JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2009, p. 932–939 Vol. 47, No. 4
0095-1137/09/$08.00+0 doi:10.1128/JCM.01449-08

Identification of Clinically Relevant Nonhemolytic Streptococci on the Basis of Sequence Analysis of 16S-23S Intergenic Spacer Region and Partial *gdh* Gene[▽]

Xiaohui Chen Nielsen,[1,2] Ulrik Stenz Justesen,[1,3] Rimtas Dargis,[1] Michael Kemp,[1] and Jens Jørgen Christensen[1]*

Department of Bacteriology, Mycology and Parasitology, Statens Serum Institut, Copenhagen,[1] Department of Clinical Microbiology, Rigshospitalet, Copenhagen,[2] and Department of Clinical Microbiology, Odense University Hospital, Odense,[3] Denmark

Received 29 July 2008/Returned for modification 30 July 2008/Accepted 23 January 2009

Nonhemolytic streptococci (NHS) cause serious infections, such as endocarditis and septicemia. Many conventional phenotypic methods are insufficient for the identification of bacteria in this group to the species level. Genetic analysis has revealed that single-gene analysis is insufficient for the identification of all species in this group of bacteria. The aim of the present study was to establish a method based on sequence analysis of the 16S-23S intergenic spacer (ITS) region and the partial *gdh* gene to identify clinical relevant NHS to the species level. Sequence analysis of the ITS region was performed with 57 NHS reference or clinical strains. Satisfactory identification to the species level was achieved for 14/19 NHS species included in this study on the basis of sequence analysis of the ITS region. *Streptococcus salivarius* and *Streptococcus vestibularis* obtained the expected taxon as the best taxon match, but there was a short maximum score distance to the next best match (distance, <10). *Streptococcus mitis*, *Streptococcus oralis*, and *Streptococcus pneumoniae* could not be unambiguously discriminated by sequence analysis of the ITS region, as was also proven by phylogenetic analysis. These five species could be identified to the group level only by ITS sequence analysis. Partial *gdh* sequence analysis was applied to the 11 *S. oralis* strains, the 11 *S. mitis* strains, and the 17 *S. pneumoniae* strains. All except one strain achieved a satisfactory identification to the species level. A phylogenetic algorithm based on the analysis of partial *gdh* gene sequences revealed three distinct clusters. We suggest that sequence analysis of the combination of the ITS region and the partial *gdh* gene can be used in the reference laboratory for the species-level identification of NHS.

Streptococci are a heterogeneous group of bacteria consisting of more than 50 species. In addition to the traditional pathogenic pneumococci and hemolytic streptococci (HS), many species of non-HS (NHS), which are part of the commensal microbiota in the human body, are known to be opportunistic pathogens that cause serious systemic and local infections. These infections include subacute infective endocarditis (24), bacteremia in immunocompromised patients (19, 44), brain abscesses (34), meningitis (35), and pneumonia (5).

There are reports describing associations between some specific species and clinical manifestations. *Streptococcus gallolyticus* and *Streptococcus lutetiensis* in the bovis group are reported to have a strong association with colon cancer (31–33). *Streptococcus anginosus* and *Streptococcus intermedius* in the anginosus group are associated with abscess formation (9, 26), and *Streptococcus mitis* and other normal oral commensals have a strong association with infective endocarditis (6, 10, 22). The precise species-level identification of NHS from relevant clinical specimens is crucial to making the right diagnosis and understanding the pathogenesis of the infection.

The conventional phenotypic tests do not always allow accurate identification. The automated systems, such as the Rapid ID 32 Strep and the Vitek 2 GP systems (BioMerieux, France), which are based on phenotypic tests, are widely applied in the clinical microbiology laboratory. The large number of species relative to the number of biochemical traits examined, the variability of several traits within species, the poor reproducibility of some tests, and the lack of sufficient phenotypic data for more recently described species in the underlying databases often result in shortcomings with regard to the exact species designations (14, 23, 29, 41). In a study by Hoshino et al. (23) with 148 strains consisting of 115 clinical isolates and 33 reference strains, the rate of correct identification by commercial kits was below 50% but varied significantly between species. The most significant problems were observed with *S. mitis*, *S. oralis*, and the 11 *Streptococcus* species that have been described since 1991. These inherent problems call for alternative means of identification.

Early and effective antimicrobial treatment can result in negative cultures with important clinical specimens, e.g., heart valve tissue or brain abscess material. This stresses the need for the possibility of performing non-culture-based molecular biology examinations (43).

Gene sequence analysis has been applied in an attempt to make an accurate species-level identification of streptococci. The target sequences have included genes encoding functional RNA (the 16S rRNA gene [3], *rnpB* [45]), protein-coding genes (*sod*A [23], *tuf* [37], *groESL* [42], *rpoB* [11], *gdh* and *ddl* [12]), and noncoding spacer regions like the intergenic spacer (ITS) region (7). However, it seems that the analysis of no single gene is sufficient for this very heterogeneous group (23, 40).

* Corresponding author. Mailing address: Department of Bacteriology, Mycology and Parasitology, Statens Serum Institut, Artillerivej 5, Copenhagen 2300, Denmark. Phone and fax: 45 32683268. E-mail: JJC@ssi.dk.

▽ Published ahead of print on 4 February 2009.

Alternatively, the performance of multilocus sequence analysis, as suggested by some studies (23), is not always realistic in the local clinical microbiology laboratory or even in reference laboratories.

The ITS region has been reported to be used for the strain typing of staphylococci and *Clostridium difficile* (15, 16). It is also suggested to be a good candidate for species identification (17, 20, 36). According to Chen et al., the intraspecies similarity scores of 11 viridans group streptococcal species were high and ranged from 0.97 to 1.0, and pairwise comparison of two species of viridans group streptococci revealed a lower level of sequence similarity between their ITS regions than between their 16S rRNA gene sequences (7). These characteristics indicate that the ITS region might constitute a more discriminative target sequence than the 16S rRNA gene for the differentiation of closely related species of NHS.

The *gdh* gene, which is about 1,500 bp in length, encodes a 45-kDa glutamate-6-phosphate dehydrogenase. It is used together with six other housekeeping genes for multilocus sequence typing of pneumococci (23). The *gdh* gene is reportedly highly conserved, as it exhibits an extremely small number of point mutations relative to the numbers in other genes (30). It has been shown that partial *gdh* sequences could be used to unambiguously differentiate *S. pneumoniae* from *S. mitis* and *S. oralis* (25, 30).

The purpose of this study was to establish a method based on ITS and partial *gdh* sequence analysis that is capable of unambiguously identifying clinically relevant NHS to the species level. At the same time, the method should be able to be easily applied in a reference laboratory.

