Identification of *Streptococcus pneumoniae* and other Mitis streptococci: importance of molecular methods

Ewa Sadowy 1 · Waleria Hryniewicz 2

Received: 20 May 2020 / Accepted: 15 July 2020 / Published online: 24 July 2020
© The Author(s) 2020

Abstract

The Mitis group of streptococci includes an important human pathogen, *Streptococcus pneumoniae* (pneumococcus) and about 20 other related species with much lower pathogenicity. In clinical practice, some representatives of these species, especially *Streptococcus pseudopneumoniae* and *Streptococcus mitis*, are sometimes mistaken for *S. pneumoniae* based on the results of classical microbiological methods, such as optochin susceptibility and bile solubility. Several various molecular approaches that address the issue of correct identification of pneumococci and other Mitis streptococci have been proposed and are discussed in this review, including PCR- and gene sequencing-based tests as well as new developments in the genomic field that represents an important advance in our understanding of relationships within the Mitis group.

Keywords *Streptococcus* · PCR · Sequencing · MLSA · WGS

Introduction

The group of Mitis streptococci, included among Viridans streptococci (VGS) [1], includes currently about different 20 species [2, 3], which is found in humans and animals. Human Mitis streptococci significantly differ in their epidemiological importance. The best-known and most studied member of the group, *Streptococcus pneumoniae* (pneumococcus), commonly colonizes human nasopharynx but is also a very important etiologic agent of the respiratory tract and invasive infections. In the case of the community-acquired pneumonia, pneumococcus represents the most common pathogen and it was estimated that pneumococcal pneumonia caused 1.5 million deaths worldwide in 2015 including about 400,000 deaths in children younger than 5 years [4]. Other infections, especially pneumococcal bacteremia and meningitis, although less common, result in high mortality and long-term neurological sequelae in survivors. While improvements in treatment reduced the case-fatality ratio of pneumococcal meningitis, it still reached 25–27% in Europe and the Americas in 2015 and was as high as 61% in Africa in the same year [5]. Children and the elderly as well as immunocompromised patients, including HIV-infected people, constitute groups, which are most vulnerable to pneumococcal infections [6]. The development of resistance to commonly used antimicrobials, seen worldwide, complicates the treatment of pneumococcal infections [7]. After the introduction of conjugated vaccines into the mass vaccination of children a significant decrease in the incidence of invasive and respiratory infections caused by serotypes targeted by vaccines was observed [4]. However, vaccination also resulted in a serotype replacement, i.e. increased proportion of pneumococci with non-vaccine serotypes in carriage and disease [8], demonstrating fast adaptation of this pathogen to the selective pressure exerted by vaccines. In consequence, the reduced incidence of pneumococcal meningitis due to vaccination might be only temporary [9]. Other Mitis streptococci are typically non-pathogenic colonizers of the human nasopharynx, where they can co-reside with *S. pneumoniae*. However, severe invasive infections, such as infective endocarditis and bacteremia caused by these bacteria are occasionally observed, mostly in immunocompromised patients [10]. Infective endocarditis is a rare but severe disease, associated with high mortality reaching 40% [11]. A significant proportion of this disease is caused by the VGS [11–13]. Neutropenic cancer patients have an increased risk of developing VGS bacteremia, mostly...
associated with *Streptococcus mitis*, which in a significant number of cases results in serious complications, such as viridans streptococcal shock syndrome (VSSS), with mortality reaching 40–100% in paediatric patients [10, 13]. Treatment of these infections is additionally complicated by high rates of antimicrobial resistance among these bacteria [10].

Several studies reported that some representatives of the Mitis group may be mistakenly identified as *S. pneumoniae*, especially in samples obtained from non-sterile body sites [14–26]. Although such situations are relatively infrequent, it has to be stressed that distinguishing between *S. pneumoniae* and other Mitis streptococci is important for the understanding of the pathogenicity of particular species as well as the correct evaluation of disease burden and antimicrobial resistance levels in pneumococci and other Mitis streptococci. Such biases, in particular an overestimation of resistance among *S. pneumoniae* due to misidentification of the other Mitis streptococci as pneumococci, were reported, e.g. in the USA and Portugal [18, 21, 27]. The phenomenon of misidentification is due to both a common evolutionary origin of these organisms as well as a horizontal gene transfer (HGT) among streptococcal species, residing in the same ecological niche [28, 29]. The issue of correct identification was addressed by several molecular and other approaches and the purpose of this review is to provide an overview of methods proposed so far, their feasibility and performance in the identification of *S. pneumoniae* and other Mitis streptococci.

**Taxonomic position of Mitis streptococci**

Bacteria included in the genus *Streptococcus* are facultatively anaerobic, Gram-positive cocci. Streptococci demonstrate different types of haemolysis on blood agar, which led to the distinction of beta- and non-beta-haemolytic (alpha and gamma) strains [30]. Analyses of sequences of 16S rRNA genes split the VGS into five subgroups, including Anginosus, Mitis, Salivarius, Bovis (Equinus) and Mutans [31–33]. Later on, the Bovis group was separated from the VGS, and the Sanguinis group was added [1]. Recently, genome-based studies introduced yet another organization of the *Streptococcus* genus, encompassing two main clades, designated Mitis-Suis and Pyogenes-Equinus-Mutans and composed of 14 subclades. The Mitis clade included Anginosus, Pneumoniae, Parasanguinis and Gordonii subclades; the Pneumoniae subclade grouped *S. pneumoniae*, *Streptococcus pseudopneumoniae*, *S. mitis*, *Streptococcus oralis*, *Streptococcus infantis*, *Streptococcus porcinus* and *Streptococcus timonensis* [3]. Another study on Mitis streptococci, based on core genomic analyses combined with phenotypic characterization, led to the distinction of eight clades, with the Mitis clade encompassing clusters of *S. pneumoniae*, *S. pseudopneumoniae* and *S. mitis*, the latter one comprised of several divergent strains [2]. This brief overview demonstrates the complexity of taxonomic relationships among streptococci.

