REVISED MANUSCRIPT

Core genome conservation of *Staphylococcus haemolyticus* limits sequence based population structure analysis.

Jorunn Pauline Cavanagh [1*], Claus Klingenberg, [1,2] Anne-Merethe Hanssen [3], Elizabeth Aarag Fredheim[1], Patrice Francois [4], Jacques Schrenzel [4], Trond Flægstad [1,2] and Johanna Ericson Sollid [5*].

[1]Pediatric Research Group, Department of Clinical Medicine, Faculty of Health Science, University of Tromsø, Tromsø, Norway,

[2] Department of Paediatrics, University Hospital of North Norway, Tromsø, Norway,

[3] Department of Medical Biology, Faculty of Health Science, University of Tromsø, Tromsø, Norway [3]

[4] Genomic research laboratory, University of Geneva Hospitals, Geneva, Switzerland

[5] Research group for host-microbe interactions, Department of Medical Biology, University of Tromsø, Tromsø, Norway.

*Corresponding author.

Johanna Sollid, Research group for host-microbe interactions, Department of Medical Biology, University of Tromsø, 9037 Tromsø, Norway. Phone: +47 77644663. Fax: 47 77645350. E-mail: johanna.e.sollid@uit.no

Jorunn Pauline Cavanagh, Pediatric Research Group, Department of Medical Biology, University of Tromsø, 9037. Tromsø, Norway. Phone; + 47 77646950. Fax: 47 77645350. E-mail: pauline.cavanagh@uit.no

Abstract

The notoriously multi-resistant *Staphylococcus haemolyticus* is an emerging pathogen causing serious infections in immunocompromised patients. Defining the population structure is important to detect outbreaks and spread of antimicrobial resistant clones. Currently, the standard typing technique is pulsed-field gel electrophoresis (PFGE). In this study we describe novel molecular typing schemes for *S. haemolyticus* using multi locus sequence typing (MLST) and multi locus variable number of tandem repeats (VNTR) analysis. Seven housekeeping genes (MLST) and five VNTR loci (MLVF) were selected for the novel typing schemes. A panel of 45 human and veterinary *S. haemolyticus* isolates was investigated. The collection had diverse PFGE patterns (38 PFGE types) and was sampled over a 20 year-period from eight countries. MLST resolved 17 sequence types (Simpsons index of diversity [SID] = 0.877) and MLVF resolved 14 repeat types (SID = 0.831). We found a low sequence diversity. Phylogenetic analysis clustered the isolates in three (MLST) and one (MLVF) clonal complexes, respectively. Taken together, neither the MLST nor the MLVF scheme was suitable to resolve the population structure of this *S. haemolyticus* collection. Future MLVF and MLST schemes will benefit from addition of more variable core genome sequences identified by comparing different fully sequenced *S. haemolyticus* genomes.

1. Introduction

Staphylococcus haemolyticus belongs to the group of coagulase-negative staphylococci (CoNS) and is part of the human normal flora of skin and mucous membranes. It is also an opportunistic pathogen and the second most frequently CoNS isolated from human blood cultures (Falcone et al., 2006). *S. haemolyticus* is primarily associated with infections in immunocompromised patients, e.g. patients with haematological disease and immature infants (Nouri et al., 2008). The ability to produce biofilm and the notoriously multi-resistance to antimicrobial agents, including glycopeptides, favours *S. haemolyticus* as an emerging cause of nosocomial infections (de Allori et al., 2006, Falcone, et al., 2006, Fredheim et al., 2009, Froggatt et al., 1989, Hiramatsu, 1998, Hope et al., 2008, Koksal et al., 2009, Schwalbe et al., 1987).

Reliable phenotypic species identification of *S. haemolyticus* is challenging (Shittu et al., 2004). Misidentification, or failure of identification of *S. haemolyticus* by conventional biochemical methods has been reported (De Paulis et al., 2003). This observation might result from structural rearrangements in the chromosome due to the presence of IS elements (Watanabe et al., 2007). Matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) has recently proven to provide a reliable and rapid tool for identification of *Staphylococcus* species (Benagli et al., 2011, Dubois et al., 2010). In a comparative study of the genomes of *S. haemolyticus* (JCSC 1435), *S. epidermidis* and *S. aureus* an average sequence identity of 78% in genes found as orthologues were detected (Takeuchi et al., 2005). In particular, the *oriC* environ contained regions common for all three species (e.g. the staphylococcal cassette chromosome -SCC) but also regions unique to

each species. Sequence similarity between resistance genes suggests that resistance determinants are readily transferred between these staphylococcal species (Froggatt, et al., 1989). When comparing different *S. haemolyticus* isolates, large scale chromosomal inversions in the *oriC* environ were reported (Watanabe, et al., 2007).

Molecular typing methods are mandatory for population structure analyses in both local and global settings. Defining the population structure and dynamics is important to detect both outbreaks of pathogenic strains as well as the establishment and spread of antimicrobial resistant clones. Feasibility of molecular typing methods depends on discriminatory power, possibility for inter-laboratory comparison and laboriousness. The current molecular typing method available for *S. haemolyticus* is genome restriction fragment pattern analysis after pulsed field gel electrophoresis (PFGE) (Ben Saida et al., 2009, Burnie et al., 1997, Tabe et al., 1998). PFGE is considered a very useful method for short term investigation of an outbreak situation. However, PFGE is labour intensive and inter-laboratory comparisons of results are difficult to achieve due to technical differences and subjective interpretation of band patterns (Murchan et al., 2003, te Witt et al., 2010, Tenover et al., 1995).

