Highly Penicillin-Resistant Multidrug-Resistant Pneumococcus-Like Strains Colonizing Children in Oeiras, Portugal: Genomic Characteristics and Implications for Surveillance

Alexandra S. Simões, Raquel Sá-Leão, Marc J. Eleveld, Débora A. Tavares, João A. Carriço, Hester J. Bootsma and Peter W. M. Hermans


Updated information and services can be found at: http://jcm.asm.org/content/48/1/238

These include:

SUPPLEMENTAL MATERIAL
http://jcm.asm.org/content/suppl/2009/12/22/48.1.238.DC1.html

REFERENCES
This article cites 48 articles, 28 of which can be accessed free at: http://jcm.asm.org/content/48/1/238#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»
Highly Penicillin-Resistant Multidrug-Resistant Pneumococcus-Like Strains Colonizing Children in Oeiras, Portugal: Genomic Characteristics and Implications for Surveillance

Alexandra S. Simões,1‡ Raquel Sá-Leão,1,2‡ Marc J. Eleveld,3 Débora A. Tavares,1 João A. Carriço,4,5 Hester J. Bootsma,3 and Peter W. M. Hermans3

Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal1; Centro de Matemática e Aplicações Fundamentais, Universidade de Lisboa, Lisbon, Portugal2; Laboratory of Pediatric Infectious Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands3; Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina de Lisboa, Lisbon, Portugal4; and Instituto de Engenharia de Sistemas e Computadores—Investigação e Desenvolvimento, Lisbon, Portugal5

Received 6 July 2009/Returned for modification 19 October 2009/Accepted 5 November 2009

While performing surveillance studies in Oeiras, Portugal, designed to describe the impact of pneumococcal conjugate vaccine on colonization, we observed an increase from 0.7% in 2003 to 5% in 2006 in the prevalence of penicillin resistance (MIC of 2 to 6 mg/liter) among presumptively identified pneumococcal isolates. Although 15 of the 22 penicillin-resistant isolates obtained in 2006 were optochin resistant, they were bile soluble and thus considered to be bona fide pneumococci. This study aimed to clarify the nature of these isolates by using a combination of phenotypic and genotypic approaches that included routine strategies for pneumococcal identification, multilocus sequence analysis (MLSA), and comparative genomic hybridization (CGH). By MLSA, all isolates were classified as “streptococci of the mitis group” that, however, were distinct from typical Streptococcus pneumoniae or Streptococcus mitis. A single isolate was identified as Streptococcus pseudopneumoniae. CGH confirmed these findings and further indicated that a considerable part of the proposed pneumococcal core genome is conserved in these isolates, including several pneumococcal virulence genes (e.g., pavA, spxB, cbpE, and cbpD). These results suggest that among pneumococci and closely related streptococci, universal unique phenotypic and genetic properties that could aid species identification are virtually impossible to define. In pneumococcal colonization studies, when atypical strains are found, MLSA and CGH are informative tools that can be used to complement routine tests. In our study, after correct identification of the penicillin-resistant true pneumococci, we found that penicillin resistance levels among pneumococci remained stable from 2003 to 2006.

Streptococcus pneumoniae is a bacterial pathogen that frequently colonizes the nasopharynx of humans, particularly young children of preschool age. Colonization is mostly asymptomatic and only rarely results in disease (3). However, when disease does occur, it may range from a mild infection such as otitis media to severe sepsis or meningitis. Globally, the morbidity and mortality associated with pneumococcal infections is extremely high. A recent report from the WHO estimated that 0.7 to 1.0 million deaths occur annually among children <5 years of age as a result of pneumococcal infections (50).

Four phenotypic characteristics are classically used in the diagnostic laboratory for the presumptive identification of S. pneumoniae: colony morphology (colonies with a depression in the center showing alpha-hemolysis on sheep blood agar), optochin susceptibility, deoxycholate (DOC) solubility (commonly referred to as bile solubility), and a positive reaction with antipneumococcal polysaccharide capsule antibodies (20). In particular, optochin susceptibility and deoxycholate solubility have been associated with high sensitivity and specificity (between 98% and 100%). However, a number of studies have reported on sporadic optochin-resistant S. pneumoniae isolates (28, 34) and rare deoxycholate-insoluble strains (32).

On the other hand, some nonpneumococcal oral streptococci (such as Streptococcus mitis) may have a colony morphology similar to that of pneumococci but are classically optochin resistant, bile insoluble, and do not react with antipneumococcal polysaccharide capsule antibodies (25). Still, optochin-susceptible nonpneumococcal isolates have been described occasionally (20), as well as isolates with positive cross-reactions with antipneumococcal polysaccharide capsule antibodies (11). Recently, a new species—Streptococcus pseudopneumoniae—was described (1). S. pseudopneumoniae isolates were found to be resistant to optochin when incubated in an atmosphere enriched in CO2 but were optochin susceptible when incubated in ambient atmosphere. As a result, S. pneumoniae isolates may be difficult to distinguish from closely related species such as S. pseudopneumoniae and S. mitis.