**Phenotypic methods and automated systems**

The preliminary identification of a streptococcal isolate is based on the morphology of colonies growing on blood agar. Pneumococci present typical colonies with a button or mucoid appearance [34]. However, a study performed on a large collection of pneumococcal clinical isolates from the respiratory tract showed that about 24% of isolates do not demonstrate this feature [18]. Classical microbiological identification of *S. pneumoniae* relies on susceptibility to optochin (ethylhydrocupreine hydrochloride) and solubility in sodium deoxycholate (bile) [35]. However, optochin susceptible and/or bile soluble isolates of other species from the Mitis group were reported in several studies [14–26]. The observed optochin susceptibility was shown to arise by a presumed transfer of *ap* genes from *S. pneumoniae* [17]. At the same time, a small number of *S. pneumoniae* isolates demonstrate optochin resistance and/or poor bile solubility [23, 36, 37]. Nevertheless, determining optochin susceptibility and bile solubility remains the key test for phenotypic identification of *S. pneumoniae* [38]. In the case of bile solubility, the use of standardized bacterial cultures and measurements of changes in absorbance determined spectrophotometrically rather than by a visual assessment was shown to improve the test performance [39].

Automated systems, such as API® rapid ID 32 Strep or the VITEK® 2, are of a limited utility for identification of Mitis streptococci [40, 41]. The matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) is an increasingly important technique for rapid identification of bacterial and fungal species [42] but the performance of this technique turned out to be problematic for some isolates belonging to the Mitis group [43, 44], although an update of reference database has recently improved the results of identification based on MALDI-TOF spectra [45].

**Immunological methods**

Polysaccharide capsule, specified by the *cps* locus, represents the major virulence factor of pneumococci [46]. Identification of capsule type, i.e. serotyping is an important step confirming the identification and characterization of a pneumococcal isolate [38]. Specific antisera, currently available for 92 pneumococcal serotypes from the Statens Serum Institute (SSI,
Copenhagen, Denmark; https://www.ssidiagnostica.com/neufeld-antisera/; 15th April 2020 date last accessed), are used in the Quellung reaction (named also Neufeld test or capsular swelling), which remains the “gold standard” for serotyping of pneumococci [38]. This test requires experienced personnel and a comprehensive panel of antisera, which in general limits its use to reference and research laboratories with a sufficiently large throughput of samples. Moreover, it was demonstrated that many representatives of S. mitis, S. oralis and S. infantis produce a capsular polysaccharide, which in some cases shows cross-reactivity in double immunodiffusion assays with antibodies recognizing several pneumococcal serotypes [47–49]. Importantly, however, tested Mitis group isolates producing capsular polysaccharide specific for serotypes 1 and 5 of S. pneumoniae reacted weakly or not at all in the classical Quellung test [48, 49], instead of demonstrating a distinct halo, typical for S. pneumoniae [14]. This difference is probably likely due to a much lower amount of polysaccharide produced by Mitis group streptococci other than S. pneumoniae [47]. Pneumotest-Latex (SSI) is a fast and easy to perform a test, however, also in this case, cross-reactions with S. mitis and S. oralis were occasionally observed for some reagent pools [50]. This test can also be applied directly to clinical samples [51]. Negative results of the Quellung test do not exclude that an isolate represents S. pneumoniae since non-serotypable (rough) pneumococci lacking capsule are observed and in fact, can be very common in some settings [52]. Such isolates may arise sporadically due to mutations in the cps locus or belong to certain lineages of S. pneumoniae devoid of the whole locus [53]. Moreover, new serotypes are constantly reported for S. pneumoniae, reaching at present 98 different ones [46]. Thus, currently available antisera may in some, presumably rare instances, yield ambiguous results in the case of pneumococci of yet uncharacterized serotypes. For more detailed information on immunological methods, the Reader may refer to the recent review on the identification of S. pneumoniae [54] that provides more detailed information on these issues.

**PCR-based methods of identification**

The development of molecular methods and increasing knowledge of S. pneumoniae virulence determinants together with a better understanding of genetic diversity of pneumococci and other Mitis streptococci led to proposing several targets for PCR and real-time PCR in order to improve species identification among Mitis streptococci, especially concerning their questionable representatives. As PCR-based approaches do not require cultured microorganisms, they find also an application in the direct detection of pneumococci in clinical materials [55, 56], including polymicrobial samples [57–60]. Thus, the right selection of appropriate target gene(s) is of great importance. Ideally, such a target should be exclusively and ubiquitously present among S. pneumoniae and absent or sufficiently different in other Mitis streptococci to allow the clear discrimination of these groups.

A number of proposed genetic targets included S. pneumoniae virulence-associated genes. The ply gene, specifying the extracellular toxin pneumolysin, was one of the first DNA targets proposed to be used for the detection of pneumococci in clinical samples [61, 62]. However, its utility for identification of S. pneumoniae turned out to be limited [63] due to the presence of the ply gene also in other Mitis streptococci [26, 64–66]. The lytA gene, which encodes LytA, the major pneumococcal autolysin responsible for the solubilization of pneumococci in bile [67] was another gene proposed for detection purposes [68]. However, counterparts of lytA were found in other Mitis streptococci [64, 65]. Observed sequence differences between lytA alleles specific for pneumococci and other Mitis streptococci, the presence of a characteristic 6-bp deletion close to the lytA 3’ end in Mitis-specific variants of the gene in particular were the basis of a PCR-restriction fragment length polymorphism (PCR-RFLP) test [69] and a real-time PCR targeting this region [55]. While these assays are considered to have a very good species specificity for S. pneumoniae [38], variants typical for S. pneumoniae in isolates from the Mitis group, as well as a Mitis-specific allele in pneumococcus, were occasionally observed [24, 70]. The psaA gene, encoding pneumococcal surface antigen A (PsaA), which constitutes a manganese ABC-type transporter is commonly found among S. pneumoniae and was proposed as a target for PCR-based detection [71], showing good performance on collections of S. pneumoniae and other Mitis streptococci [55, 72, 73]. However, also this gene turned out to be present among some isolates of Mitis streptococci [26, 66, 74, 75]. Yet another proposed virulence-associated target, the cpsA (waz) gene, which is typically the first gene in the cps operon, responsible for the biosynthesis of capsular polysaccharide [76], is absent among some non-serotypable pneumococci, as discussed above (see Immunological methods). In addition, the cpsA alleles of pneumococcal serotypes 25A, 25F and 38 are too divergent from typical pneumococcal cpsA genes [77] to yield positive results in this PCR, and such problems might also occur in the case of some novel serotypes. To additionally complicate identification issues, functional cps loci, including the conserved genes cpsABCD, were observed in a number of species of Mitis streptococci [47–49], suggesting a low utility of the cpsA target for discrimination between S. pneumoniae and other Mitis group representatives. Moreover, the similarity of some cps loci from the Mitis group streptococci to their counterparts in S. pneumoniae challenges the use of multiplex PCR for direct detection of pneumococcal serotypes, e.g. in carriage specimens [78]. Two genes, piaA and piaB encoding
components of an ABC transport system essential for iron uptake, were described as specific for *S. pneumoniae* but these genes are absent in some pneumococcal isolates, which limits the sensitivity of the test [79–81]. Moreover, *piaA* was detected in some *S. mitis* and *S. pseudopneumoniae* [66] and recently *S. pseudopneumoniae* isolates positive for *piaB* have been reported [70].

