Phenotypic and genomic characterization of pneumococcus-like streptococci isolated from HIV-seropositive patients

Truls M. Leegaard,^{1†‡} Hester J. Bootsma,^{2†} Dominique A. Caugant,^{1,3} Marc J. Eleveld,² Turid Mannsåker,^{4§} Leif Oddvar Frøholm,¹ Peter Gaustad,⁵ E. Arne Høiby¹ and Peter W. M. Hermans²

Correspondence

Truls M. Leegaard
truls.leegaard@rikshospitalet.no

¹Department of Bacteriology and Immunology, Norwegian Institute of Public Health, Oslo, Norway

²Laboratory of Pediatric Infectious Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

³Department of Oral Biology, University of Oslo, Oslo, Norway

⁴Department of Microbiology, Ullevål University Hospital Oslo, Norway

⁵Institute of Medical Microbiology, Rikshospitalet University Hospital, Oslo, Norway

Accurate differentiation between pneumococci and other viridans streptococci is essential given their differences in clinical significance. However, classical phenotypic tests are often inconclusive, and many examples of atypical reactions have been reported. In this study, we applied various phenotypic and genotypic methods to discriminate between a collection of 12 streptococci isolated from the upper respiratory tract of HIV-seropositive individuals in 1998 and 1999. Conventional phenotypic characterization initially classified these streptococci as *Streptococcus pneumoniae*, as they were all sensitive to optochin and were all bile soluble. However, they did not agglutinate with anti-pneumococcal capsular antibodies and were also far more resistant to antimicrobial agents than typeable pneumococci isolated in the same period. Genotypic characterization of these isolates and control isolates by both multilocus sequence analysis (MLSA) and comparative genomic hybridization (CGH) showed that only a single isolate was genetically considered to be a true *S. pneumoniae* isolate, and that the remaining 11 non-typable isolates were indeed distinct from true pneumococci. Of these, 10 most closely resembled a subgroup of *Streptococcus mitis* isolates genetically, while one strain was identified as a *Streptococcus pseudopneumoniae* isolate. CGH also showed that a considerable part of the proposed pneumococcal core genome, including many of the known pneumococcal virulence factors, was conserved in the non-typable isolates. Sequencing of part of the 16S rRNA gene and investigation for the presence of *ply* by PCR corroborated these results. In conclusion, our findings confirm the close relationship between streptococci of the Mitis group, and show that both MLSA and CGH enable pneumococci to be distinguished from other Mitis group streptococci.

Received 6 October 2009
Revised 27 November 2009
Accepted 1 December 2009

†These authors contributed equally to this work.

‡Present address: Institute of Medical Microbiology, Rikshospitalet University Hospital, NO-0027 Oslo, Norway.

§Present address: Department of Bacteriology and Immunology, Norwegian Institute of Public Health, Oslo, Norway.

Abbreviations: CGH, comparative genomic hybridization; EPP, estimated probability of presence; MLSA, multilocus sequence analysis; MLST, multilocus sequence typing.

The GenBank/EMBL/DDBJ accession numbers for the sequences of the non-typable isolates and for the 16S sequence data are AY624651–AY624727 and AY568578–AY568588, respectively.

INTRODUCTION

Streptococcus pneumoniae is a common and important human pathogen associated with many different types of infection, ranging from localized benign to severe systemic diseases. *S. pneumoniae* is traditionally identified by means of five tests in the diagnostic microbiological laboratory: colony morphology, microscopic morphology, optochin sensitivity, bile solubility and agglutination with anti-pneumococcal capsular polysaccharide sera. The optochin susceptibility test is the most widely used diagnostic test. Although their colony morphology can be similar, non-

pneumococcal streptococci are classically optochin-resistant and bile-insoluble, and do not react with specific pneumococcal anti-polysaccharide antibodies. Given their differences in clinical significance, accurate differentiation between *S. pneumoniae* and other closely related upper respiratory tract streptococci is imperative.

