JCM
Journals.ASM.org

Sequetyping: Serotyping *Streptococcus pneumoniae* by a Single PCR Sequencing Strategy

Marcus H. Leung,[a] Kevin Bryson,[b] Kathrin Freystatter,[a] Bruno Pichon,[c]* Giles Edwards,[d] Bambos M. Charalambous,[a] and Stephen H. Gillespie[a]*

Centre for Clinical Microbiology, University College London Medical School, Royal Free Campus, London, United Kingdom[a]; Bioinformatics Group, Department of Computer Science, University College London, London, United Kingdom[b]; Respiratory and Systemic Infection Laboratory, Health Protection Agency Centre for Infections, London, United Kingdom[c]; and Scottish Haemophilus, Legionella, Meningococcal and Pneumococcal Reference Laboratory (SHLMPRL), Stobhill Hospital, Glasgow, United Kingdom[d]

The introduction of pneumococcal conjugate vaccines necessitates continued monitoring of circulating strains to assess vaccine efficacy and replacement serotypes. Conventional serological methods are costly, labor-intensive, and prone to misidentification, while current DNA-based methods have limited serotype coverage requiring multiple PCR primers. In this study, a computer algorithm was developed to interrogate the capsulation locus (*cps*) of vaccine serotypes to locate primer pairs in conserved regions that border variable regions and could differentiate between serotypes. *In silico* analysis of *cps* from 92 serotypes indicated that a primer pair spanning the regulatory gene *cpsB* could putatively amplify 84 serotypes and differentiate 46. This primer set was specific to *Streptococcus pneumoniae*, with no amplification observed for other species, including *S. mitis*, *S. oralis*, and *S. pseudopneumoniae*. One hundred thirty-eight pneumococcal strains covering 48 serotypes were tested. Of 23 vaccine serotypes included in the study, most (19/22, 86%) were identified correctly at least to the serogroup level, including all of the 13-valent conjugate vaccine and other replacement serotypes. Reproducibility was demonstrated by the correct sequetyping of different strains of a serotype. This novel sequence-based method employing a single PCR primer pair is cost-effective and simple. Furthermore, it has the potential to identify new serotypes that may evolve in the future.

Streptococcus pneumoniae, the causative agent of otitis media, community-acquired pneumonia, meningitis, and septicemia in industrialized and developing countries, is responsible for more than 1.6 million deaths annually, with most fatalities occurring in children and the elderly (68). The interaction between *S. pneumoniae* and its human host is usually benign, consisting of colonization of the mucosal surfaces of the upper respiratory tract. For example, over 90% of children by the age of 2 years have been colonized by pneumococci (23). The pneumococcal polysaccharide capsule enables the organism to evade host defenses on the mucosal surface during carriage and systemic defenses in invasive disease (2, 29, 45, 47, 49). To date, 93 capsular serotypes have been identified, each differing in chemical structure and immunogenicity (16).

Serotyping is required to monitor epidemiological trends following the introduction of serotype-specific polysaccharide conjugate vaccines, which include serotypes commonly encountered in carriage and disease. The introduction of these vaccines has led to increased monitoring of the serotypes causing invasive disease, and it is essential that a rapid serotyping method be developed to achieve this. There are reports of increases in disease incidences caused by non-PCV7 serotypes (3, 14, 53, 56). In 2010 a 13-valent conjugate vaccine (PCV13) was introduced to replace PCV7 and expand protective coverage. Continuous surveillance is crucial to monitor vaccine efficacy and serotype replacement in different geographical regions (1, 25, 26).

Conventional serotyping by the Quellung reaction is costly, time-inefficient, and error prone (13, 27, 37). As a result, DNA-based methods targeting the capsular polysaccharide synthesis (*cps*) locus are being developed (8, 11, 19, 34, 35, 39, 50–52, 54, 59, 60, 65). The first four genes at the 5′ end of the *cps* locus are involved in the regulation of capsule production and are conserved in all serotypes (7). The middle region of the operon contains serotype-specific genes that have been used as targets of PCR-based serotyping methods (8, 11, 34, 35, 39, 50–52, 54). While specificity and sensitivity are increased by these methods, a large number of serotype-specific primer pairs are required to provide broad serotype coverage. These methods are multiplex or multistep, increasing the complexity of the assays.

A cheap, efficient, and less complex method of serotyping with high serotype coverage is desirable. The ability to identify the pneumococcal serotype from a single amplification reaction would be a major advance from currently available techniques. We describe such a technique: a robust, sequence-based molecular "sequetyping" method from a single primer pair.

MATERIALS AND METHODS

Sequence alignment of *cps* in 23-valent polysaccharide vaccine serotypes. An *in silico* interrogation was performed on the *cps* locus sequences made available by the Wellcome Trust (7). Alignment of *cps* of pneumo-

Received 15 November 2011 Returned for modification 20 December 2011
Accepted 11 March 2012
Published ahead of print 2 May 2012

Address correspondence to Bambos M. Charalambous, b.charalambous@ucl.ac.uk.

* Present address: Bruno Pichon, Microbiology Services Division, Health Protection Agency, Laboratory of Healthcare Associated Infection, London, United Kingdom; Stephen H. Gillespie, School of Medicine, University of St. Andrews, St. Andrews, Fife, United Kingdom.

Supplemental material for this article may be found at http://jcm.asm.org/.

doi:10.1128/JCM.06384-11

TABLE 1 Specificity of pneumococcal sequetyping against other bacterial species[a]

Organism	Type strain	rRNA[b]	Expected 1,061-bp band
Streptococcus pneumoniae	ATCC 49619	+	+
Streptococcus gordonii	NCTC 7865	+	−
Streptococcus anginosus	NCTC 10713	+	−
Streptococcus intermedius	NCTC 2227	+	−
Streptococcus mitis	NCTC 12261	+	−
Streptococcus mutans	NCTC 10449	+	−
Streptococcus parasanguinis	NCTC 55898	+	−
Streptococcus salivarius	NCTC 8618	+	−
Streptococcus sanguinis	NCTC 7863	+	−
Streptococcus sobrinus	NCTC 12279	+	−
Streptococcus cristatus	ATCC 51100	+	−
Streptococcus pseudopneumoniae	BAA-960	+	−
Streptococcus oralis	NCTC 11427	+	−
Haemophilus influenzae	ATCC 10211	+	−
Moraxella catarrhalis	ATCC 25238	+	−
Klebsiella pneumoniae	ATCC 13883	+	−
Escherichia coli	ATCC 51299	+	−
Escherichia faecalis	ATCC 25922	+	−