Certain other genes, such as housekeeping genes and genes of unknown function were also proposed as targets for *S. pneumoniae* detection and identification. Sequence differences between variants of the *recA* gene, occurring in *S. pneumoniae* and other Mitis streptococci allowed designing a species-specific PCR [82], which could discriminate also *S. pseudopneumoniae* [83]. Although sequencing of the 16S rRNA gene, considered as the “gold standard” for identification of many bacterial species, has a limited utility for Mitis streptococci due to high similarity of 16S rRNA genes in this group [32], a cytosine at nucleotide position 203 in the 16S rRNA genes is considered specific for pneumococci, and sequence variants in this region of the gene can be distinguished using PCR-RFLP of the 16S rRNA genes [84], although the presence of mixed bases at this position was found for some *S. pseudopneumoniae* [26]. On the basis of genomic subtractive hybridization two loci of unknown function, Spn9802 and Spn9828 were proposed to be specific for *S. pneumoniae* [85, 86]. Later studies demonstrated that a PCR targeting spn9802 apart from *S. pneumoniae* also detects *S. pseudopneumoniae* but not *S. mitis* and *S. oralis* and thus can be used to distinguish these two groups [87]. Comparative genomic analyses of *S. pneumoniae* and *S. pseudopneumoniae* led to the identification of markers specific for these two species, respectively SPN0001/SP2020, encoding a GtnR-family transcriptional regulator, and SPS0002, specifying an osmosensitive potassium channel histidine kinase/response regulator KdpDE [88]. The specificity of both markers was close to 100%, with SPN0001/SP2020 found in a few non-pneumococcal strains and SPS0002 occasionally present in *S. pneumoniae* [70, 88].

Currently, it appears that a combination of a few assays, such as PCR-RFLP or real-time PCR specific for the 6-bp deletion in the 3’ end of *lytA*, together with detection of SPN0001/SP2020 [70], possibly supplemented with analysis of the presence of *piaA/piaB* could give more reliable results than a single-gene detection. Defining only one, “perfect” target, present among all *S. pneumoniae* strains, sufficiently conserved to design universal PCR primers and at the same time not found (or at least in a very divergent form) among any other Mitis group streptococci may be very difficult to achieve, if possible at all. This situation is associated with the common genetic ancestry of the Mitis group, relatively frequent HGT among streptococcal species, facilitated by their natural competence, together with a gene loss, resulting in populations with highly diverse gene content [89, 90]. In consequence, the core genome (genes common for each genome of a species) of *S. pneumoniae* is much smaller than the pangenome (a complete set of genes of a species) and many of these genes are shared with other Mitis streptococci, including determinants of factors playing the role in interactions with the host allowing a successful survival [91–94]. The recent genomic analyses of the Mitis group streptococci are discussed in more detail below (see Genomic studies in the identification of Mitis streptococci).

**Identification based on Sanger sequencing of single and multiple genes**

The approach is based on analysis of sequence differences specific for various species, usually occurring in housekeeping genes, i.e. genes ubiquitously present among all Mitis streptococci, including *S. pneumoniae*. In comparison to PCR-based detection, Sanger sequencing is more time-consuming and represents a more expensive technology in terms of reagents and equipment. In addition, this approach requires highly qualified personnel and specialized software, although an increasing availability of commercial sequencing services makes it accessible for a growing number of laboratories. Both single-gene sequencing targets and multilocus schemes were proposed for the purposes of identification of *S. pneumoniae* and other Mitis streptococci.

The *sodA* gene encoding manganese-dependent superoxide dismutase, the *rpoB* gene encoding RNA polymerase beta-subunit and the *tuf* gene encoding elongation factor Tu were among the earliest proposed sequencing targets [95–98]. However, some later studies questioned their applicability for unambiguous identification, taking as reference results obtained with multiplex PCR, multilocus sequence analysis (MLSA, see below) or genomic analyses [16, 22, 87, 99, 100]. Other sequencing targets for identification purposes include *rnpB* (the RNase P RNA gene), *rpoA* (RNA polymerase alpha subunit gene), *recN* (recombination/repair protein gene), *gyrB* (DNA topoisomerase subunit B gene) and *rpsB* (the 30S ribosomal protein S2 gene) [80, 99, 101–103]; of these, *rnpB* and *gyrB* showed a good performance also in other studies [104].

