Whole genome sequence and comparative genomics analysis of multi-drug resistant environmental *Staphylococcus epidermidis* ST59

Zhen Xu[*1,†], Raju Misra[‡], Dorota Jamrozy[§], Gavin K Paterson[**], Ronald R Cutler[††], Mark A Holmes[‡‡], Saheer Gharbia[‡], Hermine V Mkrtchyan[§§1]

[*]Department of Public Health, National Demonstration Center for Experimental Preventive Medicine Education, Tianjin Medical University, Tianjin, China. 300070

[†]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK. E1 4NS$_2$

[‡]Natural History Museum, Core Research Laboratories, Molecular Biology, Cromwell Rd, London SW7 5BD

[§]The Wellcome Trust Sanger Institute, Cambridge, UK. CB10 1SA

[**]School of Veterinary Studies and Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian UK. EH8 9LE

[‡‡]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. CB2 1TN

[§§]Correspondence: Hermine V Mkrtchyan, School of Health, Sport and Biosciences, University of East London, Water Lane, London, UK. E15 4LZ

European Bioinformatics Institute Sequence Read Archive, accession ERR387168.

Running title: WGS analysis of *S. epidermidis* ST59

Key words: *Staphylococcus epidermidis,* antibiotic resistance, whole genome sequence, comparative analysis

Correspondence: Hermine V Mkrtchyan**,** School of Health, Sport and Biosciences, University of East London, Water Lane, London E15 4LZ, UK.

Tel: +44208234081; Email: h.mkrtchyan@uel.ac.uk.

ABSTRACT

Staphylococcus epidermidis is a major opportunistic pathogen primarily recovered from device-associated healthcare associated infections (DA-HAIs). Although *S. epidermidis* and other coagulase-negative staphylococci (CoNS) are less virulent than *Staphylococcus aureus*, these bacteria are an important reservoir of antimicrobial resistance genes and resistance-associated mobile genetic elements that can be transferred between staphylococcal species.

We report a whole genome sequence of a multidrug resistant *S. epidermidis* (strain G6_2) representing multilocus sequence type (ST) 59 and isolated from an environmental sampling of a hotel room in London, UK. The genome of *S. epidermidis* G6_2 comprises of a 2408357 bp chromosome and six plasmids, with an average G+C content of 32%. The strain displayed a multi-drug resistance phenotype which was associated with carriage of 7 antibiotic resistance genes (*blaZ, mecA, msrA, mphC, fosB, aacA-aphD, tetK*) as well as resistance-conferring mutations in *fusA* and *ileS*. Antibiotic resistance genes were located on plasmids and chromosome. Comparative genomic analysis revealed that antibiotic resistance gene composition found in G6_2 was partly preserved across the ST59 lineage.

INTRODUCTION

Staphylococcus epidermidis is a common human skin commensal, but also the most frequent pathogen among coagulase-negative staphylococci (CoNS), causing primarily device-associated healthcare associated infections (DA-HAIs). Compared with more virulent *S. aureus*, CoNS rarely produce toxins and less is known on whether the toxin genes contribute to strain virulence (Otto 2013a). *S. epidermidis* forms biofilms on medical devices and implants, from which single cells dissociate and disseminate via the bloodstream to start colonization at a different site, which might lead to sepsis, meningitis and endocarditis (Becker *et al*. 2014). In addition, *S. epidermidis* and other CoNS are believed to act as a reservoir of resistance and virulence genes for *S. aureus,* contributing to the evolution and emergence of successful clones of methicillin-resistant *S. aureus* (MRSA) (Otto 2013b).

Together with *S. aureus* and other CoNS, *S. epidermidis* accounts for 30% of hospital associated infections (Conlan *et al*. 2012). These nosocomial pathogens have developed an arsenal of strategies contributing to colonisation and infection of the hosts (Becker *et al*. 2014), while often being resistant to multiple antibiotics. Emergence of antibiotic resistant bacteria has been mostly attributed to the healthcare-associated settings (Oliveira and Tomasz 2002). However, more recently, selection of antibiotic resistance has been also associated with the community which has been linked to the misuse of antibiotics (DeLeo et al. 2010). A typical example of this is the community-acquired MRSA (CA-MRSA) which, in addition to acquiring methicillin resistance, has gradually increased the frequency of resistance determinants similarly to hospital-acquired MRSA (HA-MRSA) (Chambers 2005). There is an increasing evidence that horizontal gene transfer between closely related species may contribute to this (Otto 2013a). Recently, Méric et al showed that *S. aureus* and *S. epidermidis* share half of the genome and while homologous recombination between the two species was rare, there was an evidence of extensive MGE sharing, in particular SCC*mec*,

metal resistance and SaPIn1 elements (Méric et al. 2015). As a result, attention is now focusing on the multidrug-resistant coagulase-negative staphylococci and their rapid spread as opportunistic pathogens particularly in relation to patients with an immuno-compromised status (Morfin-Otero et al. 2012). Multidrug-resistant coagulase-negative staphylococci (MDR-CoNS) are primarily recovered from healthcare-associated medical devices, ambulatory patients and healthy animals (Becker *et al*. 2014).

