



Københavns Universitet

Review and phylogenetic analysis of *qac* genes that reduce susceptibility to quaternary ammonium compounds in *Staphylococcus* species

Wassenaar, Trudy M.; Ussery, David; Nielsen, Lene Nørby; Ingmer, Hanne

Published in:

European Journal of Microbiology and Immunology (Print)

DOI:

[10.1556/EUJMI-D-14-00038](https://doi.org/10.1556/EUJMI-D-14-00038)

Publication date:

2015

Document version

Publisher's PDF, also known as Version of record

Citation for published version (APA):

Wassenaar, T. M., Ussery, D., Nielsen, L. N., & Ingmer, H. (2015). Review and phylogenetic analysis of *qac* genes that reduce susceptibility to quaternary ammonium compounds in *Staphylococcus* species. *European Journal of Microbiology and Immunology (Print)*, 5(1), 44-61. <https://doi.org/10.1556/EUJMI-D-14-00038>

REVIEW AND PHYLOGENETIC ANALYSIS OF *qac* GENES THAT REDUCE SUSCEPTIBILITY TO QUATERNARY AMMONIUM COMPOUNDS IN *STAPHYLOCOCCUS* SPECIES

Trudy M. Wassenaar^{1,*}, David Ussery^{2,4}, Lene N. Nielsen³ and Hanne Ingmer³

¹ Molecular Microbiology and Genomics Consultants, Zotzenheim, Germany

² Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark

³ Department of Veterinary Disease Biology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

⁴ Present address: Comparative Genomics Group, Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Received: December 17, 2014; Accepted: December 23, 2014

The *qac* genes of *Staphylococcus* species encode multidrug efflux pumps: membrane proteins that export toxic molecules and thus increase tolerance to a variety of compounds such as disinfecting agents, including quaternary ammonium compounds (for which they are named), intercalating dyes and some antibiotics. In *Staphylococcus* species, six different plasmid-encoded Qac efflux pumps have been described, and they belong to two major protein families. QacA and QacB are members of the Major Facilitator Superfamily, while QacC, QacG, QacH, and QacJ all belong to the Small Multidrug Resistance (SMR) family. Not all SMR proteins are called Qac and the reverse is also true, which has caused confusion in the literature and in gene annotations. The discovery of *qac* genes and their presence in various staphylococcal populations is briefly reviewed. A sequence comparison revealed that some of the PCR primers described in the literature for *qac* detection may miss particular *qac* genes due to lack of DNA conservation. Despite their resemblance in substrate specificity, the Qac proteins belonging to the two protein families have little in common. QacA and QacB are highly conserved in *Staphylococcus* species, while *qacA* was also detected in *Enterococcus faecalis*, suggesting that these plasmid-born genes have spread across bacterial genera. Nevertheless, these *qacA* and *qacB* genes are quite dissimilar to their closest homologues in other organisms. In contrast, SMR-type Qac proteins display considerable sequence variation, despite their short length, even within the *Staphylococcus* genus. Phylogenetic analysis of these genes identified similarity to a large number of other SMR members, found in staphylococci as well as in other genera. A number of phylogenetic trees of SMR Qac proteins are presented here, starting with genes present in *S. aureus* and *S. epidermidis*, and extending this to related genes found in other species of this genus, and finally to genes found in other genera.

Keywords: biocide resistance, *qac*, MFS, *smr*, MRSA, *S. aureus*, phylogeny

Introduction

Multidrug efflux pumps are membrane proteins, typically containing multiple (four to fourteen) transmembrane domains that form a channel to transport toxic compounds out of the cell. Their substrates can be cations, quaternary ammonium compounds (*qac*), phosphonium derivatives and intercalating dyes [1–6]. These efflux pumps in part protect bacteria to a variety of toxic substances, including chemicals such as biocides and disinfectants, as well as fluoroquinolones and β -lactams [7, 8]. Repeated sub-lethal exposure can select for presence or active expression of efflux pumps, resulting in a population with reduced

susceptibility. A number of multidrug efflux pumps have been identified in *Staphylococcus aureus* and in other members of the *Staphylococcus* genus.

S. aureus is an opportunistic but potentially serious human pathogen. It colonizes over 30% of individuals and although endogenous strains are often the source of infection, *S. aureus* can also be acquired from external sources. Pathogenic strains are circulating in hospitals that can cause nosocomial infections in high-risk patient populations, but community-acquired strains can also cause aggressive infections, even in healthy individuals [9, 10]. Adding to the problem is the ability of *S. aureus* to acquire resistance to antibiotics. In recent years methicillin-

* Corresponding author: Trudy M. Wassenaar; Molecular Microbiology and Genomics Consultants, Tannestrasse 7, 55576 Zotzenheim, Germany; Phone: +49 6701 8531; Fax: +49 6701 901803; E-mail: trudy@mmgc.eu

resistant *S. aureus* (MRSA) strains have attributed to both hospital- and community-acquired infections [10, 11]. Exchange of MRSA strains between animal and human populations has also been observed, with the concerning example of sequence type ST398, which colonizes pigs in countries with a large pig-producing industry like Denmark and The Netherlands, and was shown to be transmitted from these animals to humans [12].

Another medically important staphylococcal species is *S. epidermidis*. While previously only considered to be a harmless colonizer of the skin, it is now recognized as an important opportunistic pathogen that gives rise to considerable numbers of nosocomial infections and infections associated with medical indwelling devices [13].

In *S. aureus* and *S. epidermidis*, genes encoding efflux pumps that are plasmid-borne can be variably present, and these are clinically relevant since they confer reduced susceptibility to commonly used antiseptics and disinfectants, such as cetrimide, benzalkonium chloride and chlorhexidine [6]. Their broad substrate specificity suggests they evolved to provide generic protection against toxic compounds. In clinically relevant populations their presence may have been selected for by human application of disinfectants [14, 15], although this selection could not always be experimentally demonstrated [16]. A number of these proteins are known by the name of Qac, for *quaternary ammonium compound resistance / transporter*.

Generally, multidrug efflux pumps can be divided into five protein families, depending on their energy requirements and structure [3, 6, 17]. The Qac efflux proteins of *Staphylococcus* species belong to two of these protein families: the Major Facilitator Superfamily (MFS) and the Small Multidrug Resistance (SMR) family [3]. The MFS is the largest known superfamily of secondary transporters, containing 74 protein families (known to date), transporting a variety of substrates [18]. Both the MFS and SMR protein families contain a variety of membrane transporters, all involved in symport, antiport or uniport of a range of substrates. The proton motive force of the cellular transmembrane proton gradient provides energy for such transport [3].

The scope of this review is to briefly summarize the historic literature on *qac* genes and their products from *Staphylococcus* species, and to compare the publically available genes by phylogenetic analysis using either amino acid or DNA sequences. The results of these analyses were used to identify conserved domains in the genes encoding Qac that are found across various taxonomic groups. Phenotypic characteristics are not covered here.

The Qac alphabet: from QacA to QacZ

Qac is the abbreviation for quaternary ammonium compounds, but not all genes conferring reduced susceptibility to these compounds are called *qac*, for historical reasons. Early observations of *S. aureus* resistance as a result of presence of plasmid-encoded genes were obtained from

studies using efflux of ethidium bromide as an assay system [19]. Twenty years later, in 1987, the first gene characterized from *S. aureus* responsible for this ethidium bromide efflux was therefore named *ebr* by Sasatsu and coworkers [20]. In that same year, Lyon and Skurray introduced the names QacA, QacB and QacC, in an influential review article that summarized the then current knowledge on antimicrobial resistance in *S. aureus* [1]. Based on observations published at that time, Lyon and Skurray coined the term QacA for a protein coded on pSK1 family plasmids that rendered resistance to multiple substrates: quaternary ammonium compounds, the intercalating dyes acriflavine and ethidium bromide, diamidines such as propamidine isothionate as well as diamidinodiphenylamine dihydrochloride or chlorhexidine. The protein that produced resistance to quaternary ammoniums and intercalating dyes but not to diamidines or chlorhexidine was called QacB, while the name QacC was reserved for the protein that only produced resistance to quaternary ammoniums and ethidium bromide but not to acriflavine [1]. It was recognized at that time that these phenotypes were all due to efflux proteins and, while a genetic resemblance between QacA and QacB activity had been recognized, the gene coding for QacC was known to be different [1].

The gene encoding **QacA** was sequenced from plasmid pSK1 of *S. aureus* three years later, and its similarity to transporter genes involved in sugar transport systems was recognized [21]. QacA is 514 amino acids long, though meanwhile some *qacA* genes are annotated to start from alternative start codons, producing shorter open reading frames of 504 or 486 amino acids, respectively. The protein contains 14 transmembrane domains. Although an early observation suggested that antiseptic resistance could be detected on the chromosome of *S. aureus* [22], most of the literature describes *qacA* as plasmid-encoded. The *qacA* gene requires a transcriptional regulator for correct expression, which is coded by *qacR* [23]. This regulator gene is typically located in the upstream region of *qacA*, and the two genes are divergently transcribed. QacR functions as a repressor by binding to the DNA upstream of *qacA*, thus inhibiting expression. When cationic lipophilic compounds enter the cell, these bind to the repressor, resulting in its dissociation from the DNA, which in turn enables expression of *qacA* [23].

When the gene encoding **QacB** was sequenced, from plasmid pSK23, it was found to be very similar to *qacA*, despite the observed difference in substrate specificity that had been the basis of their distinction [24]. QacA and QacB only differ at six amino acid positions, of which the non-conserved amino acid change at position 323 (Asp in QacA, Ala in QacB) was demonstrated to be responsible for the difference in specificity [24]. Similar to *qacA*, expression of the *qacB* gene is regulated by the transcriptional repressor QacR.

The proteins QacA and QacB belong to the MFS superfamily of transporters. Members of this superfamily have a variable number of transmembrane domains (also known as transmembrane segments, TMS), and it is believed they

all evolved from a simple protein that contained a hairpin structure with two TMS [18]. This prototype seems to have triplicated into a protein containing six TMS, after which duplication formed a 12-TMS protein. The MFS proteins with 14 TMS, like QacA and QacB, have two extra, centrally positioned transmembrane domains, that are believed to be the result of an intragenic duplication [18].

The gene encoding **QacC** was first recognized in pSK89, a small rolling-circle replication plasmid of *S. aureus* [1]. QacC is only 107 amino acids long, though in other bacterial genera a length of 108 or 109 amino acids has been observed. These proteins contain four transmembrane domains and form dimers in the bacterial membrane. Expression of *qacC* does not require a transcriptional regulator. Littlejohn and coworkers recognized that QacC was identical to the *ebr* gene originally discovered by Sasatsu and coworkers [2]. The same gene was also isolated from a different *S. aureus* plasmid and published under the alternative name *smr*, for staphylococcal multidrug resistance [25, 26]. The identity in acronym may have been a coincidence, but *qacC*/*smr* happens to be a member of the SMR protein family of small multidrug resistance proteins [3, 17]. The name *smr* is nowadays used as synonym for *qacC* and seems to be favoured lately; here we refer to the gene as *qacC*/*smr*. The gene can be found on both conjugative and non-conjugative (rolling-circle) plasmids and has been described for both *S. aureus* and coagulase-negative *Staphylococcus* species [27]. Grinius and coworkers pointed out that QacC was similar to an efflux protein EmrE in *E. coli* [25]. An early characterization identified a glutamate residue at position 13 as essential for functionality [26]. Similar proteins are found in Gram-positive and Gram-negative bacteria, and most work performed to reveal structural requirements of SMR transporters has been carried out with *E. coli* EmrE. Although it was established that this transporter functions as a homo-dimer, the orientation of the two subunits remains unclear; some publications reported that these are oriented in parallel in the membrane, while other studies indicated an inverted dimer to be functional (reviewed in [28]). For *Staphylococcus* QacC, the exact position of the four alpha-helices, which form the transmembrane domains, was determined by nuclear magnetic resonance spectroscopy [29].