[a] ATCC, American Type Culture Collection; NCTC, National Collection of Type Cultures. +, detected; −, not detected.
[b] rRNA PCR was performed as described by Woo et al. (67).

coccal serotypes constituting the 23-valent pneumococcal polysaccharide vaccine (PPV23) was performed to identify 500- to 1,100-bp variable regions that are bordered by conserved primer sites for PCR amplification that could differentiate between serotypes. Alignment of *cps* sequences for the 23 vaccine serotypes was performed with MAFFT version 6.624b online interface (http://mafft.cbrc.jp/alignment/server/) using the iterative refinement method (NS-i). ClustalX 2.0.10 was used to analyze the alignment generated by MAFFT and generate a q-scores file. Each nucleotide position was scored according to its conservation across all 23 serotype sequences. The q-scores file was then entered into an algorithm (PrimerFinder algorithm) to locate conserved primer-binding sites (K. Bryson, unpublished data). After identification of an optimal primer pair based on maximum conserved binding sites and sequence differentiation between all 23 serotypes, the region to be amplified by this primer pair was analyzed *in silico* for 92 published serotypes. These included 90 serotypes from the Wellcome Trust plus serotypes 6C (GenBank accession code EF538714) and 6D (accession code HM171374) (7, 10, 52). The corresponding region in the recently discovered serotype 11E (16) was unavailable at the time of the study design and is not included in the analysis.

Bacterial strains. Pneumococci with known serotypes and multilocus sequence types (STs) from upper respiratory tract samples (URT) were isolated from a cohort of asymptomatic children under the age of six in Tanzania (6, 17, 42). In addition, laboratory pneumococcal reference strains with known serotypes (some with known STs) from GR MICRO, Health Protection Agency (HPA) laboratories in the United Kingdom (Colindale and Glasgow), and the American Type Culture Collection (ATCC) were included. All pneumococcal strains were isolated and archived between the years 2002 and 2010. Nonpneumococcal reference strains were acquired from the diagnostic microbiology laboratory of the Royal Free Hospital and the Eastman Dental Institute (Table 1).

Media and culture conditions. Bacterial cells were stored in skim milk, tryptone, glucose, and glycerol (STGG) medium at −80°C and were plated onto Columbia blood agar (CBA) (Oxoid) and cultured overnight at 35°C with 5% CO_2. Pneumococcal isolates were confirmed by colony morphology, alpha-hemolysis, optochin sensitivity, bile solubility, and reaction with anticapsular antibodies as described previously (42).

Conventional and multiplex-PCR serotyping. Serotyping was performed with pooled and type sera (Statens Serum Institut, Copenhagen, Denmark) using the checkerboard method (57) in the Royal Free Hospital Microbiology Department and HPA (Colindale, United Kingdom) and the Scottish Streptococcal Reference Laboratory. For some strains, multiplex-PCR methods designed previously were employed to confirm strain serotypes by targeting serotype-specific *cps* regions (35, 51). Strains conventionally assigned as serogroup 6 but that failed to react with serotype 6B-specific type serum were serotyped by the PCR method described by Park et al. (52) to differentiate between serotypes 6A and 6C.

PCR amplification of *cpsB* for sequetyping. Genomic DNA was extracted from bacterial cells from a fresh overnight culture by the heat lysis method (42). The sequetyping primers are as follow: *cps1*, 5′-GCA ATG CCA GAC AGT AAC CTC TAT-3′, and *cps2*, 5′-CCT GCC TGC AAG TCT TGA TT-3′. To match the annealing temperature of the reverse primer, the forward primer was extended in the 5′ end by three nucleotides compared to the *in silico* primer sequence. The reaction mixture contained 2 μl of genomic DNA, 0.3 μl *Taq* polymerase (5 U/μl) (Invitrogen), 1× PCR buffer (10×), 1.5 mM $MgCl_2$, 0.6 mM deoxynucleoside triphosphates (dNTPs), and 0.4 μM each primer, made up to a final volume of 50 μl with DNase/RNase-free distilled water (Gibco). The reaction cycle consisted of an initial denaturation step at 95°C for 5 min, followed by 30 amplification cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 1 min 30 s. Amplicons from PCR were analyzed with 1.5% agarose gel electrophoresis.

16S rRNA gene primers. A second primer pair targeting the 16S rRNA gene was used as an internal control for testing species specificity of the primers (67). The reference strains listed in Table 1 were used. The thermal cycle conditions for this multiplex PCR were modified from those described by Woo et al. (67), and 0.5 μM each primer targeting the rRNA gene was used.

Nucleotide sequencing. Amplicons with the expected *cps* band sizes (~1,000 bp) were purified using the PCR purification kit (Qiagen) according to the manufacturer's instructions. Purified amplicons were cycle sequenced in both directions with the same primers as those used in the initial PCR. Cycle sequencing was performed with BigDye Sequence Terminator v.3.1 (Applied Biosystems) according to the manufacturer's protocol. Sequences were analyzed on Genetic Analyzer 3130 (Applied Biosystems) and viewed with Bionumerics software version 5 (Applied Maths).

Sequetyping by sequence comparison. The amplicon nucleotide sequences were used to interrogate the GenBank database, and if the highest BLAST bit score (usually >98% identity) gave the correct serotype, identification was defined as a "match." When the highest bit score was shared between two or more serotypes (i.e., the same amount of nucleotide variation between query and database sequences), this was defined as either serogroup specific (if the other serotype[s] was within the serogroup) or "consistent" (if a different serogroup had a matching top score). Where the query results did not correlate or were not consistent with conventional serotyping methods, the strain/serotype was defined as "misidentified." For strains with discrepant results between conventional serotyping and sequetyping, both typing methods were independently repeated blindly from frozen culture.

Nucleotide sequence accession numbers. Nucleotide accession numbers for sequences generated in this study are listed in Table 2.