Molecular population studies of pathogenic strains using multi locus sequence typing (MLST) utilize genetic diversity based on changes in relative slowly evolving housekeeping genes. The variation observed is generally due to point mutations and/or recombination (Pérez-Losada et al., 2006). Isolates with identical profiles are grouped as related, or clonal. Information of changes introduced to the slowly evolving housekeeping genes are used to describe patterns of evolution and global spread.

Multi locus variable number of tandem repeats (VNTR) analysis (MLVF) takes advantage of variation in repetitive DNA, which is found at multiple loci in most bacteria. The individual pattern of repeat units and sequence heterogeneity is a useful phylogenetic marker. Strain relatedness is based on varying number of tandem repeats and found to be an appropriate tool for investigation of short term bacterial evolution and epidemiological typing (van Belkum, 1999). Compared to PFGE and MLST, MLVF is an attractive typing method due to its simplicity, rapidity and high discriminating power (Francois et al., 2008, Francois et al., 2005, Lindstedt, 2005).

This work aimed to find a molecular typing method with a discriminatory power suitable for molecular epidemiology analyses of clinical isolates of *S. haemolyticus*, in order to answer basic questions concerning the population structure. In this report we describe the development of a MLST and a MLVF scheme, and the observation of a conserved core genome in *S. haemolyticus* (Koksal, et al., 2009).

2. Materials and methods

2.1 Strain collection

A total of 172 *S. haemolyticus* isolates were obtained from national and international collaborators. The isolates were collected during the period 1989 to 2010. The collection comprised 164 human clinical isolates (isolated in connection with clinical diagnostics), four human community acquired isolates and four isolates of veterinary clinical origin. The isolates were defined as community acquired if they were recovered within 48 hours of

hospitalisation or isolated from healthy individuals without prior hospitalisation the past year (Kaplan et al., 2005). Geographically the isolates originated from Norway (n=74), Switzerland (n=50), Japan (n=17), Germany (n=13), United Kingdom (n=12), Spain (n=3), Belgium (n=2) and Greece (n=1).

2.2 PFGE

All 172 isolates were typed by PFGE using a previously described method (Hanssen et al., 2004). The PFGE patterns were analyzed using Gel Compar software version 2.5 (Applied Maths, Ghent, Belgium). The Dice band-based similarity coefficient was calculated with a band position tolerance of 1.0%. The overall genetic relationship was determined creating a dendrogram by the unweighted pair group method with arithmetic means (UPGMA) logarithm. The isolates were assigned to different groups, where groups were defined as two or more isolates with >80% similarity (Carrico et al., 2005). The discriminatory ability of the novel MLST and MLVF schemes was calculated on a restricted collection of diverse isolates (n=45). Selection criteria were, different PFGE profiles, temporal spread and different geographic origin (Figure 1). In order to study possible geographic related clones we selected a small collection of isolates from the same geographic origin. In addition we also selected some isolates with similar PFGE band patterns. We also included veterinary and community acquired isolates in order to further evaluate the discriminatory ability. The selected isolates were investigated further as outlined below.

2.3 Species identification

Species identification was reconfirmed using a polyphasic approach. First by Gram staining, catalase test and coagulation assay by Staphaureux plus® (BioMerieux, Marcy l'Etoile, France) followed by partial 16S rRNA gene or *rpoB*gene sequencing (Drancourt and Raoult, 2002, Pettersson et al., 1997).

2.4 Antimicrobial susceptibility testing

Antimicrobial susceptibility testing to penicillin, gentamicin, erythromycin, tetracycline, vancomycin, rifampicin, and oxacillin was performed using Etest according to the manufacturer's description (AB BIODISK, Solna, Sweden). The antimicrobial breakpoints were interpreted according to the EUCAST guidelines (EUCAST, 2011).

2.5 Biofilm quantification

The biofilm producing ability of the isolates was determined by a semi-quantitative assay as described previously (Christensen et al., 1985, Klingenberg et al., 2005). Briefly, overnight cultures were diluted 1:100 in Tryptic Soy Broth (TSB, Becton Dickinson, Puls AS, Norway) with 1% glucose and incubated for 24 hours at 37°C in polystyrene microtiter plates (Nunclon, Roskilde, Denmark). The biofilm was washed 3x in phosphate buffered saline (PSB), fixed at 55°C for one hour and stained with crystal violet. Before detection the stain was dissolved with an ethanol/acetone (70:30) mixture. Optical density (OD) was measured in an ELISA reader, and isolates with an $OD_{570} \geq 0.25$ were defined as biofilm positive. *S.*

