Detection of Large Numbers of Pneumococcal Virulence Genes in Streptococci of the Mitis Group^{∇†}

Calum Johnston,¹ Jason Hinds,² Andrew Smith,³ Mark van der Linden,⁴
Johan Van Eldere,⁵ and Tim J. Mitchell^{1*}

Division of Infection and Immunity, GBRC, 120 University Place, University of Glasgow, Glasgow G12 8TA, United Kingdom¹; Bacterial Microarray Group, Division of Cellular and Molecular Medicine, St. Georges University of London, Cranmer Terrace, London SW17 0RE, United Kingdom²; Infection Research Group, Faculty of Medicine, Glasgow Dental Hospital and School, 378 Sauchiehall Street, Glasgow G2 3JZ, United Kingdom³; National Reference Centre for Streptococci, Institute of Medical Microbiology, University Hospital Aachen, Pauwelsstrasse 30, 52074 Aachen, Germany⁴; and Laboratory of Experimental Bacteriology, Department of Medical Diagnostics, University Hospital Leuven, Herestraat 49, Leuven, Belgium⁵

Received 4 September 2009/Returned for modification 25 March 2010/Accepted 24 May 2010

Seven streptococcal isolates from the mitis group were analyzed for the presence of pneumococcal gene homologues by comparative genomic hybridization studies with microarrays based on open reading frames from the genomes of *Streptococcus pneumoniae* TIGR4 and R6. The diversity of pneumolysin (*ply*) and neuraminidase A (*nanA*) gene sequences was explored in more detail in a collection of 14 *S. pseudopneumoniae* and 29 mitis group isolates, respectively. The mitis group isolates used in the microarray experiments included a type strain (NCTC 12261), two *S. mitis* isolates from the nasopharynxes of children, one *S. mitis* isolate from a case of infective endocarditis, one *S. mitis* isolate from a dental abscess, and one *S. oralis* isolate and one *S. pseudopneumoniae* isolate from the nasopharynxes of children. The results of the microarray study showed that the 5 *S. mitis* isolates had homologues to between 67 and 82% of pneumococcal virulence genes, *S. oralis* hybridized to 83% of pneumococcal virulence genes, and *S. pseudopneumoniae* hybridized to 92% of identified pneumococcal virulence genes. Comparison of the pneumolysin, mitilysin (*mly*), and newly identified pseudopneumolysin (*pply*) gene sequences revealed that *mly* and *pply* genes are more closely related to each other than either is to *ply*. In contrast, the *nanA* gene sequences in the pneumococcus and streptococci from the mitis group are closely clustered together, sharing 99.4 to 99.7% sequence identity with pneumococcal *nanA* alleles.

Streptococcus pneumoniae, also known as the pneumococcus, is an important human pathogen causing high levels of mortality worldwide. Diseases caused include meningitis, pneumonia, and otitis media. *S. pneumoniae* is a member of the viridans group streptococci (VGS) and is closely related to *Streptococcus mitis* and *Streptococcus oralis*, sharing >99% homology by 16S rRNA gene analysis (13, 23). Recently, a subset of “atypical” pneumococci were identified that were distinct from both *S. mitis* and *S. pneumoniae* by housekeeping gene analysis (48). It is likely that these isolates belonged to the species *Streptococcus pseudopneumoniae*, which has since been identified, with isolates appearing phenotypically similar to *S. mitis* but undistinguishable from *S. pneumoniae* in genotypic tests (2). As a result of the close relationships between these species, clear identification can be problematic. A recent study showed that definitive species identification can be achieved by phylogenetic analysis of 4 housekeeping genes. Furthermore, although it has been reported that comparison of superoxide dismutase genes (*sodA*) was sufficient to identify most species,

this did not distinguish between *S. pneumoniae* and *S. pseudopneumoniae* isolates (19).

Despite the close genetic relationships between these species, while *S. pneumoniae* can be a severe pathogen, *S. oralis* and *S. mitis* are usually commensal organisms of mucous membranes. However, on rare occasions, they are able to cause opportunistic disease in immunocompromised patients, particularly those with replacement heart valves, organ transplants, or cancer (4, 11, 28).

The reasons for the dramatic difference in pathogenicity observed between these species are not fully known, although the pneumococcus has a large number of well-characterized virulence factors with known roles in pathogenesis, and these may account for the differences in virulence between these species. However, VGS are naturally transformable, which allows exchange of genetic information between species by homologous recombination (16, 30). Indeed, an increase in penicillin resistance in the pneumococcal population has been attributed to acquisition of resistant penicillin-binding protein variants from *S. mitis* and *S. oralis*, resulting in a “pool” of penicillin resistance genes being transferred between these species (9, 12, 38, 41). Furthermore, homologous recombination with related VGS species has been observed in the pneumococcal virulence genes encoding IgA protease and neuraminidase A (*nanA*) (25, 37), and homologues of pneumococcal virulence factors such as pneumolysin (*Ply*) and autolysin A (*LytA*) have been identified in related VGS species (20, 32, 48).

