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

ADRIAN M. WHATMORE,1* ANDROULLA EFSTRATIOU,2 A. PAUL PICKERILL,1 KAREN BROUGHTON,2 GEOFFREY WOODARD,1 DANIEL STURGEON,1 ROBERT GEORGE,2 AND CHRISTOPHER G. DOWSON1

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL,1 and Respiratory and Systemic Infection Laboratory, Central Public Health Laboratory, London, NW9 5HT,2 United Kingdom

Received 25 August 1999/Returned for modification 17 September 1999/Accepted 10 November 1999

The oral streptococcal group (mitis phylogenetic group) currently consists of nine recognized species, although the group has been traditionally difficult to classify, with frequent changes in nomenclature over the years. The pneumococcus (Streptococcus pneumoniae), an important human pathogen, is traditionally distinguished from the most closely related oral streptococcal species Streptococcus mitis and Streptococcus oralis on the basis of three differentiating characteristics: optochin susceptibility, bile solubility, and agglutination with antipneumococcal polysaccharide capsule antibodies. However, there are many reports in the literature of pneumococci lacking one or more of these defining characteristics. Sometimes called “atypical” pneumococci, these isolates can be the source of considerable confusion in the clinical laboratory. Little is known to date about the genetic relationships of such organisms with classical S. pneumoniae isolates. Here we describe these relationships based on sequence analysis of housekeeping genes in comparison with previously characterized isolates of S. pneumoniae, S. mitis, and S. oralis. While most pneumococci were found to represent a closely related group these studies identified a subgroup of atypical pneumococcal isolates (bile insoluble and/or “acapsular”) distinct from, though most closely related to, the “typical” pneumococcal isolates. However, a large proportion of isolates, found to be atypical on the basis of capsule reaction alone, did group with typical pneumococci, suggesting that they have either lost capsule production or represent as-yet-unrecognized capsular types. In contrast to typical S. pneumoniae isolates phenotypically identified as S. mitis and S. oralis, which included isolates previously characterized in taxonomic studies, were genetically diverse. While most of the S. oralis isolates did fall into a well-separated group, S. mitis isolates did not cluster into a well-separated group. During the course of these studies we also identified a number of potentially important pathogenic isolates, which were frequently associated with respiratory disease, that phenotypically and genetically are most closely related to S. mitis but which harbor genes encoding the virulence determinants pneumolysin and autolysin classically associated with S. pneumoniae.

Streptococcus pneumoniae is a common and important human pathogen associated with pneumonia, septicaemia, meningitis, and otitis media. A number of distinct species of naturally transformable viridans or oral streptococci, which are closely related to S. pneumoniae, have been now identified, although the taxonomy and classification of these organisms has long been considered difficult (47). The most closely related species on the basis of 16S rRNA sequence are Streptococcus oralis and Streptococcus mitis, which share over 99% sequence identity with S. pneumoniae, although DNA-DNA similarity values for the entire chromosome are estimated to be less than 60% (28). S. oralis and S. mitis are usually considered to be commensals of the human oral cavity, but in recent years it has become clear that members of these species can be important pathogens. Oral streptococci, including S. mitis and S. oralis, are associated with bacterial endocarditis, especially in patients with prosthetic valves (16). In addition, S. mitis and S. oralis are now recognized as frequent causes of infection in immunocompromised patients, particularly immediately after tissue transplants, and in neutropenic cancer patients (5, 8, 9, 33).

Four phenotypic characteristics are classically used in the diagnostic laboratory for the presumptive identification of S. pneumoniae: colony morphology, optochin sensitivity, bile solubility, and agglutination with antipneumococcal polysaccharide capsule antibodies. Although their colony morphology can be very similar, nonpneumococcal oral streptococci are classically optochin resistant and bile insoluble and do not react with antipolsaccharide antibodies. It is considered important for the laboratory to differentiate between S. pneumoniae and other alpha-hemolytic oral streptococci, notably S. mitis and S. oralis, since misidentification may influence diagnosis and treatment. While these conventional methods allow identification of the majority of pneumococcal isolates, presumptive S. pneumoniae isolates may produce atypical reactions in one or more of the standard tests, and other alpha-hemolytic streptococci may give positive reactions in these tests, leading to difficulties in identification. For example, there are many ob-
servation of optochin-resistant pneumococci (1, 30, 37, 38), bile-insoluble \textit{S. pneumoniae} isolates have been reported (15, 21, 36), and nontypeable, unencapsulated pneumococci have historically been reported to comprise 2% of pneumococci isolated from normally sterile sites (10) and up to 20% of conjunctival isolates (22).

Despite these problems, most diagnostic laboratories still use these conventional identification techniques. A commercially available DNA probe test (AccuProbe Culture Identification Kit for \textit{S. pneumoniae}) has been reported to show specificity and sensitivity of 100% for pneumococci (13, 14), but it is prohibitively expensive for routine diagnostic use (36). The use of PCR in diagnostic tests, targeting genes considered to be specific targets for the pneumococcus, has been investigated. Two attractive targets for such studies have been genes encoding the putative virulence factors pneumolysin (\textit{ply}) and the major autolysin (\textit{lytA}). Such PCR-based diagnostics, while offering some promise, are not, as yet, used frequently in the routine microbiology laboratory setting.

