 10.1111/j.1469-0691.2008.02028.x

Identification of pneumococcal serotypes from culture-negative clinical specimens by novel real-time PCR

D. Tarragó[1], A. Fenoll[1], D. Sánchez-Tatay[2], L. A. Arroyo[2], C. Muñoz-Almagro[3], C. Esteva[3], W. P. Hausdorff[4], J. Casal[1] and I. Obando[2]

[1]Spanish Reference Laboratory for Pneumococci, Servicio de Bacteriología, Centro Nacional de Microbiología, Majadahonda, Madrid, [2]Fundación Reina Mercedes y Servicio de Infectología Pediátrica, Hospital Universitario Virgen del Rocío, Sevilla, [3]Servei de Microbiología, Hospital Sant Joan de Déu, Esplugues, Barcelona, Spain and [4]GlaxoSmithKline Biologicals, Rixensart, Belgium

ABSTRACT

Pneumococcal parapneumonic empyema is an increasingly common complication in children. Conventional microbiological cultures indicate bacterial causes in as few as 8% of cases; therefore, there is a vital need for new molecular methods of detection and diagnosis. The development and clinical evaluation of real-time PCR-based assays to detect the pneumococcal capsular *wzg* gene of all serotypes tested are reported here, and 24 of them have been identified in clinical specimens. Using real-time PCR assays with highly specific TaqMan MGB probes that target DNA sequences within the capsular polysaccharide gene cluster, it was possible to differentiate serotypes 1, 3, 5, 4, 6A, 6B, 7F/A, 8, 9V/A/N/L, 14, 15B/C, 18C/B, 19A, 19F/B/C, 23F and 23A. These assays showed high sensitivity (five to ten pneumococcal DNA equivalents) and they were validated with 175 clinical isolates of known serotypes. The clinical value of this approach was demonstrated by analysis of 88 culture-negative pleural fluids from children diagnosed with parapneumonic empyema in three Spanish hospitals. Pneumococcal DNA was detected in 87.5% of pleural fluids, and serotypes 1, 7F and 3 were responsible for 34.3%, 16.4% and 11.9%, respectively, of cases of parapneumonic empyema in children. Such molecular methods are critical for the diagnosis of invasive pneumococcal disease and continued epidemiological surveillance in order to monitor serotype vaccine effectiveness.

Keywords empyema, pneumococcal detection assay, pneumococcal serotype identification assay, real-time PCR, *Streptococcus pneumoniae*

Original Submission: 29 September 2007; **Revised Submission:** 27 March 2008; **Accepted:** 18 April 2008

Edited by P. Tassios

Clin Microbiol Infect 2008; **14:** 828–834

INTRODUCTION

Pneumococcal serogroup and serotype identification is currently performed by various methods involving large panels of expensive antisera, including the capsular swelling (Quellung) reaction (the traditional reference standard), latex agglutination and co-agglutination [1], and a dot blot assay [2]. However, each of these methods requires clinical isolates. As the yield of microbiological diagnostics can be as low as 8% for pleural fluids (PFs) [3] and <10% for blood [4], there is a basic need for typing methods that do not rely on clinical isolates. Recently, a conventional multiplex PCR approach for routine surveillance has been developed [5]. The advantages of real-time PCR over conventional PCR are its speed (elimination of the need for postprocessing steps that could contribute to contamination) and its wider dynamic range (allowing detection of much larger variations in concentrations of the target). The most important advantage is the lower limit of detection. The higher analytical sensitivity of real-time PCR TaqMan assays, as compared with conventional methods, would also facilitate simultaneous identification of more than

Corresponding author and reprint requests: David Tarragó, Servicio de Bacteriología, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Ctra. Majadahonda-Pozuelo km 2, 28220 Majadahonda, Madrid, Spain
E-mail: davtarrago@isciii.es

one serotype [6]. Real-time PCR-based methods for diagnosis and serotyping would represent a valuable tool to improve current methods of microbiological diagnosis and epidemiological surveillance of both pneumococcal infections and nasopharyngeal colonization [7].

