Clinical Microbiology: Open Access
ISSN: 2327-5073

Nielsen et al., Clin Microbial 2013, 2:6
DOI: 10.4172/2327-5073.1000130

Editorial Open Access

16S-23S Intergenic Spacer (ITS) Region Sequence Analysis: Applicability and Usefulness in Identifying Genera and Species Resembling Non-Hemolytic Streptococci

Xiaohui Chen Nielsen[1*], Derya Carkaci[1], Rimtas Dargis[1], Lise Hannecke[1], Ulrik Stenz Justesen[2], Michael Kemp[2], Monja Hammer[3] and Jens Jørgen Christensen[1]

[1]*Department of Clinical Microbiology, Slagelse Hospital, Slagelse, Denmark*
[2]*Department of Clinical Microbiology, Odense University Hospital, Odense, Denmark*
[3]*Department of Microbiology Diagnostics and Virology, Statens Serum Institute, Copenhagen, Denmark*

Abstract

The species that are catalase-negative, gram-positive cocci and not belonging to streptococci or enterococci have become increasingly well characterized and the number of taxonomic entities steadily growing based on molecular taxonomic studies. This complicates their identification. 16S-23S Intergenic Spacer (ITS) region sequence analysis has proven to be a useful tool for species identification of the genera *Streptococcus* and *Enterococcus*. This study investigated the possibility of using ITS sequence analysis as a common tool for species identification within the genera *Aerococcus, Abiotrophia, Alloiococcus, Dolosicoccus, Dolosigranulum, Facklamia, Granulicatella, Gemella*, *Ignavigranum, Leuconostoc*, and *Vagococcus*. ITS sequences of 29 type strains and 103 well-characterized clinical strains were determined and BLAST analysis performed for species identification. All clinical strains were convincingly identified to the species level. Phylogenetic analysis showed distinct clustering of strains with the allocated species and the respective type strains. Thus, ITS sequence analysis was useful for species identification of bacteria belonging to the genera that are catalase-negative and gram-positive cocci. Potentially, ITS could be considered as the first line identification tool for the group of catalase-negative, gram-positive cocci, including non-hemolytic streptococci, enterococci and the taxons examined in this study.

Keywords: Intergenic spacer region; Species identification; Catalase-negative, Gram-positive cocci; *Aerococcus; Gemella; Granulicatella; Abiotrophia; Leuconostoc; Facklamia*

Introduction

Catalase-negative, gram-positive cocci not belonging to streptococci or enterococci represent a group of bacteria which, over the last decades, has become increasingly well characterized. The number of taxonomic entities has been steadily growing, thereby complicating their identification. They resemble the more well-known genera of streptococci and enterococci and consequently may be mistaken as one of those. Therefore, these species often give rise to identification problems and subsequently delayed reporting on final identification to the clinicians [1]. Strains most often recognized belong to the genera *Gemella, Granulicatella*, *Abiotrophia* and *Aerococcus. Leuconostoc, Globicatella, Facklamia*, *Dolosicoccus,* and *Dolosigranulum* are also isolated from blood cultures, though less often [1].

These bacteria are usually part of the normal oral, gastrointestinal and genitourinary flora of humans, and may cause a variety of opportunistic infections. *Gemella*, *Granulicatella* and *Abiotrophia* species are recognized etiologies especially of infective endocarditis and brain abscesses [1,2]. *Aerococcus urinae* are known to cause urinary tract infections, septicaemia and infective endocarditis; the latter with a considerable mortality rate [3-5]. *Leuconostoc* has been isolated from blood cultures, particularly in immune compromised patients [6,7].

As they all may cause serious infections in humans, a rapid and secure identification method is desirable. Precise species identification of blood culture isolates helps identify the primary site of infection and can be a guide to antibiotic susceptibility, thereby having impact on the therapeutic strategy and outcome. Identification of microorganisms from patient samples has in the past mainly been based on phenotypic characteristics exhibited by the putative pathogens, which is time consuming and can produce ambiguous results. In the clinical laboratories, misidentification of *Aerococcus*, *Gemella*, *Granulicatella*, and *Abiotrophia* species as Non-Hemolytic Streptococci (NHS) often occurs. The accuracy of commercial systems (VITEK, API 32 and ATB) commonly used for clinical identification was evaluated by Woo et al. [8] and their results showed frequent misidentification of strains from the genera *Gemella* as *Streptococcus*, *Abiotrophia*, or *Granulicatella*.

For routine identification of clinical bacterial strains, Matrix-Assisted Laser Desorption/Ionization Time of Flight mass spectrometry (MALDI-ToF MS) seems promising for strains belonging to the group of bacteria examined in this study [9]. Sequence based identification methods, especially 16S rRNA gene analysis and detection/sequencing of selected genes, have revolutionized bacteriology in the last 2-3 decades [10]. 16S rRNA gene analysis has been shown to provide relatively good separation of the taxons examined in this study [11]. However, both methodologies are challenged by the closely related species in the Mitis group of NHS [12-14].

