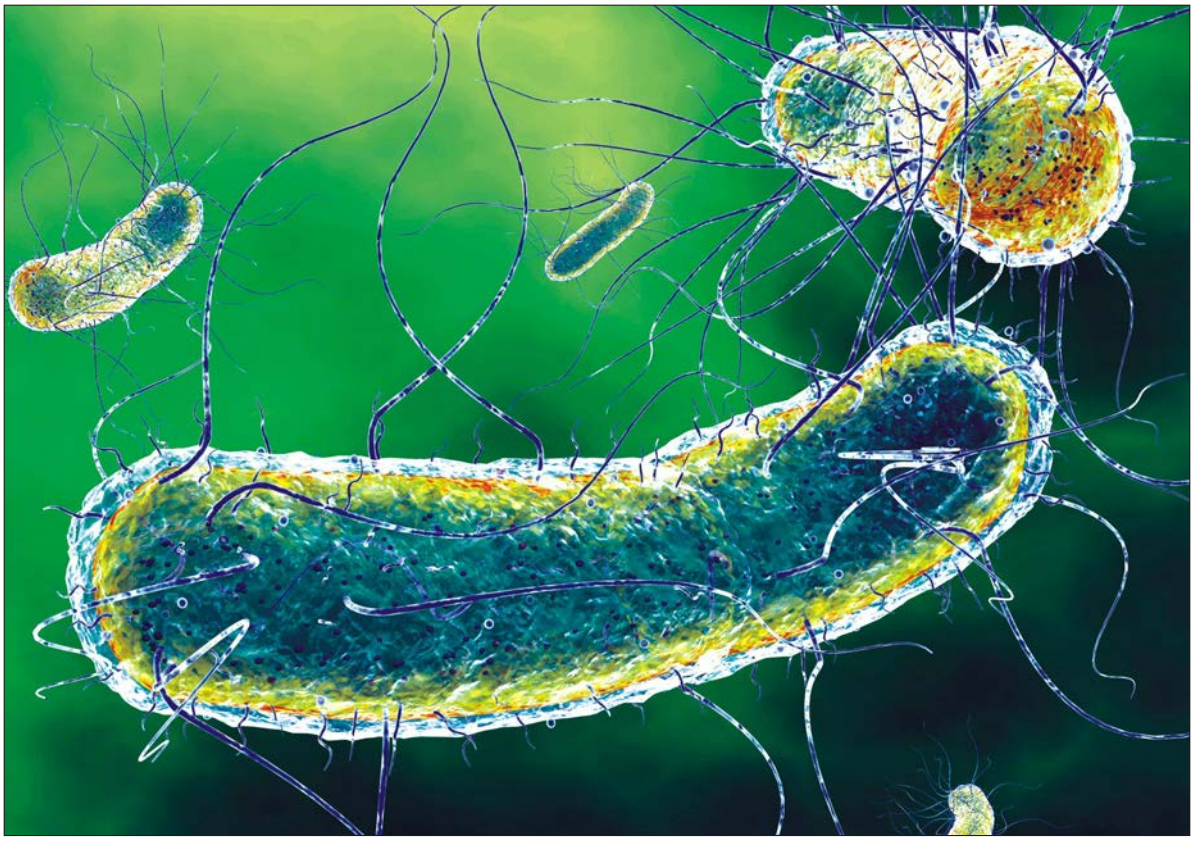

Current Research in Bacteriology

ISSN 1994-5426

science alert

ANSInet
an open access publisher
http://ansinet.com

Current Research in Bacteriology 8 (3): 48-61, 2015
ISSN 1994-5426 / DOI: 10.3923/crb.2015.48.61

Identification of Coagulase-Negative Staphylococci by *SodA* Gene Sequence Analysis

Aziyah Abdul-Aziz, Sharifah Aminah Syed Mohamad and Mohamad Faiz Foong Abdullah
Centre for Molecular Microbial Pathogenicity, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor Darul Ehsan, Malaysia

Corresponding Author: Aziyah Abdul-Aziz, Centre for Molecular Microbial Pathogenicity, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor Darul Ehsan, Malaysia

ABSTRACT

A total number of 50 internal *sodA* gene sequences with 416 nucleotides was analyzed to determine the discriminative power of the gene for the identification of the Coagulase-Negative Staphylococci (CoNS). The mean similarity between these CoNS species was found to be 80.4% indicating a good discriminative power although low sequence divergence was observed between some of the species. The relatedness between these staphylococcal species was further explored by constructing phylogenetic analysis. Six clusters was revealed, with four of these supported by significant bootstrap values of >95% while the other two clusters were less robustly supported at bootstrap values of about 92%. Following that, the identification of 200 clinical isolates of CoNS isolates were determined phenotypically by a commercial diagnostic kit and also genotypically using the *sodA* gene sequencing method. The Microgen Staph ID was found to be quite reliable for commonly isolated clinical CoNS species e.g., *S. capitis* subsp. *capitis*, *S. haemolyticus* and a majority of the *S. epidermidis* isolates. However, for the less commonly encountered species, the kit was unreliable. Identification by *sodA* sequencing method was found to be more reliable with homology values of at least 98%. However, like other nucleotide sequence based identification, the discriminative power of *sodA* at subspecies level was poor whereby the final identification had to be supplemented with biochemical tests.

Key words: Discriminative power, phenotypic, genotypic, phylogenetics, Microgen Staph ID

INTRODUCTION

The Coagulase-Negative Staphylococci (CoNS) is a group of *Staphylococcus* species which is distinguished from the more virulent *Staphylococcus aureus* by their inability to coagulate plasma. To date, there are more than 50 species of CoNS that have been identified including recent additions e.g., *S. jettensis* and *S. argenteus* (Tong *et al*., 2015). The CoNS were once considered relatively avirulent and have long been dismissed as culture contaminants since they are normal inhabitants of human skin and mucous membranes. The potential pathogenicity of CoNS in human medicine was first reported in 1958 but only in the 1970s that these organisms have become increasingly recognized as agents of clinically significant infections especially hospital-acquired opportunistic infections (Becker *et al*., 2014; Piette and Verschraegen, 2009). The CoNS are also known to be a leading cause of infections in infants. In the USA, data collected through The National Institute of Child Health and Human Development Neonatal Research Network revealed that CoNS accounted for 48% of late-onset sepsis in very low birth weight or VLBW infants (Stoll *et al*., 2002). In UK, the neonatal infection surveillance networks reported that CoNS

accounted for 54% of similar incidence in infants (Vergnano *et al*., 2011). The occurrence of CoNS have also been commonly associated with various medical devices such as prosthetic valves, cerebrospinal fluid shunts as well as intravascular, urinary and dialysis catheters (Chaudhury and Kumar, 2007; De Allori *et al*., 2006; Gatermann *et al*., 2007; Koksal *et al*., 2009; Rasigade *et al*., 2012).

