its serotype determination by a PCR-based approach as suggested previously (29). Although we did analyze genetically diverse

sets of isolates for several serotypes (data not shown), it is still possible that strains from other countries of corresponding serotypes could present confounding *cps* region sequence diversity. We targeted central *cps* targets that were most apt to be serotype specific based upon current understanding of the organization of *cps* loci and based upon our own exhaustive screening of primers against all known *cps* locus sequences. However, not all primers included in these seven reactions were entirely specific. For example, the sequence homology between serotypes within serogroups 18 and 6 prevented the development of type-specific primers. Similarly, cross-reactivity between additional targeted serotypes (7F, 7C, 9V, 38, 11A, 12F, 22F, 33F, and 35F) and other serotypes is a problem. Fortunately, strains of the cross-reacting serotypes are rarely causes of serious invasive disease in the United States (7). Nonetheless, conventional serotyping will have to be performed to differentiate such isolates, although the procedure will be greatly streamlined with the knowledge that after the major multiplex reaction serotype, there are only one to three additional serotypes to screen. The PCR assays described so far have reported similar limitations in developing type specific primers among cross-reacting serotypes (3, 13, 25). The difference between serotypes 6A and 6B has been correlated to a single nonsynonymous substitution in the putative rhamnosyl transferase gene (*wciP*) (17), making it impossible to serotype such strains by a PCR-based approach. The pyrosequencing assay is highly reliable and cost-effective (19) and could differentiate all 35 isolates of serogroup 6. Though pyrosequencing may not be readily available at all centers, it is of potential use in resolving other sets of cross-reacting serotypes, when genetic polymorphisms specific to each serotype become available.

The sequential multiplex PCR assay described here is reliable and expeditious when the first three reactions are performed simultaneously for all isolates. The first three reactions detected 54.3% of the isolates included in this study. Since these reactions include 12 primer pairs that target the predominant serotypes associated with invasive disease in the United States, they would detect ~68% of all isolates obtained through the ABCs surveillance (1999 to 2003). Similarly, the serotype prevalence data obtained through other surveillance studies also indicate that these three reactions would help to detect the predominant serotypes seen in children <5 years of age within the United States (11). Another recent study of *S. pneumoniae* infection in neonates also shows that SGT 19, 9, and 3 account for ~60% of all invasive neonatal infections (10). Therefore, this schematic approach is particularly helpful to maximize detection of predominant types using a limited number of multiplex reactions. However, it is essential to consider the changing epidemiology and the increase in certain non-vaccine types such as serogroups 15 and 33, which are reported to have increased in children <5 years of age since the introduction of PCV7 (11). Though our scheme does not allow the detection of these serotypes within the first three reactions, both of these serogroups are represented in the subsequent reactions 4 and 5 and will be detected using this system. Also, we demonstrated the flexibility of this system to adapt to new surveillance data or areas through the usage of two altered reaction formulations targeting the predominant serotypes in Asia and Africa (7). This system can also be useful

in regions where no serotype data are available and could serve as a starting point to determine serotype frequencies.

There were eight instances where discrepancies between the multiplex PCR scheme and original conventional serotyping results were observed. In these instances, the PCR-based results proved to be accurate. While these discrepant results are difficult to explain, repeat serotyping has resulted in correction of serotype data at our center (5). Such errors have been reported by other reference laboratories (12) and emphasize the need for more reliable serotyping techniques. These errors/discrepancies also highlight the subjectivity of conventional serotyping, especially when differentiating the strains within the same serogroup. Serotypes 23F/23B, 35F/35B, and 7F/7C are easily resolved using our multiplex PCR scheme, with no subjectivity in interpretation (although the 7F and 7C reactions cross-react with the rare serotypes 7A and 7B, respectively).

This multiplex PCR system can be easily introduced in most clinical laboratories. The use of crude DNA extracts obviates the need for more specialized template preparation protocols and detects the most common serotypes using a minimal number of reactions, and products can be detected easily by gel electrophoresis. We have found that we can readily assign results to 96 isolates in a single day. We envision that most reference laboratories will in the future adopt a molecular approach to capsular typing. For example, the primers used in this system could potentially be used in the development of a multiplex oligo-capture system that relies upon a universal pair of primers for "long PCR" of all *cps* operons. We are currently investigating the feasibility of such an oligo-capture system.

