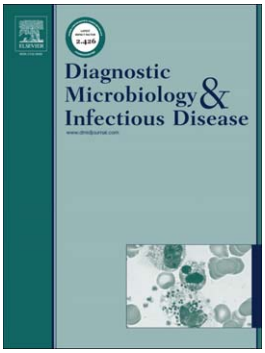
lytA-based identification methods can misidentify Streptococcus pneumoniae



lytA-based identification methods can misidentify *Streptococcus pneumonia*e

Alexandra S. Simões, Débora A. Tavares, Dora Rolo, Carmen Ardanuy, Herman Goossens, Birgitta Henriques-Normark, Josefina Linares, Hermínia de Lencastre, Raquel Sá-Leão

PII: S0732-8893(16)30072-4
DOI: doi: 10.1016/j.diagmicrobio.2016.03.018
Reference: DMB 14051

To appear in: *Diagnostic Microbiology and Infectious Disease*

Received date: 17 November 2015
Revised date: 19 February 2016
Accepted date: 20 March 2016

Please cite this article as: Simões Alexandra S., Tavares Débora A., Rolo Dora, Ardanuy Carmen, Goossens Herman, Henriques-Normark Birgitta, Linares Josefina, de Lencastre Hermínia, Sá-Leão Raquel, *lytA*-based identification methods can misidentify *Streptococcus pneumonia*e, *Diagnostic Microbiology and Infectious Disease* (2016), doi: 10.1016/j.diagmicrobio.2016.03.018

lytA-based identification methods can misidentify *Streptococcus pneumonia*e

Alexandra S. Simões[a,1,#], Débora A. Tavares[a,#], Dora Rolo[b,c,2], Carmen Ardanuy[b,c], Herman Goossens[d,e], Birgitta Henriques-Normark[f,g], Josefina Linares[b,c,], Hermínia de Lencastre[h,i], and Raquel Sá-Leão[a,*]

[a]Laboratory of Molecular Microbiology of Human Pathogens, Instituto de Tecnologia Química e Biológica António Xavier (ITQB), Universidade Nova de Lisboa (UNL), Oeiras, Portugal

[b]Microbiology Department, Hospital Universitari de Bellvitge, Universitat de Barcelona-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain

[c]CIBERES (Ciber de Enfermedades Respiratorias), ISCIII, Madrid, Spain

[d]Department of Medical Microbiology, University of Antwerp, Antwerp, Belgium

[e]Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium

[f]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden

[g]Department of Laboratory Medicine, Division of Clinical Microbiology, Karolinska University Hospital, Stockholm, Sweden

[h]Laboratory of Molecular Genetics, ITQB, UNL, Oeiras, Portugal

[i]Laboratory of Microbiology and Infectious Diseases, The Rockefeller University, New York, NY, USA

Running Tittle: *lytA*-based assays can misidentify pneumococci

*Corresponding Author.

[1] Present Address: Global Health and Tropical Medicine, Instituto de Higiene e Medicina Tropical, UNL, Lisboa, Portugal
[2] Present Address: ISGlobal, Barcelona Ctr. Int. Health Res. (CRESIB), Hospital Clínic - Universitat de Barcelona, Barcelona, Spain

Laboratory of Molecular Microbiology of Human Pathogens

Instituto de Tecnologia Química e Biológica António Xavier

Av. da República, 2780-157 Oeiras, Portugal

Tel.: + 351 214469872

E-mail address: rsaleao@itqb.unl.pt (R Sá-Leão)

[#]These authors contributed equally to this work

Word Counts for Abstract: 145

Word Counts for Body of the text: 3,556

Abstract

During surveillance studies we detected, among over 1,500 presumptive pneumococci, 11 isolates displaying conflicting or novel results when characterized by widely accepted phenotypic (optochin susceptibility and bile solubility) and genotypic (*lytA*-BsaAI-RFLP and MLST) identification methods. We aimed to determine the genetic basis for the unexpected results given by *lytA*-BsaAI-RFLP and investigate the accuracy of the WHO recommended *lytA* real-time PCR assay to classify these 11 isolates. Three novel *lytA*-BsaAI-RFLP signatures were found (one in pneumococcus and two in *S. mitis*). In addition, one pneumococcus displayed the atypical *lytA*-BsaAI-RFLP signature characteristic of non-pneumococci and two *S. pseudopneumoniae* displayed the typical *lytA*-BsaAI-RFLP pattern characteristic of pneumococci. *lytA* real-time PCR misidentified these three isolates. In conclusion, identification of pneumococci by *lytA* real-time PCR, and other *lytA*-based methodologies, may lead to false results. This is of particular relevance in the increasingly frequent colonization studies relying solely on culture-independent methods.

Keywords

S. pneumoniae; *S. pseudopneumoniae*; *lytA*; real-time PCR; identification; molecular methods.

1. Introduction

Streptococcus pneumoniae (pneumococcus) is a major human pathogen, causing a wide range of infections from otitis media to bacteremia and meningitis. Routine identification of pneumococcus (colony morphology on blood agar plates, susceptibility to optochin, cell wall lysis by 1% of sodium deoxycholate (bile solubility), and assignment of a capsular type by serotyping) is not always straightforward since some isolates may give atypical results in one or more of these assays (Arbique, et al., 2004; Balsalobre, et al., 2006; Bosshard, et al., 2004; Nunes, et al., 2008; Sá-Leão, et al., 2006; Simões, et al., 2010; Whatmore, et al., 2000).

As a human colonizer, pneumococci co-habit the nasopharynx with several other bacterial species, including its closest relatives: *S. pseudopneumoniae*, *S. mitis*, and *S. oralis*. The exchange of genetic elements between pneumococci and its closest relatives has been described and increases the difficulties in species identification (Denapaite, et al., 2010; Donati, et al., 2010; Johnston, et al., 2010). Although isolates of closely related species have been implied in disease episodes, pneumococcus is the most important disease-causing species of the mitis group (formed by pneumococcus, *S. pseudopneumoniae*, *S. mitis*, and *S. oralis*, among others) (Bochud, et al., 1994; Douglas, et al., 1993; Keith, et al., 2006; Rolo, et al., 2013). For this reason, a correct identification of pneumococcus is crucial for an accurate diagnosis and treatment. In fact, misidentification of pneumococcus could falsely increase the rates of pneumococci non-susceptible to antimicrobials since high rates of penicillin-resistant and multidrug-resistant *S. mitis* isolates have been described (Ioannidou, et al., 2001; Simões, et al., 2010; Wester, et al., 2002).