159 *epidermidis* RP62A was included as a positive control and *S. haemolyticus* 51-03 was
160 included as a negative control (Fredheim, et al., 2009, Yang et al., 2006).
161
162 2.6 DNA isolation, PCR conditions and sequencing
163
164 Template DNA was prepared by boiling, as previously described (Hanssen, et al., 2004).
165 Purified DNA was stored at -20 °C. PCRs for MLST and MLVF were performed with 25 µl
166 reaction volumes, comprising 0.4 pmol/sample of each primer, 3 µl template DNA and 12.5
167 µl of ReddyMix (Cat. no. AB-0815, ABgene, Surrey, UK). $MgCl_2$ was added to a final
168 concentration of 4.5 mM. MLST and MLVF PCRs were performed as previously described
169 (Francois, et al., 2008, Thomas et al., 2007), apart from the MLVF PCR annealing
170 temperature which was set to 55 °C. Cycle sequencing of both strands was performed as
171 previously described using the Big Dye Terminator (version 3.1) cycle sequencing kit
172 (Applied Biosystems, Warrington, UK) and analyzed on an ABI Prism 377 sequence
173 analyzer.
174
175 2.7 Design of a novel MLST scheme for *S. haemolyticus*
176
177 Internal segments of 18 genes were initially tested on five geographically diverse *S.*
178 *haemolyticus* isolates in order to find appropriate variability for the MLST scheme. The 18
179 genes tested were, i) equivalents of six of the seven loci used in the *S. epidermidis* MLST
180 scheme (*arc, aroE, gtr, mutS, pyrR, tpi*) (Thomas, et al., 2007), ii) *glp* from the *S. aureus*
181 MLST scheme (Enright et al., 2000) iii) equivalents of additional loci with reported higher

sequence divergence than the traditional MLST genes studied in *S. aureus* (*pbpB, leuB, hemH, luxS,* SH2038, SH1200, SH0328) (Cooper and Feil, 2006) and iv) four additional genes *Ribose ABC*, SH 1431, *cfxE* and SH 0871 selected from *S. haemolyticus* JCSC 1435 (Takeuchi, et al., 2005). Equivalents of *Ribose ABC* and SH 1431 were not found in the genomes of *S. epidermidis* and *S. aureus* based on comparative basic local alignment search tool (BLAST) (Altschul SF, 1990) searches. For the genes selected from the *S. epidermidis* /*S. aureus* MLST-schemes, equivalent primers were designed from the published genome of JSCS 1435 (accession number AP006716) (Takeuchi, et al., 2005). The seven gene segments that gave the highest variability were used to perform MLST on the 45 selected isolates. The primers used in the final MLST are listed in Table 1. Isolate 5MB 278-10 was excluded from the MLST analysis due to failure in amplification of one of the target genes.

2.7.1 DNA Sequence analysis

The nucleotide sequences were aligned by using Bio Edit sequence alignment editor (version 7.0.9.0) (Hall, 1999) and compared to the published sequence of JCSC 1435 in the GenBank database by using BLAST.

2.7.2 Phylogenetic analysis

Each of the selected isolates was defined by a seven digit allelic profile where each unique allelic profile defines a sequence type (ST). eBURST V3 (http://eburst.mlst.net) was used to determine the most putative relationship between isolates (Feil et al., 2004, Spratt BG, 2004).

Clonal complexes (CC) were defined using the default setting where STs that have diversified recently from a common founder and share six of seven alleles with at least one other ST in the group, are grouped in a clonal complex (Feil et al., 2003).

All analyses were performed using Molecular Evolutionary Genetics Analysis (MEGA) 4 (Tamura K, 2007). Neighbour joining (NJ) dendrograms for the individual MLST loci were created and maximum likelihood (ML) phylogentic trees were constructed for the concatenated MLST sequences of six of the seven loci *(hemh, cfxE, Ribose ABC, SH 1431, leuB* and *SH 1200)* using the general time reversible (GTR) model with 2000 bootstrap resampling replications (Lanave C, 1984). The nucleotide diversity within the major and minor CC, defined by eBURST, was calculated.

2.8 Design of a novel MLVF scheme for *S. haemolyticus*

Tandem repeat regions were detected in the published genome of JSCS 1435 (accession number AP006716) using the tandem repeats finder (http://tandem.bu.edu/trf/trf.html) (Benson, 1999). The number of putative target genes was in total 45. Nine of them contained tandem repeats and were selected for the assay. Nine PCR primer pairs targeting conserved flanking regions of repeat containing genes (orfs SH 0326, SH 0326b, SH 0999, SH 0040, SH 0040b, SH 2426, SH 01184, SH 0324 and SH 1645) were designed using Jellyfish (version 1.3 Biowire). The nine primer pairs were initially tested on five *S. haemolyticus* isolates from diverse geographical origins to find appropriate variability for the MLVF scheme. Four of the primer pairs did not generate amplicons in all strains, the remaining five

primer pairs were used to perform MLVF on the 45 selected isolates. The primers used in the final MLVF scheme are listed in Table 2.

2.8.1 DNA analysis

The PCR products were separated on a 1% agarose gel (SeaKem ® LE, Takara) with 0, 5 x TBE (Tris-borate-EDTA) buffer for 50 min at 80 V/cm. MLVF bands were visualized on an UV transilluminator, photographed and scanned. The MLVF patterns were then visually evaluated using the criteria by Sabat *et al*. (Sabat et al., 2003). Two MLVF patterns differing by one or more bands were considered distinct types.

2.8.2 Population structure

Arbitrary numbers were assigned to the different MLVF band patterns observed. The combination of numbers gives a unique fingerprint tag, or repeat type (RT) number. The results were analyzed by using the eBURST V3 algorithm (Feil, et al., 2004) (http://eburst.mlst.net/). Clonal complexes were defined as RTs that have diversified recently from a common founder sharing four of five alleles with at least one other RT in the group.

