Title: Identification of clinically relevant viridans group streptococci by phenotypic and genotypic analysis.

Authors: Cristina Teles¹, Andrew Smith², Gordon Ramage², Sue Lang^{1*}.

Address: Department of Biological and Biomedical Sciences, Glasgow Caledonian University, Glasgow, UK¹. Infection & Immunity Research Group, Glasgow Dental Hospital & School, Faculty of Medicine, University of Glasgow, Glasgow, UK².

Corresponding author: Sue Lang, Department of Biological and Biomedical Sciences, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA. Tel: +44 (0) 141 331 8092. Fax: +44 (0) 141 331 3208. Email: sue.lang@gcu.ac.uk.

Running title: Viridans groups streptococci species identification.

Category: Original article

ABSTRACT

Purpose: Two phenotypic and three molecular methods were assessed for their ability to identify viridans group streptococci (VGS) to the species level.

Methods: A panel of 23 clinical isolates, comprising strains isolated from infective endocarditis, blood cultures, pleural and peritoneal fluid, and 19 type/reference strains were analyzed. Identification was performed using two conventional phenotypic methods, API[®] rapid ID 32 Strep and VITEK[®] 2 system; and genotypic analysis of the nucleotide sequence of the housekeeping gene *sodA*, restriction patterns generated by restriction fragment length polymorphism (RFLP) of the 16S rRNA gene and multilocus sequence analysis (MLSA) of 7 housekeeping genes.

Results: The API[®] rapid ID 32 Strep accurately speciated 79% of the strains assessed, while the VITEK[®] 2 generated a successful identification for 55%, presenting limitations particularly with regard to species belonging to the mitis group. RFLP of the 16S rRNA gene correctly speciated 24% of the strains, having failed to allocate a species for 36% of the isolates examined. In contrast, sequence analysis of the *sodA* gene provided a correct identification for 95% of the strains assessed, while identification using the MLSA technique was unsuccessful due to practical limitations.

Conclusions: The results generated herein indicate that no single methodology can be used to provide an accurate identification to the species level of all VGS although, nucleotide sequence analysis of the *sodA* gene proved to be useful in providing reliable speciation.

Keywords: *Streptococcus*, MLSA, API, VITEK, RFLP, *sodA*

INTRODUCTION

Viridans group streptococci (VGS) encompass four phylogenetic clusters: the mitis, mutans, salivarius and anginosus groups, which are part of the natural human flora, being isolated mainly from the oral cavity, gastrointestinal and genitourinary tract [1]. The pathogenic potential of the species comprising the VGS is variable, with the natural competence of these organisms giving rise to a genetic exchange with *Streptococcus pneumoniae*, a species with recognized virulence traits [2]. It is believed that this interspecies gene transfer is responsible for the emerging tolerance or resistance to β -lactam antimicrobial agents amongst VGS, particularly *S. mitis* [3]. VGS can behave as opportunistic pathogens causing diseases that range from mild infections to life threatening conditions. *S. mutans* for instance, has a strong association with the development of dental caries [4], while *S. sanguinis*, *S. oralis* and *S. mitis* are responsible for 40-60% of native-valve infective endocarditis [5, 6]. Members of the anginosus group are causative agents of pyogenic abscesses in a number of body sites [7]. Other opportunistic infections caused by VGS include septicaemia in neutropenic patients [8], and occasionally cases of meningitis have been reported [9].

Currently, no recognised “gold standard” exists that is able to identify many of the species satisfactorily, which has led to an inadequate reporting of streptococcal infections, and potentially has serious implications for diagnosis and treatment [11-13]. Accurate identification of VGS to the species level is therefore important to characterise the pathogenic potential of individual species and monitor trends in antimicrobial susceptibility and emerging infections, such as septicaemia in immunocompromised individuals [14].