The multiplex PCR system described in this work is cost-effective. Accounting for the fact that we produce our own typing sera, our estimated average cost per strain by a combination of latex agglutination (for pools) and Quellung reaction is \$5, and this increases up to \$10.98 for nontypeable strains. The average cost per isolate using commercially available antisera could be as high as \$28 (3). The cost per isolate when characterized by multiplex PCR would be as low as \$2.28 if a minimum of three reactions are performed and up to \$5.32 if all seven reactions are used. Our multiplex PCR system is designed to detect more than 60% of the strains in our surveillance, within the three reactions, dramatically reducing the costs compared to conventional serotyping. It must be mentioned that a recently described flow cytometric, multiplexed serologic approach appears to be very cost-effective in that it greatly reduces labor and dependence upon typing antisera (29). This methodology also has the capability to detect secondary serotypes within mixed cultures (29), which cannot be readily accomplished using a multiplex PCR approach. Another potential weakness of DNA-based approaches is the potential to assign serotypes to strains with defective *cps* operons, although such isolates are rarely recovered from invasive infections.

Although our multiplex PCR system identifies or greatly streamlines identification of the majority of predominant serotypes circulating within the United States, the number of primers will need to be increased for a comprehensive system. We are confident that the work we present here will serve as an excellent starting point for the development of a simple system

utilizing a single *cps* locus PCR product ([http://www.sanger.ac.uk/Projects/S\\_pneumoniae/CPS/](http://www.sanger.ac.uk/Projects/S_pneumoniae/CPS/)) (2) and specific oligonucleotides for the identification of all pneumococcal serotypes.

#### ACKNOWLEDGMENTS

Rekha Pai was an International Emerging Infectious Diseases postdoctoral fellow sponsored by Eli Lilly and Company in partnership with the Association of Public Health Laboratories (APHL), CDC, and the CDC Foundation. This work was also supported through the Active Bacterial Core surveillance-Emerging Infections Program, the National Vaccine Program Office, CDC's Antimicrobial Resistance Working Group, and CDC's Opportunistic Infections Working Group.

We sincerely thank the Active Bacterial Core surveillance personnel for data collection, and Cynthia Whitney for useful initial discussion concerning this project. We gratefully acknowledge the dedicated laboratory support of Tim Bailiff, Alma Ruth Franklin, Delois Jackson, Zhongya Li, Sandra Mathis, Varja Sakota, and Shantia Williams. We also thank Terry Thompson for verifying some serotyping results and providing cost estimates.

#### REFERENCES

1. **Advisory Committee on Immunization practices.** 1997. Prevention of pneumococcal disease: recommendations of the (ACIP). *Morb. Mortal. Wkly. Rep. Recomm. Rep.* **46**:1–24.
2. **Batt, S. L., B. M. Charalambous, T. D. McHughes, S. Martin, and S. H. Gillespie.** 2005. Novel PCR-restriction fragment length polymorphism method for determining serotypes or serogroups of *Streptococcus pneumoniae* isolates. *J. Clin. Microbiol.* **43**:2656–2661.
3. **Brito, D. A., M. Ramirez, and H. de Lencastre.** 2003. Serotyping *Streptococcus pneumoniae* by multiplex PCR. *J. Clin. Microbiol.* **41**:2378–2384.
4. **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.
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. **Ghaffar, F., T. Barton, J. Lozano, L. S. Muniz, P. Hicks, V. Gan, N. Ahmad, and G. H. McCracken, Jr.** 2004. Effect of the 7-valent pneumococcal conjugate vaccine on nasopharyngeal colonization by *Streptococcus pneumoniae* in the first 2 years of life. *Clin. Infect. Dis.* **39**:930–938.
7. **Hausdorff, W. P., J. Bryant, P. R. Paradiso, and G. R. Siber.** 2000. Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. *Clin. Infect. Dis.* **30**:100–121.
8. **Henegariu, O., N. A. Heerema, S. R. Dlouhy, G. H. Vance, and P. H. Vogt.** 1997. Multiplex PCR: critical parameters and step-by-step protocol. *Bio-Techniques* **23**:504–511.
9. **Henrichsen, J.** 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **33**:2759–2762.
10. **Hoffman, J. A., E. O. Mason, G. E. Schutze, T. Q. Tan, W. J. Barson, L. B. Givner, E. R. Wald, J. S. Bradley, R. Yogev, and S. L. Kaplan.** 2003. *Streptococcus pneumoniae* infections in the neonate. *Pediatrics* **112**:1095–1102.
11. **Kaplan, S. L., E. O. Mason, Jr., W. R. Wald, G. E. Schutze, J. S. Bradley, T. Q. Tan, J. A. Hoffman, L. B. Givner, R. Yogev, and W. J. Barson.** 2004. Decrease of invasive pneumococcal infections in children among 8 children's hospitals in the United States after the introduction of the 7-valent pneumococcal conjugate vaccine. *Pediatrics* **113**:443–449.
12. **Konradsen, H. B., et al.** 2005. Validation of serotyping of *Streptococcus pneumoniae* in Europe. *Vaccine* **23**:1368–1373.
13. **Lawrence, E. R., D. B. Griffiths, S. A. Martin, R. C. George, and L. M. C. Hall.** 2003. Evaluation of semiautomated multiplex PCR assay for determination of *Streptococcus pneumoniae* serotypes and serogroups. *J. Clin. Microbiol.* **41**:601–607.
14. **Lawrence, E. R., C. A. Arias, B. Duke, D. Beste, K. Broughton, A. Efratiou, R. C. George, and L. M. Hall.** 2000. Evaluation of serotype prediction by *cpsA-cpsB* gene polymorphism in *Streptococcus pneumoniae*. *J. Clin. Microbiol.* **38**:1319–1323.
15. **Lipsitch, M.** 2001. Interpreting results from trials of pneumococcal conjugate vaccines: a statistical test for detecting vaccine-induced increases in carriage of nonvaccine serotypes. *Am. J. Epidemiol.* **154**:85–92.
16. **Lund, E.** 1960. Laboratory diagnosis of Pneumococcus infections. *Bull. W. H. O.* **23**:5–13.