2.9 Discriminatory ability and clustering concordance

Simpson's index of diversity (SID), indicating the probability of two strains sampled randomly from a population belonging to different types, was calculated to compare the discriminatory ability of MLST, MLVF and PFGE (Carrico et al., 2006, Grundmann et al.,

2001, Hunter and Gaston, 1988). Adjusted Rand (AR) indices were calculated to determine the overall concordance between the methods, corrected for the presence of chance agreement. The Wallace (W) coefficient was calculated to determine the probability that two isolates classified as the same type by one method would be classified as the same by using another typing method (Carrico, et al., 2006, Pinto et al., 2008). The concordance of the different typing techniques was calculated using the software described by (Carrico, et al., 2006) using the online tool (http://darwin.phyloviz.net/ComparingPartitions).

3.0 Results

3.1 Antimicrobial resistance and biofilm formation

Analyses of antimicrobial susceptibility and biofilm formation were included to find phenotypic similarities or differences between the isolates that could reflect genetic relationship. The results of antimicrobial susceptibility testing and the biofilm assay are presented in Figure 1. Forty of the 45 isolates displayed resistance to three or more antimicrobial agents tested and 18 were resistant to five different antimicrobial agents. Three isolates originating from Germany, Norway and the UK (MB 278-10, 2263 3461 and CN 1197) were susceptible to all antimicrobial agents tested and two isolates originating from the UK and Norway (51-72 and CN1138) were susceptible to all antimicrobial agents tested except tetracycline. Biofilm was formed by 30 of the 45 isolates according to our definition.

3.2 PFGE

The PFGE results are shown in Figure 1. Thirty eight separate PFGE types were defined among the 45 isolates. Among these 38 PFGE types there were six groups (A-F). The largest group (B) contained three isolates from Switzerland. The remaining five groups contained two isolates each; Group A (both UK), C (both UK), D (both Germany), E (from Norway and Greece) and F (both Belgium). The isolates that did not cluster in any defined group (n=32) were considered unrelated when using an 80% cut-off value.

3.3 MLST analysis

MLST of the 44 isolates resulted in 17 unique STs. eBURST grouped the isolates in one major group or clonal complex (CC), two minor CCs and six singletons. CC1 comprised 25 isolates (ST 1, 2, 3, 10 and 15), representing human clinical isolates from all eight countries included in the study and both veterinary isolates from Belgium. CC2 comprised eight isolates (ST 8, 9 and 14) from Japan and the UK including three of the community acquired non-clinical isolates from Japan and one isolate from the UK. CC3 comprised five isolates (ST 4 and 13) representing isolates from Spain, Norway and Switzerland. Six isolates (ST 6, 7, 11, 12, 16 and 17) were defined as singletons. The veterinary isolate 278-10 was not included in the eBURST analysis as no PCR product was obtainable for one of the alleles (Ribose ABC) in the MLST scheme. The MLST results are summarized in Figure 1.

3.4 MLVF analysis

296 We defined, by visual categorization of band patterns, fourteen unique RTs among the 45
297 isolates. eBURST grouped all isolates, except one of the veterinary isolates (2263-3461) in
298 one CC. Sixteen isolates originating from the UK, Norway, Switzerland, Japan and Greece
299 shared the same RT. One RT was a singleton. The MLVF results are summarized in Figure 1.
300
301 3.5 Phylogenetic analysis of MLST data
302
303 NJ dendrograms created for the individual genes used in the MLST scheme showed good
304 congruence (data not shown). All isolates except three (MB 278-10, 2263-3461 and CN 1197)
305 were grouped in one large cluster by all genes. Apart from *arcC* which grouped only one
306 isolate (CN 1197) differently. The ML tree based on the concatenate sequences of six genes,
307 excluding *arcC*, grouped the isolates in one large cluster (Figure 2). As for the NJ trees,
308 isolates MB 278-210, 2263-3461, and CN1197 were grouped separately supported by a 99%
309 bootstrap value. The global agreement between the evolutionary trees for the individual
310 MLST genes and the ML tree from the concatenated sequences suggests a low degree of
311 recombination. Comparison of the clustering obtained by eBURST and the ML tree also
312 showed a global agreement. Two minor clusters comparable to CC2 and CC3 defined by
313 eBURST were also defined in the ML tree but they were not supported by significant
314 bootstrap values (54 % and 41%; Figure 2) indicating that the clustering made by eBURST
315 might not be correct. Calculation of nucleotide diversity based on the concatenated sequences
316 within the three eBURST CCs shows a low nucleotide diversity of 0.00035, supporting the
317 uniform clustering of isolates.
318

319 3.6 Discriminatory power and clustering concordance of typing methods

320

321 The SID revealed that PFGE in our study had a higher discriminatory power than MLST and
322 MLVF (Table 3). The overall concordance (the probability that two methods cluster two
323 isolates similarly) of the different typing methods was low (Table 4). AR indices ranged from
324 0.029-0.084. The highest concordance was found between MLST and MLVF (AR = 0.084).
325 Wallace (W) coefficients were calculated to determine the directional agreement between the
326 typing methods. There was a low probability ($W = 0.333$) that two isolates with the same
327 PFGE type had the same MLST type. The directional agreement between PFGE and MLVF
328 was also low (W = 0.444). Finally, the probability of MLST to predict MLVF type and vice
329 versa was very low with a W = 0.254 and W = 0.186, respectively.

