A comparative genomics study of *Staphylococcus epidermidis* from orthopedic device-related infections correlated with patient outcome.

Virginia Post[1], Llinos G. Harris[2], Mario Morgenstern[3,4], Leonardos Mageiros[2], Matthew D. Hitchings[2], Guillaume Méric[5], Ben Pascoe[5], Samuel K. Sheppard[5], R. Geoff Richards[1], T. Fintan Moriarty[1]

[1]AO Research Institute Davos, Davos, Switzerland.

[2]Microbiology and Infectious Diseases, Institute of Life Science, Swansea University Medical School, Swansea, UK.

[3]Department of Orthopedic and Trauma Surgery, University Hospital Basel, Basel, Switzerland.

[4]Department of Trauma Surgery, Trauma Centre Murnau, Murnau, Germany.

[5]The Milner Centre for Evolution, Department of Biology & Biochemistry, University of Bath, Bath, UK.

Running Title: *Staphylococcus epidermidis* infection and patient outcome.

Correspondence: Fintan Moriarty, AO Research Institute Davos, Clavadelerstrasse 8, 7270 Davos Platz, CH 7270 Switzerland.

Phone: +41 81 414 2397; E-mail: fintan.moriarty@aofoundation.org

Abstract

Staphylococcus epidermidis has emerged as an important opportunistic pathogen causing orthopedic device-related infections (ODRIs). This study investigated the association of genome variation and phenotypic features of the infecting *S. epidermidis* isolate with the clinical outcome of the infected patient. *S. epidermidis* isolates were collected from 104 patients with ODRI. Their clinical outcome was evaluated, after an average of 26 months, as either "cured" or "not cured". The isolates were tested for antibiotic susceptibility and biofilm formation. Whole genome sequencing was performed on all isolates and genomic variation was related to features associated with "cured" and "not cured". Strong biofilm formation and aminoglycoside resistance were associated with a "not cured" outcome ($p = 0.031$ and $p < 0.001$, respectively). Based on gene-by-gene analysis, some accessory genes were more prevalent in isolates from the “not cured” group. These included: the biofilm-associated *bhp* gene; the antiseptic resistance *qacA* gene, the cassette chromosome recombinase encoding genes *ccrA* and *ccrB* and IS*256*-like transposase. This study identifies biofilm formation and antibiotic resistance as associated with poor outcome in *S. epidermidis* ODRI. Whole genome sequencing identified specific genes associated with a "not cured" outcome that should be validated in future studies.

Keywords. *Staphylococcus epidermidis;* MRSE; virulence factors; antibiotic resistance; genotype; phenotype; orthopedic device-related infections.

Introduction

Staphylococcus epidermidis is a common member of the human skin microflora, predominant in moist sites such as nares or fossae, and sebaceous areas such as the facial skin. With the advent of implanted and indwelling medical devices, *S. epidermidis* has emerged as a

prominent cause of nosocomial and device-associated infections (1, 2). The microorganism's ability to switch from a commensal to pathogenic lifestyle is facilitated by its ability to rapidly attach to, and form biofilms upon, medical devices. In the case of orthopedic device-related infections (ODRI), *S. epidermidis* accounts for up to 43% of cases and is second only to *S. aureus* as the most prevalent causative organism (1, 2).

Molecular epidemiological studies have begun to reveal information on both population structure and genetic diversity within *S. epidermidis* populations (3-5). The complete *S. epidermidis* genome is estimated at approximately 2.5 Mb and comprises 80% core genes and 20% variable genes (3, 4, 6). Three distinct phylogenetic groups (clades) are evident in the population structure of *S. epidermidis* (3, 4, 6), with at least nine globally disseminated clonal complex lineages. The most common clonal complex (CC2) contains one particularly prominent sequence type: ST2 (32% of all isolates) (5, 7).

In an attempt to identify the features that enable invasive infection in *S. epidermidis,* a number of studies have searched for features that may distinguish invasive from commensal *S. epidermidis* isolates on a genotypic and phenotypic level. Such studies have identified features such as IS*256,* folate dehydrogenase, copper remediation genes to be more common amongst invasive isolates (3, 6, 8, 9). However, clear separation between the two has proven difficult (9-11), perhaps indicating that the ability to invade the host and the ability to colonize it do not require significantly different genetically-encoded features. One possibility, that has not been explored to date however, is whether the genome/phenotype of the invasive isolates dictates the ultimate course of an infection i.e. whether the patient eventually has a successful treatment outcome, or a failed treatment outcome.

In the present study, *S. epidermidis* isolates were prospectively collected from patients with ODRI and were assigned a clinical outcome (either "cured" or "not cured") after an extended patient follow-up (FUP). Clinical outcome was then related to genome variation and phenotypes believed to be important for *S. epidermidis* virulence.

Results

Patient outcome and clinical parameters

A total of 104 patients with *S. epidermidis* ODRI were included in this study, with complete demographic information shown in Table 1. The lower extremity cohort (70 patients) included only those patients with infection of the hip, knee and upper ankle joints as well as femur, tibia and fibula. The majority of patients of the complete cohort study (n = 85, 81.7%) were considered to have had a "cured" clinical outcome at FUP.

Those considered to have a "not cured" clinical outcome at FUP were statistically more likely to have had multiple revision surgeries in comparison with "cured" outcome isolates ($p < 0.067$) (Table 2). There was no association between outcome and any of the other monitored parameters such as diabetes, chronic immunosuppression or obesity (Table 2).

Patient outcome and phenotypic properties of isolates

Antibiotic susceptibility. Antibiotic susceptibility testing of the 104 *S. epidermidis* isolates found 74% (77/104) to be multiple resistant isolates and 67.3% (70/104) were resistant to methicillin (Table 1). Rifampicin resistance was also observed in 19.2% (20/104) of the isolates, which is notable due to the critical role of this antibiotic in treating ODRI. Resistance to aminoglycosides had a statistically significant influence on a "not cured" clinical outcome ($p = 0.001$; Table 3). Further antibiotic resistance (including resistance to aminoglycosides) had no statistically significant influence on any of the other prognostic parameters such as chronic or acute ODRI. Although isolates from the group of chronic ODRIs were more often resistant to aminoglycosides than isolates from the acute ODRI group (42.7% of versus 31%), this was not statistically significant ($p = 0.276$). Furthermore, multidrug-resistance also

showed no statistically significant difference between chronic and acute ODRIs (73.3% versus 75.9%).

Biofilm formation. As shown in Table 1, 70.2% (73/104) of the isolates formed a biofilm *in vitro*. The ability to form biofilm was subdivided into weak biofilm-formers (37.5% (39/104) of the isolates), intermediate biofilm-formers (21.2%, 22/104) and strong biofilm-formers (11.5%, 12/104). The remaining isolates (29.8%, 31/104) were unable to form a biofilm under our *in vitro* conditions. A statistically significant association between biofilm forming ability and clinical "cured" versus "not cured" outcome was noted for the lower extremity cohort (p = 0.031; Table 3). Strong biofilm forming ability also resulted in the highest percentage of "not cured" outcome for the complete cohort (33.3%, p = 0.059; Table 3). However, the strength of biofilm formation had no statistically significant influence on any of the prognostic variables such as multiple revision surgery.

A description of biofilm-associated genes and relative presence with respect to *in vitro* biofilm forming ability is shown in Table 4. Among strong biofilm forming isolates the intercellular adhesion (*icaA*) gene was more prevalent (83.3%) than the accumulation-associated protein encoding *aap* (8.3%), the *bhp* (cell wall associated biofilm protein) (16.7%) and the *embp* (extracellular matrix-binding protein) (66.7%).

Patient outcome and pathogen genome variation

Relationship between virulence associated genes and patient outcome. The 104 genomes were analyzed for the presence of a selection of genes previously described as virulence factors in *S. epidermidis* (Fig 1) (12-14). Within the population as a whole (i.e. "cured" and "not cured"-outcome isolates), *aae (*vitronectin), *gehC* (lipase), *gehD* (lipase), *hlb* (β-hemolysin), *sesB (S. epidermidis* surface protein) and *sesC (S. epidermidis* surface protein)

were present in all 104 isolates. The methicillin resistance gene, *mecA* was carried by 68.3% of the isolates, whereby 69/70 of the phenotypically confirmed MRSE and 2/34 phenotypically confirmed MSSE possessed this gene. Fig 1 also shows the distribution of the known virulence genes between the 2 clinical outcome groups ("cured" and "not cured"). A trend for the presence of aminoglycoside resistance gene, *aacA* (*aac*(6`)-*aph*(2")) and *mecA* on "not cured" outcome ($p = 0.076$ and $p = 0.099$, respectively) was observed. In addition, the presence of biofilm-associated *bhp* was statistically significantly associated with a "not cured" clinical outcome in the lower extremity cohort ($p = 0.023$).

Accessory gene regulator (*agr*)-types. Overall *agr*-type I was the most prevalent type (38.5%, 40/104) among the isolates, followed by *agr*-type III (36.5%, 38/104). The distribution of the 3 different *agr*-types within the "cured" and "not cured" outcome groups are shown in Table 5, however there were no statistically significant differences ($p \geq 0.05$). The only parameter associated with *agr*-type was acute infection (Table 5). All other clinical parameters were not statistically associated with *agr*-type.

