Phylogenetic study of *Staphylococcus* and *Macrococcus* species based on partial *hsp60* gene sequences

Anita Y. C. Kwok\(^1,3\) and Anthony W. Chow\(^1,2,3\)

\(^1,2\)Division of Infectious Diseases, Departments of Medicine\(^1\), Microbiology and Immunology\(^2\), University of British Columbia, Vancouver, Canada

\(^3\)Canadian Bacterial Disease Network, Vancouver, British Columbia, Canada

A 600 bp partial *hsp60* gene sequence has been described previously as a novel genetic marker for species identification and phylogenetic studies within the genus *Staphylococcus*. In the present study, the 600 bp partial *hsp60* gene sequences of 40 validly described *Staphylococcus* species and subpecies and four *Macrococcus* species were PCR-amplified and sequenced. Phylogenetic analysis revealed excellent concordance between the unrooted dendrograms based on partial *hsp60* and 16S rRNA gene sequences. The genus *Macrococcus* is clearly separated from the genus *Staphylococcus*, but is closely related to the 'sciuri' group, the only staphylococci that are oxidase-positive. The remaining *Staphylococcus* species clustered into five broad-based subdivisions, which corresponded to the 'aureus' group, the 'epidermidis' group, the 'haemolyticus' group, the 'saprophyticus' group and the 'intermedius' group. These results agreed remarkably well with the current taxonomy of this diverse family, which is based on classical phenotypic and biochemical testing. Furthermore, pairwise sequence comparisons indicated that the *hsp60* gene is more divergent and more discriminatory than the 16S rRNA gene for species differentiation among strains of the genera *Staphylococcus* and *Macrococcus*. It is concluded that the *hsp60* gene may be an efficient alternative target for taxonomic and phylogenetic studies on members of these genera.

At the time of writing and according to the List of Bacterial Names with Standing in Nomenclature (Euzéby, 1997; http://www.bacterio.cict.fr/), the genus *Staphylococcus* consists of 36 species, most of which are coagulase-negative. Nine species of the genus also contain subdivisions with subspecies designation. Additionally, *Staphylococcus caseolyticus* has been assigned to a novel genus, *Macrococcus* (Kloos et al., 1998), which currently contains four species (*Macrococcus bovicus*, *Macrococcus caseolyticus*, *Macrococcus carousselicus* and *Macrococcus equipericius*). Members of the genus *Macrococcus* are all coagulase-negative and catalase-positive, and can be distinguished phenotypically from most staphylococci on the basis of their cellular morphology (they are 2–5–4 times larger in diameter compared to *Staphylococcus aureus*) and their positive cytochrome c oxidase reaction. Although *S. aureus* is the most common and important coagulase-positive staphylococcal species causing human disease, other staphylococci, including *Staphylococcus intermedius*, *Staphylococcus delphini*, *Staphylococcus schleiferi* subsp. coagulans, *Staphylococcus lutrae* and some strains of *Staphylococcus hyicus*, are also coagulase-positive and have been implicated in some human infections (Kloos & Bannerman, 1995; Foster et al., 1997; Martin de Nicolas et al., 1995). In addition, whereas coagulase-negative staphylococci are generally considered to have a low virulence potential, some coagulase-negative staphylococci (e.g. *Staphylococcus epidermidis*, *Staphylococcus haemolyticus*, *Staphylococcus lugdenensis*, *S. schleiferi*, etc.) have assumed increasing medical importance in hospital-acquired infections (Huebner & Goldmann, 1999; Karchmer, 2000; Richards et al., 2000). Thus, there is a pressing need to accurately identify and speciate clinically important staphylococci strains. However, despite the presence of phenotypic and genotypic differences among different *Staphylococcus* species, available methods for their identification in both clinical and reference laboratories are either cumbersome or lack both sensitivity and specificity (Birnbaum et al., 1994; Endl et al., 1984; Weinstein et al., 1998). Several molecular taxonomic methods, including DNA–DNA hybridization and 16S rRNA sequencing, as well as various PCR-based techniques, have been reported for the identification and
phylogenetic study of staphylococci (De Buyser et al., 1992; Freney et al., 1999). We have previously demonstrated the presence of hypervariable sequences within the ubiquitous and highly conserved gene encoding the bacterial 60 kDa heat-shock protein (Hsp60) (Goh et al., 1996; Kwok et al., 1999). In this communication, a 600 bp fragment of the hsp60 gene of 44 reference strains of different staphylococci and macrococci (including 34 different Staphylococcus species, nine with subspecies designation, and four Macroccocus species; Fig. 1) were PCR-amplified, and their phylogenetic relationships based on hsp60 sequences were analysed and compared with those based on 16S rRNA sequences. The two most recently described species (Staphylococcus condimenti and Staphylococcus fleurettii) were not included in this study. All strains were grown in brain–heart infusion (BHI) broth and were subcultured overnight on BHI agar plates to yield single colonies for genomic DNA extraction using the InstaGene purification kit (Bio-Rad).

