**Abstract**

**Purpose.** *Streptococcus oralis* and *Streptococcus mitis* belong to the Mitis group, which are mostly commensals in the human oral cavity. Even though *S. oralis* and *S. mitis* are oral commensals, they can be opportunistic pathogens causing infective endocarditis. A recent taxonomic re-evaluation of the Mitis group has embedded the species *Streptococcus tigurinus* and *Streptococcus dentisani* into the species *S. oralis* as subspecies. In this study, the distribution of virulence factors that contribute to bacterial immune evasion, colonization and adhesion was assessed in clinical strains of *S. oralis* (subsp. oralis, subsp. tigurinus and subsp. dentisani) and *S. mitis*.

**Methodology.** Forty clinical *S. oralis* (subsp. oralis, subsp. dentisani and subsp. tigurinus) and *S. mitis* genomes were annotated with the pipeline PanFunPro and aligned against the VFDB database for assessment of virulence factors.

**Results/Key findings.** Three homologues of pavA, psaA and lmb, encoding adhesion proteins, were present in all strains. Seven homologues of nanA, nanB, ply, lytA, lytB, lytC and iga, of importance regarding survival in blood and modulation of the human immune system, were variously present in the genomes. Few *S. oralis* subspecies specific differences were observed. iga homologues were identified in *S. oralis* subsp. oralis, whereas lytA homologues were identified in *S. oralis* subsp. oralis and subsp. tigurinus.

**Conclusion.** Differences in the presence of virulence factors among the three *S. oralis* subspecies were observed. The virulence gene profiles of the 40 *S. mitis* and *S. oralis* (subsp. oralis, subsp. dentisani and subsp. tigurinus) contribute with important new knowledge regarding these species and new subspecies.

---

**INTRODUCTION**

*Streptococcus oralis* and *Streptococcus mitis* are non-haemolytic streptococci belonging to the Mitis group, which mostly are commensals in the human oral cavity throughout life [1, 2]. Even though *S. oralis* and *S. mitis* are oral commensals, they can be opportunistic pathogens entering the bloodstream and causing infective endocarditis (IE) [3, 4]. *Streptococcus tigurinus* and *Streptococcus dentisani* are other members of the Mitis group that have likewise been isolated from the oral cavity [5, 6]. *S. tigurinus* has been described as an agent causing IE [7]. A recent taxonomic re-evaluation of the Mitis group has embedded the two newer species...
S. pneumoniae, another member of the Mitis group, is the closest relative to S. oralis and S. mitis. Besides colonizing the human nasopharynx, S. pneumoniae also causes local infections and serious life-threatening diseases, such as septicaemia, meningitis, pneumonia and, more rarely, IE [9–11]. Virulence genes contributing to colonization (e.g. nanA, nanB, lytA, lytB, lytC and ply), to evasion of the immune system (e.g. iga, cps) and to adhesion (e.g. psaA and pavA) have been discovered in S. pneumoniae [12–20]. In addition, many of these genes have been identified in S. mitis and S. oralis.

The immunoglobulin A1 (IgA1) protease has been observed in both S. oralis and S. mitis, though variously present in both species [8, 21]. The gene encoding the pneumococcal surface adhesion A (psaA) has been identified in all investigated S. mitis and S. oralis [22, 23], and horizontal psaA gene transfer has been suggested among species in the Mitis group [23]. The genes ply and lytA have both been recognized in the genomes of a minority of S. mitis genomes, but not in those of S. oralis [24, 25]. In contrast, both S. mitis and S. oralis exhibit neuraminidase activity when grown in brain–heart infusion broth [26]. A widespread presence of the gene pavA was observed in a study where all nine S. mitis and 11 S. oralis strains included hybridized with pavA, illustrating the importance of adherence and virulence protein A (PavA) for oral streptococci [25].

