Molecular typing of nasal carriage isolates of \emph{Staphylococcus aureus} from an Irish university student population based on toxin gene PCR, \textit{agr} locus types and multiple locus, variable number tandem repeat analysis

Mark M. Collery,\textsuperscript{1} Davida S. Smyth,\textsuperscript{1}† Jane M. Twohig,\textsuperscript{1} Anna C. Shore,\textsuperscript{2} David C. Coleman\textsuperscript{2} and Cyril J. Smyth\textsuperscript{1}

\textsuperscript{1}Department of Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, University of Dublin, Dublin 2, Republic of Ireland
\textsuperscript{2}Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental School and Hospital, Trinity College Dublin, University of Dublin, Dublin 2, Republic of Ireland

Forty-eight \emph{Staphylococcus aureus} isolates collected from a young, healthy, Irish university student population from 1995 to 2004 were screened for 16 enterotoxin (SE) and enterotoxin-like (SEl) genes (sea–see, seg–sei, selj–selo, selq, selu), and for the toxic shock syndrome toxin-1 gene, \textit{tst}. All of the isolates harboured at least one SE or SEl gene and 66.7\% possessed a classical SE gene (sea, seb, sec), the commonest being the seb gene. Most of the isolates (85.4\%) had a complete \textit{egc} locus (selo, selm, sei, seln, seg). The intergenic sei–seln region of the \textit{egc} locus was typed by PCR-RFLP in 34 isolates, 15 possessing pseudogenes \textit{vent}1 and \textit{vent}2 and 19 having the selu gene. The seh and sell genes, the selk–selq gene combination, and the \textit{tst} gene were each found in <15\% of isolates. The \textit{agr} genotype distribution was \textit{agr} type III, 37.5\%; \textit{agr} type I, 35.4\%; \textit{agr} type II, 25\%; and \textit{agr} type IV, 2.1\%. There was no association between SE–SEl genotype and \textit{agr} type. All \textit{tst} gene-positive isolates were of \textit{agr} type III and harboured a classical SE gene. Multiple locus, variable number tandem repeat analysis (MLVA) produced 47 different patterns. While the \textit{sdh} locus was present in all isolates, half of them lacked one or two of the \textit{sdh} gene amplifiers. Twenty isolates harboured the \textit{bbp} gene, its presence being associated with \textit{agr} type III, but not with the SE–SEl gene profile. The \textit{agr} types of isolates were associated with MLVA subclusters. Selective MLST analysis revealed seven novel sequence types and a new \textit{aroE} allele. Five clonal clusters (CCs), including CCs comprising major pandemic clones CC30, CC5 and CC22 and minor lineages CC6 and CC9, and three singletons were identified.

\textbf{INTRODUCTION}

\emph{Staphylococcus aureus} is an opportunistic pathogen that can persist in the anterior nares of healthy human carriers (Kluytmans \textit{et al}., 1997; Peacock \textit{et al}., 2001; Nouwen \textit{et al}., 2004; Wertheim \textit{et al}., 2005). The diseases caused by \textit{S. aureus} range in severity from superficial skin infections and food poisoning to life-threatening conditions such as toxic shock syndrome and septicaemia (Ferry \textit{et al}., 2005; Murray, 2005; Todd, 2005). Many molecular methods have been employed to type, differentiate, and group isolates of \textit{S. aureus} (Shopsin & Kreiswirth, 2001; Trindade \textit{et al}., 2003; Robinson & Enright, 2004; Vivoni & Moreira, 2005). While currently pulsed-field gel electrophoresis (PFGE), usually with \textit{Smal}-generated genomic DNA fragments, is considered the ‘gold’ standard method, it is time-consuming, requires specific and expensive equipment, and is relatively labour intensive (Arshad \textit{et al}., 1993; Wagner & Lai, 1994). In contrast, PCR-based techniques are simple to perform. Multiple locus, variable number tandem repeat analysis (MLVA) has been used to type \textit{S. aureus} utilizing multiplex PCRs that target intragenic regions with variable numbers of tandem repeats.

\textsuperscript{1}Present address: New York Medical College, Department of Microbiology and Immunology, Valhalla, NY 10595, USA.

\textbf{Abbreviations:} CC, clonal cluster; MLVA, multiple locus, variable number tandem repeat analysis; PFGE, pulsed-field gel electrophoresis; SE, staphylococcal enterotoxin; SEI, staphylococcal enterotoxin-like; slv, single locus variant; ST, sequence type; TSST-1, toxic shock syndrome toxin-1.
Repeats, namely clfA, clfB, sdrC, sdrD, sdrE, sspA, spa, mecA, fnbA, fnbB and cna (Sabat et al., 2003, 2006; Francois et al., 2005; Lindstedt, 2005; Malachowa et al., 2005). This method has been shown to have a high level of inter-laboratory reproducibility (Lindstedt, 2005) and a discriminatory power and clustering capacity equivalent to PFGE (Malachowa et al., 2005).

The most widely characterized of S. aureus regulatory loci is the accessory gene regulator (agr) (Novick, 2003; Robinson et al., 2005; Wright et al., 2005). Within the agr locus there is a variable region comprising the 3′-end of the agrB gene, the agrD gene and the 5′-end of the agrC gene. S. aureus has been divided into four main agr genotypes based on this sequence variation (Gilot et al., 2002; Gilot & van Leeuwen, 2004; Wright et al., 2005; Robinson et al., 2005). Associations between the agr genotype of isolates and specific staphylococcal diseases have been shown for toxic shock syndrome and staphylococcal scalded skin syndrome (Jarraud et al., 2005), but for abscesses and soft tissue infections in intravenous drug users (Monk et al., 2004).