Multi-locus sequence typing (MLST). Within the 104 isolates, 21 different Sequence Types (STs) were identified based on the 7 loci scheme for *S. epidermidis* (15) using the build-in MLST function of BIGSdb, linked with pubMLST databases (4) (Table 6). Thirty isolates could not be assigned to any known ST. While the majority of the ST2 (13/18; 72.2%) and ST5 (16/18; 88.9%) isolates were associated with "cured" outcome, all ST57 (2/2; 100%), ST89 (1/1; 100%) and S110 (1/1; 100%) were associated with a "not cured" outcome (Table 6). However, more isolates would be needed to draw statistical conclusions. The identified STs belonged to previously described 7-locus MLST clonal complexes (CC) of which the largest was CC2 (65/104; 62.5%) (4) (Table 6).

Patient outcome and accessory genomes. Further evaluation of the relative presence of accessory genes that were more prevalent in the "not cured" outcome group than in the "cured" outcome group is shown in Table 7. *S. epidermidis* isolates from the "not cured" outcome group carried the antiseptic resistance coding gene *qacA* at a statistically significant higher percentage than isolates from the "cured" outcome group (89.5% vs 27.1%; $p = 0.023$). Furthermore, the presence of the cassette chromosome recombinase encoding genes *ccrA* and *ccrB* (89.5% vs. 23.6%and 89.5% vs. 24.7%; $p = 0.042$ and $p = 0.034$, respectively) was significantly associated with the "not cured" isolate genomes.

Core and accessory genome analysis. A pan-genome of all study isolates was used to compare the genomes of the 104 clinical *S. epidermidis* isolates. ClonalFrame was used to construct ancestral genealogies, free from recombination. In order to run ClonalFrame a stringent approach to select core genes based on presence in 100 % of the 104 isolates was applied. This resulted in a reduced core genome consisting of 123 non-truncated genes. *S. epidermidis* isolates clustered into 3 clades (Fig 2), with 86% of isolates (89/104) in clade A, 9.6% (10/104 isolates) in clade B and 4.8% (5/104 isolates) in clade C (Fig 2A).

Comparing patient outcome between the clades, a trend was observed between clade A and B. Clade B consisted of a comparatively higher percentage of "not cured" outcome isolates (40%, 4/10) than clade A carrying 15.7% (14/89, Fig 2A+C). However, this trend did not quite reach statistical significance ($p = 0.08$; Fisher`s exact test). Only 1/5 (20%) Clade C isolates belonged to the "not cured" outcome group, although the low numbers of isolates precluded reliable statistical analysis (Fig 2B+D). Furthermore, clade B isolates also contained a higher percentage of moderate/strong biofilm forming isolates than clade A

isolates (40% versus 31.5%) (Fig 3A+C). In addition, the majority of clade B isolates (80%) belonged to CC2 while clade A possessed only 60.7% CC2 isolates (data not shown).

Clinical outcome, biofilm formation and antibiotic resistance phenotypes were homogeneous compared to the clonal frame ($p \geq 0.05$). In addition, the permutation test revealed a strong association between lineage and biofilm formation ($p \leq 0.0001$), resistance to methicillin ($p = 0.0002$), quinolones ($p = 0.0055$), erythromycin ($p < 0.00001$), clindamycin ($p < 0.00001$), tetracycline ($p < 0.00001$), trimethoprim/sulfonamide ($p = 0.02$), and fusidic acid ($p < 0.00001$). However, there was no association between lineage and outcome ($p = 0.09$) or resistance to penicillin ($p = 1$), aminoglycosides ($p = 0.3798$), fosfomycin ($p = 0.053$), and rifampicin ($p = 0.151$).

Discussion

S. epidermidis is a commensal microorganism that is also a frequent agent of ODRI (6, 8, 16, 17). However, little is known about the impact that genotypic and phenotypic features of the infecting pathogen can have on treatment outcome (3, 6, 8, 9). This prospective study was designed to test the hypothesis that treatment outcome in patients with *S. epidermidis* ODRI may be influenced by phenotypic or genotypic features of the infecting pathogen. Against a background of scientific studies searching for features that distinguish commensal from invasive isolates (3, 6, 8-11), or for host factors that have an influence on patient outcome (8, 18), this study advances this line of investigation by looking for bacteria-retained features that distinguish infections that result in poor treatment outcome. After prospectively collecting 104 patients, with an average 2-year follow-up (FUP), and subjecting infecting pathogens to genome sequencing and a number of phenotypic assays, we have identified a number of features associated with poor treatment outcome. Those features include biofilm formation,

aminoglycoside resistance, the cassette chromosome recombinase encoding genes *ccrA* and *ccrB*, IS*256-like* and plasmid-borne *qacA* gene, as well as the biofilm-associated *bhp* gene.

Adhesion to and biofilm formation upon biomaterial substrates are widely believed to be the primary virulence factor enabling invasive *S. epidermidis* ODRI (6, 8, 16, 17, 19). The data from our study supports this by revealing that a "not cured" clinical outcome was significantly associated with an increased ability to form biofilm *in vitro* ($p = 0.031$). Genomic analysis on the known biofilm-associated genes such as *icaA*, *aap*, *bhp* or *embp*, revealed that the only such gene found to be significantly associated with a "not cured" outcome was *bhp* in the lower extremity cohort ($p = 0.023$). Interestingly, *bhp* was most prevalent in the weak biofilm-forming isolates (52.7%) indicating that it's role may not be directly linked with biofilm forming ability, at least *in vitro*. *Bhp* has been reported to promote primary attachment to abiotic surfaces as well as intercellular adhesion during biofilm formation (20, 21). Thus, this protein might be important for rapid attachment to the implant rather than the amount of biofilm formed by the isolate *per se*. A rapid attachment clearly may be significant for early establishment of biofilm *in vivo* in "the race for the surface". This may partially explain its association with poor treatment outcome, despite the lack of association with *in vitro* biofilm forming ability.

Antibiotic resistance is a second key challenge in treatment of ODRI. Previous studies have suggested that methicillin resistance is associated with a worse treatment outcome in staphylococcal ODRIs (22-24), although a number of studies have provided contrasting findings (8, 25, 26). Methicillin resistance is due to the *mecA* gene. In the present study, resistance to methicillin showed a trend for a "not cured" patient outcome ($p = 0.082$ based on phenotypic analysis, and $p = 0.099$ for presence of *mecA* gene), supporting previously reported trends (22-24). Furthermore, the chromosome recombinase A and B encoding genes *ccrB* and *ccrA* were significantly more prevalent in "not cured" clinical outcome isolates

(89.5% vs 65.8% and 89.5% vs 64.7%; $p = 0.042$ and $p = 0.034$, respectively). These two genes are responsible for the chromosomal insertion of the genetic element called staphylococcal cassette chromosome *mec* (SCC*mec*). The SCC*mec* mobile genetic island contains the *mec* gene complex including the methicillin resistance gene *mecA*. In this study 91.8% of the *ccrA/ccrB* positive isolates possessed the *mecA* gene indicative for the presence of the mobile element SCC*mec*. Of those 67 *ccrA/ccrB*/*mecA* positive isolates, only 2 were not phenotypically resistant to methicillin, which might be due to mutations in the *mecA* gene. The 6 *ccrA*+/*ccrB*+ but *mecA* negative isolates were not phenotypically resistant to methicillin indicative of an absent SCC*mec* mobile element.

A second antibiotic class pertinent to the treatment of ODRIs is the aminoglycosides (including gentamicin and tobramycin), which are commonly used in antibiotic loaded bone cement (2, 16, 19, 27). Resistance to aminoglycosides in *S. epidermidis* isolated from patients with ODRI typically ranges from 40-65% (19, 27). In this study, 39.4% of the isolates were resistant to gentamicin/aminoglycoside and we observed an association between the “not cured” outcome and being phenotypically resistant to aminoglycosides ($p = 0.001$). The majority of the aminoglycoside resistant isolates (65.8 %) carried the *aacA (aac(6`)-aph(2")* gene that confers resistance to all aminoglycosides. This gene was also observed in a higher prevalence in the "not cured" group (42.1% versus 22.4% in "cured"; $p = 0.076$). This correlates well with other studies in terms of prevalence of the *aacA (aac(6`)-aph(2")* gene amongst aminoglycoside resistant isolates ranging between 40-92% (16, 28, 29), although how this impacted upon treatment outcome was not described for these other studies.

Our data also revealed that the antiseptic gene (quaternary ammonium compound) *qacA* gene was statistically more prevalent in the "not cured" outcome group ($p = 0.023$). The *qacA* gene is a plasmid-borne gene (pSK1 family plasmids) that confers resistance to antiseptics and disinfectants such as cetrimide, benzalkonium chloride and chlorhexidine (30-32). Our observation that the *qacA* gene was present in 67.3% of isolates (89.5% of “not cured”

isolates) within the complete cohort seems enriched compared to other studies for clinical (47%-52%) and commensal (25%) *S. epidermidis* isolates (31, 32). Despite relatively high presence in our collection and a moderate number of "not cured" isolates, there is some statistical significance to associate presence of this gene with a poor treatment outcome. Qac proteins are efflux pumps that not only protect bacteria from a variety of toxic substances but also from fluoroquinolones and β–lactams (30-32). The acquisition of such a gene/plasmid possibly from antiseptic usage within the hospital clearly provides the bacteria a survival advantage, especially in a clinical environment. Such resistant pathogens are therefore not only more difficult to clean within the hospital environment, but as we show, are also associated with a poor treatment outcome.