PCR amplification was performed using a pair of hsp60 universal primers, H279 and H280, as described previously (Goh et al., 1996). These primers, with nucleotide sequences of 5’-GAATTCGAAIIIICGGIGA(G)TGIAACIAAC-3’ and 5’-CGCGGGAATTCG(TC)(TG)II(TC)(TG)ITCCICC(G)AAIICGGGI(GC)(TC)TT-3’, respectively, amplify an anticipated 600 bp hsp60 DNA fragment. H279 had an EcoRI restriction site, while H280 had a BamHI restriction site (underlined in the above sequences). The PCR mixture contained 50 ng InstaGene DNA extract, 50 mM KCl, 10 mM Tris (pH 8.3), 1.5 mM MgCl₂ (Gibco), 200 μM each dNTP (Pharmacia Biotech), 2-5 U Taq DNA polymerase (Gibco) and 0-5 μM each of H279 and H280. The PCR thermal cycling conditions used were one cycle at 95 °C for 3 min, followed by 30 cycles at 94 °C for 30 s, 37 °C for 30 s and 72 °C for 1 min, with a final cycle at 72 °C for 10 min, to allow complete extension of all of the PCR amplicons. After PCR amplification, the amplicon was either digested with EcoRI and BamHI following cloning into pUC19 or it was directly cloned using a TA cloning vector, pCR2.1 (Invitrogen). Plasmid DNA was purified using the Wizard Plasmid Miniprep kit (Promega). For sequencing of the 600 bp region, the 600 bp hsp60 DNA fragment was PCR-amplified using modified H279 and H289 primers containing M13 primer-binding sites on the ends (Goh et al., 1997). The amplicon was then purified using the QiaQuick PCR purification kit or the QIAEX II gel extraction kit (Qiagen), prior to sequencing using an automated DNA sequencer (model 373A; Applied Biosystems).

Sequence analysis was performed with the entire 600 bp amplicon, omitting the primer sequences used to amplify the hsp60 genes. Pairwise- and multiple-sequence alignments were performed using the CLUSTAL W program, version 1.7 (Thompson et al., 1997). Phylogenetic analysis was performed using the PHYLIP package, version 3.57 (Felsenstein, 1995). The unrooted phylogenetic tree was constructed by the neighbour-joining method (Saitou & Nei, 1987), using the Jukes–Cantor one-parameter model to correct for multiple superimposed substitutions (Jukes & Cantor, 1969). The degree of data support for the tree topology was quantified by the bootstrap method, using 1000 iterations. For comparison purposes, the published 16S rRNA gene sequences of corresponding Staphylococcus and Macroccocus species were downloaded from the GenBank database, and a 16S rDNA-based phylogenetic tree was constructed using approximately the first 1300–1500 nt at the 5’ end of each available sequence.
As expected, a 600 bp PCR product was amplified from all 44 Staphylococcus and Macroccocus reference species tested using the hsp60 universal primers (data not shown). All PCR amplicons were sequenced and the resulting data were deposited in the GenBank database (accession nos in Fig. 1). DNA sequence alignments demonstrated the presence of highly conserved regions interspersed with variable segments that appeared to be randomly distributed within the 600 bp amplicons (data not shown). Pairwise sequence identity scores based on partial hsp60 gene sequences among the four Macroccocus species tested ranged from 82 to 87% (mean 83%), while those among the 40 Staphylococcus species or subspecies examined ranged from 74 to 98% (mean 82%). In contrast, the partial hsp60 sequence identity between unrelated Gram-positive (Streptococcus pyogenes and Bacillus subtilis) and Gram-negative (Escherichia coli, Campylobacter jejuni, Vibrio cholerae and Aeromonas hydrophila) bacteria ranged from 53 to 64% (data not shown). Sequence identity scores for pairwise comparisons among the macrococi were consistently higher than those for each Macroccocus species paired with a Staphylococcus species (range 71–79%; mean 74%), thus supporting the recommendation by Kloos et al. (1998) to designate Macroccocus as a novel genus that was separate from Staphylococcus. Among the 34 distinct Staphylococcus species included in the study, the most similar pairs were between members of the ‘sciuri group’ (Staphylococcus sciuri, Staphylococcus lentus, Staphylococcus pulvereri and Staphylococcus vitulinus), with sequence identity scores of 88–98% (mean 91%) and the most similar pair being S. pulvereri and S. vitulinus (98%). Among staphylococcal strains with different subspecies designations within the same species, pairwise sequence identity scores were all above 90% (ranging from 91%, between Staphylococcus capitis subsp. capitis and S. capitis subsp. urealyticus, 93% between Staphylococcus cohnii subsp. cohnii and S. cohnii subsp. urealyticus, to 98% between each pair of S. aureus subsp. aureus and S. aureus subsp. anaerobius, Staphylococcus hominis subsp. hominis and S. hominis subsp. novobiosepticus, and S. schleiferi subsp. schleiferi and S. schleiferi subsp. coagulans). Compared to the hsp60 gene sequence data, the corresponding 16S rRNA gene sequence identity scores for any given pair among the 44 Staphylococcus and Macroccocus strains were consistently higher (mean ± SEM and range among 903 pairwise comparisons were 96-34 ± 0.05% and 93–100%, respectively, for 16S rRNA gene sequences versus 80-45 ± 0.14 % and 74–98 %, respectively, for hsp60 gene sequences; P<0.0001, paired t test, two-tailed). Thus, hsp60 sequences are more discriminatory than 16S rRNA gene sequences for differentiating strains belonging to the genera Microccocus and Staphylococcus.