A variety of other gene targets including the manganese-dependent superoxide dismutase gene (*sodA)* [15,16], the heat shock protein

***Corresponding author:** Xiaohui Chen Nielsen, MD, Ph.D., Department of Clinical Microbiology, Slagelse Hospital, Ingemannsvej 18, 4200 Slagelse, Denmark, Tel: +4558559479: Fax: +4558559410: E-mail: xcn@regionsjaelland.dk

Received August 23, 2013; **Accepted** September 24, 2013; **Published** September 28, 2013

Citation: Nielsen XC, Carkaci D, Dargis R, Hannecke L, Justesen US, et al. (2013) 16S-23S Intergenic Spacer (ITS) Region Sequence Analysis: Applicability and Usefulness in Identifying Genera and Species Resembling Non-Hemolytic Streptococci. Clin Microbial 2: 130. doi: 10.4172/2327-5073.1000130

Genus and species	Type strain information	GenBank accession number for ITS	ITS region size (bp)	Best taxon match	Second best taxon match	Maximum score	Maximum score differences of the best and second best taxon matches
Genus *Aerococcus*							
A. christensenii	CCUG 28831 ^{T}	JN977130.1*	212	*A. christensenii* ^{T}	*A. urinae* ^{T}	392	229
A. sanguinicola	CCUG 43001 ^{T}	JN977131.1*	210	*A. sanguinicola* ^{T}	*A. suis* ^{T}	388	238
A. suis	CCUG 52530 ^{T}	JN977129.1*	222	*A. suis* ^{T}	*A. sanguinicola* ^{T}	411	261
A. urinae	CCUG 36881 ^{T}	JN977132.1*	207	*A. urinae* ^{T}	*A. christensenii* ^{T}	383	220
A. urinaehominis	CCUG 42038B ^{T}	JN977133.1*	204	*A. urinaehominis* ^{T}	*A. viridans* ^{T}	377	266
A. viridans	CCUG 4311 ^{T}	JN977134.1*	231	*A. viridans* ^{T}	*A. urinaehominis* ^{T}	427	316
Genus *Abiotrophia*							
A. defectiva	CCUG 27639 ^{T}	JQ012763.1*	231	*A. defectiva* ^{T}	*Facklamia hominis* ^{T}	427	328
Genus *Alloiococcus*							
A. otitis	CCUG 32997 ^{T}	JQ012769.1*	227	*A. otitis* ^{T}	*Dolosigranulum pigrum* ^{T}	420	310
Genus *Dolosicoccus*							
D. paucivorans	CCUG 39307 ^{T}	JQ012771.1*	239	*D. paucivorans* ^{T}	*Facklamia sourekii* ^{T}	438	320
Genus *Dolosigranulum*							
D. pigrum	CCUG 33392 ^{T}	JQ012770.1*	253	*D. pigrum* ^{T}	*Alloiococcus otitis* ^{T}	468	358
Genus *Facklamia*							
F. hominis	CCUG 36813 ^{T}	JQ012774.1*	270	*F. hominis* ^{T}	*F. languida* ^{T}	496	281
F. languida	CCUG 37842 ^{T}	JQ012772.1*	230	*F. languida* ^{T}	*F. hominis* ^{T}	422	207
F. miroungae	CCUG 42728 ^{T}	JQ012773.1*	233	*F. miroungae* ^{T}	*F. hominis* ^{T}	431	227
F. sourekii	CCUG 28783A ^{T}	JQ012775.1*	215	*F. sourekii* ^{T}	*Dolosicoccus paucivorans* ^{T}	398	280
Genus *Gemella*							
G. bergeri	CCUG 37817 ^{T}	JQ012764.1*	206	*G. bergeri* ^{T}	*G. morbillorum* ^{T}	372	144
G. cuniculi	CCUG 42726 ^{T}	JQ012767.1*	199	*G. cuniculi* ^{T}	*G. morbillorum* ^{T}	368	142
G. haemolysans	CCUG 37985 ^{T}	JQ697979.1*	205	*G. haemolysans* ^{T}	*G. morbillorum* ^{T}	379	66
G. morbillorum	CCUG 18164 ^{T}	JQ012768.1*	205	*G. morbillorum* ^{T}	*G. haemolysans* ^{T}	379	66
G. palaticanis	CCUG 39489 ^{T}	JQ012766.1*	184	*G. palaticanis* ^{T}	*G. haemolysans* ^{T}	340	192
G. sanguinis	CCUG 37820 ^{T}	JQ012765.1*	199	*G. sanguinis* ^{T}	*G. morbillorum* ^{T}	368	83
Genus *Granulicatella*							
G. adiacens	ATCC 49175 ^{T}	AY353083.2#	216	*G. adiacens* ^{T}	*Pediococcus pentosaceus*	399	310
G. balaenopterae	CCUG 37380 ^{T}	JQ012777.1*	204	*G. balaenopterae* ^{T}	*G. adiacens* ^{T}	377	294
G. elegans	ATCC 700633 ^{T}	JQ012776.1*	198	*G. elegans* ^{T}	*G. adiacens* ^{T}	366	201
Genus *Ignavigranum*							
I. ruoffiae	CCUG 37658 ^{T}	JQ012778.1*	197	*I. ruoffiae* ^{T}	*Facklamia languida* ^{T}	364	306
Genus *Leuconostoc*							
L. citreum	NRIC 1776 ^{T}	AB290437.1#	377	*L. citreum* ^{T}	*L. lactis*	697	177
L. lactis	NRIC 1540 ^{T}	AB290441.1#	370	*L. lactis* ^{T}	*L. carnosum*	680	175
L. mesenteroides ssp. *mesenteroides*	CCUG 30066 ^{T}	JQ012779.1*	371	*L. mesenteroides* ssp. *mesenteroides* ^{T}	*L. cremoris*	686	21
L. pseudomesenteroides	NRIC 1777 ^{T}	AB290443.1#	372	*L. pseudomesenteroides* ^{T}	*L. mesenteroides* ssp. *mesenteroides*	688	152
Genus *Vagococcus*							
V. fluvialis	CCUG 32704 ^{T}	JQ697980.1*	232	*V. fluvialis* ^{T}	*Enterococcus faecium*	429	323

CCUG, Culture Collection of the University of Göteborg, Sweden.
ATCC, American Type Culture Collection, USA.
NRIC, Nodai Research Institute Culture Collection, Japan.
*ITS sequences achieved in this study.
#Previously published ITS sequences.
^{T} Type strains.