The name *qacD* was introduced in 1991 by Littlejohn and coworkers for a gene identical to *qacC* but with different promoter sequences [2]. These authors reserved the name **QacD** for proteins residing on large conjugative plasmids, but this name is no longer in use and the gene in question is now renamed as *qacC*.

The next addition to the Qac family was **QacE**, a protein of 115 amino acids first described in *E. coli*, of which functionally active deletion derivatives (known as QacEΔ) are found in other bacterial genera. Their genes were sequenced from broad-host plasmids replicating in Gram-negative bacteria [30]. In one publication, *qacE* genes were also described in Gram-positives, including in *Staphylococcus* [31], though to our knowledge this observation has not been confirmed since.

In the early 1990s it was recognized that *qac* genes were often contained on mobile elements, notably integrons, and that they could spread from one plasmid to another by site-specific recombination. The finding of a novel integron in a plasmid of *Enterobacter aerogenes* resulted in the next member the Qac family, **QacF** [32]. Although QacF (typically 110 amino acids, with a range of 108–112) has been detected in a number of Gram-positives [31], homologues of QacF have not been described for *Staphylococcus*.

QacH was published before QacG, though the same research group coined their names. In 1998 the gene encoding QacH was published, after it was identified on plasmid p2H6 of *S. saprophyticus* [33]. The *qacH* gene encodes a protein of 107 amino acids, with strong resemblance to other SMR members, and it confers high-level ethidium bromide resistance but low-level proflavine resistance [33]. The same group identified **QacG** as the protein encoding resistance to benzalkonium chloride and ethidium bromide from plasmid pST94 of *Staphylococcus*; the species from which this plasmid was isolated was not specified [34]. QacG of *Staphylococcus* spp. is 107 amino acids long.

The next Qac protein described was **QacJ** (107 amino acids long). Its gene was discovered from a newly identified rolling-circle replicating plasmid and subsequently identified in several equine isolates of *S. aureus*, *S. simulans* and *S. intermedius* [35].

Finally, **QacZ** was the name given for a new SMR-type Qac gene identified in *Enterococcus faecalis* [36]. This gene has so far not been described for *Staphylococcus* spp.

In conclusion, although their names suggest a similarity between all Qac proteins, there is a clear division between on the one hand QacA and QacB (both members of the MFS superfamily and collectively called QacA/B here), and on the other hand QacC – J plus QacZ, which are all members of the SMR protein family. Genetically, these two protein families evolved separately; their similarity in function is an example of parallel evolution. The name QacD is no longer in use, and QacE, QacF, and QacZ are absent from *Staphylococcus* species.

Detection and distribution of *qac* genes in *Staphylococcus* populations

Since resistance to biocide and disinfectants can severely hamper disinfection measures and current hygiene strategies to reduce nosocomial infections, the presence of *qac* genes has been investigated intensively in *Staphylococcus* populations from clinical settings as well as in the community. Most studies were carried out by PCR detection of the genes, sometimes backed up with phenotypic characterization to determine the nature and level of resistance. The primers used for studies employing classical PCR are summarized in Table 1. Based on multiple alignment of currently available *qac* sequences, three of the primers that have been in use can now be considered suboptimal:

Table 1. Primers used for detection of specific *qac* genes by classical PCR as described in the literature

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')	Amplicon length*	Described by	Also used by
<i>qacA/B</i>	GCAGAAAGTGCAGAGTTCCG	CCAGTCCAATCAATGCCTG (1)	361 bp	[37]	[15], [41], [42], [45], [51]
<i>qacA/B</i>	CTATGGCAATAGGAGATATGGTGT (1)	CCACTACAGATTCCTCAGCTACATG	417 bp (2)	[38]	
<i>qacA/B</i>	GCTGCAATTTATGACAAATGTTTG	AATCCACCTACTAAAGCAG	630 bp	[14]	
<i>qacA/B</i>	ATTCCAATTGAGTGCCTTTTC	GCCCTTTCTTTAGGGTTTTC	215 bp (2)	[5]	
<i>qacC/smr</i>	AAACAATGCACACCTACCCT	AACGAACTACGCCGACTATG	157 bp	[37]	[15], [38], [51]
<i>qacC/smr</i>	ATAAGTACTGAAAGTTATTGGAAAGT	TTCCGAAAAATGTTTAAACGAAACTA	286 bp	[14]	
<i>qacC/smr</i>	CTGAAGTTATTGGAAGTGCATT	GTTATATTTAGTGGTAGGTGTTGC	139 bp (2)	[5]	
<i>qacG</i>	TTTCGTTTGGAAATTTGCTTT	AATGGCTTCTCCAAATACA	213 bp	[15]	
<i>qacG</i>	CAACAGAAATAATCGGAAT	TACATTTAAGAGCACTACA (1)	275 bp	[39]	[5], [14], [47], [51]
<i>qacJ (3)</i>	GGCCAAACATTTAGGCACACTTA	TGACTTGTATCCAAAAACGTTAAGA	232 bp (2)	[15]	
<i>qacJ (3)</i>	CTTATAATTTAGTAATAGCG	GATCCAAAAACGTTAAGA	306 bp	[39]	[47], [51]
<i>qacJ (3)</i>	CCAACATTTAGGCACACTTTATTC	CAATTAATGACTTTGATCCAAAAACG	237 bp	[5]	
<i>qacH (3)</i>	CAATAGTCAGTGAAGTAATAGGCAGTG	TGTGATGATCCGAAATGTTT	295 bp	[15]	
<i>qacH (3)</i>	ATAGTCAGTGAAGTAATAG	AGTGTGATGATCCGAAATGT	295 bp	[39]	[14], [47], [51]

* Amplicon length was prediction based on classical PCR using *qac* sequences available from public databases. Other PCR-dependent methods (realtime PCR, nested PCR, RFLP-PCR) are not included in the table

(1) Nucleotides in bold represent positions where polymorphisms have been found in sequenced genes

(2) The predicted amplicon length differs from that reported in the literature

(3) Since only two sequences of *qacH* and only one of *qacJ* were publically available in GenBank at the time of analysis, conservation of the primer sequences could not be established

the reverse primer for detection of *qacA/B* designed by Noguchi et al. [37], and used by a number of other researchers, is unlikely to amplify *qacA* of plasmid SAP070B due to a mismatch in its last 3'-nucleotide (Table 1). Moreover, the forward primer used by Mayer et al. [38] includes the polymorphic nucleotide that distinguishes *qacA* from *qacB*, and although the primer was reported to hybridize to both types, sequence analysis of the amplicon would not reveal whether the tell-tale *qacA* or *qacB* nucleotide had been present. Lastly, two mismatches in the reversed primer for *qacG* amplification designed by Bjorland et al. [39] might hamper amplification of some *qacG* genes that have since been sequenced, due to non-conserved nucleotides.

Not all *qac* genes have been found in all *Staphylococcus* populations that are frequently monitored for reduced susceptibility to quaternary ammonium compounds. *QacA* and *qacB* are mainly found in *S. aureus* and *S. epidermidis*. *QacC* is also mostly found in these two species, though it can also be present in *S. warneri*. The other Qac names were introduced for proteins of the SMR family that were discovered in *Staphylococcus* spp. other than *S. aureus* or *S. epidermidis*, but with the exception of QacH, all of these have been described for the two medically relevant species *S. aureus* and *S. epidermidis*. A brief overview of the prevalence of the different *qac* genes that were determined in various bacterial populations is given here.

Prevalence in clinical settings

Worldwide differences in *qac* presence of bacterial populations can be observed, but these genes have been reported from most continents. The presence of various *qac* genes in a clinical setting obviously concentrated on *S. aureus*, whereby in some studies MRSA and methicillin-susceptible MSSA were analysed separately.

The prevalence of *qac* genes seems to be particularly high in Asia. For instance, the genes *qacA/B* and *qacC/smr* were found in 44% and 31%, respectively, in MRSA isolated between 1998 and 1999 in various Asian countries [37]. In human clinical isolates originating in Japan, *qacA/B* was found in 32% of MRSA, though at lower frequency (7.5%) in MSSA [40]. An alarmingly high prevalence of *qacA/B* (83%) in clinical MRSA isolates was recently reported from Malaysia, although *qacC/smr* was rarely found present [41].

Compared to the Asian situation, low prevalences were reported from the North-American continent. Investigations in Toronto, Canada, identified *qacA/B* in 2% of clinical MRSA isolates from an intensive care unit, with 7% positive for *qacC/smr* [42]. A lower incidence still was found in MRSA isolated from new admissions to nursing homes in the US, where only 0.6% of over 800 samples were positive for *qacA/B* [43]. Likewise, only 0.9% of clinical MRSA samples that had been submitted to a US surveillance network were positive for *qacA/B* by real-time PCR [44]. An international study comparing clinical and environmental *S. aureus* isolates from hospitals in three US states plus

Freiburg in Germany found variable prevalences, whereby *qacC/smr* was more often detected than *qacA/B* [45].

Incidences of *qacA/B* in Europe are generally higher than reported for the US. A large European study that investigated *S. aureus* isolates from 1997–1999 reported 42% positive for *qacA/B*, again with higher frequency in MRSA (63%) than in MSSA (12%); *qacC/smr* was detected in 6% of the isolates with no difference between MRSA and MSSA [38]. A study from Scotland demonstrated that the minimal bactericidal concentration (MBC) of chlorhexidine gluconate and triclosan was 10 to 1000-fold lower than the concentration recommended for disinfection, and that this concentration was increased for *Staphylococcus* isolates that contained *qac* genes [15]. The authors demonstrated the presence of *qacA* and of *qacC/smr* (sometimes in combination) and reported absence of *qacG*, *qacH* or *qacJ* [15]. In contrast, Furi and coworkers observed effects on minimal inhibitory concentrations (MIC) but not on MBC for chlorhexidine when *qacA/B* was present [5]. High prevalence of *qacA/B* (55%) was also recently observed in a Danish study [46], where strains of *S. epidermidis* were recovered from scrub nurses who used chlorhexidine heavily, and this hardly differed for non-users (50%). The MIC/MBC of these isolates did not correlate to presence or absence of *qacA/B* [46]. A lack of standard in experimental performance can explain these conflicting results [4].

Presence of genes other than *qacA/B* and *qacC/smr* in clinical isolates has not been extensively studied, but a few studies are available. In a study from the UK of clinical MRSA isolates, apart from *qacA/B* (8%) and *qacC/smr*, which was detected at a high frequency of 44%, *qacH* was also found in 3% of the strains, while *qacG* was absent and *qacJ* was not determined [14]. The gene *qacG* was identified, however, in 11 of 21 human clinical isolates of *S. haemolyticus* from Argentina, as were *qacH* and *qacJ* (in 10 and 4 isolates, respectively) [47]. In a Swedish study at a university hospital, *qacA/B* gene was present in 62/143 isolates (43%) and *qacC/smr* in 6%, while *qacH* was found once [48].

In historical isolates of *S. epidermidis* collected in the 1960s before the introduction of chlorhexidine, however, *qacA/qacB* were completely absent; the isolates in question were all from human patients in Denmark [46]. Thus, although presence of *qacA/B* does not always affect measurable MIC or MBC, long-term use of chlorhexidine or related compounds seems to have selected for presence of these genes in current clinically relevant populations.