Multilocus sequencing approaches, which target a few (usually seven) loci, occupying separate localizations on the bacterial chromosome, overcome to a significant extent the bias introduced by the gene exchange between *S. pneumoniae* and other Mitis streptococci. Two such schemes, multilocus sequence typing (MLST) [105] and multilocus sequence analysis (MLSA) [106], have been used in numerous studies. MLST, based on sequencing of seven housekeeping loci (*aroE, gdh, gki, recP, spi, xpt and ddl*) of *S. pneumoniae* is considered to unambiguously confirm the identification of an isolate as pneumococcus [107]. In MLST, after PCR amplification, products are sequenced and the
identification of alleles and sequence types (STs) from allelic profiles is done using a large, web-accessible database (https://pubmlst.org/spneumoniae/; 15th April 2020 date last accessed), that currently contains over 15 thousand different STs and data for almost 50 thousand isolates of *S. pneumoniae* from all over the world. As a PCR-based technique, MLST can also be applied directly to clinical samples, such as cerebrospinal fluid [108]. While MLST operates on allelic profiles, MLSA is based on concatenated sequences of housekeeping loci that are used for the construction of dendrograms. The MLSA approach employing genes from the MLST schemes was used for species resolution of *Burkholderia* and *Neisseria* [109], while the identification of species within the Mitis group streptococci and other VGS seven different target loci (*map, pfl, ppaC, pyk, rpoB, sodA* and *tuf*) was selected [106]. More in the detail, sequences of loci belonging to an MLSA scheme, obtained for an analysed isolate are concatenated in a defined order, aligned with corresponding reference sequences and the alignment is used in phylogenetic clustering. The position of an isolate in the resulting tree is then examined for assigning a particular species. Unfortunately, the website described in the original paper on MLSA of VGS and dedicated for this purpose (http://www. eMLSA.net) appears to be no longer active (as of 15 April 2020). The MLSA scheme for VGS has proven its utility in a number of studies [19, 21, 22, 41, 99, 110]. This method showed a very good concordance with the results of core genome clustering [2, 26, 100] and species identification based on sequencing with the minION napore [41].

### Genomic studies in the identification of Mitis streptococci

The recent advances in whole-genome sequencing (WGS), including the availability of software for analysis of genomic data, have opened new possibilities for taxonomy and identification of Mitis streptococci. Bioinformatics tools have been developed for purposes of species identification, such as Kraken, which allows fast assigning taxonomic labels from short-read sequencing data, including metagenomic samples [111]. Such a metagenomic approach was recently applied for the detection of pathogen DNA in cerebrospinal fluid samples from patients with pneumococcal meningitis [112]. Ribosomal MLST (rMLST) was designed as a universal identification tool for bacteria, which indexes variation of the 53 genes of bacterial ribosomal proteins [113]. For identification of species and a ribosomal sequence type (rST), assembled genomic sequences are submitted to a specialized database (https://pubmlst.org/rmlst/; 20th April 2020 date last accessed) for a query. The results of rMLST turned out to be in very good agreement with the core genome-based clustering of *S. pneumoniae, S. mitis* and *S. pseudopneumoniae* [26].

Clustering based on the variability of genes belonging to the core genome of Mitis streptococci, i.e. genes present among all studied strains, has proven to be the most informative and reliable approach for analyses of phylogenetic and taxonomic relationships among Mitis streptococci [2, 91, 100]. These investigations also allowed re-assigning species of several isolates of Mitis streptococci, in which genomic data were deposited in GenBank [2], demonstrating the utility of core genome-based studies for identification. The MinION nanopore sequencer developed by Oxford Nanopore Technologies (ONT) provides long sequencing reads in real time and its small size allows performing analyses also in the field conditions. Sequencing results can be analysed using either Kraken or the “What’s in my Pot?” (WIMP) analysis pipeline [114]. The analysis of data from the MinION nanopore sequencer with WIMP and Kraken2 provided reliable identification of clinical isolates of Mitis streptococci [41], opening new possibilities in the detection and identification of these microorganisms.

Although genomic analyses are currently limited to highest-level reference centres and research laboratories, the importance and the popularity of this approach will certainly grow in the near future, with the presumable continuation of diminishing costs of sequencing and availability of more user-friendly software. Increasing amounts of genomic data for streptococci will aid the re-evaluation of known PCR and sequencing targets as well as identification of new ones for quick and cost-effective diagnostics, such as already demonstrated for SPN0001 and SPS0002 [88], discussed above. A recent comparative analysis [92] performed on 60 genomes of *S. pneumoniae, S. mitis, S. pseudopneumoniae, S. oralis* and *S. infantis* followed by a screening of over 7500 available streptococcal genomes allowed the identification of 224 genes present in at least 85% of *S. pneumoniae* genomes but absent from 80% or more genomes of *S. mitis*, and thus postulated to be associated with virulence of pneumococcus, as opposed to the commensal lifestyle of *S. mitis*. Of these 224 genes, 49 ones, which were totally absent from *S. mitis, S. pseudopneumoniae* and *S. oralis*, included genes encoding surface-located proteins PspA, CbpA and “Xisco”, the iron uptake operon *piaABCD*, genes involved in carbohydrate metabolism and others. Pangenomic analysis of 32 genomes of *S. pneumoniae, 36 genomes of S. mitis* and 13 genomes of *S. pseudopneumoniae* followed by a validation of results on larger panels of streptococcal genomes revealed ten genes, residing in five loci, specific solely for *S. pneumoniae* and nine genes unique for *S. pseudopneumoniae* [94]. Four of loci characteristics for *S. pneumoniae*, including the *piaABCD* operon, the locus containing SP2020 and two single genes, encoding a putative ABC transporter and a hypothetical protein, were common with the set of 49 genes from the previously mentioned study [92]. Additional analyses of six selected genes, longer than 500 bp, considered specific for *S. pneumoniae* and proposed as PCR targets revealed the presence of four of them.
in a few genomes of *S. pseudopneumoniae* consistent with the need to use more than a single marker for identification purposes [94].