Molecular approaches such as pulse field gel electrophoresis and multi-locus sequence typing have been widely used to evaluate the dissemination of resistant clones of bacteria (Miragaia et al. 2008). Recently, complete genome sequencing of *S. epidermidis* strains have been reported, however these are limited to commensal and nosocomial strains (Conlan *et al*. 2012; Gill et al. 2005; Zhang et al. 2003). Only one study has compared whole genome sequences of four *S. epidermidis* isolated from rice seeds with that of type strain (Chaudhry and Patil 2016). To our knowledge this is the first whole genome based study looking at MDR-CoNS isolated from general public settings.

In this study, we present the genetic features of this multidrug resistant *S. epidermidis* (strain G6_2) and compare it with six *S. epidermidis* reference genomes and 133 previously published genomes of clinical *S. epidermidis*.

MATERIAL AND METHODS

Isolates analysed in this study

Between October 2012 and April 2013, we sampled different sites in three hotels in London, UK. Permission to carry out sampling was granted by the manager/owner of each hotel and the results from each hotel were reported to each manager/owner for their information. Inanimate objects in 32 hotel rooms were sampled using COPAN dry swabs (Copan Diagnostics Inc., USA). All specimens were inoculated onto Nutrient Agar (Oxoid,

120 Basingstoke, UK) and Mannitol Salt Agar plates (Oxoid Basingstoke, UK). These cultures
121 were incubated aerobically at 37°C for 24–72 h.

122 The *S. epidermidis* G6_2 was recovered from one of the hotel rooms in April 2013 in
123 London, UK. Preliminary identification was achieved by using Matrix-assisted laser
124 desorption ionization time-flight mass-spectroscopy (Microflex LT, MALDI-TOF-MS,
125 Bruker Daltonics, Coventry, UK) as described previously (Mkrtchyan et al. 2013).

126 For comparative genomics analysis genomes of six *S. epidermidis* reference strains were
127 included: RP62A (Gill et al. 2005), ASM1192v1), ATCC12228 (Zhang et al. 2003),
128 ASM764v1), SEI (Davenport et al. 2014), CP009046), 949_S8 (Biswas et al. 2015),
129 CP010942), PM221 (Savijoki et al. 2014), HG813242), and BPH 0662 (Jyh et al. 2016),
130 NZ_LT571449) together with 129 *S. epidermidis* genomes derived from two previously
131 published collections (Roach et al. 2015; Tewhey et al. 2014).

132 **16S rRNA gene sequencing**

133 Genomic DNA of *S. epidermidis* G6_2 was prepared using a Qiagen DNA extraction kit
134 (Qiagen, Crawley, UK). 16S rRNA amplification was performed as described previously
135 (Okazaki et al. 2009), PCR products were sequenced by Eurofins MWG GmBH (Ebersberg,
136 Germany) using ABI 3730 L DNA analyser.

137 **Molecular characterization of *S. epidermidis* G6_2**

138 Carriage of the *mecA* gene was determined with PCR as described previously (Hanssen *et al*.
139 2004). SCC*mec* typing was carried out by determination of *mec* and *ccr* complexes (Kondo et
140 al. 2007). Multi locus sequence tying (MLST) has been used to determine seven
141 housekeeping genes as describe previously (Thomas et al. 2006). Sequence types were
142 determined using MLST V1.8 software (https://cge.cbs.dtu.dk/services/MLST/).

Antibiotic susceptibility testing

The antibiotic susceptibility of *S. epidermidis* G6_2 was tested against 13 antibiotics (Mast Group, Merseyside, UK) using disk diffusion methods according to BSAC guidelines (J. M. Andrews and Howe 2011). This included penicillin (1 unit), amoxicillin (10 μg), cefoxitin (10 μg), oxacillin (1 μg), cefepime (30 μg), vancomycin (5 μg), gentamicin (10 μg), streptomycin (10 μg), mupirocin (20 μg), erythromycin (15 μg), tetracycline (10 μg), fusidic acid (10 μg) and chloramphenicol (30 μg). In addition, the minimum inhibitory concentration (MIC) of the isolate to oxacillin was determined using ‘‘M.I.C. evaluators’’ (Oxoid Ltd., Basingstoke,UK).

Whole genome sequencing, assembly and comparative genomics

Genomic DNA was extracted using the MasterPure™ Gram Positive DNA Purification Kit (Cambio, Dry Drayton, UK) from overnight cultures grown from single colonies in 5 ml of tryptic soy broth overnight at 37 ^{o}C. Illumina library preparation was carried out as described previously (Quail *et al*. 2008), and genome sequencing using Hi-Seq 2000 performed following the manufacturer’s standard protocols (Illumina, Little Chesterfield, UK). The raw fastq data was quality trimmed using trimmomatic, (version 0.35) default settings, specifying a phred cutoff of Q20. Read quality was assessed using FastQC (S. Andrews 2014) and Kraken (version 0.10.5-beta) metagenomic pipeline (Wood and Salzberg 2014), including KronaTools (version 2.5) (Ondov *et al*. 2011) was used to assess library purity, that is, it was not a mixed sample and ensure the species was *S. epidermidis*. *De novo* assemblies were performed using assembler, SPAdes (version 3.5.0) (Bankevich *et al*. 2012), default PE settings, from which only contigs greater than 500 bp in length were taken for further analysis. Using the program, Andi (version 0.9.4-beta) (Haubold *et al*. 2015) the *de novo* assembled G6_2 genome along with 108 assembled Staphylococci genomes were aligned, clustered and