Since the biochemical tests commonly used are not always sufficient to distinguish S. pneumoniae from other closely related upper respiratory streptococci, molecular approaches based on amplification of ubiquitous pneumococcal genes,
such as pneumolysin (ply) and autolysin (lytA), have been proposed (27). However, homologues of lytA and ply genes have been detected in strains of closely related streptococcal species (15, 29, 37, 49). Recently, detection of piaA (which encodes a lipoprotein component of two iron ABC transporters) has been proposed as a diagnostic tool for pneumococci as it was suggested to be specific for this species (47). Some authors also described 16S rRNA and sodA as good targets for identification of S. pneumoniae although sodA does not distinguish S. pneumoniae from S. pseudopneumoniae (1, 12, 21). The construction of phylogenetic trees from the concatenated sequences of the genes used for multilocus sequence typing (MLST)—an approach commonly termed multilocus sequence analysis (MLSA)—has also been proposed as a good alternative molecular technique to differentiate pneumococci from other closely related streptococci (9, 18).

In recent years, we have been studying atypical pneumococci recovered from colonization samples collected from children attending day care centers (DCCs) in Portugal. We first described a collection of over 200 nonserotypeable pneumococci which were mostly multidrug resistant and displayed low-level resistance to penicillin. All isolates were found to be true pneumococci that lacked a capsular operon. The isolates were genetically diverse although close to half belonged to a single lineage (39). Overall, these isolates were relatively abundant in asymptomatic carriers. The second group of atypical pneumococci that we described consisted of isolates resistant to optochin. Again, all strains were found to be true pneumococci and genetically diverse, and, in this case, most expressed a pneumococcal capsular type (30).

In the current study, we describe a third set of presumptively identified atypical pneumococcal strains. These isolates were all obtained in 2006 during a cross-sectional study designed to describe the impact of the seven-valent pneumococcal conjugate vaccine in colonization. Strikingly, 5% (22 of 441) of the presumptively identified pneumococcal isolates were found to have penicillin MICs ranging from 2 to 6 mg/liter, a value that between 2001 and 2003 never exceeded 1.7% and was 0.7% in 2003. Of note, 15 of these isolates, although resistant to optochin, were bile soluble and, thus, according to routinely accepted criteria, were considered to be bona fide pneumococci. Since such high MICs of penicillin are extremely rare among this population, we initiated a detailed characterization of these isolates consisting of classical strategies for pneumococcal identification, MLSA, and comparative genomic hybridization (CGH). Ten other optochin-resistant, bile-soluble isolates with penicillin MICs ranging from 0.064 to 0.75 mg/liter were also identified and further characterized.

We report here that these isolates are not true pneumococci but have phenotypic properties and genomic determinants that are frequently associated with S. pneumoniae, challenging their correct identification by several currently accepted assays.

MATERIALS AND METHODS

Strain collection. Strains used in this study were isolated during January and February of 2006 from the nasopharynx of preschool children attending DCCs in Oeiras, Portugal. Swabs were inoculated in tryptic soy agar (TSA) containing 5% defibrinated sheep blood supplemented with 5 mg/liter gentamicin and incubated overnight in anaerobic jars at 37°C to select and identify S. pneumoniae. As controls, we used S. pneumoniae strains R6 and TIGR4 and S. pseudopneumoniae type strain ATCC BAA-960 (kindly supplied by Maria de Gloria Carvalho, Division of Bacterial Diseases, Centers for Disease Control and Prevention, Atlanta, GA). In addition, pneumococcal strains D39, GS4, 23F, OXCl41, INCl04B, and INV200 and S. mitis NCTC 12261 were used in the comparative genomic hybridization experiments described below.

Antimicrobial susceptibility testing. MICs of penicillin, ceftriaxone, erythromycin, clindamycin, tetracycline, chloramphenicol, and sulfamethoxazole-trimethoprim were determined with the Etest (AB Biodisk, Solna, Sweden) according to the manufacturer’s instructions. Results were interpreted following the recommendations and definitions of the Clinical and Laboratory Standards Institute (CLSI) (5).

Detection of antimicrobial resistance genes. Macrolide (ermB, mefE, or mef(E)) and tetracycline (tet(M)) resistance genes were screened by PCR using primers and conditions previously described (26).

DNA fingerprinting by PFGE. Preparation of chromosomal DNA, digestion with Smal endonucleases, separation of DNA fragments by pulsed-field gel electrophoresis (PFGE), and interpretation of results were carried out as previously described (40).