Studies of virulence factors in clinical strains of S. mitis and S. oralis subsp. oralis, subsp. tigurinus and subsp. dentisani have been limited. We previously sequenced the whole genome and identified 40 S. mitis and S. oralis isolated from patients with IE [27]. In this study, we identified virulence factors in these S. mitis and S. oralis genomes in order to identify the distribution of virulence genes of importance in regard to immune evasion, colonization and adhesion in S. mitis, S. oralis subsp. oralis, S. oralis subsp. dentisani and S. oralis subsp. tigurinus.

**METHODS**

**Bacterial strains**

Forty blood culture strains, S. mitis (n=12), S. oralis subsp. oralis (n=14), S. oralis subsp. tigurinus (n=8) and S. oralis subsp. dentisani (n=6) from patients with verified IE were collected retrospectively (2006–2013) from the Capital region of Denmark (RH strains), region Zealand (AE, Y and B strains) and the region of Southern Denmark (OD strains). One strain per patient was included in the study, except for one patient who contributed with two strains (B007274_11 and Y11577_11). The verification of IE was conducted by cardiologists and microbiologists according to the modified Duke criteria [28]. The 40 strains were paired-end sequenced with 100× coverage using Illumina HiSeq 2000 (BGI-Tech Solutions, Hong Kong, China) [27]. The draft genomes were de novo assembled with SPAdes [29]. The species identification was based on Multi Locus Sequence Analysis (MLSA) and core-genome phylogeny [8, 27]. The GenBank accession numbers for the 40 genomes are available through the Bioproject accession number PRJNA304678.

**Genome annotation**

The pipeline PAN-genome analysis based on FUNctional PROfiles (PanFunPro) [30] was used for prediction of both genes and functional domains in the de novo assembled genomes. Firstly, genes were predicted and translated into protein sequences using Prodigal v2.50 [31]. The translated protein sequences for each streptococcal genome were searched against the databases PfamA [32], TIGRFAM [33] and SUPERFAMILY [34], using InterProScan software [35] for prediction of functional domains. The combination of non-overlapping functional domains in the protein sequences constituted the functional profiles. Each functional profile was based on a coding sequence.

**Hierarchical clustering of species**

A presence–absence gene matrix based on the pan-genome of 40 clinical S. mitis and S. oralis strains was constructed in order to gain an impression of co-existing genes among the strains examined from the two species. The matrix was constructed using PanGenome2Abundance.pl in PanFunPro [30].

The Pearson correlation coefficient between the 40 strains, using their presence–absence functional profiles, was the basis for hierarchical clustering of the strains.

**Prediction of putative virulence genes**

Basic Local Alignment Search Tool (BLASTP) [36] was applied to search the translated protein sequences against Virulence Factors of Pathogenic Bacteria Database (VFDB) (accessed 25 August 2015), which contains various virulence factors from other streptococci, *Staphylococcus aureus* and *Enterococcus faecalis* [37–39]. The threshold for hits was an e-value <0.001, a bit score >50 and a sequence identity percentage >40%. The best hit was based on the highest bit score.

**RESULTS**

**Whole-genome sequence characterization**

The number of scaffolds from the de novo assembly ranged from 17 to 85 (S. mitis), 20–41 (S. oralis subsp. dentisani), 7–47 (S. oralis subsp. oralis) and 7–47 (S. oralis subsp. tigurinus). The estimated size of the S. mitis and S. oralis genomes ranged from 1.8–2.1 Mb. Each functional profile was considered based on a coding sequence. Between 1692 and 2083 functional profiles were predicted in the 12 S. mitis strains and 1734–2035 functional profiles were predicted in the 28 S. oralis strains. There were no subspecies-specific differences between the number of functional profiles in the 28 S. oralis strains. The GC content was slightly higher in
Number of functional profiles in the 40 strains. The heat map colours indicate the Pearson correlation coefficient among the strains: The darker the colour, the higher the correlation. The colour bars show the individual species of each particular strain: S. oralis subsp. oralis (dark blue), S. oralis subsp. tigurinus (light blue), S. oralis subsp. dentisani (green) and S. mitis (red).