Prevalence in community, food-producing animals and food

Since the importance of community-acquired *Staphylococcus* infection is drawing attention, presence of *qac* genes in community populations has also been investigated. A study from China identified few *S. aureus* but more coagulase-negative *Staphylococcus* (CNS) on an automated teller machine (ATM), as a sampling site representing

the bacterial population residing in the community. Fifteen per cent of the sampled ATMs contained *Staphylococcus* species, of which less than 1% were MRSA. Of the obtained *S. aureus*, 11% carried *qacA/B* and less than 2% had *qacC/smr*, while *qacC/smr* was found in 14% and *qacA/B* in 26% of the CNS isolates [49]. Another interesting approach to investigate community isolates was to sample the air in a Shanghai metro station but, although the primers used in this study were presented as specific for *qac*, unfortunately they detected a plasmid replication gene, so conclusions about presence of *qac* cannot be drawn [50].

Porcine *Staphylococcus* has been implied as a key population connecting community, food and clinical populations. In a study from Hong Kong, porcine MRSA contained *qacG* and *qacC/smr* but not *qacA/B*, *qacH* or *qacJ* [51]. In a recent study from Denmark that included 79 porcine *S. aureus* isolates, *qacG* was also identified. Strains containing *qacG*, as determined by PCR, were positive in an ethidium bromide efflux assay (Seier-Petersen and coworkers, unpublished data).

Coagulase-negative *Staphylococcus*, which can cause opportunistic infections, are of concern in the food-industry, in particular strains that produce enterotoxin (a virulence factor once thought to be exclusive for coagulase-positive *S. aureus*). Disinfectants are widely used in food production, so it is no surprise that *qac* genes have been found in CNS isolated from food sources. The genes *qacA/B*, *qacC/smr*, *qacG*, *qacH* and *qacJ* have all been reported from *S. epidermidis*, *S. saprophyticus*, *S. cohnii* and *S. hominis* [39, 52].

A comparison of *S. aureus qac* genes extracted from public sequence databases

In the studies described in the previous section, *qac* genes were mostly detected by PCR or real-time PCR but not

by sequencing; their presence and identity was inferred from the length of fragments generated by gene-specific primers. In contrast, genes that have been sequenced and submitted to public databases can be used for comparative analysis. In order to assess whether gene names have been used consistently, a name search was performed in the 'Gene' database at NCBI. The search revealed 6349 entries retrievable from that database with the query 'QacA', of which 198 were from *Staphylococcus* species. Only four entries were retrieved when queried with 'QacB', of which two were from *Staphylococcus*. This and other findings are summarized in Table 2.

The non-redundant protein or nucleotide databases at NCBI contain far more *qac* genes from *Staphylococcus* spp. than those identified in the 'Gene' database, but many are filed under alternative names. Nowadays, most proteins are deduced from (complete) genome sequences, and their annotation is automated, based on detected homologies to existing gene families. Depending on the program used for gene annotation, genes with homology to *qac* can receive various alternative names and annotations: MFS as synonym for QacA and Smr or EmrE for QacC were most frequently encountered (Table 2).

When the nonredundant database of NCBI was searched with QacB from pSK23 as the BlastP query, all hits to staphylococcal genes were close to 100% conserved. Some variation was observed in the start of the protein, which will be described in more detail in the next section. A hit with 99% identity was obtained to a *Listeria monocytogenes* protein and another to a protein from *Propionibacterium acnes* (though the first 124 amino acids were missing here). The most closely related protein from a different genus that was not 99% conserved over the complete aligned region (an indication that this protein evolved differently) was an 'antiseptic resistance protein' QacA from *Kocuria palustris*, with only 51% identity of the whole length of the protein.

Table 2. Search findings using the Gene database at NCBI

Gene database queried at NCBI, searched by name		Non-redundant protein database at NCBI, searched by sequence homology using PBLAST	
Search term	Nr. of entries found / in <i>Staphylococcus</i>	Query protein sequence	Protein names of hits
<i>qacA</i>	6349 / 198	QacA	QacA, QacB, MFS (99–100% identity)
<i>qacB</i>	4 / 2		
<i>qacC</i>	25 / 14	QacC	QacC, Smr, EmrE, QacG, QacH, SugE
<i>qacD</i>	1 / 1		
<i>smr</i>	5697 / 69		
<i>qacG</i>	3 / 0	QacG	QacC, Smr, EmrE, QacG, QacH, SugE
<i>qacH</i>	29 / 0	QacH	QacC, Smr, EmrE, QacG, QacH, SugE
<i>qacJ</i>	1 / 1	QacJ	QacC, Smr, EmrE, QacG, QacH, SugE
<i>ebr</i>	4474 / 0		

QacZ was not included in this search

A BlastP search with either QacC or QacG protein sequences as the query resulted in very different findings compared to the QacB Blast, as they retrieved many hits that were less strongly conserved. The hits obtained with either QacC or QacG were mostly overlapping and carried various names (Table 2); some were annotated as hypothetical. A considerable number of close relatives to QacC, QacJ or QacG were detected in genera other than *Staphylococcus*. In combination, these findings suggest that 1) the sequences of QacA/B are all highly conserved and the proteins seem to have evolved little since they spread in *Staphylococcus* species; 2) the other Qac proteins are all related to each other though they show a higher degree of sequence variation than QacA/B; and 3) QacC, J and G homologues are found in a number of other bacterial genera as well as in staphylococci, while QacA/B has only been detected in two other species outside the *Staphylococcus* genus so far.

Comparative analysis of QacA and QacB proteins and their genes

In order to compare their phylogenetic relationship, the protein sequences with >99% homology to QacB of *S. aureus* plasmid pSK23 that were identified by BlastP

as described in the previous section were aligned using ClustalW. Gene fragments missing more than 30 amino acids were discarded. This resulted in 25 different proteins, a mixture of QacA and QacB. Collectively they contained variation in 19 amino acid positions, ignoring five proteins that lacked the first few amino acids due to alternative, downstream start codons; in addition, the *L. monocytogenes* gene that was included lacked the last three residues (it was annotated as ‘partial’). A phylogenetic tree of these 25 different variants of QacA/B is shown in Fig. 1.

As can be seen in the figure, all proteins with predicted QacB substrate specificity (based on their Ala at position 323), are separated from the presumable QacA proteins. However, the genes from *S. epidermidis* and *S. aureus* are not found on separate clusters, and one protein variant was also detected in multiple *S. hominis* strains. This mixing of genes from different species within the observed clusters is to be expected, as the plasmids on which these genes reside are shared between *Staphylococcus* species.

The membrane spanning domains (TMS) of QacA/B and the loops facing outwards or inwards have been deduced from structural analysis, as summarized by Wu and coworkers [53]. Application of this structural model suggests that none of the variable amino acid positions are found in the cytoplasmic loops. Four are found within outward loops (E50D, E169D, Q170R, and D446A) but

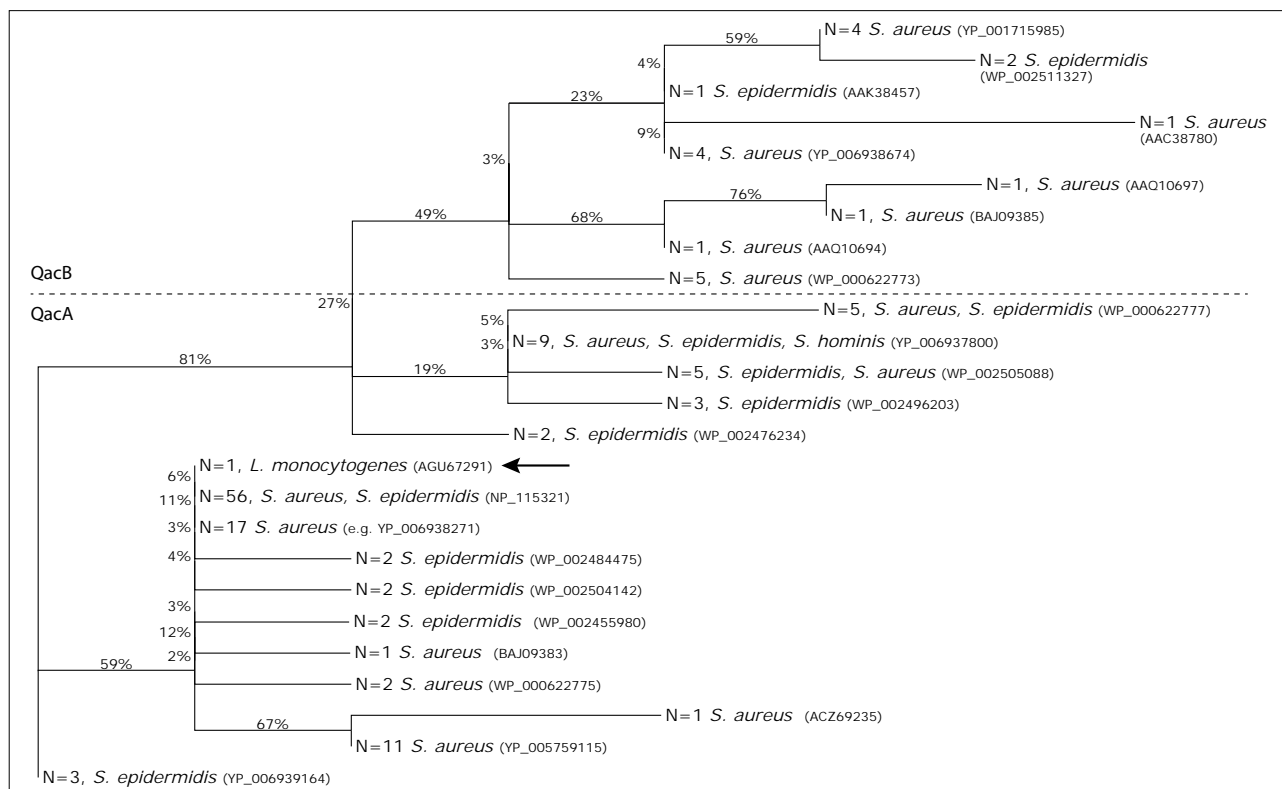


Fig. 1. Maximum likelihood phylogenetic tree of 25 different QacA/B proteins extracted from GenBank, together representing 142 submissions. Sequences shorter than 484 amino acids were removed. The arrow indicates a protein from *L. monocytogenes*, which lacks the last three amino acids. The number of entries per branch is given as N and for each branch; one accession number is given as an example in brackets. Cases where multiple entries covered proteins from different species are indicated. All proteins with the Ala at position 323, indicative of QacB substrate specificity, are above the dotted line. The tree was produced with PhyML at www.ATGC-montpellier.fr using default settings and visualized with TreeView (www.trex.uqam.ca)