visualised using PHYLIP (http://evolution.genetics.washington.edu/phylip.html) and FigTree (http://tree.bio.ed.ac.uk/ software/figtree/). Annotations were performed using the pipeline Prokka (version 1.11) (Seemann 2014). The resultant annotated genome was used for all subsequent comparative genomic studies. Carriage of antimicrobial resistance and virulence genes was assessed using the SRST 2 software (Inouye *et al*. 2014) and the ARG-ANNOT (Gupta *et al*. 2013) and VF-DB databases (Chen *et al*. 2016). Pan-genome analysis was performed using the Roary pipeline (version 3.4.2) (Page *et al*. 2015). To reconstruct phylogenetic tree, short reads were mapped against the *S. epidermidis* ATCC12228 reference genome (Zhang *et al*., 2003), using SMALT version 0.5.8 (http://www.sanger.ac.uk/science/ tools/ smalt-0). A core genome alignment was created after excluding MGE regions, variable sites associated with recombination (detected with Gubbins (Croucher *et al*. 2015) and sites with more than 5% proportion of gaps (i.e. sites with an ambiguous base). A maximum likelihood (ML) phylogenetic tree was generated with RAxML v8.2.8 (Stamatakis 2014) based on generalised time reversible (GTR) model with GAMMA method of correction for among site rate variation and 100 bootstrap (BS) replications. The phylogenetic tree was annotated using Evolview (Zhang *et al*. 2012).

Nucleotide sequence accession numbers

Reads for *S. epidermidis* G6_2 were submitted to the European Bioinformatics Institute Sequence Read Archive, accession ERR387168.

Data Availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article and its tables and figures.

RESULTS AND DISCUSSION

190 *S. epidermidis* has become a leading hospital-associated pathogen due to the increased use of
191 medical devices (Vuong *et al*. 2004). Treatment of *S. epidermidis* infections is challenging as
192 the bacteria are commonly resistant to methicillin and might also display multi-drug
193 resistance phenotype, which presents a serious public health challenge (Xu *et al*. 2015) . *S.*
194 *epidermidis*, represents an important reservoir of mobilizable genes that can be horizontally
195 transferred between staphylococci species, which has likely contributed to the development
196 of antibiotic resistance in *S. aureus* (Otto 2013a).

197 *S. epidermidis* G6_2 was isolated from a hotel room in London, UK in 2013, and the species
198 were determined by MALDI-TOF MS and 16S rRNA sequencing. Initial molecular analysis
199 revealed that the *S. epidermidis* G6_2 strain was *mecA* positive, carrying SCC*mec* type IV,
200 and represented ST59.

201 A draft genome was assembled, comprising of 53 contigs (48 >= 1kb) for the isolated *S.*
202 *epidermidis* G6_2 genome (Table S1; Table S2 and Figure S1). The assembly comprised of
203 one chromosome (2408357 bp in length) and six plasmids, annotated as pG6_2_1 to
204 pG6_2_6 (the largest, pG6_2_1, is 10570 and the smallest, pG6_2_6, is 3426 bp in length),
205 with an average G+C content of 32.02%. It has a total (chromosome and plasmids) of 2213
206 predicted protein coding sequences, of which 21.5% were annotated as hypothetical proteins
207 and 14.3% were annotated as putative functions (Table 1).

208 **Phylogenetic relationship with other *S. epidermidis* Isolates**

209 A previously described collection of 129 whole genome-sequenced *S. epidermidis* isolates
210 together with 6 reference strains was used to determine the phylogenetic relationship between
211 the G6_2 strain and other *S. epidermidis lineages*. After removal of variable sequence regions
212 corresponsing to mobile genetic elements (MGE), recombination blocks as well as sites with

more than 5% proportion of gaps, the core genome alignment contained 4262 SNP sites. Seven ST59 isolates clustered and formed a distinct clade with *S. epidermidis* G6_2 (Fig. 1).

Genotypic and phenotypic characterisation of antibiotic resistance

S. epidermidis G6_2, revealed 9 antibiotic resistance determinates across the chromosome and plasmids (Table 2). This included aminoglycoside resistance gene *aac(6') – aph(2")*, beta-lactam resistance genes *mecA* and *blaZ*, fosfomycin resistance gene *fosB*, macrolide resistance genes *mphH* and *msrA* (the latter also conferring resistance to lincosamide and streptogramin B) and tetracycline resistance gene *tet(K)*. This correlated with the results of antibiotic susceptibility testing as the strain was found resistant to 11 out of 13 antibiotics tested, demonstrating susceptibility to vancomycin and chloramphenicol only. Resistance to mupirocin and fusidic acid was associated with point mutations in chromosomally located genes, *ileS* and *fusA*, respectively. In addition to antimicrobial resistance genes, the G6_2 strain also carried plasmid-associated *qacC* gene, which encodes the multidrug resistance efflux protein and mediates resistance to biocides, and a chromosomally-inserted copper resistance operon composed of *copZ-copA-csoR* genes together with an additional copy of cobalt-zinc-cadium efflux pump gene *czcD*. The latter was distinct from the conserved chromosomal copy of *czcD* gene, and was previously identified on a number of CoNS plasmids.

