International Journal of Systematic and Evolutionary Microbiology (2007), **57**, 25–30 **DOI** 10.1099/ijs.0.64205-0

dnaJ gene sequence-based assay for species identification and phylogenetic grouping in the genus *Staphylococcus*

Mohammad Monir Shah, Hirotoshi Iihara, Makiko Noda, Sun Xiao Song, Pham Hong Nhung, Kiyofumi Ohkusu, Yoshiaki Kawamura and Takayuki Ezaki

Department of Microbiology, Regeneration and Advanced Medical Science, Gifu University Graduate School of Medicine, Japan

Correspondence
Takayuki Ezaki
tezaki@gifu-u.ac.jp

In the last few years, many attempts have been made to use conserved gene sequences for identification and for phylogenetic studies of *Staphylococcus* species. In an effort to identify a more reliable approach, a *dnaJ* gene sequence-based database was created. In this study, an approximately 883 bp portion of the *dnaJ* gene sequence from 45 staphylococcal type strains was compared with 16S rRNA and other conserved gene (*hsp60*, *sodA* and *rpoB*) sequences available in public databases. Nucleotide sequence comparisons revealed that the staphylococcal *dnaJ* gene showed higher discrimination (mean similarity 77.6 %) than the 16S rRNA (mean similarity 97.4 %), *rpoB* (mean similarity 86 %), *hsp60* (mean similarity 82 %) and *sodA* (mean similarity 81.5 %) genes. Analysis of the *dnaJ* gene sequence from 20 *Staphylococcus* isolates representing two clinically important species showed <1 % sequence divergence. Phylogenetic data obtained from the *dnaJ* gene sequence were in general agreement with those of 16S rRNA gene sequence analysis and DNA–DNA reassociation studies. In conclusion, the *dnaJ* gene sequence-based assay is an effective alternative to currently used methods, including 16S rRNA gene sequencing, for identification and taxonomical analysis of *Staphylococcus* species.

Members of the genus *Staphylococcus* are the predominant pathogens in hospital-acquired infections. Methods for accurate identification of staphylococcal (sub)species have undergone rapid development over the past two decades in response to the spread of multidrug-resistant staphylococcal strains and the growing number of novel species. Several commercial kits and automated methods based on phenotypic characteristics have been developed for the identification and detection of clinically important staphylococcal species, but the overall accuracy is low due to phenotypic differences between strains of the same species (Becker *et al.*, 2004).

Sequence-based identification of micro-organisms is becoming a useful and reliable alternative to phenotypic methods. Analysis of 16S rRNA gene sequences is the most commonly used method for identification and classification of bacteria, including staphylococci; however, the usefulness of 16S rRNA gene sequences is limited because of the high degree of sequence similarity between closely related species (Gribaldo *et al.*, 1997; Becker *et al.*, 2004; Skow *et al.*, 2005). Recently, partial sequences of the highly conserved *hsp60* (Kwok *et al.*, 1999; 2003), *sodA* (Poyart *et al.*, 2001) and *rpoB* (Drancourt & Raoult, 2002; Mellmann *et al.*, 2006) genes have been found to be useful in the identification of *Staphylococcus* at the species level. However, a universal DNA target with well-conserved DNA sequences within a given species, but with sufficient sequence variation to discriminate between species and subspecies, is needed.

DnaJ, also known as Hsp40, is a member of the heat-shock protein (Hsp) family, distributed ubiquitously in Eukarya, Bacteria and Archaea (Ang *et al.*, 1991; Macario *et al.*, 1993). The *dnaJ* gene has been shown to be more discriminative than the 16S rRNA gene for identifying species of the genus *Streptococcus* (Itoh *et al.*, 2006), *Legionella pneumophila* serogroups (Liu *et al.*, 2003) and *Mycobacterium* species (Takewaki *et al.*, 1993; Victor *et al.*, 1996). In the present study, we developed a *Staphylococcus*-specific *dnaJ* gene sequence-based assay for the identification and establishment of phylogenetic relationships among (sub)species of the genus *Staphylococcus*.

The GenBank/EMBL/DDBJ accession numbers for the *dnaJ* gene sequences of the staphylococcal strains examined in this study are listed in Table 1.

An analysis of *dnaJ* gene sequence similarity among the staphylococcal type strains used in this study is presented in a supplementary table in IJSEM Online.