RESULTS

***In silico* identification of the optimal primer pair.** *In silico* evaluation has identified a region with high sequence variation among the 23 vaccine serotypes with a 1,061-bp amplicon that is 1,351 bp downstream of *cpsA* to 224 bp downstream of *cpsC* (nucleotide positions 305125 to 306185 on the G54 genome, GenBank accession number NC011072). Subsequent *in silico* analysis of the 92 serotype sequences revealed that with each primer binding site set at a mismatch of no more than two nucleotides, up to 84 of the 92

TABLE 2 Serotypes and strains included in this study and sequetyping results based on *cpsB* sequences[a]

Serotype	Sequetype	Strain	Sequence type[b]	Isolation source[c]	Sequetype accession no.
1	1	3OP2 Sep	217	NP	JN642309
	1	H0 8114 0126	217	BC	JN660116
	1	H0 8212 0279	303	BC	JN660117
	1	H0 8078 0043	618	BC	JN660118
	1	H0 6274 0473	304	BC	JN660119
	1	H0 8228 0507	227	BC	JN660120
	1	H0 6196 0204	228	BC	JN660121
	1	H0 5122 0138	306	BC	JN660122
3	3	H0 9260 0327	180	BC	JN660145
	3	12.1690.X	180	BC	JQ743514
	3	12.1655.E	180	BC	JQ743515
	3	12.1640.H	180	BC	JQ743516
	3	12.1409.G	180	BC	JQ743517
	3	12.1259.V	180	BC	JQ743518
4	4	H0 9146 0234	246	BC	JN660124
	4	H0 8102 0237	246	BC	JN660125
	4	H0 7406 0041	246	BC	JN664256
	4	11.6349.S	205	BC	JQ743519
	4	11.5324.N	246	BC	JQ743520
	4	11.4573.Q	246	BC	JQ743521
5	5	H0 8034 0160	289	BC	JN660126
	5	02-4520	ND	BC	JQ743522
	5	04-2077	1400	BC	JQ743523
	5	07-2667	289	BC	JQ743524
	5	10-1351	4840	Eye	JQ743525
6A	6A	H0 8212 0259	65	BC	JN660127
	6A	N15	65	BC	JN680106
	6A	N94	65	BC	JN680118
	6A	N155	65	BC	JN680128
	6A	N259	65	BC	JN680138
	6A	N405	65	BC	JN680146
	6A/6B	12.1624.T	396	BC	JQ743526
	6A	11.6939.E	2467	BC	JQ743527
	6A	11.5544.Z	1876	BC	JQ743528
	6A/6B	11.5346.W	327	BC	JQ743529
	6A	11.3085.J	65	BC	JQ743530
6B	6B	H0 8052 0052	176	BC	JN660128
	6B	H0 7156 0309	3481	BC	JN660129
	6B	11NP10 Jan	4432	NP	JN642310
	6B	69OP10	4429	NP	JN642311
	6B	35NP1	4157	NP	JN642312
	6B	35NP6	854	NP	JN642313
	6B	3OP7 Sep	4373	NP	JN642314
	6B	3NP7 Mar	4368	NP	JN642315
6C	6C/D	H0 5252 0052	1390	BC	JN660130
7C	7C	1003-2	ND	NP	JN642316
	7C	2514-1	ND	NP	JN642317
7F	7F/7A	SG07F	ND	C	JN660086
	7F/7A	12.1695.Y	191	BC	JQ743531
	7F/7A	12.1675.H	191	BC	JQ743532
	7F/7A	12.1666.R	191	BC	JQ743533
	7F/7A	12.1426.R	191	BC	JQ743534
	7F/7A	12.1299.R	191	BC	JQ743535
	7F/7A	12.1276.C	191	BC	JQ743536

(Continued on following page)

TABLE 2 (Continued)

Serotype	Sequetype	Strain	Sequence type[b]	Isolation source[c]	Sequetype accession no.
8	8	H0 8342 0074	53	BC	JN660132
	8	H0 9122 0175	53	BC	JN660133
9N	9N	H0 7174 0058	66	BC	JN660134
	9N	H0 7018 0063	66	BC	JN660146
	9N	H0 9080 0083	66	BC	JN660135
9V	9V	H0 7016 0058	156	BC	JN660136
	9V	H0 8156 0265	162	BC	JN660141
	9V	11.5715.C	156	BC	JQ743537
	9V	11.3601.D	162	BC	JQ743538
	9V	11.3021.E	162	BC	JQ743539
	9V	11.2143.S	156	BC	JQ743540
10A	10A	SG10a	ND	C	JN660087
	10A	21NP1	ND	NP	JN642318
	10A	18NP10	852	NP	JN642319
10F	10F/10C	SG10f	ND	C	JN660088
11A	11A/11D/18F	SG11a	ND	C	JN660089
	11A/11D/18F	29OP10	5752	NP	JN642320
12B	12B	SG12b	ND	C	JN660091
12F	12B	SG12f	ND	C	JN660090
14	14	H0 8208 0041	9	BC	JN660147
	14	H0 8396 0107	124	BC	JN660142
	14	H0 8084 0056	124	BC	JN660143
	14	H0 7442 0047	124	BC	JN660144
	14	10-1688	ND	S	JQ743541
	14	10-2893	ND	S	JQ743542
15A	15A	SG15a	ND	C	JN660092
15B	15B	SG15b	ND	C	JN660093
16F	16F	SG16f	ND	C	JN660094
17A	34/17A	SG17a	ND	C	JN660095
17F	17F/33C	H0 9084 0082	392	BC	JN660140
	4/17F/33C/9V	22NP2	4160	NP	JN642321
	17F/33C	H0 6092 0119	392	BC	JN660137
	17F/33C	H0 8334 0064	964	BC	JN660138
18B	18B/18C	SG18b	ND	C	JN660096
18C	18B/18C	10.2671.G	113	BC	JQ743543
	18B/18C	10.2178.X	113	BC	JQ743544
	18B/18C	09.2555.S	113	BC	JQ743545
	18B/18C	09.2223.P	638	BC	JQ743546
	18B/18C	09.1742.V	2449	BC	JQ743547
	18B/18C	09.1153.K	4162	BC	JQ743548
19A	19A	35NP10 (4162)	4162	NP	JN642322
	19A	H0 9186 0354 (199)	199	BC	JN664257
	19A	12.1625.M	2081	BC	JQ743549
	19A	12.1623.F	461	BC	JQ743550
	19A	12.1579.M	450	BC	JQ743551
	19A	H0 7336 0087	276	BC	JN664258

(Continued on following page)

TABLE 2 (Continued)