330

331

332

333

334 **4. Discussion**

335

336 The mainstay for studying molecular epidemiology of *S. haemolyticus* has been PFGE. To
337 our knowledge this is the first study reporting MLST and MLVF schemes for this species and
338 to compare these typing techniques with PFGE. The discriminatory ability of the suggested
339 MLST and MLVF schemes was assessed using a diverse collection of *S. haemolyticus*. Both
340 clinical human and veterinary isolates were included. Compared with PFGE, MLST and
341 MLVF had an inferior discriminatory ability. The MLST results may even suggest that all 45

342 *S. haemolyticus* isolates were closely related. However, we believe it is unlikely that these 45
343 isolates are clonally related due to their diverse geographic origin and temporal spread.
344 MLST discriminated well between the isolates of human origin and two of the
345 isolates of veterinary origin. Two veterinary isolates, originating from Norway (MB 278-10)
346 and Germany (2263-3461), displayed a high degree of variation compared to the human
347 isolates. In contrast, the two Belgian veterinary isolates clustered together with the human
348 clinical isolates. The Belgian veterinary isolates also grouped together with the human
349 clinical isolates when comparing susceptibility to antimicrobial agents, i.e. defined as multi-
350 resistant, whereas the Norwegian and German veterinary isolate were susceptible to all
351 antimicrobials tested. An unexpected relationship was found between one isolate from the
352 UK and three community acquired isolates from Japan which all were of the same ST.
353 Phylogenetic analysis of our MLST data indicates a clonal population structure as there is
354 global congruence between the ML tree from the concatenated MLST sequences and
355 between the individual gene trees in the MLST scheme where six out of seven trees grouped
356 the isolates similar to the concatenated ML tree. The isolates were grouped in one main
357 cluster, with three isolates forming a separate cluster. The main cluster was divided in two
358 smaller clusters comparable to the CC defined by eBURST. However, low bootstrap values
359 for the smaller clusters in the ML tree indicate that the CC identified by eBURST might not
360 be correct. The low nucleotide diversity value reflects the high degree of sequence
361 conservation and suggests low levels of recombination. *S. epidermidis* and *S. aureus*, two
362 species that are closely related to *S. haemolyticus*, clearly show a different population
363 evolution. MLST population analyses of *S. epidermidis* has shown an epidemic population
364 that evolves by recombination (Miragaia et al., 2007). Analysis of *S. aureus* MLST sequence

data reveals a more clonal population evolving mainly by point mutation (Feil, et al., 2003). Our MLST data might indicate that *S. haemolyticus* has a population evolution more comparable to *S. aureus*. However, some caution must be applied when interpreting these results as our analysis is based on a restricted number of isolates. Reports of low polymorphism in housekeeping genes resulting in limited discriminatory power of MLST has previously been reported for species such as *Salmonella enterica,* serovar Typhi, *Mycoplasma pneumonia* and *Escherichia coli* (Degrange et al., 2009, Dumke et al., 2003, Fakhr et al., 2005, Noller et al., 2003).

Molecular typing by MLVF has shown to effectively discriminate homogenous bacterial populations (Noller et al., 2003, Octavia and Lan, 2009). The application of MLVF for epidemiologic studies of *S. aureus* and *S. epidermidis* has previously shown a resolution comparable to PFGE and MLST (Francois, et al., 2008, Holmes et al., 2010, Pourcel et al., 2009). The tandem repeat loci selected for MLVF are believed to be more variable than housekeeping genes for MLST due to a more diversifying selective pressure (van Belkum et al., 1997). However, in the present study the MLVF scheme was not able to discriminate between isolates of different origin. MLVF resulted in 14 RTs compared to 17 MLST STs and 38 PFGE types. Using MLVF all isolates were grouped together in one CC, except one veterinary isolates. The selection of our strain collection is biased, based on isolates which differs by PFGE. This has previously been reported to affect the discriminatory ability of MLVF (Holmes, et al., 2010, Luczak-Kadlubowska et al., 2008). Furthermore, a better resolution might have been obtained if we had targeted more than five tandem repeat loci. The search for tandem repeat loci was restricted as only one fully sequenced genome of *Staphylococcus haemolyticus* is presently available for automatic search. We found 45

putative target genes, but most of these were duplicated, poorly reliable, too short or showed a number of repeat of only one. The initial 9 primer pairs selected were considered as the maximum available number of tandem repeats containing genes for *S. haemolyticus*.

However, previously published schemes using five tandem repeat loci in *Chlamydia abortus (Laroucau et al., 2009)*, *S. epidermidis* (Johansson et al., 2006), and *Salmonella enterica* (Lindstedt et al., 2004) have shown satisfactory discrimination. Other studies comparing MLVF to MLST have also shown a good concordance between type assignment made by the two methods (Malachowa et al., 2005). In contrast, Tenover *et.al* reported that MLVF can not be used to predict PFGE type (Tenover et al., 2007).