MATERIALS AND METHODS

Bacterial strains. The 68 streptococcal strains used in our study are listed in Table 1. There were 57 strains representing 19 species of NHS and 11 strains representing 5 species of HS. Twenty-three of the strains were obtained from the American Type Culture Collection (ATCC; Manassas, VA), the National Collection of Type Strains (NCTC; London, United Kingdom), or the Culture Collection of the University of Göteborg (CCUG; Göteborg Sweden). The strains with SSI numbers were reference strains from the *Neisseria* and *Streptococcus* Reference Laboratory at the Statens Serum Institut (SSI; Copenhagen, Denmark). The strains with SK numbers were obtained from M. Kilian (Institute of Medical Microbiology and Immunology, University of Aarhus, Aarhus, Denmark). All the strains used in the study were well characterized by conventional phenotypic methods, including microscopy; the evaluation of growth characteristics; performance of the catalase test; evaluation with the Rapid ID32 Strep system; determination of the production of pyrrolidonyl aminopeptidase, leucine aminopeptidase, β-galactosidase, β-*N*-acetylglucosaminidase, β-glucosidase, α-galactosidase, alkaline phosphatase, arginine decarboxylase, urease, extracellular polysaccharide (dextran and levan), and esculetin from esculin; performance of the peroxide test; and the detection of acid production from inulin, salicin, raffinose, amygdalin, and glycogen. The reference strains from SSI, ATCC, NCTC, and CCUG were also identified by 16S rRNA gene sequence analysis at SSI (8). Most of the SK strains were also well identified by phylogenetic analysis of the nucleotide sequences of four housekeeping genes, *ddl*, *sodA*, *gdh*, and *rpoB*, at the Institute of Medical Microbiology and Immunology, University of Aarhus (23).

DNA extraction. The genomic DNA of 35 strains was extracted from the cultures by using a QIAmp DNA minikit (Qiagen, Hilden, Germany), according to the manufacturer's specifications. The genomic DNA of 33 strains was extracted by boiling the culture: one to three colonies of each strain were boiled for 10 min at 95°C in 100 μl PCR-grade water.

PCR primers. To amplify the ITS region, we designed a forward primer, primer Strep16S-1471F (5′-GTG GGA TAG ATG ATT GGG GTG AAG T-3′), the 5′ end of which is located at position 1471 of the 16S rRNA gene (*Escherichia coli* numbering). Reverse primer 6R-IGS (5′-GGG TTC CCC CAT TCG GAH AT-3′) was adapted and improved from the reverse primer of Chen et al. (7). The 5′ end of primer 6R-IGS is located at position 108 downstream of the 5′ end of the 23S rRNA gene (*E. coli* numbering).

To amplify the partial *gdh* gene, we used two primers, primer Strep-gdhF (5′-ATGGACAAACCAGCNAGYTT-3′) and primer Strep-gdhR (5′-GCT TGA GGT CCC ATR CTN CC-3′), which amplify a 660-bp amplicon.

PCR analysis of ITS region and partial *gdh* gene. PCR of the ITS region and subsequent sequence analysis were performed with all 68 streptococcal strains. PCR of the partial *gdh* gene and subsequent sequence analysis were performed only with the 39 strains belonging to the mitis group (11 *S. oralis* strains, 11 *S. mitis* strains, and 17 *S. pneumoniae* strains).

The PCR was performed with 50-μl reaction volumes consisting of 1× HotStarTaq master mix (containing final concentrations of 2.5 U HotStarTaq DNA polymerase, 1× PCR buffer, 200 μM each deoxynucleoside triphosphate; Qiagen) and 0.4 μM (final concentration) each primer. The conditions of the PCR with the primers for the ITS region (primers Strep16S-1471F and 6R-IGS) were as follows: 95°C for 15 min and 40 cycles of 94°C for 30 s, 61°C for 30 s, and 72°C for 30 s. The conditions of the PCR with the primers for the partial *gdh* region were as follows: 95°C for 15 min and 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s. Both PCRs were performed with a DNA Engine Dyad cycler (Bio-Rad). The PCR products were analyzed by 2% gel electrophoresis after staining of the gels with ethidium bromide. The PCR products were purified with spin columns (Microcon YM-100 filter units; Millipore, Billerica, MA).

Sequencing of PCR products. (i) sequencing of ITS region and sequence editing. Both DNA strands of the amplicons were sequenced on an ABI Prism 3100 Avant genetic analyzer (Applied Biosystems, Foster City, CA) with a BigDye (version 3.1) kit (Applied Biosystems). Primers Strep16S-1471F and 6R-IGS were used as sequencing primers.

After sequencing of the PCR product, the sequence had to be edited, as described by Chen et al. (7). Portions of the 16S and 23S rRNA gene regions were removed to obtain the full sequences of the ITS region with CTAAGG at the 5′ ends and AATAA at the 3′ ends of the sequences of the ITS region.

(ii) Sequencing of partial *gdh* sequence and sequence editing. Primers Strep-gdhF and Strep-gdhR were used for the sequencing PCR. The sequences were manually edited as described by M. Kilian (Institute of Medical Microbiology and Immunology, University of Aarhus) so that the sequences were more comparable in the NCBI BLAST search engine. For all the strains evaluated in this study, the 5′ ends of the *gdh* sequences were TTTAAAAACCT, whereas the 3′ ends of the *gdh* sequences were cut just before the sequence AGA ACC ATA C, so that the 3′ end of the edited sequence was TGC TTC/A TCC.

The edited sequences of the ITS region and the partial *gdh* gene were compared to sequences deposited in the NCBI database by using the BLAST search engine and by taking into consideration the percentage and number of identities, the maximum score, and E values for the best and the next best taxon matches.

Phylogenetic analysis. DNA sequences were aligned by using the ClustalW program built into the MEGA (version 4.0) program package. Phylogenetic analysis on the basis of the sequences of the ITS region and the partial *gdh* gene for the 39 strains belonging to the mitis group were performed by the neighbor-joining and minimal evolution methods in the MEGA (version 4.0) program package (downloaded from http://www.megasoftware.net). The distance between the sequences was calculated by using the Kimura two-parameter model.

Nucleotide sequence accession numbers. The GenBank accession numbers of the sequences of the ITS region and the partial *gdh* gene obtained from our study are listed in Table 1.

RESULTS

PCR amplification and determination of sequences of ITS region and partial *gdh* gene. PCR of the ITS region yielded a single band for all strains, and the sizes of the bands varied from 550 bp to 650 bp. The DNA fragment encompassed a small portion of the 16S rRNA gene region, the ITS region, and a small portion of the 23S rRNA gene region.

PCR of the partial *gdh* gene yielded a single band at 660 bp, with the size being constant independent of the species.