The species of Mitis group (or alpha-haemolytic) streptococci most closely related to *S. pneumoniae* are *Streptococcus mitis*, *Streptococcus oralis* and the new species *Streptococcus pseudopneumoniae*, reported in 2004 (Arbique *et al.*, 2004). The classification of Mitis group streptococci is not straightforward, leading to difficulties in identification. Many examples of atypical reactions to one or more of the standard tests have been reported, such as observations of optochin-resistant (Nunes *et al.*, 2008; Pikis *et al.*, 2001) and bile-insoluble (Mundy *et al.*, 1998) pneumococci. There have also been reports of optochin-susceptible alpha-haemolytic streptococci, but these isolates were bile-insoluble and did not react with pneumococcal capsular antibodies (Martin-Galiano *et al.*, 2003; Wester *et al.*, 2002). Despite these problems, most diagnostic laboratories still use the conventional identification techniques.

In addition to these classical phenotypic assays, various genotypic methods have been used to classify pneumococcal isolates, such as pulsed-field gel electrophoresis (Hermans *et al.*, 1995), PCR tests for virulence genes [e.g. pneumolysin (*ply*; Kearns *et al.*, 2000; Whatmore *et al.*, 1999)], and multilocus sequence typing [MLST; (Enright & Spratt, 1998)]. An additional approach, referred to as multilocus sequence analysis (MLSA), is to observe the distribution of a large number of strains of closely related species in sequence space and to identify clusters of strains that are well resolved from other clusters, which has been used both to differentiate within species and for taxonomic identification. With this approach, phylogenetic trees are inferred from concatenated sequences of multiple core (housekeeping) genes, using for instance the genes of the MLST scheme (Chi *et al.*, 2007; Hanage *et al.*, 2005) or alternative sets of genes (Bishop *et al.*, 2009; Kilian *et al.*, 2008). In recent years, comparative genomic hybridization (CGH) has been used to examine genomic relationships between strains for a variety of microorganisms, including *S. pneumoniae* (Obert *et al.*, 2006; Silva *et al.*, 2006). The use of CGH, together with the increasing number of sequenced genomes, has led to the identification of a candidate pneumococcal core genome, consisting of approximately 70 % of the genome (Hiller *et al.*, 2007; Obert *et al.*, 2006).

We describe a collection of oral streptococci isolated from HIV-positive patients that initially resembled isolates of *S. pneumoniae*, being both optochin-susceptible and bile-soluble, but did not react with anti-pneumococcal polysaccharide antibodies, and were far more resistant to antimicrobial agents than true pneumococci isolated at the same time and place. This strain collection and several

control strains were characterized further genetically by MLSA, 16S rRNA sequencing and PCR analysis for the presence of *ply*. In addition, we explored CGH to comprehensively compare the genomic content of the non-typable isolates to that of *S. pneumoniae*, *S. mitis* and *S. pseudopneumoniae*.

METHODS

Strains and culture conditions. The 12 non-typable streptococci characterized in this study and the control strains used for CGH are listed in Table 1. The non-typable isolates were collected during a survey of the normal flora of HIV-seropositive patients in Norway (Leegaard *et al.*, 2002). During this study, conducted from 1998 to 1999, seven carrier and five clinical isolates from HIV-seropositive patients were isolated at the same hospital. All were initially identified as pneumococci. The carrier isolates were from the naso- and oropharynx. The clinical isolates were from the upper respiratory tract and were cultured in a different laboratory from the carrier isolates. We have been unable to trace the clinical condition(s) that led to the isolation of the clinical isolates. *S. pneumoniae* was routinely grown on Columbia agar (Oxoid) plates supplemented with 5 % sheep blood (Biotrading) at 37 °C and 5 % CO₂.

Phenotypic characterization. All isolates were identified according to conventional methods, including colony morphology, optochin susceptibility (AB Biodisk) both with and without CO₂ in the incubation atmosphere, capsular serotyping by the Quellung reaction using pneumococcal antisera (Statens Seruminstitut, Copenhagen, Denmark) and the bile solubility test (Ruoff *et al.*, 2003).

Antibiotic susceptibility against benzylpenicillin, ciprofloxacin, erythromycin, cefotaxime, ceftriaxone, cotrimoxazole and tetracycline was tested with Etest (AB Biodisk) as instructed by the manufacturer.