* Corresponding author. Mailing address: Division of Infection and Immunity, University of Glasgow, B2-28, GBRC, 120 University Place, Glasgow G12 8TA, United Kingdom. Phone: 44 (0)141 330 3749. Fax: 44 (0)141 330 3727. E-mail: t.mitchell@bio.gla.ac.uk.

† Supplemental material for this article may be found at <http://jcm.asm.org/>.

[∇] Published ahead of print on 2 June 2010.

TABLE 1. Strains used in this study

Strain	Species	Isolation site	Disease	Use	Location of isolation	% Virulence genes present ^a
NCTC 12261	<i>S. mitis</i>	Reference isolate		Microarray study	Reference isolate	82.3
R75 A1	<i>S. mitis</i>	Nasopharynx	Commensal	Microarray study	Glasgow, UK	72.7
R77 C4	<i>S. mitis</i>	Nasopharynx	Commensal	Microarray study	Glasgow, UK	71.4
Col	<i>S. mitis</i>	Blood	Endocarditis	Microarray study	Glasgow, UK	77.9
990123	<i>S. mitis</i>	Pus	Dental abscess	Microarray study	Glasgow, UK	67.5
Sv29	<i>S. oralis</i>	Nasopharynx	Commensal	Microarray study	Belgium	83.1
R98 A5	<i>S. pseudopneumoniae</i>	Nasopharynx	Commensal	Microarray study	Glasgow, UK	92.2
RRR468	<i>S. pseudopneumoniae</i>	Unknown	Unknown	Ply study	Germany	NA
RRR471	<i>S. pseudopneumoniae</i>	Unknown	Unknown	Ply study	Germany	NA
RRR472	<i>S. pseudopneumoniae</i>	Unknown	Unknown	Ply study	Germany	NA
R82 E6	<i>S. pseudopneumoniae</i>	Nasopharynx	Commensal	NanA study	Glasgow, UK	NA
R83 E2	<i>S. mitis</i>	Nasopharynx	Commensal	NanA study	Glasgow, UK	NA
R89 G3	<i>S. pseudopneumoniae</i>	Nasopharynx	Commensal	NanA study	Glasgow, UK	NA
R89 G5	<i>S. pseudopneumoniae</i>	Nasopharynx	Commensal	NanA study	Glasgow, UK	NA
R98 A5	<i>S. pseudopneumoniae</i>	Nasopharynx	Commensal	NanA study	Glasgow, UK	NA

^a Determined by microarray study. NA, not applicable.

Many pneumococcal virulence factors are currently being considered vaccine candidates for a protein-based pneumococcal vaccine affording cross-serotype protection. However, the presence of homologues of pneumococcal virulence genes in related species may allow evasion of a vaccine targeting a particular virulence factor by acquisition of a gene variant by homologous recombination (6, 36). As a result, it is important not only to increase our understanding of the biological differences between closely related commensals and invasive isolates but also to provide insight into selection of appropriate targets for future vaccine design.

The aim of this study was to assess the presence of pneumococcal virulence factors in a cohort of streptococci from the mitis group at a genome-wide level, using comparative genomic hybridizations. Furthermore, the diversity and phylogenetic relatedness present in homologues of the pneumococcal virulence genes and vaccine candidates Ply and NanA were explored.

MATERIALS AND METHODS

Bacterial strains and preparation. All strains used in the study are listed in Table 1. Strains were grown in brain heart infusion (BHI) broth (Oxoid) and stored as frozen stocks on Protect beads (Technical Service Consultants). Cells were grown in 20 ml of BHI broth until mid-logarithmic phase (optical density at 600 nm [OD₆₀₀], 0.55 to 0.65). Cell lysates for hemolytic and neuraminidase assays were prepared from 20 ml of culture, as described elsewhere (26).

Gene sequencing, phylogenetic trees, and species identification. PCR amplification of *ply*, *sodA*, and *nanA* genes was carried out using flanking primer pairs 27R and S, 39I and J, and 31R and S, respectively (Table 2). Primers used to fully sequence these genes are also found in Table 2. Cleaning and sequencing of *nanA* genes were carried out at the Molecular Biology Sequencing Unit at the University of Glasgow, as described previously (21). Full gene sequences were aligned and translated into proteins by use of VectorNTI software (Invitrogen). Identification of species was done by *sodA* alignment as described previously (19). Where necessary, *S. pseudopneumoniae* isolates were further identified by optochin resistance and bile solubility (2). Phylogenetic trees were created using MEGA4 (molecular evolutionary genetics analysis) software (www.megasoftware.net).

Hemolytic and neuraminidase assays. Hemolytic assays were carried out as previously described (21). Neuraminidase assays were carried out by a previously described method (29), with modifications. Isolates were grown to mid-log phase (OD₆₀₀ of 0.6) in BHI medium (Oxoid). Induction was achieved by growth in BHI medium plus 50 µg/ml *N*-acetylneuraminic acid (Neu5Ac). Pellets from 20-ml mid-log-phase cultures were tested for neuraminidase activity, and culture supernatants were also tested, after being concentrated ×20 by use of an Amicon

centrifugal filter (Millipore). Sample plates were incubated for 30 min at 37°C. For lysates and supernatants, activity was measured in nanomoles of *p*-nitrophenol (pNP) released/mg total cell protein/hour. Isolates tested can be found in Table 1.