A total of 45 staphylococcal type strains (Table 1), 20 isolates of two clinically important species, *Staphylococcus aureus* subsp. *aureus* (n = 10) and *Staphylococcus epidermidis* (n = 10) and eight non-staphylococcal strains as negative controls (*Macrococcus caseolyticus*, *Bacillus subtilis*, *Micrococcus luteus*, *Enterococcus faecalis*, *Streptococcus pyogenes*, *Lactococcus lactis*, *Escherichia coli* and *Pseudomonas aeruginosa*) were used in this study. Some recently described species (Euzéby, 1997; http://www.bacterio.cict.fr/, updated February 2006) were not available for our study. All strains were grown on brain–heart infusion agar at 37 °C for 24 h under aerobic conditions, with the exception of *S. aureus* subsp. *anaerobius* and *Staphylococcus saccharolyticus*, which were grown on Columbia blood (5 % defibrinated sheep blood) agar under anaerobic conditions.

Genomic DNA from all strains was extracted as described previously (Ezaki *et al.*, 1990). PCR amplification and DNA sequencing were performed with the following pair of *dnaJ* degenerate primers: SA-(F) 5′-GCCAAAAGAGACTATT-ATGA-3′ and SA-(R) 5′-ATTGYTTACCYGTTTGTGTA-CC-3′. The PCR mixture was prepared with the Takara *Ex Taq* Hot Start version kit (Takara) according to the manufacturer's instructions. Reaction mixtures were first incubated at 94 °C for 3 min, followed by five cycles of 94 °C for 30 s, 45 °C for 30 s and 72 °C for 60 s and were then subjected to 30 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 60 s and completed with a final extension at 72 °C for 3 min. PCR products were purified with a Wizard SV Gel and PCR clean-up system (Promega) and sequenced with a BigDye Terminator v3.1 cycle sequencing kit on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). To exclude the possibility of sequencing errors attributable to misincorporation by *Taq* DNA polymerase, all PCR amplicons were sequenced twice with PCR products obtained from two independent rounds of PCR. For purposes of comparison, approximately 1450 bp of the 16S rRNA gene of ten *Staphylococcus* species were amplified and sequenced with 16S rRNA gene universal primers. Additional 16S rRNA and available conserved gene sequences (*hsp60*, *sodA* and *rpoB*) of corresponding *Staphylococcus* species were downloaded from the GenBank/EMBL/DDBJ database (Table 1).

Multiple sequence alignments were performed with CLUSTAL W software (Thompson *et al.*, 1994). Evolutionary distances were calculated by Kimura's two-parameter model (Kimura, 1983). Phylogenetic trees were constructed by the neighbour-joining (Saitou & Nei, 1987) and maximum-parsimony (Fitch, 1971) methods, with MEGA3 software (Kumar *et al.*, 2004) and with bootstrap values based on 1000 replications (Felsenstein, 1985). Tree figures were drawn by TREEVIEW software (Page, 1996).

In this study, an approximately 920 bp internal fragment of the *dnaJ* gene from 45 staphylococcal type strains was amplified with the use of *dnaJ* universal primers. Amplification products were not obtained from non-staphylococcal species. After we omitted the primer sequences, approximately 883 bp sequences were used for analysis and the resulting sequences were deposited in the DNA Data Bank of Japan (DDBJ) and assigned accession numbers as listed in Table 1. Phylogenetic analysis by the neighbour-joining (Fig. 1) and maximum-parsimony methods (data not shown) produced similar trees with the exception of minor differences in the tree topology of the basal branches.

dnaJ gene sequence analysis for the identification of staphylococcal species and subspecies

Analysis of DNA sequence similarity among staphylococcal type strains showed that the *dnaJ* gene sequence (mean similarity 77.6 %) (see Supplementary Table S1 available in IJSEM Online), is more discriminative than the sequences of other conserved genes such as 16S rRNA (mean similarity 97.4 %), *rpoB* (mean similarity 86 %) (Drancourt & Raoult, 2002; Mellmann *et al.*, 2006), *hsp60* (mean similarity 82 %) (Kwok *et al.*, 1999; 2003) and *sodA* (mean similarity 81.5 %) (Poyart *et al.*, 2001). In addition, the evolutionary substitution rate of the *dnaJ* sequence was much faster than that of the 16S rRNA gene sequence.