One question with PCR-based detection methods is how to validate that a culture-negative, PCR-positive result, especially with *ply*, represents the actual presence of pneumococci. One way to enhance confidence in the initial PCR result would be to test samples using a second PCR that targets an unrelated gene. Therefore, in the present study, the *wgz* real-time PCR assay was used in a dual-target approach and run in parallel with the current pneumolysin PCR assay.

In this study, novel, sensitive and specific real-time PCR assays for the detection of all pneumococci and for the specific identification of serotypes, or serogroups, 1, 3, 4, 5, 6A, 6B, 7F/A, 8, 9V/A/N/L, 14, 15B/C, 18C/B, 19A, 19F/B/C, 23A, and 23F are described. The extensive strain collection of the Spanish Reference Laboratory of Pneumococci was exploited to thoroughly evaluate the specificity of these new assays. Moreover, these assays successfully detected and identified pneumococcal serotypes in a large number of culture-negative PFs collected from children with parapneumonic empyema (PE). This work underscores the importance of ongoing serotype surveillance studies using molecular methods.

MATERIALS AND METHODS

Clinical isolates and conventional serotyping

Pneumococcal clinical isolates of known serotype and non-pneumococcal isolates (*Streptococcus mitis*, *S. oralis* and *S. gordonii*; two strains each) from the collection of the Spanish Reference Laboratory of Pneumococci were used to assess the species and the serotype or serogroup specificity and sensitivity of the real-time PCR assays. One hundred and seventy-five clinical isolates representing 35 serotypes (1, 2, 3, 4, 5, 6A, 6B, 7F, 7A, 8, 9V, 9N, 9A, 10A, 10F, 11A, 12F, 13, 14, 15B, 15C, 16F, 17F, 18C, 19A, 19F, 20, 22F, 23F, 23A, 23B, 24A, 33F, 34, 37), five strains each, were serotyped by conventional methods. This serotype panel was chosen because it comprised 95.2% of serotypes of all invasive pneumococcal isolates received at the Spanish Reference Laboratory of Pneumococci during 2006. Isolates retrieved from storage by subculture on blood agar plates (Columbia II agar base supplemented with 5% horse blood) were incubated overnight at 37°C in 5% CO_2. Conventional serotyping was performed by the Quellung reaction using rabbit polyclonal antisera from the Statens Serum Institute, Copenhagen, Denmark as previously described [1]. Dot blot serotyping was performed as previously described [2].

Culture-negative PF specimens

This study retrospectively evaluated 88 culture-negative PF samples of paediatric patients who were diagnosed with PE during 2003–2006 in three Spanish hospitals (Hospital Universitario Virgen del Rocío, Sevilla; Hospital Carlos Haya, Málaga; and Hospital Sant Joan de Deu, Barcelona) These were derived from 118 culture-negative PF samples that had been collected and stored at −80°C at participating hospitals during the study period. Genotyping by multilocus sequence typing was initially performed in these clinical samples in Oxford, Seville and Barcelona, and proved to be a difficult assay that often required multiple and repeated testing, leaving only 88 specimens for further molecular testing. The median patient age was 47 months (range 11–180 months) and the male/female ratio was 1.09 : 1. The median duration of symptoms before thoracocentesis was 6 days (range 3–23 days). Eighty-two of 88 children (93.2%) had received antibiotic treatment prior to PF sampling (oral and/or parenteral). According to parental report, 29/88 (32.9%) of patients had received at least one dose of PCV7.

Selection of the capsular gene as a target for real-time PCR assay for the detection of pneumococci

All published DNA sequences of the capsular polysaccharide gene cluster comprising 90 pneumococcal serotypes [8] and related α-haemolytic streptococci (*S. mitis*, *S. oralis* and *S. gordonii*), located with BLAST (Basic Local Alignment Search Tool) [9] and downloaded from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/), were analyzed. A PCR primer pair and TaqMan minor groove binder (MGB) probe targeting the most highly conserved capsular gene in pneumococci, the *cpsA* or *wzg* gene [10], were selected for the detection of pneumococcal DNA by real-time PCR.