The increasing significance of CoNS serves the justification for a more accurate species identification to allow a precise determination of host-pathogenic potential of each of the various CoNS species (Heikens *et al*., 2005; Poyart *et al*., 2001; Sivadon *et al*., 2005). Various methods have been proposed, both phenotypically and genotypically. In common routine diagnostics identification was based on biochemical tests and the fatty acid profile. However, while these methods are able to identify *S. aureus*, it fails to distinguish other staphylococci including the CoNS group (Martineau *et al*., 1998; Stoakes *et al*., 1994). The conventional biochemical tests for CoNS were also found to be unsatisfactory, unreliable and irreproducible. The CoNS can only be distinguished by a limited number of stable biochemical tests, hence, making identification to species level a difficult task. Similarly, the automated identification systems such as the as the ID32 STAPH® strip (bioMérieux) and the VITEK 2 identification card (bioMérieux) were also found to be unreliable due to the variable phenotypic characters as a result of expression of the different metabolic activities (Delmas *et al*., 2008; Kim *et al*., 2008).

A number of sequence-based identification methods which rely on the analysis of PCR products derived from some specific DNA targets have been developed as an alternative to phenotypic identification of *Staphylococcus* species. Analysis of 16S rRNA gene used to be the most common method used for bacteria classification and identification. However, the use of this gene is limited due to high degree of sequence similarity between closely related species (Becker *et al*., 2004; Goh *et al*., 1996; Gribaldo *et al*., 1997; Kwok *et al*., 1999). Hence, other highly conserved sequence has been exploited which includes *sodA* (Poyart *et al*., 2001), *gap* (Bergeron *et al*., 2011; Ghebremedhin *et al*., 2008), *hsp60* (Ghebremedhin *et al*., 2008; Kwok *et al*., 1999), *rpoB* (Drancourt and Raoult, 2002; Mellmann *et al*., 2006), *tuf* (Heikens *et al*., 2005) and *dnaJ* (Shah *et al*., 2007) gene sequencing methods. However, the sequences of some genes are not sufficient to discriminate between closely related *Staphylococcus* at both species and subspecies level and the databases only includes a limited number of species.

In this study, the discriminative power of the *sodA* gene for CoNS identification would first be evaluated using available *sodA* gene sequences downloaded from the GenBank database. This gene encodes for manganese-dependent superoxide dismutase, Mn-SOD. Using a pair of degenerate primers, a 429 bp long DNA fragment was amplified and characterized. The number of *sodA* sequences of CoNS analyzed in the present study is the largest so far with 50 species as compared previous partial analysis (Poyart *et al*., 2001; Ghebremedhin *et al*., 2008). Sequences from newly-reported CoNS species such as *S. nepalensis*, *S. microti*, *S. rostra*, *S. fleuretti* and subspecies of *S. succinus* were also included which gave a total of 50 species and sub-species of CoNS altogether. Following that, the identification of 200 strains of local clinical CoNS isolates will be first performed phenotypically using the commercial Microgen Staph ID (Microgen, UK) kit and subsequently compared to the *sodA* gene sequencing method.

MATERIALS AND METHODS

Analyses of the *sodA* gene sequence: DNA sequences of available CoNS sp from the GenBank database was aligned using the Clustal W software (Thompson *et al*., 1994). Similarity matrices for aligned sequences were generated using the BioEdit software (Hall, 1999) and the mean

similarity was calculated. Phylogenetic tree construction was performed by the neighbor-joining algorithm (Saitou and Nei, 1987) using the MEGA5 software (Kumar *et al*., 2004) and resampled with bootstrap values based on 1000 replications (Felsenstein, 1985). For comparison purposes, the *sodA* gene of *S. aureus* was also included in this study.

Collection of samples: A total of 200 CoNS strains were collected from the Pathology Lab of the Hospital Tuanku Ampuan Rahimah, Klang in between December, 2010 to May, 2011. These clinical strains were taken from warded patients and selected randomly from various clinical settings including blood, respiratory, pus, fluid and urine isolates. The identity of the CoNS isolates was further verified by their inability to coagulate blood plasma using the Pro-Lab Prolex™ Blue-Staph Latex Test Kit.

Phenotypic identification of the CoNS isolates: Biochemical identification to species level was performed using the Microgen Staph ID kit. The bacterial cell culture was inoculated into the 12 microwells of each strip which employs 13 standardised biochemical substrates as according to the manufacturer's instructions. Upon incubation, results based on the colour change of each well were recorded on a provided worksheet and interpreted using the Microgen™ Identification System Software (MID-60) to identify the test organisms.

Amplification of the *sodA* gene sequence: As most of these partial *sodA* sequences were generated by sequencing the PCR amplicons and deposited into the GenBank database by different groups of workers, the length of the sequences can vary greatly. Thus, the majority of these were trimmed to an internal sequence of 416 nt which corresponds to nucleotides 52-468 of the *S. epidermidis* coding region for the *sodA* gene sequence.

Amplification of the *sodA* gene sequence of the 200 CoNS isolates were performed using degenerate primers d1 (5' CCITAYICITAYGAYGCIYTIGARCC' 3) and d2 (5' ARRTARTAIGCRTGYTCCCAIACRTC' 3) which amplify a fragment of the superoxide dismutase gene that represents approximately 83% of the entire gene (Poyart *et al*., 2001).

The PCR reaction mix was prepared using Gotaq Flexi kit (Promega) in a total volume of 50 μL; 10 μL of 5X Go-Taq Buffer, 1 μL of a 200 μM concentration of each dNTP, 1 μL of 25 mM $MgCl_2$ solution, 2 μL of 10 μM of each primer, 0.5 μL of 1 U of Go-Taq DNA polymerase (Promega) and 2 μL of 150 ng of DNA as the template.

The PCR was performed in Piko Thermal Cycler (Thermo Scientific) with the following conditions: 3 min at 95°C for initial cycle, followed by 30 cycles of amplification of the followings; 60 sec of annealing at 37°C, 45 sec of elongation at 72°C and 30 sec of denaturation at 95°C. The last cycle was performed at 72°C for 10 min.

The PCR amplicons were sequenced and the DNA sequences of the gene were then identified by matching against the NCBI database using the BLAST interface.

RESULTS

Analysis of the discriminative power of the *sodA* gene sequence for the identification of CoNS isolates: The analysis was performed by first downloading the available *sodA* gene sequences from the GenBank database. However, as most of these partial *sodA* sequences were deposited into the GenBank database by different groups of workers, the length of the sequences can vary greatly. Thus, the majority of these were trimmed to an internal sequence of 416 nt which corresponds to nucleotides 52-468 of the *S. epidermidis* coding region for *sodA* gene sequence.

Similarity matrix of the partial *sodA* gene sequence of CoNS isolates: Results from the analysis showed that the mean similarity of the partial *sodA* sequences of the 50 CoNS was 80.4% as calculated from the similarity matrix shown in Fig. 1, with a minimum similarity of 56.0%. This is very similar to the value of mean similarity of 81.5% obtained by Poyart *et al*. (2001).