A survey by the National Food Centre in Dublin into microbial contamination of Irish household refrigerators revealed that 298/726 (41%) yielded S. aureus (Kennedy et al., 2005). Analysis of 157 isolates from this survey demonstrated that 101 (64%) harboured at least two out of the sea, seb, seg and sei genes encoding enterotoxins with proven emetic activity (Smyth et al., 2006). However, the animal or human origins of these enterotoxigenic S. aureus remain uncertain. While a study of S. aureus from cows, sheep, goats, rabbits and chickens had revealed isolates with enterotoxin gene profiles identical to those from Irish domestic refrigerators (Smyth et al., 2005), no data were available on the frequencies of such genes in human carriage isolates in Ireland or on the population structure of Irish nasal isolates based on molecular typing methods.

In order to address this deficiency and thus enable epidemiological analysis, the present study was undertaken using PCR-based techniques to characterize the enterotoxin (SE) and enterotoxin-like (SEI) gene profiles of S. aureus from a young, healthy Irish university student population, to determine the prevalence of the tst gene, encoding toxic shock syndrome toxin-1 (TSST-1), to group these isolates by agr typing, to determine the MLVA types and population structure, and to ascertain the frequency of the bbp gene, an allele variant of the sdrE gene amplified by the universal primers for sdr genes analysed by MLVA.

**METHODS**

**Strain collection.** As part of student laboratory exercises from 1995 to 2004, nasal swabs were taken from science students, dental students and dental nursing/dental hygiene students. These swabs were plated on Mannitol Salt agar (Oxoid) and incubated for 36–48 h. Single, well-isolated colonies with the typical appearance of S. aureus were subcultured onto Tryptic Soy agar (Oxoid). Identification of these putative isolates as S. aureus was confirmed by sugar fermentation tests (arabinose, glucose, maltose, mannitol, sucrose, trehalose, xylose), the Voges–Proskauer test, DNAse production, by a tube coagulase test using standard methods, and by the Staphaurex rapid latex test kit, which detects protein A and clumping factor (fibrinogen-binding protein) (Murex Diagnostics), according to the manufacturer’s instructions. Primers specific for the femA gene encoding the S. aureus FemA peptidyltransferase (Schneider et al., 2004) and generating a 132 bp amplimer were used for additional confirmation, as described by Mehrrota et al. (2000). Only isolates from students of the Republic of Ireland and Northern Ireland were retained.

Approximately 26–47% of the total number of nasal S. aureus isolates obtained in the years 1995, 1997, 2000 and 2002 were selected at random from those available. Only isolates from seven dental students were available for 2003, five of which were used. All of the isolates recovered during 2004 were included. The distribution of isolates by year of isolation was 1995, 7 isolates; 1997, 7 isolates; 2000, 8 isolates; 2002, 5 isolates; 2003, 5 isolates; and 2004, 16 isolates. The overall nasal carriage rate of the 492 students surveyed was 20.7%, with a range of 17–31% of the Irish students sampled in any year. These rates are very similar to those reported for medical students in three different years at Trinity College Dublin (Kingdom et al., 1983). The numbers of isolates screened in this survey reflected approximately 10% of the students in each of the three student groups surveyed for S. aureus nasal carriage (15 isolates from 146 science students, 22 isolates from 245 dental science students, and 11 isolates from 101 dental nursing/dental hygiene students).

**DNA extraction.** Preparation of genomic DNA from S. aureus was performed as previously described (Fitzgerald et al., 1997). The procedure involved lysozaphin lysis, proteinase K treatment, addition of EDTA, Sarkosyl and cetyltrimethylammonium bromide (CTAB), extraction with chloroform/isoamyl alcohol and phenol/chloroform/isoamyl alcohol, and precipitation with ethanol.

**PCR.** All primers were custom synthesized by MWG Biotech.

(i) **MLVA.** Primers used for MLVA were those described by Sabat et al. (2003). Primer pairs were tested individually and concentrations in the multiplex PCR primer mix optimized. Each multiplex PCR reaction contained 3–12 pmol of each primer. PCR was performed in a final volume of 25 μl. The PCR mix contained 16.2 μl H2O, 1 μl dNTP mix (0.2 μM each dNTP), 0.3 μl Taq DNA polymerase (5 U μl−1; Promega), 2.5 μl 10× buffer (10 mM Tris/HCl, pH 9.0), 3 μl 25 mM MgCl2, 1 μl genomic DNA sample, and 1 μl primer mix. The thermal cycling conditions were those described by Sabat et al. (2003).

(ii) **agr typing.** An agr group-specific multiplex PCR was performed using the primers and thermal cycling conditions described by Gilot et al. (2002). The reaction mix contained 7–14 pmol of each primer. The following control strains for agr typing were kindly provided by university colleagues T. J. Foster and A. Rossney: strains Newman and COL (agr type I); strains MRSA252, MW2 and NCTC 8325-4 (agr type III) and strain RN4850 (agr type IV). Strain R122 (agr type II) was from Fitzgerald et al. (2001).

(iii) **bpb gene.** The primers employed to detect the bpb gene encoding the bone sialoprotein-binding protein were those described by Tristan et al. (2003) using the thermal cycling conditions that Sabat et al. (2003) described for MLVA fingerprinting. The reaction mix contained 3–4 pmol of each primer. Strain MRSA252 was used as a positive control (Holden et al., 2004).

http://jmm.sgmjournals.org
(iv) SE and SEI genes. The standard nomenclature for SE (enterotoxin) and SEI (enterotoxin-like) genes recommended by Lima et al. (2004) is used throughout. All isolates were screened for the sea–sec, seg–sei, selj–selm, selq and tst genes using the protocols and primers described by Smyth et al. (2005). These assays were done in a series of multiplex PCRs: mix 1, sea, sec, sed, see and 16S rRNA genes; mix 2, seg, seh, sei, selj, selm and 16S rRNA genes; mix 3, sell, selm, seh, selq and 16S rRNA genes. The sea and selk genes were screened for using single PCRs. When egc locus-associated genes (seg, selm, seln, selo) appeared to be missing from any isolate using multiplex PCR mix 2 or mix 3, individual PCRs were employed to recheck for these genes. The sei–seln intergenic region has been shown to harbour either two pseudogenes or the selu gene (Letertre et al., 2003). Accordingly, isolates that were positive for the egc locus were screened by PCR-RFLP for the presence of pseudogenes \( \text{seg}^{-1} \) or the selu gene to provide data on their relative occurrence in egc\(^+\) isolates (Letertre et al., 2003; Collery & Smyth, 2007).