The G6_2 strain carried a 47-kb composite island composed of the SCC*mec* IV and a SCC element that contained plasmin-sensitive surface protein gene *pls*, spermidine N-acetyltransferase gene *speG* and a copper-translocating ATPase gene *copA*. The full sequence of this composite island was unique and did not match previously described reference genomes. However, the SCC*mec* IV sequence shared 99% identity with SCC*mec* IVa from various MRSA strains including the MRSA M1 isolated in Denmark (Larner-Svensson *et al*.

2013). The SCC element matched most closely the MRSA UCI62 strain representing ST5 (GenBank: CP018766). Carriage of *blaZ*, *tetK* and *qacC* genes was associated with plasmid sequences whereas other genes were inserted chromosomally. Elements carrying *tetK* and *qacC* matched previously reported *S. aureus* plasmids. Méric *et al* showed that hospital associated *S. aureus* and *S. epidermidis* share genes involved in pathogenicity, metal toxicity resistance and antibiotic resistance. In addition they have demonstrated that high levels of recombination of genes that might be successful in healthcare settings contribute to proliferation of subpopulations of two species (Méric *et al*. 2015).

Comparison of resistance determinant distribution revealed that the *S. epidermidis* G6_2 strain shared a common antibiotic resistance gene composition with other ST59 isolates, suggesting that the particular combination of antibiotic resistance genes found in the G6_2 strain is preserved across the ST59 lineage (Fig. 1). All ST59 isolates harboured *aac-aph*, *blaZ* and *mecA* genes, and majority contained *mphC* and *msrA* genes, whereas *tetK* was uniquely found in *S. epidermidis* G6_2. The G6_2 strain also shared the *qacC* plasmid with other ST59 isolates as well as the SCC*mec* IV sequence but not full SCC*mec*-SCC composite island, which was not detected in any other analysed *S. epidermidis* genome.

Functional genes uniquely found in *S. epidermidis* G6_2 compared with reference strains

Pan-genome analysis of the G6_2 strain and six *S. epidermidis* reference genomes revealed that 78 genes were unique to G6_2. After excluding genes found on plasmids, 64 chromosomally located genes were unique to G6_2 strain. This included a number of SCC*mec*- and SCC-associated genes as well as some of the chromosomally inserted resistance genes such as *mphC*, *msrA*, *copZ-copA-csoR* operon and the additional copy of *czcD* genes.

Comparative analysis of virulence genes

Pathogenicity of *S. epidermidis* has been linked primarily with its capacity for biofilm formation. Biofilm formation occurs by initial attachment of bacteria on both biotic and abiotic surfaces, which further accumulates into multi-layered cell agglomerates. This facilitates the internalization and persistence of *S. epidermidis* species in the host cells. Strains that facilitate this feature are therefore considered more virulent (Becker *et al*. 2014). *S. epidermidis* carries a number of virulence determinants that have been associated with its ability to attach to biotic and abiotic surfaces as well as the various phases of biofilm formation. Analysis of virulence gene composition based on the VF database, revealed a number of such virulence determinants that were detected in all or majority of analysed *S. epidermidis* isolates, including the G6_2 strain. This included the autolysin gene *atlE* (138/140), the cell wall associated fibronectin binding protein gene *ebh* (140/140), the elastin binding protein gene *ebp* (135/140), the fibrinogen binding protein genes *sdrG* (137/140) and *sdrH* (138/140), serine protease genes *sspA* (138/140) and *sspB* (138/140), lipase genes *geh* (139/140) and *lip* (138/140), and the nuclease gene *nuc* (138/140). The intercellular adhesion operon *icaADBC*, which is also associated with biofilm formation (Cramton *et al*. 1999), was variably distributed (87/140) and absent in the *S. epidermidis* G6_2 strain as well as the other ST59 isolates included in this analysis. This is in agreement with previous reports of clinical *S. epidermidis* ST59 isolates that revealed a biofilm negative phenotype (Li *et al*. 2009; Mendes *et al*. 2012; Miragaia *et al*. 2007).

In addition to the described biofilm formation-associated virulence determinants, majority of *S. epidermidis* isolates carried the hemolysin-beta gene *hlb* (136/140), which was also present in the G6_2 strain. Less common was the delta hemolysin gene *hld* (41/140), also detected in the G6_2 strain although absent in most other ST59 isolates.

