Genotypic Analysis of Invasive Streptococcus pneumoniae from Mali, Africa, by Semiautomated Repetitive-Element PCR and Pulsed-Field Gel Electrophoresis

S. M. Harrington, 1* F. Stock, 1 A. L. Kominski, 1 J. D. Campbell, 2 J. C. Hormazabal, 3 S. Livio, 2 L. Rao, 1† K. L. Kotloff, 2 S. O. Sow, 4 and P. R. Murray 1

Department of Laboratory Medicine, National Institutes of Health, Bethesda, Maryland;* Center for Vaccine Development, University of Maryland School of Medicine, Baltimore, Maryland; Microbiology Department, Instituto de Salud Publica, Santiago, Chile; and Center for Vaccine Development, Bamako, Mali

Received 8 September 2006/Returned for modification 26 October 2006/Accepted 13 December 2006

It is estimated that more than 1 million children die from pneumonia worldwide each year and that about 70% of these deaths occur in Africa and Southeast Asia (36). Streptococcus pneumoniae is the most common cause of bacterial pneumonia in children in developing countries (3, 14). Invasive pneumococcal disease (IPD) is also a major cause of morbidity and mortality, particularly in the developing world. In Mali, West Africa, children have limited access to health care and are not immunized against S. pneumoniae. Although the rates in Mali were not unlike those in other developing countries, the rates of IPD in hospitalized children in Mali were reported to be higher than those of industrialized nations (3). Furthermore, the case fatality rate due to IPD in hospitalized children in Mali was 23.6%. In the United States case fatality rates in this range were only observed in the oldest age groups in the prevaccine era (27).

Since the introduction of PCV7 in the United States in the year 2000, there has been a 75% decrease in IPD. The decline in disease due to serotypes contained in the vaccine is even more significant, with 94% fewer cases in 2003 compared to 1998 to 1999 (26). This vaccine is composed of the most common serotypes (4, 6B, 9V, 14, 18C, 19F, and 23F) known to cause IPD in the United States, Oceania, and Europe (15, 27). However, the serotype distribution in the developing world is often different. Of 106 isolates of S. pneumoniae causing IPD in Mali during a 1-year period, 78% were serotypes not included in PCV7 and 54% were serotype 5 (3). In the United States serotype 5 is uncommon; however, serotype 5 is the third, fourth, and sixth most common cause of IPD in Latin America, Asia, and Africa, respectively (15). A 5-year surveillance project conducted in six Latin American countries by the Sireva-Vigia group demonstrated that serotype 5 accounted for 9.6% of cases of IPD in the years 1994 to 1999 (9). As part of this initiative serotype 5 was found to be the second most common serotype among children less than 5 years of age in Colombia between 1994 and 1996 (32). In Africa the proportion of serotype 5 varies between 1.7% in South Africa and 14.3% in Kenya (15). A recent study conducted in 2002 and 2003 among young children in Mozambique placed serotype 1 as the most common serotype (40% of cases), with serotype 5 ranking second (10% of cases) (34). In Algeria, a country bordering Mali, serotype 5 was the most common serotype causing IPD among children, accounting for 12% of cases between the years 1996 and 2000 (25). Although these data indicate that infection with S. pneumoniae serotype 5 is com-
mon, they are far less than described for Mali in our previous report (3). Thus, serotype determination is necessary to guide vaccine implementation strategies in particular geographic regions.