Optochin susceptibility. Optochin susceptibility was tested by disk diffusion, using commercially available optochin disks (5 μg; 6 mm; Oxoid, Hampshire, England) applied to blood agar plates (Trypticase soy agar supplemented with 5% sheep blood) inoculated with colonies from overnight cultures. Plates were incubated overnight at 37°C in a 5% CO2 atmosphere. A similar assay was carried out in ambient atmosphere as described by Arbiute et al. (1). Isolates were considered to be resistant to optochin if they displayed inhibition zones smaller than 14 mm.

Capsular typing. Capsular serotyping was performed using the chessboard system (42) with specific antisera from the Statens Serum Institute (SSI; Copenhagen, Denmark). OmniSerum (SSI, Copenhagen, Denmark), a serum that contains antibodies to all known pneumococcal serotypes, was used to confirm nontypeability.

Detection of lytA (a conserved pneumococcal capsular gene present in 89 of the 91 capsular operons described to date) (2) was done by PCR using the primers cpsAF2 (5′-AGACGTTTGTGACGACTGACC-3′) and cpsAR2 (5′-GTGGGATCGGGAATCAAC-3′) with S. pneumoniae as a positive control. Both tubes were incubated at 37°C for up to 2 h. A sample was consid-
ed positive when the optical density of the cell suspension decreased more than 50% from the initial value, the assay was considered a technical control. Both tubes were incubated at 37°C for up to 2 h. A sample was considered positive when clearing of the turbidity occurred in the tube with the TIGR4 DNA with the primers L5A_Ext and L5A_Ext described above.

DOC solubility assays. DOC solubility was initially performed according to standard procedures (38): colonies from an overnight culture were suspended in 1 ml of 0.85% NaCl (wt/vol) solution to a turbidity equal to 0.5 to 1 McFarland standard. This suspension was distributed into two tubes, and three to four drops of a 10% DOC solution were added to one tube while the other served as a control. Both tubes were incubated at 37°C for up to 2 h. A sample was consid-
ered positive when clearing of the turbidity occurred in the tube with DOC but not in the control. For a more detailed characterization of the isolates, DOC solubility was reasayed in a slightly different way, as described by Llul et al. (24): 1 ml of exponentially growing cultures received 100 μl of 1 M potassium phosphate buffer (pH 8.0) and 100 μl of lysin solution (DOC at a final concentration of 1% or 0.1%). The mixtures were incubated for 15 min at 37°C. Turbidities of the solutions were read at 600 nm using an Ultraspec III instrument (Pharmacia LKB, Cambridge, United Kingdom). When the turbidity of the cell suspension decreased more than 50% from the initial value, the assay was considered positive. The experiments were repeated three times on different days.

Phy and myl PCR detection and BsaI-RFLP signatures. The presence of phy (encoding pneumolysin, a cholesterol-dependent cytolsin) or myl (a phy homologue recently identified in some S. mitis isolates that has been named mitilysin) (16) was detected by PCR. The primers used for detection of both genes were phyF (5′-TTCTGTAAACAGCTACCAAG-3′) or phyG (5′-CGATGAGTTTGTGGTATACG-3′) with phyR (5′-ACCTATCYTCTACCTGAGG-3′), yielding an internal fragment of 1,123 bp with primer phyF and 1,283 bp with primer...
phyF2. Alignment of the sequences of phy and mly alleles available at NCBI (accession numbers EF413922 to EF413926, EF413929 to EF413931, EF413933, EF413934, EF413936, EF413937, and EF413939, and EF413943 for phy alleles and EF66514 to EF66520 for mly alleles) showed that the PCR products obtained for these genes could be distinguished by differential restriction sites for BsaAI; BsaAI cuts the phy PCR product once, resulting in two fragments of 830 bp and 393 bp (or 890 bp and 393 bp), while the mly PCR product does not contain a restriction site for BsaAI. Thus, after purification of the PCR products, the fragments were digested with BsaAI and separated by agarose gel electrophoresis, and the presence of phy or mly was inferred from the pattern obtained.

**Screening for pneumococcus-specific lipoprotein piA4.** Detection of piA4, a gene encoding the lipoprotein component of the pneumococcus iron ABC transporter Pia and described as 100% conserved and uniquely found in pneumococci (47), was done by PCR. The primers used were those previously described: piaFor (5’-AGGACGATCCGCGTGAATAAT-3’) and piaArev (5’-CATGA GGCTGCTAACCCTGGTAT-3’) (47).

**Multilocus sequence typing.** Amplification of internal fragments of seven housekeeping genes—aroE, gdh, gki, recP, spi, xpt, and ddi—was done according to the MLST scheme developed by Enright and Spratt for *S. pneumoniae* (6). Sequencing reactions were conducted at Macrogen, Inc. (Seoul, South Korea). Sequencing analysis was done with DNASTar (Lasergene). Tentative allelic assignment was done at the international MLST database for *S. pneumoniae* (www.mlst.net).