The unrooted phylogenetic tree constructed from the partial hsp60 gene sequences from the 40 Staphylococcus species and subspecies (representing all but two of the entire set of 36 validly described Staphylococcus species) and the four Macroccocus species examined here is shown in Fig. 1. Bootstrap values of greater than 50% are shown at the nodal branches. The corresponding phylogenetic tree derived from the 16S rRNA gene sequences is shown in Fig. 2. The partial hsp60 gene sequences clearly separated all of the Macroccocus species (cluster 1) from the Staphylococcus species, thus supporting the recommendation that they represent a distinct genus (Kloos et al., 1998). Among the macrococi, M. equipercreus appeared to branch out very
early from other members of its genus, while *M. bovicus* and *M. carouselicus* were closely related and could be easily discriminated from *M. caseolyticus*. The phylogenetic relationships seen among these strains, all exclusively associated with animals, are consistent with their known phenotypic and genetic characteristics (de la Fuente et al., 1992; De Buyser et al., 1992). Such divergence is also consistent with data from scanning electron microscopy studies, which suggest that the cell surfaces of *M. bovicus* and *M. carouselicus* are irregular, whereas the cell surface of *M. caseolyticus* is smooth and that of *M. equiperdicus* shows small piliform projections (Kloos et al., 1998).

Immediately adjacent to the macrococcis, a tightly clustered group of *Staphylococcus* species belonging to the 'sciuri group' (*S. sciuri, S. lentus, S. pulvereri and S. vitulinus*) were seen (cluster 2, Fig. 1), all of which are also cytochrome c oxidase-positive. This suggests that the species that produce cytochrome c may be evolutionarily more closely related to macrococcis than to staphylococcis. To the best of our knowledge, such an association has not been reported previously. Also of note is the finding that *S. vitulinus* and *S. pulvereri* were paired together in both of the phylogenetic trees, each with a nodal bootstrap value of 100 % (Figs 1 and 2). Their pairwise hsp60 and 16S rRNA gene sequence identity scores were 98 and 99 %, respectively. With regard to these two species, Petrás (1998) concluded that *S. pulvereri* was in fact *S. vitulinus*, or at least a subspecies of *S. vitulinus*, after extensive biochemical tests on three different strains of *S. pulvereri* in comparison to the type strains of *S. pulvereri* and *S. vitulinus*. Our phylogenetic studies based on both hsp60 and 16S rRNA gene sequences support this view.