Identification of *Streptococcus* strains (HS and NHS) on the basis of sequence analysis of ITS region. Sequence analysis of the ITS region was performed with all 68 streptococcal strains. The sizes of the edited ITS region sequences varied from 248 to 498 bp. All species except for the *S. mitis* and *S. pneumoniae*

TABLE 1. The 68 strains of the genus *Streptococcus* used in the study

Group	Taxon	Strain identifier	GenBank accession no.	
			ITS region	Partial *gdh* gene
Hemolytic group	HS group A	ATCC 51500	EU860331[a]	
	HS group A	ATCC 700294	AE004092[c]	
	HS group A	SSI-SF 130	EU860332[a]	
	HS group B	Group B III, M781	EU860333[a]	
	HS group B	SSI 848	EU860334[a]	
	HS group C	74 Lancefield	EU860335[a]	
	HS group C	SSI 329/04 (*S. zooepidemicus*)	EU860336[a]	
	HS group G	ATCC 2394 D166B	EU860339[a]	
	HS group L	NCTC 10238 SHC	EU860340[a]	
	Streptococcus dysgalactiae	NCTC 4335	EU860341[a]	
	Streptococcus uberis	NCTC 3858	EU860355[a]	
Anginosus group	*Streptococcus anginosus*	SSI 1353/99	EU860342[a]	
	Streptococcus constellatus	NCTC 11325^T	EU860343[a]	
	Streptococcus intermedius	NCTC 11324^T	EU860344[a]	
Sanguinis group	*Streptococcus gordonii*	NCTC 3165	EU860337[a]	
		NCTC 7865^T	EU860346[a]	
	Streptococcus sanguinis	SSI 1655/99	EU860347[a]	
	Streptococcus parasanguinis	SSI 16/03	EU860350[a]	
Mitis group	*Streptococcus mitis*	SSI 992/99	EU860348[a]	EU850792[a]
	Streptococcus mitis	SK320	EU860298[a]	EU850784[a]
	Streptococcus mitis	SK599	EU860300[a]	EU850785[a]
	Streptococcus mitis	SK612	EU860301[a]	EU850786[a]
	Streptococcus mitis	SK614	EU860302[a]	EU850787[a]
	Streptococcus mitis	SK632	EU860303[a]	EU850788[a]
	Streptococcus mitis	SK648	EU860304[a]	EU850789[a]
	Streptococcus mitis	SK661	EU860305[a]	EU850790[a]
	Streptococcus mitis	SK677	EU860306[a]	EU850791[a]
	Streptococcus mitis	SK572	EU860299[a]	AB199460[b]
	Streptococcus mitis	CCUG31611^T	AY347550[b]	EU850793[a]
	Streptococcus oralis	SK100	EU860322[a]	AB199466[b]
	Streptococcus oralis	SK152	EU860323[a]	AB199487[b]
	Streptococcus oralis	SK394	EU860325[a]	AB199488[b]
	Streptococcus oralis	SK555	EU860326[a]	AB199490[b]
	Streptococcus oralis	SK570	EU860327[a]	AB199470[b]
	Streptococcus oralis	SK573	EU860328[a]	AB199471[b]
	Streptococcus oralis	SK580	EU860329[a]	AB199472[b]
	Streptococcus oralis	SSI 220/02	EU860349[a]	EU850795[a]
	Streptococcus oralis	SK155	EU860324[a]	EU850796[a]
	Streptococcus oralis	SK610	EU860330[a]	EU850797[a]
	Streptococcus oralis	CCUG 24891^T	AY347551[b]	AB199448[b]
	Streptococcus pneumoniae	SSISP1/4	EU860357[a]	EU850782[a]
	Streptococcus pneumoniae	SSISP 3/6	EU860358[a]	EU850783[a]
	Streptococcus pneumoniae	SSI6A	EU860319[a]	EU850779[a]
	Streptococcus pneumoniae	SSI6B	EU860318[a]	EU850780[a]
	Streptococcus pneumoniae	SSI14	EU860320[a]	EU850770[a]
	Streptococcus pneumoniae	SSI18A	EU860317[a]	EU850771[a]
	Streptococcus pneumoniae	SSI18B	EU860316[a]	EU850772[a]
	Streptococcus pneumoniae	SSI18C	EU860315[a]	EU850773[a]
	Streptococcus pneumoniae	SSI18F	EU860314[a]	EU850774[a]
	Streptococcus pneumoniae	SSI 19A	EU860313[a]	EU850775[a]
	Streptococcus pneumoniae	SSI 19B	EU860312[a]	EU850776[a]
	Streptococcus pneumoniae	SSI 19C	EU860311[a]	EU850777[a]
	Streptococcus pneumoniae	SSI 19F	EU860310[a]	EU850778[a]
	Streptococcus pneumoniae	ATCC 49619	EU860321[a]	EU850781[a]
	Streptococcus pneumoniae	SSI R6	AE008485[c]	AE008485[c]
	Streptococcus pneumoniae	TIGR4	AE005672[c]	AE005672[c]
	Streptococcus pneumoniae	CCUG28588^T	AY347557[b]	EU860362[a]
Salivarius group	*Streptococcus salivarius*	ATCC 9759	EU860351[a]	
	Streptococcus vestibularis	NCTC 12166^T	EU860352[a]	
	Streptococcus thermophilus	ATCC BAA-250	EU860353[a]	
Mutans group	*Streptococcus sobrinus*	NCTC 10921	DQ204559[b]	
	Streptococcus mutans	Type C 10449	EU860356[a]	
Bovis group	*Streptococcus gallolyticus* subsp. *gallolyticus*	CCUG 35224^T	EU860360[a]	
	Streptococcus infantarius subsp. *infantarius*	CCUG 43820^T	EU860359[a]	
	Streptococcus lutetiensis	CCUG 46149^T	EU860361[a]	
	Streptococcus equinus	ATCC 15351[d]	EU860345[a]	
	Streptococcus equinus	NCTC 10389	EU860338[a]	
Single species	*Streptococcus suis*	R735	EU860354[a]	

[a] Sequences obtained from our study.
[b] Sequences already published in GenBank.
[c] Published whole genome.
[d] This strain was named *S. bovis* earlier.