Serotype	Sequetype	Strain	Sequence type[b]	Isolation source[c]	Sequetype accession no.
19F	19F	ATCC 49619	ND	C	JN642323
	19F	16OP4	347	NP	JN642324
	19F	35NP8	6170	BC	JN642325
	19F	H0 8112 0101	162	BC	JN664259
	19F	H0 8242 0108	162	BC	JN664260
	19F	11.6554.S	ND	Sputum	JQ743552
	1[d]	11.6402.H	ND	Sputum	JQ743553
20	20/13	SG20	ND	C	JN660097
21	21	45OP9 Jan (1145)	1145	NP	JN642326
	21	7NP4	ND	NP	JQ009436
22A	22F/22A	SG22a	ND	C	JN660098
22F	22F/22A	SG22f	ND	C	JN660099
23B	23B	SG23b	ND	C	JN660100
	23B	7OP9	ND	NP	JN642327
23F	23F	SG23f	ND	C	JN664261
	23F	11.4091.H	ND	BC	JQ743554
	23F	11.3056.E	33	BC	JQ743555
	23F	11.2827.S	1682	BC	JQ743556
	23F	11.2737.T	1682	BC	JQ743557
	23F	11.1373.Z	6959	BC	JQ743558
24B	24B	SG24b	ND	C	JN660101
24F	24B	SG24f	ND	C	JN660102
27	27	H0 8432 0293	1475	BC	JN660139
28A	28A	SG28a	ND	C	JN660103
28F	28A	SG28f	ND	C	JN660104
29	29	SG29	ND	C	JN660105
31	31	SG31	ND	C	JN660106
33B	33B	SG33b	ND	C	JN660108
33C	35B/35C	SG33c	ND	C	JN660109
33D	33B	SG33d	ND	C	JN660110
33F	33A/33F/35A	SG33f	ND	C	JN660111
	33A/33F/35A	SG33fb	ND	C	JN660107
34	34/17A	SG34	ND	C	JN660112
35A	35A/33F/33A	SG35a	ND	C	JN642328
	35B/35C	16NP10	840	NP	JN660113
	35B/35C	55OP3	ND	NP	JQ009437
35B	35B/35C	SG35b	ND	C	JN660114
36	7F/21	SG36	ND	C	JN660115

[a] Serotypes covered in PPV23, PCV7, and PCV13 are in bold.
[b] ND, no data.
[c] NP, nasopharynx; BC, blood culture; C, control laboratory strain.
[d] Strain 11.6402 (serotype 19F) was sequetyped as serotype 1, with 98% identity (715/732) with the variation at the 5′ end of the gene. Note that the correct 19F serotype was identified at 97% homology.

serotypes could yield an amplicon (see Table S1 in the supplemental material).

As some of the amplifiable serotypes share identical interceding sequences (e.g., sequences for serotypes 2 and 41A, as well as 7B and 40 [Table 2]), a subset of 54 of the 84 amplifiable serotypes would be differentiated by *in silico* sequence analysis of the amplicon (see Table S1 in the supplemental material). However, due to the sequencing chemistry, the entire 1,061-bp region between the primers is unlikely to be sequenced successfully, because the ~200 bp at each end may yield suboptimal sequence data. *In silico* analysis was therefore repeated to determine the serotypes that could be sequetyped when the central 732-bp region of the *cpsB* amplicon was used. This predicted that 46 of the 54 serotypes could be sequetyped (see Table S1 in the supplemental material). Of the eight serotypes that putatively would not be sequetyped by *cpsB* alone, three pairs were sequetyped to the correct serogroup, namely, 6C/6D, 7F/7A, and 18B/18C. Serotypes 17F and 33C have identical *cpsB* sequences and were not predicted to be differentiated from each other.

Sequences of serotypes 25A, 25F, and 38 were predicted to be nonamplifiable, as the forward primer binding site has six mismatches, while the reverse primer binding site is absent, as the *cpsC* gene is absent in these serotypes. However, subsequent *in silico* sequence analysis of these three serotypes revealed multiple weaker reverse binding sites with at least 10 nucleotide mismatches that are present approximately 2,000 bp downstream of the forward primer, situated within the glycosyltransferase *wcyA* gene (data not shown).

Species specificity of sequetyping. Other bacterial species including a range of streptococci and common mucosal colonizers were selected to test the specificity of the selected pair of primers. An amplicon of the expected size of 1,061 bp was detected only in *S. pneumoniae* ATCC 49619 (serotype 19F). Amplification of rRNA genes as a positive control in nonpneumococcal strains confirmed the presence of a DNA template in these samples (Table 1).

Experimental evaluation of the sequetyping method based on the *cpsB* gene. The *cpsB* sequences of 138 pneumococcal strains covering 48 serotypes were tested *in vitro* (Table 2). These strains originate from Tanzania (42) and the United Kingdom, have different genetic backgrounds, and were isolated from different anatomical sample sites (Table 2). These strains included all of the PPV23 vaccine serotypes (VTs) except for serotype 2. In addition, 25 nonvaccine serotypes (NVTs) were tested. Altogether, the study evaluated 46 serotypes that were predicted by *in silico* analysis to generate an amplicon and two serotypes, 27 and 29, that were not predicted to yield an amplicon.

In silico analysis based on the entire *cpsB* gene predicted that 28 of the serotypes tested would be sequetyped, and 20 serotypes tested in this study were predicted unlikely to be differentiated (see Table S1 in the supplemental material). Of the 138 strains tested, 91 (65.9%) were sequetyped to the correct serotype. An additional 27 (19.6%) strains were sequetyped to the serogroup level, and 12 (14%) gave ambiguous results. One of 7 strains of serotype 19F (11.6402.H) was identified as belonging to serotype 1, but at lower identity (98%) than found with the true serotype 1 strains tested. It is worth noting that the correct 19F identity was a very close second best at 97% identity (Table 2). At this relatively low percent identity, we would express caution, designate the results ambiguous, and confirm them by alternative methods. In addition a serotype 33C strain (SG33c) and a serotype 36 strain (SG36) were misidentified (Table 2).

Of the 23 VTs tested, all strains of 15 serotypes (1, 3, 4, 5, 6A, 6B, 8, 9N, 9V, 10A, 14, 15B, 19A, 19F, and 23F) were correctly sequetyped (Table 2), except for 2 of 11 strains of 6A, which were sequetyped to the serogroup level, and a strain of serotype 19F, which was ambiguous as described above (Table 2). An additional four VTs were sequetyped to the correct serogroup: strains of serotypes 7F, 12F, 18C, and 22F were sequetyped as 7F/7A, 12B, 18B/18C, and 22F/22A, respectively. Ambiguous results were obtained for serotypes 11A, 17F, 20, and 33F. Except for the ambiguous serotype 19F strain, 11.6402.H, none of the VTs tested were incorrectly identified.