Here we describe the preliminary characterization of a group of “atypical” presumptive isolates of \textit{S. pneumoniae} which show aberrant reactions to bile and/or do not react with antipneumococcal polysaccharide antibodies. We have used sequencing of housekeeping genes to investigate the genetic background of these organisms in comparison to a selection of “typical” \textit{S. mitis} strain sequences previously characterized and clinical isolates of \textit{S. oralis} and \textit{S. mitis}. In addition, we describe a second group of atypical isolates identified during the course of this study. These organisms appear to be genetically and phenotypically related to \textit{S. mitis} but harbor genes encoding the virulence factors autolysin and pneumolysin, which are normally associated with pneumococci. Interestingly, many of these isolates were disease associated. The relationships of these organisms to other well-characterized oral streptococci and a preliminary characterization of the genes encoding autolysin and pneumolysin in these organisms are described here.

\section*{MATERIALS AND METHODS}

\subsection*{Strains}

A complete list of strains used in this study of genetic diversity is provided in Table 1. Isolates were routinely cultured on brain heart infusion (BHI) agar supplemented with 5% (vol/vol) sheep’s blood at 37°C and 5% CO2.

\subsection*{Preparation of chromosomal DNA.}

Chromosomal DNA was prepared from each isolate according to the method described by Sambrook et al. (37). DNA was isolated from mid-logarithmic phase cultures grown on blood agar plates. The DNA was purified by ethanol precipitation and dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5).

\subsection*{Phylogenetic analysis.}

Preliminary analysis and alignment of sequences was performed in DNAStar. Phylogenetic analysis was performed using the MEGA suite of programs (31). Phylogenetic trees were constructed by the neighbor-joining method with the Jukes-Cantor correction. The bootstrap confidence levels of internal branches, as defined for MEGA (31), were calculated from 500 random resamplings of the original sequence data with replacement.

\subsection*{Screening of isolates for pneumolysin and autolysin genes.}

All isolates were screened for the presence of the pneumolysin (\textit{ply}) gene, major autolysin (\textit{lytA}) gene, and atypical autolysin (\textit{lytA}101) encoding genes by PCR with primer sets \textit{plyAmp} and \textit{lytAmp} as described previously (46) or, in the case of \textit{lytA}101, the primer \textit{lytAmp} in combination with a novel specific reverse primer \textit{lytA}101\textit{d} (5’-CTACTTCATCGTAATCAAACCGTCAGGTTC-3’). The specificity of PCR products was confirmed by probing with digoxigenin-labeled \textit{ply} and \textit{lytA} fragments according to the manufacturer’s instructions using a nylon membrane. Probes were obtained after PCR by using the same primer sets, cloning, and confirming by southern blotting the hybridization of \textit{ply}, \textit{lytA}, and \textit{lytA}101 to the appropriate chromosomal bands. Probes were then used in a DNA:DNA hybridization assay.

\subsection*{Sequence analysis of housekeeping genes.}

\textit{S. pneumoniae} 101 from \textit{S. pneumoniae} 101 (5’-CTCTGATACGTCCGAAATCATTCTT-3’), \textit{P. aeruginosa} (5’-ACCCGGAGCGCTTTATTTTCC-3’), and \textit{P. stuartii} (5’-ATGTCGCTACTCCGGATTTTTC-3’), and \textit{S. mitis} strain \textit{xpt} (5’-GAATTTAGACGGCATGACGAT-3’) and \textit{S. milii} (5’-TATGAGATCTGCTCCAATTTAAAG-3’), respectively. PCR products were subjected to standard conditions with 32 cycles of 95°C for 1 min, X°C for 1 min, and 72°C for 1 min, where X°C represents an annealing temperature appropriate for the particular primer set used. Fragments were purified by PAGE and prepared for sequencing using an ABI 373 automated sequencing system. Sequences of 395bp (\textit{trpA}), 285bp (\textit{ply}), and 339bp (\textit{hexB}) were obtained from each isolate and used in the analysis presented here.

\section*{RESULTS AND DISCUSSION}

\subsection*{Preliminary characterization of putative virulence factors in pneumococcal isolates.}

To determine whether the putative \textit{lytA} gene was present in the pneumococcal isolate \textit{S. pneumoniae} 101, the \textit{lytA} gene was amplified by PCR using the primer sets \textit{lytAmp} (5’-CCATGGACCGCGCTTCA-3’) and \textit{lytBadn} (5’-CGCTGAATTACGTCGAAACTCTT-3’), \textit{P. aeruginosa} (5’-ACCCGGAGCGCTTTATTTTCC-3’), \textit{P. stuartii} (5’-ATGTCGCTACTCCGGATTTTTC-3’), and \textit{S. mitis} strain \textit{xpt} (5’-GAATTTAGACGGCATGACGAT-3’) and \textit{S. milii} (5’-TATGAGATCTGCTCCAATTTAAAG-3’), respectively. PCR products were subjected to standard conditions with 32 cycles of 95°C for 1 min, X°C for 1 min, and 72°C for 1 min, where X°C represents an annealing temperature appropriate for the particular primer set used. Fragments were purified by PAGE and prepared for sequencing using an ABI 373 automated sequencing system. Sequences of 395bp (\textit{trpA}), 285bp (\textit{ply}), and 339bp (\textit{hexB}) were obtained from each isolate and used in the analysis presented here.