In addition, the IS*256*-like transposase was more frequently present in "not cured" clinical outcome isolates than in "cured" outcome isolates (57.9% versus 36.5%; $p = 0.085$). Previous studies have described an association between the presence of the IS*256* element, the *aac*(6`)-*aph*(2") gene (33, 34), the *icaADBC* operon and the ability to form biofilm (11, 35, 36). Furthermore, IS*256* has been suggested as molecular marker for the molecular typing and identification of nosocomial, invasive *S. epidermidis* isolates (9-11, 36). This study provides further evidence that IS*256* is not only "enriched" within invasive isolates, but is also more prevalent in isolates with a poor treatment outcome. The increased prevalence in the "not cured" group indicates it is not a marker for infection, but rather potentially one for poor outcome however this warrants further study with a larger set of isolates.

Previous genealogical reconstruction studies of *S. epidermidis* have shown that isolates clustered into 3 phylogenetic clades (4, 6), which is consistent with the observation in this study. To date, no study has associated genotypes with clinical outcomes in ODRI. In this study, a higher number of "not cured" outcome isolates where found in clade B compared to clade A and C. Clade B was also the lineage to have the strongest biofilm forming isolates.

Harris *et al* reported in their study an association between thick biofilm being 20% more common in CC2 isolates (6). In the present study, CC2 accounted for 80% of the 10 clade B isolates with 50% of them being responsible for moderate/strong biofilm formation. Furthermore, of these moderate/strong biofilm forming CC2 clade B isolates, 75% (3/4) resulted in a "not cured" outcome. This emphasizes that clade B CC2 isolates might be more likely to result in a poor clinical outcome. However, a greater number of isolates should be analyzed in a prospective manner in order to confirm this observation and determine whether it may be a prognostic molecular marker for poor treatment outcome.

A limitation of this study was that only a single *S. epidermidis* colony from each patient was analyzed, although the infection could, at least in theory, be polyclonal. A previous study has shown that only a minority of prosthetic joint infections (28.5% (4/14 patients)) were due to polyclonal *S. epidermidis* strains (37). Any future studies should consider analyzing several colonies from each patient. Furthermore, the morphology of colonies was not recorded in this present study and so we do not know how many, if any, SCVs were present in the current collection, but this should be considered in future studies

In general, SCVs present phenotypic features such as slow growth rate and small colony morphology (38-45). Additionally, SCVs are associated with increased biofilm-forming ability, antibiotic resistance and ability to internalize and persist in osteoblasts, all of which may contribute to prolonged treatment or even treatment failure (38-45). In contrast to *S. aureus* SCVs, very little information is available on *S. epidermidis* SCVs (38, 46). Only recently has the pathogenesis of PJIs been associated with *S. epidermidis* SCVs (38, 46, 47). Furthermore, SCV colonies from the same patient showed different appearance regarding growth rate, colony size and levels of gentamicin resistance when compared to each other (38). This highlights the importance of documenting and analyzing SCVs as they may influence treatment outcome.

Patients with a "not cured" clinical outcome were more likely to have had multiple revision surgeries in comparison with "cured" outcome ($p<0.067$), which is to be expected as revision surgery is a standard intervention for failed treatment We have considered the final outcome to be "cured" or "not cured" at follow-up, regardless of the treatment steps taken in the interim period. Therefore, even though multiple revision surgeries occurred, if the patient was free of infection at FUP, we considered it to be cured. Of course, the need for multiple revisions is possibly an indicator that the infection was a greater challenge to treat, however, in a large patient population such as this, there is often a need for multiple revision surgeries to advance the healing of the fracture, or replace the device, which may occur after infection has cleared, and so such patients have also had multiple surgeries.

Overall, genome sequencing is not absolutely required to determine some of the features identified in this study as being associated with poor outcome. For example, routine antibiotic susceptibility testing and conventional *in vitro* biofilm assays are readily available to provide this information. Nevertheless, whole genome sequencing allowed us to test our hypothesis with greatest sensitivity, and also identified features that are less easily measurable in a clinical laboratory. Finally, it should be mentioned that the treatment of ODRI is achieved by antibiotic therapy and surgical removal of infected tissue. Therefore, the outcome of ODRI treatment will be influenced by these factors in addition to the host defenses and not solely upon the pathogen itself. The factors identified in this study therefore require prospective validation in further studies with larger patient cohorts in order to confirm their value as prognostic markers for ODRI treatment outcome.

Materials and methods

Ethics Statement

Institutional Review Board approval was obtained from the local ethical committee "Ethik-Kommission der Bayerischen Landesärztekammer" under approval number 12063. The study was registered with https://clinicaltrials.gov with identifier NCT02640937. Only adult patients (> 18 years) were included in this study and all patients provided informed written consent prior to inclusion in the study.

Staphylococcus epidermidis **collection**

This was a prospective study performed between November 2011 and September 2013 at the BGU Murnau, Germany, a level-one trauma center with a high volume, 70-bed unit for septic and reconstructive surgery. The phenotypic investigation of biofilm formation of a subgroup of these isolates has been previously described (8), although no genome sequence data of these isolates has been previously published.

Inclusion criteria comprised of patients treated for a confirmed *S. epidermidis* infection involving fracture fixation (FFI) or prosthetic joint infections (PJI). Most of the primary surgeries for fracture fixation or implantation of an endoprosthesis were performed in other hospitals. In cases where the patient developed an infection that was not treated/treatable at the primary center, the patients were transferred to the study site which has a specialized unit for ODRI treatment. Bacterial growth in at least two biopsies, collected at the site of interest in combination with nonunion, implant-loosening/failure or local and systemic signs suggesting a surgical site infection were requirements for the diagnosis of ODRI, as per hospital standard.

In the previously described clinical study, patient data was analyzed as a complete study cohort, but also as a cohort including only patients with infections associated with the lower limb (8). This is because there are numerous outcome measures for the lower extremity that are not available for other anatomical locations. These outcome measures include the Lower

342 Extremity Functional Score (LEFS), the Short Form-12 (SF-12) score as well as leg length
343 discrepancy (8). The remaining patients, not included in the lower extremity cohort, included
344 patients with infections at other locations such as upper extremity, pelvis and spine (Table 1).
345 At the first surgical procedure after enrolment, bone biopsies were taken from the interface
346 between implant and affected bone. Samples were placed in a sterile container with
347 thioglycollate liquid medium (bioMérieux, Hazelwood, MO, USA) and cultured for 10 days at
348 37 °C. Any growth was inoculated onto a blood agar plate (bioMérieux, Hazelwood, MO,
349 USA) for further growth and subsequent identification. All isolates were grown on Tryptone
350 Soy Agar (TSA, Oxoid, Pratteln, Switzerland) and incubated overnight at 37 °C. A single
351 colony was then taken and resuspended in 1 ml Tryptone Soy Broth (TSB, Oxoid, Pratteln,
352 Switzerland) containing 20% vol/vol glycerol for long-term storage at -80 °C. Although
353 colony morphology of culture positive samples was not described, we anticipate that SCV
354 colonies had sufficient time to emerge under standard laboratory conditions, and are not likely
355 to have been missed in the clinical routine.
356
357 **Clinical data collection**
358 Clinical data was collected from each enrolled patient. The following surgical parameters
359 were documented: affected bone or joint; type of implant; time between implantation of the
360 device and onset of symptoms; and whether the fracture was open or closed (PJIs excluded).

361 Patients were assessed for treatment outcome after an average of 26 months follow-up (FUP).
362 Patients were assigned to have had a "cured" or a "not cured" outcome at FUP. Patients had a
363 "cured" clinical outcome if they were free of infection, surgical therapy and systemic
364 antibiotic therapy ceased, and function of the affected joint or limb was restored. Patients
365 were considered to have had a "not cured" clinical outcome if at least one of the above
366 parameters was negative. Additional parameters were documented such as acute/non-acute

(chronic) infection (cut-off for onset of symptoms: six weeks), obesity (BMI ≥ 30kg/m^2), diabetes, smoking, chronic immunosuppressive conditions (diabetes mellitus, chronic alcoholism, Child`s class C cirrhosis, neoplasia, transplantation, AIDS and steroid medication) and whether multiple revision surgeries were required during treatment.

The clinical treatment strategies applied to these patients followed recent guidelines and recommendations, including guidance on antimicrobial stewardship. Therefore, treatment strategies differed between enrolled patients due to antibiotic resistance patterns, presence of implant (yes/no) and stage of treatment. The use of antibiotic loaded bone cement was not extracted from the patient records, however, in all cases of infection with a gentamicin resistant organism, any bone cement would have been loaded with vancomycin as the preferred alternative. Whether an implant was removed or retained was dependent upon the classification of the infection and the health status of the patient. In chronic infections, the implant was routinely removed in the first revision surgery whenever possible. In general, the implant was retained in acute infections if sufficient debridement was possible.