Of the 25 NVTs tested, 12 were correctly sequetyped and 8 were serogroup specific. Three NVTs gave ambiguous results, and the misidentified serotypes 33C and 36 described above are both NVTs.

From 2 to 11 strains of 23 serotypes were tested (1, 3, 4, 5, 6A, 6B, 7C, 7F, 8, 9N, 9V, 10A, 11A, 14, 17F, 18C, 19A, 19F, 21, 23B, 23F, 33F, and 35A), some with different multilocus sequence types (Table 2). All but serotypes 6A, 11A, 17F, 19F, and 35A gave correct sequetype results for the multiple strains (Table 2). Two strains of serotype 6A were sequetyped to the serogroup level, namely, 6A/6B (Table 2). Both strains of serotype 11A gave an ambiguous sequetype result of 11A/11D/18F (Table 2), and all but one of the serotype 17F strains gave an ambiguous sequetype result of 17F/33C (Table 2). The remaining serotype 17F strain also had the top sequetype score for 4/9V/17F/33C (Table 2). As described above, a serotype 19F strain was identified as serotype 1, but at a relatively low identity of 98% with 17-bp differences, which were all within the 5′ end of the gene. Of the three strains of 35A that were included in this study, two were from Tanzania and one was from the United Kingdom. The Tanzania strains (16NP10 and 55OP3) were both typed to the serogroup level (35B/35C), while the United Kingdom strain (SG35a) gave an ambiguous result (35A/33A/33F) (Table 2).

Comparison between predicted and experimental data. All the strains included in the study yielded amplicons, including those of serotypes 27 and 29, which were not predicted to be amplifiable (Table 2). The strains with these two serotypes were also sequetyped correctly.

Of the 28 serotypes in this study that were predicted to be sequetyped, 23 were correctly sequetyped. Strains with four of the remaining serotypes (7F, 24F, 28F, and 33D) were sequetyped to the correct serogroup, and the serotype 36 strain was identified as a 15B serotype (Table 2).

Twenty serotypes tested were not predicted to be sequetypeable. Remarkably, two serotype 9V strains were differentiated from 9A, a serotype predicted to have an identical *cpsB*. The remaining serotypes, as predicted, were serogroup specific, consistent, or misidentified. The single serotype 6C strain in the study was sequetyped to serotype 6C or 6D, but not to 6A or 6B.

DISCUSSION

Two well-conserved primer binding sites were located with a novel software script and were predicted to generate an amplicon in 84 of 92 serotype sequences interrogated. Furthermore, amplification of this region using a single primer pair on pneumococcal strains isolated from different geographical regions and anatomical sites has been shown to be successful in identifying a high

number of serotypes, which included those in the latest conjugate vaccine formulation, as well as the current replacement serotypes associated with invasive diseases (30).

Serotyping is conventionally based on the Quellung reaction with anticapsular sera, but this is expensive, labor-intensive, and prone to errors. Large numbers of pneumococcal cells are needed, and false reactivity may occur with different serotypes and other streptococcal species (13, 40, 66). Immunoblotting methods have been developed and have improved sensitivity and specificity but are not used routinely in most laboratories (9, 12). New methods include PCR with multiple primer sets targeting serotype-specific regions of *cps* (8, 11, 34, 35, 39, 50–52, 54), but serotype coverage is limited, which may not allow detection of replacement serotypes after conjugate vaccination (15, 56). Microarray methods for serotyping can provide coverage of all known serotypes (19, 60, 62, 65), but its high cost would limit its introduction in resource-poor regions with high burdens of disease.

A method based on serotype-specific glycosyltransferase genes was proposed (59), but only multiple strains of serogroup 6 and serotype 19F were tested, and this method is unlikely to be put into routine diagnostic use because of the requirement of serotype-specific primers. Existing methods targeting *cps* employing multiple restriction enzymes and/or a long PCR fragment, making amplification difficult and inconsistent, have been described (6, 38). A sequence-based typing system targeting the regulatory region of *cps* had been proposed (34, 35), but a low resolution was seen between cross-reactive serotypes, in addition to a lack of reproducibility in other serotypes. In those methodologies, differentiation of cross-reactive serotypes therefore required a high number of primers to amplify the serotype-specific *wzy* and *wzx* genes. This low resolution may have been due to the limited availability of published sequence data of only 11 serotypes at that time (31).

With the recent characterization of the *cps* locus of 92 serotypes (7, 10, 52) and *in silico* assessment of the capsulation region, we have identified a region unique in sequence in many serotypes that is flanked by conserved primer binding sites. The region is located at the 5′ end of the *cps* locus, a region previously known to be relatively conserved (7, 31, 44, 46, 48). It is notable that this region is present in the pneumococcus only, as no amplification is obtained with *S. mitis*, *S. oralis*, and *S. pseudopneumoniae* organisms, which have traditionally been difficult to differentiate (4, 32).

All but two of the 13 conjugate vaccine serotypes tested in our study were identified to the serotype level, while serotypes 7F and 18C were identified to the serogroup level. Serotypes such as 19A, 6A, and 15B/C are currently among the most prevalent replacement NVTs globally since the inception of PCV7 vaccination (21, 24, 28, 56). These replacement types were correctly identified. Most of the invasive pneumococcal disease (IPD) serotypes within the United Kingdom and carriage strains that include serotypes 6C, 19A, and 22F (33, 58) could be identified to at least the serogroup level. Reviewing previous epidemiological data demonstrated that our method should be able to correctly identify more than 76% of invasive pneumococcal diseases in the United Kingdom to at least the serogroup level and over 92% of penicillin-nonsusceptible clinical isolates (18, 22). Thus, we conclude that this method could be used to reduce the number of strains for which conventional serotyping methods are required and yield considerable cost savings.

Differences between the predicted *in silico* analysis data and the experimental sequetyping results were observed. Serotypes 27 and 29, which were not predicted to produce an amplicon, were in fact amplified and sequetyped *in vitro*. The values inputted into the algorithm for locating the appropriate PCR primer sites could be a possible explanation for this discrepancy. The potential primer-binding site mismatch tolerance was set to no more than two nucleotides, but additional primer mismatches may allow successful amplification. The threshold number of mismatches for successful amplification is not known and is likely to depend on the length and sequence of the primers. It is also possible that the locations of the mismatches within the primer also govern the success of amplification. Sequences for serotypes 11F, 39, and 43, which were predicted to be nonamplifiable, were amplified in practice. As for serotypes 25A, 25F, and 38, *cps* alignment suggested that although the *cpsB* gene is lacking in these strains, it is possible to amplify a larger fragment (7, 43, 64). It is possible that successful amplification of these serotypes may simply require alterations in the thermal cycling conditions without the need to employ additional primers.