Although presenting recognized limitations, phenotypic tests such as the API[®] rapid ID 32 Strep and the VITEK[®] 2 system remain the techniques most frequently used in clinical laboratories, achieving a speciation on the basis of the organisms biochemical properties [15, 16]. The demand for simple, effective and reliable methods for species identification both in the clinical and research laboratories has resulted in an increasing utilization and development of molecular based techniques [17]. Nucleotide analysis of the *sodA* gene, encoding a manganese-dependent superoxide dismutase enzyme (Mn-SOD), has been proposed as an accurate identification tool [18, 19]. Other genotypic tests previously described include analysis of restriction patterns of the total chromosome (pulsed-field gel electrophoresis) [20, 21] or within specific genes [22]. RFLP of ribosomal RNA encoding genes has been used as an identification method for several bacterial and yeast species including *Leuconostoc*, *Malassezia* and *Streptococcus* [23-25]. More recently, a MLSA technique has been developed for identification of VGS based on the nucleotide sequence analysis of seven housekeeping genes [26]. Using an electronic database comprising the nucleotide sequences acquired from 420 VGS strains, this methodology reportedly allows unknown strains to be assigned to viridans species via the internet [26].

Here we report the applicability of three molecular based techniques, as well as, two commercially available methods, widely used in routine clinical laboratories, in the identification of VGS to the species level.

MATERIALS AND METHODS

Bacterial strains and culture conditions. A collection of 23 clinical isolates, comprising 14 streptococci recovered from patients diagnosed as Duke definite for

infective endocarditis, 7 strains isolated from blood cultures, 1 pleural fluid isolate and 1 peritoneal fluid (Glasgow Caledonian University culture collection, Glasgow Dental School, Barts and The London NHS Trust) were examined (Table 1). For assessment of the accuracy of the identification provided by each method 10 type and 9 reference strains were included (Table 1). The strains designated ATCC and NCTC were obtained from the American Type Culture Collection (Manassas, USA) and National Collection of Type Cultures (London, UK), respectively.

All strains were grown at 37°C in an atmosphere enriched with 5% CO₂ on Columbia base agar (Oxoid, Hampshire, UK) supplemented with 5% (v/v) defibrinated horse blood (Oxoid) and cultivated in brain heart infusion broth (BHI) (Oxoid).

Phenotypic identification. The API[®] rapid ID 32 Strep and VITEK[®] 2 system kits were purchased from bioMérieux, UK. Phenotypic analysis by both methods was performed according to the manufacturer's instructions. The API results were obtained as a numerical profile after visual reading and the identification was achieved by manually entering the code into the identification software (apiweb – API[®] rapid ID 32 Strep V3.0, bioMérieux, France). Identification through the VITEK[®] 2 system was provided by the associated computer software (V1.02), through comparison of the biochemical profile of the strains tested against the biochemical profiles of the strains present in the software database.

Genotypic identification

DNA manipulations. Streptococcal DNA was extracted using the DNeasy[®] Tissue kit (Qiagen, West Sussex, UK) and quantified using a ND-1000 spectrophotometer (NanoDrop Technologies Inc, Wilmington, USA).

Polymerase chain reaction (PCR) analysis. DNA amplifications were performed using an Applied Biosystems 2720 thermal cycler (Applied Biosystems, Foster City, USA). All oligonucleotide primer sequences are depicted in table 2. The PCR reactions consisted of 1× NH₄ amplification buffer, 1.5mM MgCl₂, 1mM dNTP mix, 1μM of each primer, 10ng of genomic DNA and 0.025U/μl *Taq* polymerase (Bioline, London, UK) in a final volume of 50μl.

Amplification of the *sodA* gene was performed on all strains using an annealing temperature of 51.4°C.

S. oralis NCTC 11427^T and *S. oralis* 12601 were selected to assess the MLSA technique for VGS species identification. Genes *map*, *pfl* and *tuf* were amplified with 55°C annealing temperature, while primers for genes *ppaC*, *pyK* and *rpoB* were annealed at 50°C, as previously published [26]. PCR amplifications were performed with 30 cycles of denaturation (94°C, 30 seconds), annealing (X°C, 30 seconds) and extension (72°C, 30 seconds), followed by a final extension step (72°C, 5 minutes) [26]. All generated amplicons were resolved by electrophoresis using a 1.0% (w/v) SeaPlaque[®] agarose gel (Cambrex, Bio Science Rockland, Inc., USA), excised from the gel and purified using the QIAquick[®] gel extraction kit (Qiagen).