Antibiotic susceptibility testing

Antibiotic susceptibility to 28 antibiotics was determined using a Vitek2 machine (bioMérieux Vitek Inc., Hazelwood, MO, USA). The antibiotics tested were amikacin, ampicillin-sulbactam, cefotaxim, cefoxitin, cefuroxime, ciprofloxacin, clindamycin, daptomycin, erythromycin, fosfomycin, fusidic acid, gentamicin, levofloxacin, linezolid, mezlocillin, moxifloxacin, netilmicin, ofloxacin, oxacillin, penicillin, piperacillin, rifampicin, tetracycline, ticarcillin/clavulanate, tigecycline, tobramycin, trimethoprim-sulfamethoxazole and vancomycin. Multiple antibiotic resistance was defined according to the definitions of the European Committee of Antimicrobial Susceptibility Testing (EUCAST). Oxacillin resistance was considered definitive for methicillin resistance status.

392 **Biofilm formation**

393 Biofilm formation was assayed as described previously (48, 49). Briefly, overnight cultures
394 were grown in TSB and then sub-cultured in fresh TSB containing 1% glucose, to
395 approximately $1x10^6$ CFU/ml. To achieve this, bacterial density was adjusted to a target
396 optical density of known concentration using a Multiskan Go microplate reader (Thermo
397 Scientific, Zürich, Switzerland). A total of 200 µl of the bacterial suspension was incubated in
398 flat-bottomed 96-well tissue culture-treated polystyrene microtitre plates (Nuclon, Nunc A/S,
399 Denmark) for 24 h at 37 °C. Plates were rinsed with phosphate-buffered saline (PBS, Sigma-
400 Aldrich, Buchs, Switzerland) and stained with 150 µl of Gram`s crystal violet solution
401 (Sigma-Aldrich, Buchs, Switzerland). The dye bound to the attached cells was solubilized by
402 addition of 150 µl of 95% ethanol. Optical density was measured as absorbance at 595 nm
403 using the Multiskan Go microplate reader.

404 All isolates were tested in triplicate in three independent experiments. Each microtitre plate
405 also consisted of negative controls (wells without bacterial inoculation). The average OD
406 value (OD_a) was calculated for each isolate and the negative control. The results were
407 evaluated using the scale described by Stepanovic *et al*. (49), whereby isolates may fall into
408 the following four categories: no biofilm producer, weak biofilm producer, intermediate
409 biofilm producer and strong biofilm producer. Based on the OD_a values and the cut-off value
410 (OD_c), which is defined as three standard deviations (SD) above the mean OD of the negative
411 control: OD_c = average OD of negative control+(3xSD of negative control). The strength of
412 the biofilm production of each isolate was calculated as following: $OD_a \leq OD_c$ = biofilm non-
413 producer; $OD_c < OD_a \leq 2xOD_c$ = weak biofilm producer; $2xOD_c < OD_a \leq 4xOD_c$ =
414 intermediate biofilm producer and $4xOD_c < OD_a$ = strong biofilm producer. *S. epidermidis*
415 reference strain RP12 (ATCC 35983) was used as a control for strong biofilm production.

416

Genome sequencing and assembly

S. epidermidis isolates were cultured on TSA plates at 37 °C for 24h. Single colony cultures were harvested and re-suspended in 3 ml of TSB medium to minimize clumping and incubated at 37 °C with overnight shaking. Chromosomal DNA was extracted using a Qiagen QiAmp DNA mini kit (Qiagen, Hilden, Germany) following the manufacturer`s protocol using 1 µg/ml lysostaphin (Sigma-Aldrich, Buchs, Switzerland) and 2 µg/ml lysozyme (Sigma-Aldrich, Buchs, Switzerland) to facilitate cell lysis. DNA was sequenced at the Swansea University Genome Centre using a MiSeq benchtop sequencer (Illumina, San Diego, CA, USA). Sequencing libraries were prepared using Nextera XT library preparation kits (v2) and paired-end 250 bp reads generated with the MiSeq run kit (v2). Short read data was assembled using a *de novo* assembly algorithm within *Velvet* software (version 1.2.08) (50). Overall, the average number of contiguous sequences (contigs) for all 104 genomes sequenced in this study was 439 which gave rise to an average total assembled genome size of 2,436,856 bp. Short reads are available from the NCBI short read archive (SRA) associated with BioProject: PRJNA382527. Assembled genomes are also archived in the publicly accessible Staphylococcal Bacterial Isolate Genome Sequence database (BIGSdb; https://sheppardlab.com/resources/).

Genomes are archived using a gene-by-gene approach for genome alignment and comparison supported by the BLAST algorithm (51). A reference pan-genome was constructed from the clinical isolate genomes (all collected as part of this study) and the reference *S. epidermidis* RP62A (ATCC 35984) and ATCC 12228 genomes (52). Putative gene function was assigned to genes in the reference pan-genome list using RAST (Rapid Annotations using Subsystem Technology)(53) and the SEED database (54) which were cross-referenced with the *S. epidermidis* RP62A (ATCC 35984) and ATCC 12228 reference genomes before removing duplicate genes. The BLAST algorithm was used to scan all genomes for gene orthologs at each locus in the reference pan-genome. An ortholog was defined as a reciprocal best hit of

the sequence with >70% nucleotide identity over at least 50% of the alignment length. MAFFT software (55) was used to align gene orthologs on a gene-by-gene basis, and these data concatenated into contiguous sequence for each isolate genome, including gaps. A core genome of 123 genes was defined based on gene presence in all isolates (100%).

Estimating genealogies

ClonalFrame infers the clonal relationship of bacteria and the chromosomal position of homologous recombination events that disrupt a clonal pattern of inheritance (56). A stringent approach was used to estimate a reduced core genome for construction of a genealogy using ClonalFrame (version 1.2) on concatenated sequences of 104 *S. epidermidis* genomes with 100,000 iterations, half of which were discarded as burn-in. Substitution mutation and recombination regions were categorized from the output of ClonalFrame. The posterior probability of recombination and substitution at each site is calculated by ClonalFrame and recombination events were defined with a probability of recombination more than 75%, reaching 95% at any one site. The trees were visualized and annotated using MEGA6 (57).

Statistical analysis

Association amongst and between the clinical parameters, bacterial phenotypes, clades and presence/absence of genes were analyzed statistically using Chi-square test, Fisher`s exact test, Cochran-Armitage trend test or Kruskal-Wallis test as appropriate. Chi-square test was carried out to test the null hypothesis that the lineages are homogenous in their clinical outcome or resistance phenotypes. Permutation tests were performed to test the null hypothesis that there was no association between lineage and clinical outcome or resistance phenotype. Association between clinical outcome or antimicrobial resistance and lineage in the observed data was summarized by an association score. Statistical analyses were

performed using SAS software (Version 9.2; Cary, NC, USA) and SPSS (Version 10, IBM, USA) and level of significance set at $p \leq 0.05$.

Acknowledgements. The authors would like to acknowledge the contribution of Kathrin Espinoza, MSc of AO Clinical Investigation and Documentation, Dübendorf, Switzerland for statistical support.

Financial support. This work was funded by AO Trauma as part of the Clinical Priority Program, Bone Infection. SKS is supported by the Medical Research Council, project reference MR/L015080/1. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Trampuz A, Zimmerli W. 2005. Prosthetic joint infections: update in diagnosis and treatment. Swiss Med Wkly 135:243-51.
2. Trampuz A, Zimmerli W. 2006. Diagnosis and treatment of infections associated with fracture-fixation devices. Injury 37 Suppl 2:S59-66.
3. Conlan S, Mijares LA, Program NCS, Becker J, Blakesley RW, Bouffard GG, Brooks S, Coleman H, Gupta J, Gurson N, Park M, Schmidt B, Thomas PJ, Otto M, Kong HH, Murray PR, Segre JA. 2012. *Staphylococcus epidermidis* pan-genome sequence

analysis reveals diversity of skin commensal and hospital infection-associated isolates. Genome Biol 13:R64.

4. Meric G, Miragaia M, de Been M, Yahara K, Pascoe B, Mageiros L, Mikhail J, Harris LG, Wilkinson TS, Rolo J, Lamble S, Bray JE, Jolley KA, Hanage WP, Bowden R, Maiden MC, Mack D, de Lencastre H, Feil EJ, Corander J, Sheppard SK. 2015. Ecological Overlap and Horizontal Gene Transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*. Genome Biol Evol 7:1313-28.

5. Miragaia M, Thomas JC, Couto I, Enright MC, de Lencastre H. 2007. Inferring a population structure for *Staphylococcus epidermidis* from multilocus sequence typing data. J Bacteriol 189:2540-52.

6. Harris LG, Murray S, Pascoe B, Bray J, Meric G, Magerios L, Wilkinson TS, Jeeves R, Rohde H, Schwarz S, de Lencastre H, Miragaia M, Rolo J, Bowden R, Jolley KA, Maiden MC, Mack D, Sheppard SK. 2016. Biofilm Morphotypes and Population Structure among *Staphylococcus epidermidis* from Commensal and Clinical Samples. PLoS One 11:e0151240.

7. Fey PD, Olson ME. 2010. Current concepts in biofilm formation of *Staphylococcus epidermidis*. Future Microbiol 5:917-33.

8. Morgenstern M, Post V, Erichsen C, Hungerer S, Buhren V, Militz M, Richards G, Moriarty F. 2016. Biofilm formation increases treatment failure in *Staphylococcus epidermidis* device-related osteomyelitis of the lower extremity in human patients. J Orthop Res doi:10.1002/jor.23218.