DNA-DNA microarray studies and validation. In order to assess the presence of pneumococcal gene homologues in related VGS species, comparative genomic hybridization studies using *S. pneumoniae* DNA microarrays were performed on 1 *S. oralis*, 1 *S. pseudopneumoniae*, and 5 *S. mitis* isolates (Table 1). These isolates were chosen to represent a range of mitis group isolates from cases of colonization and infection (Table 1). Microarray studies were carried out using *S. pneumoniae* SPv1.1.0 microarrays, designed and produced by the Bacterial Microarray Group, St. George's University of London (BµG@S). These represent all open reading frames (ORFs) from the genomes of *S. pneumoniae* isolates TIGR4 and R6. The array design is available in BµG@Sbase (<http://bugs.sgu.ac.uk/A-BUGS-14>), under accession number A-BUGS-14, and also in ArrayExpress (accession number A-BUGS-14). DNA hybridizations were carried out per standard BµG@S protocols. Test isolate genomic DNA was labeled with Cy5 dye (GE Healthcare), and control TIGR4 genomic DNA was labeled with Cy3 dye, with reciprocal dye swaps performed to eliminate artifacts. In all cases, bacteria were incubated at 37°C in BHI medium to mid-log phase (OD₆₀₀, 0.6) before harvest of DNA for microarrays. Slides were scanned using ScanArray Express (Packard Biosciences), feature extraction and normalization were carried out using BlueFuse for Microarray 3.5 software (Bluegenome Ltd.), and gene lists were created using GeneSpring GX 7.3.1 (Agilent Technologies) software. Under the stringent microarray conditions used, sequence identities above 80% will result in hybridization, whereas identities below this level will likely result in nonhybridization. Validation of comparative genomic hybridizations was done by PCR, using the primers used to design the microarray probes.

Nucleotide sequence accession numbers. The complete sequences of Ply and NanA homologues were deposited in GenBank under accession numbers FJ440109 to FJ440116.

Microarray data accession number. Fully annotated microarray data have been deposited in BµG@Sbase under accession number E-BUGS-82 (<http://bugs.sgu.ac.uk/E-BUGS-82>) and may also be found in ArrayExpress (accession number E-BUGS-82).

RESULTS

Presence of pneumococcal gene homologues in streptococci of the mitis group. The results of the mitis group microarray study showed that a large number of pneumococcal genes had homologues in the 7 isolates tested. The five *S. mitis* isolates tested had homologues to between 72 and 83% of pneumococcal genes, as determined by hybridization to the pneumococcal gene probe. Furthermore, the *S. oralis* isolate hybridized to 84% of the genes, and the *S. pseudopneumoniae* isolate hybridized to 91.4% of them. Comparison of these hybridiza-

TABLE 2. Primers used for PCR amplification and sequencing of pneumococcal genes *sodA*, *ply*, and *nanA*

Primer	Name	DNA sequence (5'–3')	Use
Primers for <i>sodA</i>			
39I	<i>sodA</i> forward	TRCAYCATGAYAARCCACCAT	PCR amplification and sequencing of <i>sodA</i> gene fragment
39J	<i>sodA</i> reverse	ARRTARTAMGCRTGYTCCCARACRTC	PCR amplification and sequencing of <i>sodA</i> gene fragment
Primers for <i>ply</i>			
4T	<i>ply</i> internal forward	GTTGATCGTGCTCCGATGAC	Sequencing of <i>ply</i> gene
4V	<i>ply</i> internal forward	CAATACAGAAGTGAAGGCGG	Sequencing of <i>ply</i> gene
4W	<i>ply</i> internal forward	GATCATCAAGGTAAGGAAGTC	Sequencing of <i>ply</i> gene
9Y	<i>ply</i> internal forward	CGGGATCCGGCAAATAAAGCAGTAAATG ACTTT	Sequencing of <i>ply</i> gene
27R	<i>ply</i> external forward	CTTGGCTACGATATTGGC	PCR amplification and sequencing of <i>ply</i> gene
27S	<i>ply</i> external reverse	TACTTAGTCCAACCACGG	PCR amplification and sequencing of <i>ply</i> gene
27T	<i>ply</i> internal reverse	ATAAGTCATCGGAGCACG	Sequencing of <i>ply</i> gene
Primers for <i>nanA</i>			
22X	<i>nanA</i> internal forward	GAGGAACAGTATGAATCGGAGTG	Sequencing of <i>nanA</i> gene
31R	<i>nanA</i> external forward	TAACAGTACACCTTGACTGC	PCR amplification of <i>nanA</i> gene
31S	<i>nanA</i> external reverse	GTGTTTCGATAAGGATTGAGC	PCR amplification of <i>nanA</i> gene
35E	<i>nanA</i> internal forward	TGACCGCTTCTCCAGCATGC	Sequencing of <i>nanA</i> gene
35F	<i>nanA</i> internal reverse	GAAACCAATGCTTCAAATGG	Sequencing of <i>nanA</i> gene
38P	<i>nanA</i> internal forward	TCCTCAGTAATGCAGGTGGAC	Sequencing of <i>nanA</i> gene
38Q	<i>nanA</i> internal reverse	TGGAGCCTTGTTGACCAATAC	Sequencing of <i>nanA</i> gene
62F	<i>nanA</i> internal forward	CCAGAAGAGGTACAAAAACGTAGTCAAC	Sequencing of <i>nanA</i> gene
62G	<i>nanA</i> internal forward	AGGGTGATCTATACAAGGGTGACCA	Sequencing of <i>nanA</i> gene
62H	<i>nanA</i> external reverse	TGCCTTTGGTTTTTCGGAACCTT	Sequencing of <i>nanA</i> gene
62I	<i>nanA</i> internal reverse	CCCCTTCTTGAGCTAAAACAGGAGA	Sequencing of <i>nanA</i> gene
62J	<i>nanA</i> internal forward	AATCTTTACAAAGGTCAGGAATTGATT	Sequencing of <i>nanA</i> gene
62K	<i>nanA</i> internal forward	GCCGTCTTCATTCTAGTGACTGGGG	Sequencing of <i>nanA</i> gene
62L	<i>nanA</i> internal forward	CAACAAACCAAATAAAGATGGAAT	Sequencing of <i>nanA</i> gene
64I	<i>nanA</i> internal forward	GAATAAATGTCTTATTTTCAGAAATC	Sequencing of <i>nanA</i> gene
64J	<i>nanA</i> internal reverse	TTATTGTTCTCTTTTTCCCTAGC	Sequencing of <i>nanA</i> gene