Discrepancies between the predicted and experimental results, whereby some serotypes could be sequetyped despite predictions to the contrary, may be due to additional *cpsB* sequences deposited into GenBank during this study, which expanded the database available for comparison. As the *in silico* analysis was based on a representative strain of each serotype, strains of different genetic backgrounds may have variant alleles of *cpsB* (20, 31, 64). Indeed, this was true for serotype 9V, where sequetyping of our strains matched both serotypes 9A and 9V of the representative strains initially interrogated. However, additional 9V strain sequences deposited in GenBank during our study matched precisely our 9V sequence, resulting in correct identification (63). Conversely, some serotypes could not be sequetyped, contrary to the *in silico* predictions. Additional strains would be required to evaluate this method more fully.

Multiple strains of the same serotype and different STs collected over a period of 7 years could be identified correctly. All but two of the 11 serotype 6A strains tested were correctly typed, being sequetyped to either 6A or 6B. It has been shown that serotype 6A and 6B strains of unrelated genetic backgrounds can have identical *cpsB* sequences (20). A single serotype 19F strain out of 7 could not be typed correctly, showing relatively low identity with other serotype 19F sequences in the GenBank database. It is likely that this strain represents a divergent lineage of serotype 19F. It has been previously reported that intraserotype sequence variations at the *cps* regulatory region exist (64). Additional work is under way to compare differences in intraserotype variation.

The data presented are in part derived from the uncurated GenBank database. The implementation of a curated *cpsB* database may potentially increase the accuracy of this method over time, by incorporating sequence data from different laboratories.

This method is not able to replace serotyping, as some isolates cannot be identified and others can be identified only to the serogroup level. Reference serotyping will be required for some isolates until further refinements are produced. The majority of the serotypes included in the latest vaccines could be identified correctly, and more strains were identifiable at the serogroup level. This method could correctly identify greater than 85% of the strains in the study at least to the serogroup level, including identification of some of the most common serotypes seen in IPD and carriage, including types of PCV13 (21, 28, 30, 55). Therefore, the

proposed typing method could characterize commonly encountered serotypes, thus reducing cost and time compared to conventional serological and multiplex-PCR serotyping methods.

As the initial aim of this method is to provide a cost-effective alternative to conventional serotyping in resource-poor regions with minimal or no vaccine coverage, the identifiable serotypes are still likely to be common in pneumococcal disease and carriage in these regions (5, 41, 42, 61). Strains with ambiguous sequetype results and those without an amplicon can be assessed by conventional serological and other DNA-based methods, most likely targeting the *wzy* and *wzx* regions (35). While additional *in silico* analysis is required to address the misidentification of serotypes such as 33C and 36, these two serotypes are not very common, and this drawback weighs little against the effectiveness of this method in identifying prevalent serotypes. It is only with more experience of the techniques that we can understand the effect of recombination events.

This new sequetyping method has multiple advantages over other previously reported serotyping methods. It requires only a single PCR amplification. The reaction setup is straightforward, robust, and economical, requiring only crude genomic DNA from heat-lysed cells. This creates an opportunity for routine laboratories with PCR and sequencing facilities to serotype the majority of pneumococcal strains without the need for an expensive set of serological reagents. Excluding the cost of labor, sequetyping costs approximately US$6.60 per pneumococcus, which is more economical than the US$15 required for serotyping a pneumococcus by the Quellung reaction, as derived by Lalitha et al. (36). Sequetyping is likely to be a cost-effective alternative to currently available microarray technologies for serotyping. More-economical methods are present such as multiplex/multistep PCRs; however, multiple sequential reactions, in addition to the reduced serotype coverage, are likely to be time-consuming and costly when labor hours are taken into consideration. It is worth noting that the indicated sequetyping cost includes culturing on blood agar, and a culture-free amplification method would reduce sequetyping time and cost. We are currently assessing the feasibility of sequetyping clinical samples directly as well as sequetyping samples known to contain multiple serotypes.

In conclusion, our single-primer sequetyping method is simple to perform, robust, and economic. It is specific for the pneumococcus and can identify serotypes included in the latest conjugate vaccines and a high number of clinically relevant serotypes. It has the potential to identify new serotypes that may emerge in the postvaccine era and are nontypeable by standard methods. We are currently evaluating our sequetyping method on pneumococcal strains of other serotypes to fully assess the coverage of sequetyping. In addition, plans are under way to make the *in silico* primer design algorithm available to other laboratory scientists.

ACKNOWLEDGMENTS

We thank Androulla Efstratiou, Karen Broughton, Siobahn Martin, and the diagnostic laboratory personnel of the HPA (Colindale, United Kingdom) and the Scottish Streptococcal Reference Laboratory for the supply and serotyping of pneumococcal strains. We thank GR MICRO for the supply of reference pneumococcal strains. We are grateful to Derren Ready from the Eastman Dental Institute, Clare Ling from the HPA, and diagnostic laboratory personnel of the Royal Free Hospital for the nonpneumococcal strains.

We also thank the Royal Free Hospital Special Trustees for funding this project.