However, some exceptions at interspecies level were noted. Although, the majority of the partial *sodA* sequences showed remarkable discrimination, some pairs were quite similar to each other. For example, *S. condimenti* share 97% similarity in its partial *sodA* sequence with both subspecies of *S. carnosus* while *S. delphini* was 96% similar to *S. pseudintermidus*. Similarly, *S. condimenti* and *S. piscifermentans* were 95% similar to each other and the same degree of similarity was observed between *S. fleuretti* and *S. vitulinus*. In addition, only 6% sequence divergence was observed between the partial *sodA* sequences of the following pairs: *S. caprae* and *S. capitis*, both the subspecies of *S. carnosus* and *S. piscifermentans*, the three subspecies of *S. scuiri*; *S. vitulinus*, *S. pasteuri* and *S. warneri* and also between *S. saprophyticus* and *S. xylosus*.

The same trend was also observed in nucleotide-sequence identification using other genes. The 16S rDNA sequence similarity has been shown to be very high at 90-99% in 29 *Staphylococcus* species studied (Kwok *et al*., 1999). A sequence similarity of 95 and 99% was seen between *S. caprae* and *S. capitis* using the 16S rRNA and *hsp60* gene, respectively (Ghebremedhin *et al*., 2008) while the partial *rpoB* sequences of these two species were lower at 84.5% (Drancourt and Raoult, 2002) as compared to 94% similarity observed in the present study. *Staphylococcus sciuri* and *S. vitulinus* species showed 90.1-98.6% similarity in their *hsp60* gene sequences (Kwok and Chow, 2003; Kwok *et al*., 1999) while a similarity of 94% between these species were detected in the present study. Similarly, the *sodA* sequence revealed a similarity of 89% between *S. simulans* and *S. carnosus* while the sequence similarity was 87, 95 and 96% with *rpoB*, *gap* and 16s rRNA gene sequence, respectively (Ghebremedhin *et al*., 2008).

A much higher similarity was observed at subspecies level. As shown in Fig. 1, both *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticum* shared a partial *sodA* sequence similarity of 96%, making sequence based identification between these two species a difficult task. Similarly, only 1% of sequence divergence was observed between *S. hominis* subsp. *hominis* and *S. hominis* subsp. *novobiosepticus* and also between the two species of *S. schleiferi*. In addition, the partial *sodA* gene sequence failed to differentiate between the two subspecies of *S. capitis, S. carnosus, S. saprophyticus* and *S. succinus* and also the three subspecies of *S. sciuri* as the similarity among the subspecies was 100%. This finding is consistent with other conserved genes used for staphylococcal identification including *dnaJ* (Shah *et al*., 2007), *hsp60* (Kwok and Chow, 2003) and 16S rRNA (Ghebremedhin *et al*., 2008; Kwok and Chow, 2003) where majority of complete discrimination at subspecies level was unsuccessful due to of high similarities among the subspecies.

In general, the results suggest that apart from the subspecies level, the *sodA* internal sequence is able to show more species discrimination as compared to the sequence of other conserved genes such as the 16S rRNA with mean similarity of 90% (Shah *et al*., 2007), *rpoB* with mean similarity of 86% (Drancourt and Raoult, 2002; Mellmann *et al*., 2006) and *hsp60* with mean similarity of 82% (Kwok *et al*., 1999). Following general convention, a 95% identity threshold was adopted as the limit for a positive identification and unambiguous assignment to a particular species.

Phylogenetic analysis of *Staphylococcus* derived from partial *sodA* gene sequences: The relatedness between the *sodA* internal sequences of the staphylococcal species were further explored using standard phylogenetic tree. Figure 2 represents the phylogenetic tree drawn using the neighbour-joining method based on the internal sequence of the *sodA* gene.

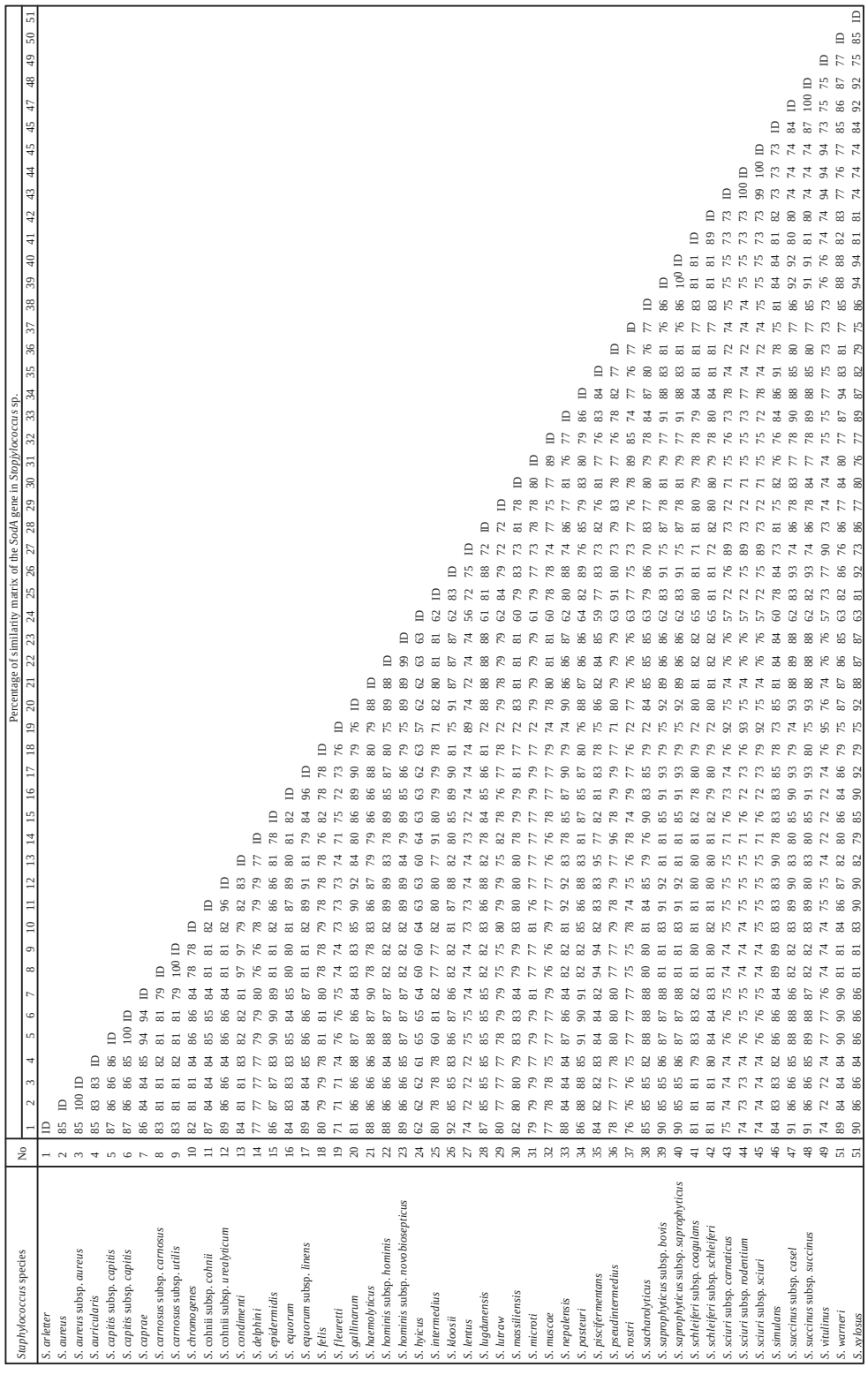

Fig. 1: Identity matrix based on the internal sequences of the *sodA* gene in *Staphylococcus* sp. The numbers representing each of the *Staphylococcus* sp correspond with the numbers displayed on top of the table. The mean similarity matrix calculated is 80.4%