In recent years, several molecular methods have been proposed to differentiate pneumococcus from closely related species. The presence of the *lytA* gene – the major autolysin and a ubiquitous virulence factor – has been proposed to identify pneumococci (Messmer, et al., 1997). However, homologues of the *lytA* gene have been detected in

strains of closely related streptococcal species (Denapaite, et al., 2010; Romero, et al., 2004; Whatmore, et al., 2000). A *lytA*-BsaAI-RFLP strategy to differentiate pneumococcus from closely related species based on signatures characteristic of pneumococcal (typical) *lytA* or non-pneumococcal (atypical) *lytA* has been proposed and successfully used (Llull, et al., 2006). Also, based on DNA sequence differences between the pneumococcal *lytA* and its homologues, real-time PCR assays for the specific identification of pneumococcus have been developed (Carvalho, et al., 2007). Nowadays, the *lytA* real-time PCR strategy developed by the CDC is currently the WHO recommended culture-independent method to detect pneumococci (Carvalho, et al., 2007; Satzke, et al., 2013).

In addition, multilocus sequence typing (MLST) and multilocus sequence analysis (MLSA) strategies have been validated as tools for reliable species identification among streptococci of the viridans group (Hanage, et al., 2005a; Bishop, et al., 2009).

During surveillance studies, we detected 11 presumptive pneumococcal isolates displaying conflicting or novel results when characterized by the combination of optochin susceptibility, bile solubility, *lytA*-BsaAI-RFLP and, MLST. In this study, we aimed to determine the genetic basis for the unexpected results given by *lytA*-BsaAI-RFLP. Also, considering the increasing and wide use of *lytA* real-time PCR for the identification of pneumococci, we also aimed to investigate the accuracy of this method in the classification of these 11 isolates.

2. Materials and methods

2.1. Ethics statement

In the present study a sub-set of bacterial isolates selected from different studies was characterized. All samples have been coded numerically upon collection and processed anonymously. No human subjects, human material or human data were used, thus excusing the requirement for an ethical approval. Approval for the original studies was obtained from: i) the Portuguese Ministry of Education; the study was registered and approved at the Health

Care Centre of Oeiras that reports to Administração Regional de Saúde (ARS; "Regional Health Administration") of Lisboa and Vale do Tejo from the Ministry of Health (PT coded isolates); signed informed consent was obtained from parents/guardians of participating children; ii) the "Comité Ètic d'Investigació Clínica del Hospital Universitari de Bellvitge" (Spain coded isolates); and iii) research sites involved in the European project GRACE (Genomics to Combat Resistance against Antibiotics in Community-acquired LRTI in Europe) obtained ethical and competent authority approval from their local organizations. Patients who fulfilled the inclusion criteria were given written and verbal information on the study and asked for informed consent (GRA coded isolates).

2.2. Study isolates

Isolates were selected from biological samples under the framework of other studies aimed to identify pneumococci. These included surveillance carriage studies performed in Portugal (obtained between 2011-2014, n=1,226 isolates), a study of lower respiratory tract infections in several European countries (obtained between 2007-2010, n=204 isolates; GRACE - Genomics to Combat Resistance against Antibiotics in Community-acquired LRTI in Europe; www.gracelrti.org), and a collection of disease isolates with atypical properties from Spain (obtained between 1991-2009, n=132; (Rolo, et al., 2013)). The assays performed for species identification are described below and are summarized in Fig. S1. Briefly, isolates were presumptively identified as pneumococci based on presence of α-hemolysis when grown in gentamycin blood agar plates and optochin susceptibility. If optochin resistance was observed, bile solubility was performed. Presumptive pneumococci were then serotyped by the Quellung reaction and/or by multiplex PCR. When the assignment of a serotype was not possible, a multiplex PCR designed to detect non-encapsulated pneumococci and *lytA*-BsaAI-RFLP were performed as described below (Simões, et al., 2011).

Among the 1,562 isolates mentioned above and presumptively identified as pneumococci, *lytA*-BsaAI-RFLP was performed for 247 isolates (99 from Portugal, 25 from GRACE, and

123 from Spain). Of these, 61 were identified as non-encapsulated pneumococci with a typical *lytA*-BsaAI-RFLP pattern (29 from Portugal and 32 from Spain) and 175 were identified as non *S. pneumoniae* with an atypical *lytA*-BsaAI-RFLP pattern (66 from Portugal, 22 from GRACE, and 87 from Spain). The 11 isolates reported here (four from Portugal, three from GRACE, and four from Spain) exhibited conflicting (or novel) results when presumptive identification based on optochin susceptibility and bile solubility was compared to the one suggested by the *lytA*-BsaAI-RFLP typing system..

2.3. Optochin susceptibility in CO_2 and in O_2 atmosphere

Optochin susceptibility was tested by disk diffusion using commercially available optochin discs (5 µg; 6 mm; Oxoid, Hampshire, England). Discs were applied to overnight cultures plated in blood agar (trypticase soy agar supplemented with 5% defibrinated sheep blood). Plates were incubated overnight at 37°C in 5% CO_2 atmosphere or in ambient atmosphere as described by Arbique et al. (Arbique, et al., 2004). Isolates were considered resistant to optochin when they displayed inhibition zones smaller than 14 mm.

2.4. Bile solubility test

The bile solubility assay was performed according to standard procedures: colonies from an overnight culture were suspended in 1 mL of a 0.85% NaCl (w/v) solution to a turbidity equal to 0.5-1 McFarland standard (Rouff, et al., 2003). This suspension was distributed into two tubes (500 µL each tube) and 200 µL of a 10% deoxycholate solution were added to one tube while the other received 200 µL of a 0.85% NaCl (w/v) solution (control). Both tubes were incubated at 37ºC for up to 2h. A sample was considered soluble in bile when clearing of the turbidity occurred in the tube with deoxycholate but not in the control.

2.5. Capsular typing

Capsular type assignment was performed by the Quellung reaction and by multiplex PCR as previously described (Brito, et al., 2003; Pai, et al., 2006; Simões, et al., 2011).

2.6. *lytA*-BsaAI-RFLP signatures

The entire *lytA* gene was amplified by PCR using primers previously described (LA5_Ext: 5'-AAGCTTTTTAGTCTGGGGTG-3' and LA3_Ext: 5'-AAGCTTTTTCAAGACCTAATAATATG-3'), yielding a PCR product of approximately 1,200 bp (Obregon, et al., 2002). Typical (characteristic of pneumococcal *lytA*) or atypical (characteristic of non-pneumococcal *lytA*) RFLP signatures were determined as described before by digesting the PCR product with BsaAI and separating the fragments by agarose gel electrophoresis (Llull, et al., 2006).