9. Yao Y, Sturdevant DE, Villaruz A, Xu L, Gao Q, Otto M. 2005. Factors characterizing *Staphylococcus epidermidis* invasiveness determined by comparative genomics. Infect Immun 73:1856-60.

514 10. Gu J, Li H, Li M, Vuong C, Otto M, Wen Y, Gao Q. 2005. Bacterial insertion
515 sequence IS*256* as a potential molecular marker to discriminate invasive strains from
516 commensal strains of *Staphylococcus epidermidis*. J Hosp Infect 61:342-8.
517 11. Montanaro L, Campoccia D, Pirini V, Ravaioli S, Otto M, Arciola CR. 2007.
518 Antibiotic multiresistance strictly associated with IS*256* and ica genes in
519 *Staphylococcus epidermidis* strains from implant orthopedic infections. J Biomed
520 Mater Res A 83:813-8.
521 12. Harris LG, Bexfield A, Nigam Y, Rohde H, Ratcliffe NA, Mack D. 2009. Disruption
522 of *Staphylococcus epidermidi*s biofilms by medicinal maggot Lucilia sericata
523 excretions/secretions. Int J Artif Organs 32:555-64.
524 13. Mack D, Davies AP, Harris LG, Knobloch JK, Rohde H. 2009. *Staphylococcus*
525 *epidermidis* Biofilms: Functional Molecules, Relation to Virulence, and Vaccine
526 Potential. Top Curr Chem 288:157-82.
527 14. Otto M. 2012. Molecular basis of *Staphylococcus epidermidis* infections. Semin
528 Immunopathol 34:201-14.
529 15. Thomas JC, Vargas MR, Miragaia M, Peacock SJ, Archer GL, Enright MC. 2007.
530 Improved multilocus sequence typing scheme for *Staphylococcus epidermidis*. J Clin
531 Microbiol 45:616-9.
532 16. Campoccia D, Montanaro L, Pirini V, Ravaioli S, Arciola CR. 2009. Prevalence of
533 genes for aminoglycoside-modifying enzymes in *Staphylococcus epidermidis* isolates
534 from orthopedic postsurgical and implant-related infections. J Biomed Mater Res A
535 88:654-63.
536 17. Otto M. 2009. *Staphylococcus epidermidis*--the 'accidental' pathogen. Nat Rev
537 Microbiol 7:555-67.

18. Spiegl U, Friederichs J, Patzold R, Militz M, Josten C, Buhren V. 2013. Risk factors for failed two-stage procedure after chronic posttraumatic periprosthetic hip infections. Arch Orthop Trauma Surg 133:421-8.

19. Uckay I, Pittet D, Vaudaux P, Sax H, Lew D, Waldvogel F. 2009. Foreign body infections due to *Staphylococcus epidermidis*. Ann Med 41:109-19.

20. Bowden MG, Chen W, Singvall J, Xu Y, Peacock SJ, Valtulina V, Speziale P, Hook M. 2005. Identification and preliminary characterization of cell-wall-anchored proteins of *Staphylococcus epidermidis*. Microbiology 151:1453-64.

21. Cucarella C, Solano C, Valle J, Amorena B, Lasa I, Penades JR. 2001. Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. J Bacteriol 183:2888-96.

22. Kilgus DJ, Howe DJ, Strang A. 2002. Results of periprosthetic hip and knee infections caused by resistant bacteria. Clin Orthop Relat Res:116-24.

23. Salgado CD, Dash S, Cantey JR, Marculescu CE. 2007. Higher risk of failure of methicillin-resistant *Staphylococcus aureus* prosthetic joint infections. Clin Orthop Relat Res 461:48-53.

24. Leung F, Richards CJ, Garbuz DS, Masri BA, Duncan CP. 2011. Two-stage total hip arthroplasty: how often does it control methicillin-resistant infection? Clin Orthop Relat Res 469:1009-15.

25. Teterycz D, Ferry T, Lew D, Stern R, Assal M, Hoffmeyer P, Bernard L, Uckay I. 2010. Outcome of orthopedic implant infections due to different staphylococci. Int J Infect Dis 14:e913-8.

26. Volin SJ, Hinrichs SH, Garvin KL. 2004. Two-stage reimplantation of total joint infections: a comparison of resistant and non-resistant organisms. Clin Orthop Relat Res:94-100.

563 27. Anguita-Alonso P, Hanssen AD, Osmon DR, Trampuz A, Steckelberg JM, Patel R.
564 2005. High rate of aminoglycoside resistance among staphylococci causing prosthetic
565 joint infection. Clin Orthop Relat Res 439:43-7.
566 28. Abbassi MS, Bouchami O, Touati A, Achour W, Ben Hassen A. 2008. Clonality and
567 occurrence of genes encoding antibiotic resistance and biofilm in methicillin-resistant
568 *Staphylococcus epidermidis* strains isolated from catheters and bacteremia in
569 neutropenic patients. Curr Microbiol 57:442-8.
570 29. Cherifi S, Byl B, Deplano A, Nonhoff C, Denis O, Hallin M. 2013. Comparative
571 epidemiology of *Staphylococcus epidermidis* isolates from patients with catheter-
572 related bacteremia and from healthy volunteers. J Clin Microbiol 51:1541-7.
573 30. Wassenaar TM, Ussery D, Nielsen LN, Ingmer H. 2015. Review and phylogenetic
574 analysis of qac genes that reduce susceptibility to quaternary ammonium compounds
575 in *Staphylococcus* species. Eur J Microbiol Immunol (Bp) 5:44-61.
576 31. Prag G, Falk-Brynhildsen K, Jacobsson S, Hellmark B, Unemo M, Soderquist B.
577 2014. Decreased susceptibility to chlorhexidine and prevalence of disinfectant
578 resistance genes among clinical isolates of *Staphylococcus epidermidis*. APMIS
579 122:961-7.
580 32. Taheri N, Ardebili A, Amouzandeh-Nobaveh A, Ghaznavi-Rad E. 2016. Frequency of
581 Antiseptic Resistance Among *Staphylococcus aureus* and Coagulase-Negative
582 Staphylococci Isolated From a University Hospital in Central Iran. Oman Med J
583 31:426-432.
584 33. Culebras E, Martinez JL. 1999. Aminoglycoside resistance mediated by the
585 bifunctional enzyme 6'-N-aminoglycoside acetyltransferase-2''-O-aminoglycoside
586 phosphotransferase. Front Biosci 4:D1-8.
587 34. Thomas WD, Jr., Archer GL. 1989. Mobility of gentamicin resistance genes from
588 staphylococci isolated in the United States: identification of Tn*4031*, a gentamicin

589 resistance transposon from *Staphylococcus epidermidis*. Antimicrob Agents
590 Chemother 33:1335-41.
591 35. Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K, Ziebuhr W. 2004. The bacterial
592 insertion sequence element IS*256* occurs preferentially in nosocomial *Staphylococcus*
593 *epidermidis* isolates: association with biofilm formation and resistance to
594 aminoglycosides. Infect Immun 72:1210-5.
595 36. Ziebuhr W, Krimmer V, Rachid S, Lossner I, Gotz F, Hacker J. 1999. A novel
596 mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence
597 for control of the polysaccharide intercellular adhesin synthesis by alternating
598 insertion and excision of the insertion sequence element IS*256*. Mol Microbiol
599 32:345-56.
600 37. Galdbart JO, Morvan A, El Solh N. 2000. Phenotypic and molecular typing of
601 nosocomial methicillin-resistant *Staphylococcus aureus* strains susceptible to
602 gentamicin isolated in france from 1995 to 1997. J Clin Microbiol 38:185-90.
603 38. Bogut A, Niedzwiadek J, Koziol-Montewka M, Strzelec-Nowak D, Blacha J,
604 Mazurkiewicz T, Marczynski W, Plewik D. 2014. Characterization of *Staphylococcus*
605 *epidermidis* and *Staphyloccocus warneri* small-colony variants associated with
606 prosthetic-joint infections. J Med Microbiol 63:176-85.
607 39. Hirschhausen N, Block D, Bianconi I, Bragonzi A, Birtel J, Lee JC, Dubbers A,
608 Kuster P, Kahl J, Peters G, Kahl BC. 2013. Extended *Staphylococcus aureus*
609 persistence in cystic fibrosis is associated with bacterial adaptation. Int J Med
610 Microbiol 303:685-92.
611 40. Proctor RA, Kriegeskorte A, Kahl BC, Becker K, Loffler B, Peters G. 2014.
612 *Staphylococcus aureus* Small Colony Variants (SCVs): a road map for the metabolic
613 pathways involved in persistent infections. Front Cell Infect Microbiol 4:99.