In conclusion, this study is the first analysis of the genome of *S. epidermidis* isolated from the general public environment and harbouring a cassette of resistance genes to an array of antimicrobials. The comparison of *S. epidermidis* G6_2 genome with clinical reference strains revealed its antibiotic resistance and virulence gene arsenal. Resistance genes were carried on both bacterial chromosome and plasmids. We established that *S. epidermidis* G6_2 harbours 12 virulence genes, and delta hemolysin gene *hld* (41/140) is known to be detected in the G6_2 strain but absent in most other ST59 isolates. In addition, 9 antibiotic resistance determinants which are responsible for the resistance to 12 antibiotics, including streptomycin, gentamicin, penicillin, oxacillin, amoxicillin, cefoxitin, cefepime, erythromycin, fosfomycin, tetracycline, fusidic acid, mupirocin, have been identified in *S. epidermidis* G6_2. Additional whole genome sequence and comparative genomics analysis are warranted to further our understanding of the origin and evaluation of multidrug resistant isolates from different ecological niches.

ACKNOWLEDGEMENTS

Genome sequencing was supported by Medical Research Council (MRC) partnership grant G1001787/1.

References

306 Andrews, J. M., and R. A. Howe. 2011 BSAC standardized disc susceptibility testing method
307 (Version 10). J Antimicrob Chemoth. 66 (12):2726–57.

308 Andrews, S. 2011 FastQC a quality control tool for high throughput sequence data.
309 Cambridge, UK: Babraham Institute, 175–76.

310 Bankevich, A, S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin *et al*. 2012 SPAdes: a new
311 genome assembly algorithm and its applications to single-cell sequencing. J
312 Computational Biol. 19 (5): 455–77.

313 Becker, K, C. Heilmann, and G. Peters. 2014 Coagulase-negative staphylococci. Clin
314 Microbiol Rev. 27 (4):870–926.

315 Biswas, R. K., M. M. Kock, T. Adelowotan, W. Strasheim, T. G. Mahomed *et al*., 2015
316 Draft genome sequences of five clinical methicillin-resistant *Staphylococcus aureus*
317 isolates and a methicillin-resistant *Staphylococcus epidermidis* isolate. Genome
318 Announcements. 3 (4):e00836-15.

319 Chambers, H. F., 2005 Community-associated MRSA--resistance and virulence converge.
320 New Engl J Med. 352 (14):1485–87.

321 Chaudhry, V., and P. B Patil, 2016 Genomic investigation reveals evolution and lifestyle
322 adaptation of endophytic *Staphylococcus epidermidis*. Sci Rep. 6:19263.

323 Chen, Lihong, Dandan Zheng, Liu Bo, Yang Jian, and Jin Qi. 2016 VFDB 2016: hierarchical
324 and refined dataset for big data analysis—10 years on. Nucleic Acids Res. 44:D694–97.

325 Conlan, S., H. H. Kong, and J. A. Segre, 2012 Species-level analysis of DNA sequence data
326 from the NIH human microbiome project. PloS One. 7 (10):e47075.

Cramton, S. E., C. Gerke, N. F. Schnell, W. W. Nichols, and F. Gotz, 1999 The intercellular adhesion (Ica) locus is present in *Staphylococcus aureus* and is required for biofilm formation. Infect. Immun. 67 (10):5427–33.

Croucher, N. J., A. J. Page, T. R. Connor, A. J. Delaney, J. A. Keane *et al*., 2015 Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using gubbins. Nucleic Acids Res. 43 (3):e15.

Davenport, K. W., H. E. Daligault, T. D. Minogue, K. A. Bishoplilly, S. M. Broomall et al., 2014 Complete genome assembly of *Staphylococcus epidermidis* AmMS 205. Genome Announcements. 2 (6):e01059-14.

DeLeo, F. R., M. Otto, B. N. Kreiswirth, and H. F. Chambers, 2010 Community-associated meticillin-resistant *Staphylococcus aureus*. Lancet. 375 (9725):1557–68.

Gill, S. R., D. E. Fouts, G. L. Archer, E. F. Mongodin, R. T. Deboy *et al*., 2005 Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J Bacter. 187 (7):2426–38.

Gupta, S. K., B. R. Padmanabhan, S. M. Diene, R. Lopez-Rojas, M. Kempf *et al*., 2013 ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation), a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Ch. 58 (1):212–20.

Hanssen, A., G. Kjeldsen, and J. U. Ericson Sollid, 2004 Local variants of staphylococcal cassette chromosome *mec* in sporadic methicillin-resistant *Staphylococcus aureus* and methicillin-resistant coagulase-negative staphylococci: evidence of horizontal gene transfer? Antimicrob Agents Ch. 48 (1):285–96.

Haubold, B., F Klötzl, and P Pfaffelhuber, 2015 Andi: fast and accurate estimation of evolutionary distances between closely related genomes. Bioinformatics. 31 (8):1169-75.

Inouye, M., H. Dashnow, L. A. Raven, M. B. Schultz, B. J. Pope *et al*., 2014 SRST2: rapid genomic surveillance for public health and hospital microbiology labs. Genome Med. 6 (11):90.

Jyh, L., I R Monk, S J Pidot, S Singh, C. Kyl *et al*., 2016 Functional analysis of the first complete genome sequence of a multidrug resistant sequence Type 2 *Staphylococcus epidermidis*. Microbial Genomics. 2 (9): e000077.