2.7. Real-time PCR targeting *lytA and piaA*

The presence of the *lytA* gene was tested by real-time PCR using previously described primers (lytA_F: 5'-ACGCAATCTAGCAGATGAAGCA-3' and lytA_R: 5'-TCGTGCGTTTTAATTCCAGCT-3') and probe (5'-FAM-GCCGAAAACGCTTGATACAG GGAG-3'-BHQ1) (Carvalho, et al., 2007). The presence of the *piaA* gene was tested by real-time PCR using previously described primers (piaF: 5'-CATTGGTGGCTTAGTAAGTGCAA-3' and piaR: 5'-TACTAACACAAGTTCCTGATAAGGCAAGT-3') and probe (5'-FAM-TGTAAGCGGAAAAGCAGGCCTTACCC-3'-BHQ1) (Trzcinski, et al., 2013). The assays were carried out in a final volume of 25 µL using the FastStart TaqMan Probe Master (Roche) containing 2.5 µL of 0.2 ng/µL DNA, 0.15 µM of each primer, and 0.075 µM of probe. The assay was performed three times on different days and DNA from *S. pneumoniae* TIGR4 (positive control) and *S. pseudopneumoniae* ATCC BAA-960 (negative control) was used in every run. DNA was amplified with CFX96 real time system (Bio-Rad) with the cycling parameters previously described (Carvalho, et al., 2007). Samples were considered positive when cycle threshold (Ct) values were below 35.

2.8 DNA sequencing

Sequencing reactions needed for the methods described below were conducted at Macrogen, Inc. (Amsterdam, The Netherlands). Sequencing analysis was done with DNAStar (Lasergene).

2.9. *lytA* sequencing analysis

lytA PCR products of 1,200 bp were obtained as described above (Obregon, et al., 2002). Sequencing was conducted at Macrogen, Inc. (Amsterdam, The Netherlands) and subsequent analysis of the sequences was done with DNAstar (Lasergene). *lytA* sequences were also obtained for strains TIGR4 (pneumococcus, NCBI accession number AE 005672.3) and ATCC BAA-960 (*S. pseudopneumoniae*, NCBI accession number AM113495.1), to be used for comparison. Nucleotide sequences of *lytA* gene described in this study were deposited at the GenBank database with the accession numbers KT253593-KT253603.

2.10. Multilocus sequence typing (MLST)

Amplification of internal fragments of the seven housekeeping genes – *aroE*, *gdh*, *gki, recP*, *spi*, *xpt*, and *ddl* - was done according to the MLST scheme developed by Enright and Spratt for *S. pneumoniae* (Enright and Spratt, 1998). Sequencing analysis was done with DNAStar (Lasergene). Allele number assignment was done at the international MLST database for *S. pneumoniae* (www.mlst.net).

2.11. Multilocus sequence analysis (MLSA) for viridans group streptococci

Amplification of internal fragments of the seven housekeeping genes – *map*, *pfl*, *ppaC*, *pyk*, *rpoB*, *sodA*, and *tuf* – was done according to the scheme developed and validated for viridans group streptococci by Bishop et al., except for the primer sodA-dn, that had an R to Y substitution (5'-A**Y**RTARTAMGCRTGYTCCCARACRTC-3') based on published sequences of strains TIGR4 (*S. pneumoniae*, accession number AE005672.3), B6 (*S. mitis*, accession number NC_013853.1), and IS7493 (*S. pseudopneumoniae*, accession number

CP002925.1) (Bishop, et al., 2009). Phylogenetic analysis of the concatenated sequences of strains analyzed in this study and the ones deposited at the eMLSA database (427 strains of the viridans group of *Streptococcus*; http://www.emlsa.net/) was performed using MEGA6.06 (http://www.megasoftware.net/): sequences were aligned by ClustalW using default parameters (gap opening penalty of 15 and gap extension penalty of 6.66 for both pairwise and multiple alignment; IUB as DNA weight matrix, with a transition weight of 0.5). A minimum-evolution phylogenetic tree was constructed using default parameters (maximum composite likelihood was used as the substitution model of nucleotides, with transitions and transversions as the substitutions to include; uniform rates among sites; homogeneous pattern among lineages; complete deletion of gaps and missing data and close-neighbor-interchange was used as the ME heuristic method, based on an initial tree by neighbor-joining). Different concatenated sequences of the strains analyzed were arbitrarily named viridans MLSA profiles 1 to 10 and species assignment was inferred based on clustering analysis of the study isolates with the strains from the MLSA database.

3. Results

3.1. Phenotypic and genotypic characterization of the strains

Characteristics of the strains studied are described in Table 1, which also shows a summary of all results. All strains grew in gentamycin blood agar and displayed pneumococcus-like colony morphology albeit some were optochin resistant. The presence of *cpsA* or of other capsular genes (screened by PCR serotyping as described in the Material and Methods section) could not be detected in any of the strains, suggesting the absence of a pneumococcal capsule.

3.2. Atypical and novel *lytA*-BsaAI-RFLP patterns found in pneumococci

One pneumococcal strain (GRA218B) displayed an atypical *lytA*-BsaAI-RFLP pattern, usually associated with non-pneumococci (Fig. 1) (Llull, et al., 2006). This strain was both optochin susceptible and bile soluble, belonged to ST8073, and clustered with pneumococci

by MLSA (Table 1, Fig. 2). Sequence analysis of the *lytA* gene confirmed a 98% similarity of the amino acid sequence with the LytA homologue of ATCC BAA-960, the *S. pseudopneumoniae* strain used as reference (Fig. S2).

Three other pneumococcal strains (GRA036B, Spain6220, and Spain7582) displayed a *lytA*-BsaAI-RFLP pattern not previously described (pattern A, Fig. 1). These strains were both optochin susceptible and bile soluble. Also, these strains belonged to ST508 and clustered with pneumococci by MLSA (Table 1, Fig. 2). Such similarities between these strains suggest that they belong to a common lineage that may be disseminated. Sequence analysis of the *lytA* gene revealed a 60bp deletion at 724bp, resulting in the loss of 20 amino acids (KKIAEKWYYFDGEGAMKTGW). Apart from this deletion, the LytA amino acid sequence shared a 99% similarity with that of TIGR4, the pneumococcal strain used as reference (Fig. S2).

3.3. Typical and novel *lytA*-BsaAI-RFLP patterns found in non-pneumococci

Two non-pneumococcal strains (Spain2270 and Spain9880) displayed a typical *lytA*-BsaAI-RFLP pattern, usually associated with pneumococci (Fig. 1) (Llull, et al., 2006). These strains were optochin resistant in CO_2 atmosphere, optochin susceptible in aerobic conditions, and bile soluble. These strains could not be assigned to an ST according to the *S. pneumoniae* MLST database and clustered with *S. pseudopneumoniae* by MLSA (Table 1, Fig. 2). Such similarities between these strains suggest that they belong to a common lineage. Sequence analysis of the *lytA* gene confirmed a 99% similarity of the amino acid sequence with the LytA of TIGR4, the pneumococcal strain used as reference (Fig. S2).