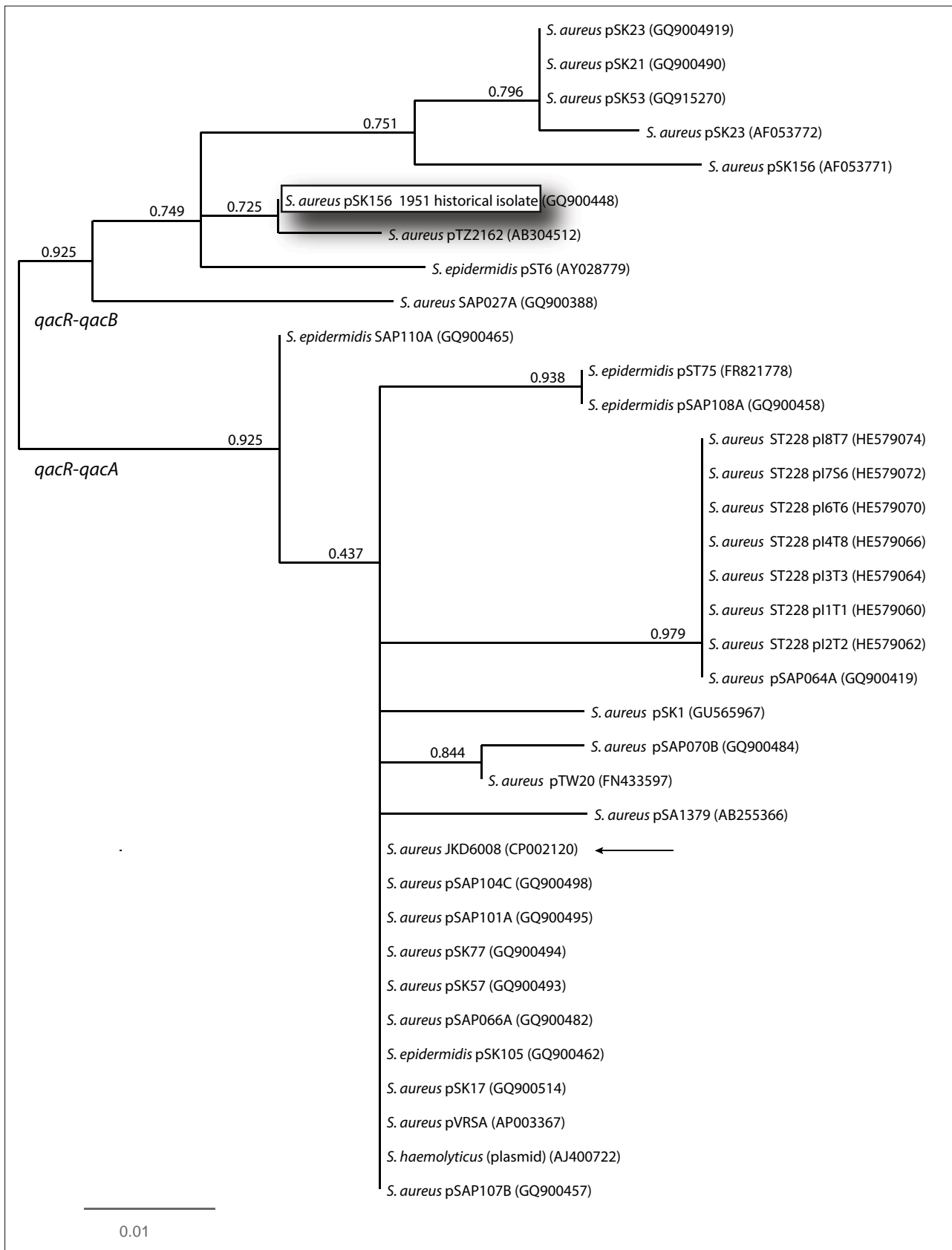


Fig. 2. Maximum likelihood phylogenetic tree of DNA sequences of the *qacA/B-qacR* locus, spanning both open reading frames and their intervening sequence. Sequences were identified by BlastN with 2289 nt from pSK156, a historical isolate from 1951 (shaded box). Hits were included only if they covered the complete query length and their accession number is shown in brackets. The arrow indicates a locus that is located on the chromosome of *S. aureus* strain JKD6008. The first node of the tree separates the *qacA* locus from *qacB*. The tree was produced with PhyML at www.trex.uqam.ca using default settings and visualized with TreeView

most variation occurs within the TMSs, in particular in TMS-5 (three variable positions) and TMS-6 and 12 (2 positions each). Finally, three variable positions are found at the border of a TMS, two of which are positioned at the border of a cytoplasmatic loop and a TMS (V137L for TMS-5 and V498F for TMS-14) and A290T is located where TMS-9 changes to an outside loop. A number of these variants have been described before by Nakaminami et al. [8]. These authors stated that QacB was more heterogeneous in sequence than QacA. This trend was weakly observed in our dataset: in total 36 mutations were observed in 9 QacB sequences, compared to 30 mutations in 16 QacA sequences. Most of the variation described by Nakaminami and coworkers did not have consequences to functionality or specificity of the protein, with the exception of A320E in TMS-10 of QacB, which resulted in reduced susceptibility to fluoroquinolones [8]. Several site-directed mutagenesis studies have been performed for QacA, but other than TMS-10, which seems essential for substrate specificity (this is where D323A is located that distinguishes QacA from QacB), few amino acid positions could be identified as essential for functionality [24, 53–57]. This suggests that the protein is relatively robust.

For evolutionary investigations, a DNA sequence comparison is more informative than protein analysis, as it also captures synonymous mutations. An analysis based on

DNA sequences was extended for the complete *qacA/B-qacR* locus, as QacA/B cannot be functional without its regulator, which means that the two genes have evolved together. DNA sequences covering the open reading frames of *qacR*, *qacA/B*, and their intervening sequence were collected by BlastN, using 2289 nucleotides of pSK156, a historical isolate from 1951, as the query [58]. The resulting sequences were aligned and a phylogenetic tree was constructed for hits covering the complete query (Fig. 2).

The historical isolate pSK156 is not positioned at the root of the tree shown in Fig. 2, indicating it does not resemble an ancestor of the genes captured here. The tree clearly splits *qacA*-bearing loci from those with *qacB*. Apparently, these two genetic lineages are distant enough to be detected as two separate clades when *qacR* and *qacA/B* DNA sequences are analyzed together. Moreover, another pSK156 sequence that is present in the database is not identical to the historical isolate (for the region analysed, eight mismatches and one single-nucleotide gap were detected). Whether these differences are due to sequencing errors or represent true biological variation could not be assessed. The *qac* locus from SequenceType ST228 isolates are all identical, as can be expected, since these are all representatives of a clone that recently spread globally. A chromosomal *qac* locus, present on an integrated mobile element, was indistinguishable from loci on plas-

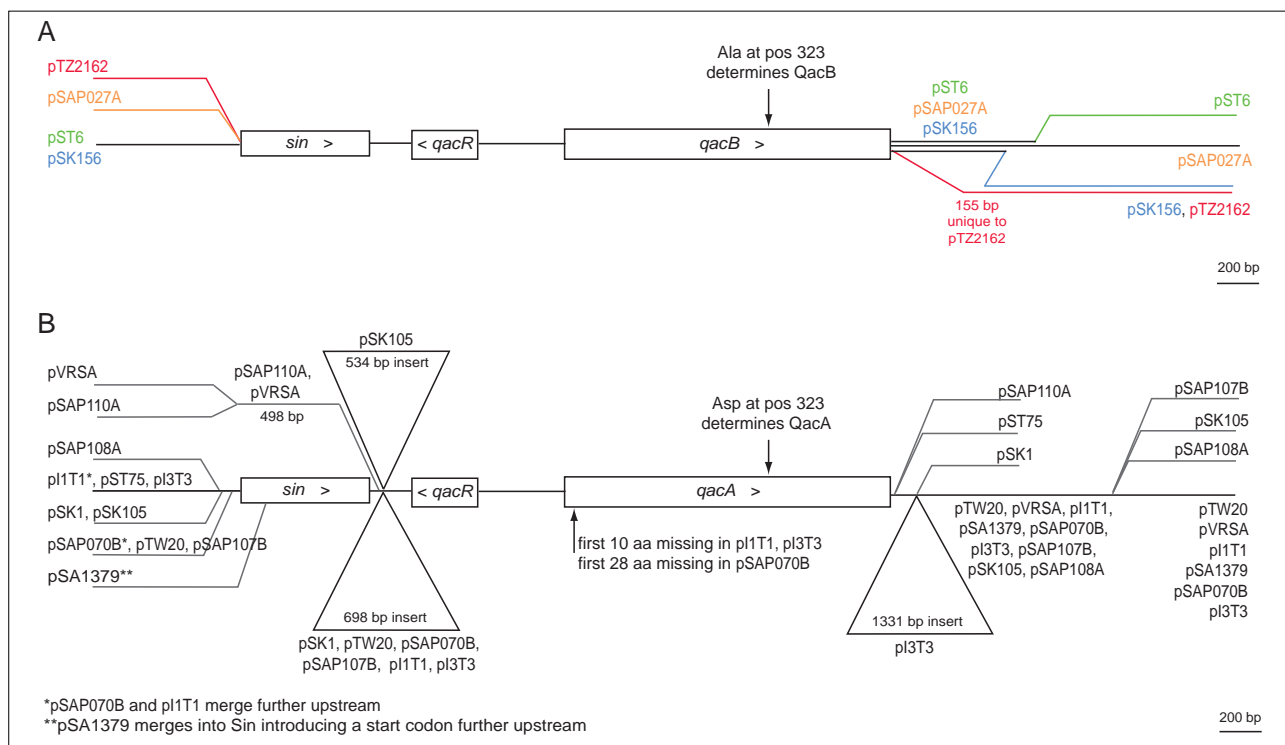


Fig. 3. Compilation of the extended *qac* locus with flanking regions of 16 completely sequenced plasmids. In panel A, the relevant part of four plasmids with *qacB* are shown in colour for clarity. Downstream of *qacR*, to the left in the figure, a recombinase *sin* gene is conserved. Beyond this, homology is lost. Directly downstream of *qacB*, pTZ2162 diverges from the other three plasmids, while pSK156 merges with this after 155 nucleotides. Further downstream still, pST6 diverges. Black represents sequences also found in QacA plasmids. In panel B, twelve plasmids with *qacA* are included. The Sin recombinase is absent in pVRSA and pSAP110A, which share 498 bp before they diverge. Open reading frames other than *sin*, *qacR*, and *qacA/B* are not shown. The figure is drawn to scale

mids from either *S. aureus*, *S. epidermidis* or *S. haemolyticus*.

The sequences flanking the *qacA/B-qacR* locus were also analysed, as these could potentially reveal information on how the locus spread from one plasmid to another. For this, the DNA sequences of 16 completely sequenced plasmids (selected on the basis that they were not identical) were compared for extended regions of both flanks, until homology was lost. A compilation of the results is shown in Fig. 3.

Even comparing only four extended *qacB* loci, shown in panel A of Fig. 3, already produces a complicated picture. Of note is a conserved gene coding Sin recombinase flanking *qacR*; the presence of this recombinase has been observed before [59], and a role of this recombinase in transfer of the *qac* locus can be postulated. However, no evidence of a conserved recombination site was identified downstream of *qacB*, where sequences diverged (pTZ2162) and converged again (pSK156, pTZ2162).

An even more complicated picture resulted when 12 *qacA*-bearing plasmids were compiled, as shown in Fig. 3B. Two different events have taken place in a number of plasmids between *sin* and *qacR* that resulted in insertions, while two plasmids have lost the *sin* gene completely (pSAP110A and pVRSA). As was observed for *qacB*-bearing plasmids, the plasmids shown in Fig. 3B diverge upstream of *sin*. The downstream flank of *qacA*, however, shows more conformity, and at least one *qacB* plasmid (pSAP027A) shares sequences with a number of *qacA* plasmids. Nevertheless, the evolutionary history of these plasmids can no longer be deduced, as too many recom-

ination events have occurred over time, one overprinting another, to recognize the order of events.

Comparative analysis of *qac* genes belonging to the SMR family

For a similar phylogenetic comparison of QacC, homologues to QacC from *Staphylococcus* species were obtained by BlastP with the gene from pSK89 as the query. Using a cut-off of identity over at least 60% of the query length, this search identified not only proteins annotated as QacC and its synonym Smr, but also QacG, QacH, and proteins annotated with alternative names (Table 2). Therefore, the search was extended with queries of the other *qac* genes, and all hits above the cutoff-value, for proteins with a minimum length of 103 and a maximum of 113 amino acids were collected. Records specified as partial sequences were ignored. A phylogenetic tree with the genes annotated as specific *qac* genes, plus their closest relatives, is shown at the top of Fig. 4.