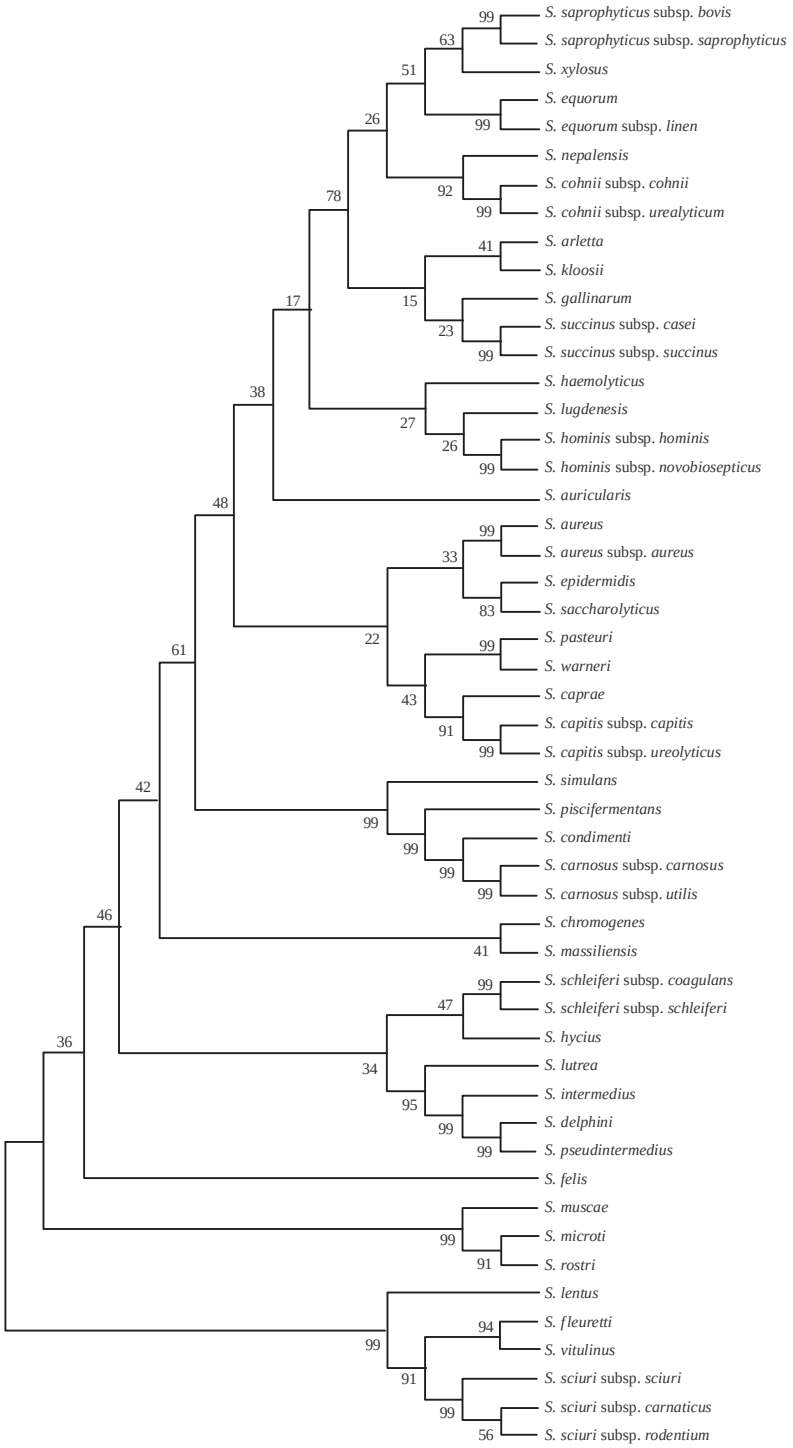

Fig. 2: Phylogenetic tree based on partial *sodA* gene sequence of 50 *Staphylococcus* species and subspecies. Phylogenetic tree based on the *sodA* internal sequence showing relationships among the 51 staphylococcal strains. The tree was established using the neighbour-joining method from an analysis of sequences listed in Fig. 1. The percentage of bootstrap value indicated at each branch was calculated from 1000 resamplings. Each of the shaded area represents a cluster

In agreement with Poyart *et al*. (2001), three major clusters supported by significant bootstrap values of >95% were observed. These are the *S. sciuri*, *S. simulans* and the *S. intermedius* clusters. The *S. sciuri* group consists of the three subspecies of *S. sciuri* together with *S. lentus, S. vitulinus* and the newly reported *S. fleuretti*. As noted by Poyart *et al*. (2001), members of this cluster differ from the other staphylococcal species by the presence of an extra codon-GCT-at nucleotide number 220 of the *sodA* internal sequence (or nucleotide 274 of the corresponding *sodA* gene of *S. epidermidis*). This codon codes for an extra alanine residue at amino acid number 92 of the corresponding *S. epidermidis sodA* protein. All members of the *S. sciuri* cluster were also observed to carry two unique signature sequences within their *sodA* gene, first is the CCTTC moiety at nucleotide 226 and the second is a GTTCTAC at nucleotide 257.

The second cluster consists of *S. simulans*, *S. piscifermentans*, *S. condimenti*, *S. carnosus* subsp. *carnosus* and *S. carnosus* subsp. *utilis*. This cluster is supported by a bootstrap value of 99% and has a unique feature of an extra codon-AAA-at nucleotide 349 (corresponding to nucleotide 403 of the *sodA* coding region), resulting in an extra lysine residue at amino acid number 135 of the corresponding protein. The third, *S. intermedius* cluster was supported by a bootstrap value of 95% and consists of *S. intermedius*, *S. pseudintermedius*, *S. lutrae* and *S. delphini*.

The phylogenetic tree also revealed the presence of an additional three clusters. The fourth cluster, the *S. muscae* cluster, consists of *S. muscae*, *S. rostri* and *S. microti* and robustly supported by a bootstrap value of 99%. The fifth and sixth clusters consist of *S. caprae*, *S. capitis* subsp. *capitis* and *S. capitis* subsp. *urealyticum* in one and *S. nepalensis*, *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticum* in the other, were less robustly supported at bootstrap values of 91 and 92%, respectively.

Identification of CoNS isolates by microgen staph ID: Using the Microgen Staph ID biochemical test kit, a total of 14 species of CoNS were identified, with *S. epidermidis* accounting for the highest number of isolates at 75, followed by *S. haemolyticus*, *S. hominis* and *S. capitis* with 48, 17 and 14 isolates, respectively.