The 16S rRNA gene was amplified as a positive control for \( S. \) aureus DNA in all multiplex PCRs (Monday & Bohach, 1999; Smyth et al., 2005). DNA from SE and SEI gene-positive control strains was included with these multiplex PCR tests: RF122 (sec-.bovine, sell, tst) (Fitzgerald et al., 2001); NY19 (seb), FR913 (sea, sec3, see, tst), FR5699 (seh), FR472 (sed, sec, sei, selj), kindly provided by G. A. Bohach, University of Idaho, Moscow, USA, and A900322 (selo, selm, sei, \( \text{seg}^{-1} \), \( \text{seg}^{+} \), seh, sell, selj, sell, selk) (Collery & Smyth, 2007). Strains A900322 and RF122 were used as controls for pseudogenes \( \text{seg}^{-1} \) and the seh gene, respectively, in PCR-RFLP (Collery & Smyth, 2007).

(v) MLST. MLST was performed selectively on isolates from MLVA clusters and subclusters by PCR amplification of internal fragments of seven housekeeping genes by using a previously described procedure and primers (Enright et al., 2000). Sequencing of both DNA strands was performed commercially by Lark Technologies, using an automated Applied Biosystems 373A DNA sequencer, dye-labelled terminators and the primers used for PCR amplification. Analysis of sequences was performed using the BioNumerics software version 4.6 (Applied Maths). The alleles at each of the seven housekeeping loci were identified by comparing the sequences obtained from the test isolates with sequences held in the MLST database (http://saureus.mlst.net) (Table 1). This database was also used to identify the allelic profile and hence the sequence type (ST) of each isolate. STs were assigned to clonal complexes (CCs) using the clustering algorithm eBURST (based upon related sequence types) (Feil et al., 2004; http://eburst.mlst.net). Using eBURST, isolates with specific STs are assigned to a particular CC if they are related to at least one other ST in that CC at six out of the seven MLST loci used. Isolates that do not share alleles at six of the seven MLST loci with any other ST in the MLST database are deemed singletons.

Computer-assisted analyses of MLVA band profiles. All PCR products were analysed on 2% (w/v) agarose gels, stained with ethidium bromide and visualized under UV light. For the purposes of generating dendograms, PCR products from MLVA of \( S. \) aureus reference strain NCTC 8325-4, a derivative of a human isolate NCTC 8325 that is cured of prophages (Novick, 1967), were run in the first, middle and last lanes of gels as a universal gel standard. Computer-assisted analyses of MLVA band profiles were performed using the DENDRON software package version 2.4 (Soltech) as described previously (Soll, 2000; Gee et al., 2002). Similarity coefficients (\( S_{\text{AB}} \)) based on band position alone were calculated in a pairwise combination of isolate patterns according to the formula \( S_{\text{AB}} = 2E / (2E + a + b) \), where \( E \) is the number of bands shared by isolates A and B, \( a \) is the number of bands unique to isolate A, and \( b \) is the number of bands unique to isolate B. An \( S_{\text{AB}} \) of 1.00 represents identical patterns, and \( S_{\text{AB}} \) ranging from 0.01 to 0.99 represent patterns with increasing numbers of bands at the same positions. Dendograms were generated based on \( S_{\text{AB}} \) values using an unweighted pair-group method (Soll, 2000).

Statistics. Statistical significance was calculated using online calculators for the Fisher’s exact test (http://www.matforsk.no/ola/fisher.htm) or \( \chi^2 \) test (http://www.physics.csbsju.edu/stats/contingency_NROW_NCOLUMN_form.html).

RESULTS

SE, SEI and TSST-1 genes

All of the 48 nasal isolates of \( S. \) aureus screened for 16 SE and SEI genes possessed at least one of these; 66.7% had a classical SE gene (sea, seb, sec), of which 28 had a single gene and the remaining four had two of these genes in combination (3 \( \text{sea}^{+} \) seb\(^+\), 1 seb\(^+\) sec\(^+\) ) (Table 1). The commonest SE gene was seb, which was present in 75% of those possessing a classical SE gene. Of the isolates, 85.4% possessed the complete egc locus (selo, selm, sei, seln, seg); two other isolates lacked the sei gene and the selm and selo genes, respectively, as determined by both multiplex PCR and the use of primers for single genes (Table 1). The seh, Table 1. Prevalence of enterotoxin genes, enterotoxin-like genes and the tst gene among 48 human nasal isolates of \( S. \) aureus

<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of isolates</th>
<th>Percentage of isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>sea*</td>
<td>8</td>
<td>16.7</td>
</tr>
<tr>
<td>seb*</td>
<td>24</td>
<td>50.0</td>
</tr>
<tr>
<td>sec†</td>
<td>4</td>
<td>8.3</td>
</tr>
<tr>
<td>sed</td>
<td>43</td>
<td>89.6</td>
</tr>
<tr>
<td>see</td>
<td>4</td>
<td>8.3</td>
</tr>
<tr>
<td>seg</td>
<td>42</td>
<td>87.5</td>
</tr>
<tr>
<td>seh</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>selj</td>
<td>3</td>
<td>6.3</td>
</tr>
<tr>
<td>sell</td>
<td>7</td>
<td>14.6</td>
</tr>
<tr>
<td>selm</td>
<td>43</td>
<td>89.6</td>
</tr>
<tr>
<td>seln</td>
<td>42</td>
<td>87.5</td>
</tr>
<tr>
<td>selo</td>
<td>42</td>
<td>87.5</td>
</tr>
<tr>
<td>selq</td>
<td>3</td>
<td>6.3</td>
</tr>
<tr>
<td>tst</td>
<td>6</td>
<td>12.5</td>
</tr>
<tr>
<td>egc‡</td>
<td>41</td>
<td>85.4</td>
</tr>
<tr>
<td>Incomplete egc§</td>
<td>2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