**Conclusions**

The correct identification of species belonging to the Mitis group of streptococci is important for clinical practice, epidemiology and studies on the biology of these organisms. Several molecular approaches have been proposed over time to supplement traditional microbiological methods. The performance of PCR-based methods, targeting a single gene varied considerably, mostly due to the presence of counterparts of some of these “pneumococcal” genes in other Mitis streptococci, in many cases presumably associated with the high frequency of HGT. The circulation of genuine pneumococci lacking certain target genes represents another issue that complicates PCR-based detection and identification. Using the combination of a few selected targets significantly increases the probability of correct identification by PCR, which represents a fast and cost-effective method, available for many laboratories. Multilocus sequencing (such as the MLST and MLSA approaches) and genomic sequencing provide a clear distinction among species of the Mitis group and it can be expected that in the near future the ongoing advances in WGS and downstream data analysis will make such approaches accessible for an increasing number of laboratories. Genomic studies have already greatly improved our understanding of relationships within the group of Mitis streptococci and they will also aid to better define consensus targets for fast, cost-effective and specific detection of *S. pneumoniae* and other Mitis streptococci.

**Funding information** This work was supported by the grant 2013/10/M/NZ2/00259 from Narodowe Centrum Nauki (NCN), Poland, and by Narodowy Program Ochrony Antybiotyków (NPOA) from the Ministry of Health, Poland.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

---

**References**

1. Facklam R (2002) What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin Microbiol Rev 15: 613–630
2. Jensen A, Scholz CFP, Kilian M (2016) Re-evaluation of the taxonomy of the mitis group of the genus *Streptococcus* based on the whole genome phylogenetic analyses, and proposed reclassification of *Streptococcus dentisani* as *Streptococcus oralis* subsp. *dentisani* comb. nov., *Streptococcus tigurinus* as *Streptococcus oralis* subsp. *tigurinus* comb. nov., and *Streptococcus oligofermentans* as a later synonym of *Streptococcus cristatus*. Int J Syst Evol Microbiol 66:4803–4820. https://doi.org/10.1099/ijsem.0.001433
3. Patel S, Gupta RS (2018) Robust demarcation of fourteen different species groups within the genus *Streptococcus* based on genome-based phylogenies and molecular signatures. Infect Genet Evol 66:130–151. https://doi.org/10.1016/j.meegid.2018.09.020
4. Troeger C, Forouzanfar M, Rao PC, Khalil I, Brown A, Swartz S, Fullman N, Mosser J, Thompson RL, Reiner RC Jr, Abajobir A, Alam N, Alemayohu MA, Amare AT, Antonio CA, Asayesh H, Avelino E, Barac A, Beshir MA, Boneya DJ, Brauer M, Dandonna L, Dandonna R, Fitchett JRA, Gebrehiwot TT, Hailu GB, Hotez PJ, Kasaean C, Khoja T, Kissoon N, Knibbs L, Kumar GA, Rai RK, El Razeq HMA, Mohammed MSK, Nielson K, Oren E, Osman A, Patton G, Qorbani M, Roba HS, Sartorius B, Savic M, Shigematsu M, Sykes B, Swaminathan S, Topor-Madry R, Ukwaja K, Werdecker A, Yonemoto N, El Sayed ZM, Lim SS, Naghavi M, Vos T, Hay SI, Murray CJL, Mokdad AH (2017) Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the global burden of disease study 2015. Lancet Infect Dis 17:1133–1161. https://doi.org/10.1016/S1473-3099(17)30396-1
5. Wahl B, O’Brien KL, Greenbaum A, Majumder A, Liu L, Chu Y, Lukšiè I, Nair H, McAlister DA, Campbell H, Rudan I, Black R, Knoll MD (2018) Burden of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. Lancet Glob Health 6:e744–e757
6. van Aalst M, Løtsh C, Spijker K, van der Meer JT, Langendam MW, Goorhuis A, Grobusch MP, de Bree GIJ (2018) Incidence of invasive pneumococcal disease in immunocompromised patients: a systematic review and meta-analysis. Travel Med Infect Dis 24: 89–100. https://doi.org/10.1016/j.tmaid.2018.05.016
7. Cherazard R, Epstein M, Doan TL, Salim T, Bharti S, Smith MA (2017) Antimicrobial resistant *Streptococcus pneumoniae*: prevalence, mechanisms, and clinical implications. Am J Ther 24:e361–e369. https://doi.org/10.1097/MTJ.0000000000000551
8. Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease following pneumococcal vaccination: a discussion of the evidence. Lancet 378:1962–1963. https://doi.org/10.1016/S0140-6736(10)62225-8
9. Koelman DLH, Brouwer MC, van de Beek D (2020) Resurgence of pneumococcal meningitis in Europe and Northern America. Clin Microbiol Infect 26:199–204. https://doi.org/10.1016/j.cmi.2019.04.032
10. Doern CD, Burnham CAD (2010) It's not easy being green: the viridans group streptococci, with a focus on pediatric clinical manifestations. J Clin Microbiol 48:3829–3835. https://doi.org/10.1128/JCM.01563-10
11. Beynon RP, Bahl VK, Prendergast BD (2006) Infective endocarditis. BMJ 333:334–339
12. Murdoch DR, Corey GR, Hoen B, Miró JM, Fowler VG Jr, Bayer AS, Karchmer AW, Olaison L, Pappas PA, Moreillon P,
of *Streptococcus pneumoniae* from other mitis group streptococci. Sci Rep 7:7167. doi:10.1038/s41598-017-0772-x

40. Teles C, Smith A, Ramage G, Lang S (2011) Identification of clinically relevant viridans group streptococci by phenotypic and genotypic analysis. Eur J Clin Microbiol Infect Dis 30:243–250. doi:10.1007/s10096-010-1076-y

41. Imai K, Nemoto R, Kodama M, Tarumoto N, Sakai J, Kawamura T, Ibeuchi K, Mitsukata K, Murakami T, Maesaki S, Fujiwara T, Hayakawa S, Hoshino T, Seki M, Maeda T (2020) Rapid and accurate species identification of mitis group streptococci using the MiniON Nanopore sequencer. Front Cell Infect Microbiol 10:11. doi:10.3389/fcimb.2020.00011

42. Murray PR (2010) Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: usefulness for taxonomy and epidemiology. Clin Microbiol Infect 16:1626–1630. https://doi.org/10.1111/j.1469-0691.2010.03364.x