17. **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.
18. **Morona, J. K., R. Morona, and J. C. Paton.** 1999. Analysis of the 5' portion of the type 19A capsule locus identifies two classes of *cpsC*, *cpsD*, and *cpsE* genes in *Streptococcus pneumoniae*. *J. Bacteriol.* **181**:3599–3605.
19. **Pai, R., J. Limor, and B. Beall.** 2005. Use of pyrosequencing to differentiate *Streptococcus pneumoniae* serotypes 6A and 6B. *J. Clin. Microbiol.* **43**:4820–4822.
20. **Pai, R., M. R. Moore, T. Pilishvili, R. E. Gertz, C. G. Whitney, B. Beall et al.** 2005. Post vaccine genetic structure of *Streptococcus pneumoniae* serotype 19A from children in the United States. *J. Infect. Dis.* **192**:1988–1995. (First published 1 November 2005; doi:10.1086/49803.)
21. **Paton, J. C., and J. K. Morona.** 2000. *Streptococcus pneumoniae* capsular polysaccharide, p. 201–213. In V. A. Fischetti, R. P. Novick, J. J. Ferretti, D. A. Portnoy, and J. I. Rood (ed.), *Gram-positive pathogens*. American Society for Microbiology, Washington, D.C.
22. **Pelton, S. I.** 2000. Acute otitis media in the era of effective pneumococcal conjugate vaccine: will new pathogens emerge? *Vaccine* **19**:S96–S99.
23. **Robinson, K. A., W. Baughman, G. Rothrock, N. L. Barrett, M. Pass, C. Lexau, B. Lamaske, K. Stefonek, B. Barnes, J. Patterson, E. R. Zell, A. Schuchat, C. G. Whitney et al.** 2001. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States 1995–1998. *JAMA* **285**:1729–1735.
24. **Ronaghi, M., and E. Elahi.** 2002. Pyrosequencing for microbial typing. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **782**:67–72.
25. **Rubin, L. G., and A. Rizvi.** 2004. PCR-based assays for detection of *Streptococcus pneumoniae* serotypes 3, 14, 19F and 23F in respiratory specimens. *J. Med. Microbiol.* **53**:595–602.
26. **Ruoff, K. L., R. A. Whiley, and D. Beighton.** 1999. *Streptococcus*, p. 283–296. In P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. H. Tenover (ed.), *Manual of clinical microbiology*, 7th ed. American Society for Microbiology, Washington, D.C.
27. **Veenhoven, R., D. Bogaert, C. Uiterwaal, C. Brouwer, H. Kiezebrink, J. Bruin, E. Ijzerman, P. Hermans, R. de Groot, B. Zegers, W. Kuis, G. Rijkers, A. Schilder, and E. Sanders.** 2003. Effect of conjugate pneumococcal vaccine followed by polysaccharide pneumococcal vaccine on recurrent otitis media: a randomized study. *Lancet* **361**:2189–2195.
28. **Whitney, C. G., M. M. Farley, J. Hadler, L. H. Harrison, N. M. Bennett, R. Lynfield, A. Reingold, P. R. Cieslak, T. Pilishvili, D. Jackson, R. R. Facklam, J. H. Jorgensen, A. Schuchat et al.** 2003. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N. Engl. J. Med.* **348**:1737–1746.
29. **Yu, J., J. Lin, W. H. Benjamin, Jr., K. B. Waites, C.-H. Lee, and M. H. Nahm.** 2005. Rapid multiplex assay for serotyping pneumococci with monoclonal and polyclonal antibodies. *J. Clin. Microbiol.* **43**:156–162.