**Multilocus sequence analysis.** Phylogenetic analysis of MLST data was done by concatenating the sequences of all MLST loci except ddi to obtain one single sequence of 2,758 bp (9). One isolate failed to yield an amplification product for a single locus (gdi for PT5645a), and, in this case, a shorter contig was constructed for MLST analysis of *S. pneumoniae* (393 bp). Strains belonging to *S. mitis*, *S. pneumoniae*, and *S. oralis*, previously described by Hanage et al. and Chi et al. (4, 9) were also used. Sequence alignment was performed using the ClustalW algorithm included in MEGA, version 4, build 4025 (43). A bootstrap consensus tree inferred from 1,050 replicates was taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the maximum composite likelihood method.

**Comparative genomic hybridization.** Microarrays used in this study have been described in detail before and contain PCR amplicons representing 2,087 open reading frames (ORFs) of *S. pneumoniae* TIGR4, all spotted in duplicate (10). Chromosomal DNA for CGH was isolated from bacterial cultures by cetyltrimethylammonium bromide extraction as described previously (44). In all cases, TIGR4 genomic DNA was used as the reference sample. For CGH, 400 ng of DNA from each strain was fluorescently labeled using the BioPrime Array CGH Genomic Labeling System (Invitrogen). The concentrations of nonfluorescent nucleotides in the 50–µl reaction mixtures were 0.2 mM each of dATP, dCTP, and dGTP and 0.1 mM dTTP, and fluorescent nucleotide analogs (Cy5-dUTP or Cy3-dUTP, Amersham Biosciences) were added to a final concentration of 0.1 mM. The reaction mixture was incubated overnight at 37°C, and the reaction was stopped by addition of 5 µl of 0.5 M EDTA (pH 8.0). Reactions were cleaned up using the purification module of the BioPrime kit, after which yields and incorporation of dye were verified using a Nanodrop ND-1000 (Nanodrop Technologies). Labeled sample and reference DNA were combined and precipitated with ethanol in the presence of 0.3 M sodium acetate (NaAc; pH 5.2). Dried samples were resuspended in 65 µl of Slidehyb buffer 1 (Ambion) and applied to the microarrays underneath a lifter slip (Erie Scientific Co.). Following overnight hybridization at 45°C, arrays were washed with 2× SSC (1× SSC = 0.15 M NaCl plus 0.015 M sodium citrate; Invitrogen) containing 0.25% SDS for 5 min, followed by two washes in 1× SSC and 0.5× SSC for 5 min each. Finally, slides were dipped into H2O and dried by centrifugation. Two replicate hybridizations (dye swap) were performed for all strains.

**CGH data acquisition and analysis.** Dual-channel array images were acquired on a GenePix 4000AL microarray scanner and analyzed with GenePix Pro software (Axon Instruments, Union City, CA). Automatically flagged spots, spots with low-background-subtracted signal intensities (sum of median Cy3 and Cy5 net signals of <−1,000) and spots with >40% saturated pixels were filtered out of all data sets prior to analysis. Slides used for CGH of pneumococcal strains were normalized by an array-based Lowess transformation. Due to the highly skewed distribution of log ratios with the atypical strains, array-based Lowess transformation could not be performed with these arrays. Instead, we selected a subset of genes (list available on request) that gave consistent hybridization signals for all strains and calculated a normalization factor so that the mean ratio across this gene set was 1. Next, average normalized log ratios were calculated for genes with at least three measurements per strain in the dye swap pair. As a final selection criterion, only genes with valid data for 15 out of the 25 atypical strains (total, 1,838) were used. Designation of genes in each strain as present, divergent, or absent was performed using the graded assignment categorization option of the GACK software (22). GACK uses the ratio distribution per strain to calculate an estimated probability of presence (EPP) value for each gene without the need for arbitrarily defined cutoffs. A calculated EPP of 100% indicates that a gene is present (assigned 0.5), an EPP of 0% indicates a gene is absent (assigned −0.5), and intermediate EPP values indicate that a gene is divergent (assigned a value between −0.5 and 0.5 on a linear scale, where a value of 0 indicates that a gene has a 50% chance of being divergent or present). Clustering was based on the GACK scores and was done with the Euclidian distance metric and average linkage, using TIGR Multixperiment Viewer (MeV; http://www.tm4.org/mev.html).

**Analysis of comC allelic variation.** Chromosomal DNA from each strain was obtained using a Wizard Genomic DNA Purification Kit (Promega, Madison, WI) according to the manufacturer’s instructions. comC PCR products of 337 bp were obtained as previously described by Whatmore and coworkers (48) using the primers comCf (5’-TGAAGTGTGAGAGAATCTT-3’) and comCr (5’-C TTTCATTTTATTGACCT-3’). Sequencing was conducted at Macrogen, Inc. (Seoul, South Korea), and subsequent analysis of obtained sequences was done with DNASTar (Lasergene).