Some of these identifications were however unconvincing. *Staphylococcus chromogenes* is known as an animal pathogen and is also one of the predominant CoNS species in bovine mastitis. To date, no case of human infection caused by this bacterium has been reported. Similarly, *S. hyicus*, *S. intermedius* and *S. schleiferi* subsp. *coagulans* have been known as part of the commensal flora of various animals while infection in human is very rare (Bes *et al*., 2002; Casanova *et al*., 2011; Fudaba *et al*., 2005). Hence, results from the identification by the Microgen™ Identification System Software were doubtful.

In general, although some of the isolates displayed excellent to good differentiation at a probability of >90%, some identification were poor with probabilities of less than 80%, adding to the uncertainties of the actual identity of these species of CoNS.

Amplification and identification of CoNS via the *sodA* gene sequence: Figure 3 shows some representative of the result of amplification of the *sodA* gene of the 200 CoNS clinical isolates. A single amplicon of 429 bp was observed in all of the staphylococcal species studied. In total, the *sodA* gene identified *S. epidermidis* with 74 isolates followed by *S. haemolyticus* (60), *S. hominis* (36) and *S. capitis* (17). The remaining isolates include *S. cohnii*, *S. lugdenensis*, *S. saprophyticus*, *S. sciuri* and *S. warneri*.

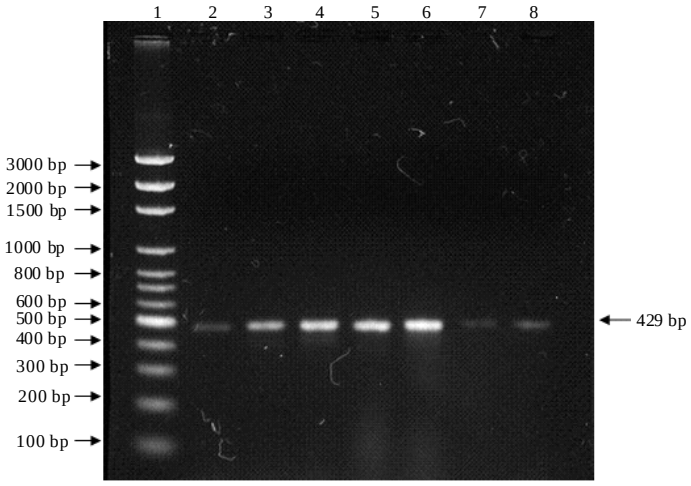

Fig. 3: Amplification of the *sodA* gene sequence in the CoNS isolates. Lane 1 : 100 bp DNA ladder, Lane 2: ATCC 35984 (positive control, lane 3-8 are samples of CoNS isolates displaying positive amplification for the *sodA* gene with an expected amplicon of 429 bp

Comparison of results from Microgen Staph ID and the *sodA* sequencing method: Table 1 shows the differences in results of the identification of some of the CoNS isolates via the Microgen Staph ID kit and *sodA* sequence analysis. Both the Microgen Staph ID and *sodA* sequence results were in agreement for the majority of the common CoNS isolates. However, for less common species, a high number of discrepancies between both sets of data were observed.

Both the Microgen Staph ID and *sodA* sequence results were in agreement for the majority of the *S. epidermidis*. However, 14.9% or 11 of the isolates identified by Microgen as *S. epidermidis* were identified as either *S. hominis* subsp. *hominis* or *S. hominis* subsp. *novobiosepticus* with each exhibiting a BLAST percentage identity of 99% except for isolate B213 which was identified as *S. capitis* subsp. *ureolyticus* instead with percentage identity of 100%. Similarly, almost all of the *S. haemolyticus* isolates were correctly identified by Microgen except for two isolates. Isolate B172 and P15 gave a low percentage of probability of 67.4 and 86.7%, respectively which were later identified as *S. hominis* subsp. *hominis* using *sodA* sequences.

Both the Microgen Staph ID kit and *sodA* analysis correctly identified all of the *S. capitis* subsp. *capitis* isolates with percentage of probability ranging from 86.47-99.24% and percentage of identity of at least 99%. However, the kit was appears to misidentify *S. capitis* subsp. *ureolyticus*. These isolates were misidentified either as *S. epidermidis* (B213), *S. schleiferi* subsp. *coagulans* (P27) or *S. xylosus* (B111) albeit the high percentage probability of between 80-96% generated by the Microgen software. For *S. hominis* subsp. *hominis*, a total of 40% or 12 of isolates were correctly identified by the Microgen Staph ID as reconfirmed by the *sodA* sequencing method while the remaining eight were misidentified.

Identification by Microgen Staph ID for the rest of the CoNS isolates was correspondingly unreliable. The kit failed to identify any of the *S. cohnii*, *S. hominis* subsp. *novobiosepticus*, *S. saprophyticus* and *S. sciuri* isolates. Results by BLAST showed that none of the clinical isolates were either *S. chromogenes*, *S. hyicus*, *S. intermedius*, *S. lentus* or *S. schleiferi* subsp. *coagulans* which are known to be associated with infections in animals. Similarly, none of the clinical isolates were *S. simulans* or *S. xylosus* as identified by the Microgen Staph ID kit, although these two species are generally known to be related to humans.

Table 1: Discrepancies between identification by biochemical tests and *sodA* gene sequence analysis