MLST and MLSA. The non-typable isolates were subjected to the MLST scheme developed for *S. pneumoniae*, as described by Enright & Spratt (1998). Tentative allele number assignment was performed by comparing sequences to those in the pneumococcal MLST database (<http://spneumoniae.mlst.net/>). Nucleotide sequence data were deposited in GenBank with accession numbers AY624651–AY624727. MLST analysis of *S. pseudopneumoniae* ATCC BAA-960 and *S. oralis* NCTC 11427 was performed in the same way, except that *spi* was amplified from *S. pseudopneumoniae* with primers *spi*-3up and *spi*-3-dn (Porat *et al.*, 2006), and from *S. oralis* with primers *spi*-up2 and *spi*-dn2 (Chi *et al.*, 2007). Previously described (Chi *et al.*, 2007; Hanage *et al.*, 2005) MLST sequences of *S. pneumoniae*, *S. pseudopneumoniae*, *S. mitis* and *S. oralis* strains were retrieved from GenBank or the MLST database.

The sequences of six of the seven MLST loci (*ddl* was not included) were concatenated using the software available at the MLST website (<http://spneumoniae.mlst.net/>). Phylogenetic analysis of the resulting contigs of 2751 bp was conducted using MEGA version 4 (Tamura *et al.*, 2007). Sequence alignment was performed using CLUSTAL W, and sequences were treated as protein-coding data. The bootstrap test of phylogeny was used as the analysis method, with 1050 bootstrap replications and the minimum evolution algorithm. 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.

CGH. Microarrays used in this study have been described before and contain duplicate PCR amplicons representing 2087 ORFs of *S. pneumoniae* TIGR4 (Hendriksen *et al.*, 2007). Chromosomal DNA was isolated from bacteria grown in GM17-broth by cetyltrimethylammonium bromide extraction (Hermans *et al.*, 1995). In all cases,

Table 1. Bacterial strains used in this study

NT, Non-typable; –, unencapsulated derivative of D39 ; NA, not applicable.

Strain	Serotype	Reference
0459	NT	This study
0463	NT	This study
0568	NT	This study
1060	NT	This study
1265	NT	This study
2020	NT	This study
2022	NT	This study
2084	NT	This study
2134	NT	This study
2136	NT	This study
2146	NT	This study
2158	NT	This study
<i>S. pneumoniae</i> D39	2	NCTC 7466
<i>S. pneumoniae</i> R6	–	Hoskins <i>et al.</i> (2001)
<i>S. pneumoniae</i> TIGR4	4	Tettelin <i>et al.</i> (2001)
<i>S. pneumoniae</i> G54	19F	Dopazo <i>et al.</i> (2001)
<i>S. pneumoniae</i> 23F	23	Hiller <i>et al.</i> (2007)
<i>S. pneumoniae</i> INV104B	1	Hiller <i>et al.</i> (2007)
<i>S. pneumoniae</i> INV200	14	Hiller <i>et al.</i> (2007)
<i>S. pneumoniae</i> OXC141	3	Hiller <i>et al.</i> (2007)
<i>S. mitis</i>	NA	NCTC 12261; courtesy of Dr R. Sá-Leão
<i>S. pseudopneumoniae</i>	NA	ATCC BAA-960; courtesy of Dr R. Sá-Leão
<i>S. oralis</i>	NA	NCTC 11427; courtesy of Dr R. Sá-Leão

TIGR4 DNA was used as the reference sample. DNA (400 ng) from each strain was labelled overnight at 37 °C with the BioPrime array CGH genomic labelling system (Invitrogen) as recommended, using 0.2 mM dATP, dCTP and dGTP, 0.1 mM dTTP and 0.1 mM Cy3- or Cy5-labelled dUTP (Amersham Biosciences). Reactions were stopped by adding 5 µl 0.5 M EDTA (pH 8.0) and cleaned up with the purification module of the kit; yields and dye-incorporation were verified using the Nanodrop ND-1000 (Nanodrop Technologies). Labelled sample and reference DNA were combined and precipitated using standard methods. Dried samples were resuspended in 65 µl Slidehyb buffer 1 (Ambion), applied to microarrays underneath a lifterslip (Erie Scientific) and hybridized at 45 °C overnight. Arrays were washed with consecutive 5 min washes in 2 × SSC (Invitrogen) containing 0.25 % SDS, 1 × SSC and 0.5 × SSC. Finally, slides were dipped into water and dried by centrifugation. Two replicate hybridizations (dye swap) were performed for all strains.