However, at the interspecies level, the *dnaJ* gene sequence showed remarkable discrimination, (71.1–88.8 %), except for the most similar pairs such as *S. condimenti* and *S. carnosus* (90.9 %), *S. intermedius* and *S. delphini* (92.3 %) and *S. pulvereri* and *S. vitulinus* (99.3 %). Recently, *S. pulvereri* was reclassified as a later heterotypic synonym of *S. vitulinus* (Švec *et al.*, 2004) and this was also inferred in our *dnaJ* gene sequence analysis (see Supplementary Table S1). All staphylococcal species showed more than 90 % similarity in their 16S rRNA gene sequences. *S. epidermidis*, *S. muscae*, *S. saccharolyticus*, *S. sciuri* and *S. vitulinus* species showed 90.1–98.6 % similarity in their *hsp60 gene* sequences (Kwok *et al.*, 1999; 2003). The species *S. capitis*, *S. caprae*, *S. carnosus*, *S. cohnii*, *S. condimenti*, *S. delphini*, *S. equorum*, *S. gallinarum*, *S. intermedius*, *S. kloosii*, *S. pasteuri*, *S. piscifermentans*, *S. saprophyticus*, *S. sciuri*, *S. simulans*, *S. vitulinus*, *S. warneri* and *S. xylosus* showed 90–96.8 % similarity in their *sodA* gene sequences (Poyart *et al.*, 2001). When the *rpoB* gene sequence was examined, *S. auricularis*, *S. capitis*, *S. caprae*, *S. carnosus*, *S. cohnii*, *S. condimenti*, *S. chromogenes*, *S. equorum*, *S. fleurettii*, *S. gallinarum*, *S. haemolyticus*, *S. hominis*, *S. hyicus*, *S. kloosii*, *S. nepalensis*, *S. pasteuri*, *S. piscifermentans*, *S. saprophyticus*, *S. sciuri*, *S. succinus*, *S. vitulinus*, *S. warneri* and *S. xylosus* showed 90–96.5 % gene sequence similarity (Drancourt & Raoult, 2002; Mellmann *et al.*, 2006). These results indicate that the *dnaJ* sequence may constitute a more useful target sequence than that of other conserved genes, including the 16S rRNA gene, in the discrimination of closely related species.

At the subspecies level, *dnaJ* gene sequences can discriminate only *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticus* (7 % sequence divergence) and *S. aureus* subsp. *aureus* and *S. aureus* subsp. *anaerobius* (3.7 % sequence divergence).

Table 1. List of staphylococcal type strains and accession numbers investigated by *dnaJ* and 16S rRNA gene sequences

GTC, Gifu Type Culture Collection; ATCC, American Type Culture Collection; DSM, Deutsche Sammlung von Mikrooganismen und Zellkulturen; CCM, Czechoslovak Collection of Microorganisms; NCTC, National Collection of Type Cultures; JCM, Japan Collection of Microorganisms; PCM, Polish Collection of Microorganisms.