TABLE 2. Identification of HS and NHS strains on the basis of ITS sequence data[a]

Taxonomic groups and species	No. of strains	Best taxon match	ITS size (bp)	Identity		Maximum score	Maximum score difference from next best taxon match	Identification level
				No. of base pairs	%			
Hemolytic group	11			284–496	97–100	563–971	18–422	Species
S. pyogenes	3	*S. pyogenes*	364–391	391	100	775	365–381	Species
S. agalactiae	2	*S. agalactiae*	295–372	295–372	100	585–735	48–49	Species
S. dysgalactiae	4	*S. dysgalactiae*	301	285–369	97–100	565–648	58–79	Species
S. equi	1	*S. equi*	498	496	99	971	767	Species
S. uberis	1	*S. uberis*	431	331	99	630	422	Species
Anginosus group	3			222–319	96–100	389–632	38–394	Species
S. anginosus	1	*S. anginosus*	292	222	96	389	38	Species
S. constellatus subsp. *constellatus*	1	*S. constellatus*	396	319	100	632	394	Species
S. intermedius	1	*S. intermedius*	347	270	100	535	143	Species
Salivarius group	3			272–350	99–100	533–694	8–36	Species
S. salivarius	1	*S. salivarius*	273	272	99	533	8	Group
S. vestibularis	1	*S. vestibularis*	350	350	100	694	8	Group
S. thermophilus	1	*S. thermophilus*	365	365	100	659	38	Species
Mutans group	2			229–388	99–100	454–755	232–543	Species
S. mutans	1	*S. mutans*	389	388	99	755	543	Species
S. sobrinus	1	*S. sobrinus*	407	229	100	454	232	Species
Bovis group	5			273–274	100	493–495	6–47	Species
S. gallolyticus subsp. *gallolyticus*	1	*S. gallolyticus* subsp. *gallolyticus*	274	274	100	495	6	Species
S. infantarius subsp. *infantarius*	1	*S. infantarius* subsp. *infantarius*	273	273	100	493	47	Species
S. lutetiensis	1	*S. lutetiensis*	273	273	100	493	14	Species
S. equinus	2	*S. equinus*	364	284	100	563	18	Species
Mitis group	39			194–248	99–100	351–492	0–17	Group
S. mitis	7	*S. mitis*	248–249	248–249	99–100	444–448	1–11	Group
S. mitis	4	*S. pneumoniae*	248–249	239–248	100	432–448	5–6	
S. oralis	11	*S. oralis*	246	194–246	99–100	351–472	7–17	Group
S. pneumoniae	17	*S. pneumoniae*	248	247–248	99–100	443–492	0–5	Group
Sanguinis group	4			246–336	98–100	476–628	127–397	Species
S. gordonii	2	*S. gordonii*	323–324	246–247	99–100	476–488	127–133	Species
S. sanguinis	1	*S. sanguinis*	341	294	100	583	397	Species
S. parasanguinis	1	*S. parasanguinis*	341	336	98	628	253	Species
S. suis	1	*S. suis*	419	419	100	831	617	Species

[a] Data are compiled for strains belonging to the same species.

strains achieved the expected taxon as the best taxon match. *S. anginosus* had only a 96% sequence identity. Low maximum score differences from the second best taxon match (difference, <10) were obtained for the *S. salivarius* and *S. vestibularis* strains, although the percent sequence identities were high (99 to 100%), and the next best taxon matches belonged to the salivarius group. Among the 39 strains in the mitis group, all 11 *S. oralis* strains achieved the expected taxon as the best taxon match, although maximum score differences to the next best taxon match (*S. mitis* or *S. pneumoniae*) were relatively low (differences, 7 to 17). Four of the 11 strains of *S. mitis* (strains SK612, SK614, SK648, and SK661) had *S. pneumoniae* as the best taxon match and *S. mitis* as the second best taxon match. However, the maximum score differences between the two taxons were very low (differences, 5 to 6). Six of the 11 *S. mitis* strains and the 17 *S. pneumoniae* strains had the expected taxon as the best taxon match, and the next best taxon matches were *S. pneumoniae* and *S. mitis*, respectively. The maximum score differences between the two best taxon matches were low (differences, 0 to 11) (Table 2).

Identification of strains belonging to the mitis group on the basis of sequence analysis of partial *gdh* gene. Sequence analysis of the partial *gdh* gene was performed with the 39 strains belonging to the mitis group. All of the edited sequences of the partial *gdh* gene were 431 bp in length (Table 3). Among these strains, 38 strains achieved satisfactory identification to the expected taxon with a long maximum score distance to the next best taxon (difference, 18 to 195). Only one *S. mitis* strain (strain SK611) could not be allocated to a single species on the basis of partial *gdh* gene sequence analysis, as the maximum scores for *S. mitis* and *S. pseudopneumoniae* in both cases were 710. Three other *S. mitis* strains (strains SK612, SK614, and SK648) that could not be allocated to the expected taxon on the basis of sequence analysis of the ITS region achieved the correct identification on the basis of the subsequent sequence analysis of the partial *gdh* gene.

TABLE 3. Identifications of the 39 strains of *S. mitis*, *S. pneumoniae*, and *S. oralis* on the basis of the partial *gdh* sequences[a]

Taxonomic species	No. of strains	Best taxon match	Next best taxon match	Identity		Maximum score	Maximum score difference from next best taxon match
				No. of base pairs	%		
S. mitis	10	*S. mitis*	*S. pneumoniae*	410–419	98.5–100	708–722	18–27
S. mitis	1	*S. mitis*	*S. pseudopneumoniae*	410	97	702	0
S. oralis[b]	11	*S. oralis*	*S. mitis*	430–431	99–100	778	36–195
S. pneumoniae	17	*S. pneumoniae*	*S. mitis*	423–434	99–100	764–780	67–91

[a] Data are compiled for strains belonging to the same species. The partial *gdh* gene sequences are all 431 bp in length.
[b] The *gdh* sequences of strains SSI 220/202, SK155, and SK610 and type strain CCUG24891 were achieved in our laboratory. The *gdh* sequences of strains SK100, SK152, SK555, SK394, SK570, SK573, and SK580 were downloaded from M. Kilian's website (www.immi.au.dk/service/download/kilian).

No misidentification was observed on the basis of sequence analysis of the ITS region and the partial *gdh* gene.

Phylogenetic analysis on the basis of the sequences of the ITS region and partial *gdh* gene. The phylogenetic relationship of *S. mitis*, *S. oralis*, and *S. pneumoniae* derived from the sequences of the ITS region is presented in Fig. 1. The 39 strains failed to form any distinct clusters according to the taxon. The evolutionary distances for 38 strains were within 0.002. The distance between the last *S. oralis* strain (strain SK152) to the other strains was only 0.005. Therefore, sequence analysis of the ITS region was insufficient for discrimination of the three species *S. oralis*, *S. pneumoniae*, and *S. mitis* from each other.

The phylogenetic relationship of *S. mitis*, *S. oralis*, and *S. pneumoniae* on the basis of the sequences of the partial *gdh* genes of the 39 strains are presented in Fig. 2. *S. oralis*, *S. pneumoniae*, and *S. mitis* were separated as three distinct clusters, with the type strains represented in each of the distinct clusters. The *S. pneumoniae* and the *S. mitis* clusters had a shorter distance to each other than to the *S. oralis* cluster. The *S. mitis* cluster was much more scattered than the *S. pneumoniae* cluster. The 11 *S. mitis* strains formed three subclusters within the *S. mitis* cluster.