Five other non-pneumococcal strains displayed two *lytA*-BsaAI-RFLP patterns not previously described (patterns B and C, Fig. 1). All these strains were bile soluble and while one was susceptible to optochin, the others were resistant. None of the strains could be assigned to an ST according to the *S. pneumoniae* MLST database and they all clustered with *S. mitis*

by MLSA (Table 1, Fig. 2). Sequence analysis of the *lytA* gene of strains displaying pattern B (GRA254A, PT8543, PT8638 and PT9018) revealed a silent C329T mutation resulting in an extra BsaAI cutting site. On the other hand, sequence analysis of the *lytA* gene of the strain displaying pattern C (PT8238) revealed a silent C163T mutation resulting in the loss of the BsaAI cutting site. Apart from these mutations, the LytA amino acid sequences of these strains were 97-98% similar to that of the homologue of ATCC BAA-960, the *S. pseudopneumoniae* strain used as reference.

3.4. Misidentifications by real-time PCR targeting *lytA* and *piaA*

Pneumococcal identification by *lytA* real-time PCR failed for three of the 11 strains in this study: the pneumococcal strain harboring the *S. pseudopneumoniae* homologue of the *lytA* gene (no amplification) and the two *S. pseudopneumoniae* strains harboring (the pneumococcal) *lytA* (Ct 24 for both, Table 1). For all other cases described, the mutations found were located outside the annealing sites of the real-time PCR primers and probe and no misidentifications by real-time PCR were found (Fig. S2).

However, when pneumococcal identification by real-time PCR targeting *lytA* was complemented by real-time PCR targeting *piaA*, only the pneumococcal strain harboring the *S. pseudopneumoniae* homologue of the *lytA* gene remained misidentified, as this strain was also negative for *piaA* (Table 1).

4. Discussion

In this study we aimed to determine the genetic basis for the unexpected results given by *lytA*-BsaAI-RFLP and to investigate the accuracy of the *lytA* real-time PCR to classify 11 α-hemolytic streptococcal isolates displaying conflicting or novel results when characterized by the combination of optochin susceptibility, bile solubility, *lytA*-BsaAI-RFLP, and MLST. The main findings of our work were: (i) the identification of one pneumococcal carriage strain harboring a non-pneumococcal homologue of *lytA*; (ii) the identification of two invasive

(cerebrospinal fluid and sputum) *S. pseudopneumoniae* strains harboring the characteristic pneumococcal *lytA*; and (iii) the misidentification of these three strains just referred to above by the commonly used *lytA* real-time PCR. In addition, novel *lytA*-BsaAI-RFLP patterns were identified and these were due to deletions or point mutations in *lytA*.

This is, to the best of our knowledge, the first report of a pneumococcal isolate harboring a non-pneumococcal homologue of *lytA.* Isolates with similar properties may have passed undetected in culture-independent studies relying solely on detection of *lytA* to identify pneumococci. On the other hand, non-pneumococcal isolates originating false positive results by the *lytA* real time PCR strategy have been described (Carvalho, et al., 2013). In particular, two copies of *lytA* (one copy of pneumococcal *lytA* and one copy of the non-pneumococcal homologue) were described in the genome of a clinical isolate (IS7493) of *S. pseudopneumoniae* (Shahinas, et al., 2011). However, *lytA* real-time PCR amplification of strain IS7493 did not retrieve a positive result (Tavares et al., unpublished).

The use of *piaA* real-time PCR to complement *lytA* real-time PCR has been proposed as *piaA* has been described as a pneumococcal specific non-ubiquitous gene that appears to be present in the majority of pneumococcal isolates (a notable exception are some non-encapsulated pneumococci) (Tavares, et al., 2015; Trzcinski, et al., 2013; Whalan, et al., 2006). Whalan et al. found this gene in all encapsulated pneumococci tested (39 isolates covering 27 serotypes) and in six out of eight (75%) non-typeable pneumococci (Whalan, et al., 2006). However, in a recent study of non-typeable pneumococci circulating in Portugal, we have only detected *piaA* in 12 out of 35 (34%) non-typeable pneumococci suggesting absence of this gene is common among these strains (Tavares, et al., 2015). Nevertheless, the strategy of targeting *piaA* in addition to *lytA* clearly enhances the specificity of pneumococcal identification (Trzcinski, et al., 2013). Even so, in the present study, one pneumococcal strain could not be identified based on these two real-time PCR assays.

The occurrence of genetic exchange between oral *Streptococcus* species has been well documented and horizontal gene transfer has been suggested as an important attenuator of putative species barriers (Chi, et al., 2007; Denapaite, et al., 2010; Donati, et al., 2010; Hanage, et al., 2005b; Johnston, et al., 2010). In fact, a smooth transition between pneumococcus species and its close relatives has been proposed (Hakenbeck, et al., 2001). On the other hand, Kilian et al. have proposed that both pneumococcus and *S. mitis* have evolved divergently from a pathogenic common ancestral: while pneumococcus maintained most of the ancestral virulence genes, *S. mitis* evolved to become a commensal (Kilian, et al., 2008, Kilian, et al., 2014).

It is important to highlight that, although exceptions have been reported here and elsewhere (Arbique, et al., 2004; Balsalobre, et al., 2006; Bosshard, et al., 2004; Nunes, et al., 2008; Sá-Leão, et al., 2006; Simões, et al., 2010; Whatmore, et al., 2000), susceptibility to optochin, bile solubility, and assignment of a capsular type are, each one of them, excellent presumptive methods to identify the majority of pneumococcal isolates. In this report, all but one *S. mitis* strain were correctly identified based on optochin susceptibility. For dubious cases, the assignment of specific *lytA*-BsaAI-RFLP signatures and a multiplex PCR strategy have been proposed (Llull, et al., 2006; Simões, et al., 2011). However, this study suggests that with these methods, although rarely, misidentification can still occur. The assignment of a sequence type by the *S. pneumoniae* MLST scheme or the viridans MLSA scheme, although very useful as tools for species identification, are time-consuming and expensive for routine laboratories, thus often being used only in selected cases (Bishop, et al., 2009; Hanage, et al., 2005). Alternatively, the determination of a pneumococcal-specific sequence signature of 16S rRNA has also been proposed as an inexpensive identification tool (Scholz, et al., 2012).

One possible limitation of our study is that we did not systematically study the frequency at which isolates with the new or conflicting characteristics described in this study occur in

collections of pneumococcal clinical and carriage isolates. This aim was beyond the scope of this study. Although all isolates were characterized by optochin susceptibility and bile solubility, *lytA*-BsaAI-RFLP was applied only with isolates presumptively identified as pneumococci and for which a serotype could not be assigned. Still, based on our experience, the frequency of such isolates appears to be low. In particular, among the colonization isolates from Portugal (n=1,226) we have systematically performed the *lytA*-BsaAI-RFLP assay for presumptive pneumococcal isolates for which a capsular type could not be assigned (n=99). Of these, four *S. mitis* isolates (described in this study), corresponding to c.a. 4.0% of the tested samples, had unusual *lytA* patterns. In addition, in our collections the unusual pattern A was found in three of 61 (4.9%) *S. pneumoniae* isolates tested by the *lytA*-BsaAI-RFLP.