*Three strains possessed the sea and seb genes in combination. The sea gene is carried by serotype F lysogenic converting bacteriophage (Carroll et al., 1995).
†One isolate possessed the seb and sec genes in combination.
‡Complete egc locus comprising the selo, selm, sei, seln and seg genes.
§egc locus lacking the sei gene or the selo and selo genes by multiplex PCR and the use of primers for individual genes.
selk, sell, selq and tst genes were each found in <15% of isolates (Table 1).

The sei–seln intergenic region of the egc locus generally harbours two pseudogenes or the selu gene (Letertre et al., 2003). Of the 43 isolates with a complete or incomplete egc locus, 15 (34.9%) possessed pseudogenes \( \psi_{ent1} \) and \( \psi_{ent2} \) and 19 (44.2%) the selu gene as shown by PCR-RFLP (Collery & Smyth, 2007). None of the remaining 9 egc\(^+\) isolates, including the two with incomplete egc loci, gave PCR products with the primers of Letertre et al. (2003) despite several repeated attempts. Four of these ‘non-typable’ isolates lacked classical SE genes (1 egc only, 1 egc seb, 2 with incomplete egc loci). The other five ‘non-typable’ isolates all possessed the seb gene in combination with the egc locus (2 seb egc; 1 seb egc selk selq; 1 seb sea egc; 1 seb sec egc sell tst). Thomas et al. (2006) described isolates with egc loci harbouring insertion sequences and transposable genes, as well as an atypical egc locus with two new SEI genes, designated seln2 and selv, generated by a deleted deletion in pseudogenes \( \psi_{ent1} \) and \( \psi_{ent2} \) and by recombination between selhn and sei, respectively. It is not known whether the primers described by Letertre et al. (2003) would type the sei–seln intergenic regions of such variant isolates.

The rates of occurrence of classical SE genes varied among the isolates examined from each of the six sampling years: 1995, 1 seb\(^+\), 14.3%; 1997, 5 seb\(^+\), 2 sea\(^+\) seb\(^+\), 100%; 2000, 7 seb\(^+\), 87.5%; 2002, 2 sea\(^+\), 1 seb\(^+\), 60%; 2003, 2 seb\(^+\), 3 sec\(^+\), 100%; and 2004, 3 sea\(^+\), 1 sea\(^+\) seb\(^+\), 4 seb\(^+\), 1 seb\(^+\) sec\(^+\), 56.3%. The differences in the numbers of isolates possessing classical SE genes between sampling years are statistically significant for 1995 versus 1997 and 1995 versus 2000 (Fisher’s exact test, 2-tailed, \( P=0.005 \) and \( P=0.01 \), respectively). Humphreys et al. (1989) found that only 11.1% of 27 nasal isolates obtained from healthy medical students prior to arrival in hospitals for their clinical training and from biochemistry laboratory technicians in Dublin in the mid 1980s produced a classical SE (all SED-positive). Moreover, they found a statistically significant lower rate of SEB production among 17 community-acquired versus 35 hospital-acquired septicaemia isolates (\( P<0.01 \)). While the prior study tested for SEs and the present study for SE genes, the enterotoxin genotype data imply a real change in the rate of occurrence of seb\(^+\) isolates in the community from the late 1990s compared to the mid 1980s through mid 1990s, coupled with a more frequent occurrence of sea\(^+\) and sec\(^+\) isolates. Carriage of the sed gene-bearing plasmid may have imposed a fitness cost that was ameliorated by its loss.

Sixteen of the 22 dental students (72.7%), 5 of the 11 dental nursing/hygiene students (45.5%) and 9 of the 15 science students (60.0%) yielded isolates bearing one or two classical SE genes. The differences in the carriage rates of isolates bearing classical SE genes between these student groups are not statistically significant. The somewhat higher occurrence of isolates bearing SE genes in the science and dental student groups may reflect the gender balance of these groups (see Methods) since male gender has been shown to be a candidate factor for persistent carriage in a student community (Bischoff et al., 2004). The numbers of isolates bearing sea: seb: sec enterotoxin genes among nasal isolates from the dental, dental nurse/hygienist and science student groups were 3:13:3, 1:4:0 and 4:7:1, respectively. The differences in the frequencies of these genes between the three student groups are not statistically significant.

### SE, SEI and TSST-1 gene combinations

The two commonest gene combinations were the seb gene with the egc locus and the egc locus alone (complete and incomplete), found in 13 isolates (27.1%) and in 12 isolates (25%), respectively (Table 2). Twenty-seven of the 32 isolates (84.4%) possessing a classical SE gene also had a complete egc locus. Of the five isolates possessing a classical SE gene but lacking the egc locus, two possessed the sea and seb genes together (one also with the selk and selq genes), and three possessed the seb gene only. The egc locus was present in 6 of the 8 isolates with the sea gene (75%) – 5 of these carried the selu gene, in 19 of the 24 isolates with the seb gene (79.2%) – 9 of these carried the selu gene, and in all 4 isolates with the sec gene – none of these carried the selu gene. Only 4 of the 14 isolates with a complete egc locus and lacking a classical enterotoxin gene had the selu gene. Neither of the isolates with an incomplete egc locus possessed a classical SE gene, additional SE or SEI genes, the tst gene or the selu gene.