43. Ikyannikova LN, Filimonova AV, Malakhova MV, Savinova T, Filimonova O, Ilina EN, Dubovickaya VA, Sidorenko SV, Goverun VM (2012) Discrimination between *Streptococcus pneumoniae* and *Streptococcus mitis* based on sorting of their MALDI mass spectra. Clin Microbiol Infect 19:1066–1071. https://doi.org/10.1111/j.1469-0691.2011.12113

44. Zhou M, Yang Q, Kudinha T, Zhang L, Xiao M, Kong F, Zhao Y, Xu YC (2016) Using Matrix-Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) complemented with selected 16S rRNA and *gyrB* genes sequencing to practically identify clinical important Viridans Group Streptococci (VGS). Front Microbiol 7:1328. https://doi.org/10.3389/fmicb.2016.01328

45. Harju J, Lange C, Kostzrewa M, Maier T, Rantakokko-Jalava K, Haanpää M (2017) Improved differentiation of *Streptococcus pneumoniae* and other *S. mitis* group streptococci by MALDI Biotyper using an improved MALDI Biotyper database content and a novel result interpretation algorithm. J Clin Microbiol 55:914–922. https://doi.org/10.1128/JCM.01990-16

46. Paton JC, Trappetti C (2019) *Streptococcus pneumoniae* capsaicin polysaccharide. Microbiol Spectr 7. https://doi.org/10.1128/microbiolspec.GPP3-0019-2018

47. Skov Sørensen UB, Yao K, Yang Y, Tettelin H, Kilian M (2016) Improved differentiation of *Streptococcus pneumoniae* and a novel result interpretation algorithm. J Clin Microbiol 55:892–898. https://doi.org/10.1128/JCM.02243-09

48. Lessa FC, Milucky J, Rouphael NG, Bennett NM, Talbot HK, Hayakawa S, Hoshino T, Seki M, Maeda T (2020) Rapid and accurate species identification of mitis group streptococci using the MiniON Nanopore sequencer. Front Cell Infect Microbiol 10:11. doi:10.3389/fcimb.2020.00011

49. Pimenta F, Gertz RE Jr, Park SH, Kim E, Moura I, Milucky J, Lessa FC, Milucky J, Rouphael NG, Bennett NM, Talbot HK, Hayakawa S, Hoshino T, Seki M, Maeda T (2020) Rapid and accurate species identification of mitis group streptococci using the MiniON Nanopore sequencer. Front Cell Infect Microbiol 10:11. doi:10.3389/fcimb.2020.00011

50. Sloved HC, Kaitlof M, Skovsted IC, Kern MB, Espersen F (2004) Simple, rapid latex agglutination test for serotyping of pneumococci (*Pneumotest-Latex*). J Clin Microbiol 42:2518–2225

51. Kuch A, Gołębiewska A, Wałko I, Ronkiewicz P, Markowska M, Hryniewicz W, Skoczynska A (2014) usefulness of pneumotest-latex for direct serotyping of *Streptococcus pneumoniae* isolates in clinical samples. J Clin Microbiol 52:2647–2649. https://doi.org/10.1128/JCM.00451-14

52. Chewapreecha C, Harris SR, Croucher NJ, Turner C, Martinen P, Cheng L, Pessia A, Anensen DM, Mather AE, Page AJ, Salter SJ, Harris D, Nosten F, Goldblatt D, Corander J, Parkhill J, Turner P, Bentley SD (2014) Dense genomic sampling identifies highlights of pneumococcal recombination. Nat Genet 46:305–309. https://doi.org/10.1038/ng.2895

53. Hilty M, Wüthrich D, Salter SJ, Engel H, Campbell S, Sá-Leão R, de Lancastre H, Hermans P, Sadowy E, Turner P, Chewapreecha C, Diggie M, Pluschke G, McGee L, Eser ÖK, Low DE, Smith-Vaughan H, Endimiani A, Küffer M, Dupasquier M, Beaudong E, Weber J, Bruggmann R, Hanage WP, Parkhill J, Hathaway LJ, Mühlemann K, Bentley SD (2014) Global phylogenomic analysis of nonencapsulated *Streptococcus pneumoniae* reveals a deep branching classic lineage that is distinct from multiple sporadic lineages. Genome Biol Evol 6:3281–3294. https://doi.org/10.1093/gbe/evu263

54. Varghese R, Jayaraman R, Veeraraghavan B (2017) Current challenges in the accurate identification of *Streptococcus pneumoniae* and its serogroups/serotypes in the vaccine era. J Microbiol Methods 141:48–54. https://doi.org/10.1016/j.mimet.2017.07.015

55. da Gloria CM, Tondella ML, McCaustland K, Weidlich L, McGee J, Mayer LW, Steigerwalt A, Whaley M, Facklam RR, Fields B, Carlone G, Ades EW, Dagan R, Sampson JS (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

56. Book M, Lehmann LE, Zhang X, Stüber F (2013) Monitoring infection: from blood culture to polymerase chain reaction (PCR). Best Pract Res Clin Anaesthesiol 27:279–288. https://doi.org/10.1016/j.bpa.2013.06.010

57. da Gloria CM, Pimenta FC, Jackson D, Roundtree A, Ahmad Y, Millar EV, O’Brien KL, Whitney CG, Cohen AL, Beall BW (2010) Revisiting pneumococcal carriage by use of broth enrichment and PCR techniques for enhanced detection of carriage and serotypes. J Clin Microbiol 48:1611–1618. https://doi.org/10.1128/JCM.02423-09

58. Strání K, Herrmann B, Abdeldaim G, Oleçon P, Holmberg H, Mölling P (2014) Comparison of *lytA* and *nasA* as qualitative samples and of the PCR gene targets *lytA* and *Spn9802* for quantitative PCR for rapid detection of pneumococcal pneumonia. J Clin Microbiol 52:83–89. https://doi.org/10.1128/JCM.01742-13