DISCUSSION

We describe a method for species-level identification by combining sequence analysis of the ITS region and sequence analysis of the partial *gdh* gene that is capable of identifying 24 clinically relevant streptococcal species (19 NHS, 5 HS).

In our study, sequence analysis of the ITS region was used as a first-line tool for the identification of species in the genus *Streptococcus*. All 11 strains (100%) belonging to the five HS species achieved the correct species as the best taxon match on the basis of sequence analysis of the ITS region. The maximum score differences from the next best taxon match varied from 18 to 422. These differences are large enough for the differentiation of species. Of the strains belonging to the 19 NHS species, 53/57 strains (94%) achieved the correct species designation as the best taxon match.

Members of the mutans, sanguinis, and anginosus groups achieved unambiguous identifications with high identity scores, and the differences in the maximum scores from the best to the next best taxon match were significant (232 to 543, 127 to 397, and 38 to 394, respectively).

A high degree of heterogeneity within the *S. anginosus* species has been reported previously (27). It was suggested that the species contains several subspecies or new species. In our study, we included only one strain of *S. anginosus*. Therefore, the low percent identity (96%) probably reflects the heterogeneity in this species. Further molecular taxonomic studies are needed to explore this heterogeneity within strains belonging to this species.

The type strain of *S. gallolyticus* subsp. *gallolyticus* achieved the expected taxon as the best taxon match, although the maximum score difference from the next best taxon match, *S. macedonicus*, was only 6. The taxonomy of the bovis group has undergone dynamic changes in the last two decades. Recent studies based on DNA-DNA hybridization, 16S rRNA gene sequencing, and *sodA* gene sequencing revealed that the *S. bovis*/*S. equinus* complex consists of five clusters. *S. gallolyticus* and *S. macedonicus* belong to one cluster. The DNA-DNA hybridization data from the same study revealed that the genomes of *S. macedonicus* and *S. gallolyticus* display >70% homology, which supports the hypothesis that *S. macedonicus* and *S. gallolyticus* are a single species (38). The four other strains belonging to the bovis group included in this study, *S. infantarius* subsp. *infantarius*, *S. lutetiensis*, and *S. equinus*, achieved unambiguous species identification on the basis of sequence analysis of the ITS region. Strain ATCC 15351 was previously named *S. bovis*. The high degrees of similarity by both DNA-DNA hybridization and 16S rRNA gene sequencing brought the conclusion that the species *S. equinus* and *S. bovis* belong to a single species. The name *S. equinus* has nomenclatural priority. Therefore, *S. bovis* is no longer a recognized taxon (39).

Among the members of the salivarius group, only *S. salivarius* is commonly identified from a variety of human infections (4, 13, 21). *S. thermophilus* has been isolated only from dairy products. *S. vestibularis* was identified from the human oral cavity, and its association with human infections has not been confirmed. In our study, both *S. salivarius* and *S. vestibularis* achieved the correct taxon as the best taxon match, although the maximum score distance to the next taxon was only 8. The second best taxon for *S. salivarius* was *S. vestibularis* and vice versa. The failure of *S. salivarius* to produce extracellular polysaccharides on sucrose-containing agar is helpful in securing a correct distinction between these two species.

The differences in maximum scores between the best and the next best taxon matches were very small for the strains belonging to the mitis group (range of differences, 0 to 17), often making it impossible to allocate the strain examined to a specific species. Four strains of *S. mitis* had *S. pneumoniae* as the best taxon match, followed very closely by *S. mitis* (maximum score differences, 5 to 6). The phylogenetic analysis of the

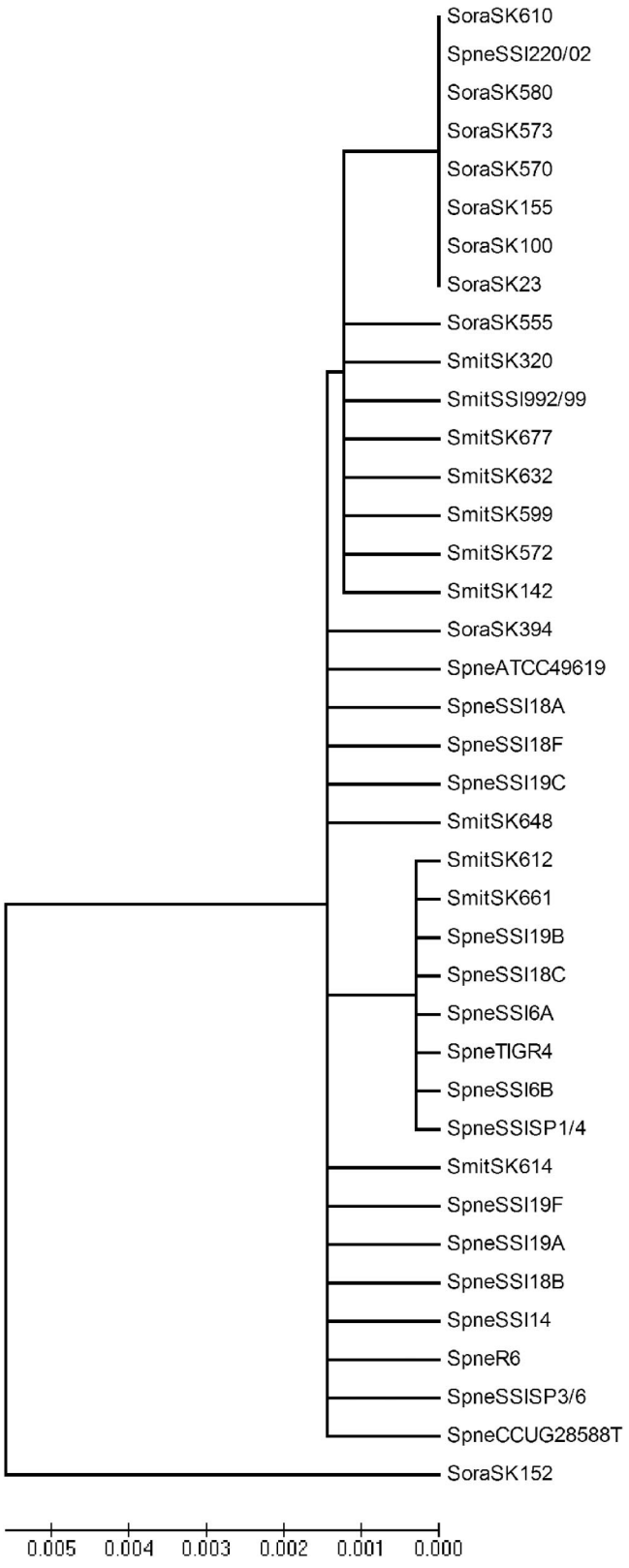

FIG. 1. Phylogenetic tree determined on the basis of the sequences of the ITS regions of 11. *S. mitis* (Smit) strains, 11 *S. oralis* (Soral) strains, and 17 *S. pneumoniae* (Spneu) strains obtained by the unrooted neighbor-joining method in the MEGA (version 4.0) program package. The scale bar indicates the evolutionary distance between the sequences determined by calculation of the percent sequence divergence. It clearly demonstrates that *S. mitis*, *S. pneumoniae*, and *S. oralis* are genetically closely related species and cannot be discriminated from each other on the basis of the sequences of their ITS regions.