CGH data acquisition and analysis. Array images were acquired on a GenePix 4000AL microarray scanner and analysed with GenePix Pro software (Axon Instruments). Automatically flagged spots, spots with low signal intensities (sum of median Cy3 and Cy5 net signals <1500) and spots with >40 % saturated pixels were filtered out of all datasets. Slides used for CGH of pneumococci were normalized by an array-based Lowess transformation. Due to the highly skewed distribution of log ratios with the non-typable strains, similar normalization could not be performed for these strains. Instead, we selected a subset of genes (available on request) with consistent hybridization signals throughout, and calculated a normalization factor so that the mean ratio across this gene set was 1. Mean normalized log ratios were calculated for genes with at least three measurements per strain in the dye-swap pair. Finally, only genes with valid data for 10 of the 12 non-typable strains (1961 total) were used. Genes in each strain were

designated as present, divergent or absent using the graded assignment categorization option of the GACK software (Kim *et al.*, 2002). GACK uses the ratio distribution per strain to calculate the estimated probability of presence (EPP) for each gene without the need for arbitrarily defined cut-offs. Genes are then assigned values on a linear scale between –0.5 (EPP 0%; gene is absent) and 0.5 (EPP 100%; gene is present), 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 Euclidean distance metric and average linkage, using TIGR MultiExperiment Viewer (<http://www.tm4.org/mev.html>).

Pneumolysin gene PCR. The non-typable isolates were screened for the presence of the *ply* gene by PCR using primers *ply* up and *ply* dn as described by Whatmore *et al.* (1999).

16S rRNA sequencing. A fragment of about 450 bp of the 16S rRNA gene of the 12 non-typable isolates was amplified using the generic 16S primers pA and pD and sequenced using the primers pB and pC, as described by Rogall *et al.* (1990). The 16S sequence data were deposited in GenBank (accession numbers AY568578–AY568588).

RESULTS

Phenotypic characterization

The 12 strains isolated from HIV-seropositive patients were found to be non-typable using pooled capsular antisera. These non-typable streptococcal isolates were all positive in the bile solubility test, and repeated optochin testing

Table 2. Characteristics of 12 non-typable streptococci

NG, No growth. All isolates were bile soluble and tested positive for *ply* by PCR.

Isolate no.	Origin	Optochin sensitivity (zone; mm)*		% identity to TIGR4 16S rRNA gene
		With CO ₂	In ambient air	
0459	Clinic	13	18	99.8
0463	Clinic	13	21	99.8
0568	Clinic	15	21	99.4
1060	Clinic	10	15	99.4
1265	Clinic	13	19	98.9
2020	Carriage	13	NG	99.6
2022	Carriage	14	19	99.6
2084	Carriage	14	27	98.1
2134	Carriage	14	NG	99.8
2136	Carriage	14	16	99.2
2146	Carriage	14	NG	99.6
2158	Carriage	14	21	99.2

*Optochin sensitivity in ambient air: ≥ 12 mm, *S. pneumoniae*; no zone, α -haemolytic streptococci. Incubation in CO₂ gives zones 1–3 mm smaller (according to manufacturer).

Fig. 1. Antibiotic susceptibility of true *S. pneumoniae* [black bars; from Leegaard *et al.* (2002)] and closely related streptococcal isolates from HIV-positive patients (white bars; from this study) to benzylpenicillin (a), ciprofloxacin (b), erythromycin (c), cefotaxime (d), ceftriaxone (e), cotrimoxazole (f) and tetracycline (g).

Fig. 2. Phylogenetic relationship of strains based on MLST. This tree was inferred from concatenated sequences (2751 bp) of the MLST loci except that of *ddl*, using the method of minimum evolution. The non-typable isolates are indicated by strain number and the control strains by symbols: ●, *S. pneumoniae*; ●, *S. pseudopneumoniae*; ▲, *S. mitis*; ■, *S. oralis*. The type strains are indicated by arrows. Bar, substitutions per site.

Fig. 3. Phylogeny of strains based on CGH. Each column represents a strain, indicated on top. Non-typable strains are indicated in bold font. Each row corresponds to a gene, arranged in TIGR4 genome order from top to bottom. The dotted arrows indicate gene clusters specific for TIGR4. The 14 main regions of diversity (RD) between pneumococci and the non-typable strains are numbered and indicated by black arrows to the right. The colours indicate the status of the ORFs, as follows: blue, present (GACK score 0.5); yellow, absent (GACK score -0.5); grey, no data.

confirmed their optochin susceptibility, both with and without CO₂ (Table 2). Furthermore, comparison of antibiotic susceptibilities of these 12 non-typable isolates to those of 18 true pneumococci isolated from the same HIV patient population at the same time showed that the non-typable isolates were far more resistant to all tested antibiotics than true pneumococci (Fig. 1).