**Fig. 1.** Hierarchical clustering of Pearson correlation coefficients determined from the presence/absence of functional profiles in the 40 strains.

**S. oralis** (40.75–41.50 %) than in **S. mitis** (39.71–40.28 %). Number of scaffolds, N50, the longest sequences and the number of functional profiles in the 40 **S. mitis** and **S. oralis** genomes are presented in the online Supplementary Material.

When clustering the strains based on the presence/absence of the functional profiles, a tight cluster containing **S. mitis** was identified (Fig. 1). The **S. oralis** strains clustered into three sub-clusters, which were congruent with earlier observed sub-clusters based on core-gene phylogeny [27]. Furthermore, the sub-clustering of **S. oralis** was congruent with the division of the strains into the three subspecies, **S. oralis** subsp. oralis, subsp. tigurinus and subsp. dentisani [8].

Two **S. oralis** strains (**S. oralis** B007274_11 and **S. oralis** Y11577_11) with high correlation were isolated from the same patient over a 24 h period, and should be considered as the same strain.

**Virulence genes present in S. mitis and S. oralis subsp. oralis, subsp. tigurinus and subsp. dentisani.**

In order to determine the presence of virulence genes in **S. mitis** and **S. oralis** subsp. oralis, subsp. tigurinus and subsp. dentisani, the functional profiles based on coding sequences in the 40 strains were aligned against the VFDB database. The number of strains that contained the putative virulence genes and the protein sequence identity to the VFDB reference sequence are specified in Table 1. Genes encoding proteins homologous to Adherence and virulence protein A (PavA), Laminin-binding protein (Lmb) and Pneumococcal surface adhesion A (PsaA) were identified in all 40 strains.

Homologues of the seven genes nanA, nanB, ply, lytA, lytB, lytC, and iga that have been associated with bacterial survival in blood and immune evasion were variously present in the genomes [12, 16, 17, 24]. Both nanA and nanB gene
homologues were identified in *S. mitis* RH50275_09 and *S. mitis* RH50738_11; these were the only strains containing both neuraminidase genes. The *nanA* and *nanB* homologues were neighbours. None of the *S. mitis* strains contained *lytA* and *ply* gene homologues simultaneously. *iga* homologues were identified in all 14 *S. oralis* subspp. *oralis*, whereas *lytA* homologues only were identified in *S. oralis* subspp. *oralis* and subspp. *tigurinus*.

Poly saccharide capsule production (CPS) has been described as being important in regard to bacterial avoidance of phagocytosis [19, 40]. Genes encoding homologues of *cps4* from *S. pneumoniae* TIGR4 were identified in both *S. mitis* and *S. oralis*. *cps4A* gene homologues were present in all 40 strains whereas genes homologous to *cps4B*, *cps4C* and *cps4D* were variously present in the genomes. Eight *S. mitis* strains and 22 *S. oralis* strains contained homologues of the four capsular genes *cps4A*, *cps4B*, *cps4C* and *cps4D*. Furthermore, 22 *S. oralis* strains and one *S. mitis* strain contained a gene homologous to *cps4I*. One *S. oralis* subspp. *dentisani* strain, RH9883_08, contained genes homologous to *cps4E*, *cps4F*, *cps4J*, *cps4K* and *cps4L*.

In summary, three genes homologous to the adhesion genes *psaA*, *lmb* and *pavA* were identified in all 40 strains. The presence of the seven putative virulence genes (homologues of *nanA*, *nanB*, *ply*, *lytA*, *lytB*, *lytC* and *iga*), important for immune evasion and colonization in the 40 *S. mitis* and *S. oralis* genomes, were not coherent. A few *S. oralis* subspecies-specific differences were observed. All 14 *S. oralis* subspp. *oralis* contained an *iga* homologue, whereas homologues of *lytA* only were identified in *S. oralis* subspp. *oralis* and *S. oralis* subspp. *tigurinus*. Homologues of *nanB* and *ply* were only identified in *S. mitis*. Furthermore, homologues of the *cps4* genes were identified variously in *S. oralis* and *S. mitis* strains, but none of the strains included a full capsular locus compared to the VFDB reference *S. pneumoniae* TIGR4 genome.