The future is heading towards automated screenings of unprocessed samples. Currently, *lytA* real-time PCR is the culture-independent methodology of choice and the association with a second gene such as *piaA* as proposed by Trzcinski et al. is an interesting strategy (Carvalho Mda, et al., 2007; Satzke, et al., 2013; Trzcinski, et al., 2013). Whole genome sequencing (WGS) is being increasingly used and an approach combining automated extraction of WGS information with MLST-extended schemes will undoubtedly reveal itself very useful for the unambiguous classification of strains as the ones described in this study (Sabat, et al., 2013).

In conclusion, identification of pneumococci based on *lytA* detection, including real-time PCR, may lead to false results. This is of particular relevance in the increasingly frequent colonization studies relying solely on culture-independent methods targeting *lytA*.

Funding

This work was funded by Fundação para a Ciência e a Tecnologia, Portugal, through grants PTDC/BIA-MIC/64010/2006 and PTDC/BIA-BEC/098289/2008 to RSL,

SFRH/BD/70147/2010 to DAT, SFRH/BD/27325/2006 to ASS, and Pest-OE/EQB/LAO004/2011 to Laboratório Associado de Oeiras. The funding agency had no involvement in the study design, collection, analysis, and interpretation of data, writing of the article, nor in the decision to submit the study for publication.

Conflicts of interest

None to declare.

Acknowledgments

We thank Sónia T. Almeida for assistance with MLST.

References

Arbique JC, Poyart C, Trieu-Cuot P, Quesne G, Carvalho Mda G, Steigerwalt AG, Morey RE, Jackson D, Davidson RJ, Facklam RR (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 doi: 10.1128/JCM.42.10.4686-4696.2004.

Balsalobre L, Hernandez-Madrid A, Llull D, Martin-Galiano AJ, Garcia E, Fenoll A, de la Campa AG (2006) Molecular characterization of disease-associated streptococci of the mitis group that are optochin susceptible. *J Clin Microbiol* 44:4163-4171 doi: 10.1128/JCM.01137-06.

Bosshard PP, Abels S, Altwegg M, Bottger EC, Zbinden R (2004) Comparison of conventional and molecular methods for ientification of aerobic catalase-negative gram-positive cocci in the clinical laboratory. *J Clin Microbiol* 42:2065-2073 doi: 10.1128/JCM42.5.2065-2073.2004.

Bishop CJ, Aanensen DM, Jordan GE, Kilian M, Hanage WP, Spratt BG (2009) Assigning strains to bacterial species via the internet. *BMC Biol* 7:3 doi: 10.1186/1741-7007-7-3.

Bochud PY, Calandra T, Francioli P (1994) Bacteremia due to viridans streptococci in neutropenic patients: a review. *Am J Med* 97:256-264 doi: 0002-9343(94)90009-4

Brito DA, Ramirez M, de Lencastre H (2003) Serotyping *Streptococcus pneumoniae* by multiplex PCR. *J Clin Microbiol* 41:2378-2384 doi: 10.1128/JCM.41.6.2378-2384.2003.

Carvalho Mda G, Tondella ML, McCaustland K, Weidlich L, McGee L, 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 doi: 10.1128/JCM.02498-06.

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

Chi F, Nolte O, Bergmann C, Ip M, Hakenbeck R (2007) Crossing the barrier: evolution and spread of a major class of mosaic *pbp2x* in *Streptococcus pneumoniae*, *S. mitis* and *S. oralis*. *Int J Med Microbiol* 297:503-512 doi: 10.1016/j.ijmm.2007.02.009.

Denapaite D, Bruckner R, Nuhn M, Reichmann P, Henrich B, Maurer P, Schahle Y, Selbmann P, Zimmermann W, Wambutt R, Hakenbeck R (2010) The genome of *Streptococcus mitis* B6--what is a commensal? *PLoS One* 5:e9426 doi: 10.1371/journal.pone.0009426.

Donati C, Hiller NL, Tettelin H, Muzzi A, Croucher NJ, Angiuoli SV, Oggioni M, Hotopp JCD, Hu FZ, Riley DR, Covacci A, Mitchell TJ, Bentley SD, Kilian M, Ehrlich GD, Rappuoli R, Moxon ER, Masignani V (2010) Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Gen Biol* 11:R107 doi: 10.186/gb-2010-11-10-r107.

Douglas CW, Heath J, Hampton KK, Preston FE (1993) Identity of viridans streptococci isolated from cases of infective endocarditis. *J Med Microbiol* 39:179-182 doi: 10.1099/00222615-39-3-179.

Enright MC, Spratt BG (1998) A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* 144:3049-3060 doi: 10.1099/00221287-144-11-3049.

Hanage WP, Kaijalainen T, Herva E, Saukkoriipi A, Syrjanen R, Spratt BG (2005a) Using multilocus sequence data to define the pneumococcus. *J Bacteriol* 187:6223-6230 doi: 10.1128/JB.187.17.6223-6230.2005.

Hanage WP, Fraser C, Spratt BG (2005b) Fuzzy species among recombinogenic bacteria. *BMC Biol* 7:6 doi: 10.1186/1741-7007-3-6.

Hakenbeck R, Balmelle N, Weber B, Gardes C, Keck W, de Saizieu A (2001) Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect Immun* 69:2477-2486 doi: 10.1128/IAI.69.4.2477-2486.2001.

Ioannidou S, Tassios PT, Kotsovili-Tseleni A, Foustoukou M, Legakis NJ, Vatopoulos A (2001) Antibiotic resistance rates and macrolide resistance phenotypes of viridans group streptococci from the oropharynx of healthy Greek children. *Int J Antimicrob Agents* 17:195-201 doi: 10.1016/S0924-8579(00)00338-1.

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

Keith ER, Podmore RG, Anderson TP, Murdoch DR (2006) Characteristics of *Streptococcus pseudopneumoniae* isolated from purulent sputum samples. *J Clin Microbiol* 44:923-927 doi: 10.1128/JCM.44.3.923-927.2006.

Kilian M, Poulsen K, Blomqvist T, Havarstein L, Bek-Thomsen M, Tettelin H, Sorensen UBS (2008) Evolution of *Streptococcus pneumoniae* and its close commensal relatives. PLoS One 3:e2683 doi: 10.1371/journal.pone.0002683.

Kilian M, Riley DR, Jensen A, Bruggemann H, Tettelin H (2014) Parallel evolution of *Streptococcus pneumoniae* and *Streptococcus mitis* to pathogenic and mutualistic lifestyles. *MBio* 5:e01490-01414 doi: 10.1128/mBio.01490-14.