Amplification of *ply* gene and 16S rRNA analysis

The *ply* gene, encoding the major pneumococcal virulence factor pneumolysin, was detected by PCR in all 12 non-typable isolates and in *S. pseudopneumoniae* ATCC BAA-960, but not in *S. mitis* NCTC 12261. Analysis of the 16S rRNA sequences of the 12 non-typable isolates showed that they displayed 98.1–99.8% identity to TIGR4 16S (Table 2), similar levels as *S. mitis* NCTC 12261 and *S. pseudopneumoniae* ATCC BAA-960 (99.6%; data not shown).

Genotypic characterization by MLST/MLSA

Characterization of the 12 non-typable isolates by MLST indicated that the alleles of 11 of the 12 isolates were at least 3–5% divergent from those of true pneumococci, which made it unlikely that these 11 strains were *S. pneumoniae*. One isolate (0463), however, appeared to be a true non-typable strain of *S. pneumoniae* and was assigned sequence type ST-1054. To further examine the relationship between the non-typable strains and other closely related *Streptococcus* spp., we aligned concatenated sequences of six of the seven MLST alleles (*ddl* was not included) with those of 70 *S. pneumoniae*, *S. pseudopneumoniae*, *S. mitis* and *S. oralis* strains, including the type strains for each non-pneumococcal species. Phylogeny based on MLSA clearly placed our non-typable isolates apart from *S. pneumoniae* and *S. oralis*, with the exception of strain 0463 (Fig. 2). One strain, 1060, clustered with *S. pseudopneumoniae* isolates (the type strain and those described by Hanage *et al.*, 2005), while the 10 remaining non-typable isolates appeared to be most closely related to a subgroup of the *S. mitis* isolates, albeit not the *S. mitis* type strain.

Genomic characterization by CGH

To examine the variation in gene content between the non-typable isolates and other Mitis group streptococci, we performed CGH using *S. pneumoniae* TIGR4 as reference strain, six other true pneumococcal strains (six clinical isolates and unencapsulated strain R6) and the *S. mitis* NCTC 12261 and *S. pseudopneumoniae* ATCC BAA-960 type strains. CGH data obtained with *S. oralis* appeared to be unreliable and were not included in this analysis. Consistent with the MLSA findings, the phylogeny dendrogram based on the GACK absent/present/divergent gene scores for the 1961 TIGR4 genes that passed our selection criteria showed that the non-typable strains

clearly clustered separate from the pneumococcal strains and *S. mitis* NCTC 12261, with the exception of isolate 0463 (Fig. 3). Also in line with the MLSA analysis, isolate 1060 appeared to be most similar to *S. pseudopneumoniae* ATCC BAA-960. A few TIGR4 gene clusters were absent or divergent in all strains examined, including the control pneumococci, such as the loci encoding the serotype 4 capsule locus and the RlrA pathogenicity islet (Fig. 3).

Of the 1961 genes used for the phylogenetic analysis, 1414 belong to the pneumococcal core genome proposed by Obert *et al.* (2006). Of these, 444 were divergent or absent in at least one of the 11 non-typable strains that clustered together, with an average of 212 per strain. In total, 52 core genes were divergent or absent in all 11 isolates. In *S. pseudopneumoniae* ATCC BAA-960 and *S. mitis* NCTC 11427, 220 and 293 core genes, respectively, were absent or divergent. Interestingly, most of these genes were grouped into 14 TIGR4 chromosomal regions of diversity (Figs 3 and 4). They are predicted to encode proteins with a variety of functions, but most prominent classes include transport (e.g. PTS systems) and metabolism.

To assess whether the non-typable strains contain factors known to be important for the infection potential of *S. pneumoniae*, we focused on CGH results for some of the major pneumococcal virulence genes (Fig. 5). Two targets frequently used in PCR diagnostics, pneumolysin (*ply*) and autolysin (*lytA*), were found to be present in all strains examined. Major differences between pneumococci and the non-typable strains appeared to reside in the PiaA iron ABC transporter system, in genes encoding the surface proteins CbpG and hyaluronidase (*hyl*), and, to a lesser extent, the choline-binding protein PcpA and the neuraminidases NanA and NanB.