Selection of capsular genes as targets for real-time PCR typing assays and group design

Molecular typing was based on serotypes assigned by conventional serotyping. Sequences for each serotype and each capsular gene were searched with BLAST for all similar sequences in the GenBank database (http://www.ncbi.nlm.nih.gov). All retrieved sequences were re-aligned and analyzed by MEGALIGN (Lasergene® v7.2; DNASTAR Inc., Madison, WI, USA) and MACAW software (National Center for Biotechnology Information, National Library of Medicine, version 2.0.5) to reveal gene-specific differences in sequence conservation between serotypes. Consequently, all 90 serotypes were grouped for each capsular gene according to sequence similarity. Target genes were selected to distinguish polymorphisms among relevant serotypes PCV7 serotypes (4, 6B, 9V, 14, 18C, 19F and 23F), the 10-valent and 13-valent vaccines serotypes that add to PCV7 emerging or replacing serotypes in IPD [11] such as 1, 3, 5, 6A, 7F and 19A (Table 1).

Primers for the PCR typing assays allowed serotype grouping according to sequence similarity and relevance of serotype; for example, a duplex real-time PCR assay for serotypes 1 and 5 was designed on the basis of sequence similarity of the *ugd* gene and the observation that both serotypes are highly prevalent in children with PE [12]. A single real-time PCR was designed for serotype 3, because genes within its capsular locus are either highly conserved (*wzg*) or specific (*cap3B*; *wchE*; *cps3S*). Using real-time PCR for

typing serotypes 6A and 6B, it was possible to detect a single non-synonymous polymorphism at codon 195 within *wciP*, the rhamnosyl transferase gene, which has previously been shown to be responsible for the expression of serotype 6A or 6B [13]. Both this non-synonymous polymorphism and an additional one are included in the specific 6A and 6B TaqMan MGB probes. Furthermore, where possible, the more prevalent serotypes in some specimens of IPD were combined in a single real-time PCR assay. Some serotypes were indistinguishable from other closely related serotypes, most of which belonged to the same serogroup (Table 1). Groups of serotypes analyzed by real-time PCR assays and selected genes of each group are shown in Table 1.

Oligonucleotide primers and TaqMan MGB probe designs

In the preliminary analysis of sequences, very minor variability within different strains of the same serotype (a few nucleotides throughout the whole gene sequence) were found. Specifically designed primers and probes were therefore created to avoid those variable positions in sequences. Consensus or degenerate oligonucleotide primers and probes were designed for each real-time PCR group using PRIMER EXPRESS 3.0 (Applied Biosystems, Applera Hispania SA, Tres Cantos, Madrid, Spain). Serotype-specific TaqMan MGB probes were labelled with FAM or VIC fluorochromes. Primers were synthesized by Sygma-Genosys (Cambridge, UK) and probes by Applied Biosystems (Applera Hispania SA). The oligonucleotide primers and probes used in this study, their target genes, specificity, sequences, and the numbered base positions of probes are shown in Table 1.

DNA extraction from clinical specimens and clinical isolates

DNA was extracted from 200 μL of PF using the QIAamp DNA mini kit (Qiagen, Izasa, Barcelona, Spain) according to the protocol provided by the manufacturer. This protocol minimizes the presence of PCR inhibitors in the final sample. DNA was isolated from serotyped clinical isolates by suspension of *S. pneumoniae* strains at 0.1 MacFarland unit in MilliQ water, heated to 90°C for 30 min, and subsequently cooled on ice and centrifuged for 5 min at 5000 *g*. to pellet non-disrupted bacteria and cell debris. DNA samples were stored at −20°C. The concentrations of the DNA extracted from the bacterial cultures were determined by the Nanodrop method (Nanodrop Technologies, Wilmington, Delaware, USA). A 2.5-μL aliquot of supernatant containing free DNA was used as the template for PCR.

Pneumolysin gene PCR assay for the detection of pneumococci

Detection of pneumococci in PF was performed by a previously described method based on PCR with *ply* [4].