Llull D, Lopez R, Garcia 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 doi: 10.1128/JCM.44.4.1250-1256.2006.

Messmer TO, Whitney CG, Fields BS (1997) Use of polymerase chain reaction to identify pneumococcal infection associated with hemorrhage and shock in two previously healthy young children. *Clin Chem* 43:930-935.

Nunes S, Sá-Leão R, de Lencastre H (2008) Optochin resistance among *Streptococcus pneumoniae* strains colonizing healthy children in Portugal. *J Clin Microbiol* 46:321-324 doi: 10.1128/JCM.02097-07.

Obregon V, Garcia P, Garcia E, Fenoll A, Lopez R, Garcia JL (2002) Molecular peculiarities of the *lytA* gene isolated from clinical pneumococcal strains that are bile insoluble. *J Clin Microbiol* 40:2545-2554 doi: 10.1128/JCM.40.7.2545-2554.2002.

Pai R, Gertz RE, Beall B (2006) Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. *J Clin Microbiol* 44:124-131 doi: 10.1128/JCM.44.1.124-131.2006.

Rolo D, Simões AS, Domenech A, Fenoll A, Linares J, de Lencastre H, Ardanuy C, Sa-Leão R (2013) Disease isolates of *Streptococcus pseudopneumoniae* and non-typeable *S. pneumoniae* presumptively identified as atypical *S. pneumoniae* in Spain. *PLoS One* 8:e57047 doi: 10.1371/journal.pone.0057047.

Romero P, Lopez R, Garcia E (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 doi: 10.1128/JB.186.24.8229-8239.2004.

Rouff K, Whiley RA, Beighton D (2003) *Streptococcus*. In P. R. Murray, E. J. Barron, J. H. Jorgensen, M. A. Pfaller & R. H. Yolken (Eds, *Manual of Clinical Microbiology* (8th ed., pp. 405-421). Washington, D.C.: American Society for Microbiology.

Sá-Leão R, Simões AS, Nunes S, Sousa NG, Frazão N, de Lencastre H (2006) Identification, prevalence and population structure of non-typable *Streptococcus pneumoniae* in carriage samples isolated from preschoolers attending day-care centres. *Microbiology* 152:367-376 doi: 10.1099/mic.0.28596-0.

Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl J, Laurent F, Grundmann H, Friedrich AW (2013) Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill* 18:20380.

Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, Henao-Restrepo AM, Leach AJ, Klugman KP, Porter BD, Sá-Leão R, Scott JA, Nohynek H, O'Brien KL (2013) Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* 32:165-179 doi: 10.1016/j.vaccine.2013.08.062.

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

Shahinas D, Tamber GS, Arya G, Wong A, Lau R, Jamieson F, Ma JH, Alexander DC, Low DE, Pillai DR (2011) Whole-genome sequence of *Streptococcus pseudopneumoniae* isolate IS7493. *J Bacteriol* 193:6102-6103 doi: 10.1128/JB.06075-11.

Simões AS, Sá-Leão R, Eleveld MJ, Tavares DA, Carriço JA, Bootsma HJ, Hermans PW (2010) Highly penicillin-resistant multidrug-resistant pneumococcus-like strains colonizing children in Oeiras, Portugal: genomic characteristics and implications for surveillance. *J Clin Microbiol* 48:238-246 doi: 10.1128/JCM.01313-09.

Simões AS, Valente C, de Lencastre H, Sá-Leão R (2011) Rapid identification of noncapsulated Streptococcus pneumoniae in nasopharyngeal samples allowing detection of co-colonization and reevaluation of prevalence. *Diagn Microbiol Infect Dis* 71:208-216 doi: 10.1016/j.diagmicrobio.2011.07.009.

Tavares DA, Simões AS, Bootsma HJ, Hermans PW, de Lencastre H, Sá-Leão R (2015) Non-typeable pneumococci circulating in Portugal are of *cps* type NCC2 and have genomic features typical of encapsulated isolates. *BMC Genomics* 15:863 doi: 10.1186/1471-2164-15-863.

Trzcinski K, Bogaert D, Wyllie A, Chu ML, van der Ende A, Bruin JP, van den Dobbelsteen G, Veenhoven RH, Sanders EA (2013) Superiority of trans-oral over trans-nasal sampling in detecting *Streptococcus pneumoniae* colonization in adults. *PLoS One* 8:e60520 doi: 10.1371/journal.pone.0060520.

Wester CW, Ariga D, Nathan C, Rice TW, Pulvirenti J, Patel R, Kocka F, Ortiz J, Weinstein RA (2002) Possible overestimation of penicillin resistant *Streptococcus pneumoniae* colonization rates due to misidentification of oropharyngeal streptococci. *Diagn Microbiol Infect Dis* 42:263-268 doi: 10.1016/S0732-8893(01)00358-3.

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 doi: 10.1128/JB.188.3.1031-1038.2006.

Whatmore AM, Efstratiou A, Pickerill 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 doi: 10.1128/IAI.68.3.1374-1382.2000.

Figures

Figure 1. *lytA*-BsaAI-RFLP patterns. (A) Schematic representation of the DNA fragments produced by digestion of *lytA* gene with BsaAI. Solid triangles show the restriction sites of BsaAI. Open triangles represent deletions. (B) Separation of BsaAI-digested fragments by agarose gel electrophoresis.

Figure 2. Genetic relationships of the strains analyzed in this study and *S. pneumoniae*, *S. pseudopneumoniae*, and *S. mitis* strains deposited at the eMLSA database. Orange circle – GRA036B and Spain7582, light green circle – Spain6220, dark green circle – GRA218B, dark blue circle – GRA254A, blue-green circle – PT8543, maroon circle – PT8638, cyan circle – PT9018, gray circle – PT8238, rose circle – Spain 2270, red circle – Spain9880, green triangle – *S. pneumoniae* reference strain TIGR4, purple triangle – *S. pseudopneumoniae* reference strains ATCC-BAA 960 and IS7493, blue triangle – *S. mitis* reference strain DSM12643, no markers – *S. pneumoniae*, *S. pseudopneumoniae*, and *S. mitis* strains deposited at eMLSA database.