DISCUSSION

The taxonomy and classification of *S. pneumoniae* and other Mitis group streptococci have long been considered difficult (Whiley & Beighton, 1998). The species most closely related to *S. pneumoniae*, *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 they can be important pathogens (Douglas *et al.*, 1993; Lucas *et al.*, 1997). Here, we describe a collection of 12 streptococci isolated from HIV-seropositive patients, which by conventional phenotypic characteristics, i.e. colony morphology, optochin susceptibility and bile solubility, were classified as pneumococci. However, the capsular swelling reaction was negative and the isolates displayed increased resistance to various antibiotics compared with confirmed pneumococci isolated under the same circumstances at the same time, indicating that these isolates were not true pneumococci. Based on our clinical experience, it is our impression that clinical diagnostic laboratories regularly encounter streptococcal isolates, especially from the upper airways, where it is questionable whether they are true pneumococci or

Fig. 4. Pneumococcal core genes not found in the non-typable strains. The 14 main regions of diversity (RD) are shown with TIGR4 locus indication and putative gene function given to the right. The colours indicate the status of the ORFs, as described in the legend for Fig. 3.

Fig. 5. Gene content variation in pneumococcal virulence factors. CGH data for selected genes encoding major pneumococcal virulence factors are shown. The colours indicate the status of the ORFs, as described in the legend for Fig. 3.

not. Patients infected with HIV are particularly susceptible to bacterial infections with encapsulated organisms, including *S. pneumoniae*, but we do not believe that these isolates are unique to this patient group. The isolates described here are just good examples of the sort of streptococcal isolates in question.

In recent years, a number of genotypic tools have been developed that could aid the correct classification of *S. pneumoniae* and other alpha-haemolytic streptococci, such as amplification of supposed pneumococcus-specific genes like *ply* and *lytA* (Kearns *et al.*, 2000). However, these PCR-based tests have shown variable specificity and sensitivity

for identification of *S. pneumoniae* (Kaijalainen *et al.*, 2002; Kilian *et al.*, 2008) and we detected *ply* by PCR in all strains. A common tool used in bacterial species identification, 16S rRNA sequencing, cannot reliably differentiate pneumococci from Mitis group streptococci, as they show sequence identities of over 99% (Kawamura *et al.*, 1995; Kilian *et al.*, 2008). Similar levels of 16S rRNA sequence identity were observed for the 12 non-typable isolates in our study.

To establish whether these non-typable strains were true pneumococci, we characterized the isolates with two molecular typing methods, MLST/MLSA and CGH, and

compared them to various control strains. MLSA is not generally used to discriminate between species, but has been successfully used for phylogenetic analysis of typable and non-typable pneumococci (Berron *et al.*, 2005; Chi *et al.*, 2007; Hanage *et al.*, 2005; Kilian *et al.*, 2008). On the basis of MLSA, 11 of the non-typable strains were found to be clearly distinct from *S. pneumoniae*, similar to the phylogenetic analysis based on CGH. Characterization of the non-typable strains at the genomic level by CGH showed that a considerable part of the proposed pneumococcal core genome is conserved in these isolates. Main differences appeared to reside in loci encoding proteins involved in nutrient uptake, metabolism and transcriptional regulation. Interestingly, many of the known pneumococcal virulence factors (reviewed by Bergmann & Hammerschmidt, 2006) were detected by CGH, with a few notable exceptions: CbpG, a surface protein that functions as an adhesin to eukaryotic cells; hyaluronidase, which is thought to play a role in bacterial penetration through host tissue; and the Pia iron-uptake system. MLSA showed that, genotypically, 10 of the non-typable isolates were more closely related to a subgroup of *S. mitis* isolates, while one isolate was most related to *S. pseudopneumoniae* ATCC BAA-960. The latter strain was somewhat distinct from *S. pseudopneumoniae* phenotypically, however, as isolates of the latter species are reported to be bile-insoluble and optochin-resistant when incubated in a CO₂-rich atmosphere (Arbique *et al.*, 2004). Given the high genetic diversity of the non-pneumococcal streptococci of the Mitis group (Kilian *et al.*, 2008), the non-typable strains examined in this study are likely to represent one of the multiple separate lineages within the *S. mitis* group.