Table 1: Strain information and data from the BLAST examination of the ITS regions for type strains belonging to the genera *Aerococcus, Abiotrophia, Alloiococcus, Dolosicoccus, Dolosigranulum, Facklamia, Granulicatella*, *Gemella, Ignavigranum, Leuconostoc*, and *Vagococcus*. GenBank accession numbers (both those obtained from this study and those previously published in GenBank) are listed.

groESL [17,18], ribosomal protein *rpoB* [19], and recombination and repair protein *recN* [20] have also been used for species identification of the genus *Streptococcus* with promising results for most of the species, except for the Mitis group. However, other members of the catalase-negative, gram-positive cocci not belonging to *Streptococcus* and *Enterococcus* have only rarely been investigated [17,19,21].

The ribosomal 16S-23S Intergenic Spacer (ITS) region has been suggested as a good candidate for bacterial identification and strain typing [22,23] In a previous study by our group, the feasibility of using the ITS sequence to identify clinical strains of NHS was established [24] ITS sequence analysis was suggested as a first line identification tool for the NHS group. However, a housekeeping gene, glucose-dehydrogenase (*gdh*), would also have to be analysed in order to safely differentiate between *S. mitis*, *S. oralis* and *S. pneumoniae*. Early and effective antimicrobial treatment can result in negative cultures from 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. Sequence based methods, which can elegantly separate relevant taxons, will be natural candidates for this purpose. ITS sequence analyses have also proven to be useful in species identification of enterococcal strains [23]. For that reason, we found it of interest to expand ITS sequencing to other members of the catalase-negative, gram-positive cocci that resemble NHS morphologically. The purpose of this study was to investigate the possible role of ITS sequence analysis as a common key for the identification of clinical strains of NHS, enterococcal, and the NHS-like taxons examined in this study.

Materials and Methods

Type strains

Twenty five type strains belonging to 11 genera were received from the Culture Collection, University of Göteborg, Sweden (CCUG) (Table 1). Strains were grown and maintained on 5% Danish horse blood agar plates and stored at -80 °C in 10% glycerol broth (Statens Serum Institut, Copenhagen, Denmark). These strains underwent PCR, sequencing and subsequent editing to determine the ITS sequence (see the paragraph about "Sequencing of ITS region and sequence editing" for details). Sequences from four other type strains with published ITS sequences were also included in the study (Table 1).

Clinical strains

A total of 103 clinical strains of gram-positive, catalase-negative cocci were included in this study belonging to the following genera: *Aerococcus* (n=37), *Abiotrophia* (n=9), *Granulicatella* (n=14), *Gemella* (n=30), *Leuconostoc* (n=5) and *Facklamia* (n=8) (Table 2). Among these; 75 strains were from the Reference Laboratory at Statens Serum Institut (SSI), Copenhagen, Denmark. These strains were sent from local departments of clinical microbiology in Denmark for identification, from March 2000 to June 2010. Conventional phenotypic analysis, partial 16S rRNA (a 526 bp stretch) gene sequence analysis [12, 25] and MALDI-ToF MS analysis (Bruker Biotyper, Germany) [9] were performed to characterize these strains. Furthermore, 28 clinical strains were purchased from CCUG and MALDI-ToF MS was performed to confirm identification of these strains. These included *A. defectiva* (n=7), *F. hominis* (n=4), *F. languida* (n=2), *F. sourekii* (n=2),

Genus and species	n	ITS region size (bp)	Best taxon match	Second best taxon match	Maximum score	Maximum score differences of the best and second best taxon matches	Query coverage [%]	Maximum identity [%]	Identification level
Genus Aerococcus									
A. christensenii	3	212-233	*A. christensenii* ^{T}	*A. urinae* ^{T}	370-392	216-229	100	98-100	Species
A. sanguinicola	5	210	*A. sanguinicola* ^{T}	*A. suis* ^{T}	388	238	100	100	Species
A. urinae	27	207-208	*A. urinae* ^{T}	*A. christensenii* ^{T}	377-383	218-220	99-100	99-100	Species
A. viridans	2	228-252	*A. viridans* ^{T}	*A. urinaehominis* ^{T}	363-416	233-308	99-100	95-99	Species
Genus *Abiotrophia*									
A. defectiva	9	229-231	*A. defectiva* ^{T}	*Facklamia hominis* ^{T}	396-417	294-318	99-100	99-100	Species
Genus *Facklamia*									
F. hominis	4	266-269	*F. hominis* ^{T}	*F. languida* ^{T}	453-483	223-263	100	97-99	Species
F. languida	2	230	*F. languida* ^{T}	*F. hominis* ^{T}	399-411	180-198	100	98-99	Species
F. sourekii	2	215	*F. sourekii* ^{T}	*Dolosicoccus paucivorans* ^{T}	389	271	100	100	Species
Genus *Gemella*									
G. bergeri	4	206-209	*G. bergeri* ^{T}	*G. morbillorum* ^{T}	316-381	94-150	99-100	94-100	Species
G. haemolysans	8	202-208	*G. haemolysans* ^{T}	*G. morbillorum* ^{T}	335-368	22-66	100	97-99	Species
G. haemolysans	2	202	*G. haemolysans*	*G. morbillorum*	315	6	100	95	Species
G. morbillorum	11	205-207	*G. morbillorum* ^{T}	*G. haemolysans* ^{T}	315-374	66-88	100	99-100	Species
G. sanguinis	5	198-199	*G. sanguinis* ^{T}	*G. morbillorum* ^{T}	336-368	43-83	100	97-100	Species
Genus Granulicatella									
G. adiacens	9	218	*G. adiacens* ^{T}	*Pediococcus pentosaceus*	379-390	214-298	99	98-99	Species
G. elegans	5	197-198	*G. elegans* ^{T}	*G. adiacens* ^{T}	311-338	144-171	100	96-98	Species
Genus Leuconostoc									
L. citreum	1	375	*L. citreum*	*L. lactis*	693	177	100	100	Species
L. lactis	2	370	*L. lactis* ^{T}	*L. citreum*	658-664	148-165	99	99	Species
L. pseudo-mesenteroides	2	370	*L. pseudo-mesenteroides* ^{T}	*L. mesenteroides*	652	129	99	99	Species

^{T} Type strains.