\subsection*{Sequence alignment and analysis of housekeeping genes.}

Sequence alignment and analysis of housekeeping genes were performed in DNAStar. Phylogenetic analysis was performed using the MEGA suite of programs (31). Phylogenetic trees were constructed by the neighbor-joining method with the Jukes-Cantor correction. The bootstrap confidence levels of internal branches, as defined for MEGA (31), were calculated from 500 random resamplings of the original sequence data with replacement.

\subsection*{Screening of isolates for pneumolysin and autolysin genes.}

All isolates were screened for the presence of the pneumolysin (\textit{ply}) gene, major autolysin (\textit{lytA}) gene, and atypical autolysin (\textit{lytA}101) encoding genes by PCR with primer sets \textit{plyAmp\textit{lytAmp}} and \textit{lytAmp\textit{lytAmp}d} as described previously (46) or, in the case of \textit{lytA}101, the primer \textit{lytAmp} in combination with a novel specific reverse primer \textit{lytA}101\textit{d} (5’-CTACTTCATCGTAATCAAACCGTCAGGTTC-3’). The specificity of PCR products was confirmed by probing with digoxigenin-labeled \textit{ply} and \textit{lytA} fragments according to the manufacturer’s instructions using a nylon membrane. Probes were obtained after PCR by using the same primer sets, cloning, and confirming by southern blotting the hybridization of \textit{ply}, \textit{lytA}, and \textit{lytA}101 to the appropriate chromosomal bands. Probes were then used in a DNA:DNA hybridization assay.

\subsection*{Sequence analysis of housekeeping genes.}

\textit{S. pneumoniae} 101 from \textit{S. pneumoniae} 101 (5’-CTCTGATACGTCCGAAATCATTCTT-3’), \textit{P. aeruginosa} (5’-ACCCGGAGCGCTTTATTTTCC-3’), \textit{P. stuartii} (5’-ATGTCGCTACTCCGGATTTTTC-3’), and \textit{S. mitis} strain \textit{xpt} (5’-GAATTTAGACGGCATGACGAT-3’) and \textit{S. milii} (5’-TATGAGATCTGCTCCAATTTAAAG-3’), respectively. PCR products were subjected to standard conditions with 32 cycles of 95°C for 1 min, X°C for 1 min, and 72°C for 1 min, where X°C represents an annealing temperature appropriate for the particular primer set used. Fragments were purified by PAGE and prepared for sequencing using an ABI 373 automated sequencing system. Sequences of 395bp (\textit{trpA}), 285bp (\textit{ply}), and 339bp (\textit{hexB}) were obtained from each isolate and used in the analysis presented here.

\subsection*{Preparation of chromosomal DNA.}

Chromosomal DNA was prepared from each isolate as described previously (45).

\subsection*{Sequence alignment and analysis of housekeeping genes.}

Sequence alignment and analysis of housekeeping genes were performed in DNAStar. Phylogenetic analysis was performed using the MEGA suite of programs (31). Phylogenetic trees were constructed by the neighbor-joining method with the Jukes-Cantor correction. The bootstrap confidence levels of internal branches, as defined for MEGA (31), were calculated from 500 random resamplings of the original sequence data with replacement.
### TABLE 1. Characteristics of strains examined in this study

<table>
<thead>
<tr>
<th>Strain no.</th>
<th>Other ID</th>
<th>Identification</th>
<th>Origin</th>
<th>Site of isolation</th>
<th>Latex agglutination/sero-type</th>
<th>Optochin</th>
<th>Bile solubility</th>
<th>Gen-Probe</th>
<th>Presence of hyA hy4101 ply</th>
<th>Source or reference</th>
</tr>
</thead>
</table>
#### Group 1 (typical S. pneumoniae)
- PC15 1011 S. pneumoniae UK Throat 23F S + + + + + 35
- 11B 912 S. pneumoniae UK Throat 15B S + + + + + 35
- PC19 1012 S. pneumoniae UK Throat 35A S + + + + + 35
- CL1 (E26) 355 S. pneumoniae Uruguay Blood 14 S + ND + + + 35
- CL26 (Sp8) 30 S. pneumoniae Spain Blood 13 S + ND + + + 35
- R6 749 S. pneumoniae US R S + + + + 2
- 44A 938 S. pneumoniae UK Throat 6A S + + + + + 35
- 80013 873 S. pneumoniae Kenya Throat 1 S + + + + + 35
- COL1 (Sp9) 555 S. pneumoniae Spain Blood 8 S + ND + + + 35

#### Group 2a (S. oralis and S. mitis characterized in previous taxonomic studies)
- NCTC 11427 542 S. oralis
- NCTC 7864 627/571 S. oralis
- OPA1 626 S. oralis
- PP53 619 S. oralis
- HV51 621 S. mitis
- NCTC 10712 567 S. mitis
- NS51T 620 S. mitis
- K208 622 S. mitis
- OSS1 624 S. mitis