The amplification of the 16S rRNA gene was performed with 35 cycles of denaturation (94°C for 1 minute), annealing (60°C for 1 minute) and extension (72°C for 1.5 minutes), followed by a final extension step at 72°C for 10 minutes [22]. The resulting products were resolved by electrophoresis using a 1% (w/v) agarose gel (Invitrogen, Paisley, UK), after which the amplicons were recovered and purified from the PCR reaction mixture using QIAquick[®] PCR Purification Kit (Qiagen).

Sequencing analysis. Nucleotide sequence analysis of the *sodA* gene was performed based on the nucleotide sequence analysis of a single DNA strand of all VGS examined. Evaluation of MLSA was performed by sequencing both DNA strands of the amplified fragments of the *sodA*, *map*, *pfl*, *tuf*, *ppaC*, *pyK* and *rpoB* genes of *S. oralis* NCTC 11427^T and *S. oralis* 12601 (infective endocarditis isolate). Sequencing was performed using 15ng of each amplicon, 3.5μM primer and 50°C annealing temperature or 64μM primer and 45°C annealing temperature by The Sequencing Service (College of Life Sciences, University of Dundee, UK) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer. Species identification was achieved through nucleotide comparison against published sequences available in the microbial genomes database using the Basic Local Alignment Search Tool (BLAST) available online (<http://www.ncbi.nlm.nih.gov/BLAST/>).

RFLP of the 16S rRNA gene. RFLP of the amplified 16S rRNA genes was carried out as described in previous studies [22], with modifications regarding the restriction enzyme digestion conditions. Double restriction enzyme digestions were performed with *Hae*III and *Hpa*II (New England Biolabs, Hitchin, UK) according to the manufacturer's instructions and the resulting restriction fragments were resolved by electrophoresis using a 2% (w/v) agarose gel (Invitrogen). A library of the restriction patterns was produced for the reference strains and identification of the clinical isolates was achieved using Dice similarity coefficient. A dendrogram was then constructed showing the clusters generated by the Unweighted pair group method with arithmetic mean (UPGMA, BioNumerics software, Applied Maths NV, Belgium).

RESULTS

Species identification through API[®] rapid ID 32 Strep system. The API[®] rapid ID 32 Strep system provided an accurate identification to the species level for 33 strains (79%), without the need to perform additional tests. Eight strains (19%) were given a correct identification to the group level, these included 7 strains from the mitis group, which presented a number of unusual reactions, and type strain *S. sobrinus* ATCC 33478^T was given an inconclusive identification, stating *sobrinus/S. downei*. Since the biochemical range of this method is unable to differentiate between these two species further tests were recommended.

Species identification through the Vitek[®] 2 system. The results generated by the Vitek[®] 2 accurately speciated 23 strains (55%) while, 9 isolates of the mitis group were allocated to the correct group (21%). The result of all *S. oralis* and *S. mitis* was presented as *S. mitis/S. oralis* with a high level of confidence, since this methodology does not differentiate between these species. Assignment of *S. sanguinis* reference and type strains to *S. mitis/S. oralis* occurred in two instances (*S. sanguinis* NCTC 7863 and *S. sanguinis* ATCC10556^T), possibly due to a negative reaction in both arginine tests, which is characteristic of *S. mitis* and *S. oralis* strains [27]. *S. parasanguinis* ATCC 15912^T was speciated as *S. gordonii*, presenting a single obvious atypical reaction (methyl- β -D-glucopyranoside positive). No identification was generated for 8 strains (19%) represented in the software database, including isolates from the mitis, salivarius and anginosus group, with 2 (5%) isolates having been misidentified.

Species identification through sequence analysis of the *sodA* gene. The PCR amplification of an internal fragment of the *sodA* gene was successful for all strains,

resulting in a single amplicon with the expected size of approximately 440bp. Sequence analysis of the *sodA* gene resulted in the correct speciation of 40 strains (95%), with *S. sobrinus* ATCC 33478^T having been incorrectly assigned to *S. thermophilus* with 89% nucleotide similarity and failure to generate sequence data occurred for one isolate, despite repeated attempts.