Table 2: ITS region sequence BLAST analysis results of the 103 clinical strains belonging to the genera *Aerococcus, Abiotrophia, Facklamia, Granulicatella*, *Gemella* and *Leuconostoc*.

G. haemolysans (n=4), *G. sanguinis* (n=4), *G. adiacens* (n=1), and *G. elegans* (n=4).

DNA extraction

The genomic DNA of all strains was extracted by heating one to three colonies of each strain for 10 min at 95 °C in 100 µl PCR-grade water.

PCR amplification of ITS region

To amplify the ITS region, we used primers Strep16S-1471F (5'-GTG GGA TAG ATG ATT GGG GTG AAG T-3') and 6R-IGS (5'-GGG TTC CCC CAT TCG GAH AT-3') as previous described [24]. The PCR was performed with 50 µl reaction volumes consisting of 25 µl of Brilliant II SYBR Green master mix (Agilent Technologies) and 0.5 µM (final concentration) of each primer and 2 µl of the DNA template. The PCR program was: 94°C for 10 minutes followed by 35 cycles of 94°C for 30 seconds, 61°C for 30 seconds, 72°C for 30 seconds. PCR was performed on Mx3005P (Stratagene, Agilent Technologies). The PCR products were analyzed both by real-time amplification and melting curves in the program MxPro (Stratagene, Agilent Technologies) and by capillary electrophoresis system QIAxcel (Qiagen).

Sequencing of ITS region and sequence editing

Amplicons were sequenced at Eurofins MWG Operon (Germany) and GATC Biotech (Germany). The primers Strep16S-1471F and 6R-IGS were used as sequencing primers. Results from sequencing were analyzed with CLC Main Workbench v6. The forward and reverse sequence reads were assembled to obtain the consensus sequence of the ITS regions. The regions belonging to 16S and 23S rRNA genes were removed to obtain full-length sequences of the ITS region with CTAAGG at the 5-prime and TTAAGT/C at the 3-prime ends of the sequences of the ITS region.

Blast

The edited sequences of the ITS regions from both the type and clinical strains were compared to sequences deposited in the NCBI database by using the BLAST search engine (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and by taking into consideration % identity (number of identical bases between the query and the subject sequence in the database), the Maximum score (indication of alignment concordance), and E values (indication of statistical significance of a given alignment) for the best and the second best taxon matches. The Maximum score difference between the best and second best taxon match at a minimum of 10 was used as the criteria for species differentiation.

Phylogenetic analysis

Intraspecies distances were calculated by performing alignment of ITS regions achieved from the clinical strains and the corresponding type strain with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2). The alignment was then used to compute the pairwise distance calculations with Kimura-2-parameter model in the Molecular Evolutionary Genetic Analysis (MEGA) 5.0 program package (http://www.megasoftware.net). Interspecies distances were calculated in the same way with ITS sequences achieved from type strains belonging to the same genus. Phylogenetic analysis on the basis of the sequences of the ITS region for each genus including the type and clinical strains were performed by Neighbour-Joining method (MEGA 5.0). The robustness of the phylogenetic tree was determined with 1000 bootstrap replicates.

Results

Amplification and sequence analysis of the ITS region for type strains and determination of editing sites

One predominant amplicon was achieved from all 25 type strains. The sizes of ITS PCR products varied between 184 and 377 bp. Sequences with the ITS region and part of the 16S and 23S region were generated. Published ITS sequences of four other type strains were downloaded. Alignment of these 25 ITS sequences and earlier published streptococcal ITS sequences [24], revealed the editing sites of 5'- and 3'-ends to be CTAAGG and TTAAGT/C, respectively. The edited ITS sequences of the 25 type strains were submitted to GenBank and accession numbers are listed in Table 1.

Amplification and sequence analysis of the ITS region for the clinical strains

For most of the clinical strains, only one amplicon was detected. For two *A. urinae* strains, more than one amplicon was detected according to the dissociation curve analysis, although in all strains only one product was detected by the QIAxcel capillary electrophoresis system. Figure 1 presents an example with one such *A. urinae* strain. The forward and reverse sequences were assembled and a consensus sequence was achieved for all clinical strains (CLC Main Workbench). All the sequences were edited as described in Materials and Methods. The sizes of the edited ITS sequences varied between 197-375 bp (Table 1).

ITS region sequences BLAST results of the type strains

ITS sequences from all the type strains were sent to BLAST to achieve the knowledge of how the ITS will perform as a candidate target for species identification. All the type strains had achieved their own strains as the first taxon match (data not shown). The Maximum score distance from the first to the second best taxon match varied between 21 and 358 (Table 1).

Species identification based on ITS sequence BLAST

The species distribution and BLAST results of the clinical strains are shown in (Table 2).