#### Group 2b (commensal and/or clinical isolates of S. oralis and S. mitis)
- AR37 1078 S. oralis UK Blood, endocarditis – R – ND – – – 4, 29
- COL21 PN93/447 Streptococcus sp. UK Eye, eye disease – R – ND – – – This study
- COL25 PN93/1003 S. oralis UK Blood, endocarditis – R – ND – – – This study
- AR5 1072 S. oralis UK Blood, endocarditis – R – ND – – – D. Beighton
- AR13 1076 S. oralis UK Endocarditis – R – ND – – – D. Beighton
- AC1372 1039 S. oralis UK Blood, endocarditis – R – ND – – – D. Beighton
- C17 1073 S. oralis UK Neutropenia – R – ND – – – D. Beighton
- COL19 PN93/800 S. oralis UK Blood – R – ND – – – This study
- COL22 PN93/1264 Streptococcus sp. UK NK – R – ND – – – This study
- M1 Tb(2) 10 S. mitis UK Normal flora – R – ND – – – D. Beighton
- M3 Tb(2) 13 S. mitis UK Normal flora – R – ND – – – D. Beighton
- M4 Tb(2) 14 S. mitis UK Normal flora – R – ND – – – D. Beighton
- M2 Tb(3) 12 S. mitis UK Normal flora – R – ND – – – D. Beighton
- CL22 103 S. mitis Spain Lower respiratory tract – R – ND – – – This study
- AC1374 1040 S. mitis UK Blood culture – R – ND – – – D. Beighton
- AC1362 1042 S. oralis UK Blood culture – R – ND – – – D. Beighton

#### Group 2c (atypical oral streptococci)
- COL18 PN93/454 Streptococcus sp. UK Sputum, chest infection – S – – + + + This study
- COL15 PN92/1139 S. mitis UK Sputum, chest infection – S – – + + + This study
- COL16 PN93/952 Streptococcus sp. UK Bronchial lavage, pneumonia, HIV+ – S – – + + + This study
- 806 COL28 PN93/918 Streptococcus sp. UK NK – R – – + + + This study
- COL20 PN93/776 Streptococcus sp. UK NK – S/R – – + + + This study
- COL24 PN93/656 Streptococcus sp. UK Sputum – S/R – – + + + This study
- COL17 PN91/2475 Streptococcus sp. UK NK – R – – + + + This study
- 764 COL6 PN93/779 S. pneumoniae UK NK – R – – + + + S. Gillespie

#### Group 3 (putative atypical pneumococci)
- 1376 COL1 PN97/3197 S. pneumoniae UK r S + ND + – – This study
- COL3 PN92/1207 S. pneumoniae UK NK r S + ND + – – This study
- COL3 PN92/944 S. pneumoniae UK Eye, eye disease r S + + + + + This study
- COL6 PN93/779 S. pneumoniae UK Blood r S + + + + + This study

(Continued on following page)
Group 2b consisted of both commensal and pathogenic isolates and included in this study as reference strains. For ease of understanding, the strains in Table 1 have been split into several groups. Group 1 consisted of typical S. pneumoniae isolates fulfilling all the conventional criteria of a pneumococcus being serotypable, optochin sensitive, and bile soluble, and the well-characterized laboratory strain R6. Group 2a contained strains of S. mitis and S. oralis that had been used previously in taxonomic studies (4) included in this study as reference strains. Group 2b consisted of both commensal and pathogenic isolates of the most closely related streptococci. Of these isolates were genetically distinct from pneumococci, r, strains which autoagglutinated when latex agglutination-capsular serotyping was performed (rough, nontypeable strains); –, no results against capsular typing sera (smooth, nontypeable strains).

RESULTS

Strains included in this study. A complete list of strains examined in this study and their reactions in classical tests used to distinguish pneumococci from oral streptococci is given in Table 1. One aim of this study was to investigate the genetic relationships between the typical capsular and atypical acapsular and/or bile-insoluble pneumococci sometimes encountered in the diagnostic laboratory. As it became clear that some of these isolates were genetically distinct from pneumococci, isolates of the most closely related streptococci (S. mitis and S. oralis) were included as reference strains. For ease of understanding, the strains in Table 1 have been split into several groups. Group 1 consisted of typical S. pneumoniae isolates fulfilling all the conventional criteria of a pneumococcus being serotypable, optochin sensitive, and bile soluble, and the well-characterized laboratory strain R6. Group 2a contained strains of S. mitis and S. oralis that had been used previously in taxonomic studies (4) included in this study as reference strains. Group 2b consisted of both commensal and pathogenic isolates classified by the suppliers as either S. mitis or S. oralis. All strains in these latter two groups fulfilled conventional criteria for these organisms in being latex agglutination negative, optochin resistant, and bile insoluble. Group 2c consisted of a group of organisms we have called atypical oral streptococci: this designation is based largely upon the presence of pneumococcal virulence factor genes as described later. However, many of these latter isolates were originally obtained because they proved difficult to classify and were considered unusual; many showed some sensitivity to optochin and displayed aberrant biochemical reactions (data not shown). Group 3 consisted of organisms described as atypical pneumococci largely on the basis of being acapsular, either because they autoagglutinated (rough strains) or because they simply failed to react with any pneumococcal antisera. However, some also showed reduced sensitivity to optochin and some were also bile insoluble. The prototype atypical S. pneumoniae strain 101/87, originally described by Diaz et al. (13) on the basis of reactivity with a lytA probe (21), was included within this group as a reference strain.

Genetic relationships between isolates. In order to determine the genetic relationships between all of the isolates examined in this study, we examined sequence fragments from three housekeeping genes: xpt, recP, and hexB. By using the neighbor-joining method, dendrograms of genetic relationships between isolates were constructed with sequence data from each of the three genes individually to examine tree topology (data not shown). After this, the sequence data of all three genes were combined and treated as a single contiguous fragment in order to obtain an overall picture of genetic relationships between these isolates, as illustrated in Fig. 1.