tion percentages to those for pneumococcal isolates hybridized in a previous study (unpublished data) showed differences in genetic content between the species, implying that *S. pseudopneumoniae* is more closely related to *S. pneumoniae* than the other species are (Fig. 1).

FIG. 1. Comparison of percentages of genes hybridized to pneumococcal microarrays for different VGS species (reference genome, *S. pneumoniae* TIGR4). *S. pneumoniae* hybridization results are representative of hybridization values from previous studies of *S. pneumoniae* strains, using the same arrays, in our laboratory.

Pneumococcal virulence gene homologues present in streptococci of the mitis group. For comparison, pneumococcal virulence genes were considered to be the 231 genes identified in a signature-tagged mutagenesis screen as essential for pneumococcal pneumonia in a murine model (15). Our results showed that homologues to 218/231 virulence genes were present in at least one mitis group isolate, many of which are also being considered as pneumococcal vaccine candidates (42). The presence of homologues to the most significant pneumococcal virulence genes is shown in Fig. 2, while a full spreadsheet, with hybridization results from all 7 isolates aligned to the pneumococcal gene list, can be found in Fig. S1 in the supplemental material. A number of capsular genes were present in the *S. oralis* (Sv29) and *S. mitis* (R77 C4) isolates. The capsular genes were not those of any known pneumococcal serotype, and the strains did not react with pneumococcal typing serum (data not shown). Conversely, 13/231 essential virulence genes did not hybridize with any of the mitis group isolates tested (Table 3). It is possible that these genes are unique to the pneumococcus but also that high levels of variation prevented the hybridization of these genes to the corresponding microarray probes. Many of the genes were hypothetical in function, but included were the transcriptional regulator gene *rlrA*, present in the RlrA islet encoding the pneumococcal pilus

Virulence gene	Presence in						
	NCTC12261	R75 A1	R77 C4	Col	990123	Sv29	R98 A5
	<i>S. mitis</i>	<i>S. mitis</i>	<i>S. mitis</i>	<i>S. mitis</i>	<i>S. mitis</i>	<i>S. oralis</i>	<i>S. pseudopneumoniae</i>
<i>sodA</i>							
<i>ply</i>							
<i>nanA</i>							
<i>nanB</i>							
<i>pspA</i>							
<i>psaA</i>							
<i>pspC</i>							
<i>phtA</i>							
<i>phtB</i>							
<i>phtD</i>							
<i>lytA</i>							
<i>lytC</i>							
<i>cbpA</i>							
<i>cbpC</i>							
<i>cbpD</i>							
<i>pppA</i>							
<i>ppmA</i>							
<i>slrA</i>							
<i>piaA</i>							
<i>piuA</i>							
<i>rlrA</i>							
<i>cpsA</i>							
<i>cpsB</i>							
<i>cpsC</i>							
<i>cpsD</i>							
<i>psrP</i>							
<i>licA</i>							
<i>licB</i>							
<i>licC</i>							
<i>licD1</i>							
<i>licD2</i>							
<i>tacF</i>							
<i>ciaR</i>							
<i>ciaH</i>							
<i>ccpA</i>							
<i>codY</i>							
<i>endA</i>							
<i>comD</i>							
<i>comE</i>							
<i>pbp2a</i>							
<i>pbp2x</i>							
SP1321							
SP1326							
SP1328							
SP2159							
SP2167							

FIG. 2. Presence of pneumococcal virulence gene homologues in related VGS species by comparative genomic hybridization. The presence of a homologue by hybridization is represented by a black box, and its absence is shown by a white box. The final five genes shown are essential virulence genes of unknown function present on identified regions of pneumococcal diversity (33).