As part of the assessment of S. pneumoniae serotypes, it is important to determine whether clonal dissemination of the predominant serotypes has occurred. Pulsed-field gel electrophoresis (PFGE) is generally considered to be the gold standard for bacterial strain typing; however, it has some disadvantages. PFGE takes several days to perform, is technically demanding, and requires specialized gel electrophoresis equipment and software for computer-assisted analysis. Without rigorous standardization, comparison of PFGE results over time and between laboratories is difficult. The multiplex sequence typing (MLST) method has been developed for strain typing of S. pneumoniae (11). The major advantage of this method is that it largely overcomes the problems with reproducibility between laboratories by comparing sequence data for a set of housekeeping genes. MLST has been used to compare isolates between labs around the world (29, 38). However, this method is also labor-intensive and requires sophisticated sequencing equipment. Repetitive-element PCR (rep-PCR) was described in 1993 as a rapid method for strain typing that relies on amplification of the sequences between various repetitive elements interspersed throughout the genome (37). Studies have utilized manual rep-PCR assays for typing of S. pneumoniae (5, 13, 35), but the method is now commercially available as a standardized, semiautomated technique (16). Kits containing standard primers and PCR master mix reagents are marketed by Bacterial Barcodes (bioMérieux, Inc., Athens, GA). Gel electrophoresis has been replaced by separation in a microfluidics chamber, and digitized gel images are generated and compared by using proprietary computer software. Although this method is limited by the requirement of significant expenditure for reagents and equipment (28), it has two major advantages. First, results are generated more rapidly than for either PFGE or MLST. A batch of 12 samples is easily performed within a single day. Second, standardized, digitized gel images are stored for comparison between runs and between laboratories. This method also can be used with a wide range of organisms, including methicillin-resistant Staphylococcus aureus (25, 27), vancomycin-resistant Enterococcus spp. (24), mycobacterial species (4), and fungi (17, 18). Because the rep-PCR microbial typing system could be applied to comparison of isolates between laboratories, it might have advantages over PFGE for comparison of S. pneumoniae for international studies where comparisons across borders are desirable.

There were two major goals of the present study. The first was to determine the extent of clonal dissemination of isolates in serotypes 2 and 5. Related to this aim was the question of whether the serotype 5 isolates from Mali were genotypically similar to those from Latin America. The second goal was to compare the results of strain typing of these isolates by PFGE and the Bacterial Barcodes rep-PCR method. The Mali cohort contains S. pneumoniae isolated within a single hospital over a relatively short time frame. Based on serotyping results, this set of organisms is likely to contain both genetically related and unrelated isolates and was therefore considered to be a group that would be appropriate for comparing the two methods.

(This study was presented in part at the 2006 National Meet-

ing of the American Society for Microbiology, Orlando, FL, abstr. C-141.)

MATERIALS AND METHODS

Bacterial isolates. One hundred S. pneumoniae isolates were available from all cases of IPD from pediatric patients less than 15 years of age hospitalized in Bamako, Mali, between 1 June 2002 and 1 July 2003. All isolates were identified in Mali and confirmed as S. pneumoniae at the University of Maryland Center for Vaccine Development. Serotyping was also performed previously by the Quellung reaction at the S. pneumoniae Reference Laboratory at the Instituto de Salud Publica de Chile, Santiago, Chile (3). S. pneumoniae available for study included 1 isolate that was serotype 1, 12 isolates that were serotype 2, 58 isolates that were serotype 5; 2 isolates were serotype 6A, 3 isolates that were serotype 6B, 7 isolates that were serotype 7F, 1 isolate that was serotype 9A, 4 isolates that were serotype 9V, 1 isolate that was serotype 12F, 1 isolate that was serotype 14, 6 isolates that were serotype 19F, and 4 isolates that were serotype 23F. CLB34 is a serotype 5 strain that is representative of Colombia5-19, a clone initially characterized in Colombia and later shown to be highly prevalent in a number of Latin American countries (12, 32). This isolate was kindly provided by Herminia de Lencastre (The Rockefeller University, New York, NY).