59. Gillis HD, Lang ALS, ELSherif M, Martin I, Hatchette TF, McNeil SA, LeBlanc JJ, Serious Outcomes Surveillance (SOS) Network of the Canadian Immunization Research (CIRN) study. BMJ Open 7:e015008. https://doi.org/10.1136/bmjopen-2016-013008

60. Sundell N, Andersson LM, Brittain-Long R, Sundvall PD, Alsiö A, Lindh M, Gustavsson L, Westin J (2019) PCR detection of respiratory pathogens in asymptomatic and symptomatic adults. J Clin Microbiol 57:e00716–e00718. https://doi.org/10.1128/JCM.00716-18

61. Virolainen A, Salo P, Jero J, Karma P, Eskola J, Leinonen M (1994) Comparison of PCR assay with bacterial culture for detecting *Streptococcus pneumoniae* from nasopharyngeal swabs collected for viral studies in Canadian adults hospitalised with community-acquired pneumonia: a Serious Outcomes Surveillance (SOS) Network of the Canadian Immunization Research (CIRN) study. BMJ Open 7:e015008. https://doi.org/10.1136/bmjopen-2016-013008

62. Salo P, Ortvist A, Leinonen M (1995) Diagnosis of bacteremic pneumococcal pneumonia by amplification of pneumolysin gene in middle ear fluid of children with acute otitis media. J Clin Microbiol 32:2667–2670

63. Kaijalainen T, Rintamäki S, Herva E, Leinonen M (2002) Evaluation of genotypic and conventional methods in the identification of *Streptococcus pneumoniae*. J Microbiol Methods 52:111–118
64. Whatmore AM, Efstratiou A, Pickering AP, Broughton K, Woodard G, Sturgeon D, George R, Dowson CG (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

65. Neelameggha C, Klasing CH, Klomberg DM, de Valk HA, Mouton JW (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 42:4345–4357

66. Johnston C, Hinds J, Smith A, van der Linden M, Van Eldere J, Mitchell TJ (2010) Detection of large numbers of pneumococcal virulence genes in streptococci of the mitis group. J Clin Microbiol 48:2762–2769. https://doi.org/10.1128/JCM.01746-09

67. Mosser JL, Tomasz A (1970) Choline-containing teichoic acid as a structural component of pneumococcal cell wall and its role in sensitivity to lysis by an autolytic enzyme. J Biol Chem 245:287–298

68. Gillespie SH, Ullman C, Smith MD, Emery V (1994) Detection of Streptococcus pneumoniae in sputum samples by PCR. J Clin Microbiol 32:1308–1311

69. Llull D, López R, García E (2006) Characteristic signatures of the lytA gene provide a basis for rapid and reliable diagnosis of Streptococcus pneumoniae infections. J Clin Microbiol 44:1250–1256

70. Tavares DA, Handem S, Carvalho RJ, Paulo AC, de Lencastre H, Morrison KE, Lake D, Crook J, Carlone GM, Ades E, Facklam RR (2004) Comparison of four polymerase chain reaction assays for specificity in the identification of Streptococcus pneumoniae. J Clin Microbiol 42:4345–4357

71. Verhelst R, Kaaijalainet N, De Baere E, Verschraegen G, Claey G, Van Smaey L, De Ganck C, Vaneechoutte M (2003) Comparison of five genotypic techniques for identification of optochin-resistant pneumococci-like isolates. J Clin Microbiol 41:3521–3525

72. Messmer TO, Sampson JS, Stinson A, Wong B, Carlon GM, Facklam RA (2000) Confirmation of pyaA in all 90 serotypes of Streptococcus pneumoniae by PCR and potential of this assay for identification and diagnosis. J Clin Microbiol 38:434–437

73. Jado I, Fenoll A, Casal J, Pérez A (2001) Identification of the psaA gene, coding for pneumococcal surface adhesin A, in viridans group streptococci other than Streptococcus pneumoniae. Clin Diagn Lab Immunol 8:895–898

74. Zhang Q, Ma Q, Su D, Li Q, Yao W, Wang C (2010) Identification of horizontal gene transfer and recombination of PsaA gene in Streptococcus mitis group. Microbiol Immunol 54:313–319. https://doi.org/10.1111/j.1348-0421.2010.00216.x

75. Park HK, Lee SJ, Yoon JW, Shin JW, Shin HS, Kook JK, Myung SC, Kim W (2010) Identification of the cpsA gene as a specific marker for the discrimination of Streptococcus pneumoniae from viridans group streptococci. J Med Microbiol 59:1146–1152. https://doi.org/10.1099/jmm.0.017798-0

76. Marvoldi A, Aanensen DM, Godoy D, Skovsted IC, Kaltoft MS, Reeves PR, Bentley SD, Spratt BG (2007) Genetic relatedness of the Streptococcus pneumoniae capsular biosynthetic loci. J Bacteriol 189:7841–7855. https://doi.org/10.1128/JB.00836-07

77. da Gloria CM, Pimenta FC, Moura I, Roundtree A, Gertz RE Jr, Li Z, Jagero G, Bigogo G, Junghae Me, Conklin L, Feikin DR, Breiman RF, Whitney CG, Beall BW (2013) Non-pneumococcal mitis-group streptococci confound detection of pneumococcal capsular serotype-specific loci in upper respiratory tract. PeerJ 1:e97. https://doi.org/10.7717/peerj.97

78. Whalan RH, Funnell SG, Bowler LD, Hudson MJ, Robinson A, Dowson CG (2006) Distribution and genetic diversity of the ABC transporter lipoproteins PiuA and PiaA within Streptococcus pneumoniae and related streptococci. J Bacteriol 188:1031–1038

79. Wessels E, Schelfaut JJ, Bernards AT, Claas EC (2012) Development of a real-time PCR assay for the specific detection and identification of Streptococcus pseudopneumoniae using the recA gene. Clin Microbiol Infect 18:1089–1096. https://doi.org/10.1111/j.1469-0691.2011.03684.x

80. Scholz CF, Poulsen K, Kilian M (2012) Novel molecular method for identification of Streptococcus pneumoniae applicable to clinical microbiology and 16S rRNA sequence-based microbiome studies. J Clin Microbiol 50:1968–1973. https://doi.org/10.1128/JCM.00365-12