614 41. Sendi P, Proctor RA. 2009. *Staphylococcus aureus* as an intracellular pathogen: the
615 role of small colony variants. Trends Microbiol 17:54-8.
616 42. Singh R, Ray P, Das A, Sharma M. 2009. Role of persisters and small-colony variants
617 in antibiotic resistance of planktonic and biofilm-associated *Staphylococcus aureus*: an
618 in vitro study. J Med Microbiol 58:1067-73.
619 43. Tuchscherr L, Kreis CA, Hoerr V, Flint L, Hachmeister M, Geraci J, Bremer-Streck S,
620 Kiehntopf M, Medina E, Kribus M, Raschke M, Pletz M, Peters G, Loffler B. 2016.
621 *Staphylococcus aureus* develops increased resistance to antibiotics by forming
622 dynamic small colony variants during chronic osteomyelitis. J Antimicrob Chemother
623 71:438-48.
624 44. Tuchscherr L, Loffler B. 2016. *Staphylococcus aureus* dynamically adapts global
625 regulators and virulence factor expression in the course from acute to chronic
626 infection. Curr Genet 62:15-7.
627 45. von Eiff C, McNamara P, Becker K, Bates D, Lei XH, Ziman M, Bochner BR, Peters
628 G, Proctor RA. 2006. Phenotype microarray profiling of *Staphylococcus aureus menD*
629 and *hemB* mutants with the small-colony-variant phenotype. J Bacteriol 188:687-93.
630 46. Perez K, Patel R. 2017. *Staphylococcusepidermidis* Small-Colony Variants Are
631 Induced by Low pH and Their Frequency Reduced by Lysosomal Alkalinization. J
632 Infect Dis 215:488-490.
633 47. Tande AJ, Osmon DR, Greenwood-Quaintance KE, Mabry TM, Hanssen AD, Patel R.
634 2014. Clinical characteristics and outcomes of prosthetic joint infection caused by
635 small colony variant staphylococci. MBio 5:e01910-14.
636 48. Post V, Wahl P, Uckay I, Ochsner P, Zimmerli W, Corvec S, Loiez C, Richards RG,
637 Moriarty TF. 2014. Phenotypic and genotypic characterisation of *Staphylococcus*
638 *aureus* causing musculoskeletal infections. Int J Med Microbiol 304:565-76.

639 49. Stepanovic S, Vukovic D, Hola V, Di Bonaventura G, Djukic S, Cirkovic I, Ruzicka
640 F. 2007. Quantification of biofilm in microtiter plates: overview of testing conditions
641 and practical recommendations for assessment of biofilm production by staphylococci.
642 APMIS 115:891-9.
643 50. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly
644 using de Bruijn graphs. Genome Res 18:821-9.
645 51. Sheppard SK, Jolley KA, Maiden MC. 2012. A Gene-By-Gene Approach to Bacterial
646 Population Genomics: Whole Genome MLST of Campylobacter. Genes (Basel)
647 3:261-77.
648 52. Meric G, Yahara K, Mageiros L, Pascoe B, Maiden MC, Jolley KA, Sheppard SK.
649 2014. A reference pan-genome approach to comparative bacterial genomics:
650 identification of novel epidemiological markers in pathogenic Campylobacter. PLoS
651 One 9:e92798.
652 53. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes
653 S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA,
654 McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R,
655 Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid
656 annotations using subsystems technology. BMC Genomics 9:75.
657 54. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S,
658 Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED
659 and the Rapid Annotation of microbial genomes using Subsystems Technology
660 (RAST). Nucleic Acids Res 42:D206-14.
661 55. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version
662 7: improvements in performance and usability. Mol Biol Evol 30:772-80.
663 56. Didelot X, Falush D. 2007. Inference of bacterial microevolution using multilocus
664 sequence data. Genetics 175:1251-66.

57. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-9.

Figure legends

Fig 1. Graph showing the percentage of genes present in the whole collection (black) and present in the two outcome groups. "Cured" outcome (dashed) and "not cured" outcome (grey), asterisk indicates statistically significance $p \leq 0.05$.

Fig 2. Population structure of *S. epidermidis* isolates constructed from 123 core genes and implemented in ClonalFrame. In (**A**) all 104 isolates of the complete cohort study and in (**B**) all 70 isolates of the lower extremity are labelled according to the clinical follow-up (FUP) outcome: "not cured" (black circle) and "cured" (open circle). **(C)** and **(D)** shows the percentage distribution of "cured" and "not cured" outcome in the three clades A, B, C with (**C**) showing the complete cohort and (**D)** the lower extremity cohort

Fig 3. Population structure of *S. epidermidis* isolates constructed from 123 core genes and implemented in ClonalFrame. In (**A**) all 104 isolates of the complete cohort and in (**B**) of the lower extremity cohort are labelled according to the ability to form a biofilm: biofilm negative (open circles), weak biofilm (grey circles) and moderate to strong biofilm formers (black circles). (**C)** and **(D**) shows the percentage distribution of the strength of biofilm formation in the three clades A, B, C with (**C**) showing the complete cohort and (**D**) the lower extremity cohort.

688 **Table 1. Patient health status, infection characteristics, bacteriology, clinical course and**
689 **outcome.**

	Complete study cohort	Lower extremity cohort
Total, n (%)*	104 (100.0)	70 (100.0)
Clinical course and infection outcome, n (%)		
Multiple revision surgeries	89 (85.6)	68 (97.1)
Clinical outcome cured	85 (81.7)	53 (75.7)
Health status n (%)		
Obesity[1]	44 (42.3)	32 (45.7)
Smoking	28 (26.9)	17 (24.3)
Diabetes	14 (13.5)	8 (11.4)
Chronic immunosuppression	25 (24.0)	14 (20.0)
Infection characteristics, n (%)		
Infection after fracture fixation (FFI)	78 (75.0)	44 (62.9)
Prosthetic joint infections (PJI)	26 (25.0)	26 (37.1)
Acute infection	29 (27.9)	20 (28.6)
Closed facture[2]	53 (67.9)	25 (56.8)
Open fracture	25 (32.1)	19 (43.2)
Type of implant, n (%)		
Internal fixator	6 (5.8)	0 (0)
Prosthetic joint	26 (25.0)	26 (37.1)
Nail	23 (22.1)	20 (28.6)
Plate	40 (38.5)	22 (31.4)
Screw	8 (7.7)	2 (2.9)
K-wire	1 (1.0)	0 (0)
Localization, n (%)		
Spine	6 (5.8)	NA
Upper extremity	7 (6.7)	NA
Pelvis	7 (6.7)	NA
Tibia	2 (1.9)	NA
Clavicle	3 (2.9)	NA
Hip joint	10 (9.6)	10 (14.3)
Femur	7 (6.7)	7 (10.0)
Knee joint	18 (17.3)	18 (25.7)
Lower leg including upper ankle joint	35 (33.7)	35 (50.0)
Lower ankle joint including foot	9 (8.7)	NA
Bacteriological evaluation, n (%)		
Methicillin resistance	70 (67.3)	52 (74.3)
Multi-drug resistance	77 (74.0)	56 (80.0)
Biofilm formation, n (%)		
Non	31 (29.8)	24 (34.3)
Weak	39 (37.5)	26 (37.1)
Intermediate	22 (21.2)	11 (15.7)
Strong	12 (11.5)	9 (12.9)

690 *Single *S. epidermidis* isolate from each patient.
691 [1]Obesity defined as body mass index BMI > 30.
692 [2]PJI not included.
693 NA: not applicable

694 **Table 2. Association between prognostic factors and cure status for the complete study**
695 **cohort.**

	Cured			
	No	Yes		
	n (%)	n (%)	Odds ratio for cured[1] (95%-Confidence Interval)	*p*-value
Total number of patient*	19 (18.3)	85 (81.7)		
Infection type			0.92 (0.30;2.86)	1.000†
FFI	14 (17.9)	64 (82.1)		
PJI	5 (19.2)	21 (80.8)		
Fracture			0.39 (0.12;1.27)	0.126†
Closed	7 (13.2)	46 (86.8)		
Open	7 (28.0)	18 (72.0)		
Acute infection			1.56 (0.44;7.08)	0.463††
No	15 (20.0)	60 (80.0)		
Yes	4 (13.8)	25 (86.2)		
Multiple-revision surgery			0.12 (0.00;2.04)	0.067†
No	0 (0.0)	15 (100.0)		
Yes	19 (21.3)	70 (78.7)		
Obesity			0.78 (0.25;2.42)	0.621††
No	10 (16.7)	50 (83.3)		
Yes	9 (20.5)	35 (79.5)		
Smoking			0.76 (0.23;2.74)	0.613††
No	13 (17.1)	63 (82.9)		
Yes	6 (21.4)	22 (78.6)		
Diabetes			0.50 (0.14;1.81)	0.281†
No	15 (16.7)	75 (83.3)		
Yes	4 (28.6)	10 (71.4)		
Chronic immunosuppression			0.46 (0.16;1.34)	0.233†
No	12 (15.2)	67 (84.8)		
Yes	7 (28.0)	18 (72.0)		

[1]For calculation of odds ratios involving cells with 0 observations, the 0.5 zero-cell correction was applied.

*Each patient had 1 *S. epidermidis* isolate.