results of sequence analysis of the ITS region revealed that *S. oralis* did not form a distinct cluster in relation to *S. mitis* and *S. pneumoniae*. It is well known that members of the mitis group are also closely related on the basis of their 16S rRNA gene sequences (18, 28).

On the basis of these results, we concluded that the ITS region can be used for the species-level identification of strains belonging to the hemolytic, anginosus, mutans, and sanguinis groups. Strains belonging to the salivarius group and the mitis group can be identified only to the group level.

Sequence analysis of the partial *gdh* gene proved to be useful in separating *S. mitis*, *S. oralis*, and *S. pneumoniae*, which are otherwise hard to differentiate from each other. In this study, 17/17 *S. pneumoniae* strains (100%), 11/11 *S. oralis* strains (100%), and 10/11 *S. mitis* strain (90.9%) were correctly identified. Only one *S. mitis* strain could not be discriminated from *S. pseudopneumoniae*. The maximum score difference between these two taxons was 0. This is probably because of the taxonomic changes that have been made in recent years. *S. pseudopneumoniae* is a relatively new taxon that was first described in 2004 and that is closely related to *S. mitis* and *S. pneumoniae* (1). Phylogenetic analysis based on the *sodA* gene sequence showed that the strains assigned to the species *S. pseudopneumoniae* were more closely aligned with *S. mitis* than with *S. pneumoniae* (1). The phylogenetic analysis based on concatenated partial sequences of the *ddl*, *gdh*, *rpoB*, and *sodA* genes showed that *S. pseudopneumoniae* is included within the pneumoniae-mitis-pseudopneumoniae cluster (29). This species was, however, not included in our study.

The three distinct clusters of *S. mitis*, *S. pneumoniae*, and *S. oralis* in the phylogenetic algorithm based on partial *gdh* sequences proved that *gdh* sequence analysis is capable of discriminating these three genetically closely related species from each other. As a result, we have a tool for the species-level identification of these species.

The phylogenetic algorithm also gave us information about the genetic relationship between the three species. There was a much longer distance from the *S. oralis* cluster to the *S. mitis* and *S. pneumoniae* clusters than between the *S. mitis* and *S. pneumoniae* clusters. This suggests that *S. mitis* and *S. pneumoniae* are genetically more closely related to each other than they are to *S. oralis*. The *S. mitis* cluster was much more scattered than the clusters of the other two species, and it formed several subclusters. This suggests that *S. mitis* is genetically more heterogeneous. This is in accordance with the recent observations of Bek-Thomsen et al., who observed that the range of interstrain *gdh* sequence distances was significantly larger for *S. mitis* than for what was found among *S. pneumoniae* strains (2).

In this study, we present a reliable method for the identification of clinically relevant streptococci, with a focus on NHS, to the species level. The method is easy to perform in a laboratory that has sequencing facilities. By sequence analysis of

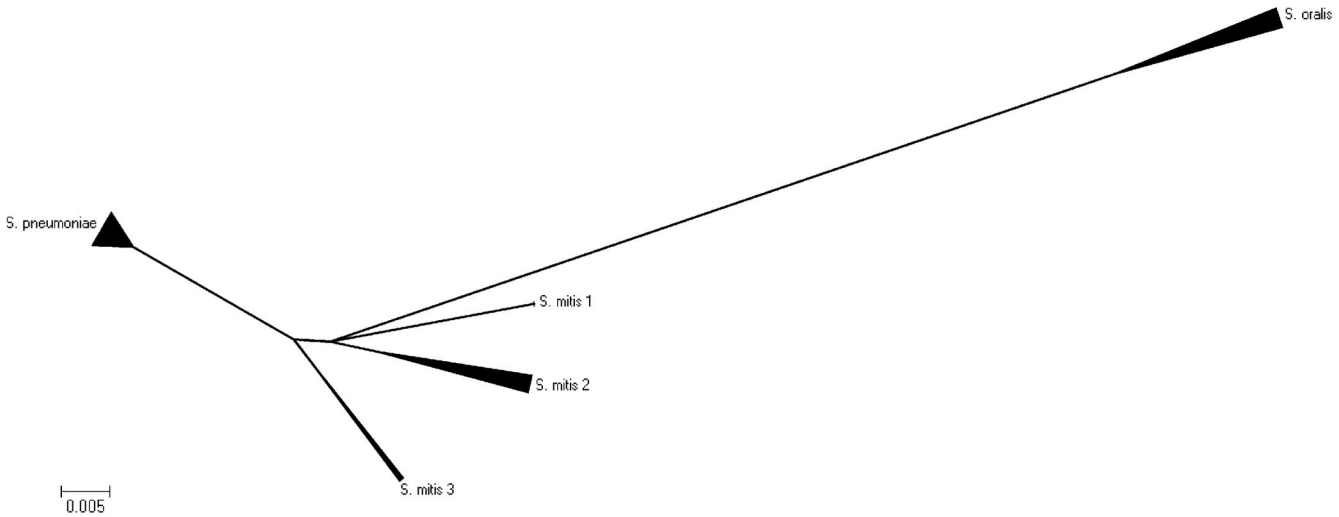

FIG. 2. Minimal evolution algorithm (suppressed) obtained by using the MEGA (version 4.0) program and based on the partial *gdh* gene sequences of 11 *S. oralis strains*, 17 *S. pneumoniae* strains, and 13 *S. mitis* strains. It shows that the three species form three distinct clusters. The *S. oralis* cluster has a longer distance to the two other clusters, indicating that *S. pneumoniae* and *S. mitis* are genetically more closely related on the basis of *gdh* gene evolution. There are three subclusters within the *S. mitis* cluster, indicating that the species *S. mitis* contains a heterogeneous group of strains.

the ITS region, all HS strains and most NHS strains could be identified to the species level. *S. mitis*, *S. pneumoniae*, and *S. oralis* could not be unambiguously discriminated from each other by sequence analysis of the ITS region. A second sequence analysis based on the partial *gdh* gene distinguished these three species from each other. Only one *S. mitis* strain could not be unambiguously discriminated from *S. pseudopneumoniae*, which is probably because of the new nomenclature change. The phylogenetic tree based on the *gdh* gene sequences clearly shows that *S. oralis*, *S. mitis*, and *S. pneumoniae* form three distinct clusters. If colonies are available, sequencing of the ITS region and the partial *gdh* gene can both be completed within 24 to 72 h.