Table 1. Properties of strains characterized in this study[a]

Strain	Clinical source	Country and year of isolation	Patient age (years) and gender	Optochin susceptibility		Bile solubility	*S. pneumoniae* MLST pattern (ST)[b]	*lytA*		*piaA* real-time PCR (Ct)	Viridans MLSA classification[d]	
				in 5% CO_2	in O_2			RFLP pattern[c]	Real-time PCR (Ct)		Profile	Species assignment
GRA036B	NP	SP, 2008	69, M	S	S	Pos	13, 8, 65, 1, 60, 16, 6 (508)	A	Pos (23)	Pos (24)	1	*S. pneumoniae*
Spain6220	Sputum	SP, 2002	71, M	S	S	Pos	13, 8, 65, 1, 60, 16, 6 (508)	A	Pos (24)	Pos (27)	2	*S. pneumoniae*
Spain7582	Sputum	SP, 2005	61, M	S	S	Pos	13, 8, 65, 1, 60, 16, 6 (508)	A	Pos (24)	Pos (28)	1	*S. pneumoniae*
GRA218B	NP	SP, 2010	29, F	S	S	Pos	8, 10, 84, 1, 2, 14, 59 (8073)	Atypical	Neg	Neg	3	*S. pneumoniae*
GRA254A	Sputum	BE, 2009	53, M	S	S	Pos	242(99%), 67(97%), 48(97%), 38(97%), 393(95%), 46(97%), nd	B	Neg	Neg	4	*S. mitis*
PT8543	NP	PT, 2011	4, M	R	R	Pos	106(98%), 414(98%), 318(96%), 195(97%), 193(96%), nd, 2(98%)	B	Neg	Neg	5	*S. mitis*
PT8638	NP	PT, 2011	5, M	R	R	Pos	59(97%), 94(98%), 3(96%), 195(97%), 40(95%), nd, 2(98%)	B	Neg	Neg	6	*S. mitis*
PT9018	NP	PT, 2012	4, M	R	R	Pos	242(98%), nd, 51(98%), 80(98%), 313(99%), 179(95%), 107(97%)	B	Neg	Neg	7	*S. mitis*
PT8238	NP	PT, 2011	3, M	R	R	Pos	88(95%), 94(98%), 318(97%), 67(96%), 182(96%), 153(94%), 545(96%)	C	Neg	Neg	8	*S. mitis*
Spain2270	CSF	SP, 2009	64, M	R	S	Pos	103(99%), 381(98%), 94(97%), 37(99%), 389(99%), nd, 447(97%)	Typical	Pos (24)	Neg	9	*S. pseudopneumoniae*
Spain9880	Sputum	SP, 2009	53, M	R	S	Pos	103(99%), 381(98%), 94(97%), 37(99%), 389(99%), nd, 447(97%)	Typical	Pos (24)	Neg	10	*S. pseudopneumoniae*

[a]NP, nasopharynx; CSF, cerebrospinal fluid; SP, Spain; BE, Belgium; PT, Portugal; S, susceptible; R, resistant; pos, positive; neg, negative.

[b]Allelic profile corresponds to the following genes: *aroE, gdh, gki, recP, spi, xpt,* and *ddl.* When a % is indicated it means that the allele of the strain is divergent from all the alleles described at the *S. pneumoniae* MLST database as of April 2015. The information indicates the similarity with the closest match.

[c]Typical, characteristic of pneumococcal *lytA;* atypical, characteristic of non-pneumococcal *lytA*; A, B and C, new patterns described here.

[d]Different concatenated sequences were arbitrarily named viridans MLSA profiles 1 to 10; and species assignment was inferred based on clustering analysis of the study isolates with the strains from the MLSA database (http://www.emlsa.net/).