Strain no.	*Staphylococcus* species and subspecies	Strain	GenBank accession nos	
			***dnaJ* gene**	**16S rRNA gene**
1	*S. arlettae*	GTC 383^T (=DSM 20672^T)	AB234056	–
2	*S. aureus* subsp. *anaerobius*	GTC 371^T (=ATCC 35844^T)	AB234057	–
3	*S. aureus* subsp. *aureus*	GTC 286^T (=ATCC 12600^T)	AB234058	–
4	*S. auricularis*	GTC 326^T (=ATCC 33753^T)	AB234059	–
5	*S. capitis* subsp. *capitis*	GTC 287^T (=ATCC 27840^T)	AB234060	–
6	*S. capitis* subsp. *ureolyticus*	GTC 727^T (=ATCC 49326^T)	AB234061	AB233325
7	*S. caprae*	GTC 378^T (=CCM 3573^T)	AB234062	–
8	*S. carnosus* subsp. *carnosus*	GTC 384^T (=DSM 20501^T)	AB234063	–
9	*S. carnosus* subsp. *utilis*	GTC 1232^T (DSM 11676^T)	AB234064	AB233329
10	*S. chromogenes*	GTC 482^T (=NCTC 10530^T)	AB234065	–
11	*S. cohnii* subsp. *cohnii*	GTC 288^T (=ATCC 29974^T)	AB234066	–
12	*S. cohnii* subsp. *urealyticus*	GTC 728^T (=ATCC 49330^T)	AB234067	AB233328
13	*S. condimenti*	GTC 1233^T (=DSM 11674^T)	AB234068	–
14	*S. delphini*	GTC 476^T (=ATCC 49171^T)	AB234319	–
15	*S. epidermidis*	GTC 289^T (=ATCC 14990^T)	AB234069	–
16	*S. equorum*	GTC 381^T (=DSM 20674^T)	AB234070	–
17	*S. felis*	GTC 509^T (=JCM 7469^T)	AB234071	–
18	*S. fleurettii*	GTC 1999^T (=DSM 13212^T)	AB234072	AB233330
19	*S. gallinarum*	GTC 379^T (=CCM 3572^T)	AB234073	–
20	*S. haemolyticus*	GTC 290^T (=ATCC 29970^T)	AB234074	–
21	*S. hominis* subsp. *hominis*	GTC 485^T (=ATCC 27844^T)	AB234075	–
22	*S. hominis* subsp. *novobiosepticus*	GTC 1228^T (=ATCC 700236^T)	AB234076	AB233326
23	*S. hyicus*	GTC 305^T (=ATCC 11249^T)	AB234077	–
24	*S. intermedius*	GTC 266^T (=ATCC 29663^T)	AB234078	–
25	*S. kloosii*	GTC 382^T (=DSM 20676^T)	AB234079	–
26	*S. lentus*	GTC 267^T (=ATCC 29070^T)	AB234080	–
27	*S. lugdunensis*	GTC 458^T (=ATCC 43809^T)	AB234081	–
28	*S. lutrae*	GTC 1248^T (=DSM 10244^T)	AB234082	AB233333
29	*S. muscae*	GTC 704^T (=ATCC 49910^T)	AB234083	–
30	*S. pasteuri*	GTC 705^T (=ATCC 51129^T)	AB234084	–
31	*S. piscifermentans*	GTC 585^T (=JCM 6057^T)	AB234085	–
32	*S. pulvereri**	GTC 722^T (=PCM 2443^T)	AB234086	–
33	*S. saccharolyticus*	GTC 181^T (=ATCC 14953^T)	AB234087	–
34	*S. saprophyticus* subsp. *bovis*	GTC 843^T (=CCM 4410^T)	AB234088	AB233327
35	*S. saprophyticus* subsp. *saprophyticus*	GTC 265^T (=ATCC 15305^T)	AB234089	–
36	*S. schleiferi* subsp. *coagulans*	GTC 510^T (=JCM 7470^T)	AB234320	AB233334
37	*S. schleiferi* subsp. *schleiferi*	GTC 483^T (=ATCC 43808^T)	AB234321	–
38	*S. sciuri* subsp. *carnaticus*	GTC 1227^T (=ATCC 700058^T)	AB234322	AB233331
39	*S. sciuri* subsp. *rodentium*	GTC 844^T (=CCM 4657^T)	AB234323	AB233332
40	*S. sciuri* subsp. *sciuri*	GTC 291^T (=ATCC 29062^T)	AB234324	–
41	*S. simulans*	GTC 292^T (=ATCC 27848^T)	AB234325	–
42	*S. succinus*	GTC 1256^T (=ATCC 700337^T)	AB234326	–
43	*S. vitulinus*	GTC 729^T (=ATCC 51145^T)	AB234327	–
44	*S. warneri*	GTC 293^T (=ATCC 27836^T)	AB234328	–
45	*S. xylosus*	GTC 294^T (=ATCC 29971^T)	AB234329	–