Different bacterial populations exhibit varying rates of genetic change. In populations where no or little recombination has taken place the population will appear as clonal whereas highly recombining strains will appear as non-clonal (Spratt and Maiden, 1999). A major challenge for molecular typing methods is to select molecular markers that are sufficiently diverse enabling identification of variants of closely related bacteria (Maiden, 2006). In the present, study only four of the 45 isolates was clustered together by all three methods and we found very low values for the AR and the Wallace coefficient. We believe that the low variability observed by MLST and MLVF reflects a high degree of core genome conservation in *S. haemolyticus*, indicating a low rate of recombination. A diversifying selection may instead be due to accumulation of point mutations. The lack of congruence between the typing methods can also be explained by different detection levels. PFGE displays variation found in the total genome, whereas MLST and MLVF reveal variation found in short fragments of the core genome.

The observed core genome conservation contradicts the previously reported genome plasticity of *S. haemolyticus* indicated by the rapid acquisition of resistance genes as well as phenotypic variability (Watanabe, et al., 2007). Sequencing of *S. haemolyticus* JCSC 1435 revealed a large proportion of IS elements which is believed to contribute to the large scale inversions and deletions observed in JCSC 1435, mostly associated with the *oriC* environ (Takeuchi, et al., 2005, Watanabe, et al., 2007). This region contains integrated copies of SCC and IS elements. If genetic diversity mainly depends on mobile genetic elements and rearrangements in discrete regions (e.g. *oriC* environ) the changes will be detected by PFGE but not by MLST and MLVF, as the selected genes used in the MLST and MLVF schemes are not located in the *oriC* environ.

The results from this study show that neither the MLST nor the MLVF scheme could resolve the population structure of the *S. haemolyticus* collection. We suggest that there is potential for MLST and MLVF as epidemiologic tools by inclusion of more variable genes, in order to increase their discriminatory power. However, comparative genome analyses and the possibility to detect genes with higher variation are limited by the fact that there currently still is only one fully sequenced genome published (Takeuchi, et al., 2005). Full genome sequence based analysis is now possible for bacterial populations exhibiting levels of nucleotide diversity too low for resolution by MLST (Baker et al., 2010). Further molecular studies, including deep sequencing of the entire bacterial genome, are needed to provide high-resolution spatial and genetic data on *S. haemolyticus* epidemiology.

Acknowledgements

We thank the following for kindly providing isolates to this study: Dr. Holger Rohde, Universitätsklinikum Hamburg-Eppendorf, Germany; Dr. Stefan Schwarz Friedrich- Loeffler - Institutt, Neustadt-Mariensee, Germany; Dr. Marianne Sunde, Veterinærinstituttet, Oslo, Norway; Dr.Russel Hope at the Bacteraemia Resistance Surveillance Programme (BSAC), GB; Dr. Teruyo Ito, Department of Microbiology and Infection Control Science, Juntendo University, Tokyo, Japan; Dr. Rafael Cantón, Microbiologia Hopital Ramón y Cajal, Madrid, Spain; Dr. Nuno Cerca, Centre of Biological Engineering, Institute of Biotechnology and Bioengineering. University of Minho, Portugal; Dr. Jerry Pier, Medicine, Microbiology and Molecular Genetics Brigham and Women's Hospital, Harvard Medical School, USA, and Dr. Wannes Vanderhaegen, Ghent University, Faculty of Veterinary Medicine, Department of Pathology, Bacteriology and Poultry diseases, Belgium and Sandie Rodrigo-Humair, Genomic research laboratory, University of Geneva Hospitals, Geneva, Switzerland. We also thank Ed Feil for critically commenting on results and manuscript.

REFERENCES

Altschul SF, G.W., Miller W, Myers EW, Lipman DJ, 1990. Basic local alignment search tool. Journal of Molecular Biology. Oct 5;215, 403-410.

Baker, S., Hanage, W.P., Holt, K.E., 2010. Navigating the future of bacterial molecular epidemiology. Current Opinion in Microbiology. 13, 640-645.

Ben Saida, N., Marzouk, M., Ferjeni, A., Boukadida, J., 2009. A three-year surveillance of nosocomial infections by methicillin-resistant Staphylococcus haemolyticus in newborns reveals the disinfectant as a possible reservoir. Pathologie Biologie. 57, e29-e35.

Benagli, C., Rossi, V., Dolina, M., Tonolla, M., Petrini, O., 2011. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for the Identification of Clinically Relevant Bacteria. PLoS ONE. 6, e16424.

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucl. Acids Res. 27, 573-580.

Burnie, J., Naderi-Nasab, M., Loudon, K., Matthews, R., 1997. An epidemiological study of blood culture isolates of coagulase- negative staphylococci demonstrating hospital-acquired infection. J. Clin. Microbiol. 35, 1746-1750.

Carrico, J.A., Pinto, F.R., Simas, C., Nunes, S., Sousa, N.G., Frazao, N., de Lencastre, H., Almeida, J.S., 2005. Assessment of Band-Based Similarity Coefficients for Automatic Type and Subtype Classification of Microbial Isolates Analyzed by Pulsed-Field Gel Electrophoresis. J. Clin. Microbiol. 43, 5483-5490.