Real-time PCR assay for the detection of pneumococci and real-time PCR assays for serotype determination

Standard real-time PCR conditions recommended by the manufacturer were used in all experiments, with minor modifications. Real-time PCR assays were carried out in a final 25-μL reaction volume in optical tubes or on optical 96-well plates (Applied Biosystems), depending on the

Table 1. Oligonucleotide primers and Taqman minor groove binder (MGB) probes for each group of serotypes

Serotype/serogroup specificity of real-time PCR[a]	Target gene	Forward/reverse primer sequences (5′–3′)	TaqMan MGB probe[b] (nucleotide position)[c]
All	*wzg*(*cpsA*)	GCTCCTAAGACGCTCTAAGAATCAGTCT/CGACACCGAACTAATAGGACCAT	Pneumo FAM-TCTATGTTAGTGGAATTGAC (3743–3762)
1 and 5	*ugd*(*cpsK*)	TKYWGTSGARGCWGTYATTGA/TCCCGTAAAAAYTCAGGACTAAASA	St1 FAM-CAGTTATGAAATATAATCCTGATGCC (13 837–13 861)
			St5 VIC-ATTGCGGTTAATTTGAAG (13 023–13 040)
3	*wchE*(*cap3*; *cpsS*)	GGACGCCTAGAACCTTGAGTGA/TTGTCGTTACCCCACCTATTTTTT	St3 FAM-TTGCTGAAGCCTTTTGT (7426–7442)
4[d]	*wcil* (*cpsF*; *wchF*)	CGGCACGCAAACCAATTAT/CATCTCGTTCCGGACTAACA	St4 FAM-CAGGAGATGCTAAAATA (7118–7134)
6A and 6B	*wciP*(*cpsS*)	GCTAGAGATGGTTCCTTCAGTTGAT/CATACTCTAGTGCAAACTTTGCAAAAT	St6A FAM-CTGGCTCATGATAGTT (8872–8887)
			St6B VIC-ACTGTCTCATGATAATT (8769–8785)
7F/A	*wcwA*	AAGCACAGTGCGTGAACAAT/AAAATCTCCCTGTCCCTTCC	St7F/A FAM-CTATTCCAGAAGAATCTC (8327/8298–8344/8315)
8 and 9V/A/N/L	*ugd* (*cpsK*)	ACTTACCTWGCKTTGCGTGTTTCTTA/CRTASCCRAAGGAWGGATTGTT	St8 FAM-CACGAATCGGCTCTC (12 710–12 724)
			St9V/A/N/L VIC-CCACGTATAGGAAATCAT (17 406–17 423/17 089–17 106/15 134–15 151/15 134–15 151)
14 and 15B/C	*wzy*	TGTAYGARGAATCYCTWAAAGCT/TGAYWCCTGCKCCAAGT	St14 FAM-AATGGAATGGAAATGTTAC (7950–7968)
			St15B/C VIC-TAATGGGCTGGGAATT (7953–7968)
19A and 19F/B/C	*wchO*	AATKCKGTRTTTATGGGRGTTGG/AGAGACGTTTAGGCTCATTWGC	St19A FAM-ATGCAAAATGCTCACCTAG (7424–7442)
			St19F/B/C VIC-ATGCAAAAGTCAAATTTAGA (9047–9066/6772–6791/8670–8689)
18C/B	*gct*	CCCTGAAACTAGTTGGGAACA/TTCCAATCATCACCCATTACA	St18C/B FAM-AAAGTCAGATGTTAAAGACTAC (15 849–15 870)
23F and 23A	*wchV*	CTGGCCAAGATATTTAAAAGAGAGT/AATTYCGCATCAGAGTATGCAA	St23F FAM-TTGCTCTTCGAAAAATGT (10 118–10 135)
			St23A VIC-TTGCTCCTCGAAAAGTGT (9554–9571)

[a]St/St means that it is not possible to distinguish between these serotypes; for example, 7F/A means that pneumococci may express either the 7F or 7A serotype.
[b]TaqMan MGB probes labelled with FAM or VIC fluorochromes.
[c]Nucleotide position refers to sequences from GenBank accession numbers[8]: CR926497, CR931637, CR931634, CR931635, CR931638, CR931639, CR931643/CR931640, CR931644, CR931648/CR931645/CR931647/CR931646, CR931662, CR931664/CR931665, CR931675, CR931678/CR931676/CR931677, CR931673/CR931672, CR931685, and CR931683, respectively.
[d]Strain 600/62 (serotype 4) and also Eddy no. 72 (serotype 45).