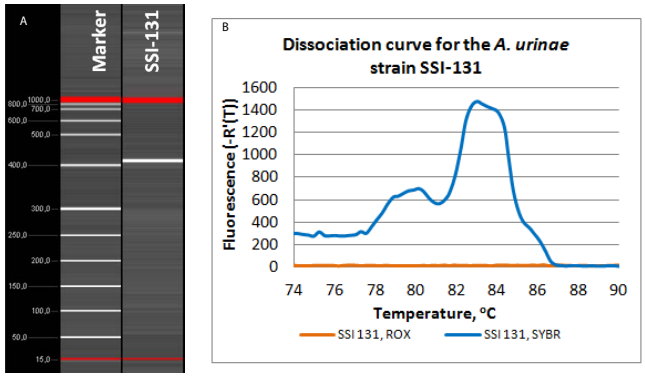


Figure 1: a: QIAxcel electrophoresis image of ITS amplicon of one *A. urinae* strain (SSI-131). This sample only contained a 427 bp amplicon. The red fragments in the beginning and in the final part of each lane are the 15 bp and 1000 bp alignment markers used for the QIAxcel alignment. The QX DNA Size Marker 50-800 bp was used as DNA ladder.
b: Dissociation curve of ITS amplicon from the same strain. Blue curve: SYBR Green fluorescence curve. The major peak had a melting temperature of 82-84 °C, while the minor peak had a melting temperature of 78-80 °C.

For all 37 *Aerococcus* strains, of which *A. urinae* (n=27) dominated in number, best taxon matches were in agreement with the presumed species identifications. The large difference in Maximum score values between the best and second best taxon matches of 216-308 made identifications convincing. Similarly, the best taxon matches obtained for the 36 clinical strains belonging to the genera *Abiotrophia*, *Facklamia*, *Granulicatella* and *Leuconostoc* were in agreement with the presumed species identifications. The difference in Maximum score values between the best and second best taxon matches were large (129-318). For the genus *Gemella*, all clinical strains (n=30) obtained best taxon matches that were in agreement with the presumed species identifications. However, two strains that were designated as *G. haemolysans* only had a difference of six in Maximum score values between the best and the next best taxon match (*G. morbillorum*) (Table 2). In no cases did misidentification occur.

Phylogenetic analysis

Phylogenetic analysis based on sequences of the ITS regions of strains belonging to the genera *Aerococcus* (Figure 2a), *Granulicatella* and *Abiotrophia* (Figure 2b), *Facklamia* (Figure 2c), *Gemella* (Figure 2d), and *Leuconostoc* (Figure 2e), showed distinct clustering of each species with clinical strains and the corresponding type strain. The two ambiguous *G. haemolysans* strains were clearly allocated to the *G. haemolysans* cluster (Figure 2d).

Interspecies and intraspecies distances

Pairwise comparisons of the ITS sequences were performed for type strains belonging to the same genus to calculate the interspecies distances. Intraspecies distances were calculated among the strains belonging to the same species. The interspecies distances among the type strains of *Aerococcus*, *Facklamia*, *Granulicatella*, and *Leuconostoc* were in the range of 0.067-0.266. The intraspecies distances for the strains belonging to these four genera were all less than 0.047, for some species it was zero (Table 3). The interspecies distance between the type strains of *G. haemolysans* and *G. morbillorum* was only 0.025. The interspecies distances among the type strains of *G. bergeri*, *G. cuniculi*, *G. palaticanis*, and *G. sanguinis* were in the range of 0.038 to 0.113. The intraspecies distances were less than 0.056 for the *G. bergeri* strains, less than 0.016 for the *G. haemolysans* strains, less than 0.036 for the *G. morbillorum* strains, and less than 0.026 for the *G. sanguinis* strains (Table 3).

Genus	Species	n	Intraspecies distances	Interspecies distances*
Aerococcus	*A. christensenii*	3	< 0.014	0.121-0.266
	A. sanguinicola	5	0.0	
	A. urinae	27	0.0	
	A. viridans	2	< 0.031	
Abiotrophia	*A. defectiva*	9	< 0.013	-
Facklamia	*F. hominis*	4	< 0.015	0.087-0.233
	F. languida	2	< 0.017	
	F. sourekii	2	0.0	
Granulicatella	*G. adiacens*	9	< 0.024	0.211
	G. elegans	5	< 0.047	
Gemella	*G. bergeri*	4	< 0.056	0.025-0.113
	G. haemolysans	10	< 0.016	
	G. morbillorum	13	< 0.036	
	G. sanguinis	5	< 0.026	
Leuconostoc	*L. citreum*	1	0.0	0.067-0.133
	L. lactis	2	< 0.047	
	L. pseudomesenteroides	2	< 0.014	

*Calculations based on multiple sequence alignments and pairwise distance measurements of ITS sequences of the type strains.

Table 3: Inter- and intraspecies distances of the genus *Aerococcus, Abiotrophia* (only intraspecies), *Facklamia*, *Granulicatella*, *Gemella*, and *Leuconostoc* (Kimura-2-parameter substitution model (MEGA 5)).