In conclusion, our results clearly highlight the need for a more specific, genomic-based differentiation between pneumococci and closely related streptococci.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the laboratory technicians for their technical assistance, especially Anne Ramstad Alme, Gro Lemark, Torill Alvestad and Anne-Marie Klem. This study was in part funded by a grant from The Norwegian Ministry of Health and Social Affairs.

REFERENCES

- Arbique, J. C., Poyart, C., Trieu-Cuot, P., Quesne, G., Carvalho, M. G., Steigerwalt, A. G., Morey, R. E., Jackson, D., Davidson, R. J. & Facklam, R. R. (2004). Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov. *J Clin Microbiol* **42**, 4686–4696.
- Bergmann, S. & Hammerschmidt, S. (2006). Versatility of pneumococcal surface proteins. *Microbiology* **152**, 295–303.
- Berron, S., Fenoll, A., Ortega, M., Arellano, N. & Casal, J. (2005). Analysis of the genetic structure of nontypeable pneumococcal strains isolated from conjunctiva. *J Clin Microbiol* **43**, 1694–1698.
- Bishop, C. J., Aanensen, D. M., Jordan, G. E., Kilian, M., Hanage, W. P. & Spratt, B. G. (2009). Assigning strains to bacterial species via the internet. *BMC Biol* **7**, 3.
- Chi, F., Nolte, O., Bergmann, C., Ip, M. & Hakenbeck, R. (2007). Crossing the barrier: evolution and spread of a major class of mosaic pbp2x in *Streptococcus pneumoniae*, *S. mitis* and *S. oralis*. *Int J Med Microbiol* **297**, 503–512.
- Dopazo, J., Mendoza, A., Herrero, J., Caldara, F., Humbert, Y., Friedli, L., Guerrier, M., Grand-Schenk, E., Gandin, C. & other authors (2001). Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate. *Microb Drug Resist* **7**, 99–125.
- Douglas, C. W., Heath, J., Hampton, K. K. & Preston, F. E. (1993). Identity of viridans streptococci isolated from cases of infective endocarditis. *J Med Microbiol* **39**, 179–182.
- Enright, M. C. & Spratt, B. G. (1998). A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**, 3049–3060.
- Hanage, W. P., Kaijalainen, T., Herva, E., Saukkoriipi, A., Syrjanen, R. & Spratt, B. G. (2005). Using multilocus sequence data to define the pneumococcus. *J Bacteriol* **187**, 6223–6230.
- Hendriksen, W. T., Silva, N., Bootsma, H. J., Blue, C. E., Paterson, G. K., Kerr, A. R., de Jong, A., Kuipers, O. P., Hermans, P. W. & Mitchell, T. J. (2007). Regulation of gene expression in *Streptococcus pneumoniae* by response regulator 09 is strain dependent. *J Bacteriol* **189**, 1382–1389.
- Hermans, P. W., Sluif, M., Hoogenboezem, T., Heersma, H., van Belkum, A. & de Groot, R. (1995). Comparative study of five different DNA fingerprint techniques for molecular typing of *Streptococcus pneumoniae* strains. *J Clin Microbiol* **33**, 1606–1612.
- Hiller, N. L., Janto, B., Hogg, J. S., Boissy, R., Yu, S., Powell, E., Keefe, R., Ehrlich, N. E., Shen, K. & other authors (2007). Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. *J Bacteriol* **189**, 8186–8195.
- Hoskins, J., Alborn, W. E., Jr, Arnold, J., Blaszcak, L. C., Burgett, S., DeHoff, B. S., Estrem, S. T., Fritz, L., Fu, D. J. & other authors (2001). Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J Bacteriol* **183**, 5709–5717.
- Kaijalainen, T., Rintamäki, S., Herva, E. & Leinonen, M. (2002). Evaluation of gene-technological and conventional methods in the identification of *Streptococcus pneumoniae*. *J Microbiol Methods* **51**, 111–118.
- Kawamura, Y., Hou, X. G., Sultana, F., Miura, H. & Ezaki, T. (1995). Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. *Int J Syst Bacteriol* **45**, 406–408.
- Kearns, A. M., Wheeler, J., Freeman, R., Seiders, P. R., Perry, J., Whatmore, A. M. & Dowson, C. G. (2000). Pneumolysin detection identifies atypical isolates of *Streptococcus pneumoniae*. *J Clin Microbiol* **38**, 1309–1310.
- Kilian, M., Poulsen, K., Blomqvist, T., Havarstein, L. S., Bek-Thomsen, M., Tettelin, H. & Sørensen, U. B. (2008). Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS One* **3**, e2683.
- Kim, C. C., Joyce, E. A., Chan, K. & Falkow, S. (2002). Improved analytical methods for microarray-based genome-composition analysis. *Genome Biol* **3**, RESEARCH0065.
- Leegaard, T. M., Caugant, D. A., Froholm, L. O., Hoiby, E. A., Rønning, E. J., Sandven, P. & Bruun, J. N. (2008). Do HIV-seropositive patients become colonised with drug-resistant microorganisms? *Eur J Clin Microbiol Infect Dis* **21**, 856–863.