Apart from strains characterized as typical pneumococci, the overall picture provided by the data was one of extensive genetic diversity. Only two major groups were strongly supported by bootstrapping. First, many of the S. pneumoniae isolates fall into a subgroup, labeled A on the dendrogram, containing all isolates characterized as typical pneumococci—the identical group was also supported by bootstrap values of at least 90% when each of the three genes was considered individually. This group is relatively conserved compared with the remaining organisms: the mean nucleotide diversity between members of this group is 1.42% (range, 0 to 3.0%). A second strongly supported group, labeled group C, contained a much more diverse group of organisms predominantly identified as S. oralis and included the S. oralis type strain. With the exception of strain AC1372, which harbored a distant xpt gene, all of these strains fell within the same group in trees constructed by using the three individual housekeeping genes. A feature of this group was the extensive nucleotide diversity relative to that of isolates of typical S. pneumoniae, with a mean diversity of 16.65% (range, 8.7 to 21.2%).

<table>
<thead>
<tr>
<th>Strain no.</th>
<th>Other ID</th>
<th>Identificationa</th>
<th>Originb</th>
<th>Site of isolationc</th>
<th>Latex agglutination/seryotypea</th>
<th>Optochina</th>
<th>Bile solubility</th>
<th>Geno-Probec</th>
<th>Presence of d, e, f lytA lytA101 ply</th>
<th>Source or referencec</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL5</td>
<td>PN93/832</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Blood</td>
<td>r</td>
<td>S/R +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL7</td>
<td>PN93/356</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Nasal, respiratory infection</td>
<td>r</td>
<td>S/R +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL8</td>
<td>PN93/707</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Sputum</td>
<td>–</td>
<td>S/R +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL11</td>
<td>R93/608</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>CSF, meningitis</td>
<td>–</td>
<td>S +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL12</td>
<td>PN93/950</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Blood</td>
<td>–</td>
<td>S +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL14</td>
<td>PN93/789</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Blood, pneumonia</td>
<td>–</td>
<td>S +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL9</td>
<td>PN93/904</td>
<td>S. pneumoniaea</td>
<td>UK</td>
<td>Blood</td>
<td>–</td>
<td>S/R +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>X158</td>
<td></td>
<td>Streptococcus sp.</td>
<td></td>
<td>r</td>
<td></td>
<td>S</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL27</td>
<td>PN93/135</td>
<td>Streptococcus sp. UK</td>
<td>Sputum, chest infect; HIV+</td>
<td>r</td>
<td>S/R –</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>COL26</td>
<td>PN93/403</td>
<td>Streptococcus sp. UK</td>
<td>Synovial fluid, septic arthritis</td>
<td>r</td>
<td>S/R –</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>This study</td>
</tr>
<tr>
<td>151/87</td>
<td>86027</td>
<td>S. pneumoniaea b</td>
<td>Spain</td>
<td>Blood, pneumonia</td>
<td>–</td>
<td>R –</td>
<td>ND</td>
<td>–</td>
<td>+</td>
<td>15</td>
</tr>
</tbody>
</table>

a Unless stated otherwise, identities are based on the phenotypic tests described in this table in combination with API biotyping and additional biochemical tests when necessary. Where strain profiles were not consistent with any one species or where discordant identities were obtained on separate occasions, strains are not given a species designation. T, type strain.

b UK, United Kingdom.

c ND, not done.

d The presence of lytA, lytA101, and ply was determined by PCR and confirmatory blotting.

e O. Brighton, Joint Microbiology Research Unit, Guy’s, King’s, and St. Thomas’ Dental Institute, London, United Kingdom; S. Gillespie, Department of Medical Microbiology, Royal Free Hospital, London, United Kingdom.

f The species identities of group 2a strains of S. oralis and S. mitis and S. pneumoniae 101/87 are those of strain suppliers (4, 15).
of organisms. Apparently most closely related to, but clearly distinct from, the typical pneumococci are a number of organisms (Col5 to 101/87) classified as atypical pneumococci either on the basis of being acapsular or displaying aberrant reactions to bile and/or optochin. Thus, there appears to be a genuine genetic distinction between typical pneumococci and some of the organisms that have historically been described as atypical pneumococci. The majority of the remaining 20 organisms were identified as S. mitis. This loose association of genetically diverse isolates included commensal S. mitis isolates, isolates associated with invasive disease, isolates previously characterized in taxonomic studies, and the S. mitis type strain. Once again, these organisms were characterized by their extensive genetic diversity. The mean nucleotide diversity within group B (see Fig. 1), which contains most of these organisms, is 16.2% (range, 0.02 to 21.2%).

Genetic relationships of acapsular putative pneumococci. One initial aim of this study was to examine whether acapsular pneumococci are genetically distinct from typical serotypable pneumococci. Of the 16 acapsular organisms included in group 3, nine were closely related to typical capsular pneumococci and presumably represented organisms that either do not express their capsule or else represent as-yet-unrecognized capsular types. However, the remaining organisms (Col5, Col1, 86027, Col27, Col26, 1916, and 101/87) were clearly genetically distinct from typical capsular organisms.