REFERENCES

1. **Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JA.** 2008. The descriptive epidemiology of *Streptococcus pneumoniae* and *Haemophilus influenzae* nasopharyngeal carriage in children and adults in Kilifi district, Kenya. Pediatr. Infect. Dis. J. **27**:59–64.
2. **Abeyta M, Hardy GG, Yother J.** 2003. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. Infect. Immun. **71**:218–225.
3. **Aguiar SI, Brito MJ, Goncalo-Marques J, Melo-Cristino J, Ramirez M.** 2010. Serotypes 1, 7F and 19A became the leading causes of pediatric invasive pneumococcal infections in Portugal after 7 years of heptavalent conjugate vaccine use. Vaccine **28**:5167–5173.
4. **Arbique JC, et al.** 2004. Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov. J. Clin. Microbiol. **42**:4686–4696.
5. **Arifeen SE, et al.** 2009. Invasive pneumococcal disease among children in rural Bangladesh: results from a population-based surveillance. Clin. Infect. Dis. **48**(Suppl 2):S103–S113.
6. **Batt SL, Charalambous BM, McHugh TD, Martin S, Gillespie SH.** 2005. Novel PCR-restriction fragment length polymorphism method for determining serotypes or serogroups of *Streptococcus pneumoniae* isolates. J. Clin. Microbiol. **43**:2656–2661.
7. **Bentley SD, et al.** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet. **2**:e31. doi:10.1371/journal.pgen.0020031.
8. **Billal DS, Hotomi M, Tasnim S, Fujihara K, Yamanaka N.** 2006. Evaluation of serotypes of *Streptococcus pneumoniae* isolated from otitis media patients by multiplex polymerase chain reaction. ORL J. Otorhinolaryngol. Relat. Spec. **68**:135–138.
9. **Bogaert D, et al.** 2004. Colony blot assay: a useful method to detect multiple pneumococcal serotypes within clinical specimens. FEMS Immun. Med. Microbiol. **41**:259–264.
10. **Bratcher PE, Kim KH, Kang JH, Hong JY, Nahm MH.** 2010. Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical and serological characterization. Microbiology **156**:555–560.
11. **Brito DA, Ramirez M, de Lencastre H.** 2003. Serotyping *Streptococcus pneumoniae* by multiplex PCR. J. Clin. Microbiol. **41**:2378–2384.
12. **Bronsdon MA, et al.** 2004. Immunoblot method to detect *Streptococcus pneumoniae* and identify multiple serotypes from nasopharyngeal secretions. J. Clin. Microbiol. **42**:1596–1600.
13. **Browne K, Miegel J, Stottmeier KD.** 1984. Detection of pneumococci in blood cultures by latex agglutination. J. Clin. Microbiol. **19**:649–650.
14. **Brueggemann AB, Pai R, Crook DW, Beall B.** 2007. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. PLoS Pathog. **3**:e168. doi:10.1371/journal.ppat.0030168.
15. **Brugger SD, Frey P, Aebi S, Hinds J, Muhlemann K.** 2010. Multiple colonization with *S. pneumoniae* before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine. PLoS One **5**:e11638. doi:10.1371/journal.pone.0011638.
16. **Calix JJ, Nahm MH.** 2010. A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene. J. Infect. Dis. **202**:29–38.
17. **Charalambous BM, Oriyo NM, Gillespie SH.** 2008. How valid is single-colony isolation for surveillance of *Streptococcus pneumoniae* carriage? J. Clin. Microbiol. **46**:2467–2468.
18. **Cooke B, et al.** 2010. Antibiotic resistance in invasive *Streptococcus pneumoniae* isolates identified in Scotland between 1999 and 2007. J. Med. Microbiol. **59**:1212–1218.
19. **Donkor ES, Bishop CJ, Antonio M, Wren B, Hanage WP.** 2011. High levels of recombination among *Streptococcus pneumoniae* isolates from the Gambia. mBio **2**:e00040–00011. doi:10.1128/mBio.00040-11.
20. **Elberse K, et al.** 2011. Sequence diversity within the capsular genes of *Streptococcus pneumoniae* serogroup 6 and 19. PLoS One **6**:e25018. doi:10.1371/journal.pone.0025018.
21. **Fenoll A, et al.** 2009. Temporal trends of invasive *Streptococcus pneumoniae* serotypes and antimicrobial resistance patterns in Spain from 1979 to 2007. J. Clin. Microbiol. **47**:1012–1020.
22. **Gladstone RA, Jefferies JM, Faust SN, Clarke SC.** 2011. Continued control of pneumococcal disease in the UK—the impact of vaccination. J. Med. Microbiol. **60**:1–8.
23. **Gray BM, Converse GM, III, Dillon HC, Jr.** 1980. Epidemiologic studies