### RESULTS

**Presumptive identification of atypical streptococci.** The 25 strains described in this study were all isolated from selective agar medium (containing gentamicin) routinely used to recover pneumococci while inhibiting the growth of other bacterial species that colonize the nasopharynx. All strains displayed a pneumococcus-like colony morphology, and although they were resistant to optochin, they were bile soluble. Therefore, they were presumptively identified as optochin-resistant pneumococci.

**Abnormally high levels and prevalence of penicillin and multidrug resistance.** Antimicrobial susceptibility testing showed that all but one of the 25 strains were nonsusceptible to penicillin (MIC higher than 0.06 mg/liter), and 64% were multidrug resistant (defined as nonsusceptibility to three or more classes of antimicrobial agents) (Table 1). Among the 19 isolates resistant to macrolides, 3 contained ermB, 14 contained either mefA or mefE, and two contained both ermB and either mefA or mefE; the 7 isolates resistant to tetracycline contained tetM (Table 1). Furthermore, penicillin MICs of 15 strains were high, ranging from 2 to 6 mg/liter, values that are extremely rare in Portuguese colonization samples and could have important implications. These results prompted a detailed characterization to clarify whether the 25 strains were true pneumococci.

**Phenotypic characterization.** No pneumococcal capsule could be detected in any of the strains by the Quellung reaction: eight isolates showed no positive reaction, and the remaining ones autoagglutinated. In addition, cpsD could not be detected by PCR in any of the strains, suggesting its absence.

To determine if the isolates showed the CO2-dependent optochin phenotype described for *S. pseudopneumoniae* (1), we examined optochin susceptibility in 5% CO2 and ambient atmosphere in parallel. Twenty isolates were fully resistant to optochin in both environments (no halos were obtained). Four isolates (PT5295b, PT5557b, PT3645, and PT5729) had increased susceptibility in ambient atmosphere but were still considered resistant according to the proposed interpretation criteria. One isolate (PT5479) had results identical to those described for *S. pseudopneumoniae*; i.e., it was resistant to
<table>
<thead>
<tr>
<th>Strain Code</th>
<th>MIC (mg/liter)</th>
<th>Resistance Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT5804</td>
<td>&lt;256</td>
<td>0.38, 0.094</td>
</tr>
<tr>
<td>PT5798b</td>
<td>256</td>
<td>0.064, 0.38, 2</td>
</tr>
<tr>
<td>PT5796b</td>
<td>256</td>
<td>0.064, 0.38, 2</td>
</tr>
<tr>
<td>PT5794b</td>
<td>256</td>
<td>0.064, 0.064, 3</td>
</tr>
<tr>
<td>PT5793b</td>
<td>256</td>
<td>0.064, 0.38, 3</td>
</tr>
<tr>
<td>PT5645b</td>
<td>256</td>
<td>0.064, 0.064, 0.094</td>
</tr>
<tr>
<td>PT5590</td>
<td>256</td>
<td>0.064, 0.064, 0.253</td>
</tr>
<tr>
<td>PT5645</td>
<td>256</td>
<td>0.064, 0.064, 0.38</td>
</tr>
</tbody>
</table>

**TABLE 1**: Properties of strains characterized in this study.
optochin when incubated in 5% CO₂ but susceptible when incubated in ambient atmosphere.

Deoxycholate solubility assays at 1% and 0.1% final concentrations following the protocol proposed by Obregon et al. (32) were carried out. These authors reported that streptococci of the mitis group (SMG) harboring an atypical lytA did not lyse with 1% DOC solution but lysed at 0.1%; i.e., in the latter case the turbidity of the cell suspension decreased at least 50% after 15 min at 37°C. In our study, all but one strain lysed in the presence of both concentrations of DOC, a result reminiscent of typical pneumococci and similar to control strains R6 and TIGR4. The only exception found (PT5479) was the strain that, based on optochin susceptibility, was tentatively classified as *S. pseudopneumoniae*. The same behavior was observed for *S. pseudopneumoniae* control strain ATCC BAA960.

**Screening for pneumococcus-specific genes.** Screening of *lytA* by PCR yielded a product with the expected size in 18 isolates (Table 1). By BsaI-RFLP, a signature identical to the atypical *lytA* found in some streptococci of the mitis group was obtained for 17 isolates (data not shown) (24). For one strain (PT5534b) a restriction pattern could not be obtained (Table 1). In the remaining seven isolates no PCR product could be obtained despite several repeated attempts under different experimental conditions. However, by Southern blotting of SmaI-PFGE patterns, in all but one strain (PT5479), at least one fragment hybridized with a *lytA*-specific probe, suggesting that some form of this gene was present (Table 1).

PCR screening for *ply* or *mly* indicated that all 25 strains harbored one of these allelic variants. Subsequent BsaI-RFLP analysis indicated that five isolates harbored pneumolysin and 20 harbored mitilysin (Table 1). PCR screening for the *piaA* gene suggested that it was absent in all 25 strains characterized in this study.