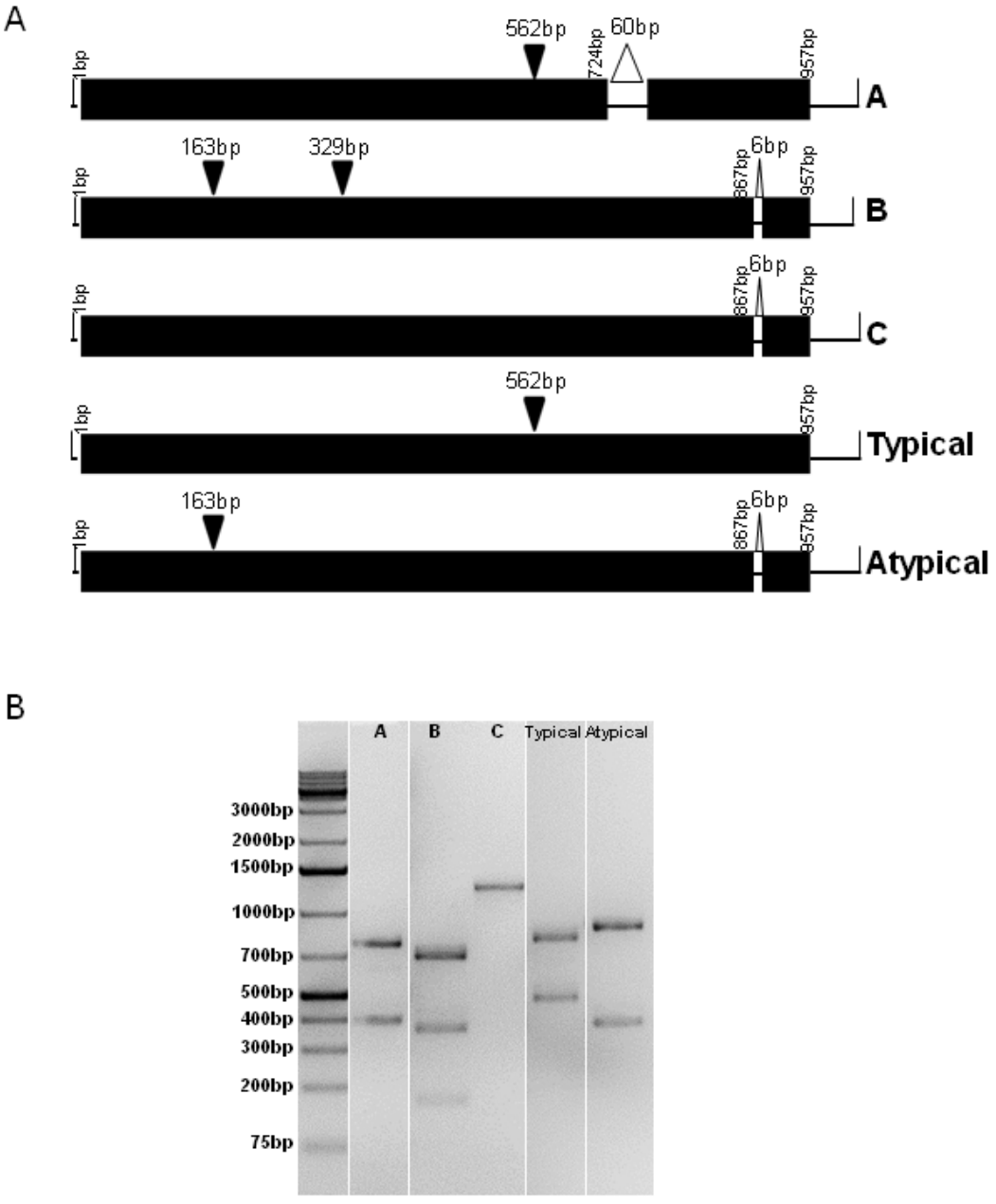


Figure 1

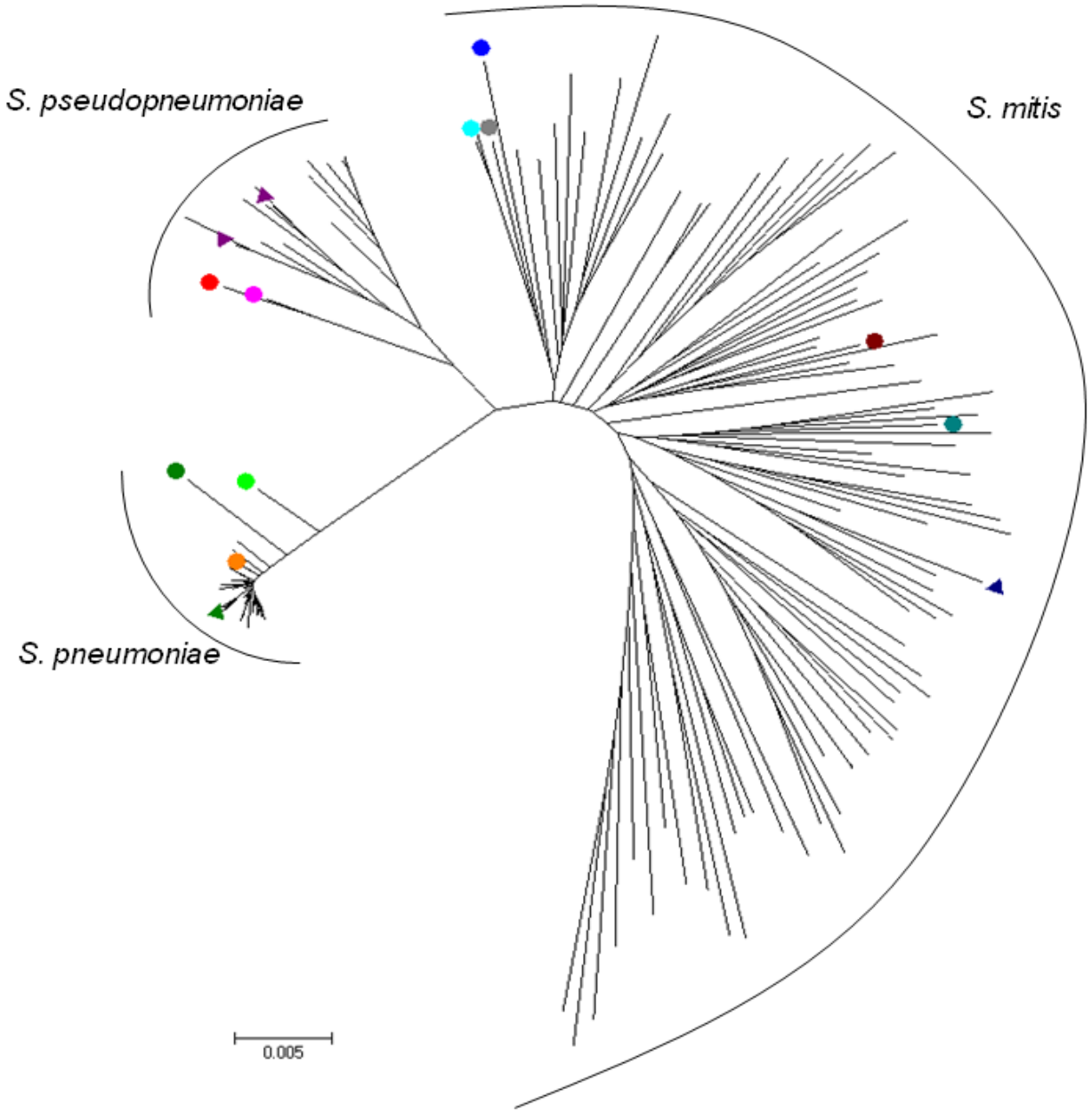
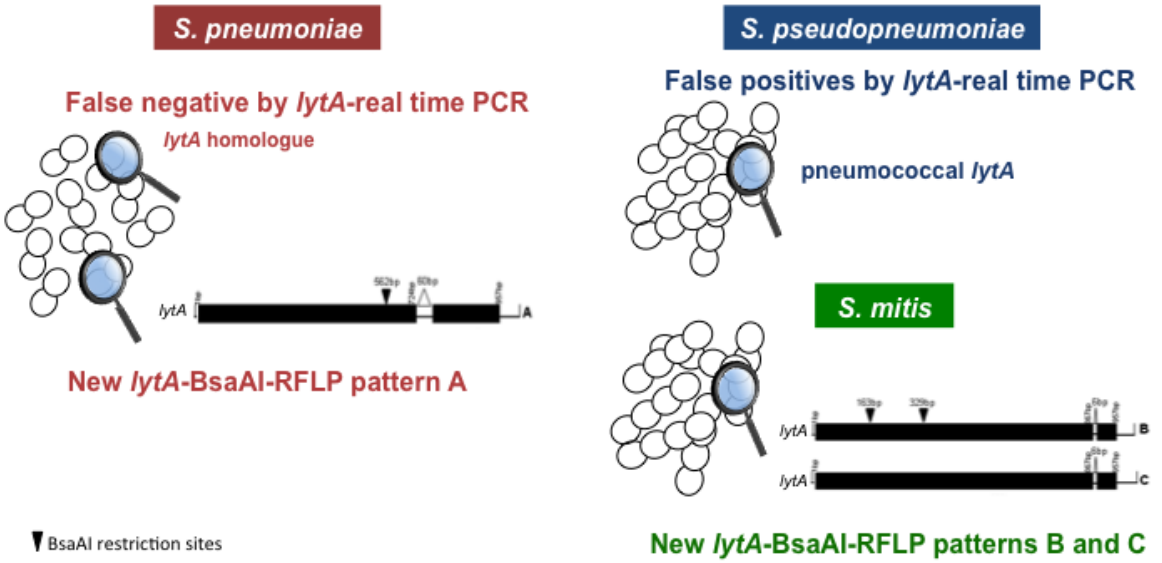


Figure 2



Graphical Abstract

Highlights

- Identification of pneumococci by *lytA*-based methodologies may lead to false results.
- A pneumococcal carriage strain harbouring a *lytA* homologue is described.
- Two invasive *S. pseudopneumoniae* harbouring the pneumococcal *lytA* are described.
- These strains are misidentified by the WHO recommended *lytA*-real time PCR assay.
- Three novel *lytA*-BsaAI-RFLP signatures are described.