of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. J. Infect. Dis. **142**:923–933.
24. **Hanage WP, et al.** 2010. Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete. Epidemics **2**:80–84.
25. **Hanquet G, et al.** 2010. Pneumococcal serotypes in children in 4 European countries. Emerg. Infect. Dis. **16**:1428–1439.
26. **Hill PC, et al.** 2008. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clin. Infect. Dis. **46**:807–814.
27. **Holmberg H, Danielsson D, Hardie J, Krook A, Whiley R.** 1985. Cross-reactions between alpha-streptococci and Omniserum, a polyvalent pneumococcal serum, demonstrated by direct immunofluorescence, immunoelectroosmophoresis, and latex agglutination. J. Clin. Microbiol. **21**:745–748.
28. **Huang SS, et al.** 2009. Continued impact of pneumococcal conjugate vaccine on carriage in young children. Pediatrics **124**:e1–e11. doi:10.1542/peds.2008–3099.
29. **Hyams C, Camberlein E, Cohen JM, Bax K, Brown JS.** 2010. The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect. Immun. **78**:704–715.
30. **Jefferies JM, Tee WS, Clarke SC.** 2011. Molecular analysis of *Streptococcus pneumoniae* clones causing invasive disease in children in Singapore. J. Med. Microbiol. **60**:750–755.
31. **Jiang SM, Wang L, Reeves PR.** 2001. Molecular characterization of *Streptococcus pneumoniae* type 4, 6B, 8, and 18C capsular polysaccharide gene clusters. Infect. Immun. **69**:1244–1255.
32. **Kawamura Y, Hou XG, Sultana F, Miura H, Ezaki T.** 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus.* Int. J. Syst. Bacteriol. **45**:406–408.
33. **Kaye P, et al.** 2009. Invasive pneumococcal disease (IPD) in England & Wales after 7-valent conjugate vaccine (PCV7); potential impact of 10 and 13-valent vaccines. Abstr. 27th Annu. Meet. Eur. Soc. Paediatric Infect. Dis., Brussels, Belgium.
34. **Kong F, Gilbert GL.** 2003. Using *cpsA-cpsB* sequence polymorphisms and serotype-/group-specific PCR to predict 51 *Streptococcus pneumoniae* capsular serotypes. J. Med. Microbiol. **52**:1047–1058.
35. **Kong F, et al.** 2005. A molecular-capsular-type prediction system for 90 *Streptococcus pneumoniae* serotypes using partial *cpsA-cpsB* sequencing and wzy- or wzx-specific PCR. J. Med. Microbiol. **54**:351–356.
36. **Lalitha MK, et al.** 1996. Serotyping of *Streptococcus pneumoniae* by agglutination assays: a cost-effective technique for developing countries. Bull. World Health Organ. **74**:387–390.
37. **Lalitha MK, Thomas K, Kumar RS, Steinhoff MC.** 1999. Serotyping of *Streptococcus pneumoniae* by coagglutination with 12 pooled antisera. J. Clin. Microbiol. **37**:263–265.
38. **Lawrence ER, et al.** 2000. Evaluation of serotype prediction by *cpsA-cpsB* gene polymorphism in *Streptococcus pneumoniae.* J. Clin. Microbiol. **38**:1319–1323.
39. **Lawrence ER, Griffiths DB, Martin SA, George RC, Hall LM.** 2003. Evaluation of semiautomated multiplex PCR assay for determination of *Streptococcus pneumoniae* serotypes and serogroups. J. Clin. Microbiol. **41**:601–607.
40. **Lee CJ, et al.** 1984. Capsular polysaccharides of nongroupable streptococci that cross-react with pneumococcal group 19. J. Immunol. **133**:2706–2711.
41. **Leimkugel J, et al.** 2005. An outbreak of serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are characteristic of *Neisseria meningitidis* meningitis epidemics. J. Infect. Dis. **192**:192–199.
42. **Leung MH, Oriyo NM, Gillespie SH, Charalambous BM.** 2011. The adaptive potential during nasopharyngeal colonisation of *Streptococcus pneumoniae.* Infect. Genet. Evol. **11**:1989–1995.
43. **Mavroidi A, et al.** 2007. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. J. Bacteriol. **189**:7841–7855.
44. **McEllistrem MC, Noller AC, Visweswaran S, Adams JM, Harrison LH.** 2004. Serotype 14 variants of the France 9V(-3) clone from Baltimore, Maryland, can be differentiated by the *cpsB* gene. J. Clin. Microbiol. **42**:250–256.
45. **Morona JK, Miller DC, Morona R, Paton JC.** 2004. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA*, *cpsB*, and *cpsD* have on virulence of *Streptococcus pneumoniae.* J. Infect. Dis. **189**:1905–1913.
46. **Morona JK, Morona R, Paton JC.** 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.
47. **Morona JK, Morona R, Paton JC.** 2006. Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. Proc. Natl. Acad. Sci. U. S. A. **103**:8505–8510.
48. **Morona JK, Morona R, Paton JC.** 1997. Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway. Mol. Microbiol. **23**:751–763.
49. **Nelson AL, et al.** 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect. Immun. **75**:83–90.
50. **O'Halloran DM, Cafferkey MT.** 2005. Multiplex PCR for identification of seven *Streptococcus pneumoniae* serotypes targeted by a 7-valent conjugate vaccine. J. Clin. Microbiol. **43**:3487–3490.
51. **Pai R, Gertz RE, Beall B.** 2006. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. J. Clin. Microbiol. **44**:124–131.
52. **Park IH, Park S, Hollingshead SK, Nahm MH.** 2007. Genetic basis for the new pneumococcal serotype, 6C. Infect. Immun. **75**:4482–4489.
53. **Pelton SI, et al.** 2007. Emergence of 19A as virulent and multidrug resistant pneumococcus in Massachusetts following universal immunization of infants with pneumococcal conjugate vaccine. Pediatr. Infect. Dis. J. **26**:468–472.
54. **Rubin LG, Rizvi A.** 2004. PCR-based assays for detection of *Streptococcus pneumoniae* serotypes 3, 14, 19F and 23F in respiratory specimens. J. Med. Microbiol. **53**:595–602.
55. **Sa-Leao R, et al.** 2009. Changes in pneumococcal serotypes and antibiotypes carried by vaccinated and unvaccinated day-care centre attendees in Portugal, a country with widespread use of the seven-valent pneumococcal conjugate vaccine. Clin. Microbiol. Infect. **15**:1002–1007.
56. **Singleton RJ, et al.** 2007. Invasive pneumococcal disease caused by nonvaccine serotypes among Alaska native children with high levels of 7-valent pneumococcal conjugate vaccine coverage. JAMA **297**:1784–1792.
57. **Sørensen UB.** 1993. Typing of pneumococci by using 12 pooled antisera. J. Clin. Microbiol. **31**:2097–2100.
58. **Tocheva AS, et al.** 2011. Declining serotype coverage of new pneumococcal conjugate vaccines relating to the carriage of *Streptococcus pneumoniae* in young children. Vaccine **29**:4400–4404.
59. **Tomita Y, Okamoto A, Yagi T, Hasegawa Y, Ohta M.** 2011. Capsular-type prediction by phylogenetic tree of glycosyltransferase gene polymorphism in *Streptococcus pneumoniae.* Open Access Bioinformatics **3**:67–73.
60. **Tomita Y, Okamoto A, Yamada K, Yagi T, Hasegawa Y, Ohta M.** 2011. A new microarray system to detect Streptococcus pneumoniae serotypes. J. Biomed. Biotechnol. **2011**:352736.
61. **Traore Y, et al.** 2009. Incidence, seasonality, age distribution, and mortality of pneumococcal meningitis in Burkina Faso and Togo. Clin. Infect. Dis. **48**(Suppl 2):S181–S189.
62. **Turner P, et al.** 2011. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. J. Clin. Microbiol. **49**:1784–1789.
63. **van Selm S, et al.** 2002. Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V. Microbiology **148**:1747–1755.
64. **Varvio SL, Auranen K, Arjas E, Makela PH.** 2009. Evolution of the capsular regulatory genes in *Streptococcus pneumoniae.* J. Infect. Dis. **200**:1144–1151.
65. **Wang Q, et al.** 2007. Development of a DNA microarray to identify the *Streptococcus pneumoniae* serotypes contained in the 23-valent pneumococcal polysaccharide vaccine and closely related serotypes. J. Microbiol. Methods **68**:128–136.
66. **Werno AM, Murdoch DR.** 2008. Medical microbiology: laboratory diagnosis of invasive pneumococcal disease. Clin. Infect. Dis. **46**:926–932.
67. **Woo PC, Cheung EY, Leung K, Yuen K.** 2001. Identification by 16S ribosomal RNA gene sequencing of an *Enterobacteriaceae* species with ambiguous biochemical profile from a renal transplant recipient. Diagn. Microbiol. Infect. Dis. **39**:85–93.
68. **World Health Organization.** 2007. Pneumococcal conjugate vaccine for childhood immunization—WHO position paper. Wkly. Epidemiol. Rec. **82**:93–104.