**Genotyping by PFGE and MLST/MLSA.** PFGE analysis showed a high level of diversity between the 25 atypical strains as 23 distinct PFGE patterns were identified, and only two pairs of strains displayed the same profile (Table 1).

Sequence analysis of the seven housekeeping genes included in the *S. pneumoniae* MLST scheme showed that among the 174 alleles obtained, only 7 had been previously deposited in the pneumococcal database (i.e., the alleles that are assigned numbers in Table 1). All other alleles diverged up to 9% from the closest available match. The simultaneous occurrence of six to seven novel alleles in each strain suggested that they were not authentic pneumococci (9).

The phylogenetic tree resulting from the MLSA clearly placed the strains described in this study apart from *S. pneumoniae* and *S. oralis* (Fig. 1). In particular, all strains but one clustered together with the isolates of the streptococci of the mitis group described by Chi et al. (4). However, among the SMG, the strains described here formed a separate group that diversified close to the root of the tree. One strain, PT5479, clustered together with several *S. pseudopneumoniae* isolates described by Hanage et al. (9).

**Genomic characterization by CGH.** To obtain insights into the variation of gene content between the isolates described in this study and *S. pneumoniae*, we performed CGH using *S. pneumoniae* TIGR4 as a reference strain. In addition to the 25 atypical isolates, we also used seven true pneumococcal strains (six clinical isolates and strain R6), *S. mitis* NCTC 12261, and *S. pseudopneumoniae* ATCC BAA960. Clustering of strains based on the GACK absent/present/divergent gene scores for the 1,838 TIGR4 genes that passed our selection criteria showed that all 25 atypical strains clustered together and that they appeared to be more similar to streptococcus of the mitis group than to the cluster of true pneumococcal strains, consistent with the MLSA findings. CGH also confirmed the close relationship of the two strain pairs with identical PFGE pattern, PT5525b/PT5557b, and PT5787b/PT5790b. A few gene clusters appeared to be absent or divergent in all strains examined, including the pneumococcal strains such as the TIGR4-specific capsule locus, the *rha*A pathogenicity islet, a cluster of hypothetical ORFs, and a fructose phosphotransferase system (PTS).

Of the 1,838 genes used for the phylogenetic analysis, 1,365 belonged to the pneumococcal core genome proposed by Obert et al. (31). CGH analysis indicated that 394 pneumococcal core genes were divergent or absent in at least one of the 25 atypical isolates, with an average of 157 per strain. In total, 53 core genes were divergent or absent in all 25 isolates. In *S. pseudopneumoniae* and *S. mitis*, respectively, 135 and 318 core genes were absent or divergent. Interestingly, most of the absent or divergent core genes were grouped into 12 TIGR4 chromosomal regions of diversity (see Table S1 in the supplemental material). These genes are predicted to encode proteins with a variety of functions, but most prominent classes are involved in transport (e.g., ABC transporters and PTS systems) and energy metabolism.

To assess whether the atypical strains contained factors known to be important for the infection potential of *S. pneumoniae*, we focused on the CGH results for some of the major pneumococcal virulence genes (Fig. 2). As already demonstrated by the PCR-RFLP analysis described above, *ply* or *mly* (ply-like) sequences could be detected in all of the atypical strains. The *lytA* gene was found to be present in 11 of the 25 strains while it was classified as absent or divergent in the remainder, including five of the seven strains that failed to yield a *lytA* PCR product. Major differences between *S. pneumoniae* and the atypical strains appeared to reside in the PiaA iron ABC transporter system (as was observed by PCR) and in genes encoding the choline-binding protein surface proteins ChpG and PcpA and the neuraminidases NanA and NanB (Fig. 2). By contrast, other genes involved in virulence in *S. pneumoniae*, such as *pavA*, *spxB*, *chpE*, and *chpD*, were found to be present in the atypical isolates as well.

**Allelic variation of competence stimulating peptide (comC).** Most members of the mitis group are naturally competent for genetic transformation, and CGH indicated that the known competence genes were indeed present in our isolates. To examine potential sequence diversity of *comC* in more detail, we sequenced a *comC* PCR product from 23 strains. A total of 14 different alleles based on the amino acid sequence of ComC were identified. Of these, only four have been previously described (Table 2) (21, 35).

**DISCUSSION.**

This study describes several phenotypic and genotypic properties of a collection of streptococcal isolates obtained in a single cross-sectional study aimed to describe the impact of...
PCV7 in colonization. Based on routinely used assays such as colony morphology, bile solubility, and optochin susceptibility, these isolates were initially classified as optochin-resistant, bile-soluble pneumococci. However, they did not agglutinate with anti-pneumococcal capsular antibodies, and several isolates displayed unusually high MICs to penicillin (2 to 6 mg/liter), indicating a sudden increase in the prevalence of penicillin resistance from 1.7% to 5% over a period of 3 years.