Each of the six isolates containing the tst gene harboured at least one classical SE gene; all had the egc locus, three the selu gene, and three the sell gene. Of the three isolates possessing the selk–selq gene combination, two also had the seb gene and the egc locus (one with the selu gene), and the third isolate had the sea and seb genes. Of the four isolates with the selh gene, all had the egc locus – one with the selu gene, and two with the sea gene. All of the seven isolates with the seln gene had the egc locus – only one with the selu gene, three of which had classical SE genes – one with seb, one with seb and sec and one with sec – and the tst gene.

### agr genotyping

The nasal isolates were agr genotyped (Table 2). The two commonest agr genotypes were types I and III. Only one strain was found to be of agr type IV. Eight of the agr type I isolates, seven of the agr type II isolates, 14 of the agr type III isolates and the single agr IV isolate harboured classical SE genes. These differences in the rates of occurrence of classical SE gene-positive isolates between agr genotypes I, II and III are not significant, nor are the rates of occurrence of individual classical SE genes between these agr types.

With respect to the commonest SE–SEI gene combination, namely seb egc, three isolates were of agr type I, four of agr type II, five of agr type III, and one of agr type IV (Table 2).
Table 2. Prevalence of agr genotypes and SE, SEI and TSST-1 genotypes among 48 human nasal isolates of S. aureus

<table>
<thead>
<tr>
<th>agr genotype</th>
<th>SE, SEI, TSST-1 genes</th>
<th>No. of isolates</th>
<th>Percentage of isolates*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sor+, seh</td>
<td>1</td>
<td>17</td>
<td>35.4%</td>
</tr>
<tr>
<td>sor+, seh, selκ, selq+</td>
<td>1</td>
<td>1</td>
<td>5.9%</td>
</tr>
<tr>
<td>sor+, seh, ege$</td>
<td>1</td>
<td>1</td>
<td>5.9%</td>
</tr>
<tr>
<td>seh, egc</td>
<td>3</td>
<td>17.6%</td>
<td></td>
</tr>
<tr>
<td>seh, selκ, selq, egc</td>
<td>2</td>
<td>2</td>
<td>11.8%</td>
</tr>
<tr>
<td>sec, sell, egc</td>
<td>2</td>
<td>11.8%</td>
<td></td>
</tr>
<tr>
<td>seh, egc</td>
<td>2</td>
<td>11.8%</td>
<td></td>
</tr>
<tr>
<td>sell, egc</td>
<td>1</td>
<td>9.9%</td>
<td></td>
</tr>
<tr>
<td>egc</td>
<td>4</td>
<td>23.5%</td>
<td></td>
</tr>
<tr>
<td>Type II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>seh</td>
<td>3</td>
<td>25.0%</td>
<td></td>
</tr>
<tr>
<td>seh, egc</td>
<td>4</td>
<td>33.3%</td>
<td></td>
</tr>
<tr>
<td>egc</td>
<td>3</td>
<td>25.0%</td>
<td></td>
</tr>
<tr>
<td>egc, sell</td>
<td>2</td>
<td>16.7%</td>
<td></td>
</tr>
<tr>
<td>Type III</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sor+, ege</td>
<td>1</td>
<td>37.5%</td>
<td></td>
</tr>
<tr>
<td>sor+, seh, ege</td>
<td>2</td>
<td>2</td>
<td>11.1%</td>
</tr>
<tr>
<td>sor+, ege, tst</td>
<td>2</td>
<td>2</td>
<td>11.1%</td>
</tr>
<tr>
<td>seh, ege</td>
<td>5</td>
<td>27.8%</td>
<td></td>
</tr>
<tr>
<td>seh, ege, tst</td>
<td>1</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>seh, sell, ege, tst</td>
<td>1</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>sell, ege, tst</td>
<td>1</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>ege</td>
<td>3</td>
<td>16.7%</td>
<td></td>
</tr>
<tr>
<td>Type IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>seh</td>
<td>1</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

*Gene distribution within agr genotype calculated as percentage of isolates of that agr type.
†The sec gene is carried by serotype F lysogenic converting bacteriophage (Carroll et al., 1995).
§The sek and seq genes occur along with the seh gene on a pathogenicity island (Orwin et al., 2002).
$Complete egc locus comprising the selo, selm, sei, seln and seg genes.
||Incomplete egc locus lacking either the sei gene or the selm and selo genes.

Of those with the complete egc locus alone, four isolates were of agr type I, and three each of agr types II and III. Of the seven harbouring the egc locus and the sell gene with or without a SE gene, three were of agr type I and four of agr type II. All six isolates with the tst gene were of agr type III. The difference in the rates of occurrence of the tst gene between agr types I, II and III is significant ($X^2$ test, $P=0.017$). The three isolates that contained the sek–selq gene combination were of agr type I. Both the seh and sell genes were only associated with agr types I and III (Table 2).

The isolates from dental students were distributed as follows: agr type I, 10, 45.5 %; agr type II, 5, 22.7 %; and agr type III, 7, 31.8 %. For the dental nurse/hygienist group the distribution was agr type I, 4, 36.4 %; agr type II, 2, 18.2 %; agr type III, 4, 36.4 %; and agr type IV, 1, 9.1 %. In the case of the science students the distribution was agr type I, 3, 20.0 %; agr type II, 5, 33.3 %; and agr type III, 7, 46.7 %.