On the basis of the results from this study, we conclude that the combination of sequence analysis of the ITS region and sequence analysis of the partial *gdh* gene is a potential tool for the identification of clinically relevant streptococci in a clinical microbiology reference laboratory.

ACKNOWLEDGMENT

We thank Mogens Kilian, Institute of Medical Microbiology and Immunology, University of Aarhus, for his kind advice in choosing the housekeeping gene *gdh* for sequencing and for supplying us with reference and clinical strains.

REFERENCES

1. **Arbique, J. C., C. Poyart, P. Trieu-Cuot, G. Quesne, M. G. Carvalho, A. G. Steigerwalt, R. E. Morey, D. Jackson, R. J. Davidson, and R. R. Facklam.** 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.
2. **Bek-Thomsen, M., H. Tettelin, I. Hance, K. E. Nelson, and M. Kilian.** 2008. Population diversity and dynamics of *Streptococcus mitis*, *Streptococcus oralis*, and *Streptococcus infantis* in the upper respiratory tracts of adults, determined by a nonculture strategy. Infect. Immun. **76:**1889–1896.
3. **Bosshard, P. P., S. Abels, M. Altwegg, E. C. Bottger, and R. Zbinden.** 2004. Comparison of conventional and molecular methods for identification of aerobic catalase-negative gram-positive cocci in the clinical laboratory. J. Clin. Microbiol. **42:**2065–2073.
4. **Carley, N. H.** 1992. *Streptococcus salivarius* bacteremia and meningitis following upper gastrointestinal endoscopy and cauterization for gastric bleeding. Clin. Infect. Dis. **14:**947–948.
5. **Carratala, J., B. Roson, A. Fernandez-Sevilla, F. Alcaide, and F. Gudiol.** 1998. Bacteremic pneumonia in neutropenic patients with cancer: causes, empirical antibiotic therapy, and outcome. Arch. Intern. Med. **158:**868–872.
6. **Catto, B. A., M. R. Jacobs, and D. M. Shlaes.** 1987. *Streptococcus mitis*. A cause of serious infection in adults. Arch. Intern. Med. **147:**885–888.
7. **Chen, C. C., L. J. Teng, and T. C. Chang.** 2004. Identification of clinically relevant viridans group streptococci by sequence analysis of the 16S-23S ribosomal DNA spacer region. J. Clin. Microbiol. **42:**2651–2657.
8. **Christensen, J. J., R. Dargis, M. S. Kaltoft, K. Andresen, and M. Kemp.** 2006. Ribosomal DNA sequencing of streptococci: usefulness in species identification? Int. Congr. Ser. **1289:**155–158.
9. **Claridge, J. E., III, S. Attorri, D. M. Musher, J. Hebert, and S. Dunbar.** 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:**1511–1515.
10. **Douglas, C. W., J. Heath, K. K. Hampton, and F. E. Preston.** 1993. Identity of viridans streptococci isolated from cases of infective endocarditis. J. Med. Microbiol. **39:**179–182.
11. **Drancourt, M., V. Roux, P. E. Fournier, and D. Raoult.** 2004. *rpoB* gene sequence-based identification of aerobic gram-positive cocci of the genera *Streptococcus*, *Enterococcus*, *Gemella*, *Abiotrophia*, and *Granulicatella*. J. Clin. Microbiol. **42:**497–504.
12. **Garnier, F., G. Gerbaud, P. Courvalin, and M. Galimand.** 1997. Identification of clinically relevant viridans group streptococci to the species level by PCR. J. Clin. Microbiol. **35:**2337–2341.
13. **Gautam, M., K. B. Chopra, D. D. Douglas, R. A. Stewart, and S. Kusne.** 2007. *Streptococcus salivarius* bacteremia and spontaneous bacterial peritonitis in liver transplantation candidates. Liver Transplant. **13:**1582–1588.
14. **Gorm, J. T., K. H. Bossen, and B. Bruun.** 1999. Evaluation of the Rapid ID 32 Strep system. Clin. Microbiol. Infect. **5:**417–423.
15. **Gurtler, V.** 1993. Typing of *Clostridium difficile* strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. J. Gen. Microbiol. **139:**3089–3097.
16. **Gurtler, V., and H. D. Barrie.** 1995. Typing of *Staphylococcus aureus* strains by PCR-amplification of variable-length 16S-23S rDNA spacer regions: characterization of spacer sequences. Microbiology **141**(Pt 5)**:**1255–1265.
17. **Gurtler, V., and V. A. Stanisich.** 1996. New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology **142**(Pt 1)**:**3–16.
18. **Haanpera, M., J. Jalava, P. Huovinen, O. Meurman, and K. Rantakokko-Jalava.** 2007. Identification of alpha-hemolytic streptococci by pyrosequencing the 16S rRNA gene and by use of VITEK 2. J. Clin. Microbiol. **45:**762–770.
19. **Han, X. Y., M. Kamana, and K. V. Rolston.** 2006. Viridans streptococci isolated by culture from blood of cancer patients: clinical and microbiologic analysis of 50 cases. J. Clin. Microbiol. **44:**160–165.
20. **Henry, T., P. C. Iwen, and S. H. Hinrichs.** 2000. Identification of *Aspergillus*