Further attempts to characterize these isolates by PCR screening for lytA and plyA were not conclusive as most strains appeared to have some form of lytA and plyA. Additional analyses showed that the isolates contained atypical lytA genes, as defined by Llull et al. (24); five had pneumolysin, the others had the recently described mitilysin (16), and all lacked cpsA and piaA. However, the latter two genes are known to be absent in some nonencapsulated true pneumococci (47), and,
thus, their usefulness in defining the species is limited when they are absent. DNA fingerprinting by PFGE indicated that the strains were genetically diverse. Epidemiological data also suggested that they were mostly unrelated as they came from seven different day care centers.

Isolates distinct from pneumococci but that cannot be resolved from them by optochin susceptibility, bile solubility, or the presence of genes such as \textit{lytA} and \textit{plyA}, have been described in the literature (9, 14, 23, 32, 45, 49).

To establish the relationship between our strains and true pneumococci, we characterized them by MLSA and CGH and compared results with those obtained for control strains of \textit{S. pneumoniae}, \textit{S. pseudopneumoniae}, \textit{S. mitis}, and \textit{S. oralis}. On the basis of MLSA, the isolates were identified as streptococci of the mitis group, a classification that has been proposed for a group of evolutionarily related streptococcal isolates for which species classification challenges currently accepted criteria (32). In the SMG group, one cluster included several \textit{S. pseudopneumoniae} isolates described by Hanage et al. and strain PT5479 isolated in this study (9). The latter was the single strain that had been identified as \textit{S. pseudopneumoniae} based on optochin susceptibility assays. CGH also indicated

![Gene content variation in pneumococcal virulence factors. CGH data for selected genes encoding major pneumococcal virulence factors are shown. The colors indicate the status of the ORFs, as follows: blue, present; yellow, absent; gray, no data.](image)

**FIG. 2.** Gene content variation in pneumococcal virulence factors. CGH data for selected genes encoding major pneumococcal virulence factors are shown. The colors indicate the status of the ORFs, as follows: blue, present; yellow, absent; gray, no data.

---

**TABLE 2.** \textit{comC} alleles of strains characterized in this study

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Amino acid sequence of \textit{comC}</th>
<th>Reference</th>
<th>Allele/strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT5714</td>
<td>MKNTVKLEQFVALKEKDLQNIKKGESRISDILLDLFLQFKK</td>
<td>35</td>
<td>CSP8</td>
</tr>
<tr>
<td>PT5790</td>
<td>MKNTVKLEQFVALKEKDLQNIKKGESRISDILLDLFLQFKK</td>
<td>35</td>
<td>CSP6.1</td>
</tr>
<tr>
<td>PT5479, PT5779</td>
<td>MKNTVKLEQFVALKEKDLQNIKKGESRISDILLDLFLQFKK</td>
<td>35</td>
<td>CSP6.1</td>
</tr>
<tr>
<td>PT5274, PT5525c, PT5736b, PT5787b, PT5790b, PT5794b</td>
<td>MKNTVKLEQFVALKEKDLQNIKKGESRISDILLDLFLQFKK</td>
<td>35</td>
<td>CSP6.1</td>
</tr>
<tr>
<td>PT5645</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5790b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5798b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5804</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5346b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5729</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5590b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5532, PT5796b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
<tr>
<td>PT5645b</td>
<td>MKNTVKLEQFVTLEKDLQIKIQGSESRLSRLLDFQFKQIQ</td>
<td>21</td>
<td>SK612</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Sequence of the mature CSP is given in boldface.
that the atypical strains are indeed distinct from true pneumo-
cocci. However, it showed that a considerable part of the pro-
posed pneumococcal core genome (88% on average) is con-
served in these isolates. Absent regions corresponded mainly
to loci encoding proteins involved in nutrient uptake and me-
tabolism. Interestingly, many of the known pneumococcal vir-
ulence factors were detected by CGH, with a few notable ex-
ceptions: CbpG, a surface protein that functions as an ad-
hesin to eukaryotic cells, the neuraminidases NanA and NanB,
and the Pia iron-uptake system.

Why species definition is sometimes difficult (if not impos-
sible) for some nonhemolytic streptococcal isolates is a matter
of debate. Killian et al. (21) have recently proposed that the
ancestor of the pneumoniae-mitis-pseudopneumoniae group
was a pneumococcus-like bacterium with all the properties
associated with virulence. The commensal streptococci sub-
sequently evolved from this pathogen by genome reduction, thus
explaining the random (but syntenic) presence of pneumococ-
cal virulence genes in representatives of this group. Others
have suggested that horizontal gene transfer between naturally
competent streptococci that share the same ecological niche
significantly contributes to attenuation of putative barriers be-
tween these species, resulting in not only mosaic genes but also
mosaic genomes (1, 4, 7, 8). While our results based on CGH
do not enable us to favor any of these theories, they do
enhance our perception of how closely related these spe-
cies are.