Kondo, Y., T. Ito, X. X. Ma, S. Watanabe, B. N. Kreiswirth et al., 2007 Combination of multiplex PCRs for staphylococcal cassette chromosome *mec* type assignment: rapid identification system for *mec*, *ccr*, and major differences in junkyard regions. Antimicrob Agents Ch. 51 (1):264–74.

Larnersvensson, H., P. Worning, M. D. Bartels, H. L. Hestbjerg, K. Boye et al., 2013 Complete genome sequence of *Staphylococcus aureus* strain M1, a unique t024-ST8-IVa danish methicillin-resistant *S. aureus* clone. Genome Announcements. 1 (3):S69-82.

Li, M., X. Wang, Q. Gao, and Y. Lu. 2009 Molecular characterization of *Staphylococcus epidermidis* strains isolated from a teaching hospital in Shanghai, China. J Med Microbiol 58 (4):456–61.

Mendes, R. E., L. M. Deshpande, A. J. Costello, and D. J. Farrell, 2012 Molecular epidemiology of *Staphylococcus epidermidis* clinical isolates from U.S. hospitals. Antimicrob Agents Ch. 56 (9):4656–61.

372 Méric, G., M. Miragaia, M. de Been, K. Yahara, B. Pascoe *et al*., 2015 Ecological overlap
373 and horizontal gene transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*.
374 Genome Biol and Evol. 7 (5):1313–28.

375 Miragaia, M, J. A. Carriço, J. C. Thomas, I. Couto, M. C. Enright *et al*., 2008 Comparison of
376 molecular typing methods for characterization of *Staphylococcus epidermidis*: proposal
377 for clone definition. J Clin Microbiol. 46 (1):118–29.

378 Miragaia, M., J C Thomas, I Couto, M C Enright, and H de Lencastre, 2007 Inferring a
379 population structure for *Staphylococcus epidermidis* from multilocus sequence typing
380 Data. J Bacteriol. 189 (6):2540–52.

381 Mkrtchyan, H. V., C. A. Russell, N. Wang, and R. R. Cutler. 2013 Could public restrooms be
382 an environment for bacterial resistomes? PLoS ONE. 8 (1):e54223.

383 Morfin-Otero, R., M. A. Martínez-Vázquez, D. López, E. Rodríguez-Noriega, and E. Garza-
384 González, 2012 Isolation of rare coagulase-negative isolates in immunocompromised
385 patients: *Staphylococcus gallinarum*, *Staphylococcus pettenkoferi* and *Staphylococcus*
386 *pasteuri*. Annals of Clinical & Laboratory Science. 42 (2):182–85.

387 Okazaki, M., K. Ohkusu, H. Hata, H. Ohnishi, K. Sugahara *et al*., 2009 *Mycobacterium*
388 *Kyorinense* sp. nov., a novel, slow-growing species, related to mycobacterium celatum,
389 isolated from human clinical specimens. IJSEM. 59 (6):1336–41.

390 Oliveira, D. C, and A. L. H Tomasz. 2002 Secrets of success of a human pathogen: molecular
391 evolution of pandemic clones of meticillin-resistant *Staphylococcus aureus*. Lancet
392 Infect Dis. 2 (3):180–89.

393 Ondov, B. D., N. H. Bergman, and A. M. Phillippy, 2011 Interactive metagenomic
394 visualization in a web browser. Bmc Bioinformatics. 12 (1):1–10.

Otto, M., 2013a Coagulase-negative staphylococci as reservoirs of genes facilitating MRSA infection: staphylococcal commensal species such as *Staphylococcus epidermidis* are being recognized as important sources of genes promoting MRSA colonization and virulence. BioEssays. 35 (1):4–11.

Otto, M., 2013b Staphylococcal infections: mechanisms of biofilm maturation and detachment as critical determinants of pathogenicity. Annu Rev of Med. 64:175–88.

Page, A. J., C. A. Cummins., M. Hunt, V. K. Wong, S. Reuter *et al*., 2015 Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31 (22):3691–93.

Quail, M. A., I. Kozarewa, F. Smith, A. Scally, P. J. Stephens *et al*., 2008 A large genome center's improvements to the illumina sequencing system. Nat Methods. 5 (12):1005–10.

Roach, D. J., J. N. Burton, C. Lee, B. Stackhouse, S. M Butler-Wu *et al*., 2015 A Year of Infection in the Intensive Care Unit: Prospective Whole Genome Sequencing of Bacterial Clinical Isolates Reveals Cryptic Transmissions and Novel Microbiota. PLoS Genet. 11 (7):e1005413.

Savijoki, K., A. Iivanainen, P. Siljamäki, P. K. Laine, L. Paulin *et al*., 2014 Genomics and proteomics provide new insight into the commensal and pathogenic lifestyles of bovine- and human-associated *Staphylococcus epidermidis* strains. J of Proteome Res. 13. (8):3748-3762.

Seemann, T., 2014 Prokka: rapid prokaryotic genome annotation. Bioinformatics. 30 (14):2068-9.