**Staphylococcus pulvereri* reclassified as a later synonym of *Staphylococcus vitulinus*.

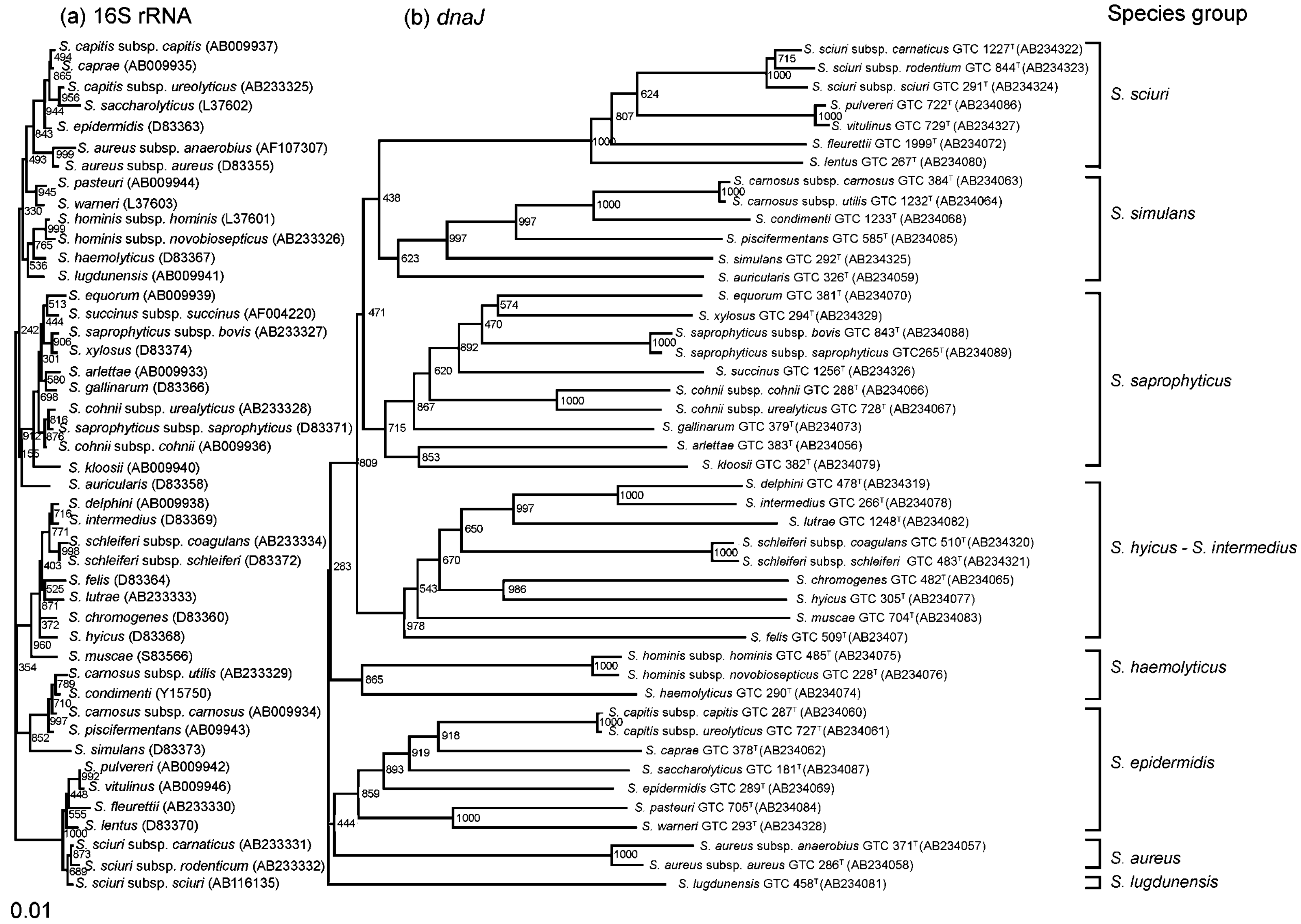

Fig. 1. Phylogenetic trees (unrooted) based on 16S rRNA (a) and *dnaJ* (b) gene sequences showing the relations among 45 recognized *Staphylococcus* species and subspecies. The tree was constructed with the neighbour-joining method (Saitou & Nei, 1987). Bootstrap values, indicated at branches, were calculated from 1000 resamplings. Bar, 0.01 substitutions per nucleotide position.