PFGE. Isolates were subcultured onto Trypticase soy sheep blood agar plates and incubated at 35°C in 5% CO2 for no more than 16 h to prevent autolysis. For extraction, growth from solid media was resuspended in 1 M NaCl, 10 mM Tris-Cl (pH 7.6) to achieve approximately the turbidity of a 1.0 McFarland turbidity standard. Extraction of genomic DNA in 0.8% InCert agarose plugs (Cambrex Corp., East Rutherford, NJ) was performed by standard methods and identified to Myco-Extract (Megazyme, Sigma, St. Louis, MO) and an extraction buffer of 1 mg/ml of mutanolysin and 10 mg/ml of mutanolysin/ml of bacterial lysis (21). DNA was digested with 2 U of the restriction endonuclease Smal at 30°C. PFGE was performed in a CHEF DR III (Bio-Rad Laboratories, Hercules, CA) apparatus with ramped pulse times from 5 to 35 sec for 24 h at 6 V/cm. S. aureus NCTC 8325 was included in at the first and last lanes of each gel as a molecular weight standard for gel-to-gel comparisons (33).

Similarity of DNA fingerprints was determined with BioNumerics software v2.5 (Applied Maths, Kortrijk, Belgium) using the Pearson’s coefficient (20). Dendrograms were generated by using the unweighted pair group method of average linkages. PFGE fingerprint profiles were interpreted according to the guidelines proposed by Tenover et al. (33) in which isolates that have no band differences are considered “indistinguishable” or the “same” strain or genotype. If one to three band differences were present, isolates were considered “related” to one another or subtypes of a common strain. When ≥4 band differences were present, isolates were considered “unique” strain types (33). Each unique PFGE strain type is designated by the letter R followed by a number. Subtypes within a given strain type are further designated with a lowercase letter. Numbers used for molecular strain type designations are independent of serotype numbers.

rep-PCR. Using the UltraClean microbial DNA isolation kit from Mo Bio Laboratories (Bacterial Barcodes), genomic DNA was extracted from several colonies of an overnight subculture on Trypticase soy sheep blood agar. PCRs utilized 25 ng of purified DNA as template and reagents in the Enterococcus DiversiLab DNA fingerprinting kit (Bacterial Barcodes). For the best discrimination of S. pneumoniae, this kit was recommended by the manufacturer. Cycling conditions for PCR included an initial denaturation of 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 90 s, with a final extension at 72°C for 5 min. PCR products were separated in a microfluidics DNA chip (Bacterial Barcodes) in the Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA) according to the protocol provided by Bacterial Barcodes.

rep-PCR fingerprint profiles were compared by using the Pearson correlation coefficient with DiversiLab v3.1 software (Bacterial Barcodes) that assesses both band position and band intensity. Interpretation of fingerprint patterns was developed for the present study based upon the best fit of the data to known serotype groupings and guidelines suggested by Bacterial Barcodes (16; Stacie Frye, unpublished data). “Indistinguishable” isolates or those with the same genotype had a similarity index (SI) of ≥97% and no obvious band differences. Isolates were characterized as being “similar” or subtypes if they had ≥97% similarity and one band different. “Different” strains had <97% similarity or two or more bands different. Each unique rep-PCR strain or genotype is designated by the letter “R,” followed by a number. Subtypes within a given strain type are further designated with a lowercase letter. Numbers used for molecular strain type designations are independent of serotype numbers.

Strain analysis. Evaluation of the S. pneumoniae isolates by PFGE and the Bacterial Barcodes rep-PCR method was performed in the National Institutes of Health clinical center microbiology laboratory. The genotyping results for PCV7 and PCV7-related serotypes are considered separately in the present study be-
cause the epidemiology associated with these serogroups in Mali appears to be different from non-PCV7 serotypes. Although not included in PCV7, serotypes 6A and 9A are in the same serogroups as vaccine serotypes 6B and 9V, respectively, and are evaluated as PCV7-related serotypes.