Relationship of phenotype to genetic background. Not one of the classical phenotypic criterion was found to satisfactorily distinguish the genetic groups uncovered in this study. As stated above, many of the acapsular pneumococci are clearly typical pneumococci. Many of the atypical pneumococci and the atypical oral streptococci displayed aberrant reactions to optochin and bile. Recently, use of the GenProbe test has been reported to provide accurate and sensitive identification of pneumococci (13, 14), and we therefore used this test on a subset of our isolates. As might be expected, this test, with a nucleic acid probe to a housekeeping gene, provided the most accurate reflection of genetic relationships determined in this study. In addition to all of the typical pneumococci tested, all of the organisms clustering in the atypical pneumococcal group
The pneumolysin gene (ply) was studied by using PCR primers (Table 1, Fig. 1). The specificity of PCR products was confirmed by probing PCR products transferred to a membrane with the *S. pneumoniae* R6 ply gene at high stringency. All isolates falling into the *S. pneumoniae* grouping (both typical and atypical) possessed ply and, as expected, isolates within the *S. oralis* grouping did not contain ply. Unexpectedly, a number of isolates which fell within the probe chromosomal DNA (digested with *Pvu*II, which is predicted not to cut within ply on the basis of the published sequence) of some of the group 2c organisms (764, Col15, Col17, Col16, Col18, and Col20) and two of the atypical pneumococci, Col26 and Col27 (data not shown). This probe hybridized with all of the chromosomal DNAs and, in agreement with the extensive genetic diversity among these organisms demonstrated by sequence analysis of housekeeping genes, all isolates displayed distinct restriction profiles. An additional interesting finding was that multiple bands in digests from some of these strains hybridized to the *ply* probe. Whether these represent the presence of multiple copies of pneumolysin or merely sequence diversity within a single copy of *ply* requires further investigation.

As final confirmation of the presence of *ply* in nonpneumococcal oral streptococcal isolates, hemolytic titer assays were performed to compare the hemolytic activity associated with some of the atypical *S. mitis* strains with both pneumococci and “conventional” *S. mitis* isolates (Table 2). Both typical pneumococcal controls (strain PC19 and the well-characterized laboratory strain D39) and an atypical pneumococcus (Col27) displayed substantial intracellular and extracellular hemolytic titers. All strains genetically allied to *S. mitis*, but in which ply had been detected (Col15, Col17, Col20, and 764), also displayed hemolytic activity in both cytoplasmic and extracellular fractions. The specificity of the assay for pneumolysin was confirmed by the inclusion of a ply-deficient D39 mutant strain in these assays. In contrast, *S. mitis* isolates in which ply was not detected by genetic approaches, as well as several *S. oralis* isolates, lacked any hemolytic activity under the assay conditions used in this study.

**Distribution of the autolysin gene.** The distribution of the autolysin gene (*lytA*) was studied by using PCR primers specific for either the typical *lytA* sequence (primers *lytA*up and *lytA*dn) or the atypical *lytA*101 sequence (primers *lytA*up and *lytA*101dn) previously reported from the bile insoluble Spanish *S. pneumoniae* isolate 101/87 (Table 1). Once again, specificity was confirmed by using either R6 *lytA* or 101/87 *lytA*101 to probe membrane-bound PCR products. As expected, all isolates that form part of the pneumococcal group (typical and atypical) were found to possess *lytA*. Indeed, all but one (764) of the isolates in which *ply* was detected were also found to possess *lytA*. No isolates were found to contain *lytA* but lack *ply*. In contrast to *lytA*, a *lytA*101-specific product was detected in only the strain *S. pneumoniae* 101/87, from which this gene was originally isolated (15), and one of the atypical *S. mitis* strains, Col17.

**Sequencing of the *lytA* gene found in atypical isolates.** It might be expected that isolates possessing *lytA* would display the characteristic bile soluble phenotype associated with this gene (42). This was clearly not the case for some isolates of atypical pneumococci (e.g., Col26 and Col27) and the atypical oral streptococci. Thus, in order to examine the relationship of *lytA* seen in the bile insoluble atypical pneumococci (Col26 and Col27) and some of the atypical oral streptococci (Col15, Col18, Col20, Col16, and Col17) with the known *lytA* sequence (24), PCR products obtained from these strains were sequenced in full. The relationships between these sequences are illustrated in a dendrogram (Fig. 2) comparing the sequences determined here with the original published sequence of *lytA* from strain Rst7, nine *lytA* allelic variants recently reported in typical pneumococci (44), and *lytA*101 (101/87). All of the sequences were up to 20% divergent from the published Rst7 *lytA* sequence and other allelic variants seen in typical pneumococci but were much more closely related to *lytA*101 sequence (5 to 7% divergent). Interestingly, there was between 0.3 and 4.5% divergence within this group, somewhat higher than the levels of divergence than that reported between allelic variants in typical pneumococci (44).

**Relationship of genetic background and clinical association.** Many of the *S. oralis* isolates included in this study were blood isolates obtained from immunocompromised patients. However, these isolates involved in infection appeared to be genetically diverse and widely distributed among the isolates characterized in previous taxonomic studies. Similarly, there was no clear separation of the *S. mitis* strains containing *lytA* and *ply* and associated with disease from the commensal isolates of *S. mitis* and those characterized previously in taxonomic studies.