species using internal transcribed spacer regions 1 and 2. J. Clin. Microbiol. **38:**1510–1515.
21. **Hoecker, J. L., L. K. Pickering, D. Groschel, and S. Kohl.** 1978. *Streptococcus salivarius* sepsis in children with malignancies. J. Pediatr. **92:**337–338.
22. **Horaud, T., and F. Delbos.** 1984. Viridans streptococci in infective endocarditis: species distribution and susceptibility to antibiotics. Eur. Heart J. **5**(Suppl. C)**:**39–44.
23. **Hoshino, T., T. Fujiwara, and M. Kilian.** 2005. Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients. J. Clin. Microbiol. **43:**6073–6085.
24. **Hsu, R. B., and F. Y. Lin.** 2006. Effect of penicillin resistance on presentation and outcome of nonenterococcal streptococcal infective endocarditis. Cardiology **105:**234–239.
25. **Ip, M., F. Chi, S. S. Chau, M. Hui, J. Tang, and P. K. Chan.** 2006. Use of the housekeeping genes, *gdh* (*zwf*) and *gki*, in multilocus sequence typing to differentiate *Streptococcus pneumoniae* from *Streptococcus mitis* and *Streptococcus oralis*. Diagn. Microbiol. Infect. Dis. **56:**321–324.
26. **Jacobs, J. A., H. G. Pietersen, E. E. Stobberingh, and P. B. Soeters.** 1995. *Streptococcus anginosus*, *Streptococcus constellatus* and *Streptococcus intermedius*. Clinical relevance, hemolytic and serologic characteristics. Am. J. Clin. Pathol. **104:**547–553.
27. **Jacobs, J. A., C. S. Schot, and L. M. Schouls.** 2000. The *Streptococcus anginosus* species comprises five 16S rRNA ribogroups with different phenotypic characteristics and clinical relevance. Int. J. Syst. Evol. Microbiol. **50**(Pt 3)**:**1073–1079.
28. **Kawamura, Y., X. G. Hou, F. Sultana, H. Miura, and T. Ezaki.** 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. Int. J. Syst. Bacteriol. **45:**406–408.
29. **Kilian, M., K. Poulsen, T. Blomqvist, L. S. Havarstein, M. Bek-Thomsen, H. Tettelin, and U. B. Sorensen.** 2008. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. PLoS One **3:**e2683.
30. **Kiratisin, P., L. Li, P. R. Murray, and S. H. Fischer.** 2005. Use of housekeeping gene sequencing for species identification of viridans streptococci. Diagn. Microbiol. Infect. Dis. **51:**297–301.
31. **Klein, R. S., M. T. Catalano, S. C. Edberg, J. I. Casey, and N. H. Steigbigel.** 1979. *Streptococcus bovis* septicemia and carcinoma of the colon. Ann. Intern. Med. **91:**560–562.
32. **Klein, R. S., R. A. Recco, M. T. Catalano, S. C. Edberg, J. I. Casey, and N. H. Steigbigel.** 1977. Association of *Streptococcus bovis* with carcinoma of the colon. N. Engl. J. Med. **297:**800–802.
33. **Kok, H., R. Jureen, C. Y. Soon, and B. H. Tey.** 2007. Colon cancer presenting as *Streptococcus gallolyticus* infective endocarditis. Singapore Med. J. **48:**e43–e45.
34. **Lu, C. H., W. N. Chang, Y. C. Lin, N. W. Tsai, P. C. Liliang, T. M. Su, C. S. Rau, Y. D. Tsai, C. L. Liang, C. J. Chang, P. Y. Lee, H. W. Chang, and J. J. Wu.** 2002. Bacterial brain abscess: microbiological features, epidemiological trends and therapeutic outcomes. Q. J. Med. **95:**501–509.
35. **Moller, K., E. H. Frederiksen, J. H. Wandall, and P. Skinhoj.** 1999. Meningitis caused by streptococci other than *Streptococcus pneumoniae*: a retrospective clinical study. Scand. J. Infect. Dis. **31:**375–381.
36. **Olsen, R. J., P. A. Cernoch, C. M. Austin, E. A. Graviss, D. H. Farkas, and G. A. Land.** 2007. Validation of the MycoAlign system for *Mycobacterium* spp. identification. Diagn. Microbiol. Infect. Dis. **59:**105–108.
37. **Picard, F. J., D. Ke, D. K. Boudreau, M. Boissinot, A. Huletsky, D. Richard, M. Ouellette, P. H. Roy, and M. G. Bergeron.** 2004. Use of *tuf* sequences for genus-specific PCR detection and phylogenetic analysis of 28 streptococcal species. J. Clin. Microbiol. **42:**3686–3695.
38. **Poyart, C., G. Quesne, and P. Trieu-Cuot.** 2002. Taxonomic dissection of the *Streptococcus bovis* group by analysis of manganese-dependent superoxide dismutase gene (*sodA*) sequences: reclassification of '*Streptococcus infantarius* subsp. *coli*' as *Streptococcus lutetiensis* sp. nov. and of *Streptococcus bovis* biotype 11.2 as *Streptococcus pasteurianus* sp. nov. Int. J. Syst. Evol. Microbiol. **52:**1247–1255.
39. **Schlegel, L., F. Grimont, E. Ageron, P. A. Grimont, and A. Bouvet.** 2003. Reappraisal of the taxonomy of the *Streptococcus bovis*/*Streptococcus equinus* complex and related species: description of *Streptococcus gallolyticus* subsp. *gallolyticus* subsp. nov., *S. gallolyticus* subsp. *macedonicus* subsp. nov. and *S. gallolyticus* subsp. *pasteurianus* subsp. nov. Int. J. Syst. Evol. Microbiol. **53:**631–645.
40. **Simmon, K. E., L. Hall, C. W. Woods, F. Marco, J. M. Miro, C. Cabell, B. Hoen, M. Marin, R. Utili, E. Giannitsioti, T. Doco-Lecompte, S. Bradley, S. Mirrett, A. Tambic, S. Ryan, D. Gordon, P. Jones, T. Korman, D. Wray, L. B. Reller, M. F. Tripodi, P. Plesiat, A. J. Morris, S. Lang, D. R. Murdoch, and C. A. Petti.** 2008. Phylogenetic analysis of viridans group streptococci causing endocarditis. J. Clin. Microbiol. **46:**3087–3090.
41. **Spelleberg, B., and C. Brandt.** 2007. *Streptococcus*, p. 412–429. *In* P. R. Murray, E. J. Baron, J. H. Jorgensen, M. L. Landry, and M. A. Pfaller (ed.), Manual of clinical microbiology, 9th ed. ASM Press, Washington, DC.
42. **Teng, L. J., P. R. Hsueh, J. C. Tsai, P. W. Chen, J. C. Hsu, H. C. Lai, C. N. Lee, and S. W. Ho.** 2002. *groESL* sequence determination, phylogenetic analysis, and species differentiation for viridans group streptococci. J. Clin. Microbiol. **40:**3172–3178.
43. **Tsai, J. C., L. J. Teng, and P. R. Hsueh.** 2008. Direct detection of bacterial pathogens in brain abscesses by polymerase chain reaction amplification and sequencing of partial 16S ribosomal deoxyribonucleic acid fragments. Neurosurgery **62**(Suppl. 2)**:**547–555.
44. **Tunkel, A. R., and K. A. Sepkowitz.** 2002. Infections caused by viridans streptococci in patients with neutropenia. Clin. Infect. Dis. **34:**1524–1529.
45. **Westling, K., I. Julander, P. Ljungman, M. Vondracek, B. Wretlind, and S. Jalal.** 2008. Identification of species of viridans group streptococci in clinical blood culture isolates by sequence analysis of the RNase P RNA gene, *rnpB*. J. Infect. **56:**204–210.