number of samples to be tested per run. All assays with each sample were performed in triplicate, using the Master Mix PCR without AmpErase UNG 2X (Applied Biosystems), according to the instructions of the manufacturer, with 2.5 μL of sample DNA. The primer and probe concentrations for each of the 11 assays were optimized, and, in accordance with the experimentally optimized concentrations, 900 nM primers and 200 nM TaqMan MGB probes were used for all subsequent experiments. A non-template control and the corresponding serotype-positive DNA control chosen from the serotype panel were included in every run. DNA was amplified with the Applied Biosystems 7300 Real-time PCR System (Applied Biosystems) by using the following cycling parameters: 95°C for 10 min, followed by 40 cycles at 95°C for 15 s and 60°C for 1 min (total time: 1 h 45 min). Amplification data were analyzed by instrument software (SDS, Applied Biosystems). Negative results were defined as those with cycle threshold (C_T) values >40. A sample was considered positive if at least two of three yielded a positive result within the <40-cycle cut-off. For assessments of the lower limits of detection, serial ten-fold dilutions (equivalent to from *c.* 5000 to five bacteria) of purified DNA from the corresponding serotype-positive DNA control, chosen from the serotype panel, were prepared, and aliquots were tested using all 11 real-time PCR protocols. Specificity determinations were made by testing at 5 mg/L the DNAs extracted from all *S. pneumoniae* isolates of the serotype panel and six non-pneumococcal isolates (*S. mitis*, *S. oralis* and *S. gordonii*, two strains each) by all 11 assays.

RESULTS

Development of a real-time PCR assay for the detection of the *S. pneumoniae wzg* gene and its validation with clinical isolates

A real-time PCR assay was developed for the detection of a highly specific sequence in the pneumococcal *wzg* gene. This assay allowed the detection of a minimum of 12–24 fg of DNA, corresponding to five to ten bacteria. Theoretical analytical specificity was ascertained using a BLAST [9] search of the probe and primer sequences, in which no exact match was found with any non-pneumococcal sequence. DNA from each of 175 pneumococcal strains representing 35 serotypes (see Materials and methods for a list) was detected using *wzg*-targeting real-time PCR with the same detection limit for each serotype. To test the analytical specificity of the specific pneumococcal TaqMan MGB probe, DNA from three related α-haemolytic streptococci carrying homologous, but not identical, *wzg* genes (two strains each of *S. mitis*, *S. oralis* and *S. gordonii*) was also tested. In addition, two culture-positive PFs that yielded *Staphylococcus aureus* and *S. pyogenes* isolates were tested. No reporter fluorescent signal was detected with DNA extracted from any of these non-pneumococcal strains, demonstrating high specificity of the probe sequence used.

Development and validation of real-time PCR assays for the determination of pneumococcal serotypes or serogroups

Ten groups of real-time PCR typing assays covering 24 serotypes were developed. In 2006, 80.9% of all invasive isolates received at the Spanish Reference Laboratory of Pneumococci belonged to one of these serotypes. The 35 different serotypes, represented by five strains each, were used to pretest the specificity of the real-time PCR typing assays. The strains of all the serotypes included in the assays were correctly typed by the technique described here, whereas the strains of serotypes not included were negative in all real-time PCR typing assays. Although we believe that the method is largely validated with the 175 strains, we further note that in ten cases of culture-positive PE tested, the serotype determined by the Quellung technique matched those detected with serotype real-time PCR assays directly, using the corresponding PF. All PCR products were analyzed by gel electrophoresis and yielded amplicons of the expected size, and the sequencing results perfectly matched the corresponding published sequences, confirming the specificity of the real-time PCR assays. All ten assays allowed the detection of a minimum of 12–24 fg of DNA, corresponding to five to ten bacteria.