The differences in the rates of occurrence of agr genotypes I, II and III between these student groups are not significant. Nor are the differences in the rates of occurrence of the agr genotypes by year of isolation significant.

bbp gene

When MLVA fingerprinting was performed according to Sabat et al. (2003) on the 48 nasal isolates, the PCR band profiles of 28 isolates (58.3 %) revealed only five or six bands against the seven anticipated on the basis of the primer sets used to amplify seven genes, while a few strains produced eight bands (Fig. 1). These MLVA patterns were highly reproducible. The differences in the numbers of bands were subsequently associated with differences in the presence or absence of sdr gene products. Sabat et al. (2006) reported that the PCR product of the sdc gene was always present, but that one or both of the sdrD and sdrE gene PCR products was absent in 58.9 % of meticillin-sensitive isolates from asymptomatic nasal carriers in the UK, the Netherlands and Poland. In the latter respect the present observations mirror that report.

Universal forward and reverse primers for the sdc, sdrD and sdrE genes are used in the MLVA multiplex PCR of Sabat et al. (2003). The bbp gene is an allelic variant of the sdc gene (Tung et al., 2000). The A-regions of the Bbp and SdrE proteins show 76 % amino acid sequence identity and the B-regions, 95–96 % identity (Tung et al., 2000). Using the BLAST program (http://www.ncbi.nlm.nih.gov/blast/), the universal forward and reverse sdr primers of Sabat et al. (2003) have 100 % matches for bbp gene sequences of S. aureus strains 024 (Tung et al., 2000), MRSA252 (Holden et al., 2004) and NN1 (Otsuka et al., 2006).

As the primers used by Sabat et al. (2003, 2006) and Malachowa et al. (2005) do not distinguish the sdrE gene from the bbp gene, the frequency of the bbp gene among the nasal isolates was determined using specific primers that do not match sdrCDE gene sequences (Tristan et al., 2003). Twenty isolates were positive for the bbp gene. None of the bbp† isolates yielded five bands on MLVA fingerprinting, while 11 gave six bands, six gave seven bands, and three gave eight bands, indicating that the bbp and sdrE genes were both present in the latter three MLVA groups of isolates. The bbp gene was present in one agr type I isolate (5.9 %), in 4 agr type II isolates (33.3 %), in 14 out of 18 agr type III isolates (77.8 %), and in the one agr type IV isolate. The differences in the frequencies of the bbp gene between agr type III and agr types I and II are statistically significant (Fisher’s exact test, 2-tailed, $P=0.00002$ and $P=0.025$, respectively). There were no associations between the occurrence of the bbp gene and the egc locus or the bbp gene and classical SE genes either taken together or separately.
MLVA typing

The 48 *S. aureus* nasal isolates produced 47 MLVA patterns. $S_{AB}$ values were computed for all possible pairwise combinations of isolates and the data used to construct a dendrogram showing the relationships between the nasal isolates (Fig. 2). The mean $S_{AB}$ for 1128 pairwise comparisons was 0.25. Examination of the dendrogram reveals that the isolates were divided into two clusters, designated I and II, comprising 11 and 37 of the isolates with mean $S_{AB}$ of 0.15 ± 0.14 and 0.29 ± 0.21, respectively. Seven of the cluster I isolates were of *agr* type I, three of *agr* type II, and one of *agr* type III. These isolates formed a heterogeneous population with limited relatedness.

MLVA cluster II could be subdivided into four subclusters, designated IIA to IId, at $S_{AB}$ nodes of 0.37, 0.24 and 0.12. There was a marked distribution of isolates according to *agr* type within these subgroups. Subgroup IIA contained 13 isolates of *agr* type III and the single isolate of *agr* type IV. Subgroup IIB comprised the remaining four isolates of *agr* type III, two isolates of *agr* type I and one isolate of *agr* type II. Subgroups IIC and IID each contained eight isolates, with four each of *agr* types I and II. Within MLVA cluster II, 21 of the 37 isolates formed nine pairs and one triplet of the same *agr* types at different $S_{AB}$ nodes. In MLVA cluster I, four of the 11 isolates paired with respect to *agr* type at different $S_{AB}$ nodes.

Subgroups IIA and IIB (mean $S_{AB}$ 0.71) are more related to each other than to subgroups IIC andIID (mean $S_{AB}$ 0.58). The distribution of isolates of *agr* types I and II combined compared to *agr* types III and IV combined within MLVA subgroups IIA and IIB combined versus subgroups IIC and IID combined is highly statistically significant (Fisher’s exact test, 2-tailed, $P=0.00033$). Because of the strong association of the *bbp* gene with *agr* type III, the distribution of *bbp* + isolates within MLVA subgroups IIA and IIB combined versus subgroups IIC and IID combined is statistically significant (Fisher’s exact test, 2-tailed, $P=0.00033$)

Of the isolates of *agr* type III that possessed the *tst* gene, four were in subgroup IIA and two in subgroup IIB. However, there was no association between the MLVA clusters or the four subclusters and the SE–SEI genotypes of isolates or between the MLVA clusters or the four subclusters and their year of isolation or the student group from which they were isolated.

MLST typing

On the basis of the MLVA dendrogram, 13 of the Irish nasal isolates were selected for MLST typing – four from subcluster IIA (all *agr* type III) and two from subcluster IIB (both *agr* type III), two from subcluster IIC and two from subcluster IID (three *agr* type II, one *agr* type I), and three from cluster I (all *agr* type I), reflecting 1 : 3 to 1 : 4 of the isolates in any cluster or subcluster. Selection within clusters or subclusters was based on diversity of enterotoxin gene profile and *agr* type.

The 13 isolates investigated were found to consist of 12 different STs, including seven novel STs (ST932, ST933, ST934, ST935, ST936, ST937 and ST938) (Fig. 2, Table 3). One novel *aroE* allele was found in ST937 and was designated 157. Using eBURST these 12 STs belonged to five CCs (CC30, $n=5$; CC6, $n=2$; CC5, $n=1$; CC22, $n=1$; CC9, $n=1$), and three were singletons.