**DISCUSSION**

Assessment of virulence factors in clinically important *S. mitis* and *S. oralis* subspp. *oralis*, subspp. *tigurinus* and subspp. *dentisani* has not been widely investigated.

In the present study, functional profiles were extracted from 40 IE clinical strains of *S. mitis* and *S. oralis* subspp. *oralis*, subspp. *tigurinus* and subspp. *dentisani* using the pipeline PanFunPro [30]. We previously used PanFunPro for extraction of a Mitis group streptococcal core-genome for evaluation of core-genome phylogeny [27]. The core-genome phylogeny revealed a sub-clustering of *S. oralis* into three subclusters [27]. Sub-clustering of *S. oralis* was later illustrated by Jensen *et al.* [8] using core-genome phylogeny, and it was proposed that the species *S. tigurinus* and *S. dentisani* should be reassigned as subspecies in *S. oralis*. Core-genome phylogeny was the basis for identification of clinical IE strains in the present study and, Fig. 1 clearly illustrates clustering of the *S. oralis* strains into the three subspecies.

The clustering of the three *S. oralis* subspecies strains in Fig. 1 based on the pan-genome indicates that differences may occur between subspecies other than in the core-genes. By using a sequence identity percentage >40% at protein level, few subspecies-specific differences in virulence factors were observed among the three subspecies *S. oralis* subspp. *oralis*, subspp. *tigurinus* and subspp. *dentisani*. The threshold at 40% sequence identity was based on findings in a study by Rost [41], who showed that 90% of the protein pairs were homologous when using a cut-off of roughly 30% sequence identity. Furthermore, 40% sequence identity has previously been used for protein identification in the Mitis group [42].

Alignment of the functional profiles against the VFDB database revealed that *iga* homologues were present in all 14 *S. oralis* subspp. *oralis* and in seven out of 12 *S. mitis*. The *iga* gene encoding IgA1 protease, which cleaves the human immunoglobulin A1 in the hinge region, has been variously

<table>
<thead>
<tr>
<th>Genes</th>
<th>Product</th>
<th><em>S. oralis</em> subspecies</th>
<th><em>S. mitis</em> subspecies</th>
<th><em>S. oralis</em> Identity %†</th>
<th><em>S. mitis</em> Identity %†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>oralis</em></td>
<td><em>tigurinus</em></td>
<td><em>dentisani</em></td>
<td><em>oralis</em></td>
</tr>
<tr>
<td><em>pavA</em></td>
<td>Adherence and virulence protein A</td>
<td>14/14</td>
<td>8/8</td>
<td>6/6</td>
<td>12/12</td>
</tr>
<tr>
<td><em>lmb</em></td>
<td>Laminin-binding surface protein</td>
<td>14/14</td>
<td>8/8</td>
<td>6/6</td>
<td>12/12</td>
</tr>
<tr>
<td><em>psaA</em></td>
<td>Pneumococcal surface adhesion A</td>
<td>14/14</td>
<td>8/8</td>
<td>6/6</td>
<td>12/12</td>
</tr>
<tr>
<td><em>nanA</em></td>
<td>Neuraminidase A</td>
<td>14/14</td>
<td>7/8</td>
<td>6/6</td>
<td>7/12</td>
</tr>
<tr>
<td><em>nanB</em></td>
<td>Neuraminidase B</td>
<td>0/14</td>
<td>0/8</td>
<td>0/6</td>
<td>6/12</td>
</tr>
<tr>
<td><em>ply</em></td>
<td>Pneumolysin</td>
<td>0/14</td>
<td>0/8</td>
<td>0/6</td>
<td>2/12</td>
</tr>
<tr>
<td><em>lytA</em></td>
<td>Autolysin</td>
<td>4/14</td>
<td>3/8</td>
<td>0/6</td>
<td>5/12</td>
</tr>
<tr>
<td><em>lytB</em></td>
<td>Cell Wall Hydrolase</td>
<td>14/14</td>
<td>8/8</td>
<td>6/6</td>
<td>11/12</td>
</tr>
<tr>
<td><em>lytC</em></td>
<td>Cell Wall Hydrolase</td>
<td>5/14</td>
<td>6/8</td>
<td>3/6</td>
<td>11/12</td>
</tr>
<tr>
<td><em>iga</em></td>
<td>IgA1 protease</td>
<td>14/14</td>
<td>0/8</td>
<td>0/6</td>
<td>7/12</td>
</tr>
</tbody>
</table>