RESULTS

Non-PCV7 serotypes. The results with PFGE and rep-PCR were very similar for each of the five non-PCV7 serotypes. Both methods delineated only one genotype within each of these serotypes (Table 1). Each method further distinguished subtypes within the two major serotypes (serotypes 2 and 5). For serotype 2, five isolates were designated PFGE strain type P-2, with an additional three that were subtype P-2a and four that were subtype P-2b. Seven isolates were rep-PCR type R-2, two were subtype R-2a, and three were subtype R-2b (Table 1 and Fig. 1). These rep-PCR subtypes were ≥97% similar to R-2 but differed by minor bands in the high-molecular-weight region (right side of gel image). These bands were generally less reproducible and more difficult to interpret than the rest of the bands on the gel image. Notably, the methods did not place the same isolates in related subtype groups. For example, iso-

TABLE 1. PFGE and rep-PCR genotypes for five non-PCV7 serotypes

<table>
<thead>
<tr>
<th>Serotype</th>
<th>Total no. of strains</th>
<th>PFGE</th>
<th>rep-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of strains</td>
<td>Strain type</td>
<td>No. of strains</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>P-1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>5</td>
<td>P-2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td>P-2a</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>3</td>
<td>P-2b</td>
</tr>
<tr>
<td>5</td>
<td>58</td>
<td>28</td>
<td>P-3</td>
</tr>
<tr>
<td>6A</td>
<td>1</td>
<td>P-3a</td>
<td>1</td>
</tr>
<tr>
<td>7A</td>
<td></td>
<td>1</td>
<td>P-3b</td>
</tr>
<tr>
<td>8A</td>
<td></td>
<td>3</td>
<td>P-3c</td>
</tr>
<tr>
<td>9A</td>
<td></td>
<td>12</td>
<td>P-3d</td>
</tr>
<tr>
<td>10A</td>
<td></td>
<td>13</td>
<td>P-3e</td>
</tr>
<tr>
<td>7F</td>
<td>7</td>
<td>7</td>
<td>P-4</td>
</tr>
<tr>
<td>12F</td>
<td>1</td>
<td>1</td>
<td>P-5</td>
</tr>
</tbody>
</table>

a Each strain type P or R followed by a number represents a unique strain type for PFGE or rep-PCR, respectively. Subtypes within a strain type are designated by lowercase letters. Comparison of typing patterns can only be made within a method (or down a column).

FIG. 1. PFGE (A) and rep-PCR (B) similarity analysis and gel images for serotype 2 S. pneumoniae. Organism identification number (ID), PFGE genotype (PFGE), and rep-PCR genotype (PCR) results are shown to the right. Molecular weight standards for rep-PCR are given in base pairs.
lates in rep-PCR type R-2 could be PFGE type P-2, P-2a, or P-2b.

All 58 serotype 5 isolates were also related to each other by either method. Again, both PFGE and rep-PCR further discriminated a portion of the isolates as subtypes (Table 1). For PFGE the predominant strain was designated type P-3, and there were 28 isolates with this pattern. There were 13 S. pneumoniae in a group designated P-3e, 12 in P-3d, 3 in P-3c, and 1 in each of the two other groups. Figure 2 shows a PFGE gel image with isolates representative of these strain subtypes.

Of 58 serotype 5 isolates, 57 were available for rep-PCR analysis. By rep-PCR, 36 isolates were included in the group with the predominant pattern, R-3. Seven isolates were included in each of three related subtypes, designated R-3a, R-3b, and R-3c. Like serotype 2, the rep-PCR subtypes within serotype 5 differed from one another by minor bands in the high-molecular-weight region of the image (not shown).

PCV7 and PCV7-related serotypes. Of the 21 isolates from the seven serotypes related to or within PCV7, there were between 1 and 6 isolates per serotype. PFGE and rep-PCR distinguished the same isolates as unique genotypes within serotypes (Table 2 and Fig. 3). For serotype 9V, two isolates designated PFGE type P-10 differed by one rep-PCR band and were, therefore, designated as similar, or subtypes by rep-PCR (R-10 and R-10a). One serotype 9A isolate grouped by both PFGE and rep-PCR as a genotype present in serotype 9V. (See isolate 1371, PFGE type P-9 and rep-PCR type R-9.) Otherwise, genotyping was more discriminatory than serotyping, i.e., there were between one and three genotypes within each serotype.