Discussion

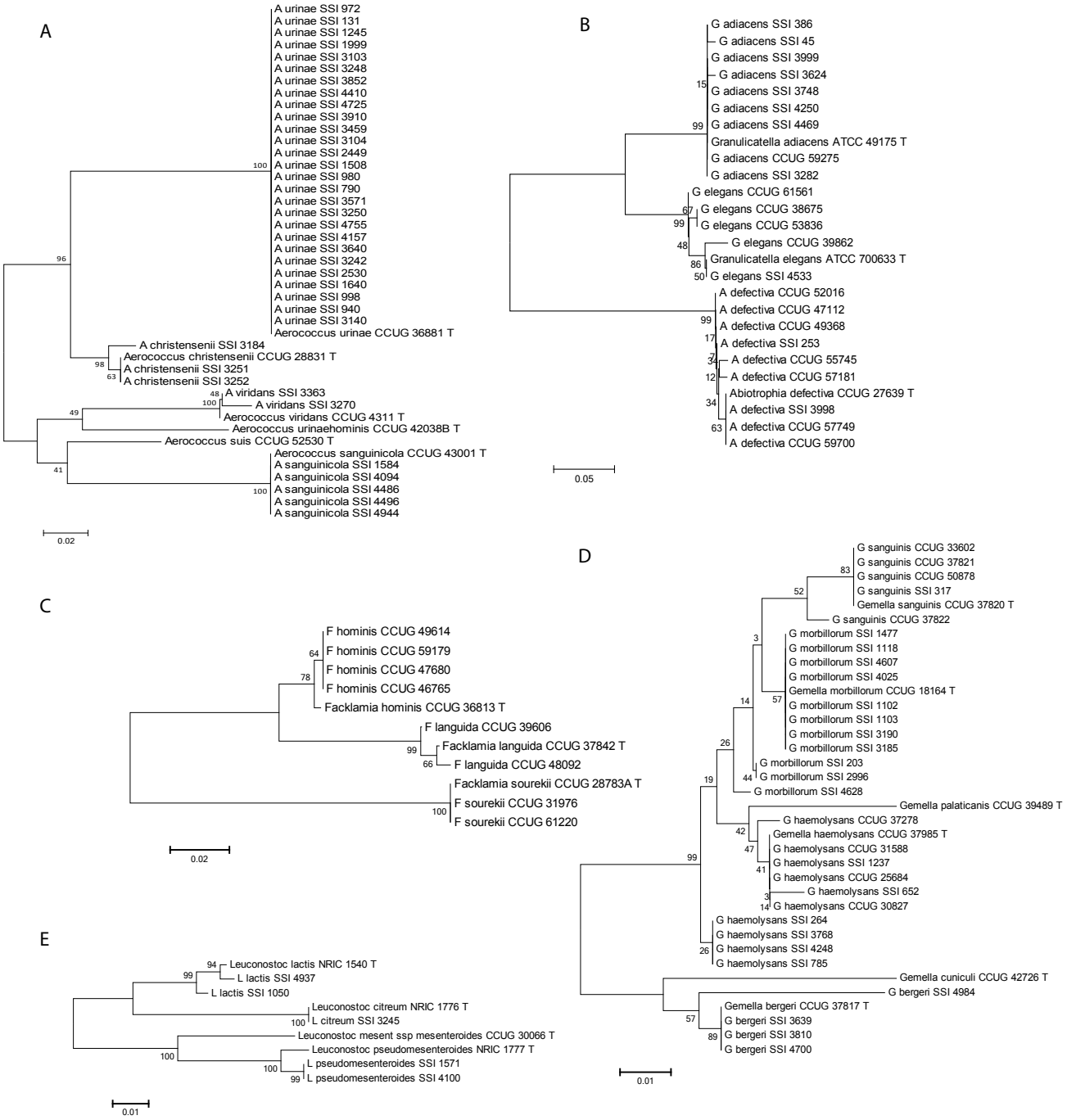
We describe a method for species-level identification by ITS sequence analysis for the strains belonging to 11 genera that are catalase-negative, gram-positive cocci and not streptoocci nor enterococci.

In our study, ITS sequences were determined for a total of 25 type strains; and ITS sequences were downloaded for further four type strains. The re-BLAST results for ITS region sequences of these 29 type strains showed a large distance from the first to the second best taxon match. This indicates that the ITS region has great interspecies divergence and is suitable as a species identification target (Table 1).

103 clinical strains from 17 species of the genera *Aerococcus*, *Abiotrophia*, *Facklamia*, *Granulicatella*, *Gemella*, and *Leuconostoc* were examined and identified based on ITS sequence analysis. All clinical strains, irrespective of obtained Maximum score value, were allocated to the expected species (Table 2). Large interspecies divergence, high intraspecies homology, and the distinct clustering demonstrated by the phylogentic analysis supported that the ITS region is a good target for species identification of strains belonging to the genera *Aerococcus*, *Abiotrophia*, *Facklamia*, *Granulicatella*, and *Leuconostoc*.

In the genus *Gemella*, 28 out of the 30 clinical strains were allocated to the expected species by BLAST. Two *Gemella* strains achieved *G. haemolysans* as first taxon match, though the Maximum score distance to the second taxon match (*G. morbillorum)* was only six, which was too short to make an unambiguous conclusion. The interspecies divergence between *G. haemolysans* and *G. morbillorum,* based on ITS sequences was as small as 0.025. This contributes to the difficulty in differentiating strains of *G. morbillorum* and *G. haemolysans*. An earlier publication from our group applied MALDI-ToF MS for species identification on 23 strains from the same collection. It resulted in unreliable identifications for 14 of 23 *Gemella* strains, the number of which was reduced considerably after in-house database extension [9]. However, more species of *G. morbillorum* and *G. haemolysans* had short distances to second best taxon match illustrating the close relationship for some of the species. Species identification for *Gemella* strains based on *rpoB* and *groESL* genes showed similar difficulties [12,18,19] However, the phylogentic analysis based on ITS sequences showed two distinct clusters of *G. morbillorum* and *G. haemolysans* strains, and the two specific strains in question were allocated to the *G. haemolysans* cluster (Figure 2d). Therefore, in this case, ITS sequence analysis with the combination of BLAST and phylogenetic analysis was sufficient to identify these two *G. haemolysans* strains to the species level.

Using the heterogeneity of the 16S-23S ITS region has become more common over the past years for identification and typing purposes of bacteria [26]. It is known that bacterial genomes can contain several *rrn* operons, e.g. *Enterococcus* has three copies of this operon [27] and *Streptococcus pneumoniae* has four operons [28]. The ITS sequence tends to present a mosaic organization of blocks, highly conserved at intra- and interspecies level, within the genus *Streptococcus* [28]. But in other genera, both length and copy number can vary from strain to strain [26]. Gurtler et al. established typing of *Clostridium difficile* strains by PCR amplification of variable length of ITS regions, which is still the standard for *C. difficile* PCR ribotyping [29]. Therefore,



CCUG: Culture Collection of the University of Göteborg, Sweden.
ATCC: American Type Culture Collection, USA.
NRIC: Nodai Research Institute Culture Collection, Japan.
SSI: Reference Laboratory Statens Serum Institut, Denmark.
T: Type strains.