Of the four isolates in subcluster IIA, two were ST30 (isolates MC15 and MC26) and one was ST484 (isolate MC19), a single locus variant (slv) of both ST30 and ST36 at the *pta* locus. The remaining isolate MC24 was of ST938, a slv of ST30 at the *gmk* locus. Three of the isolates formed pairs with other isolates with $S_{AB}$s of >0.82 by MLVA and...
the fourth isolate, MC15, had an $S_{AB}$ of 0.55 to one of these pairs. All belonged to CC30. Of those examined in subcluster IIb, isolate MC14 was ST936, a singleton with no close relatives in the database, and isolate MC16 was ST933, a slv of ST39 at the $gmk$ locus and a member of CC30. Thus, of the six examined $agr$ type III isolates within subclusters IIa and IIb, five belonged to CC30.

Of the two isolates in subcluster IIc, strain MC29 ($agr$ type I) was of ST6 and the predicted ancestor of a small clonal complex, CC6. Strain MC38 ($agr$ type II) was ST935, a slv of ST5 at the $glpF$ locus, and belonged to CC5. Isolate MC31 ($agr$ type II) in subcluster IId was ST109, a slv of ST9 at $aroE$, and thus belonged to CC9. The other isolate MC38 ($agr$ type II) was ST937 and was predicted by eBURST to be a singleton.

Of the three isolates in subcluster I (all $agr$ type I), strain MC49 was ST217, a slv of ST22 at the $tpi$ locus, and belonged to CC22. Isolate MC43 was ST932, a slv of ST6 at the $pta$ locus, and belongs to CC6. Isolate 40 was ST934 and predicted by eBURST to be a singleton. These isolates were distantly related by MLVA ($S_{AB}$s of 0.04–0.14).

Thus, the more closely related isolates by MLVA (subclusters IIa and IIb) belonged to the same CC, while those in more distantly related MLVA groups (cluster I and subclusters IIc and IId) belonged to diverse CCs. Indeed, Malachowa et al. (2005) found the overall correlation between MLVA clusters and MLST CCs to be lower than between MLVA and PFGE clusters and MLVA and $spa$ typing clusters.

**DISCUSSION**

Few data are available on SE and TSST-1 production by or on the presence of SE, $SEL$ and TSST-1 genes in Irish human isolates of *S. aureus* (Humphreys et al., 1989; Aucken et al., 2006) or in animal isolates (Fitzgerald et al., 2000; Smyth et al., 2005, 2006). Herein 100% of nasal isolates from a young, healthy, Irish university student population collected over a 10-year period harboured a SE and/or $SEL$ gene, 66.7% possessing at least one classical SE gene and 89.6%, a complete or incomplete $egc$ locus. In contrast to the findings of Humphreys et al. (1989) on SED production, no isolate possessed the $sed$ gene.
Five groups in Germany, Japan, Poland, New Zealand and Bulgaria have screened nasal carriage isolates for 12–18 of the SE and SEI genes in S. aureus discovered at the time of the present study (Becker et al., 2003, 2004; Omoe et al., 2005; Bania et al., 2006; Boerema et al., 2006; Nashev et al., 2007). Comparison of these studies with the findings herein brings out several points. The percentage of isolates possessing an SE or SEI gene varied from 75 to 100%. The proportion bearing a classical SE gene range from 30 to 67%. The predominant classical SE gene varied from country to country – Ireland, seC; Germany, sea, sec; Japan, seC; New Zealand, seB; Poland, sec; and Bulgaria, sea. Moreover, of isolates with a classical SE gene, those possessing a complete (and incomplete) egc locus ranged from 38 to 84% between these studies. Differences in the nature of the study population screened for nasal carriage, i.e. healthy individuals versus patients, and differences in PCR screening methodologies, e.g. primers, amplification cycles, single vs multiplex PCR (Monday & Bohach, 1999; Jarraud et al., 2001; Becker et al., 2003, 2004; Chen et al., 2004; Smyth et al., 2005; Bania et al., 2006), may contribute to such variations, and as indeed may the makes of the thermal cyclers used (Saunders et al., 2001; Schoder et al., 2005).

Differences in the frequencies of incomplete egc loci among nasal isolates are particularly notable – 38.1% (Becker et al., 2003, 2004), 12.4% (Omoe et al., 2005), 3% (Boerema et al., 2006), 0% (Bania et al., 2006), 1.1% (Nashev et al., 2007) and 4.2% herein. In an extensive study of human isolates collected during 2001–2003 from supplicative infections, acute toxaemia and asymptomatic nasal carriage, Thomas et al. (2006) found incomplete egc loci to be uncommon (3.8% of egc gene-positive isolates). While the Polish, New Zealand and Bulgarian investigations (Bania et al., 2006; Boerema et al., 2006; Nashev et al., 2007) tested for egc locus genes individually, neither the Japanese study nor the German studies (Becker et al., 2003, 2004; Omoe et al., 2005) stated that they rechecked their egc loci with ‘missing’ genes using individual sets of primers. Detection of egc genes by PCR may be limited by primer specificity for individual alleles (Monday & Bohach, 1999; Jarraud et al., 2001; Becker et al., 2003, 2004; Smyth et al., 2005; Bania et al., 2006) as sequence variants of the seg, sei and seh genes have been described (Abe et al., 2000; Letertre et al., 2003; Blaiotta et al., 2004).

The frequency of the seh gene in nasal isolates is generally low: 4.3% (Jarraud et al., 2002), 16.7% and 4.1% in Japan (Omoe et al., 2002, 2005), 3.8% in Poland (Bania et al., 2006), 6.2% in Germany (Becker et al., 2003), 20% in New Zealand (Boerema et al., 2006), 6.7% in Bulgaria (Nashev et al., 2007) and 8.3% herein. The selq gene is present on the SaP13 pathogenicity island directly 5′ of the selk gene and together with the seh gene (Orwin et al., 2002). In nasal isolates the frequency of the selk selq gene combination is low: 5% in Poland (Bania et al., 2006), 6.2% in Japan (Omoe et al., 2005) and 6.3% herein. The rates of occurrence of the tst gene in nasal isolates are fairly consistent – 12.4% in Japan (Omoe et al., 2005), 12.5% in Poland (Bania et al., 2006), 22.4% in Germany (Becker et al., 2003), 24.3% in the USA (Mehrotra et al., 2000) and 13.7% herein.