Capsular type switching can occur for S. pneumoniae isolates when there is a genetic exchange that alters the capsule such that different antisera are reactive with the capsule, but the genomic content is largely maintained (10, 20). In the present study there were two isolates, 2389 and 2937, which share the genotype of strains within serotype 5. However, serotyping determined these to be types 6B and 23F, respectively (Fig. 3). One serotype 9A isolate grouped by both genotyping methods with two serotype 9V isolates. (See results for isolates 1371, 1976, and 1967 in Fig. 3.) It is unclear whether this is truly capsular type switching or whether the genotyping methods were unable to discriminate within these closely related serotypes.

Comparison to the Colombian clone. Previous reports have demonstrated the predominance of a serotype 5 clone in Colombia, South America, designated Colombia\textsuperscript{5-19} (32). CLB34, a chloramphenicol-susceptible, tetracycline-resistant isolate having the Colombia\textsuperscript{5-19} PFGE pattern was compared to the Mali serotype 5 isolates. As shown in Fig. 2, this isolate is indistinguishable from Mali PFGE type P-3d. Consistent with this, CLB34 was rep-PCR type R-3 (not shown). Not only has this strain type disseminated throughout Latin America, it was the predominating strain in Mali. In the original description of this clone, 38 of 39 isolates with indistinguishable Smal PFGE patterns were resistant to both chloramphenicol and tetracycline. Susceptibility to chloramphenicol was tested for 36 of 58 (62\%) of the serotype 5 isolates from Mali by the Kirby-Bauer method (data not shown). Only 2 of these 36 isolates were chloramphenicol resistant. However, 12 of 12 serotype 5 isolates (PFGE type P-3d) were resistant to tetracycline. The study by Gamboa et al. describing the dissemination of serotype 5 isolates in Latin America demonstrated that 78\% of serotype 5 isolates with the Colombia\textsuperscript{5-19} PFGE pattern were chloramphenicol resistant. However, those that are

![FIG. 2. Representative PFGE results for serotype 5 isolates. M, molecular weight standard S. aureus NCTC 8325. P-3 represents the predominant genotype. The letters above the figure designate PFGE subtypes within genotype P-3. Cl is the Colombia\textsuperscript{5-19} clone, isolate CLB 34.](http://jcm.asm.org/)
subtypes of this pattern were resistant <5% of the time (12). Thus, isolates in this genetic lineage may be chloramphenicol resistant but are not exclusively so.

**DISCUSSION**

One of the challenges with implementing a new strain typing system is determining the criteria for interpretation. For PFGE, visual examination of band differences (based on the criteria of Tenover et al.) with or without the addition of computer-assisted analysis has been applied to most evaluations (28, 33). However, in some studies the number of isolates is so large that a similarity index cutoff that corresponds to established criteria for band differences is applied (6, 22). Relatively few papers utilizing the Bacterial Barcodes rep-PCR method have been published, and no clear interpretive guidelines have been established. Suggested guidelines are available from the company, based on reproducibility studies performed with *E. coli*, *S. aureus*, and *E. faecium* and an analysis of well-characterized isolates of *N. meningitidis* (16). From these studies replicates of the same strain or isolates that were known to be epidemiologically related had a similarity index (SI) of 97 to 98%. Other evaluations have been done to de-
termine SI interpretive guidelines for different genera (4, 24, 28). When using this system for strain typing of *S. aureus*, one study suggests that using a lower SI (such as 85%) would generate groupings for this organism that agree better with PFGE data and known epidemiologic relationships (28). However, the study by Shutt et al. utilized an SI of >95% and one band difference to group similar strains of *S. aureus* (30). An SI of >95% with only one band difference was used for enterococci to distinguish related strains (24) and an SI of >95% grouped a set of well-characterized mycobacteria strains accurately (4). For this analysis of *S. pneumoniae*, isolates were considered the same strain type if the SI was at least 97% and there were no obvious band differences. Using these criteria and excluding strains for which capsular type switching occurred, there was only one instance in which rep-PCR grouped isolates (e.g., isolates 1371 and 1967) as the same genotype when they were of different serotypes. However, PFGE also grouped these isolates together. Isolates were characterized as similar if they had an SI of ≥97% and only one band different. Isolates that had <97% similarity to one another were of different serotypes and differed by two or more bands. An SI of 97% was chosen as the point separating similar and different genotypes because it distinguished between serotypes most accurately with the least requirement for objective scrutiny of band differences. Although the computer-generated similarity analysis component of the Bacterial Barcodes rep-PCR method makes it less subjective than PFGE, some isolates could be incorrectly classified unless some operator analysis of band differences is performed visually. Based upon our experience and the published literature, it appears that the SI cutoff delineating related strains may be different depending on the organism.