Figure 2: Distance based Neighbor-Joining (NJ) phylogenetic trees on ITS sequences for each genus including the type and clinical strains were performed by Molecular Evolutionary Genetic Analysis (MEGA) software v 5.0. The significance test of the phylogenetic trees was based on 1000 bootstrap replicates.
a: Genus *Aerococcus.* It shows distinct clustering for the clinical strains and the corresponding type strains of the species *A. urinae, A. christensenii, A. sanguinicola,* and *A. viridians.*
b: Genera *Granulicatella* and *Abiotrophia.* It shows distinct clustering for the clinical strains and the corresponding type strains of the species *A. defectiva*, *G. adiacens*, and *G.elegans*.
c: Genus *Facklamia*. It shows distinct clustering of the clinical strains and type strains of *F. hominis*, *F. languida*, and *F. sourekii.*
d: Genus *Gemella.* It shows distinct clustering of the clinical strains and type strains of *G. bergeri*, *G. sanguinis*, *G. morbillorum,* and *G. haemolysans*. The two strains that were designated as *G. haemolysans* but with very short Maximum score distance to the second best match taxon lies clearly in the *G. haemolysans* cluster.
e: Genus *Leuconostoc.* It shows distinct clustering for the clinical strains and the corresponding type strains of the species *L. citreum*, *L. lactis*, and *L. pseudomesenteroides*.

a challenge for species identification based on ITS sequence analysis could be that some strains generate more than one amplicon, though only one may be dominant. It is important to understand the variability of ITS sequences in a given genome to gain insights into bacterial taxonomy. Tung et al. [23] investigated the applicability of ITS for identification of *Abiotrophia*, *Enterococcus*, *Granulicatella* and *Streptococcus*. The correct species identification rate by ITS sequence analysis for the 217 clinical strains belonging to these four genera was 98.2%. Except for the genus *Streptococcus*, all the other genera produced more than one amplicon. This made it necessary to perform agarose gel separation and purification of the amplicons before sequencing. Even after gel purification, two strains resulted in mixed sequences. Cloning was necessary for achieving unambiguous sequences. In our study, by optimizing the PCR conditions using SYBR Green master mix and changing the annealing temperature, it was possible to generate one predominant PCR amplicon for all the analyzed strains, and DNA sequences were obtained for these amplicons.

ITS sequence analysis seems promising for species identification of the strains that are catalase-negative, gram-positive, and resembles NHS. Of great importance, no species misidentifications were suggested by ITS sequence analysis among these strains. In a former study, the same method was applied for the genus *Streptococcus* and was shown to be sufficient for species identification of most streptococcal species [24]. For *S. mitis*, *S. oralis*, and *S. pneumonia* strains, it was necessary to examine an additional housekeeping gene, *gdh*. ITS sequence analysis has also been shown to be a useful tool for species identification of the genus *Enterococcus*.

Several other gene targets have been applied for species level identification for this group of bacteria. Drancourt et al. showed that the *rpoB* gene was useful in achieving species identification of the genus *Streptococcus* and related genera. However, a 99.4% similarity between *G. haemolysans* and *G. morbillorum* in their partial *rpoB* gene sequence was observed in the same study [19] besides, there were only 1-3 strains for each included species belonging to *Abiotrophia*, *Granulicatella* and *Gemella* [19]. Hung et al. developed a multiplex PCR attempting to differentiate strains of these three genera by the different size of *groESL* PCR products. This method could only achieve identification to the genus level. High intraspecies variation was observed in the *groESL* gene sequences among *G. haemolysans* isolates [17]. None of these two studies included *Aerococcus*, *Facklamia*, and *Leuconostoc*. Bosshard et al. concluded that 16S rRNA gene sequence analysis was an effective mean for identification of 171 clinical isolates of catalase-negative, gram-positive cocci. However, only a limited number of species were included from the genus *Aerococcus*, *Gemella*, *Enterococcus*, and *Streptococcus* [21]. Dynamic taxonomy development has happened in this area in the past ten years. Therefore, studies that include more species are necessary to support this conclusion.

In conclusion, ITS sequence analysis might be considered as a common identification key for bacteria that are catalase-negative, gram-positive cocci. Potentially, ITS sequence analysis can also be useful in detecting bacteria directly from clinical specimens, which are culture negative, as with direct 16S rRNA gene analysis on specimens.

Acknowledgements

Thanks to all departments of clinical microbiology in Denmark for submitting clinical strains to the Reference Laboratory at Statens Serum Institut, Copenhagen, Denmark. Thanks to Annemarie Hesselbjerg and Elsa Vilhelmsen from the Department of Microbiology Diagnostics and Virology for the excellent technical support.

This research was supported by grants from Oda and Hans Svenningsens Foundation, Grosserer L. F. Foghts Foundation, and from The Regional and the Local Research Foundation of Region Zealand, Denmark.