TABLE 3. List of pneumococcal virulence genes not identified in any of the VGS isolates tested, with names and functions where known

Gene no.	Name	Role/function
SP0461	<i>rlrA</i>	Positive regulator of pneumococcal pilus expression (3)
SP1143		Hypothetical protein
SP1321	<i>ntpK</i>	V-type sodium ATP synthase, subunit K
SP1328		Sodium:solute symporter family protein
SP1343		Prolyl oligopeptidase family protein
SP1344		Conserved hypothetical protein
SP1433		AraC family transcriptional regulator
SP1434		Permease/ATP-binding protein of ABC transporter
SP1760		Hypothetical protein
SP1770		Glycosyl transferase, family 8
SP1771		Glycosyl transferase, family 2/glycosyl transferase family 8
SP1772	<i>psrP</i>	Cell wall surface anchor family protein
SP1952		Hypothetical protein

(3), and genes in the *PsrP* islet (33). Both of these islets were fully absent from all VGS isolates tested.

Pneumolysin homologues in streptococci of the mitis group.

Previous work characterized a Ply homologue termed mitilysin (Mly) in a subset of *S. mitis* isolates, with the protein sequences of 7 alleles reported (20). We report the characterization of a further related toxin in *S. pseudopneumoniae* isolates, referred to as pseudopneumolysin (pPly). A collection of 14 *S. pseudopneumoniae* isolates were screened for the presence of pPly by PCR, with 13 positive isolates (data not shown). Of these, three were chosen for further study (Table 1). Full *pply* gene sequencing of 3 isolates showed that compared to wild-type *ply* allele 1 from isolate D39 (21, 45), mutations were present at 74 positions. Translation of these sequences resulted in only 7 amino acid changes, with many synonymous mutations present, and 3 unique pPly alleles. Of the 7 amino acid changes, 2 were shared by both Ply and Mly (N14D and K224R), while another

FIG. 5. Minimum evolution tree of *nanA* DNA alleles and homologues in related streptococcal species. Pneumococcal *nanA* alleles include those from sequenced pneumococcal strains. The tree was drawn using MEGA4 software and demonstrates the relatedness of these enzymes at the DNA level.

ing a transcriptional regulator of the AraC family, a sodium symporter, and an ABC transporter, as well as a number of essential genes of unknown function. However, many of the proteins in these regions remain hypothetical, and further study may uncover genes of interest as pneumococcal vaccine candidates if they are found to be fully absent from mitis group isolates.

There was increased hybridization between the *S. oralis* isolate and the *S. pneumoniae* array compared to that for the *S. mitis* strains. This was unexpected but may have been due to the fact that the *S. oralis* isolate was chosen for study as an

isolate of interest because it was seen to possess at least some pneumococcal capsular genes (data not shown). Therefore, it is possible that the *S. oralis* isolate used in this study is not representative of the *S. oralis* population, but the result shows that this level of similarity is possible between strains distinguished as *S. pneumoniae* and *S. oralis*. Hybridization of further *S. oralis* strains would shed light on this subject.

There are a number of drawbacks to the use of microarray technology for comparative genomic hybridization, and these must be appreciated in this study. First, only information on the presence or absence of genes present in the reference isolate (in this case, *S. pneumoniae* TIGR4) is available. Furthermore, nonhybridization of a particular gene does not confirm its absence from the test genome, as it may be a result of diversity in the target region of the probe. This is particularly important for comparative genomic hybridization across species, where levels of diversity are likely to be higher. As a result, further study is required to confirm whether nonhybridized genes in this study are indeed absent or simply display increased diversity in the tested isolates. For example, this is likely to be the case for the IgA proteases, which have been shown to display high levels of variation between mitis group isolates (37) and did not hybridize to the pneumococcal probes for any isolates. However, this does allow a certain degree of certitude for positive results, which must share at least 80% sequence identity with the array probes.

Two recent publications have also begun to address the genetic relationships between *S. pneumoniae* and related VGS species. One used microarrays generated from a sequenced *S. mitis* isolate to hybridize DNAs from a number of related species (10). The second study was a study of population genetics further exploring the evolutionary relationships between these species (24). Our results are in agreement with both studies, showing that there is far greater diversity in the *S. mitis* species than in *S. pneumoniae* and that many pneumococcal virulence factors are present in this population. However, there are a number of contrasting results presented. Denapait and colleagues were unable to identify virulence factors such as *pspA* and *pspC* in the *S. mitis* population (10). This may be due to the diverse nature of these genes (5, 18) or simply to the variation in virulence gene content in the *S. mitis* population. Furthermore, Kilian and colleagues were able to identify capsular genes in the sequence of *S. mitis* strain NCTC 12261, while these were not identified in our microarray study (24). This may result from increased diversity in this region compared to that in *S. pneumoniae*, as a sequence identity of <80% would result in nonhybridization to the probe under our conditions.