Stamatakis, A., 2014 RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30 (9):1312–13.

418 Tewhey, R., B. Gu, T. Kelesidis, C. Charlton, A. Bobenchik et al., 2014 Mechanisms of
419 linezolid resistance among coagulase-negative Staphylococci determined by whole-
420 genome sequencing. mBio. 5 (3):e00894-14.

421 Thomas, J. C., M. R. Vargas, M. Miragaia, S. J. Peacock, G. L. Archer et al., 2006 Improved
422 multilocus sequence typing scheme for *Staphylococcus epidermidis*. J Clin Microl. 45
423 (2):616–19.

424 Vuong, C., S. Kocianova, Y. Yao, A. B. Carmody, and M. Otto, 2004 Increased colonization
425 of indwelling medical devices by quorum-sensing mutants of *Staphylococcus*
426 *epidermidis* in vivo. J infect dis. 190 (8):1498–1505.

427 Wood, D. E., and S. L. Salzberg, 2014 Kraken: ultrafast metagenomic sequence classification
428 using exact alignments. Genome Biol. 15 (3):R46.

429 Xu, Z., V. M. Hermine , and R. R. Cutler. 2015 Antibiotic resistance and *mecA*
430 characterization of coagulase-negative staphylococci isolated from three hotels in
431 London, UK. FrontMicrobiol. 9 (6):947..

432 Zhang, H., S. Gao, M. J. Lercher, S. Hu, and W. Chen. 2012 EvolView, an online tool for
433 visualizing, annotating and managing phylogenetic trees. Nucleic Acids Res. 40:569–72.

434 Zhang, Y., S. Ren, H. Li, Y. Wang, G Fu *et al*., 2003 Genome-based analysis of virulence
435 genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). Mol
436 Microbiology. 49 (6):1577–93.