References

1. Ruoff KL (2011) *Aerococcus, Abiotrophia*, and other aerobic, catalase-negative, Gram-positive cocci. Manual of clinical microbiology. (10th edn). American Society for Microbiology, Washington DC.
2. Brouqui P, Raoult D (2001) Endocarditis due to rare and fastidious bacteria. Clin Microbiol Rev 14: 177-207.
3. Cattoir V, Kobal A, Legrand P (2010) *Aerococcus urinae* and *Aerococcus sanguinicola*, two frequently misidentified uropathogens. Scand J Infect Dis 42: 775-780.
4. Christensen JJ, Skov R (2002) *Aerococcus urinae* and *Aerococcus sanguinicola*. In Antimicrobial Therapy and Vaccines .Lippincott, Williams and Wilkins.,
5. Senneby E, Petersson AC, Rasmussen M (2012) Clinical and microbiological features of bacteraemia with *Aerococcus urinae*. Clin Microbiol Infect 18: 546-550.
6. Lee MR, Huang YT, Lee PI, Liao CH, Lai CC, et al. (2011) Healthcare-associated bacteraemia caused by Leuconostoc species at a university hospital in Taiwan between 1995 and 2008. J Hosp Infect 78: 45-49.
7. Ishiyama K, Yamazaki H, Senda Y, Yamauchi H, Nakao S (2011) Leuconostoc bacteremia in three patients with malignancies. J Infect Chemother 17: 412-418.
8. Woo PC, Lau SK, Fung AM, Chiu SK, Yung RW, et al. (2003) Gemella bacteraemia characterised by 16S ribosomal RNA gene sequencing. J Clin Pathol 56: 690-693.
9. Christensen JJ, Dargis R, Hammer M, Justesen US, Nielsen XC, et al. (2012) Matrix-assisted laser desorption ionization-time of flight mass spectrometry analysis of Gram-positive, catalase-negative cocci not belonging to the Streptococcus or Enterococcus genus and benefits of database extension. J Clin Microbiol 50: 1787-1791.
10. Kolbert CP, Persing DH (1999) Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. Curr Opin Microbiol 2: 299-305.
11. Woo PC, Ng KH, Lau SK, Yip KY, Fung AM, et al. (2003) Usefulness of the MicroSeq 500 16S ribosomal DNA-based bacterial identification system for identification of clinically significant bacterial isolates with ambiguous biochemical profiles. J.Clin.Microbiol 41: 1996-2001.
12. Christensen JJ, Dargis R, Kaltoft MS, Andresen K, Kemp M (2006) Ribosomal DNA sequencing of streptococci: Usefulness in species identification? International Congress Series 1289: 155-158.
13. Ikryannikova LN, Lapin KN, Malakhova MV, Filimonova AV, Ilina EN, et al. (2011) Misidentification of alpha-hemolytic streptococci by routine tests in clinical practice. Infect Genet Evol 11: 1709-1715.
14. Sistek V, Boissinot M, Boudreau DK, Huletsky A, Picard FJ, et al. (2012) Development of a real-time PCR assay for the specific detection and identification of Streptococcus pseudopneumoniae using the recA gene. Clin Microbiol Infect 18: 1089-1096.
15. Kitten T, Munro CL, Zollar NQ, Lee SP, Patel RD (2012) Oral streptococcal bacteremia in hospitalized patients: taxonomic identification and clinical characterization. J Clin Microbiol 50: 1039-1042.
16. Poyart C, Quesne G, Coulon S, Berche P, Trieu-Cuot P (1998) Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. J Clin Microbiol 36: 41-47.
17. Hung WC, Tseng SP, Chen HJ, Tsai JC, Chang CH, et al. (2010) Use of groESL as a target for identification of Abiotrophia, Granulicatella, and *Gemella* species. J Clin Microbiol 48: 3532-3538.
18. Teng LJ, Hsueh PR, Tsai JC, Chen PW, Hsu JC, et al. (2002) groESL sequence determination, phylogenetic analysis, and species differentiation for viridans group streptococci. J Clin Microbiol 40: 3172-3178.
19. Drancourt M, Roux V, Fournier PE, Raoult D (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.
20. Glazunova OO, Raoult D, Roux V (2010) Partial recN gene sequencing: a new tool for identification and phylogeny within the genus Streptococcus. Int J Syst Evol Microbiol 60: 2140-2148.

21. Bosshard PP, Abels S, Altwegg M, Böttger EC, Zbinden R (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.

22. Tung SK, Teng LJ, Vaneechoutte M, Chen HM, Chang TC (2006) Array-based identification of species of the genera Abiotrophia, Enterococcus, Granulicatella, and Streptococcus. J Clin Microbiol 44: 4414-4424.

23. Tung SK, Teng LJ, Vaneechoutte M, Chen HM, Chang TC (2007) Identification of species of *Abiotrophia, Enterococcus, Granulicatell*a and *Streptococcus* by sequence analysis of the ribosomal 16S-23S intergenic spacer region. J Med Microbiol 56: 504-513.

24. Nielsen XC, Justesen US, Dargis R, Kemp M, Christensen JJ (2009) Identification of clinically relevant nonhemolytic Streptococci on the basis of sequence analysis of 16S-23S intergenic spacer region and partial gdh gene. J Clin Microbiol 47: 932-939.

25. Christensen JJ, Andresen K, Justesen T, Kemp M (2005) Ribosomal DNA sequencing: experiences from use in the Danish National Reference Laboratory for Identification of Bacteria. APMIS 113: 621-628.

26. Gürtler V, Stanisich VA (1996) New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology 142 : 3-16.

27. Gürtler V, Rao Y, Pearson SR, Bates SM, Mayall BC (1999) DNA sequence heterogeneity in the three copies of the long 16S-23S rDNA spacer of *Enterococcus faecalis* isolates. Microbiology 145 : 1785-1796.

28. Gianninò V, Santagati M, Guardo G, Cascone C, Rappazzo G, et al. (2003) Conservation of the mosaic structure of the four internal transcribed spacers and localisation of the rrn operons on the *Streptococcus pneumoniae* genome. FEMS Microbiol Lett 223: 245-252.

29. Gürtler V (1993) Typing of Clostridium difficile strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. J Gen Microbiol 139: 3089-3097.