Carrico, J.A., Silva-Costa, C., Melo-Cristino, J., Pinto, F.R., de Lencastre, H., Almeida, J.S., Ramirez, M., 2006. Illustration of a Common Framework for Relating Multiple Typing Methods by Application to Macrolide-Resistant Streptococcus pyogenes. J. Clin. Microbiol. 44, 2524-2532.

Christensen, G.D., Simpson, W.A., Younger, J.J., Baddour, L.M., Barrett, F.F., Melton, D.M., Beachey, E.H., 1985. Adherence of coagulase-negative staphylococci to plastic tissue culture plates: a quantitative model for the adherence of staphylococci to medical devices. J Clin Microbiol. 22, 996-1006.

Cooper, J.E., Feil, E.J., 2006. The phylogeny of Staphylococcus aureus - which genes make the best intra-species markers? Microbiology. 152, 1297-1305.

de Allori, M.C., Gaudioso, Jure, M.A., Romero, C., de Castillo, M.E.C., 2006. Antimicrobial Resistance and Production of Biofilms in Clinical Isolates of Coagulase-Negative Staphylococcus Strains. Biological & Pharmaceutical Bulletin. 29, 1592-1596.

De Paulis, A.N., Predari, S.C., Chazarreta, C.D., Santoianni, J.E., 2003. Five-Test Simple Scheme for Species-Level Identification of Clinically Significant Coagulase-Negative Staphylococci. J. Clin. Microbiol. 41, 1219-1224.

Degrange, S., Cazanave, C., Charron, A., Renaudin, H., Bebear, C., Bebear, C.M., 2009. Development of Multiple-Locus Variable-Number Tandem-Repeat Analysis for Molecular Typing of Mycoplasma pneumoniae. J. Clin. Microbiol. 47, 914-923.

Drancourt, M., Raoult, D., 2002. rpoB Gene Sequence-Based Identification of Staphylococcus Species. J. Clin. Microbiol. 40, 1333-1338.

493 Dubois, D., Leyssene, D., Chacornac, J.P., Kostrzewa, M., Schmit, P.O., Talon, R., Bonnet,
494 R., Delmas, J., 2010. Identification of a Variety of Staphylococcus Species by Matrix-
495 Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. Journal of Clinical
496 Microbiology. 48, 941-945.
497 Dumke, R., Catrein, I., Pirkl, E., Herrmann, R., Jacobs, E., 2003. Subtyping of Mycoplasma
498 pneumoniae isolates based on extended genome sequencing and on expression profiles.
499 International Journal of Medical Microbiology. 292, 513-525.
500 Enright, M.C., Day, N.P.J., Davies, C.E., Peacock, S.J., Spratt, B.G., 2000. Multilocus
501 Sequence Typing for Characterization of Methicillin-Resistant and Methicillin-Susceptible
502 Clones of Staphylococcus aureus. J. Clin. Microbiol. 38, 1008-1015.
503 EUCAST, T.E.C.o.A.S.T.-. 2011. Breakpoint tables for interpretation of MICs and zone
504 diameters v 1.3, EUCAST.
505 Fakhr, M.K., Nolan, L.K., Logue, C.M., 2005. Multilocus Sequence Typing Lacks the
506 Discriminatory Ability of Pulsed-Field Gel Electrophoresis for Typing Salmonella enterica
507 Serovar Typhimurium. J. Clin. Microbiol. 43, 2215-2219.
508 Falcone, M., Giannella, M., Raponi, G., Mancini, C., Venditti, M., 2006. Teicoplanin use and
509 emergence of *Staphylococcus haemolyticus*: is there a link? Clinical Microbiology and
510 Infection. 12, 96-97.
511 Feil, E.J., Cooper, J.E., Grundmann, H., Robinson, D.A., Enright, M.C., Berendt, T., Peacock,
512 S.J., Smith, J.M., Murphy, M., Spratt, B.G., Moore, C.E., Day, N.P.J., 2003. How Clonal Is
513 Staphylococcus aureus? J. Bacteriol. 185, 3307-3316.
514 Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. eBURST: Inferring
515 Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from
516 Multilocus Sequence Typing Data. J. Bacteriol. 186, 1518-1530.
517 Francois, P., Hochmann, A., Huyghe, A., Bonetti, E.-J., Renzi, G., Harbarth, S., Klingenberg,
518 C., Pittet, D., Schrenzel, J., 2008. Rapid and high-throughput genotyping of Staphylococcus
519 epidermidis isolates by automated multilocus variable-number of tandem repeats: A tool for
520 real-time epidemiology. Journal of Microbiological Methods. 72, 296-305.
521 Francois, P., Huyghe, A., Charbonnier, Y., Bento, M., Herzig, S., Topolski, I., Fleury, B.,
522 Lew, D., Vaudaux, P., Harbarth, S., van Leeuwen, W., van Belkum, A., Blanc, D.S., Pittet,
523 D., Schrenzel, J., 2005. Use of an Automated Multiple-Locus, Variable-Number Tandem
524 Repeat-Based Method for Rapid and High-Throughput Genotyping of Staphylococcus aureus
525 Isolates. J. Clin. Microbiol. 43, 3346-3355.
526 Fredheim, E.G.A., Klingenberg, C., Rohde, H., Frankenberger, S., Gaustad, P., Flægstad, T.,
527 Sollid, J.E., 2009. Biofilm formation by Staphylococcus haemolyticus. J. Clin. Microbiol.,
528 JCM.01891-01808.
529 Froggatt, J.W., Johnston, J.L., Galetto, D.W., Archer, G.L., 1989. Antimicrobial resistance in
530 nosocomial isolates of Staphylococcus haemolyticus. Antimicrob. Agents Chemother. 33,
531 460-466.
532 Grundmann, H., Hori, S., Tanner, G., 2001. Determining Confidence Intervals When
533 Measuring Genetic Diversity and the Discriminatory Abilities of Typing Methods for
534 Microorganisms. J. Clin. Microbiol. 39, 4190-4192.
535 Hall, T.A., 1999. Bio Edit:a user -friendly biological sequence alignment editor and analysis
536 program for windows 95/98/NT, Nucl.acids.Symp.ser. 41:95-98.
537 Hanssen, A.-M., Kjeldsen, G., Sollid, J.U.E., 2004. Local Variants of Staphylococcal
538 Cassette Chromosome mec in Sporadic Methicillin-Resistant Staphylococcus aureus and