The presence of at least a partial capsule locus in two isolates tested here, an *S. mitis* and an *S. oralis* isolate, demonstrates that remnants of capsule loci may be present in other VGS species. This is supported by the findings of Kilian and colleagues, who were able to identify a capsule locus in an *S. mitis* strain, as discussed previously (24). It would therefore be very relevant to carry out a similar study of capsule locus presence/diversity and expression in related VGS species compared to the pneumococcus, whose capsule locus is very well characterized. This study could provide further important information in relation to the use of the current pneumococcal vaccine, which targets the polysaccharide capsule, as the pres-

ence and expression of capsular genes in related species could allow vaccine escape to occur.

The results of the comparative genomic hybridization analysis allowed selection of homologues of pneumococcal vaccine candidates for further study in order to assess the diversity present in and genetic relationships between these genes. Recently, we identified and characterized a Ply homologue, termed mitilysin (20), in a subset of *S. mitis* isolates. Our results identify and further characterize a Ply homologue in *S. pseudopneumoniae* which is more closely related to *mly* than to *ply*. This segregation is in agreement with phenotypic traits such as bile solubility, lack of capsule, and optochin resistance, which all link *S. pseudopneumoniae* to *S. mitis* (2). *ply* is one of the genes used routinely to identify *S. pneumoniae* by PCR-based detection in clinical samples and, along with *lytA*, is considered one of the best targets to distinguish the pneumococcus (8). These genes have been shown to be well conserved in the pneumococcus (21, 47), but the identification of relatively similar homologues in related VGS species in this and other studies shows that these genes may not be ideal candidates for identifying the pneumococcus by PCR (10, 24, 39).

In comparing the sequences of *nanA* gene homologues identified in *S. mitis* and *S. pseudopneumoniae* isolates to pneumococcal *nanA* alleles, there was no species segregation, in contrast to the case for the *ply* alleles. These results further the argument that a smooth continuum may exist between these species, with pneumococcal virulence genes and homologues dispersed in isolates identified as *S. mitis* or *S. pseudopneumoniae* (14). Furthermore, the fact that *nanA* alleles are distributed in a non-species-specific manner may indicate that there has been genetic exchange between the species at this locus, which would have occurred after the species diverged from a common ancestor. Indeed, genetic exchange at the *nanA* locus in the pneumococcus has been documented between *S. pneumoniae* and *S. oralis*, in agreement with this hypothesis (25).

Our results demonstrate the presence of a NanA homologue in a subset of *S. mitis* and *S. pseudopneumoniae* isolates. However, sialidase genes have previously been found in *S. oralis*, a close relation of the pneumococcus (7, 25). The sequences of the protein alleles reported in our work were very similar to that of pneumococcal NanA, and this may indicate that tight control of the NanA structure is more important for biological activity than that of Ply and the Mly homologue. NanA has a highly defined role in pneumococcal colonization (29, 43, 44), and the presence of NanA in commensal relatives probably highlights its role in adherence and colonization of mucosal surfaces as well. In contrast, Ply has a much less clear role in pneumococcal colonization but has well-defined roles in invasive diseases such as pneumonia and meningitis (17, 22, 40, 46). Therefore, it may be expected that NanA would be important for the colonization of commensal organisms, whereas the role of a cytolytic toxin remains unclear. However, the presence of these homologues in the mitis group and their possible transfer into the pneumococcal population under the selective pressure of a vaccine demonstrate that the study of virulence factor homologues in related species is important in the context of selecting protein vaccine candidates. Overall, the diversity present in homologues of Ply and NanA remains relatively low, implying that the presence of these homologues in related

species may not pose a problem if Ply or NanA is targeted with a pneumococcal vaccine. Further study, including study of further mitis group streptococcal isolates for diversity analysis and antigenic studies, would help to confirm this conclusion.

In conclusion, genomic microarray analysis has demonstrated the presence of a large number of important pneumococcal virulence genes in the commensal mitis group streptococcal population. Specific virulence factors such as RlrA and PsrP appeared absent from related species and may provide a useful starting point from which to explore mechanisms explaining the different biological behaviors between these groups of organisms. This approach may also inform future vaccine strategies for selection of targets less likely to be compromised by horizontal recombination between species.

ACKNOWLEDGMENTS

We are grateful to Julie Galbraith at the Molecular Biology Sequencing Unit at the University of Glasgow for sequence generation and advice. We thank the Scottish Haemophilus, Legionella, Meningococcus and Pneumococcus Reference Laboratory (SHLMPL) for supply of strains and for performing serotyping and multilocus sequence typing.

C.J. was funded by an MRC studentship.