Species identification through RFLP of the 16S rRNA gene. Amplification of the 16S rRNA gene of all VGS strains generated a single amplicon with the expected size of approximately 1500bp. With the similarity threshold set at 90%, analysis of the RFLP revealed 11 clusters incorporating 27 strains (64%), 24% of which were correctly speciated (10/42) and 21% of isolates (9/42) were allocated to the correct VGS group while 19% (8/42) were misidentified (fig. 1). The remaining 15 strains (36%) examined, including species of the mutans, mitis and anginosus group, were excluded from the generated clusters despite the presence of other members of the same species and/or VGS group (fig. 1).

Species identification through MLSA. The PCR amplification of the 7 housekeeping genes was successful for both strains examined, generating single amplicons of the expected sizes. Using the forward primers the nucleotide sequence generated for 3 genes (3/7) of the type strain *S. oralis* NCTC 11427^T and 2 genes (2/7) of the clinical isolate *S. oralis* 12601 did not reach the length required by the MLSA software and both strains failed to yield sequence results using the *ppaC*-F primer. When analyzing the complementary DNA strand using the reverse primers, the *ppaC* gene of *S. oralis* NCTC 11427^T and *S. oralis* 12601 remained below the number of base pairs necessary to proceed with the analysis. In addition, no nucleotide sequence was obtained for the

pfl and *pyk* genes of the type strain and *map*, *pyk*, *rpoB* and *sodA* genes of the clinical isolate. When modifications were introduced to the published protocol, an improvement was seen for 2 of the genes of interest however, this was not beneficial overall as it caused no improvement on either strand of the *ppaC* gene and decreased the number of base pairs acquired for the *sodA* gene. Since the required read length of the 7 housekeeping genes was not obtained for either of the strains examined, no analysis could be performed.

Identification concordance between the methods assessed. Comparison of the identification allocated to each strain by the techniques assessed, revealed a species level agreement between the two phenotypic methods for 18 strains (43%). In contrast, the results generated by the genotypic methods were concordant for 10 isolates (24%). The species allocated by the VITEK[®] 2 and sequence analysis of the *sodA* gene were similar for 21 strains (50%) however, when the identifications produced by the VITEK[®] 2 were compared with those of the RFLP of the 16S rRNA gene, only 6 strains (14%) were allocated to the same species. The API[®] rapid ID 32 Strep and the RFLP of the 16S rRNA gene were concordant for 8 strains (19%) while, the highest levels of agreement were obtained when comparing the API[®] rapid ID 32 Strep and sequence analysis of the *sodA* gene, with 32 isolates (76%) being similarly speciated.

DISCUSSION

The results presented in this study reflect the complex taxonomy of the VGS which has lead to difficulties in defining a “gold standard” technique that would provide accurate identification to the species level. The API[®] rapid ID 32 Strep system remains

one of the most frequently used methodologies in the clinical laboratory for streptococci identification to the species level [15]. Previous reports demonstrated the low sensitivity of this technique for speciation of members of the anginosus and mitis group [6, 28-30], which is in agreement with the results found in the present study as all the discrepant speciations by this method occurred with members of the mitis group. Further limitations of this system, and others of this nature, are associated with the subjectivity inherent with the interpretation of the colorimetric reactions which can interfere with the identification obtained. In addition, the availability of an up-to-date database incorporating the taxonomic re-evaluations of this diverse group of bacteria is crucial in acquiring accurate results [31]. Despite being a simple and non time consuming method, the API[®] rapid ID 32 Strep only provides reliable identification to the group level.

The VITEK[®] 2 system is an automated machine that is capable of identifying numerous bacteria, including VGS, based on kinetic analysis of the fluorescence, turbidity and colorimetric readings of 43 tests measuring enzymatic activities, carbon source and antibiotic resistance [32]. Difficulties in the speciation of VGS members using the VITEK[®] 2 have been reported [17, 33] and the results generated by this technique in the present assessment reinforce these findings in that only 55% of the strains examined were accurately speciated. In addition, the supplementary tests required for two of the most clinically relevant VGS species, *S. mitis* and *S. oralis*, proves to be a major limitation also extending the time required to obtain identification. In agreement with previous reports, the results provided by the VITEK[®] 2 suggest this methodology may be better suited for identification to the species level of organisms other than VGS, such as β -haemolytic streptococci [14].