For the PCV7 and PCV7-related serotypes, PFGE and rep-PCR performed equivalently. This finding is consistent with at least one evaluation of *Staphylococcus aureus* typing and one in which VRE were characterized (24, 30); however, Ross et al. found rep-PCR to be less discriminatory than PFGE for MRSA (28). The Bacterial Barcodes system exhibited resolving power equal to or better than the established RFLP methods for mycobacteria. Furthermore, rep-PCR was able to characterize some mycobacterial species for which other strain typing methods have not been established (4). Thus, although the Bacterial Barcodes system appears to be as discriminating as other molecular typing systems, additional evaluations for individual genera and/or species are indicated.

With the rep-PCR method it was difficult to define relationships among subtypes within the larger clusters of serotypes 2 and 5. There were some minor band differences in the high-molecular-weight region, and these were more difficult to define consistently since the Bacterial Barcodes method interprets both band position and intensity. In addition, depending on the number of isolates in the analysis and the order in which they are considered by the analysis software, isolates could be grouped somewhat differently. This contributed to the difficulty in defining clear subgroups within large clusters. However, when computer analysis is used to interpret PFGE data for large numbers of isolates, small shifts in band positions can also cause isolates to group incorrectly. Unlike the Bacterial Barcodes software, the BioNumerics software used in the present study to analyze the PFGE data allows the user to set a percent position tolerance around each band position. Although this position tolerance may be chosen somewhat subjectively, it allows the user to make adjustments to the data to increase the correlation to expected results.

For serotypes 2 and 5, PFGE grouped more isolates as subtypes of the predominating strain types than did rep-PCR, but this difference between the methods was not significant. However, PFGE and rep-PCR generated different groups of subtypes within both serotype 2 and serotype 5. It is not surprising that the methods would separate different isolates into subtype groups. Although these methods potentially evaluate the entire chromosome, they measure different genetic components. It would be expected that minor genetic changes would be detected by the two methods differently. These changes likely reflect the beginnings of genetic drift but probably do not indicate more significant evolutionary changes that would cause isolates to be classified as distinct genotypes.

Given the large proportion of highly related serotype 5 isolates, it is most likely that the serotype 5 isolates represent an outbreak in which this particular clone circulated throughout Bamako during the study period. Another explanation for this high degree of strain relatedness may be that there are simply fewer genetic differences within serotype 5 isolates. Of 172 serotype 5 isolates studied from six countries in Latin America over a period of 6 years, all were either the PFGE Columbia-19 strain type or related to this strain (12). Studies have demonstrated that there is generally more clonal diversity among PCV7 serotypes compared to non-PCV7 serotypes. For example, serotypes 6B, 14, 19F, and 23F are more genetically diverse than serotypes 1, 5, or 7 (1, 31, 38). Our results are consistent with these findings. Although only 21 isolates were PCV7 or PCV7-related serotypes, for each PCV7-serotype group, more than one isolate there were two or three different strain types (Table 2). There were no multiple genotypes for any of the non-PCV7 serotypes (Table 1).