Indeed, known competence genes required for natural for-
formation were detected in the atypical isolates by CGH. Fur-
thermore, several different comC alleles, coding for the com-
petence stimulating peptide (CSP), were found by sequence
analysis, most of which appeared to be novel. Cross-commu-
nication between isolates requires recognition of the CSP by a
ComD receptor. Cross talk between different alleles due to
promiscuous ComD has been described (17). Whether that is
the case for the isolates described in this study has not been
investigated but could potentially contribute to the ambiguous
nature of these strains.

High rates of resistance to penicillin and high MICs are
frequent among S. mitis isolates (13, 32, 33). Therefore, mis-
identification of SMG isolates such as those described here can
result in overreported levels of penicillin resistance, a problem
previously identified (36, 46). In our study, after ruling out the
15 SMG isolates, penicillin resistance levels among pneumo-
cocci were found to have remained stable over time.

It is unclear why we have identified these SMG isolates only
now, as we have been conducting surveillance studies since
1996. Most (15 of 25) were picked from the primary selective
blood agar plate as secondary isolates that displayed colony
morphologies distinct from the majority of the sample but still
resembled pneumococci. Our recent focus on noncapsulated
pneumococci and optochin resistance, as well as an interest in
multiple colonization, most likely contributed to an increased
propensity to pick such colonies. Alternatively (and not mutu-
ally exclusively), it could be that changes in the nasopharynx
ecosystem due to introduction of the seven-valent pneumococ-
cal conjugate vaccine facilitated colonization by SMG, thereby
increasing its abundance in the population. The prevalence of
these “confounding” organisms remains unknown, and their

contribution to pneumococcal evolution remains to be ascer-
tained.

In conclusion, in the era of multivalent pneumococcal con-
jugate vaccines where surveillance is very important, research-
ers need to be aware that routinely used tests to identify pneu-
 mococci may not be sufficient when atypical isolates are found.
This is critical in colonization studies since closely related streptococci inhabit the same niche and can be highly resistant
to antibiotics. On the other hand, although S. pseudopneu-
moniae has been implicated in chronic obstructive pulmonary
disease (COPD) (19), it is not clear whether the atypical con-
founding organisms described in this study may cause disease.
This will be the subject of future research, aiming at increasing
the insight into the ecological and clinical roles of these strains.

ACKNOWLEDGMENTS

This work was partially supported by projects Grace (contract
PLS18226) and Pneumopath (contract Health-F3-2009-222983) from
the European Commission and projects PTDC/BIA-MIC/64101/2006
and PTDC/SAU-ESA/65048/2006 from Fundação para a Ciência e
Tecnologia (FCT). A.S.S. was supported by FCT (SFRH/BD/27325/
2006). H.J.B. was supported by iOP Genomics grant IGE03002 from
the Dutch Ministry of Economic Affairs. This publication made use of
the Multi Locus Sequence Typing website (http://www.mlst.net) at
Imperial College London developed by David Aanensen and Man-
Suen Chan and funded by the Wellcome Trust.

We report that we have no conflicts of interest.

REFERENCES

1. Arbique, J. C., C. Poyart, P. Trieu-Cuot, G. Quesne, M. D. S. Carvalho, A. G.
Steigerwalt, R. E. Morey, D. Jackson, R. J. Davidson, and R. R. Facklam.
2004. Accuracy of phenotypic and genotypic testing for identification of
Streptococcus pneumoniae and description of Streptococcus pseudopneu-
2. Bentley, S. D., D. M. Ansen, A. Mavroidis, D. Saunders, E. Rabbinowitsch,
M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C.
Skovsted, M. S. Kaltstoft, B. Barrett, P. R. Reeves, J. Parkhill, and B. G.
Spratt. 2006. Genetic analysis of the capsular biosynthetic locus from all 90
barrier: evolution and spread of a major class of mosaic pmb2x in Strepto-
509.
for antimicrobial susceptibility testing; 19th informational supplement.
CLSI/NCCLS M100-S19. Clinical and Laboratory Standards Institute,
Wayne, PA.
for Streptococcus pneumoniae: identification of clones associated with serious
9. Hanage, W. P., T. Kaijalainen, E. Herva, A. Saukkoriipi, R. Syjänen, and
B. G. Spratt. 2005. Using multilocus sequence data to define the pneumo-
Regulation of gene expression in Streptococcus pneumoniae by response
1985. Detection of C polysaccharide in Streptococcus pneumoniae in the
sputa of pneumonia patients by an enzyme-linked immunosorbent assay.
phenotypic analyses to identify nonhemolytic streptococci isolated from bac-
Legakis, and A. Vatopoulos. 2001. Antibiotic resistance rates and macrolide