*Number of strains in which the genes are present.
†Percentage of identical amino acids obtained using BLASTP.

---

Rasmussen *et al.*, *Journal of Medical Microbiology* 2017;66:1316–1323
identified in *S. mitis* and *S. oralis* strains [8, 21, 42, 43]. IgA1 is a predominant immunoglobulin present on the mucosal surfaces [44], and cleavage of this limits the host humoral response and thereby promotes colonization of *S. pneumoniae* [12]. Recently, Jensen et al. [8] described that iga is only present in *S. oralis* subsp. *oralis* and not in *S. oralis* subsp. *tigrinus* and subsp. *dentisani*, which is in accordance with the findings in the present study. These findings are further supported by Conrads et al., who used the former nomenclature and identified iga in *S. oralis* but not in *S. tigrinus* [45]. Another subspecies difference was observed among *S. oralis* subsp. *oralis*, subsp. *tigrinus* and subsp. *dentisani* in the present study (Table 1). Homologues of *lytA* were identified only in strains of *S. oralis* subsp. *oralis* and subsp. *tigrinus*. Conrads et al. did not include *S. dentisani* in their study, but they identified *lytA* in some *S. oralis* and *S. tigrinus* strains, congruent with the present results [45]. *lytA* encodes the autolytic cell wall hydrolase Autolysin (*LytA*), which appears to be necessary for the release of the cell cytoplasmic-located protein pneumolysin (*Ply*) [46]. Pneumolysin (*Ply*), encoded by the gene *ply*, is a pore-forming toxin that induces cell death by apoptosis. It is suggested to be an important factor in the initial establishment in nasal colonization and for development of septicaemia [13, 14, 47]. The two genes *lytA* and *ply* have been localized simultaneously in all analysed *S. pneumoniae* genomes [24, 42] and in *S. tigrinus AZ_3a* [45]. In contrast, *lytA* and *ply* have only been identified in three out of 31 *S. mitis* genomes [24] and in none of the examined *S. oralis* genomes [24, 42]. In the present study, only two *S. mitis* genomes contained genes homologous to *ply* and five genomes contained genes homologous to *lytA* (Table 1). *lytA* and *ply* homologues were not present simultaneously in any *S. mitis* strain, indicating that the presence and potential cooperation of *lytA* and *ply* is not a precondition for *S. mitis* virulence.

Other cell wall hydrolases, *LytB* and *LytC*, encoded by *lytB* and *lytC*, are important in the colonization of *S. pneumoniae* in the nasopharynx, and they contribute to bacterial avoidance of phagocytosis mediated by neutrophils and alveolar macrophages [16, 48]. In the present study, *lytB* homologues were identified in all 28 *S. oralis* strains whereas genes homologous to *lytC* were identified in 14 of the *S. oralis* strains distributed among all three subspecies (Table 1). In contrast, genes homologous to both *lytB* and *lytC* were identified in the majority (11 out of 12) of *S. mitis* strains. In strains where both genes were present, *lytB* and *lytC* homologues were located in different loci, indicating that these genes are not transcribed together.