Isolate	Microgen staph ID	Probability (%)	*SodA* sequence genotyping	ID (%)
B170	*S. chromogenes*	99.92	*S. haemolyticus*	99
B118	*S. chromogenes*	95.80	*S. haemolyticus*	100
B9	*S. chromogenes*	91.02	*S. hominis* subsp. *hominis*	100
B194	*S. chromogenes*	87.90	*S. hominis* subsp. *hominis*	100
B207	*S. chromogenes*	75.32	*S. hominis* subsp. *hominis*	99
B10	*S. chromogenes*	71.33	*S. hominis* subsp. *hominis*	100
B204	*S. chromogenes*	71.33	*S. hominis* subsp. *hominis*	99
U3	*S. cohnii* subsp. *urealyticum*	86.41	*S. cohnii* subsp. *cohnii*	100
B213	*S. epidermidis*	96.22	*S. capitis* subsp. *ureolyticus*	100
B128	*S. epidermidis*	79.90	*S. hominis* subsp. *novobiosepticus*	99
B6	*S. epidermidis*	99.48	*S. hominis* subsp. *novobiosepticus*	99
B40	*S. epidermidis*	99.48	*S. hominis* subsp. *novobiosepticus*	99
B48	*S. epidermidis*	99.48	*S. hominis* subsp. *novobiosepticus*	99
B134	*S. epidermidis*	99.48	*S. hominis* subsp. *novobiosepticus*	99
B16	*S. epidermidis*	98.19	*S. hominis* subsp. *hominis*	99
B205	*S. epidermidis*	94.33	*S. hominis* subsp. *hominis*	99
B50	*S. epidermidis*	85.25	*S. hominis* subsp. *hominis*	99
B100	*S. epidermidis*	85.25	*S. hominis* subsp. *hominis*	99
B172	*S. haemolyticus*	67.36	*S. hominis* subsp. *hominis*	99
P15	*S. haemolyticus*	86.70	*S. hominis* subsp. *hominis*	99
B61	*S. hominis* subsp. *hominis*	53.20	*S. epidermidis*	99
R15	*S. hominis* subsp. *hominis*	53.20	*S. epidermidis*	100
B51	*S. hominis* subsp. *hominis*	99.38	*S. saprophyticus* subsp. *saprophyticus*	99
P26	*S. hominis* subsp. *hominis*	97.32	*S. saprophyticus* subsp. *saprophyticus*	99
U1	*S. hominis* subsp. *hominis*	91.12	*S. saprophyticus* subsp. *saprophyticus*	99
B136	*S. hyicus*	52.05	*S. cohnii* subsp. *urealyticum*	99
B193	*S. hyicus*	99.98	*S. haemolyticus*	99
P12	*S. hyicus*	81.86	*S. haemolyticus*	99
B11	*S. intermedius*	84.27	*S. haemolyticus*	99
B146	*S. lentus*	92.90	*S. sciuri*	100
B185	*S. lugdenensis*	72.02	*S. haemolyticus*	99
B162	*S. lugdenensis*	98.37	*S. hominis* subsp. *hominis*	100
B177	*S. lugdenensis*	74.95	*S. hominis* subsp. *hominis*	100
P27	*S. schleiferi* subsp. *coagulans*	81.83	*S. capitis* subsp. *ureolyticus*	100
B182	*S. schleiferi* subsp. *coagulans*	97.73	*S. epidermidis*	99
B163	*S. schleiferi* subsp. *coagulans*	93.85	*S. epidermidis*	98
R27	*S. simulans*	72.19	*S. haemolyticus*	99
B49	*S. simulans*	51.46	*S. haemolyticus*	99
B111	*S. xylosus*	79.92	*S. capitis* subsp. *ureolyticus*	100
B154	*S. xylosus*	99.94	*S. cohnii* subsp. *urealyticum*	99
R33	*S. xylosus*	99.61	*S. haemolyticus*	100
P3	*S. xylosus*	99.20	*S. haemolyticus*	99

Identification of the CoNS subspecies: As the partial *sodA* gene sequence method failed to resolve the CoNS isolates at subspecies level, their final identification had to rely on biochemical characteristics. *Staphylococcus capitis* subsp. *ureolyticus* can be differentiated from subsp. *Capitis* by three biochemical tests (Bannerman and Kloos, 1991) which were the presence of urease and the production of acid from sucrose and mannitol (variable for subsp. *Capitis*). The results for these tests were then extracted from the Microgen Staph ID test as they were included as part of the biochemical tests of the kit. Isolates with positive results were thus confirmed as *S. capitis* subsp. *ureolyticus*.

The partial *sodA* sequencing method also failed to discriminate between the subspecies of *S. hominis* with only 1% difference in the sequences by three nucleotide changes. However, biochemically, *S. hominis* subsp. *novobiosepticus* can be differentiated from subsp. *hominis* by the lack of arginine dehydrogenase activity and failure to produce acid aerobically from trehalose and N-acteylglucosamine (Kloos *et al*., 1998). *Staphylococcus cohnii* subsp. *cohnii* can be differentiated

from subsp. *urealyticum* by the lack of urease and b-glucuronidase activities which are positive for subsp. *urealyticum*. In addition, *S. cohnii* subsp. *cohnii* failed to produce acid from N-acetylglucosamine, another trait that is positive for subsp. *Urealyticum* (Kloos and Wolfshohl, 1991).

Similarly, *S. saprophyticus* subsp. *novis* can be phenotypically distinguished from *S. saprophyticus* subsp. *saprophyticus* on the basis of nitrate reduction and the ability to produce PYR (Hajek *et al*., 1996) which were part of the tests included in the Microgen Staph ID kit. All three clinical isolates of *S. saprophyticus* were unable to reduce nitrate and did not produced PYR. Thus, they were identified as *S. saprophyticus* subsp. *saprophyticus*. However, the phenotypic characteristics to distinguish between the three subspecies of *S. sciuri* are tedious with a number of variables as outlined by Kloos and friends (Kloos *et al*., 1997). Hence, other methods have been used to differentiate these subspecies which include ribotyping delineation and DNA-DNA liquid hybridization.

DISCUSSION

The use of *sodA* gene as a nucleic acid targets provides an alternative technique for the accurate identification and classification of *Staphylococcus* species. Unfortunately, one perfect identification system for *Staphylococcus* does not exist. Like other conserved gene used in staphylococcal identification, *sodA* gene sequence does have some drawbacks as it fails to discriminate some of the closely related bacteria. This was shown in the present study where the sequence failed to differentiate some *Staphylococcus* isolates at their subspecies level. As such, the identity of these isolates was confirmed via selective biochemical tests instead.

Alternatively, other nucleotide targets can be used instead. For example, *dnaJ* gene sequence displayed a similarity of 90.9% between *S. condimenti* and *S. carnosus* as compared to 97% similarity observed in the present study. At subspecies level, with 7% divergence, *dnaJ* (Shah *et al*., 2007) and *hsp60* (Kwok *et al*., 1999) gene sequences can be used to discriminate *S. cohnii* subsp. *cohnii* from *S. cohnii* subsp. *urealyticus* as compared to only 4% divergence using *sodA* gene as observed in this study. Similarly, *hsp60* gene (Kwok *et al*., 1999) showed 9% sequence divergence between *S. capitis* subsp. *capitis* and *S. capitis* subsp. *ureolyticus*.

Accurate identification of staphylococcal isolates is crucial for the correct management of staphylococcal infections. The emergence of clinically significant CoNS and the expanding number of staphylococcal species and subspecies due to new discoveries have led to the need of a more reliable and reproducible method for accurate identification of this species. In addition, the identification of CoNS species is essential to establish epidemiological trends, to determine the cause of specific infections and analysis of treatment failures (De Paulis *et al*., 2003). Microgen Staph ID appears to be inadequate for the identification of uncommon CoNS species, although it could identify the major clinical species. In contrast, with homology values of at least 98%, the genotypic PCR-amplicon sequencing based technique via the *sodA* gene proved to be a more reliable, easy and valuable approach to identify the various species of CoNS. However, in some closely related isolates staphylococcal identification via *sodA* alone is insufficient and further tests are essential to allow discrimination at subspecies level and confirmed the identity of the isolates.