Study of culture-negative PFs from children with PE

The 88 PF specimens were analyzed using both *ply*-targeting and *wzg*-targeting PCR assays. Among 88 PFs, 74 were positive using *ply* PCR and 72 were positive using *wzg* real-time PCR. There were eight discrepancies between the two assay methods: five samples were *ply*-positive but *wzg*-negative, and three were *ply*-negative but *wzg*-positive. Therefore, putative pneumococcal DNA (either *ply*-positive or *wzg*-positive, or both) was detected in 77/88 (87.5%) of culture-negative samples. However, in two of the five *ply*-positive/*wzg*-negative PFs, unrelated pathogens (*S. pyogenes* and *Stapylococcus aureus*) grew, suggesting

false-positive PCR results. Thus, if 75 PFs represent the number of true pneumococcal samples, then the sensitivity of either *ply* PCR or *wzg* PCR would be 72/75 (96%). Overall, among these 75 PFs, sufficient material was available for serotype real-time PCR with 67 samples, and the serotype was successfully identified in 52 samples. Thus, the maximum specificity of *wzg* PCR could be as high as 100%, as there is no evidence that any of the 72 *wzg*-positive samples were not true pneumococci and, in fact, two of three *ply*-negative/*wzg*-positive samples could be serotyped, indicating that they are true pneumococci. Alternatively, if one were to assume that either being *wzg*-positive/*ply*-positive or being a serotypeable sample is sufficient to define a true pneumococcus, then the minimum specificity would be 71/72 (98.6%) ($wzg^+/ply^+ = 69$ plus the proportion of serotypeable samples among the wzg^+/ply^- samples= 2). Results based on the 88 PFs are shown in Fig. 1. Serotype distributions are shown in Table 2.

Table 2. Molecular study of 67 culture-negative pleural fluids: results of serotype real-time PCR

Serotype	Serotype real-time PCR; no. of positive pleural fluids (%)
1	23 (34.3)
3	8 (11.9)
5	2 (2.9)
7F/A	11 (16.4)
8	1 (1.4)
14	3 (4.4)
19A	3 (4.4)
19F/B/C	1 (1.4)
Unknown[a]	15 (22.3)
Total	67 (100)

[a]Unknown indicates that the serotype could not be identified by real-time PCR as one of the 24 serotypes.

DISCUSSION

Real-time PCR offers an opportunity to re-address the problem of the diagnosis of infections due to *S. pneumoniae* and to determine the responsible serotypes. The assays described in this article proved to be highly sensitive and specific in detecting and identifying pneumococcal serotype-specific DNA. The serotype real-time PCR method produced unambiguous results that perfectly matched Quellung reaction and dot blot test results. Although DNA from only 35 of the 90 known serotypes was tested, the serotype specificity of each group of real-time PCR was supported by comparison of the selected gene sequence with the CPS gene cluster sequences for all 90 known serotypes available in the GenBank database. Similarly, the real-time PCR detection assay gave positive results with the panel of 35 serotypes and negative results with strains of *S. mitis*, *S. gordonii* and *S. oralis*. As oligonucleotide primers and probes were designed specifically for *S. pneumoniae* DNA sequences, and did not match any sequence of *Staphylococcus aureus* or *S. pyogenes*, extensive specificity tests were not included in the assays for these pathogens. However, because those are relevant pathogens in other