In two studies of nasal carriage isolates, the agr type distributions were: agr type I, 55.2% and 52.3%, versus 35.4% herein; agr type II, 25.9% and 29.2%, versus 25.0% herein; agr type III, 17.2% and 10.8%, versus 37.5% herein; and agr type IV, 0% and 7.7%, versus 2.1% herein (Cespedes et al., 2005; Lina et al., 2003). In a study of nasal isolates from children and their guardians, Shopsin et al. (2003) reported that the distribution of the four agr groups was agr type I, 41.6%; agr type II, 24.7%; agr type III, 33.8%; and agr type IV, 0%; this is somewhat similar to that found in the present study. All of the Irish isolates possessing the tst gene were of agr type III and belonged to MLVA subgroups IIA and IIB, four of these isolates possessing the bbp gene.

The diversity in the numbers of bands produced on MLVA fingerprinting had been unexpected, as such observations had not been specifically annotated in the original report of Sabat et al. (2003). Subsequently, Sabat et al. (2006) reported that, while all 142 MLVA-typed nasal isolates possessed the sdrC gene, only 55 (38.7%) of these were both sdrD- and sdrE-positive, 76 (53.5%) had the sdrE gene but not the sdrD gene, 3 (2.1%) had the sdrD but not the sdrE gene, and 8 (5.6%) lacked both the sdrD and sdrE genes, i.e. 94.4% would yield six or seven bands versus 40 out of 48 nasal isolates (83.3%) yielding six to eight bands herein. The percentage of isolates herein possessing the bbp gene (41.7%) is somewhat higher than that reported by Tristan et al. (2003) for nasal isolates (20.7%).

Three isolates yielding eight bands on MLVA typing were bbp-positive. While the sdrE and bbp alleles have been found separately in isolates (Otsuka et al., 2006; Stephens et al., 2006; Taneike et al., 2006), their dual occurrence has only been previously recorded in a single community-acquired MRSA isolate NN13 (Otsuka et al., 2006). However, seven isolates belonging to CC30 of MLST types ST36, ST30 and ST39 have been found to possess the sdrE and bbp genes with bbp arranged tandemly immediately downstream of sdrE (Dr Edward Feil, University of Bath, personal communication). The three isolates in the present study possessing the sdrE and bbp alleles all had the seh gene and the egc locus with the selu gene (Letertre et al., 2003; M. M. Collery, unpublished data), were of agr type III and grouped in MLVA subclusters IIA and IIB, two of them being confirmed to belong to CC30 by MLST.

The association seen between MLVA subgroups and agr genotype has not been previously recorded, although MLVA clusters have been shown to match well with those obtained by PFGE and to be consistent with spa typing (Malachowa et al., 2005). PFGE patterns and agr types have been shown to be linked within certain CCs (Goerke et al., 2005). PFGE has also revealed a pronounced diversity among nasal carriage isolates and that isolates (non-MRSA) from
unrelated individuals yield unique patterns (Hu et al., 1995; Cespedes et al., 2005).

The Irish nasal carriage isolates investigated in the present study belonged to diverse genotypes as determined by MLST analysis. They yielded 12 distinct STs, seven of which had not been described previously. Furthermore, 3 of the 13 isolates were singletons and the remainder belonged to five CCs, the most common being CC30. Interestingly, three of these CCs comprised major pandemic clones, i.e. CC30, CC5 and CC22, whereas the other two, CC6 and CC9, belonged to minor lineages. Grundmann et al. (2002) previously identified S. aureus isolates of CC30, CC22, CC5 and CC9 among nasal carriage isolates in the UK, similar to the present study. They also identified CC30 as the commonest, accounting for 25.6% of the nasal isolates examined. In addition, Peacock et al. (2002) reported that 26–33% of nasal isolates from a cohort of renal dialysis outpatients on any of three sampling occasions belonged to CC30. Based on the selective MLST typing of isolates within subclusters IIA and IIB herein, a similar overall percentage of nasal isolates would be predicted to belong to CC30.

As the outbreak incidence of food-borne disease declines, monitoring sporadic cases of staphylococcal food poisoning will be an important challenge for the 21st century (Mead et al., 1999). The molecular genetic definition of a population of nasal carriage isolates from a healthy Irish population together with data on molecular genotyping of animal isolates (Smyth et al., 2005) and of S. aureus from contaminated Irish domestic refrigerators (Smyth et al., 2006) should add to the knowledge base for epidemiological analysis of staphylococcal food poisoning incidents in Ireland and of the roles, which are possibly underestimated, of the more newly described SEs and possibly SEIs in this toxin-mediated syndrome.

ACKNOWLEDGEMENTS

Mark Collery was in receipt of a studentship from the Sarah Purser Medical Research Fund and of a Trinity College postgraduate research studentship. Davida Smyth was supported by a Teagasc Walsh Fellowship. Anna Shore was supported by Health Research Board Grant TRA/2006/4 and by the Microbiology Research Unit, Dublin Dental School & Hospital. Thanks to Karsten Becker and Ed Feil, who provided helpful additional and unpublished data, respectively, to our University & Hospital. Thanks to Karsten Becker and Ed Feil, who provided helpful additional and unpublished data, respectively, to our University & Hospital. Thanks to Karsten Becker and Ed Feil, who provided helpful additional and unpublished data, respectively, to our University & Hospital.

REFERENCES