The remaining subspecies, *S. capitis* subsp. *capitis*, *S. capitis* subsp. *ureolyticus*, *S. carnosus* subsp. *carnosus*, *S. carnosus* subsp. *utilis*, *S. hominis* subsp. *hominis*, *S. hominis* subsp. *novobiosepticus*, *S. saprophyticus* subsp. *bovis*, *S. saprophyticus* subsp. *saprophyticus*, *S. schleiferi* subsp. *coagulans*, *S. schleiferi* subsp. *schleiferi*, *S. sciuri* subsp. *rodentium*, *S. sciuri* subsp. *sciuri* and *S. sciuri* subsp. *carnaticus*, showed more than 97.3 % gene sequence similarity and did not allow for complete discrimination at the subspecies level. This finding was consistent with that for other conserved genes, including the *hsp60* gene which showed 9 % sequence divergence between *S. capitis* subsp. *capitis* and *S. capitis* subsp. *ureolyticus* and 7 % sequence divergence between *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticus*; the *sodA* gene which showed 4.4 % sequence divergence between *S. cohnii* subsp. *cohnii* and *S. cohnii* subsp. *urealyticus* and the *rpoB* gene which showed 12.7 % sequence divergence between *S. aureus* subsp. *aureus* and *S. aureus* subsp. *anaerobius* (Drancourt & Raoult, 2002), although it was recently reported that obligately anaerobic *S. aureus* could not be identified on the basis of *rpoB* sequences (Peake *et al.*, 2006). However, the remaining subspecies of *Staphylococcus* showed more than 98 % gene sequence similarity for the *hsp60*, *sodA* and *rpoB* genes (Poyart *et al.*, 2001; Drancourt & Raoult, 2002; Kwok & Chow, 2003; Mellmann *et al.*, 2006).

In addition, 20 isolates of *Staphylococcus* representing two clinically important species, *S. aureus* subsp. *aureus* and *S. epidermidis*, were used for the analysis of intraspecies sequence variation. Sequence divergence was 0.7 % among *S. aureus* isolates and 0.4 % among *S. epidermidis* isolates (data not shown) indicating that a *dnaJ* gene sequence database would be useful in the identification of clinical isolates.

Use of the *dnaJ* and other conserved genes to compare phylogenetic relationships among *Staphylococcus* species

Taxonomic studies based on DNA–DNA reassociation and 16S rRNA and *hsp60* gene sequence analysis have indicated that genealogically distinct species groups exist in the genus *Staphylococcus* (Kloos & Schleifer, 1986; Takahashi *et al.*, 1999). However, in many cases, low resolution and lack of congruity was observed to delineate these groupings. For example, DNA–DNA reassociation studies indicated the existence of nine staphylococcal species groups, whereas 16S rRNA and *hsp60* gene sequence-based analyses indicated twelve and six genogroups, respectively (Takahashi *et al.*, 1999; Kwok & Chow, 2003). Moreover, both DNA–DNA reassociation and 16S rRNA gene sequence-based studies included *Staphylococcus caseolyticus* in their groups. *Staphylococcus caseolyticus* has since been reclassified as *Macrococcus caseolyticus* (Kloos *et al.*, 1998), thus limiting the number of species included in DNA–DNA reassociation studies and 16S rRNA gene sequence-derived genogroups.

Species groupings have also been observed on the basis of phenotypic characteristics. For example, novobiocin resistance and oxidase activity can differentiate the *S. saprophyticus* and *S. sciuri* groups from other groups. However, other well-described phenotypic characteristics are not shared by all species due to their diverse phenotypic relationships (Kloos *et al.*, 1991).