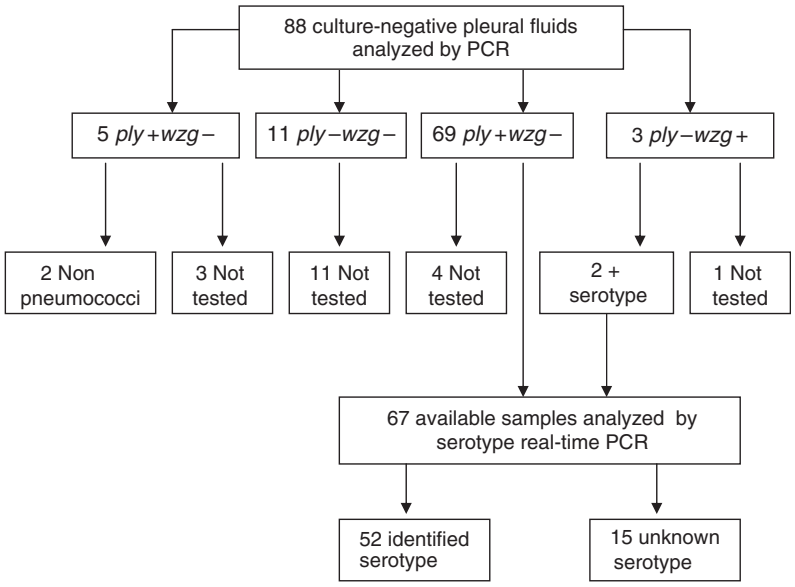

Fig. 1. Diagrammatic representation of the molecular analysis of 88 parapneumonic empyema cases without microbiological pneumococcal isolates. Eighty-eight samples were tested for pneumococcal DNA by *ply* PCR and *wzg* real-time PCR. Twenty-one pleural fluid samples were not tested by serotype real-time PCR because they were negative according to both *ply* PCR and *wzg* real-time PCR, they did not contain pneumococci, or there was an insufficient amount of sample. Subsequently, 67 samples were analyzed by serotype real-time PCR, 52 of which were identified to the serotype level.

areas and *Staphylococcus aureus* has recently emerged as the leading cause of PE in certain areas of the USA [14], it is worth noting that two cases of PE in which the PFs yielded *Staphylococcus aureus* and *S. pyogenes*, respectively, both gave negative results using *wzg* real-time PCR. Although there are many reports of false-positive *ply*-targeted PCRs [15,16], detection of both *wzg* DNA and *ply* DNA sequences, combined with real-time PCR serotype identification, makes it likely that these are indeed true-positive results. Furthermore, it seems that the results of these combined assays could be as reliable as the combination of the recently described real-time PCR assays targeting the *lytA*, *ply* and *psaA* genes for identification of *S. pneumoniae* as the aetiological agent of infections [17].

Carvalho *et al.* [17] reported that *ply*-targeted PCRs showed positive results when pneumococcus-like viridans group streptococci were tested. In this regard, an additional target on the capsular *wzg* gene, exploited by the real-time detection assay described here, may contribute to clarifying the molecular identification of *S. pneumoniae*, particularly when identification is attempted with non-sterile specimens such as sputum. However, it should be noted that because *ply* PCR was performed in Seville and Barcelona, the *wzg* real-time PCR and serotype real-time PCR assays were performed retrospectively, using limited volumes of DNA extracts shipped to the reference laboratory in Madrid. Unfortunately, there was not always sufficient sample volume to repeat the test. Moreover, as *ply* PCR and *wzg* real-time PCR were always performed sequentially, the samples to be tested using *wzg* real-time PCR have been frozen and thawed more frequently than those tested with *ply* PCR. These procedures compromise the quantity and quality of DNA and could account for some of the few discrepancies among the methods observed in this study. Therefore, the possibility that the three *ply*-positive and *wzg*-negative PCRs could also have arisen from pneumococcus-positive pleural fluids cannot be conclusively ruled out, and further studies are needed for clarification.

As the *wzg*-targeted PCR assay was used to detect pneumococci in clinical specimens with potential inhibitors of PCR reactions, it would be desirable to include an internal control to detect false negatives. Toward this end, we are working on an exogenous active reference by constructing an internal control plasmid that will be used to distinguish true target negatives from PCR inhibition.

Routine clinical practices such as administration of antibiotics prior to thoracocentesis could conceivably lead to an overestimation of the role of resistant serotypes in PE, and an underestimation of the importance of susceptible serotypes such as 1, 3, 5 and 7F. Serotypes 1, 7F and 3 were detected in 34.3%, 16.4% and 11.9%, respectively, of the *ply*-positive or *wzg*-positive, culture-negative PE cases, which suggests that these values more accurately reflect the actual serotype distribution in cases of empyema from our area than those obtained by conventional non-molecular methods. In this regard, 47 of the 52 (90.3%) culture-negative PF samples for which serotype results were obtained by real-time PCR came from children who had previously undergone intravenous antibiotic therapy for a median duration of 3 days (range 1–10 days). This demonstrates the ability of this method to detect and identify the serotype despite prolonged antibiotic therapy.

The real-time PCR technique described here would permit analysis of a large number of clinical specimens in a single day, the scale being limited by the equipment available. Although the usefulness of the assays has been demonstrated with culture-negative pleural fluids, they may be equally useful for the analysis of culture-negative cerebrospinal fluid and blood samples. The simplicity and flexibility of this approach mean that it is likely to be reproducible in other laboratories, and it should be possible to add more serotypes by designing new real-time PCR assays using the same strategy. The application of this technique has improved the diagnostic and reference activities of our laboratory, and has extended its diagnostic capability to culture-negative specimens. We believe that this technique has the potential to be a robust alternative to serological serotyping and will be of value to reference centres and researchers investigating serotype replacement and the population biology of *S. pneumoniae*.