REFERENCES

- Aguiar, S. I., I. Serrano, F. R. Pinto, J. Melo-Cristino, and M. Ramirez. 2008. The presence of the pilus locus is a clonal property among pneumococcal invasive isolates. *BMC Microbiol.* **8**:41.
- Arbique, J. C., C. Poyart, P. Trieu-Cuot, et al. 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.
- Barocchi, M. A., J. Ries, X. Zogaj, et al. 2006. A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. U. S. A.* **103**:2857–2862.
- Beighton, D., A. D. Carr, and B. A. Oppenheim. 1994. Identification of viridans streptococci associated with bacteraemia in neutropenic cancer patients. *J. Med. Microbiol.* **40**:202–204.
- Brooks-Walter, A., D. Briles, and S. Hollingshead. 1999. The *pspC* gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infect. Immun.* **67**:6533–6542.
- Brueggemann, A. B., R. Pai, D. W. Crook, and B. Beall. 2007. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLoS Pathog.* **3**:e168.
- Byers, H. L., E. Tarelli, K. A. Homer, and D. Beighton. 2000. Isolation and characterisation of sialidase from a strain of *Streptococcus oralis*. *J. Med. Microbiol.* **49**:235–244.
- Carvalho, M. G., M. L. Tondella, K. McCaustland, L. Weidlich, L. McGee, L. W. Mayer, A. Steigerwalt, M. Whaley, R. R. Facklam, B. Fields, G. Carlone, E. W. Ades, R. Dagan, and J. S. Sampson. 2007. Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. *J. Clin. Microbiol.* **45**:2460–2466.
- Chi, F., O. Nolte, C. Bergmann, M. Ip, and R. Hakenbeck. 2007. Crossing the barrier: evolution and spread of a major class of mosaic *phb2x* in *Streptococcus pneumoniae*, *S. mitis* and *S. oralis*. *Int. J. Med. Microbiol.* **297**:503–512.
- Denapate, D., R. Bruckner, M. Nuhn, P. Reichmann, B. Henrich, P. Maurer, Y. Schahle, P. Selbmann, W. Zimmermann, R. Wambutt, and R. Hakenbeck. 2010. The genome of *Streptococcus mitis* B6—what is a commensal? *PLoS One* **5**:e9426.
- Douglas, C. W., J. Heath, K. K. Hampton, and F. E. Preston. 1993. Identity of viridans streptococci isolated from cases of infective endocarditis. *J. Med. Microbiol.* **39**:179–182.
- Dowson, C. G., T. J. Coffey, and B. G. Spratt. 1994. Origin and molecular epidemiology of penicillin-binding-protein-mediated resistance to beta-lactam antibiotics. *Trends Microbiol.* **2**:361–366.
- Facklam, R. 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. *Clin. Microbiol. Rev.* **15**:613–630.
- Hakenbeck, R., N. Balmelle, B. Weber, C. Gardes, W. Keck, and A. de Saizieu. 2001. Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect. Immun.* **69**:2477–2486.
- Hava, D. L., and A. Camilli. 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* **45**:1389–1406.