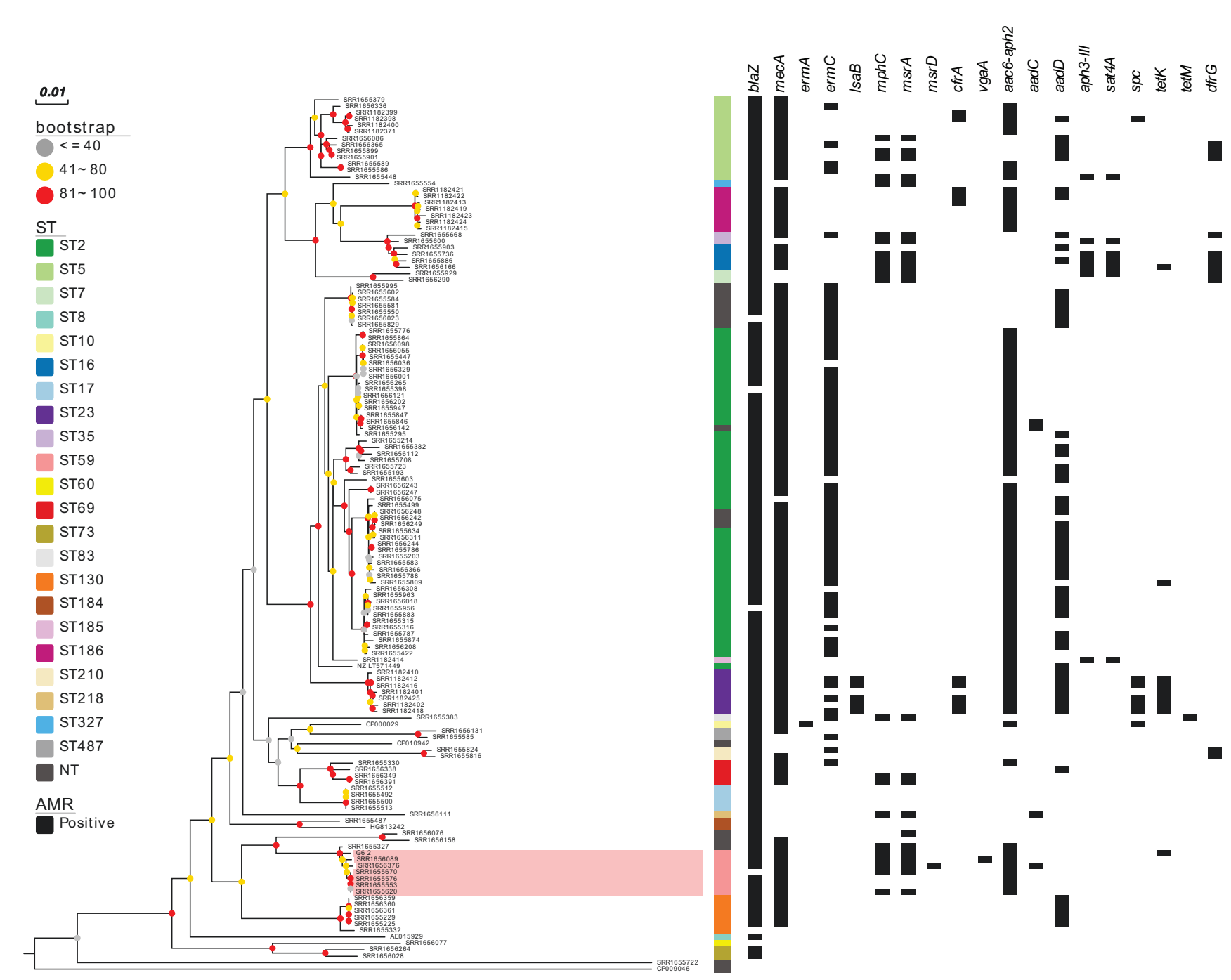
437

438

439 Figure legend

440 **Figure 1. Core-genome mid-point rooted phylogenetic tree of 136 *S. epidermidis* isolates.** The tree nodes are annotated with bootstrap value ranges based on 100
441 replicates. The tree is also annotated with the sequence type (ST) assignment and antimicrobial resistance gene (AMR) carriage. Gene names above the annotation are
442 grouped in accordance with the corresponding antimicrobial class (beta-lactams: *blaZ*, *mecA*; macrolides, lincosamides and streptogramines: *ermA*, *ermC*, *lsaB*, *mphC*, *msrA*,
443 *msrD*, *cfrA* and *vgaA*; aminoglycosides: *aac6-aph2*, *aadC*, *aadD*, *aph3*-III, *sat4A* and *spc*; tetracyclines: *tetK* and *tetM*; trimpethoprim: *dfrG)*. The ST59 cluster that contains
444 the G6_2 strain is highlighted in pink.

445


0.01
bootstrap
< = 40
41 ~ 80
81 ~ 100
ST
ST2
ST5
ST7
ST8
ST10
ST16
ST17
ST23
ST35
ST59
ST60
ST69
ST73
ST83
ST130
ST184
ST185
ST186
ST210
ST218
ST327
ST487
NT
AMR
Positive
blaZ
mecA
ermA
ermC
lsaB
mphC
msrA
msrD
cfrA
vgaA
aac6-aph2
aadC
aadD
aph3-III
sat4A
spc
tetK
tetM
dfrG

1 # TABLES

2 **Table 1 Comparative general features of *S. epidermidis* G6_2 and the reference strains.**

	RP62a	ATCC 12228	SEI	949_S8	PM221	BPH 0662	G6_2
Chromosome[a]							
Length of sequences (bp)	2616530	2499279	2538314	2339868	2490012	2793003	2408357
G+C content	32.10%	32.10%	32.10%	32.00	32.10%	32.00%	32.02%
Protein coding region	2391	2419	2504	2119	2399	2699	2213
Ribosomal RNAs							4
16S	6	5	6	-[c]	6	5	1
23S	6	5	6	-[c]	6	5	1
5S	7	6	7	5	7	6	2
Transfer RNAs	59	60	58	56	59	59	60
Plasmids[b]							
Length of sequences (bp)	P1:27310	P1:4439	P1:37688	-[c]	P1:4439	P1:45804	P1：10570
		P2:4679			P2:11152	P2:2366	P2：4909
		P3:8007			P3:33094		P3：4588
		P4:17261			P4:58811		P4：4576

	P5:24370	P5：4271
	P6:6585	P6：3426

a Chromosome section includes: the length of the chromosome, G+C content of the chromosome, protein coding region, ribosomal RNA and transfer RNAs numbers.
b Plasmids section includes: the length of each plasmid and the number of plasmids. P - Plasmid. Numbers - the number of plasmids.
c '-' No data available in Genbank file. Draft assembly.

Table 2 Genotypic and phenotypic characterisation of antibiotic resistance in *S. epidermidis* G6_2

Product	Gene name	Accession number	Location	Function	Class of antibiotic	Antibiotics

		(Identity %)				
Aminoglycoside-modifying enzymes	*aac(6')-aph(2'')*	M13771 (100)	plasmid	Aminoglycoside resistance	Aminoglycoside	Gentamycin streptomycin
β-lactamase	*blaZ*	AJ302698 (100)	plasmid	Beta-lactam resistance		Penicillin oxacillin
Penicillin-binding protein 2a	*mecA*	AB505628 (100)	Chromosome	Beta-lactam resistance	Beta-lactam	Amoxcillin cefepime cefoxitin
Fosfomycin resistance protein	*fosA*	ACHE01000077 (100)	Chromosome	Fosfomycin resistance	Phosphonic	Fosfomycin
Macrophage scavenger receptors	*msr(A)*	X52085 (98.98)	plasmid	Macrolide, Lincosamide and Streptogramin B	Microlide	Erythromycin

				resistance		
Inactivating enzymes	*mph(C)*	AF167161 (100)	plasmid	Macrolide resistance		
Tetracycline efflux pump	*tet(K)*	U38428 (99.93)	plasmid	Tetracycline resistance	Tetracycline	Tetracycline
Isoleucyl RNA synthetase	*ileS*	-	-	Fusidic acid resistance	Fusidic acid	Fusidic acid
Elongation factor G	*fusA*	-	-	Monoxycarbolic resistance	Monoxycarbolic	Mupirocin