The remaining *Staphylococcus* species formed a broadly based cluster that contained five distinct subdivisions, which corresponded to the 'aureus group' (group a), the 'epidermidis group' (group b), the 'haemolyticus group' (group c), the 'saprophyticus group' (group d) and the 'intermedius group' (group e) (Fig. 1). These results agreed remarkably well with the current taxonomy of this diverse family, which is based on DNA–DNA hybridization data and classical phenotypic and biochemical testing (Kloos, 1997). Of interest is the observation that the 'non-*S. aureus* coagulate-positive staphylococcis, including *S. intermedius*, *S. delphini, S. lutrae, S. schleiferi* subsp. coagulans and some strains of *S. hyicus*, appeared to be closely related and were grouped together (group e, Fig. 1) (Foster et al., 1997; Kloos & Bannerman, 1995).

There was remarkable concordance between the phylogenetic trees constructed from the partial hsp60 and 16S rRNA gene sequences. Thus, both trees revealed the same three major clusters, with high bootstrap values of 81–100 % (Figs 1 and 2). In both trees, the *Macrococcus* species occurred in a tight cluster that was clearly separated from the *Staphylococcus* species. Members of the 'sciuri group' formed the second major cluster in both trees. However, there were some differences as well as similarities in the hierarchy and clustering patterns among members within the third major cluster. For example, there was general agreement that *S. schleiferi, S. hyicus, Staphylococcus chromogenes, S. delphini, S. intermedius, S. lutrae* and *Staphylococcus felis* were related to each other in both trees (nodal bootstrap values of 77 and 85 %, respectively). Similarly, *S. epidermidis* appeared to be related to *Staphylococcus saccharolyticus* and *S. capitis* in both trees, *S. hominis* was related to *S. haemolyticus*, and *Staphylococcus piscifermentans, Staphylococcus carnosus* and *Staphylococcus simulans* were related to each other. However, whereas *S. aureus* and its subspecies were grouped with *S. epidermidis* in the 16S rDNA-based tree (both within group c, nodal bootstrap value of 86 %; Fig. 2), this relationship was not apparent in the hsp60 phylogenetic tree (located within groups a and b, respectively; Fig. 1). Other minor differences were also observed: *Staphylococcus warneri* was paired with *Staphylococcus pasteurii* in the 16S rDNA tree (within group c, Fig. 2) but it was paired with *S. lugdunensis* in the hsp60 tree (within group a, Fig. 1).

hsp60 genes, which encode highly conserved housekeeping proteins that assist in proper protein folding (also known as molecular chaperonins), are ubiquitous in both prokaryotes and eukaryotes (Craig et al., 1993; Ellis, 1999; Goh et al., 1996). In a previous study, we demonstrated with a limited number of *Staphylococcus* strains that the phylogenetic tree constructed from hsp60 gene sequences agreed better with DNA–DNA hybridization data than with 16S rRNA gene sequence data (Kwok et al., 1999). In the current study, these observations were further substantiated by expanding the number of reference strains tested to include all but two of the recent additions to the genus *Staphylococcus* (*n*=40) and four *Macrococcus* species. Furthermore, we demonstrated unambiguously that hsp60 gene sequences are more discriminatory than 16S rDNA sequences for species differentiation within these two genera. In addition to their usefulness for discriminating *Staphylococcus* and *Macrococcus* species, hsp60 sequences have also been shown to be an efficient molecular tool for the accurate identification of members of the genera *Streptococcus* (Goh et al., 1998) and *Enterococcus* (Goh et al., 2000), as well as *Enterobacteriaceae* (Wong & Chow, 2002), *Vibrionaceae* (Kwok et al., 2002) and *Mycobacterium* species (Ringuet et al., 1999; Steingrube et al., 1995). The International Committee on Systematic Bacteriology has recently proposed minimal standards for the taxonomic description of novel *Staphylococcus* species, which include both phenotypic and genotypic criteria (Freney et al., 1999). It has also been suggested that DNA sequencing of highly conserved housekeeping or other genes may supplant DNA–DNA reassociation or 16S rRNA gene sequence data for taxonomic analyses of ecologically distinct populations (Palys et al., 1997; Wong & Chow, 2002). Based on data obtained from the present study, we suggest that the hsp60...
gene may be a useful alternative to DNA–DNA hybridization or 16S rRNA sequencing for taxonomic classification and phylogenetic studies of members of the genera *Staphylococcus* and *Macrococcus*.

**Acknowledgements**

This work was supported in part by a grant from the Canadian Bacterial Diseases Network National Centers of Excellence Programme. We are grateful to Dr W. Kloos of North Carolina State University, Raleigh, NC, USA, for providing some of the strains studied.

**References**


http://ij.sgmjournals.org