The sequence analysis of an internal fragment of the *sodA* gene resulted in an accurate speciation of 95% of all strains assessed, revealing the high reliability of this

method for VGS speciation. Traditionally, the sequencing of this housekeeping gene has been carried out by cloning the amplified fragment into a vector prior to sequencing [18, 19, 34]. In this study, elimination of this step made this technique less time-consuming and more cost-effective, without significant consequence to its efficiency. GenBank includes a comprehensive and freely available microbial genomic database which, despite not guaranteeing the accuracy and quality of the data deposited, is an up-to-date database which incorporates the *sodA* nucleotide sequences of VGS reference/type strains and clinical isolates [18, 19, 34]. These factors have contributed to the high levels of reliable identification generated by this methodology.

The sequence analysis of the 16S rRNA gene has been widely applied for speciation of streptococci having greatly contributed to a better understanding of the classification of such a complex bacterial group [35]. The high level of nucleotide similarities between closely related species, particularly *S. oralis*, *S. mitis* and *S. pneumoniae* soon became evident, questioning the discriminatory capability of this methodology [18]. Restriction enzyme digestion of the 16S rRNA gene has been suggested as an alternative method to gene sequence analysis for streptococcal speciation, potentially offering more precise results [22]. The use of a single restriction enzyme (*HpaII*) for discrimination of mutans group species was previously described [22] however, in that study, not all strains presented distinctive patterns requiring a second restriction endonuclease (*HaeIII*) to allow differentiation. Given the number and diversity of strains being assessed in the present evaluation, a simultaneous double restriction reaction was applied. Traditionally, species identification was performed by visual comparison of the restriction patterns resolved by agarose gel electrophoresis however, computer analysis is less subjective and a more practical approach for high throughput of strains [25]. Analysis of the constructed dendrogram demonstrated the

limited capacity of RFLP of the 16S rRNA gene to speciate VGS (24%) and therefore, low concordance with the other methodologies evaluated. This unsatisfactory results may have been influenced by the low resolution, or inability to visually detect, small restriction fragments. Other limitations of this technique have also been reported. These included the inability to differentiate between *S. parasanguinis* and *S. mitis*, despite the use of three separate restriction endonuclease combinations, due to the sequence homology of the 16S rRNA gene shared amongst strains within a VGS group [25]. Based on the findings generated using the panel of strains assessed, RFLP of the 16S rRNA genes was considered unsuitable for streptococcal identification to the species level.

Unlike the MLST technique successfully applied in the genetic differentiation between serotypes of *S. aureus* and *S. pneumoniae*, MLSA has recently been described for VGS species identification based on the sequence data acquired for 7 housekeeping genes [26]. The results obtained in the present study suggest that MLSA of VGS species, at present is not a technique sufficiently robust to allow inter-lab variability and cannot be used routinely.

In conclusion, no individual method assessed was able to unequivocally assign the correct species to all VGS strains. The difficulties in achieving accurate speciation were particularly noticeable in members of the mitis group, which correlates with the genetic homology shared among the species included in this VGS group. Comparison of the data provided indicated that nucleotide sequence analysis of the *sodA* gene is a useful and reliable technique for VGS species identification.