### TABLE 2. Comparison of extracellular and intracellular hemolytic titers of representative isolates examined in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Strain ID</th>
<th>Cytoplasmic</th>
<th>Extracellular</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. pneumoniae</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>PC19</td>
<td>729</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>D39</td>
<td>81</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>D39ply</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical</td>
<td>Col27</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td><em>S. mitis</em> group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>NS51T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical</td>
<td>Col15</td>
<td>243</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Col17</td>
<td>729</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Col20</td>
<td>729</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>764</td>
<td>27</td>
<td>10</td>
</tr>
<tr>
<td><em>S. oralis</em> group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical</td>
<td>NCTC 11427T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Col21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Col25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Hemolytic titer is expressed as the reciprocal of the lowest dilution at which complete lysis of horse blood was seen. Extracellular titers were multiplied by 10 such that they represent a cell number comparable to that used to obtain cytoplasmic titers. –, No hemolytic activity detected.
FIG. 2. Dendrogram of genetic relationships between lytA sequences examined in this study constructed from gene sequence data by using the neighbor-joining method. Only bootstrap values exceeding 90% are shown. The scale represents the number of nucleotide substitutions per site. The upper group of the tree consists of the previously published lytA sequences from strain Rst7, a typical pneumococcal isolate, and a number of allelic variants of lytA recently reported from typical pneumococci (44). The lower group contains the sequence lytA101 from the classical bile-insoluble atypical pneumococcus 101/87 and the lytA sequences determined from atypical pneumococci and atypical oral streptococci examined in this study.

DISCUSSION

The primary aim of this study was to examine whether there is a genetic basis to “difficult” or “atypical” isolates sometimes submitted to diagnostic laboratories which have been tentatively identified as pneumococci or oral streptococci but display aberrant phenotypic features. Since the dendrogram of genetic relationships constructed in this study is based on only three housekeeping genes from species that are naturally transformable, we are reluctant to make phylogenetic interpretations based on these data. Recombination is known to occur frequently within pneumococci with the species displaying an epidemic population structure (20, 26, 35). In addition interspecies recombination events between pneumococci and S. oralis have been seen at least in genes under intense selective pressures, such as penicillin-binding-protein encoding genes (17–19). This possibility of recombination makes phylogenetic interpretation potentially hazardous. Having said this, a number of observations are consistent for all three genes examined when considered individually and in the overall analysis, allowing some confidence in them. Two groupings were strongly supported in all trees. These were a subgroup containing the typical pneumococci and the clear separation of the S. oralis group. The atypical organisms (Col5, Col1, Col27, Col24, Col26, 1916, and 86027) were generally placed as the organisms most closely related to pneumococci. In the case of all gene fragments considered individually and the data as a whole the remaining strains (largely identified as S. mitis) formed a relatively diverse group of organisms with few consistent branches.

One obvious feature of the data is the extent of diversity seen within the S. oralis and S. mitis groups when compared to typical pneumococci. This may reflect sampling strategy, although we deliberately selected typical pneumococci thought from previous studies (35) to represent the breadth of genetic diversity within the species. The diversity seen in oral streptococci is consistent with a number of previous reports. For example, among 101 isolates of S. mitis examined, 93 distinct PvuII ribotypes were reported (23). Similarly, restriction fragment length polymorphism analysis of S. mitis populations in and between individuals found limited sharing of genotypes among family members and some 6 to 13 types in individual subjects (27). Repetitive extragenic palindromic PCR has been used to study S. oralis diversity with populations in individuals found to be heterogeneous at a single time point and highly variable when they are monitored longitudinally (3). Recently, PCR-based fingerprinting was also used to examine the relationships between oral streptococci isolated from the blood of neutropenic cancer patients, and all isolates were found to display distinct fingerprint patterns (48).

It is clear from the housekeeping gene sequence data that many isolates that fail to react in the latex agglutination system represent entirely typical pneumococci. It is still not clear why these organisms fail to react with pneumococcal antisera. They may represent as-yet-uncharacterized capsular serotypes, they may possess capsule genes that are nonfunctional or not expressed or, alternatively, they may lack the genes encoding the capsule biosynthetic pathway. We are currently investigating these alternative possibilities using the organisms described in this study. However, a proportion of apparently acapsular isolates did not group with other pneumococci and represent genetically divergent organisms. Thus, there does appear to be a genuine genetic basis to previously described atypical pneumococci which can both fail to react in capsular typing and/or show aberrant biochemical reactions. In light of the extensive genetic diversity within these organisms, it is not surprising that separation on the basis of phenotypic criteria can prove notoriously difficult. The Gen-Probe test was the diagnostic test that appeared to most closely match genetic relationships between isolates with all typical pneumococci and virtually all atypical isolates as far out as Col26 tested giving a positive reaction. Since the basis of this test is hybridization to a housekeeping gene, the correlation with the data presented here is not unexpected. Mundy et al. (36) reported recently that many isolates which show a discordant combination of reactions in capsular typing, optochin sensitivity, and bile solubility tests represent pneumococci on the basis of the Gen-Probe test. In that study the authors suggested the use of Gen-Probe to unequivocally identify isolates showing such discordant pheno-
typic reactions. Our results indicate that many of the isolates identified as pneumococci by Gen-Probe could actually represent organisms which are genetically rather divergent from typical pneumococci.