Again, most genes were represented by multiple submissions. QacC was by far the most often recorded: a total of 76 Genbank entries were identified, with one variant being conserved in 70 cases. This predominant form of QacC, whose sequence is shown in panel B of Fig. 4 with the positions of the TMSs indicated as determined by Poget et al. [29], has so far been found in *S. aureus*, *S. epidermidis*, *S. warneri* and in *Enterococcus faecalis*. Two variants of QacG protein were identified, either from

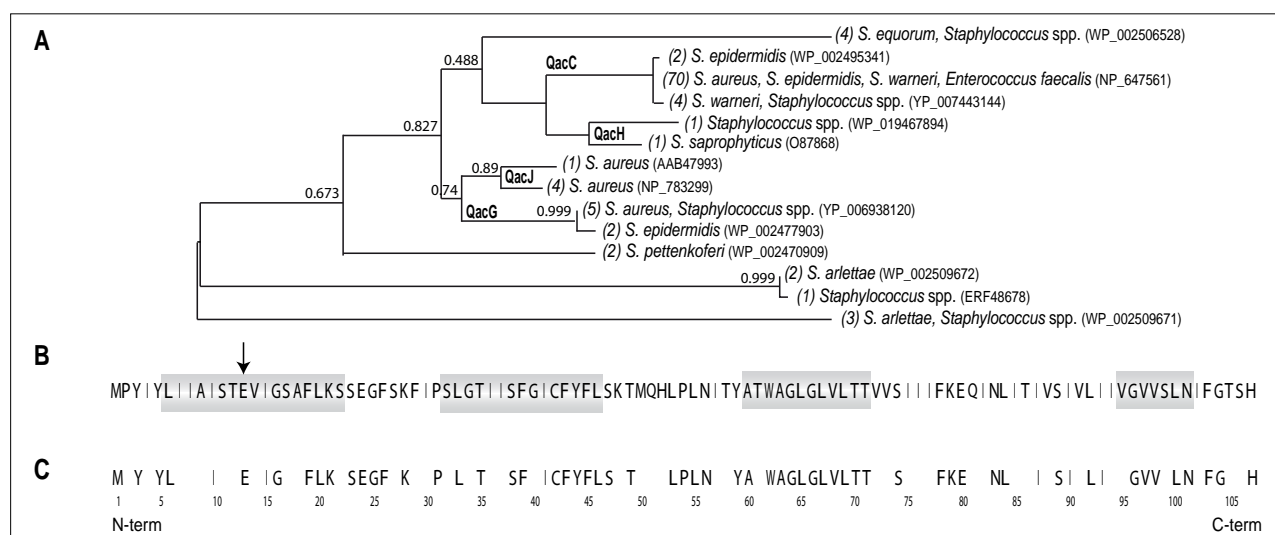


Fig. 4. Comparison of 14 different QacC, QacG, QacJ and QacH proteins, together with their closest homologues belonging to the SMR family from various *Staphylococcus* species. Panel A: Phylogenetic Neighbour Joining tree including the closest homologues of QacC, QacG, QacH and QacJ. Numbers between brackets indicate the number of entries in GenBank for a particular amino acid sequence, followed by the species in which these proteins were reported, with one accession number as an example. Bootstrap values are indicated. The tree was produced with PhyML at <http://phylogeny.lirmm.fr> using default settings. Panel B: Amino acid sequence for the most commonly recorded QacC protein (represented 70 times in GenBank). The boxes indicate alpha-helix transmembrane segments. The arrow points out Glutamate at position 13 that is essential for functionality. Panel C: Conserved amino acid positions based on a comparison of all proteins except for the two outliers from panel A. A sequence logo plot for these proteins is available as *Supplementary Fig. S1*

S. epidermidis or from *S. aureus* (both variants would be amplified with the *qacG* primers described by Smith et al. [15], shown in Table 1). Of the two variants of QacH, the one isolated from ‘*Staphylococcus* spp.’ was not annotated as such, but based on its close resemblance it is likely to represent QacH (note, however, that neither of the *qacH*-specific primer pairs shown in Table 1 would amplify this gene). The same applies to a variant of what is likely QacJ, which would again not be amplified by any of the *qacJ*-specific primer pairs shown in Table 1.

A comparison of the amino acid sequences from the main cluster containing all genes annotated as Qac plus the gene of *S. pettenkoferi* (the top 11 proteins in the tree) identified the conserved positions as shown in panel C of Fig. 4. A logo plot of this comparison is available as supplementary Fig. S1. The four TMSs are relatively strongly conserved, in particular the third TMS, while the loops display more variation. Whether this conservation is related to functionality of the protein has not yet been established. Even though the genes from the two bottom branches of the tree (isolated from *S. arlettae* and *Staphylococcus* spp.) are more distant, they are also likely to confer resistance to a broad variety of quaternary ammonium compounds.

The proteins represented in Fig. 4 are not the only members of the SMR family present in *Staphylococcus* species: the searches described above identified many more homologues that were more distantly related and variably annotated as quaternary ammonium compound transporters, SugE, Smr or with alternative descriptors. A tree was composed of all identified members of the SMR family from *Staphylococcus* species (Fig. 5, a version with each branch annotated is available as Supplementary Fig. S2). Four clusters were recognized, of which Cluster 4 contained the QacC proteins that were already shown in Fig. 4. The fact that all QacC proteins were clustered together and separated from all other SMR genes is in accordance with their specialized phenotype. SMR proteins from pathogenic *S. aureus* were only found in this Cluster 4, while protein genes from *S. epidermidis* were

found in Clusters 1, 3 and 4. Clusters 2 and 3 mainly contained proteins derived from coagulase-negative species, although *S. delphini* and *S. intermedius* were also represented. Interestingly, the only species containing proteins from all four clusters was *S. arlettae*.

In a review article by Bay and coworkers, the SMR family was subdivided into three subclasses, based on sequence comparisons across bacterial species [17]. These authors defined the subclass of small multidrug pumps (SMP) to which QacC-J of *Staphylococcus* and EmrE of *E. coli* belong, the SUG family with SugE of *E. coli* as the prototype (the name stands for suppressor of groEL mutation) and the group of paired small multidrug resistance proteins (PSMR) which produce a heterodimer of two proteins in the membrane [17]. In an attempt to identify if any of the four clusters we had identified coincided with these subclasses, representatives for each subclass from *E. coli* (EmrE, SugE, and the pair YdgE and YdgF for PSMR) were included in the analysis (results shown in Fig. S2). This identified that Cluster 3 was most similar to SugE but not EmrE of *E. coli*, while Cluster 4 with QacC was distantly related to *E. coli* EmrE, YdgE and YdgF. Only three genes (two from *S. arlettae* and one from *Staphylococcus* spp.) bore resemblance to the PSMR subclass of YdgE/F. That proteins belonging to the SugE-like Cluster 2 and the QacC Cluster 4 are transporter proteins is obvious, but they do not share phylogenetic characteristics to identify the three subclasses as defined by the *E. coli* genes. Whether the substrate specificity of each member within these two clusters is conserved remains unclear and would require phenotypic characterization. The function of Clusters 1 and 2 protein genes is mostly annotated as SugE or EmrE, for those cases where a functional annotation is specified, but phenotypic analysis would be required to justify these annotations.

For each of the four clusters of the tree shown in Fig. 5, the protein alignment was analysed to identify conserved domains. This identified telltale amino acids that were typically conserved in members of each cluster, summa-

Table 3. Conserved domains within clusters of the Staphylococcal SMR genes

	TMS1 (starting from pos. 9)	Extended TMS3 (including parts of surrounding loops)
Cluster 1	F I V AGx E xGV xLNE L V I	A II I K IPMGTA Y AIWTG I GT GGT GM FYxES G LV L T
Cluster 2	M I AGx F E LG xxNxY V V	I S V AV AI K P M TAY A WTG I G GG LGMxxYxES L G I TA TL T
Cluster 3	I Axxx E xWVxGL V	V I A TV LPVGT Y AxFV G G xG xxxxx F F S L T VT
Cluster 4	A I TA S S Ixx E IG FLKx EGF K S V SS A T	T T V LPLNx Y A WAGLGLVLTxxS xx F KE A S I

The clusters are numbered according to Fig. 5. Positions showing variation with 2 dominant amino acid residues are indicated. Positions with variation in 3 or more amino acids are shown as x. Amino acids conserved in all four clusters are shown in bold

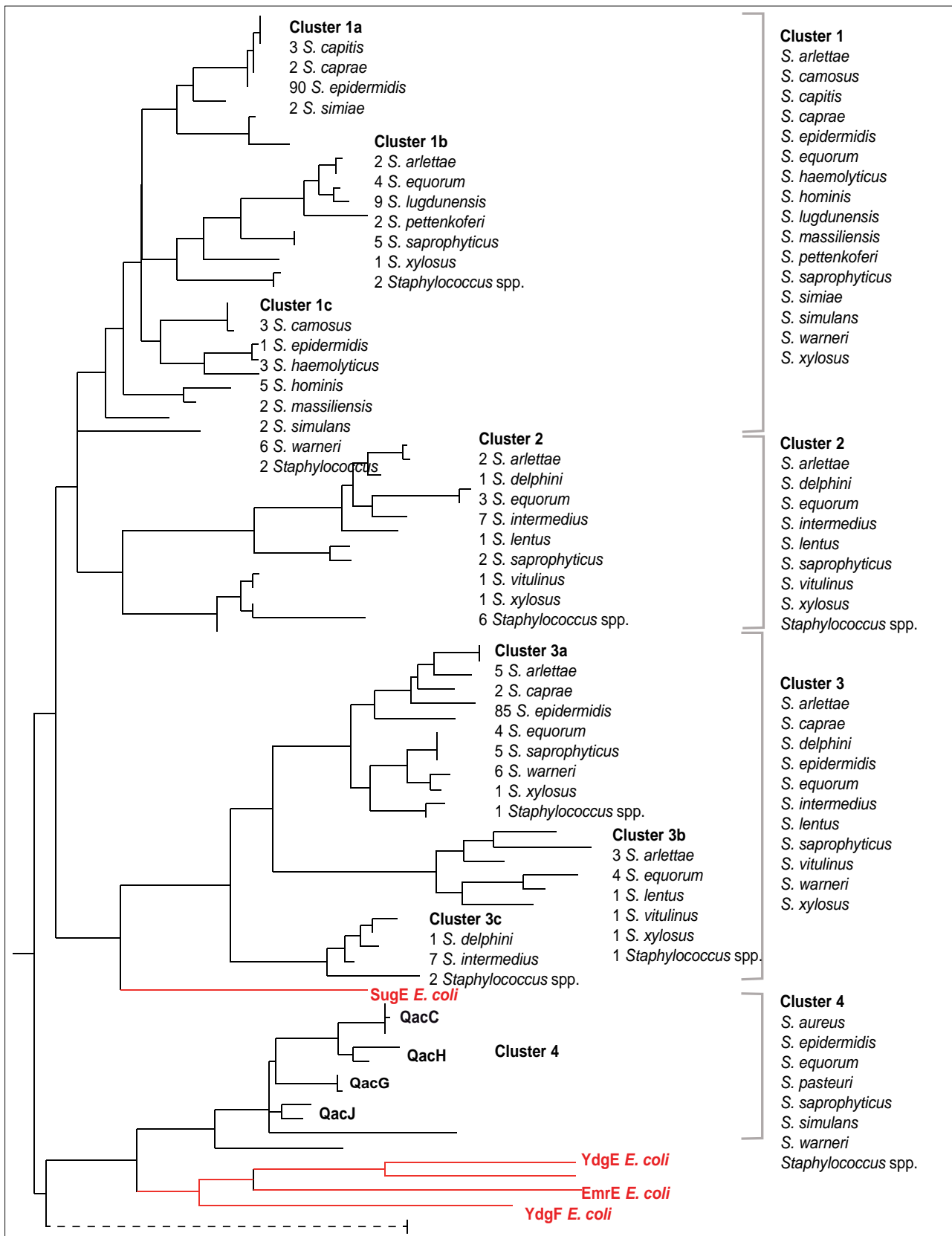


Fig. 5. Phylogenetic tree of all homologues belonging to the SMR protein family identified from *Staphylococcus* species. Four clusters, numbered 1 to 4, are indicated and the species represented in each cluster are listed alphabetically to their right. Staphylococcal Qac proteins all belong to Cluster 4. The four SMR genes of *E. coli* were added for comparison, shown in red. The dotted branches at the bottom identify *S. arlettae* proteins that possibly belong to the paired small multidrug resistance proteins. The tree was produced with PhyML at <http://phylogeny.lirmm.fr> using default settings. A complete version with annotated branches is available as *Supplementary Fig. S2*

rized in Table 3. The results are also available as logoplots in Supplementary Fig. S3. The two domains that are specifically conserved per cluster coincide with the predicted TMS1 and TMS3, the latter extended with parts of its surrounding loops (Table 3). As indicated in Table 3, a few amino acids are conserved in all four clusters, and these include the glutamate (E) at position 13, which was shown to be essential for QacC functionality. Based on its conservation it is likely to be essential for all SMR proteins, though not specific for QacC function.