317

318

REFERENCES

- [1] Alam S, Brailsford SR, Whiley RA, Beighton D (1999) PCR-Based methods for genotyping viridans group streptococci. *J Clin Microbiol* 37 (9):2772-2776
- [2] Hakenbeck R (2007) *Molecular Biology of Streptococci*. Horizon Bioscience, Wymondham, U.K.
- [3] Amoroso A, Demares D, Mollerach M, Gutkind G, Coyette J (2001) All detectable high-molecular-mass penicillin-binding proteins are modified in a high-level beta-lactam-resistant clinical isolate of *Streptococcus mitis*. *Antimicrob Agents Chemother* 45 (7):2075-2081
- [4] Oho T, Yamashita Y, Shimazaki Y, Kushiya M, Koga T (2000) Simple and rapid detection of *Streptococcus mutans* and *Streptococcus sobrinus* in human saliva by polymerase chain reaction. *Oral Microbiol Immunol* 15 (4):258-262
- [5] Beynon RP, Bahl VK, Prendergast BD (2006) Infective endocarditis. *Bmj* 333 (7563):334-339
- [6] Chen CC, Teng LJ, Chang TC (2004) Identification of clinically relevant viridans group streptococci by sequence analysis of the 16S-23S ribosomal DNA spacer region. *J Clin Microbiol* 42 (6):2651-2657
- [7] Claridge JE, 3rd, Attorri S, Musher DM, Hebert J, Dunbar S (2001) *Streptococcus intermedius*, *Streptococcus constellatus*, and *Streptococcus anginosus* ("*Streptococcus milleri* group") are of different clinical importance and are not equally associated with abscess. *Clin Infect Dis* 32 (10):1511-1515
- [8] Tunkel AR, Sepkowitz KA (2002) Infections caused by viridans streptococci in patients with neutropenia. *Clin Infect Dis* 34 (11):1524-1529

343 [9] Shenep JL (2000) Viridans-group streptococcal infections in
344 immunocompromised hosts. *Int J Antimicrob Agents* 14 (2):129-135

345 [10] Simmon KE, Hall L, Woods CW, Marco F, Miro JM, Cabell C, Hoen B, Marin
346 M, Utili R, Giannitsioti E, Doco-Lecompte T, Bradley S, Mirrett S, Tambic A, Ryan S,
347 Gordon D, Jones P, Korman T, Wray D, Reller LB, Tripodi MF, Plesiat P, Morris AJ,
348 Lang S, Murdoch DR, Petti CA (2008) Phylogenetic analysis of viridans group
349 streptococci causing endocarditis. *J Clin Microbiol* 46 (9):3087-3090

350 [11] Boudewijns M, Bakkers JM, Sturm PD, Melchers WJ (2006) 16S rRNA gene
351 sequencing and the routine clinical microbiology laboratory: a perfect marriage? *J Clin*
352 *Microbiol* 44 (9):3469-3470

353 [12] Petti CA, Polage CR, Schreckenberger P (2005) The role of 16S rRNA gene
354 sequencing in identification of microorganisms misidentified by conventional methods.
355 *J Clin Microbiol* 43 (12):6123-6125

356 [13] Yuan S, Astion ML, Schapiro J, Limaye AP (2005) Clinical impact associated
357 with corrected results in clinical microbiology testing. *J Clin Microbiol* 43 (5):2188-
358 2193

359 [14] Haanpera M, Jalava J, Huovinen P, Meurman O, Rantakokko-Jalava K (2007)
360 Identification of alpha-hemolytic streptococci by pyrosequencing the 16S rRNA gene
361 and by use of VITEK 2. *J Clin Microbiol* 45 (3):762-770

362 [15] Jayarao BM, Oliver SP, Matthews KR, King SH (1991) Comparative evaluation
363 of Vitek gram-positive identification system and API Rapid Strep system for
364 identification of *Streptococcus* species of bovine origin. *Vet Microbiol* 26 (3):301-308

365 [16] Lupetti A, Barnini S, Castagna B, Capria AL, Nibbering PH (2010) Rapid
366 identification and antimicrobial susceptibility profiling of Gram-positive cocci in blood
367 cultures with the Vitek 2 system. *Eur J Clin Microbiol Infect Dis* 29 (1):89-95

368 [17] Innings A, Krabbe M, Ullberg M, Herrmann B (2005) Identification of 43
 369 *Streptococcus* species by pyrosequencing analysis of the rnpB gene. J Clin Microbiol 43
 370 (12):5983-5991

371 [18] Hoshino T, Fujiwara T, Kilian M (2005) Use of phylogenetic and phenotypic
 372 analyses to identify nonhemolytic streptococci isolated from bacteremic patients. J Clin
 373 Microbiol 43 (12):6073-6085