81. Suzuki N, Seki M, Nakano Y, Kiyoura Y, Maeno M, Yamashita Y (2005) Discrimination of Streptococcus pneumoniae from viridans group streptococci by genomic subtractive hybridization. J Clin Microbiol 43:434–437

82. Suzuki N, Yuyama M, Maeda S, Ogawa H, Mashiko K, Kiyoura Y (2006) Genotypic identification of presumptive Streptococcus pneumoniae by PCR using four genes highly specific for S. pneumoniae. J Med Microbiol 55:709–714

83. Wessels E, Schelfaut JJ, Bernards AT, Claas EC (2012) Evaluation of several biochemical and molecular techniques for identification of Streptococcus pneumoniae and Streptococcus pseudopneumoniae and their detection in respiratory samples. J Clin Microbiol 50:1171–1177. https://doi.org/10.1128/JCM.06609-11

84. Croxen MA, Lee TD, Azana R, Hoang LM (2018) Use of genomics to design a diagnostic assay to discriminate between Streptococcus pneumoniae and Streptococcus pseudopneumoniae. Microb Genom 4. https://doi.org/10.1099/mgen.0.000175

85. Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiolii SV, Oggoni M, Dunning Hotopp JC, Hu FZ, Riley DR, Covacci A, Mitchell TJ, Bentley SD, Kilian M, Ehrlich GD, Rappuoli R, Moxon ER, Massignani V (2010) Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species. Genome Biol 11:R107. https://doi.org/10.1186/gb-2010-11-10-r107

86. Andam CP, Hanage WP (2015) Mechanisms of genome evolution of Streptococcus. Infect Genet Evol 33:334–342. https://doi.org/10.1016/j.meegid.2014.11.007

87. Velsko IM, Perez MS, Richards VP (2019) Resolving phylogenetic relationships for Streptococcus mitis and Streptococcus oralis through core- and pan-genome analyses. Genome Biol Evol 11:1077–1087. https://doi.org/10.1093/gbe/evz049
92. Kilian M, Tettelin H (2019) Identification of virulence-associated properties by comparative genome analysis of Streptococcus pneumoniae, S. pseudopneumoniae, S. mitis, three S. oralis sub-species, and S. infantis. mBio 10:e01985–e01919. https://doi.org/10.1128/mBio.01985-19

93. Salvadori G, Junges R, Morrison DA, Petersen FC (2019) Competence in Streptococcus pneumoniae and close commensal relatives: mechanisms and implications. Front Cell Infect Microbiol 9:94. https://doi.org/10.3389/fcimb.2019.00094

94. Gonzales-Siles L, Karlsson R, Schmidt P, Salvà-Serra F, Jaén-Luchoro D, Skovbjerg S, Moore ERB, Gomila M (2020) A panageneome approach for discerning species-unique gene markers for identifications of Streptococcus pneumoniae and Streptococcus pseudopneumoniae. Front Cell Infect Microbiol 10:222. https://doi.org/10.3389/fcimb.2020.00222

95. Poyart C, Quesne G, Coulon S, Berche P, Trieu-Cuot P (1998) Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. J Clin Microbiol 36:41–47

96. Kawamura Y, Whiley RA, Shu SE, Saukko P, Poyart C, Quesne G, Coulon S, Berche P, Trieu-Cuot P (1998) Use of rpoB gene polymorphisms accurately assign in-vasive viridans group streptococci. J Clin Microbiol 36:41–47

97. Rasmussen LH, Dargis R, Hojholjt K, Christensen JJ, Skovgaard O, Justensen US, Rosenvinge FS, Moser C, Lukjancenko O, Rasmussen S, Nielsen XC (2016) Whole genome sequencing as a tool for phylogenetic analysis of clinical strains of mitis group streptococci. Eur J Clin Microbiol Infect Dis 35:1615–1625. https://doi.org/10.1007/s10096-016-2700-2

98. Westling K, Juhlander I, Jungman P, Vondracek M, Wretlind B, Jalal S (2008) Identification of species of viridans group streptococci in clinical blood culture isolates by sequence analysis of the RNAse P RNA gene, rnpB. J Inf Secur 56:204–210

99. Park HK, Yoon JW, Shin JW, Kim KY, Kim W (2010) rpoA is a useful gene for identification and classification of Streptococcus pneumoniae from the closely related viridans group streptococci. FEMS Microbiol Lett 305:58–64. https://doi.org/10.1111/j.1574-6968.2010.01913.x

100. Glazunova OO, Raoult D, Roux V (2010) Partial reCN gene sequencing: a new tool for identification and phylogeny within the genus Streptococcus. Int J Syst Evol Microbiol 60:2140–2148. https://doi.org/10.1099/ijs.0.018176-0

101. Maeda Y, Goldsmith CE, Coulter WA, Mason C, Dooley JS, Lowery CJ, Millar BC, Moore JE (2011) Comparison of five gene loci (rnpB, 16S rRNA, 16S-23S rRNA, sodA and dnaJ) to aid the molecular identification of viridans-group streptococci and pneumococci. Br J Biomed Sci 68:190–196

102. Kilian M, Tettelin H, Jordan GE, Kilian M, Hanage WP, Spratt BG (2009) Assigning strains to bacterial species via the internet. BMC Biol 7:3. https://doi.org/10.1186/1741-7007-7-3

103. Hanage WP, Kaialalainen T, Herva E, Saukkoriipi A, Syrjänen R, Spratt BG (2005) Using multilocus sequence data to define the pneumococcus. J Bacteriol 187:6223–6230

104. Hargreaves AJ, Nutman BT (2014) Metagenomic analysis of the gut microbiome reveals broad-based associations of species with Enterobacteriaceae. J Med Microbiol 63:204–210

105. Hargreaves AJ, Nutman BT (2014) Metagenomic analysis of the gut microbiome reveals broad-based associations of species with Enterobacteriaceae. J Med Microbiol 63:204–210

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.