Based on the findings summarized in Fig 5 and Table 3, it is questionable whether the nominal distinction of four different protein types within one cluster (QacC, QacG, QacJ and QacH) and the indiscriminate annotation of all other proteins as EmrE/SugE is reflective of the genetic variation and conservation within this large family. Genetically speaking, all genes of Cluster 4 are close enough to be given the same name, as they are more closely related to each other than to the other clusters. The current nomenclature therefore seems to be over-discriminatory, and is not a true reflection of genetic relationship.

Detection of *qac* genes in other bacterial genera

The *QacA/B* locus has so far been detected in *S. aureus* and *S. epidermidis* mainly, and in a few cases *qacA* was found in *S. hominis*. Of concern is, however, that *qacA* was also detected in a *L. monocytogenes* sequence, submitted directly to GenBank (Accession number AGU57291). The information provided was that the source of the *L. mono-*

cytogenes strain was frozen food from China. Whether the sequence was obtained from chromosomal or plasmid DNA was not recorded. Unfortunately, the sequence only partially covers *qacA*, and no flanking sequences are currently available to assess if the gene resides on a mobile element; it is even possible that the sequenced sample was contaminated with Staphylococcal DNA. However, the recent detection (by PCR) of *qacA/B* in two *E. faecalis* isolates, one from cattle and one from human blood [60], suggest that QacA/B is being exchanged between bacterial genera. The direction of such transfer is not known, but since these genes have so far only been described in Staphylococcus species, it is tempting to assume the transfer was to, and not from *E. faecalis*.

A similar situation is found for the SMR family members of *Staphylococcus* species, where *qacC/smr* has been found in *E. faecalis* (shown in Fig. 5). In this case, the gene was identified from a complete genome sequence but it was found on a short contig that has not (yet) been assembled into context. Bischoff and coworkers also produced a *qacC* amplicon from *E. faecalis* from human blood, and another from cheese [59]. None of the other *qac* members of the SMR family have so far been recorded in other genera: thus far, most of the SMR genes described for *Staphylococcus* species seem to be distinct from those in other bacteria.

Figure 6 shows a phylogenetic tree of the closest relatives of staphylococcal SMR. These proteins share over 55% identity to either staphylococcal QacC or QacH and were identified by BlastP. An SMR member of each cluster of staphylococcal proteins was included for comparison. Although no further proteins identical to a *Staphylococcus* gene were identified, the similarity of some of these

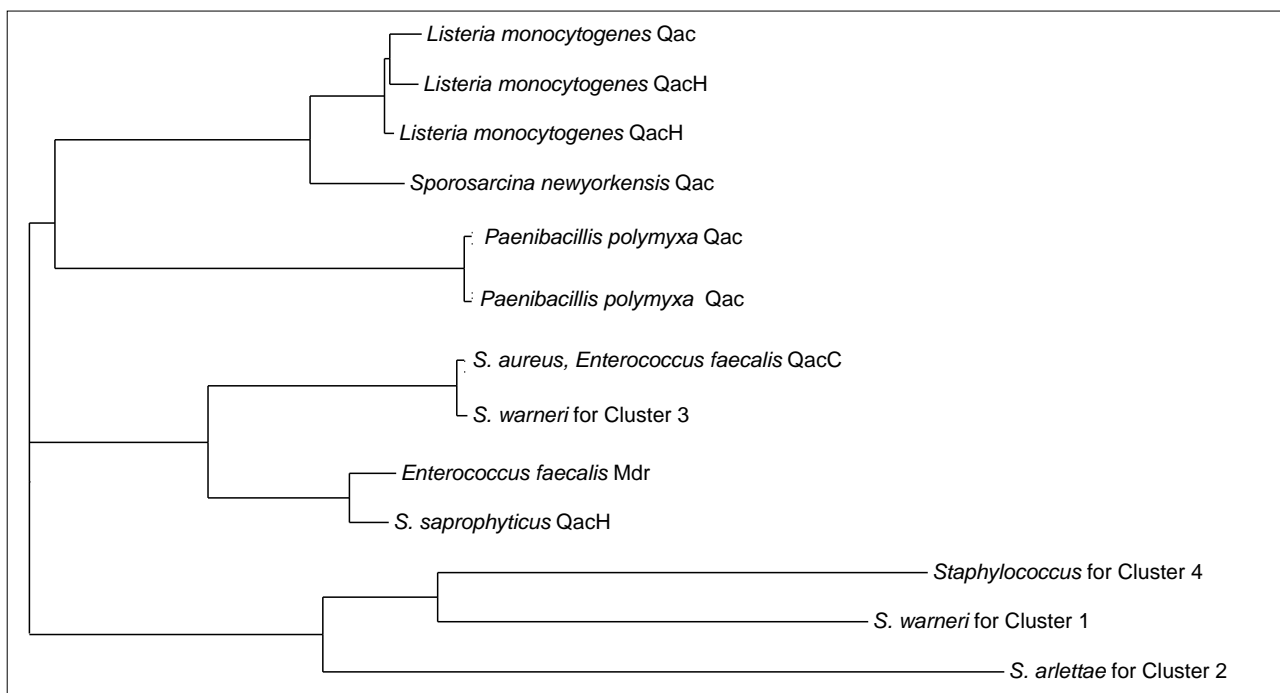


Fig. 6. Closest relatives to SMR proteins of *Staphylococcus* species identified in other genera. The three branches at the bottom relate to Fig. 4. The tree was produced with PhyML at <http://phylogeny.lirmm.fr> using default settings

proteins from other genera is strong enough to predict they could be potentially shared, in particular between staphylococci and *E. faecalis*, as these species are known to exchange genes.

Conclusions

The nomenclature of efflux pumps resulting in reduced susceptibility to quaternary ammonium compounds and their genes has historical roots and unfortunately can be confusing and ambiguous. The universal name ‘Qac’ misses the fact that these proteins represent two major families, separating QacA and QacB (MFS superfamily members) from all other Qac proteins that belong to the SMR superfamily. These two major protein families have little in common, other than a similarity in efflux function with some shared substrates. The *qacA* and *qacB* genes show limited sequence variation and are mainly found in *S. aureus* and *S. epidermidis*. Apart from an identical gene found in *E. faecalis*, their closest homologues in different genera are quite distinct. The MFS-type *qac* genes, on the other hand, are less conserved within the *Staphylococcus* genus, and similar genes are present in other bacterial genera as well. Although QacC, QacG, QacH and QacJ reflect different branches of a phylogenetic tree, their nomenclatural distinction may be over-precise, in view of all other SMR-type efflux pump genes. The name ‘Smr’ has recently been favoured over QacC in *Staphylococcus* species, but this potentially leads to confusion, as ‘SMR’ is also the acronym for the larger protein family to which QacC belongs. Compared to all staphylococcal proteins of this SMR family, the amino acid sequences of QacC-J all cluster together, while a wide variety of other proteins of this family are represented in *Staphylococcus* species that are more distantly related to QacC. Many of these other SMR members do not alter susceptibility to quaternary ammonium compounds. The confusing nomenclature will not change soon, but the phylogenetic analyses presented here may at least assist in recognition of key sequence features of the SMR members, and to a better identification of these important genes.

Acknowledgements

The authors thank all scientists who submitted their valuable sequences to public databases, without which this study could not have been performed. This study was supported by the Danish Research Council for Strategic Research (2101-08-0030).

References

1. Lyon BR, Skurray R: Antimicrobial resistance of *Staphylococcus aureus*: genetic basis. *Microbiol Rev* 51, 88–134 (1987)
2. Littlejohn TG, DiBerardino D, Messerotti LJ, Spiers SJ, Skurray RA: Structure and evolution of a family of genes encoding antiseptic and disinfectant resistance in *Staphylococcus aureus*. *Gene* 101, 59–66 (1991)
3. Putman M, van Veen HW, Konings WN: Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* 64, 672–693 (2000)
4. Buffet-Bataillon S, Tattevin P, Bonnaure-Mallet M, Jolivet-Gougeon A: Emergence of resistance to antibacterial agents: the role of quaternary ammonium compounds – a critical review. *Int J Antimicrob Agents* 39, 381–389 (2012)
5. Furi L, Ciusa ML, Knight D, Di Lorenzo V, Tocci N, Cirasola D, Aragones L, Coelho JR, Freitas AT, Marchi E, Moce L, Visa P, Northwood JB, Viti C, Borghi E, Orefici G; BIOHYPO Consortium, Morrissey I, Oggioni MR: Evaluation of reduced susceptibility to quaternary ammonium compounds and bisbiguanides in clinical isolates and laboratory-generated mutants of *Staphylococcus aureus*. *Antimicrob Agents Chemother* 57, 3488–3497 (2013)
6. Costa SS, Viveiros M, Amaral L, Couto I: Multidrug Efflux Pumps in *Staphylococcus aureus*: an Update. *Open Microbiol J* 7, 59–71 (2013)
7. Fuentes DE, Navarro CA, Tantaleán JC, Araya MA, Saavedra CP, Pérez JM, Calderón IL, Youderian PA, Mora GC, Vásquez CC: The product of the *qacC* gene of *Staphylococcus epidermidis* CH mediates resistance to beta-lactam antibiotics in gram-positive and gram-negative bacteria. *Res Microbiol* 156, 472–477 (2005)
8. Nakaminami H, Noguchi N, Sasatsu M: Fluoroquinolone efflux by the plasmid-mediated multidrug efflux pump QacB variant QacBIII in *Staphylococcus aureus*. *Antimicrob. Agents Chemother* 54, 4107–4111 (2010)
9. Lowy FD: *Staphylococcus aureus* infections. *N Engl J Med* 339, 520–532 (1998)
10. Taylor AR: Methicillin-resistant *Staphylococcus aureus* infections. *Prim Care* 40, 637–654 (2013)
11. Uhlemann A-C, Otto M, Lowy FD, DeLeo FR: Evolution of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus*. *Infect Genet Evol* 21, 563–574 (2014)
12. Fluit AC: Livestock-associated *Staphylococcus aureus*. *Clin Microbiol Infect* 18, 735–744 (2012)
13. Otto M: *Staphylococcus epidermidis* – the “accidental” pathogen. *Nat Rev Microbiol* 7, 555–567 (2009)
14. Vali L, Davies SE, Lai LLG, Dave J, Amyes SGB: Frequency of biocide resistance genes, antibiotic resistance and the effect of chlorhexidine exposure on clinical methicillin-resistant *Staphylococcus aureus* isolates. *J Antimicrob Chemother* 61(3), 524–532 (2008)
15. Smith K, Gemmell CG, Hunter IS: The association between biocide tolerance and the presence or absence of *qac* genes among hospital-acquired and community-acquired MRSA isolates. *J Antimicrob Chemother* 61, 78–84 (2008)
16. Cervinkova D, Babak V, Marosevic D, Kubikova I, Jaglic Z: The role of the *qacA* gene in mediating resistance to quaternary ammonium compounds. *Microb Drug Resist* 19, 160–167 (2013)
17. Bay DC, Turner RJ: Diversity and evolution of the small multidrug resistance protein family. *BMC Evol Biol* 9, 140 (2009)
18. Reddy VS, Shlykov MA, Castillo R, Sun EI, Saier MH: The major facilitator superfamily (MFS) revisited. *FEBS J* 279, 2022–2035 (2012)