In the present study, *dnaJ* gene sequence analysis was compared with 16S rRNA gene sequence and DNA–DNA reassociation studies in an effort to determine more reliable species relationships in the genus *Staphylococcus*. Phylogenetic analysis based on *dnaJ* gene sequences yielded eight distinct species groups with relatively high bootstrap values (80–100 %); the *S. sciuri* group, *S. simulans* group, *S. saprophyticus* group, *S. hyicus–intermedius* group, *S. haemolyticus* group, *S. epidermidis* group, *S. aureus* group and *S. lugdunensis* group (Fig. 1). Interestingly, these groups showed nearly identical relationships with those from 16S rRNA gene sequence and DNA–DNA reassociation studies, with the exception of some minor differences in the basal branches. In the *dnaJ* gene sequence tree, the *S. simulans* species group (*S. carnosus* subsp. *carnosus*, *S. carnosus* subsp. *utilis*, *S. condimenti*, *S. piscifermentans* and *S. simulans*) produced a monophyletic clade with a high bootstrap value of 99 % at each node, except for the branching of *S. auricularis*. This finding is inconsistent with the data from the 16S rRNA gene sequence and DNA–DNA reassociation studies and requires further analysis. In addition, *S. warneri* and *S. pasteuri* produced a deep subline within the *S. epidermidis* group consistent with *sodA* gene sequence analysis and DNA–DNA reassociation studies, whereas *S. simulans* and *S. carnosus* and *S. warneri* and *S. epidermidis* produced separate species groups in the 16S rRNA gene sequence tree, with low resolution. However, in a previous study (Takahashi *et al.*, 1999) 16S rRNA gene sequence-based analysis yielded twelve genogroups, possibly due to the limited number of species used or the low discriminatory power to produce reliable branches. Moreover, similar to the 16S rRNA gene, the *dnaJ* gene sequence analysis data also showed minor differences with those of DNA–DNA reassociation studies, inferring close relationships of the *S. hyicus–intermedius* and *S. epidermidis* species groups.

In contrast to the topologies of the *dnaJ* and 16S rRNA gene sequence trees, *hsp60* gene sequences yielded two cluster groups. Cluster 1 contained the *S. aureus* group (group a), the *S. epidermidis* group (group b), the *S. haemolyticus* group (group c), the *S. saprophyticus* group (group d) and the *S. intermedius* group (group e). Cluster 2 contained the *S. sciuri* group (Kwok & Chow, 2003). However, in the *hsp60* gene sequence tree, species relationships in the *S. aureus* and *S. haemolyticus* groups were not apparent compared with those of other species groups and showed discordance with the *dnaJ* and 16S rRNA gene sequence trees. In the *sodA* gene phylogenetic tree, *S. sciuri*, *S. simulans*, *S. epidermidis*, *S. saprophyticus*, *S. hyicus–intermedius* and *S. lugdunensis* species groups showed good concordance with the *dnaJ* gene sequence tree topology, with the exception of *S. haemolyticus*, *S. hominis* and *S. schleiferi* (Poyart *et al.*, 2001).

In the *rpoB* tree, the positions of *S. cohnii*, *S. haemolyticus*, *S. hominis* and *S. warneri* inferred an inconsistent relationship with those of the *dnaJ* and 16S rRNA gene sequence trees (Drancourt & Raoult, 2002). Nevertheless, the *dnaJ* gene sequence-based phylogenetic analysis showed remarkable concordance with the DNA–DNA reassociation, 16S rRNA, *hsp60* and *sodA* gene sequence-based analysis and showed greater power to discriminate species relationships in the genus *Staphylococcus*.

In conclusion, our *dnaJ* gene sequence-based assay offers a reliable alternative to currently used methods, including the 16S rRNA gene sequence-based assay, for the identification and phylogenetic analysis of almost all staphylococci at the species level and, with some exceptions, at the subspecies level. We are currently studying the utility of this *dnaJ* gene sequence-based assay in the identification and phylogenetic analysis of other bacterial genera.

References

Ang, D., Liberek, K., Skowyra, D., Zylicz, M. & Georgopoulos, C. (1991). Biological role and regulation of the universally conserved heat shock proteins. *J Biol Chem* **266**, 24233–24236.

Becker, K., Harmsen, D., Mellmann, A., Meier, C., Schumann, P., Peters, G. & von Eiff, C. (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.

Drancourt, M. & Raoult, D. (2002). *rpoB* gene sequence-based identification of *Staphylococcus* species. *J Clin Microbiol* **40**, 1333–1338.

Euzéby, J. P. (1997). List of bacterial names with standing nomenclature: a folder available on the Internet. *Int J Syst Bacteriol* **47**, 590–592.

Ezaki, T., Saidi, S. M., Liu, S. L., Hashimoto, Y., Yamamoto, H. & Yabuuchi, E. (1990). Rapid procedure to determine the DNA base composition from small amounts of Gram-positive bacteria. *FEMS Microbiol Lett* **67**, 127–130.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* **20**, 406–416.

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

Itoh, Y., Kawamura, Y., Kasai, H., Shah, M. M., Nhung, P. H., Yamada, M., Sun, X., Koyana, T., Hayashi, M. & other authors (2006). *dnaJ* and *gyrB* gene sequence relationship among species and strains of genus *Streptococcus*. *Syst Appl Microbiol* **29**, 368–374.