ACKNOWLEDGEMENTS

We thank M. D. Vicioso, I. Hernandez and O. Robledo for their technical assistance.

TRANSPARENCY DECLARATION

This work was supported by Ministerio de Sanidad y Consumo, Instituto de Salud Carlos III, Spanish Network for the Research in Infectious Diseases (REIPI RD06/0008) and grants CP05/00068 and PI05/0924 from Fondo de Investigaciones Sanitarias. W. P. Hausdorff is employed by GlaxoSmithKline, a developer of pneumococcal conjugate vaccines. The other authors declare no conflict of interest.

REFERENCES

1. Henrichsen J. Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* 1995; **33**: 2759–2762.
2. Fenoll A, Jado I, Vicioso D, Casal J. Dot blot assay for the serotyping of pneumococci. *J Clin Microbiol* 1997; **35**: 764–766.
3. Thomson AH, Hull J, Kumar MR, Wallis C, Balfour Lynn IM. Randomised trial of intrapleural urokinase in the treatment of childhood empyema. *Thorax* 2002; **57**: 343–347.
4. Salo P, Ortqvist A, Leinonen M. Diagnosis of bacteremic pneumococcal pneumonia by amplification of pneumolysin gene fragment in serum. *J Infect Dis* 1995; **171**: 479–482.
5. Pai R, Gertz RE, Beall B. Sequential multiplex pcr approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. *J Clin Microbiol* 2006; **44**: 124–131.
6. Huebner RE, Dagan R, Porath N, Wasas AD, Klugman KP. Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. *Pediatr Infect Dis J* 2000; **19**: 1017–1020.
7. O'Brien KL, Nohynek H. Report from a WHO working group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. *Pediatr Infect Dis J* 2003; **22**: 133–140.
8. Bentley SD, Aanensen DM, Mavroidi A *et al*. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* 2006; **2**: e31.
9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403–410.
10. Garcia E, Llull D, Munoz R *et al*. Current trends in capsular polysaccharide biosynthesis of *Streptococcus pneumoniae* genetic bases and medical relevance of capsular polysaccharide biosynthesis in pathogenic streptococci. *Res Microbiol* 2000; **151**: 429–435.
11. Byington CL, Korgenski K, Daly J, Ampofo K, Pavia A, Mason EO. Impact of the pneumococcal conjugate vaccine on pneumococcal parapneumonic empyema. *Pediatr Infect Dis J* 2006; **25**: 250–254.
12. Obando I, Arroyo LA, Sanchez-Tatay D *et al*. Molecular epidemiology of paediatric invasive pneumococcal disease in southern Spain after the introduction of heptavalent pneumococcal conjugate vaccine. *Clin Microbiol Infect* 2007; **13**: 347–348.
13. Mavroidi A, Godoy D, Aanensen DM, Robinson DA, Hollingshead SK, Spratt BG. Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J Bacteriol* 2004; **186**: 8181–8192.
14. Schultz KD, Fan LL, Pinsky J *et al*. The changing face of pleural empyemas in children: epidemiology and management. *Pediatrics* 2004; **113**: 1735–1740.
15. Suzuki N, Seki M, Nakano Y, Kiyoura Y, Maeno M, Yamashita Y. Discrimination of *Streptococcus pneumoniae* from viridans group streptococci by genomic subtractive hybridization. *J Clin Microbiol* 2005; **43**: 4528–4534.
16. Whatmore AM, Efstratiou A, Pickerill AP *et al*. Genetic relationships between clinical isolates of *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*: characterization of 'atypical' pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes. *Infect Immun* 2000; **68**: 1374–1382.
17. Carvalho Mda G, Tondella ML, McCaustland K *et al*. Evaluation and improvement of real-time PCR assays targeting lyta, ply, and psaa genes for detection of pneumococcal DNA. *J Clin Microbiol* 2007; **45**: 2460–2466.