Neuraminidases A and B (NanA and NanB), encoded by *nanA* and *nanB*, respectively, are enzymes that have also been claimed as being important for colonization, and both enzymes seem to be essential for survival in blood [17]. Intravenous infection with *nanA* and *nanB* mutants in mice revealed a progressive clearance of bacteria in blood within 48 h compared to wild types, which persisted for longer. In a previous study, *nanA* was identified using PCR in all strains of *S. oralis* (*n* = 23) and *S. mitis* (*n* = 10) [49], while only *nanB* was identified in strains of *S. mitis* by hybridization [25]. Genes homologous to *nanA* were identified in 27 strains of *S. oralis* and seven strains of *S. mitis* in the present study (Table 1). Genes homologous to *nanB* were only observed in six *S. mitis* strains in concordance with previous studies. Homologues of both *nanA* and *nanB* were only identified simultaneously in two *S. mitis* strains. In these strains, *nanA* and *nanB* homologues were neighbours, indicating that these two genes may belong to a *nanAB* locus which has been described in *S. pneumoniae* [50]. Furthermore, the dispersed presence of *nanA* and *nanB* in *S. mitis* and *S. oralis* indicates that these two genes are not essential for bacterial survival in blood.

Adhesion of bacterial cells to fibronectin may contribute to development of IE [51]. Fibronectin is an extracellular matrix protein secreted by a variety of cells and it is also present in saliva and blood [52, 53]. *S. pneumoniae* adheres to immobilized fibronectin by the fibronectin-binding surface protein PavA encoded by the gene *pavA*, and it was demonstrated that *pavA* mutants had reduced ability to adhere to human epithelial and endothelial cells [18, 54]. A study of cell surface proteins in *S. pneumoniae*, *S. mitis* and *S. oralis* showed that all 21 strains were hybridized with *pavA* using microarray [55], and in another study *pavA* was identified in all *S. tigrinus* strains [45]. *lmb* encoding the lipoprotein Lmb is another gene contributing to adhesion, described for *Streptococcus agalactiae* as a protein that mediates bacterial attachment to human laminin, promoting transfer of bacteria to the bloodstream and colonization of damaged epithelium [56]. The same study illustrated the presence of *lmb* in all 11 examined *S. agalactiae* serotypes, confirming the importance of this gene [56]. *psaA* encoding another lipoprotein, PsaA, also contributing to bacterial adhesion, was likewise identified in all serotypes of *S. pneumoniae* [20]. The virulence properties of *psaA* were described using in vitro studies where *psaA*+ mutants illustrated significantly less virulence compared to the wild type when inoculated intranasally and intraperitoneally in mice [57]. Although *S. pneumoniae* as *S. agalactiae* strains are associated with IE cases, they are mostly associated with non-IE infections [11, 58].

In our study, genes homologues to *pavA*, *lmb* and *psaA* were identified in all 40 strains and these genes have been proven important for bacterial adhesion [54, 56, 59]. The presence of these genes across different species could be the result of horizontal gene transfer, as previously suggested by Zhang et al. for *psaA* [23].

Capsular polysaccharides (CPS) are indispensable for the virulence of *S. pneumoniae* by forming an inert shield, which prevents phagocytosis [19, 40]. Currently, 97 serologically and structurally distinct CPS types have been recognized [60]. The encapsulated serotype 4, *S. pneumoniae* TG4 strain was used as reference in the present study to examine the presence of capsule loci in the 40
strains investigated. The cps locus in TIGR4 include the genes cps4A–cps4L [61]. A cps4A homologue was identified in all 40 clinical strains (Table 1). Only one S. oralis subsp. dentisani strain (RH9883_08) contained genes homologous to cps4E, cps4F, cps4J, cps4K and cps4L. Serotype switching between S. mitis strains and the strain S. pneumoniae TIGR4 has previously been reported [62], and this may also be possible for S. oralis subsp. dentisani. Skov et al. [63] identified complete cps loci in 74% of 66 investigated S. mitis strains and in 95% of 20 investigated S. oralis strains, including the subspecies tigurinus and dentisani. These authors confirmed capsule expression using antigenic analyses and demonstrated serological identities with different pneumococcal serotypes [63]. In the present study, eight S. mitis strains and 22 S. oralis strains contained genes homologous to cps4A, cps4B, cps4C and cps4D. cpsB–cpsD have been found essential for encapsulation in S. pneumoniae, whereas cpsA influenced the level of CPS produced [64]. The presence of cps4A, cps4B, cps4C and cps4D homologues in the eight S. mitis and 22 S. oralis strains indicates that these strains may be able to express capsule proteins. However, identification of capsular genes is not synonymous with capsule expression. Similar antigenic analyses, as conducted by Skov et al. [63], may possibly elucidate whether the IE strains in the present study express capsules.