- Lucas, V. S., Beighton, D., Roberts, G. J. & Challacombe, S. J. (1997). Changes in the oral streptococcal flora of children undergoing allogeneic bone marrow transplantation. *J Infect* **35**, 135–141.
- Martin-Galiano, A. J., Balsalobre, L., Fenoll, A. & de la Campa, A. G. (2003). Genetic characterization of optochin-susceptible viridans group streptococci. *Antimicrob Agents Chemother* **47**, 3187–3194.
- Mundy, L. S., Janoff, E. N., Schwebke, K. E., Shanholtzer, C. J. & Willard, K. E. (1998). Ambiguity in the identification of *Streptococcus pneumoniae*. Optochin, bile solubility, quellung, and the AccuProbe DNA probe tests. *Am J Clin Pathol* **109**, 55–61.
- Nunes, S., Sa-Leao, R. & de Lencastre, H. (2008). Optochin resistance among *Streptococcus pneumoniae* strains colonizing healthy children in Portugal. *J Clin Microbiol* **46**, 321–324.
- Obert, C., Sublett, J., Kaushal, D., Hinojosa, E., Barton, T., Tuomanen, E. I. & Orihuela, C. J. (2006). Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. *Infect Immun* **74**, 4766–4777.
- Pikis, A., Campos, J. M., Rodriguez, W. J. & Keith, J. M. (2001). Optochin resistance in *Streptococcus pneumoniae*: mechanism, significance, and clinical implications. *J Infect Dis* **184**, 582–590.
- Porat, N., Greenberg, D., Givon-Lavi, N., Shuval, D. S., Treffer, R., Segev, O., Hanage, W. P. & Dagan, R. (2006). The important role of nontypable *Streptococcus pneumoniae* international clones in acute conjunctivitis. *J Infect Dis* **194**, 689–696.
- Rogall, T., Flohr, T. & Bottger, E. C. (1990). Differentiation of *Mycobacterium* species by direct sequencing of amplified DNA. *J Gen Microbiol* **136**, 1915–1920.
- Ruoff, K. L., Whiley, R. A. & Beighton, D. (2003). Streptococcus. In *Manual of Clinical Microbiology*, pp. 405–421. Edited by P. R. Murray, E. J. Barron, J. H. Jorgensen, M. A. Pfaller & R. H. Tenen. Washington, DC: American Society for Microbiology.
- Silva, N. A., McCluskey, J., Jefferies, J. M., Hinds, J., Smith, A., Clarke, S. C., Mitchell, T. J. & Paterson, G. K. (2006). Genomic diversity between strains of the same serotype and multilocus sequence type among pneumococcal clinical isolates. *Infect Immun* **74**, 3513–3518.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H. & other authors (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**, 498–506.
- Wester, C. W., Ariga, D., Nathan, C., Rice, T. W., Pulvirenti, J., Patel, R., Kocka, F., Ortiz, J. & Weinstein, R. A. (2002). Possible overestimation of penicillin resistant *Streptococcus pneumoniae* colonization rates due to misidentification of oropharyngeal streptococci. *Diagn Microbiol Infect Dis* **42**, 263–268.
- Whatmore, A. M., King, S. J., Doherty, N. C., Sturgeon, D., Chanter, N. & Dowson, C. G. (1999). Molecular characterization of equine isolates of *Streptococcus pneumoniae*: natural disruption of genes encoding the virulence factors pneumolysin and autolysin. *Infect Immun* **67**, 2776–2782.
- Whiley, R. A. & Beighton, D. (1998). Current classification of the oral streptococci. *Oral Microbiol Immunol* **13**, 195–216.

Edited by: M. Kilian