Perhaps the most surprising finding of this study was the characterization of isolates phenotypically and genetically allied to S. mitis harboring genes encoding the putative virulence factors pneumolysin and autolysin normally associated with pneumococci. There has been at least one previous report of a lytA probe reacting with optochin-resistant and bile-insoluble streptococci (21), and the use of this gene as a probe to identify atypical pneumococci has been suggested. Our results suggest that using lytA probes to identify difficult organisms as atypical pneumococci could also select organisms which are genetically more closely related to S. mitis. We are not aware of previous reports of the presence of the pneumolysin encoding gene in oral streptococci. Both lytA and ply were absent from previously characterized S. mitis isolates, as well as commensal isolates obtained in this study. Comparative assays of hemolytic activity associated with S. mitis isolates harboring ply and lacking ply (basing the species identity on observed genetic relationships described in this study) suggested that the hemolysin is actively expressed by the ply-containing S. mitis strains. In virtually all cases where a strain history was available, the S. mitis isolates harboring lytA and/or ply were associated with respiratory disease. It is thus tempting to speculate that these genes may enhance the pathogenic potential of these organisms relative to the typical commensal organisms. The detection of ply and lytA in disease-associated putative S. mitis strains provided the impetus for us to include S. oralis strains associated with invasive disease in immunocompromised patients in this study. However, there was no evidence of the presence of lytA or ply in members of this group. It is interesting to consider these findings in the light of virulence studies showing that ply is important in murine intranasal infection (7, 11) and of a recent signature-tagged mutagenesis study (38) which suggested that lytA is important for establishing pneumonia but is not necessary in septicemia. Our data may reflect this in that lytA-containing S. mitis isolates were associated with pneumonia and respiratory tract disease, while isolates of S. oralis from the blood of neutropenic patients did not possess lytA. Perhaps the debilitated state of these patients allows ready access of “commensal” isolates to the bloodstream, bypassing the need for any true invasive step. The detection of “pneumococcal” virulence factor genes in S. mitis also has implications for attempts to use them as diagnostic targets in PCR (25, 40, 41, 43). Clearly, these organisms could cross-react in such tests, and the community needs to be aware that organisms other than “typical” pneumococci may be identified by such tests although, since they may be pathogenic, detection of these organisms is not necessarily an argument against the development of such tests.

We also performed a preliminary characterization of the ply and lytA genes found in S. mitis isolates. Interestingly, although PCR products were obtained from these strains by using a lytA primer set and, apart from one isolate, not with the lytA101 primer set, their lytA genes were found to be much more closely related to lytA101 than to the typical lytA sequence. The activity of LysA101 is known to be inhibited by sodium deoxycholate, and this property is believed to be responsible for the bile-insoluble phenotype of the host strain 101/87 (15). While the bile insolubility of these strains could reflect their possession of a lytA101-like sequence, other possibilities, such as an autolysin refractory cell wall, cannot be ruled out without further experimentation. Probing Southern blots of the same strains with ply suggested the possibility of substantial diversity of ply within S. mitis isolates despite the fact that evidence to date suggests that ply in S. pneumoniae is a relatively conserved gene (32, 34). An alternative interpretation is that the S. mitis isolates may possess multiple copies of this gene. We are currently investigating both of these possibilities. In light of these results it is interesting to speculate on the evolutionary history of this group of organisms. Clearly, the organisms allied to S. mitis but harboring lytA and ply are genetically diverse, as demonstrated by housekeeping gene sequencing, distinct restriction profiles when probing for ply, and lytA sequencing. This may reflect the fact that these genes are moving freely between organisms with distinct genetic backgrounds. In this respect it is interesting to note all but one strain appeared to harbor both genes and that ply and lytA are closely linked in the pneumococcal genome lying some 7 kb apart. An alternative, though not mutually exclusive, scenario is that these genes represent ancient characteristics that have been lost by some organisms related to the S. mitis group.

We believe it will be important to consider the existence of these atypical organisms in strategies of vaccination against S. pneumoniae. These organisms appear to be of pathogenic potential and may be refractory to immune responses generated by immunogens of typical pneumococci. It is also possible that the atypical organisms we have characterized in this study may act as source of DNA in recombination events, generating new alleles of pneumococcal genes under high selective pressure (17–19). In light of this possibility, it is interesting to note that some of the S. mitis isolates harboring lytA and ply (Col15, Col16, and Col18) have been shown to possess competence-stimulating peptides much more closely related to those of pneumococci than to those reported from other isolates of S. mitis (45). This raises the possibility that these organisms may show an increased likelihood to donate DNA to pneumococci in horizontal gene transfer events simply because their own competence-stimulating peptides may induce a degree of competence in pneumococci. Although this idea requires experimental confirmation, this could mean that these organisms might serve as a pool of variant DNA which could be readily imported into S. pneumoniae genes put under the selective pressure for diversity that inclusion in a potential vaccine might impose. It is thus crucial to examine the nature and extent of genetic diversity in these previously understudied organisms when considering potential vaccine targets in S. pneumoniae.

ACKNOWLEDGMENTS

This work was supported by a project grant from The Wellcome Trust (045171Z/95/Z). A.M.W. is supported by a Wellcome Trust Research Fellowship in Biodiversity.

We are extremely grateful to David Brightton, Rob Whiley, and Steven Gillespie for providing some of the strains included in this study.

REFERENCES