The former species S. dentisani, now S. oralis subsp. dentisani, was originally isolated from the oral cavity [65]. A recent study conducted by López-López et al. confirmed this by identifying S. dentisani in metagenomic sequences from 118 healthy individuals [6]. Apart from the ability to colonize the oral cavity, these authors demonstrated that S. dentisani affects growth of the oral pathogens Streptococcus mutans, Streptococcus sobrinus and Prevotella intermedia, illustrating a probiotic feature of S. dentisani. Based on their findings, they proposed clinical trials to test the potential of S. dentisani in promoting human oral health [6]. In the present study, the isolation of six strains from IE patients clearly demonstrates that S. oralis subsp. dentisani is an agent responsible for IE. This new knowledge is important, as experimental inoculation of S. dentisani into the oral cavity of healthy humans may affect their likelihood of developing IE.

**Conclusion**

In the present study, we describe for the first time that S. oralis subsp. dentisani is able to cause infective IE. The hierarchical clustering based on the pan-genome illustrates clustering of the S. oralis strains into subsp. oralis, subsp. dentisani and subsp. tigurinus, indicating that differences may occur between the subspecies other than in the core-genomes.

Alignment of 40 clinical S. oralis (subsp. oralis, subsp. dentisani and subsp. tigurinus) and S. mitis genomes against the VFDB database revealed genes in the genomes homologous to virulence genes that contribute to bacterial avoidance of the immune system, colonization and adhesion. Three genes homologous to psaA, pavA and lmb that contribute to adhesion were identified in all strains. The presence of adhesion genes in all strains indicates the importance of adhesion properties in S. mitis and S. oralis. Seven genes (homologues of nanA, nanB, ply, lytA, lytB, lytC and iga) contributing to colonization and evasion of the immune system were variably identified in the strains. iga homologues were identified in S. mitis and in all 14 S. oralis subsp. oralis, whereas lytA homologues were identified in S. mitis, S. oralis subsp. oralis and S. oralis subsp. tigurinus, indicating subspecies-specific differences in S. oralis virulence. Genes homologous to the capsular genes cps4 in S. pneumoniae TIGR4 were variably identified in all 40 strains. However, none of the strains contained a full cps4 locus compared to S. pneumoniae TIGR4. The virulence gene profiles of the 40 clinical S. mitis and S. oralis (subsp. oralis, subsp. dentisani and subsp. tigurinus) contribute with important knowledge about the virulence of these species and new subspecies. However, further elucidation using expression and in vivo studies is necessary before the clinical relevance of these three new subspecies can be established.

**Funding information**

This work was supported by the Danish Heart Foundation (12–04.R90-A4026-22720 and 15-R99-A6040-22951), The A.P. Møller Foundation for the Advancement of Medical Science, The Foundation of Hans and Norah Buchard, The Foundation of Aase and Ejner Danielsen and The Region Zealand Foundation for Health Research.

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**Ethical statement**

Recognition of the streptococcal strains is part of the routine diagnostic protocol at the Departments of Clinical Microbiology in the Capital region of Denmark, region Zealand and the region of Southern Denmark. The strains were analysed anonymously in a retrospective manner, and ethical approval and informed consent were thus not required.

**References**