Kimura, M. (1983). *The Natural Theory of Molecular Evolution*, pp. 41–46. Cambridge: Cambridge University Press.

Kloos, W. E. & Schleifer, K.-H. (1986). Genus IV *Staphylococcus*. In *Bergey's Manual of Systematic Bacteriology*, vol. 2, pp. 1013–1035. Edited by P. H. A. Sneath, N. S. Mair, M. E. Sharpe & J. G. Holt. Baltimore, MD: Williams & Wilkins.

Kloos, W. E., Schleifer, K.-H. & Götz, F. (1991). The genus *Staphylococcus*. In *The Prokaryotes*, 2nd edn. pp. 1369–1420. Edited by A. Balows, H. G. Truper, M. Dworkin, W. Harder & K. H. Schleifer. New York: Springer.

Kloos, W. E., Ballard, D. N., George, C. G., Webster, J. A., Hubner, R. J., Ludwig, W., Schleifer, K. H., Fiedler, F. & Schubert, K. (1998). Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., and *Macrococcus bovicus* sp. nov. and *Macrococcus carouselicus* sp. nov. *Int J Syst Bacteriol* **48**, 859–877.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

Kwok, A. Y. & Chow, A. W. (2003). Phylogenetic study of *Staphylococcus* and *Macrococcus* species based on partial *hsp60* gene sequences. *Int J Syst Evol Microbiol* **53**, 87–92.

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

Liu, H., Li, Y., Huang, X., Kawamura, Y. & Ezaki, T. (2003). Use of the *dnaJ* gene for the detection and identification of all *Legionella pneumophila* serogroups and description of the primers used to detect 16S rDNA gene sequences of major members of the genus *Legionella*. *Microbiol Immunol* **47**, 859–869.

Macario, A. J., Dugan, C. B., Clarens, M. & Conway de Macario, E. (1993). *dnaJ* in Archaea. *Nucleic Acids Res* **21**, 2773.

Mellmann, A., Becker, K., von Eiff, C., Keckevoet, U., Schumann, P. & Harmsen, D. (2006). Sequencing and staphylococci identification. *Emerg Infect Dis* **12**, 333–336.

Page, R. D. (1996). TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**, 357–358.

Peake, S. L., Peter, J. V., Chan, L., Wise, R. P., Butcher, A. R. & Grove, D. I. (2006). First report of septicemia caused by an obligately anaerobic *Staphylococcus aureus* infection in a human. *J Clin Microbiol* **44**, 2311–2313.

Poyart, C., Quesne, G., Boumaila, C. & Trieu-Cuot, P. (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.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Skow, A., Mangold, K. A., Tajuddin, M., Huntington, A., Fritz, B., Thomson, R. B., Jr & Kaul, K. L. (2005). Species-level identification of staphylococcal isolates by real-time PCR and melt curve analysis. *J Clin Microbiol* **43**, 2876–2880.

Švec, P., Vancanneyt, M., Sedláček, I., Engelbeen, K., Štětina, V., Swings, J. & Petráš, P. (2004). Reclassification of *Staphylococcus pulvereri* Zakrzewska-Czerwińska *et al.* 1995 as a later synonym of *Staphylococcus vitulinus* Webster *et al.* 1994. *Int J Syst Evol Microbiol* **54**, 2213–2215.

Takahashi, T., Satoh, I. & Kikuchi, N. (1999). Phylogenetic relationships of 38 taxa of the genus *Staphylococcus* based on 16S rRNA gene sequence analysis. *Int J Syst Bacteriol* **49**, 725–728.

Takewaki, S., Okuzumi, K., Ishiko, H., Nakahara, K., Ohkubo, A. & Nagai, R. (1993). Genus-specific polymerase chain reaction for the mycobacterial *dnaJ* gene and species-specific oligonucleotide probes. *J Clin Microbiol* **31**, 446–450.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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.

Victor, T. C., Jordaan, A. M., Van Schalkwyk, E. J., Coetzee, G. J. & Van Helden, P. D. (1996). Strain-specific variation in the *dnaJ* gene of mycobacteria. *J Med Microbiol* **44**, 332–339.