Contributing to our serotype distribution and genotyping results may be the fact that all isolates were from patients with invasive disease. A study comparing serotype and genotype conducted among children in Southern Israel demonstrated that serotypes 1 and 5 were much more frequently isolated from blood cultures than serotypes 6B and 23F. The latter serotypes were more prevalent among nasopharyngeal isolates from healthy, colonized children. Serotypes 1 and 5 were comprised of only 1 clone, whereas serotypes 6B and 23F contained 10 and 11 clones, respectively (23). A meta-analysis of studies among children from seven diverse populations suggests that some serotypes, such as serotypes 1 and 5, have a greater invasive disease potential and that isolates of serotypes 6, 19F, and 23F are among those more likely to colonize (2). Interestingly, most of the serotype 1 and 5 isolates from this analysis were from Kenya. Sjöström et al. present similar data from adults from western populations and suggest that isolates of serotypes that are more likely to colonize are often opportunistic pathogens, primarily causing disease in individuals with underlying illness (31).

One of the most interesting results of this analysis was the relationship between the genotype of the serotype 5 isolates from Mali and that of those from Latin America. The Colombia-19 clone was indistinguishable from one of the subtypes of the predominant serotype 5 PFGE clone in Mali. The Colombian clone was isolated as early as 1994 and was reported to be
circulating in the year 2000 (12). Serotype 5 *S. pneumoniae* isolated in Israel between 1995 and 1999 were also reported to have this same PFGE pattern (12, 23). Thus, this clone has disseminated to multiple continents and is a dominant strain causing IPD. This suggests that this clone is perhaps more fit for survival and dissemination or might be more virulent, causing more invasive disease. Given the fact that most countries in Africa and the Americas have reported proportions of up to 10% for serotype 5, it is likely that the proportion of 54% serotype 5 from 2002 to 2003 in Mali represents an outbreak of a dominant, virulent clone that is highly adapted for survival. As has been suggested by Gamboa et al., this clone may have been more recently established and, thus, there has been less time on an evolutionary scale for differentiation (12). It will be interesting to determine whether increased divergence of this clone occurs in future episodes of IPD caused by serotype 5 *S. pneumoniae*.

The data from Africa (3, 25, 34) and those from Latin America (9, 12) emphasize the need to develop vaccines that will be effective against the serotypes that are predominant in the target population. In addition to PCV7, pneumococcal conjugate vaccines have recently been formulated that contain 9 (7, 19), 11 (8), or 13 serotypes. These new-generation vaccines contain antigen to serotype 5 and are in various stages of clinical trials. Such vaccines might be particularly important in countries such as Mali and those from Latin America where serotype 5 is common.

Our results demonstrate that the Bacterial Barcodes rep-PCR method was able to delineate *S. pneumoniae* with a discriminatory power equal to that of PFGE. The rep-PCR method could be used as a stand-alone method for this species, the only qualification being that we found subtype groupings to be more difficult to define for rep-PCR. The serotype 5 isolates causing IPD in Mali were all clonally related by both methods. These data support the conclusion that an outbreak occurred during the study period. Furthermore, we showed that one of the subtypes of the outbreak strain is the same genotype as the highly prevalent Colombia5-19 strain. The Bacterial Barcodes rep-PCR method has the potential to increase the speed of analysis of strains from different countries since DNA fingerprints can be archived and compared without the exchange of organisms. However, this type of international comparison has not been undertaken at this time. Although the expense of the method is a drawback, if automated rep-PCR were utilized by international reference centers the dissemination of highly prevalent and perhaps virulent strains such as the Colombian clone would almost certainly be more rapidly defined.

ACKNOWLEDGMENTS

We thank Stacie Frye from Bacterial Barcodes, Inc., for helpful advice about the interpretation of the rep-PCR results and the preparation of the rep-PCR figures. We are grateful to William G. Merz for thoughtful review of the manuscript.

This study was funded in part by grants from the Bill and Melinda Gates Foundation, the Rockefeller Foundation and the Intramural Research Program at the NIH Clinical Center.

REFERENCES