374 [19] Poyart C, Quesne G, Coulon S, Berche P, Trieu-Cuot P (1998) Identification of
 375 streptococci to species level by sequencing the gene encoding the manganese-dependent
 376 superoxide dismutase. J Clin Microbiol 36 (1):41-47

377 [20] Baseggio N, Mansell PD, Browning JW, Browning GF (1997) Strain
 378 differentiation of isolates of streptococci from bovine mastitis by pulsed-field gel
 379 electrophoresis. Mol Cell Probes 11 (5):349-354

380 [21] Mineyama R, Yoshino S, Maeda N (2006) DNA fingerprinting of isolates of
 381 *Streptococcus mutans* by pulsed-field gel electrophoresis. Microbiol Res

382 [22] Sato T, Hu JP, Ohki K, Yamaura M, Washio J, Matsuyama J, Takahashi N
 383 (2003) Identification of mutans streptococci by restriction fragment length
 384 polymorphism analysis of polymerase chain reaction-amplified 16S ribosomal RNA
 385 genes. Oral Microbiol Immunol 18 (5):323-326

386 [23] Jang J, Kim B, Lee J, Han H (2003) A rapid method for identification of typical
 387 *Leuconostoc* species by 16S rDNA PCR-RFLP analysis. J Microbiol Methods 55
 388 (1):295-302

389 [24] Mirhendi H, Makimura K, Zomorodian K, Yamada T, Sugita T, Yamaguchi H
 390 (2005) A simple PCR-RFLP method for identification and differentiation of 11
 391 *Malassezia* species. J Microbiol Methods 61 (2):281-284

392 [25] Rudney JD, Neuvar EK, Soberay AH (1992) Restriction endonuclease-fragment
393 polymorphisms of oral viridans streptococci, compared by conventional and field-
394 inversion gel electrophoresis. J Dent Res 71 (5):1182-1188

395 [26] Bishop CJ, Aanensen DM, Jordan GE, Kilian M, Hanage WP, Spratt BG (2009)
396 Assigning strains to bacterial species via the internet. BMC Biol 7:3

397 [27] Facklam R (2002) What happened to the streptococci: overview of taxonomic
398 and nomenclature changes. Clin Microbiol Rev 15 (4):613-630

399 [28] Freney J, Bland S, Etienne J, Desmonceaux M, Boeufgras JM, Fleurette J (1992)
400 Description and evaluation of the semiautomated 4-hour rapid ID 32 Strep method for
401 identification of streptococci and members of related genera. J Clin Microbiol 30
402 (10):2657-2661

403 [29] Jacobs JA, Schot CS, Bunschoten AE, Schouls LM (1996) Rapid species
404 identification of "*Streptococcus milleri*" strains by line blot hybridization: identification
405 of a distinct 16S rRNA population closely related to *Streptococcus constellatus*. J Clin
406 Microbiol 34 (7):1717-1721

407 [30] Limia A, Alarcon T, Jimenez ML, Lopez-Brea M (2000) Comparison of three
408 methods for identification of *Streptococcus milleri* group isolates to species level. Eur J
409 Clin Microbiol Infect Dis 19 (2):128-131

410 [31] Bosshard PP, Abels S, Altwegg M, Bottger EC, Zbinden R (2004) Comparison
411 of conventional and molecular methods for identification of aerobic catalase-negative
412 gram-positive cocci in the clinical laboratory. J Clin Microbiol 42 (5):2065-2073

413 [32] Ligozzi M, Bernini C, Bonora MG, De Fatima M, Zuliani J, Fontana R (2002)
414 Evaluation of the VITEK 2 system for identification and antimicrobial susceptibility
415 testing of medically relevant gram-positive cocci. J Clin Microbiol 40 (5):1681-1686

416 [33] Brigante G, Luzzaro F, Bettaccini A, Lombardi G, Meacci F, Pini B, Stefani S,
417 Toniolo A (2006) Use of the Phoenix automated system for identification of
418 *Streptococcus* and *Enterococcus* spp. J Clin Microbiol 44 (9):3263-3267

419 [34] Poyart C, Berche P, Trieu-Cuot P (1995) Characterization of superoxide
420 dismutase genes from gram-positive bacteria by polymerase chain reaction using
421 degenerate primers. FEMS Microbiol Lett 131 (1):41-45