The genotypic method was found to be more reliable than the phenotypic method for the identification of CoNS. Microgen Staph ID kit correctly identified all of the *S. capitis* subsp. *capitis* isolates and also a majority of both *S. haemolyticus* and *S. epidermidis* isolates. However, the kit appears to be unreliable for the remaining isolates. Only 40% of the *S. hominis* subsp. *hominis*

isolates were correctly identified. The same kit was unable to identify *S. capitis* subsp. *Ureolyticus*, *S. hominis* subsp. *novobiosepticus*, *S. saprophyticus*, *S. sciuri* and *S. cohnii* subsp. *cohnii* although, these species were included in their database. For other species identified in this study such as *S. cohnii* subsp. *urealyticum*, *S. lugdenensis* and *S. warneri*, the number of the isolates belonging to these species was too limited to be conclusive. In addition, the BLAST results of the *sodA* gene showed that none of the clinical isolates were *S. chromogenes*, *S. hyicus*, *S. intermedius*, *S. lentus*, *S. schleiferi* subsp. *coagulans*, *S. simulans* or *S. xylosus* species as indicated by the Microgen software. Without performing a third test for validation, it is difficult to claim that the gene sequence method is more reliable. However, circumstantial observation i.e., the fact that some of the less common species as identified by the Microgen kit are very rarely isolated from clinical samples support this.

Similar findings have also been reported on the superiority of genotypic over phenotypic identification of CoNS. In 2004, a study was conducted to compare the identification of 168 CoNS isolates using both the *sodA* sequencing method and ID32 STAPH, a commercial system based on biochemical reactions. The ID32 STAPH failed to identify almost 40% of the isolates and the system was also found to be reliable only for *S. epidermidis* isolates. As such, the final identification had to rely on the *sodA* gene sequence analysis (Sivadon *et al*., 2004). In another study, phenotypic identification was performed on 47 clinical isolates of CoNS using two systems, the API Staph ID test and BD Phoenix Automated Microbiology system. Although the API Staph ID test was more reliable than the BD Phoenix Automated Microbiology system, the genotypic means via 16S rRNA, *tuf* and *sodA* gene sequencing method was found to be a more superior identification method for CoNS isolates (Heikens *et al*., 2005). Similarly, the 5' end of the 16S rDNA of 81 type and reference strains of staphylococcal were sequenced and the results revealed four ambiguities as compared to 13 (23.6%) and 19 (34.5%) isolates misidentified by ID 32 Staph and VITEK 2 (Becker *et al*., 2004).

While the conventional identification schemes based on physiological and biochemical tests are relatively time-consuming and unreliable, the commercial "rapid" identification systems share the problems of failure not only to identify commonly encountered bacteria but also in identifying uncommon isolates. A closer examination on the Staph ID identification matrix showed that the biochemical data was only generated from a limited number of individual isolates, especially in uncommon species thus bringing its reliability into question. The lack of adequate reference strains in the accompanying databases has also been reported in other commercial systems (Renneberg *et al*., 1995; Spanu *et al.,* 2003) hence providing ambiguous results or presenting two or more suggestions for identification with a comparable safety level. It is disturbing to know that even at a high percentage of probability of more than 99%, there is no guarantee of a correct identification of the CoNS isolates by Microgen Staph ID. This poor performance of the system could be due to the fact that phenotypic differences were established on only 13 biochemical substrates plus two preliminary tests which were insufficient to generate enough conclusive data for discrimination as there are more than 40 types of CoNS species and subspecies available. In addition, the system claimed to be able to identify 26 species and subspecies of *Staphylococcus*. However, there are more than 50 species and subspecies of *Staphylococcus* that has been documented, implying that the system would not be able to identify at least half of the *Staphylococcus* species. Moreover, because Microgen Staph ID is not an automated system, readings were recorded manually which was based on the colour development of the fermentation of the substrates. Hence, there were occasional problems in deciding whether some tests were weakly positive or negative which may contribute to the inaccuracy of the results.

REFERENCES

Bannerman, T.L. and W.E. Kloos, 1991. *Staphylococcus capitis* subsp. *ureolyticus* subsp. nov. from human skin. Int. J. Syst. Bacteriol., 41: 144-147.

Becker, K., C. Heilmann and G. Peters, 2014. Coagulase-negative staphylococci. Clin. Microbiol. Rev., 27: 870-926.

Becker, K., D. Harmsen, A. Mellmann, C. Meier, P. Schumann, G. Peters and C. von Eiff, 2004. Development and evaluation of a quality-controlled ribosomal sequence database for 16S ribosomal DNA-based identification of *Staphylococcus* species. J. Clin. Microbiol., 42: 4988-4995.

Bergeron, M., O. Dauwalder, M. Gouy, A.M. Freydiere and M. Bes *et al*., 2011. Species identification of staphylococci by amplification and sequencing of the *tuf* gene compared to the *gap* gene and by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Eur. J. Clin. Microbiol. Infect. Dis., 30: 343-354.

Bes, M., J. Freney, J. Etienne and V. Guerin-Faublee, 2002. Isolation of *Staphylococcus schleiferi* subspecies *coagulans* from two cases of canine pyoderma. Vet. Rec., 150: 487-488.

Casanova, C., L. Iselin, N. von Steiger, S. Droz and P. Sendi, 2011. *Staphylococcus hyicus* bacteremia in a farmer. J. Clin. Microbiol., 49: 4377-4378.

Chaudhury, A. and A.G. Kumar, 2007. *In vitro* activity of antimicrobial agents against oxacillin resistant staphylococci with special reference to *Staphylococcus haemolyticus*. Indian J. Med. Microbiol., 25: 50-52.

De Allori, M.C.G., M.A. Jure, C. Romero and M.E.C. de Castillo, 2006. Antimicrobial resistance and production of biofilms in clinical isolates of coagulase-negative *Staphylococcus* strains. Biol. Pharmaceut. Bull., 29: 1592-1596.

De Paulis, A.N., S.C. Predari, C.D. Chazarreta and J.E. Santoianni, 2003. Five-test simple scheme for species-level identification of clinically significant coagulase-negative staphylococci. J. Clin. Microbiol., 41: 1219-1224.

Delmas, J., J.P. Chacornac, F. Robin, P. Giammarinaro, R. Talon and R. Bonnet, 2008. Evaluation of the Vitek 2 system with a variety of *Staphylococcus* species. J. Clin. Microbiol., 46: 311-313.

Drancourt, M. and D. Raoult, 2002. *RpoB* gene sequence-based identification of *Staphylococcus* species. J. Clin. Microbiol., 40: 1333-1338.

Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39: 783-791.

Fudaba, Y., K. Nishifuji, L.O. Andresen, T. Yamaguchi, H. Komatsuzawa, M. Amagai and M. Sugai, 2005. *Staphylococcus hyicus* exfoliative toxins selectively digest porcine desmoglein 1. Microbial Pathogen., 39: 171-176.

Gatermann, S.G., T. Koschinski and S. Friedrich, 2007. Distribution and expression of macrolide resistance genes in coagulase-negative staphylococci. Clin. Microbiol. Infect., 13: 777-781.

Ghebremedhin, B., F. Layer, W. Konig and B. Konig, 2008. Genetic classification and distinguishing of *Staphylococcus* species based on different partial *gap*, 16S rRNA, *hsp60*, *rpoB*, *sodA* and *tuf* gene sequences. J. Clin. Microbiol., 46: 1019-1025.

Goh, S.H., S. Potter, J.O. Wood, S.M. Hemmingsen, R.P. Reynolds and A.W. Chow, 1996. HSP60 gene sequences as universal targets for microbial species identification: Studies with coagulase-negative staphylococci. J. Clin. Microbiol., 34: 818-823.