19. Johnston LH, Dyke KG: Ethidium bromide resistance, a new marker on the staphylococcal penicillinase plasmid. *J Bacteriol* 100, 1413–1414 (1969)
20. Sasatsu M, Shima K, Shibata Y, Kono M: Nucleotide sequence of a gene that encodes resistance to ethidium bromide from a transferable plasmid in *Staphylococcus aureus*. *Nucleic Acids Res* 17, 10103 (1989)
21. Rouch DA, Cram DS, DiBerardino D, Littlejohn TG, Skurray RA: Efflux-mediated antiseptic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins. *Mol Microbiol* 4, 2051–62 (1990)
22. Gillespie MT, Lyon BR, Skurray RA: Gentamicin and antiseptic resistance in epidemic methicillin-resistant *Staphylococcus aureus*. *Lancet* 1(8636), 503 (1989)
23. Peters KM, Sharbeen G, Theis T, Skurray RA, Brown MH: Biochemical characterization of the multidrug regulator QacR distinguishes residues that are crucial to multidrug binding and induction of *qacA* transcription. *Biochem* 48, 9794–9800 (2009)
24. Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA: Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity. *Proc Natl Acad Sci USA* 93, 3630–3635 (1996)
25. Grinius L, Dreguniene G, Goldberg EB, Liao CH, Projan SJ: A staphylococcal multidrug resistance gene product is a member of a new protein family. *Plasmid* 27, 119–129 (1992)
26. Grinius LL, Goldberg EB: Bacterial multidrug resistance is due to a single membrane protein which functions as a drug pump. *J Biol Chem* 269, 29998–30004 (1994)
27. Leelaporn A, Firth N, Paulsen IT, Hettiaratchi A, Skurray RA: Multidrug resistance plasmid pSK108 from coagulase-negative staphylococci; relationships to *Staphylococcus aureus qacC* plasmids. *Plasmid* 34, 62–67 (1995)
28. Schuldiner S: EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim Biophys Acta* 1794, 748–762 (2009)
29. Poget SF, Harris R, Cahill SM, Girvin ME: 1H, 13C, 15N backbone NMR assignments of the *Staphylococcus aureus* small multidrug-resistance pump (Smr) in a functionally active conformation. *Biomol NMR Assign* 4, 139–142 (2010)
30. Paulsen IT, Littlejohn TG, Rådström P, Sundström L, Sköld O, Swedberg G, Skurray RA: The 3' conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants. *Antimicrob Agents Chemother* 37, 761–768 (1993)
31. Kazama H, Hamashima H, Sasatsu M, Arai T: Distribution of the antiseptic-resistance gene *qacE* delta 1 in gram-positive bacteria. *FEMS Microbiol Lett* 165, 295–299 (1998)
32. Ploy MC, Courvalin P, Lambert T: Characterization of In40 of *Enterobacter aerogenes* BM2688, a class 1 integron with two new gene cassettes, *cmlA2* and *qacF*. *Antimicrob Agents Chemother* 42, 2557–2563 (1998)
33. Heir E, Sundheim G, Holck AL: The *Staphylococcus qacH* gene product: a new member of the SMR family encoding multidrug resistance. *FEMS Microbiol Lett* 163, 49–56 (1998)
34. Heir E, Sundheim G, Holck AL: The *qacG* gene on plasmid pST94 confers resistance to quaternary ammonium compounds in staphylococci isolated from the food industry. *J Appl Microbiol* 86, 378–88 (1999)
35. Bjorland J, Steinum T, Sunde M, Waage S, Heir E: Novel plasmid-borne gene *qacJ* mediates resistance to quaternary ammonium compounds in equine *Staphylococcus aureus*, *Staphylococcus simulans*, and *Staphylococcus intermedius*. *Antimicrob Agents Chemother* 47, 3046–3052 (2003)
36. Braga TM, Marujo PE, Pomba C, Lopes MFS: Involvement, and dissemination, of the enterococcal small multidrug resistance transporter QacZ in resistance to quaternary ammonium compounds. *J Antimicrob Chemother* 66, 283–286 (2011)
37. Noguchi N, Suwa J, Narui K, Sasatsu M, Ito T, Hiramatsu K, Song JH: Susceptibilities to antiseptic agents and distribution of antiseptic-resistance genes *qacA/B* and *smr* of methicillin-resistant *Staphylococcus aureus* isolated in Asia during 1998 and 1999. *J Med Microbiol* 54, 557–565 (2005)
38. Mayer S, Boos M, Beyer A, Fluit AC, Schmitz FJ: Distribution of the antiseptic resistance genes *qacA*, *qacB* and *qacC* in 497 methicillin-resistant and -susceptible European isolates of *Staphylococcus aureus*. *J Antimicrob Chemother* 47, 896–897 (2001)
39. Bjorland J, Steinum T, Kvitle B, Waage S, Sunde M, Heir E: Widespread distribution of disinfectant resistance genes among *Staphylococci* of bovine and caprine origin in Norway. *J Clin Microbiol* 43, 4363–4368 (2005)
40. Alam MM, Kobayashi N, Uehara N, Watanabe N: Analysis on distribution and genomic diversity of high-level antiseptic resistance genes *qacA* and *qacB* in human clinical isolates of *Staphylococcus aureus*. *Microb Drug Resist* 9, 109–121 (2003)
41. Shamsudin MN, Alreshidi MA, Hamat RA, Alshrari AS, Atshan SS, Neela V: High prevalence of *qacA/B* carriage among clinical isolates of methicillin-resistant *Staphylococcus aureus* in Malaysia. *J Hosp Infect* 81, 206–208 (2012)
42. Longtin J, Seah C, Siebert K, McGeer A, Simor A, Longtin Y, Low DE, Melano RG: Distribution of antiseptic resistance genes *qacA*, *qacB*, and *smr* in methicillin-resistant *Staphylococcus aureus* isolated in Toronto, Canada, from 2005 to 2009. *Antimicrob Agents Chemother* 55, 2999–3001 (2011)
43. McDanel JS, Murphy CR, Diekema DJ, Quan V, Kim DS, Peterson EM, Evans KD, Tan GL, Hayden MK, Huang SS: Chlorhexidine and mupirocin susceptibilities of methicillin-resistant *Staphylococcus aureus* from colonized nursing home residents. *Antimicrob Agents Chemother* 57, 552–558 (2013)
44. Mc Gann P, Milillo M, Kwak YI, Quintero R, Waterman PE, Lesho E: Rapid and simultaneous detection of the chlorhexidine and mupirocin resistance genes *qacA/B* and *mupA* in clinical isolates of methicillin-resistant *Staphylococcus aureus*. *Diagn Microbiol Infect Dis* 77, 270–272 (2013)
45. Kosmidis C, Schindler BD, Jacinto PL, Patel D, Bains K, Seo SM, Kaatz GW: Expression of multidrug resistance efflux pump genes in clinical and environmental isolates of *Staphylococcus aureus*. *Int J Antimicrob Agents* 40, 204–209 (2012)
46. Skovgaard S, Larsen MH, Nielsen LN, Skov RL, Wong C, Westh H, Ingmer H: Recently introduced *qacA/B* genes in *Staphylococcus epidermidis* do not increase chlorhexidine

- MIC/MBC. *J Antimicrob Chemother* 68, 2226–2233 (2013)
47. Correa JE, De Paulis A, Predari S, Sordelli DO, Jeric PE: First report of *qacG*, *qacH* and *qacJ* genes in *Staphylococcus haemolyticus* human clinical isolates. *J Antimicrob Chemother* 62, 956–960 (2008)
 48. Prag G, Falk-Brynhildsen K, Jacobsson S, Hellmark B, Unemo M, Söderquist B. Decreased susceptibility to chlorhexidine and prevalence of disinfectant resistance genes among clinical isolates of *Staphylococcus epidermidis*. *APMIS* 122, 961–967 (2014)
 49. Zhang M, O'Donoghue M, Boost M V: Characterization of staphylococci contaminating automated teller machines in Hong Kong. *Epidemiol Infect* 140, 1366–1371 (2012)
 50. Zhou F, Wang Y: Characteristics of antibiotic resistance of airborne *Staphylococcus* isolated from metro stations. *Int J Environ Res Public Health* 10, 2412–2426 (2013)
 51. Wong TZ, Zhang M, O'Donoghue M, Boost M: Presence of antiseptic resistance genes in porcine methicillin-resistant *Staphylococcus aureus*. *Vet Microbiol* 162, 977–979 (2013)
 52. Heir E, Sundheim G, Holck AL: Identification and characterization of quaternary ammonium compound resistant staphylococci from the food industry. *Int J Food Microbiol* 48, 211–219 (1999)
 53. Wu J, Hassan KA, Skurray RA, Brown MH: Functional analyses reveal an important role for tyrosine residues in the staphylococcal multidrug efflux protein QacA. *BMC Microbiol* 8, 147 (2008)
 54. Xu Z, O'Rourke BA, Skurray RA, Brown MH: Role of transmembrane segment 10 in efflux mediated by the staphylococcal multidrug transport protein QacA. *J Biol Chem* 281, 792–799 (2006)
 55. Hassan KA, Galea M, Wu J, Mitchell BA, Skurray RA, Brown MH: Functional effects of intramembraneous proline substitutions in the staphylococcal multidrug transporter QacA. *FEMS Microbiol. Lett* 263, 76–85 (2006)
 56. Hassan KA, Skurray RA, Brown MH: Transmembrane helix 12 of the *Staphylococcus aureus* multidrug transporter QacA lines the bivalent cationic drug binding pocket. *J Bacteriol* 189, 9131–9134 (2007)
 57. Hassan KA, Souhani T, Skurray RA, Brown MH: Analysis of tryptophan residues in the staphylococcal multidrug transporter QacA reveals long-distance functional associations of residues on opposite sides of the membrane. *J Bacteriol.* 190, 2441–2449 (2008)
 58. Paulsen IT, Brown MH, Skurray RA: Characterization of the earliest known *Staphylococcus aureus* plasmid encoding a multidrug efflux system. *J Bacteriol* 180, 3477–3479 (1998)
 59. Sidhu MS, Heir E, Sørum H, Holck A: Genetic linkage between resistance to quaternary ammonium compounds and beta-lactam antibiotics in food-related *Staphylococcus* spp. *Microb Drug Resist* 7, 363–371 (2001)
 60. Bischoff M, Bauer J, Preikschat P, Schwaiger K, Mölle G, Hölzel C: First detection of the antiseptic resistance gene *qacA/B* in *Enterococcus faecalis*. *Microb Drug Resist* 18, 7–12 (2012)