422 [35] Kawamura Y, Hou XG, Sultana F, Miura H, Ezaki T (1995) Determination of
423 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and
424 phylogenetic relationships among members of the genus *Streptococcus*. Int J Syst
425 Bacteriol 45 (2):406-408

426 [36] Whiley RA, Beighton D (1991) Emended descriptions and recognition of
427 *Streptococcus constellatus*, *Streptococcus intermedius*, and *Streptococcus anginosus* as
428 distinct species. Int J Syst Bacteriol 41 (1):1-5

429

430

431

432

433

434 **Table 1** – VGS species assessed in this study and reference strains included for
 435 evaluation of the methods accuracy.

Phylogenetic group	Species	Number of strains	Reference strain included
Mutans	<i>S. mutans</i>	4	<i>S. mutans</i> NCTC 10449 ^T
	<i>S. sobrinus</i>	1	<i>S. sobrinus</i> ATCC 33478 ^T
Salivarius	<i>S. salivarius</i>	6	<i>S. salivarius</i> NCTC 8618 ^T
Anginosus	<i>S. anginosus</i>	4	<i>S. anginosus</i> NCTC 10713 ^T
	<i>S. constellatus</i>	3	<i>S. constellatus</i> subsp <i>constellatus</i> ATCC 27823 ^T
Mitis	<i>S. mitis</i>	3	<i>S. mitis</i> NCTC 12261 ^T
	<i>S. sanguinis</i>	6	<i>S. sanguinis</i> ATCC 10556 ^T
	<i>S. parasanguinis</i>	4	<i>S. parasanguinis</i> ATCC 15912 ^T
	<i>S. gordonii</i>	7	<i>S. gordonii</i> NCTC 7865 ^T
	<i>S. oralis</i>	2	<i>S. oralis</i> NCTC 11427 ^T
	<i>S. cristatus</i>	1	None included
	<i>S. oligofermentans</i>	1	None available

436
 437

438 **Table 2** – PCR primers to amplify *sodA*, 16S rRNA, *map*, *pfl*, *tuf*, *ppaC*, *pyK* and *rpoB*
439 genes.

			Expected	
Primers	Sequence (5' to 3')	Gene amplified	product size (bp)	Origin
sodA-F	T <u>R</u> CAYCATGAY <u>A</u> ARCACCAT*	sodA	436	[18]
sodA-R	A <u>R</u> R <u>T</u> A <u>R</u> TAMGC <u>R</u> TGYTCCCA <u>R</u> AC <u>R</u> TC*			
8UA	AGAGTTTGATCCTGGCTCAG	16S	1521	[22]
1492R	TACGGGTACCTTGTTACGACTT	rRNA		
map-F	GC <u>W</u> GACTC <u>W</u> TGTTGGGC <u>W</u> TATGC	map	459	[26]
map-R	TTA <u>R</u> TAAGTTCY <u>T</u> TCTTC <u>D</u> CCTTG			
pfl-F	AACGTTGCTTACTCTAAACAAACTGG	pfl	458	
pfl-R	ACTTC <u>R</u> TGGAAGACACGTTG <u>W</u> GTC*			
tuf-F	GTTGAAATGGAAATCCGTGACC	tuf	531	
tuf-R	GTTGAAGAATGGAGTGTGACG			
ppaC-F	GACCAY <u>A</u> AATGAATTY <u>C</u> ARCAATC*	ppaC	683	
ppaC-R	TGAGGN <u>A</u> CM <u>A</u> CTTGTTT <u>S</u> TTACG*			
pyk-F	GCGGT <u>W</u> GAA <u>W</u> TCCGTGGTG*	pyK	557	
pyk-R	GCAAG <u>W</u> GCTGGGAAAGGAAT*			
rpoB-F	A <u>A</u> RYTIGG <u>M</u> CCTGAAGAAAT*	rpoB	740	
rpoB-R	TG <u>I</u> A <u>R</u> TTT <u>R</u> TCATCAACCATGTG*			

440 * The underlined degenerate base sites correspond to positions within the gene that may
441 differ between species.
442
443