Gribaldo, S., B. Cookson, N. Saunders, R. Marples and J. Stanley, 1997. Rapid identification by specific PCR of coagulase-negative staphylococcal species important in hospital infection. J. Med. Microbiol., 46: 45-53.

Hajek, V., H. Meugnier, M. Bes, Y. Brun and F. Fiedler *et al*., 1996. *Staphylococcus saprophyticus* subsp. *bovis* subsp. nov., isolated from bovine nostrils. Int. J. Syst. Bacteriol., 46: 792-796.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acid Symp. Ser., 41: 95-98.

Heikens, E., A. Fleer, A. Paauw, A. Florijn and A.C. Fluit, 2005. Comparison of genotypic and phenotypic methods for species-level identification of clinical isolates of coagulase-negative staphylococci. J. Clin. Microbiol., 43: 2286-2290.

Kim, M., S.R. Heo, S.H. Choi, H. Kwon and J.S. Park *et al*., 2008. Comparison of the MicroScan, VITEK 2 and Crystal GP with 16S rRNA sequencing and MicroSeq 500 v2. 0 analysis for coagulase-negative staphylococci. BMC Microbiol., Vol. 8. 10.1186/1471-2180-8-233

Kloos, W.E. and J.F. Wolfshohl, 1991. *Staphylococcus cohnii* subspecies: *Staphylococcus cohnii* subsp. *cohnii* subsp. nov. and *Staphylococcus cohnii* subsp. *urealyticum* subsp. nov. Int. J. Syst. Bacteriol., 41: 284-289.

Kloos, W.E., D.N. Ballard, J.A. Webster, R.J. Hubner and A. Tomasz *et al*., 1997. Ribotype delineation and description of *Staphylococcus sciuri* subspecies and their potential as reservoirs of methicillin resistance and staphylolytic enzyme genes. Int. J. Syst. Bacteriol., 47: 313-323.

Kloos, W.E., D.N. Ballard, C.G. George, J.A. Webster and R.J. Hubner *et al*., 1998. Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., *Macrococcus bovicus* sp. nov. and *Macrococcus carouselicus* sp. nov. Int. J. Syst. Bacteriol., 48: 859-877.

Koksal, F., H. Yasar and M. Samasti, 2009. Antibiotic resistance patterns of coagulase-negative staphylococcus strains isolated from blood cultures of septicemic patients in Turkey. Microbiol. Res., 164: 404-410.

Kumar, S., K. Tamura and M. Nei, 2004. $MEGA_3$: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform., 5: 150-163.

Kwok, A.Y., S.C. Su, R.P. Reynolds, S.J. Bay, Y. Av-Gay, N.J. Dovichi and A.W. Chow, 1999. Species identification and phylogenetic relationships based on partial HSP60 gene sequences within the genus *Staphylococcus*. Int. J. Syst. Bacteriol., 49: 1181-1192.

Kwok, A.Y.C. and A.W. Chow, 2003. Phylogenetic study of Staphylococcus and Macrococcus species based on partial hsp60 gene sequences. Int. J. Syst. Evol. Microbiol., 53: 87-92.

Martineau, F., F.J. Picard, P.H. Roy, M. Ouellette and M.G. Bergeron, 1998. Species-specific and ubiquitous-DNA based assays for rapid identification of *Staphylococcus aureus*. J. Clin. Microbiol., 36: 618-623.

Mellmann, A., K. Becker, C. von Eiff, U. Keckevoet, P. Schumann and D. Harmsen, 2006. Sequencing and staphylococci identification. Emerg. Infect. Dis., 12: 333-336.

Piette, A. and G. Verschraegen, 2009. Role of coagulase-negative staphylococci in human disease. Vet. Microbiol., 134: 45-54.

Poyart, C., G. Quesne, C. Boumaila and P. Trieu-Cuot, 2001. Rapid and accurate species-level identification of coagulase-negative staphylococci by using the *sodA* gene as a target. J. Clin. Microbiol., 39: 4296-4301.

Rasigade, J.P., O. Raulin, J.C. Picaud, C. Tellini and M. Bes *et al*., 2012. Methicillin-resistant *Staphylococcus capitis* with reduced vancomycin susceptibility causes late-onset sepsis in intensive care neonates. PLoS One, Vol. 7. 10.1371/journal.pone.0031548

Renneberg, J., K. Rieneck and E. Gutschik, 1995. Evaluation of Staph ID 32 system and Staph-Zym system for identification of coagulase-negative staphylococci. J. Clin. Microbiol., 33: 1150-1153.

Saitou, N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406-425.

Shah, M.M., H. Iihara, M. Noda, S.X. Song and P.H. Nhung *et al*., 2007. *dnaJ* gene sequence-based assay for species identification and phylogenetic grouping in the genus *Staphylococcus*. Int. J. Syst. Evol. Microbiol., 57: 25-30.

Sivadon, V., M. Rottman, J.C. Quincampoix, V. Avettand and S. Chaverot *et al*., 2004. Use of *sodA* sequencing for the identification of clinical isolates of coagulase negative staphylococci. Clin. Microbiol. Infect., 10: 939-942.

Sivadon, V., M. Rottman, S. Chaverot, J.C. Quincampoix and V. Avettand *et al*., 2005. Use of genotypic identification by *sodA* sequencing in a prospective study to examine the distribution of coagulase-negative *Staphylococcus* species among strains recovered during septic orthopedic surgery and evaluate their significance. J. Clin. Microbiol., 43: 2952-2954.

Spanu, T., M. Sanguinetti, D. Ciccaglione, T. DInzeo, L. Romano, F. Leone and G. Fadda, 2003. Use of the VITEK 2 system for rapid identification of clinical isolates of Staphylococci from bloodstream infections. J. Clin. Microbiol., 41: 4259-4263.

Stoakes, L., M.A. John, R. Lannigan, B.C. Schieven, M. Ramos, D. Harley and Z. Hussain, 1994. Gas-liquid chromatography of cellular fatty acids for identification of staphylococci. J. Clin. Microbiol., 32: 1908-1910.

Stoll, B.J., N. Hansen, A.A. Fanaroff, L.L. Wright and W.A. Carlo *et al*., 2002. Late-onset sepsis in very low birth weight neonates: The experience of the NICHD Neonatal Research Network. Pediatrics, 110: 285-291.

Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22: 4673-4680.

Tong, S.Y.C., F. Schaumburg, M.J. Ellington, J. Corander and B. Pichon *et al*., 2015. Novel staphylococcal species that form part of a *Staphylococcus aureus*-related complex: The non-pigmented *Staphylococcus argenteus* sp. nov. and the non-human primate-associated *Staphylococcus schweitzeri* sp. nov. Int. J. Syst. Evol. Microbiol., 65: 15-22.

Vergnano, S., E. Menson, N. Kennea, N. Embleton and A.B. Russell *et al*., 2011. Neonatal infections in England: The NeonIN surveillance network. Arch. Dis. Childhood-Fetal Neonatal Edn., 96: F9-F14.