†Chi-Square test

††Fishers exact test

Table 3: Association between bacterial phenotype and clinical cured status

	Complete cohort (n=104)				Lower extremity cohort (n=70)			
	Cured				Cured			
	No	Yes			No	Yes		
	n (%)	n (%)	Odds ratio for cured (95%-Confidence Interval)	*p*-value	n (%)	n (%)	Odds ratio for cured (95%-Confidence Interval)	*p*-value
Biofilm formation				0.059†††				0.031†††
No	3 (9.7)	28 (90.3)			3 (12.5)	21 (87.5)		
Weak	7 (17.9)	32 (82.1)	0.49 (0.08;2.43)		6 (23.1)	20 (76.9)	0.48 (0.07;2.64)	
Intermediate	5 (22.7)	17 (77.3)	0.36 (0.05;2.19)		4 (36.4)	7 (63.6)	0.25 (0.03;1.96)	
Strong	4 (33.3)	8 (66.7)	0.21 (0.03;1.62)		4 (44.4)	5 (55.6)	0.18 (0.02;1.51)	
Antibiotic resistance*								
Methicillin			0.33 (0.06;1.29)	0.082††			0.54 (0.14;2.16)	0.529†
No	3 (8.8)	31 (91.2)			3 (16.7)	15 (83.3)		
Yes	16 (22.9)	54 (77.1)			14 (26.9)	38 (73.1)		
Aminoglycosides			0.17 (0.04;0.56)	<.001††				0.051††
No	5 (7.9)	58 (92.1)			5 (14.3)	30 (85.7)	0.32 (0.08;1.17)	
Yes	14 (34.1)	27 (65.9)			12 (34.3)	23 (65.7)		

*Not all antibiotic resistances are listed. Others tested showed no statistical significance.

†Chi-Square test, ††Fishers exact test, ††† Cochran-Armitage Trend Test

Table 4. Biofilm-associated genes and biofilm formation.

Presence of Biofilm-associated genes	Complete cohort study (n=104) Strength of biofilm *in vitro* n; (%)				Lower cohort study (n=70) Strength of biofilm *in vitro* n; (%)			
	None (n=31)	Weak (n=39)	Intermediate (n=22)	Strong (n=12)	None (n=24)	Weak (n=26)	Intermediate (n=11)	Strong (n=9)
icaA	7 (22.6)	7 (18.0)	8 (36.4)	10 (83.3)	7 (29.2)	5 (19.2)	6 (54.5)	8 (88.9)
aap	1 (3.2)	0 (0)	0 (0)	1 (8.3)	1 (4.2)	0 (0)	0 (0)	0 (0)
bhp	7 (22.6)	10 (25.7)	0 (0)	2 (16.7)	4 (16.7)	9 (34.6)	0 (0)	1 (11.1)
embp	28 (90.3)	31 (79.5)	15 (68.2)	8 (66.7)	21 (87.5)	20 (76.9)	6 (54.5)	6 (66.7)

Table 5. Association between the *agr*-types and clinical outcome.

Outcomes	*agr*-type[1]			*p*-value[$]
	I	II	III	
	n (%)	n (%)	n (%)	
Clinical outcome				0.946
Not cured	7 (36.8)	5 (26.3)	7 (36.8)	
Cured	33 (39.8)	19 (22.9)	31 (37.3)	
Acute Infection				0.002
Non acute (chronic)	26 (35.1)	13 (17.6)	35 (47.3)	
Acute	14 (50.0)	11 (39.3)	3 (10.7)	

[1] Two isolates, not belonging to any of the 3 *agr* groups were excluded for statistical reasons.

[$]Chi-Square test

Table 6. MLST of the 104 clinical *S. epidermidis* isolates.

ST[a]	CC[b]	No of isolates; n (%)	Clinical outcome	
			"Cured"	"Not cured"
2	2	18 (17.3)	13 (72.2)	5 (27.8)
5	2	18 (17.3)	16 (88.9)	2 (11.1)
7	2	1 (1.0)	1 (100)	0 (0)
23	2	4 (3.9)	4 (100)	0 (0)
57	2	2 (1.9)	0 (0)	2 (100)
59	2	6 (5.8)	5 (83.3)	1 (16.7)
73	2	1 (1.0)	1 (100)	0 (0)
83	2	1 (1.0)	1 (100)	0 (0)
87	2	4 (3.9)	2 (50)	2 (50)
88	2	1 (1.0)	1 (100)	0 (0)
89	2	1 (1.0)	0 (0)	1 (100)
130	2	6 (5.8)	6 (100)	0 (0)
184	2	1 (1.0)	1 (100)	0 (0)
384	2	1 (1.0)	1 (100)	0 (0)
19	147	2 (1.9)	2 (100)	0 (0)
32	S32	2 (1.9)	2 (100)	0 (0)
110	S110	1 (1.0)	0 (0)	1 (100)
167	S167	1 (1.0)	1 (100)	0 (0)
297	S297	1 (1.0)	1 (100)	0 (0)
490	S490	1 (1.0)	1 (100)	0 (0)
528	S528	1 (1.0)	1 (100)	0 (0)
n/a	n/a	30 (28.9)	25 (83.3)	5 (16.7)
total		104 (100)	85	19

[a]Sequence types determined using the built-in MLST function of BIGSdb, linked with pubMLST databases. Lack of information denotes the possible truncation of a corresponding MLST locus at the end of a contig.

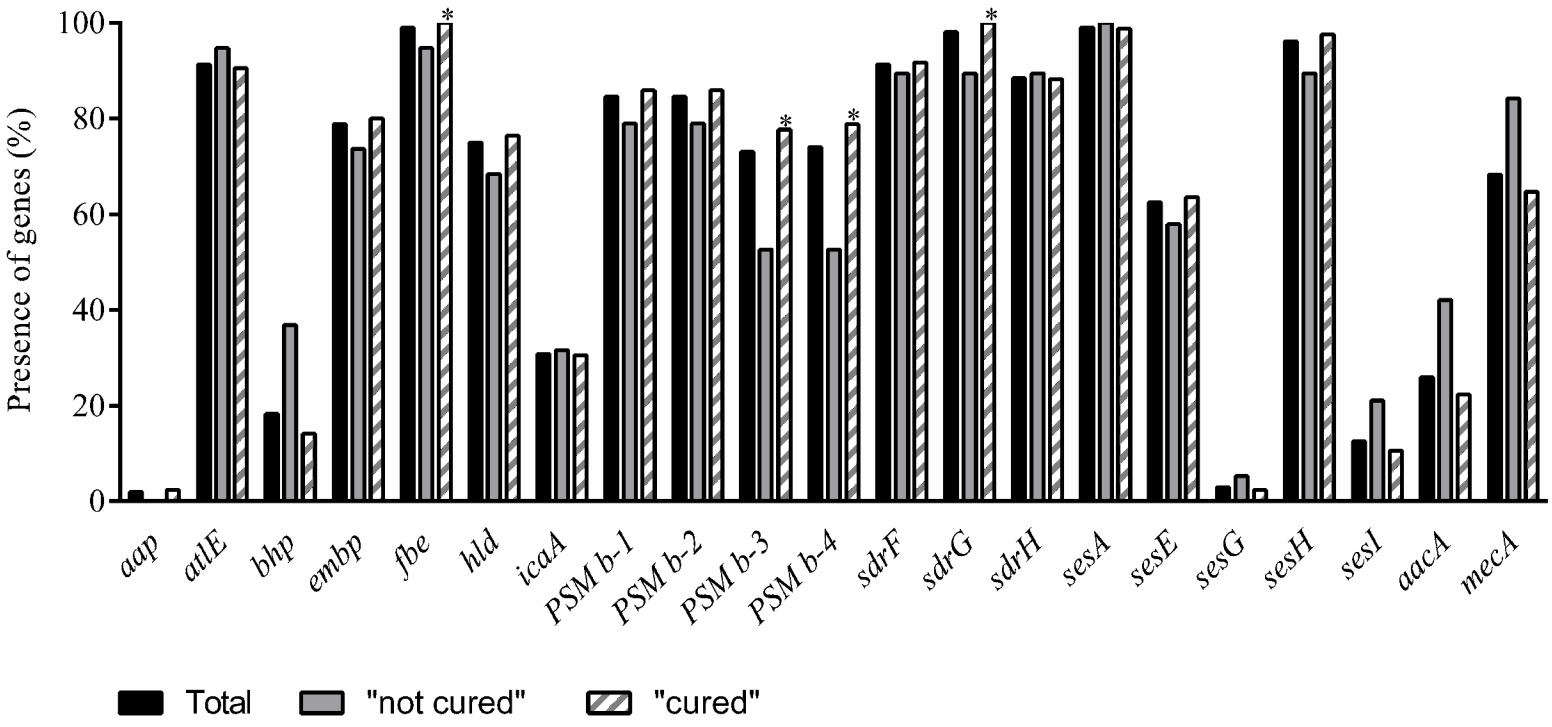
[b]Clonal complexes were obtained from previously described data (Meric et al.2015).