Supplement

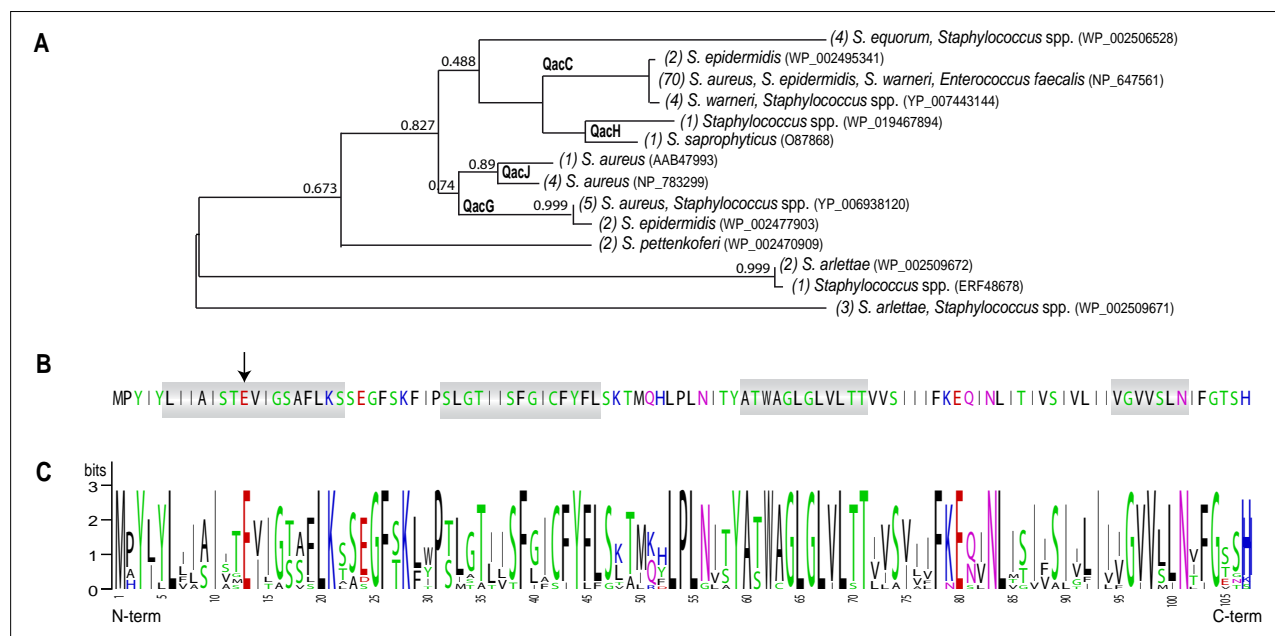


Fig. S1. Comparison of 14 different QacC, QacG, QacJ and QacH proteins, together with their closest homologues belonging to the SMR family from various *Staphylococcus* species. Panel A: Phylogenetic Neighbour Joining tree including the closest homologues of QacC, QacG, QacH and QacJ. Numbers between brackets indicate the number of entries in GenBank for a particular amino acid sequence, after the species in which these genes were reported are given, followed by an example accession number. Bootstrap values are indicated. The tree was produced with PhyML at <http://phylogeny.lirmm.fr> using default settings. Panel B: Amino acid sequence for the most commonly recorded QacC protein (represented 70 times in GenBank). The boxes indicate alpha-helix transmembrane segments. The arrow points out Glutamate at position 13 that is essential for functionality. Panel C: Sequence logo plot based on a comparison of all proteins except for the two outliers from panel A

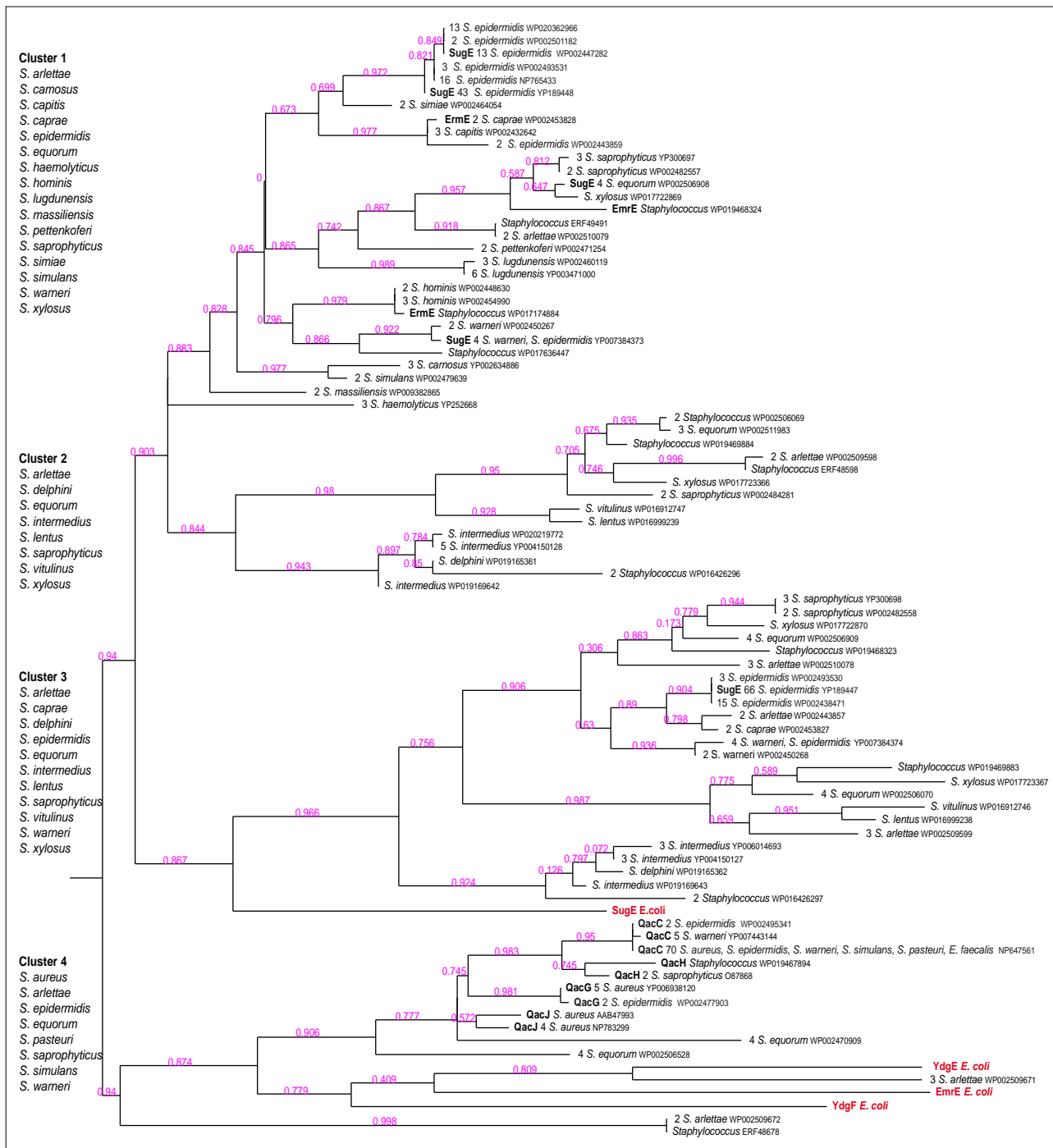


Fig. S2. Phylogenetic tree of all homologues belonging to the SMR protein family identified from *Staphylococcus* species. Each branch gives the species for which this gene was found and their numbers when >1, followed by an example accession number. Bootstrap values are indicated. Four clusters, numbered 1 to 4, are indicated and the species represented in each cluster are listed alphabetically to the left. Staphylococcal Qac proteins all belong to Cluster 4. The four SMR genes of *E. coli* were added for comparison, shown in red. The dotted branches at the bottom identify *S. arlettae* proteins that possibly belong to the paired small multidrug resistance proteins. The tree was produced with PhyML at <http://phylogeny.lirmm.fr> using default settings

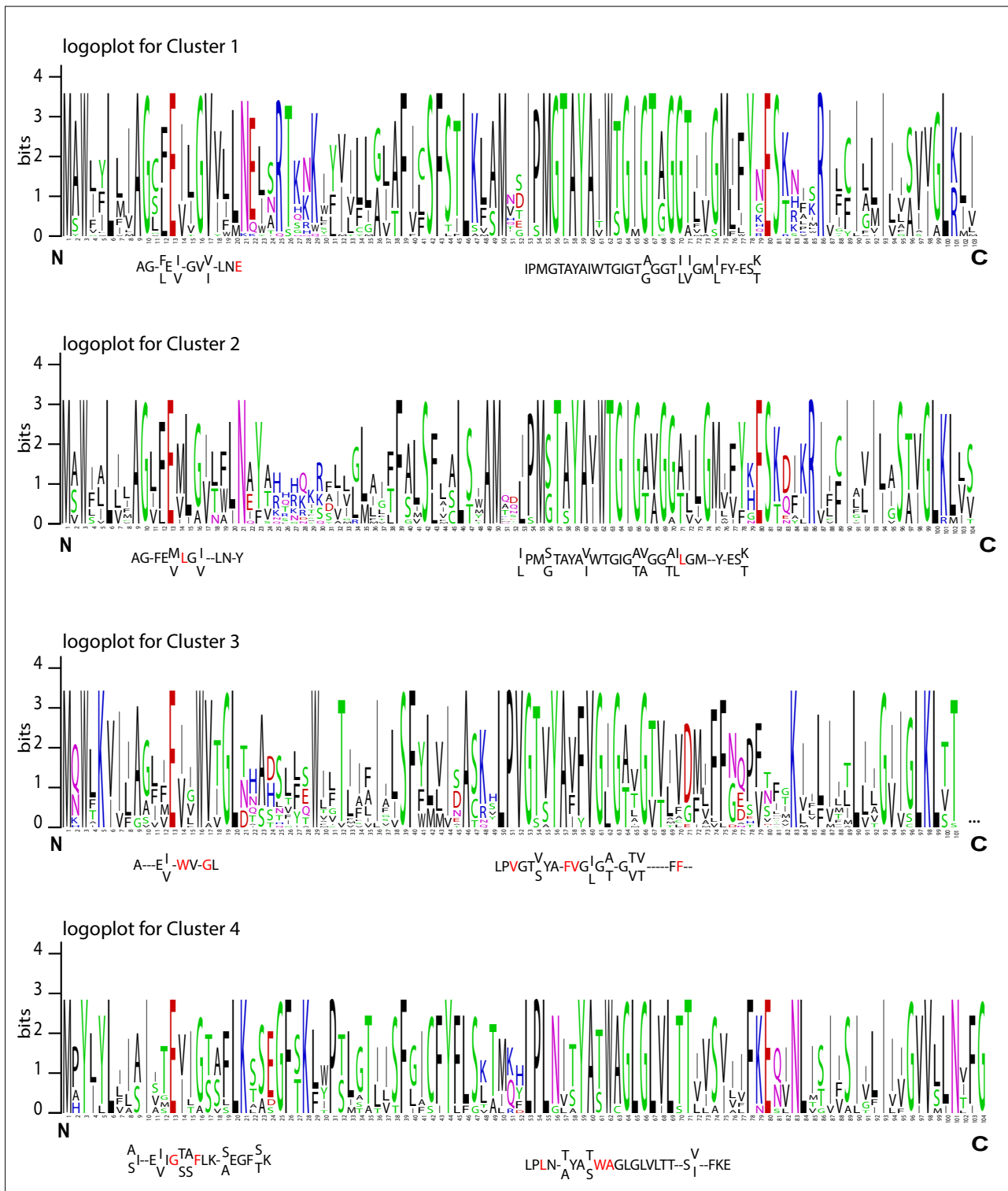


Fig. S3. Logoplot for the four clusters formed by SMR genes of *Staphylococcus* spp. The *qacC/Smr* genes belong to Cluster 4. Consensus sequences for two domains that are conserved within each cluster are indicated. The plots were produced at weblogo.berkeley.edu