16. Havarstein, L. S., P. Gaustad, I. F. Nes, and D. A. Morrison. 1996. Identification of the streptococcal competence-pheromone receptor. *Mol. Microbiol.* **21**:863–869.
17. Hirst, R. A., B. Gosai, A. Rutman, C. J. Guerin, P. Nicotera, P. W. Andrew, and C. O'Callaghan. 2008. *Streptococcus pneumoniae* deficient in pneumolysin or autolysin has reduced virulence in meningitis. *J. Infect. Dis.* **197**:744–751.
18. Hollingshead, S., R. Becker, and D. Briles. 2000. Diversity of PspA: mosaic genes and evidence for past recombination in *Streptococcus pneumoniae*. *Infect. Immun.* **68**:5889–5900.
19. Hoshino, T., T. Fujiwara, and M. Kilian. 2005. Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients. *J. Clin. Microbiol.* **43**:6073–6085.
20. Jefferies, J., L. Nieminen, L. A. Kirkham, C. Johnston, A. Smith, and T. J. Mitchell. 2007. Identification of a secreted cholesterol-dependent cytolysin (mitilysin) from *Streptococcus mitis*. *J. Bacteriol.* **189**:627–632.
21. Jefferies, J. M., C. H. Johnston, L. A. Kirkham, et al. 2007. Presence of nonhemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks. *J. Infect. Dis.* **196**:936–944.
22. Jounblat, R., A. Kadioglu, T. J. Mitchell, and P. W. Andrew. 2003. Pneumococcal behavior and host responses during bronchopneumonia are affected differently by the cytolytic and complement-activating activities of pneumolysin. *Infect. Immun.* **71**:1813–1819.
23. Kawamura, Y., X. G. Hou, F. Sultana, H. Miura, and T. Ezaki. 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.
24. Kilian, M., K. Poulsen, T. Blomqvist, L. S. Havarstein, M. Bek-Thomsen, H. Tettelin, and U. B. Sorensen. 2008. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS One* **3**:e2683.
25. King, S. J., A. M. Whatmore, and C. G. Dowson. 2005. NanA, a neuraminidase from *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part through recombination with *Streptococcus oralis*. *J. Bacteriol.* **187**:5376–5386.
26. Kirkham, L. A., J. M. Jefferies, A. R. Kerr, Y. Jing, S. C. Clarke, A. Smith, and T. J. Mitchell. 2006. Identification of invasive serotype 1 pneumococcal isolates that express nonhemolytic pneumolysin. *J. Clin. Microbiol.* **44**:151–159.
27. Reference deleted.
28. Lucas, V. S., D. Beighton, G. J. Roberts, and S. J. Challacombe. 1997. Changes in the oral streptococcal flora of children undergoing allogeneic bone marrow transplantation. *J. Infect.* **35**:135–141.
29. Manco, S., F. Hernon, H. Yesilkaya, J. C. Paton, P. W. Andrew, and A. Kadioglu. 2006. Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect. Immun.* **74**:4014–4020.
30. Morrison, D. A. 1997. Streptococcal competence for genetic transformation: regulation by peptide pheromones. *Microb. Drug Resist.* **3**:27–37.
31. Moschioni, M., C. Donati, A. Muzzi, et al. 2008. *Streptococcus pneumoniae* contains 3 *rlrA* pilus variants that are clonally related. *J. Infect. Dis.* **197**:888–896.
32. Neeleman, C., C. H. Klaassen, D. M. Klomberg, H. A. de Valk, and J. W. Mouton. 2004. Pneumolysin is a key factor in misidentification of macrolide-resistant *Streptococcus pneumoniae* and is a putative virulence factor of *S. mitis* and other streptococci. *J. Clin. Microbiol.* **42**:4355–4357.
33. Obert, C., J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. I. Tuomanen, and C. J. Orihuela. 2006. Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. *Infect. Immun.* **74**:4766–4777.
34. Reference deleted.
35. Reference deleted.
36. Pai, R., M. R. Moore, T. Pilishvili, R. E. Gertz, C. G. Whitney, and B. Beall. 2005. Postvaccine genetic structure of *Streptococcus pneumoniae* serotype 19A from children in the United States. *J. Infect. Dis.* **192**:1988–1995.
37. Poulsen, K., J. Reinholdt, C. Jespersgaard, K. Boye, T. A. Brown, M. Hauge, and M. Kilian. 1998. A comprehensive genetic study of streptococcal immunoglobulin A1 proteases: evidence for recombination within and between species. *Infect. Immun.* **66**:181–190.
38. Reichmann, P., A. Konig, J. Linares, F. Alcaide, F. C. Tenover, L. McDougal, S. Swidsinski, and R. Hakenbeck. 1997. A global gene pool for high-level cephalosporin resistance in commensal *Streptococcus* species and *Streptococcus pneumoniae*. *J. Infect. Dis.* **176**:1001–1012.
39. Romero, P., R. Lopez, and E. Garcia. 2004. Characterization of LytA-like N-acetylmuramoyl-L-alanine amidases from two new *Streptococcus mitis* bacteriophages provides insights into the properties of the major pneumococcal autolysin. *J. Bacteriol.* **186**:8229–8239.
40. Rubins, J. B., D. Charboneau, J. C. Paton, T. J. Mitchell, P. W. Andrew, and E. N. Janoff. 1995. Dual function of pneumolysin in the early pathogenesis of murine pneumococcal pneumonia. *J. Clin. Invest.* **95**:142–150.
41. Sibold, C., J. Henriksen, A. Konig, C. Martin, L. Chalkley, and R. Hakenbeck. 1994. Mosaic *pbpX* genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from *pbpX* genes of a penicillin-sensitive *Streptococcus oralis*. *Mol. Microbiol.* **12**:1013–1023.
42. Tai, S. S. 2006. *Streptococcus pneumoniae* protein vaccine candidates: properties, activities and animal studies. *Crit. Rev. Microbiol.* **32**:139–153.
43. Tong, H. H., M. A. McIver, L. M. Fisher, and T. F. DeMaria. 1999. Effect of lacto-N-neotetraose, asialoganglioside-GM1 and neuraminidase on adherence of otitis media-associated serotypes of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Microb. Pathog.* **26**:111–119.
44. Tong, H. H., L. E. Blue, M. A. James, and T. F. DeMaria. 2000. Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chinchilla model. *Infect. Immun.* **68**:921–924.
45. Walker, J. A., R. L. Allen, P. Falmagne, M. K. Johnson, and G. J. Boulnois. 1987. Molecular cloning, characterization, and complete nucleotide sequence of the gene for pneumolysin, the sulfhydryl-activated toxin of *Streptococcus pneumoniae*. *Infect. Immun.* **55**:1184–1189.
46. Wellmer, A., G. Zysk, J. Gerber, T. Kunst, M. Von Mering, S. Bunkowski, H. Eiffert, and Nau. 2002. Decreased virulence of a pneumolysin-deficient strain of *Streptococcus pneumoniae* in murine meningitis. *Infect. Immun.* **70**:6504–6508.
47. Whatmore, A. M., and C. G. Dowson. 1999. The autolysin-encoding gene (*lytA*) of *Streptococcus pneumoniae* displays restricted allelic variation despite localized recombination events with genes of pneumococcal bacteriophage encoding cell wall lytic enzymes. *Infect. Immun.* **67**:4551–4556.
48. Whatmore, A. M., A. Efstratiou, A. P. Pickerill, K. Broughton, G. Woodard, D. Sturgeon, R. George, and C. G. Dowson. 2000. Genetic relationships between clinical isolates of *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*: characterization of “atypical” pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes. *Infect. Immun.* **68**:1374–1382.