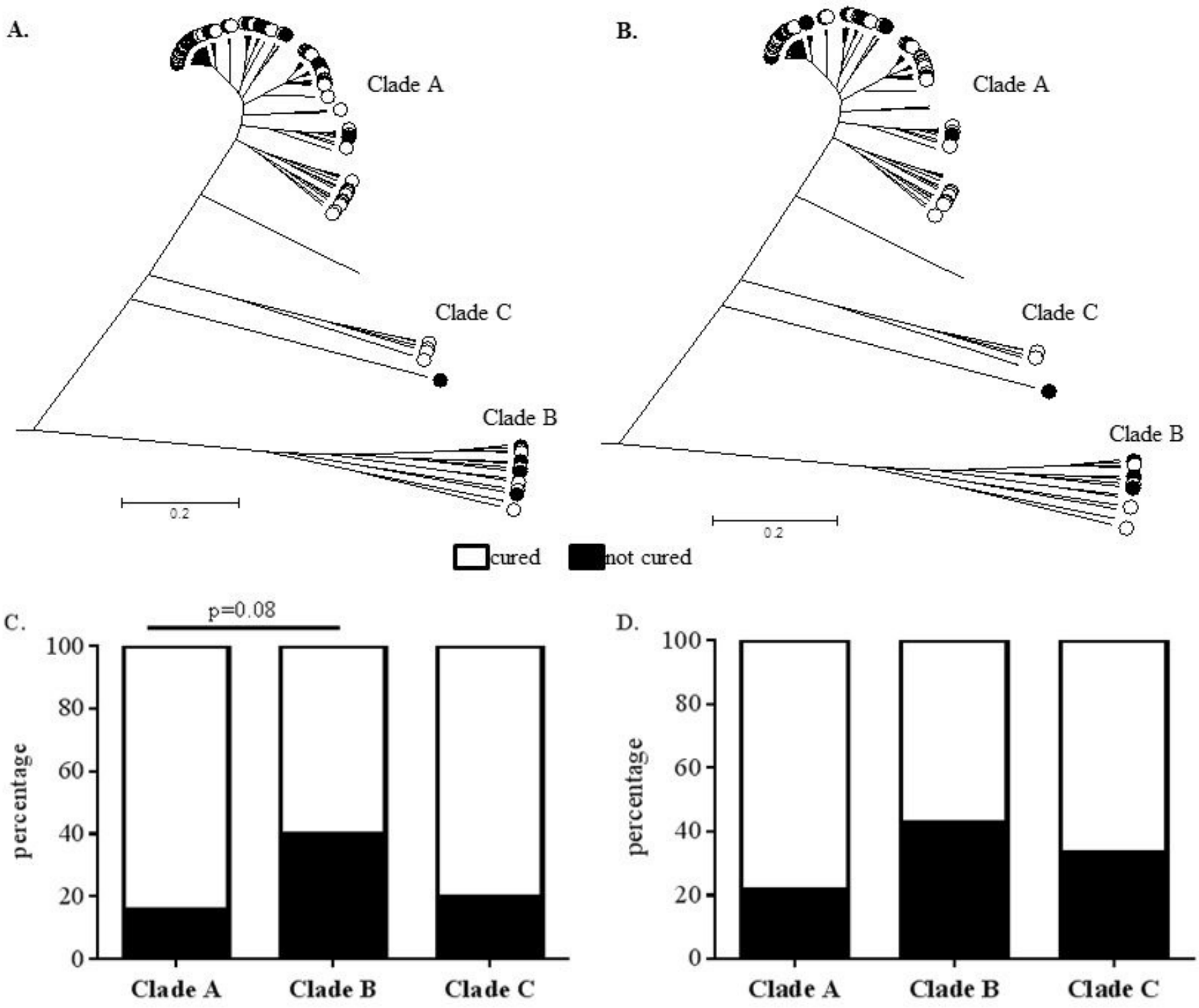
722 **Table 7. Relative over-representation of accessory genes in the "Not cured" outcome isolates (> 20% difference)**

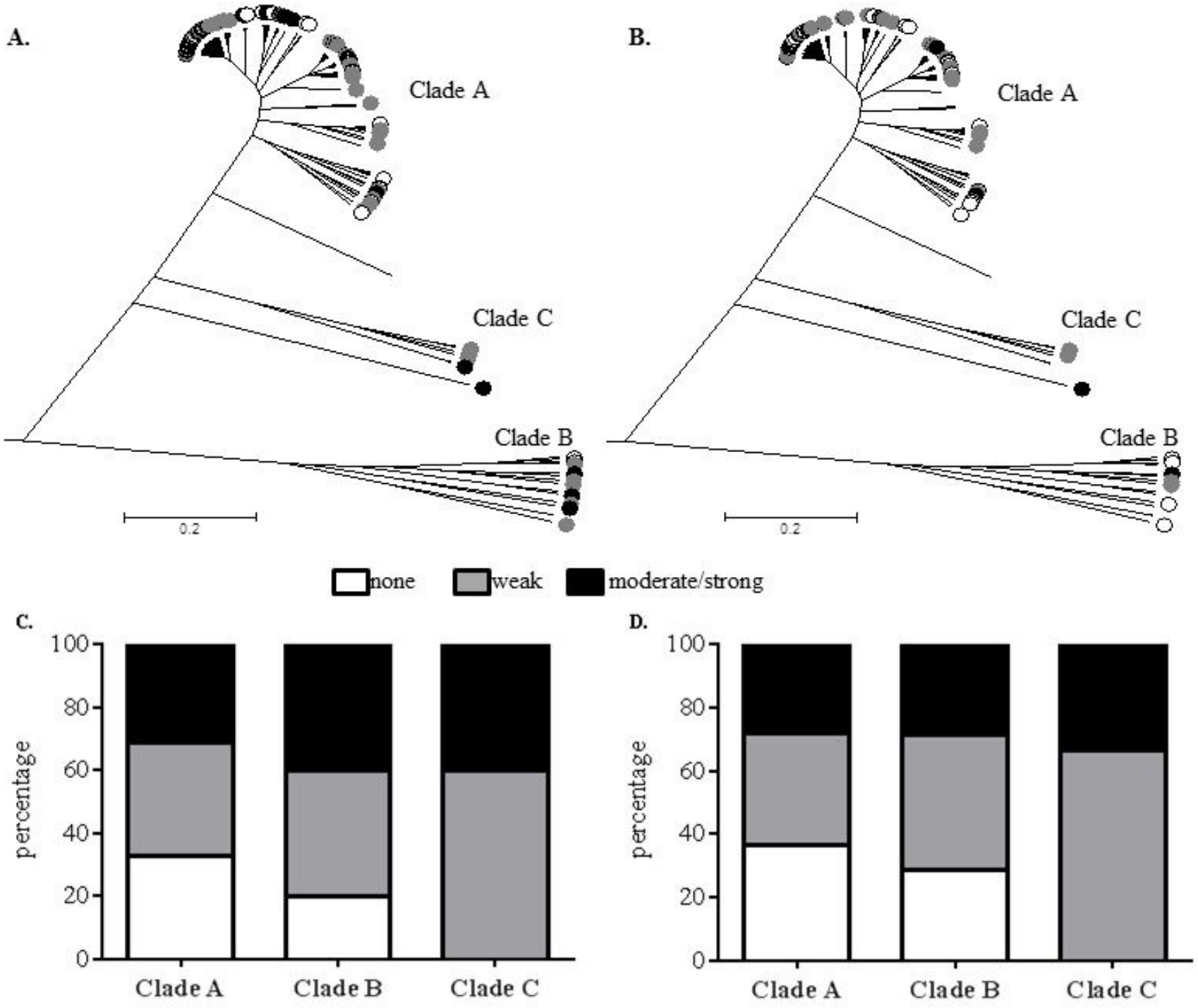
Loci	Description	No. isolates "Not cured" (n = 19)	Prevalence "Not cured" %	No. isolates "cured" (n = 85)	Prevalence "Cured"%	Difference % "Not cured" vs "Cured"	*p*-value
SERP0915	IS256-like transposase	11	57.9	31	36.5	21.4	0.085
SERP1222	Transposase	9	47.4	16	18.8	28.5	0.008
SERP1586	Acetyltransferase, GNAT family	9	47.4	22	25.8	21.5	0.064
SERP2498	Cassette chromosome recombinase A (*ccrA)*	17	89.5	56	65.8	23.6	0.042
SERP2499	Cassette chromosome recombinase B (*ccrB)*	17	89.5	55	64.7	24.7	0.034
id1043_1239	Hypothetical protein	10	52.6	25	29.4	23.3	0.053
id1044_0888	Phage protein	8	42.1	15	17.6	24.5	0.02
id1044_0895	Phage antirepressor protein	8	42.1	17	20	22.1	0.041
id1044_1909	Antiseptic resistance protein QacA	17	89.5	53	62.4	27.1	0.023
id1044_2610	unknown	8	42.1	15	17.6	24.5	0.02
id1048_0369	Replication-associated protein	11	57.9	29	34.1	23.7	0.054
id1632_0817	Zn-dependent hydroxyacylglutathione hydrolase	14	73.7	44	51.7	21.9	0.082

723

724


Presence of genes (%)
100
80
60
40
20
0
*
*
*
*
aap
atlE
bhp
embp
fbe
hld
icaA
PSM b-1
PSM b-2
PSM b-3
PSM b-4
sdrF
sdrG
sdrH
sesA
sesE
sesG
sesH
sesI
aacA
mecA
Total
"not cured"
"cured"


A.
Clade A
Clade C
Clade B
0.2
B.
Clade A
Clade C
Clade B
0.2
cured
not cured
C.
p=0.08
100
80
60
40
20
0
percentage
Clade A
Clade B
Clade C
D.
100
80
60
40
20
0
percentage
Clade A
Clade B
Clade C


A.
Clade A
Clade C
Clade B
0.2
B.
Clade A
Clade C
Clade B
0.2
none
weak
moderate/strong
C.
percentage
100
80
60
40
20
0
Clade A
Clade B
Clade C
D.
percentage
100
80
60
40
20
0
Clade A
Clade B
Clade C