539 Methicillin-Resistant Coagulase-Negative Staphylococci: Evidence of Horizontal Gene
540 Transfer? Antimicrob. Agents Chemother. 48, 285-296.
541 Hiramatsu, K., 1998. Vancomycin resistance in staphylococci. Drug Resistance Updates. 1,
542 135-150.
543 Holmes, A., Edwards, G.F., Girvan, E.K., Hannant, W., Danial, J., Fitzgerald, J.R.,
544 Templeton, K.E., 2010. Comparison of Two Multilocus Variable-Number Tandem-Repeat
545 Methods and Pulsed-Field Gel Electrophoresis for Differentiating Highly Clonal Methicillin-
546 Resistant Staphylococcus aureus Isolates. J. Clin. Microbiol. 48, 3600-3607.
547 Hope, R., Livermore, D.M., Brick, G., Lillie, M., Reynolds, R., 2008. Non-susceptibility
548 trends among staphylococci from bacteraemias in the UK and Ireland, 2001–06. Journal of
549 Antimicrobial Chemotherapy. 62, ii65-ii74.
550 Hunter, P.R., Gaston, M.A., 1988. Numerical index of the discriminatory ability of typing
551 systems: an application of Simpson's index of diversity. J. Clin. Microbiol. 26, 2465-2466.
552 Johansson, A., Koskiniemi, S., Gottfridsson, P., Wistrom, J., Monsen, T., 2006. Multiple-
553 Locus Variable-Number Tandem Repeat Analysis for Typing of Staphylococcus epidermidis.
554 J. Clin. Microbiol. 44, 260-265.
555 Kaplan, S.L., Hulten, K.G., Gonzalez, B.E., Hammerman, W.A., Lamberth, L., Versalovic, J.,
556 Mason, E.O., 2005. Three-Year Surveillance of Community-AcquiredStaphylococcus aureus
557 Infections in Children. Clinical Infectious Diseases. 40, 1785-1791.
558 Klingenberg, C., Aarag, E., Ronnestad, A., Sollid, J.E., Abrahamsen, T.G., Kjeldsen, G.,
559 Flaegstad, T., 2005. Coagulase-negative staphylococcal sepsis in neonates. Association
560 between antibiotic resistance, biofilm formation and the host inflammatory response. Pediatr
561 Infect Dis J. 24, 817-822.
562 Koksal, F., Yasar, H., Samasti, M., 2009. Antibiotic resistance patterns of coagulase-negative
563 staphylococcus strains isolated from blood cultures of septicemic patients in Turkey.
564 Microbiological Research. 164, 404-410.
565 Lanave C, P.G., Saccone C, Serio G, 1984. A new method for calculating evolutionary
566 substitution rates. Journal of Molecular Evolution. 20, 86-93.
567 Laroucau, K., Vorimore, F., Bertin, C., Yousef Mohamad, K., Thierry, S., Hermann, W.,
568 Maingourd, C., Pourcel, C., Longbottom, D., Magnino, S., Sachse, K., Vretou, E., Rodolakis,
569 A., 2009. Genotyping of Chlamydophila abortus strains by multilocus VNTR analysis.
570 Veterinary Microbiology. 137, 335-344.
571 Lindstedt, B.-A., 2005. Multiple-locus variable number tandem repeats analysis for genetic
572 fingerprinting of pathogenic bacteria. Electrophoresis. 26, 2567-2582.
573 Lindstedt, B.-A., Vardund, T., Aas, L., Kapperud, G., 2004. Multiple-locus variable-number
574 tandem-repeats analysis of Salmonella enterica subsp. enterica serovar Typhimurium using
575 PCR multiplexing and multicolor capillary electrophoresis. Journal of Microbiological
576 Methods. 59, 163-172.
577 Luczak-Kadlubowska, A., Sabat, A., Tambic-Andrasevic, A., Payerl-Pal, M., Krzyszton-
578 Russjan, J., Hryniewicz, W., 2008. Usefulness of Multiple-Locus VNTR Fingerprinting in
579 detection of clonality of community- and hospital-acquired ;Staphylococcus aureus isolates.
580 Antonie van Leeuwenhoek. 94, 543-553.
581 Maiden, M.C.J., 2006. Multilocus Sequence Typing of Bacteria. Annual Review of
582 Microbiology. 60, 561-588.
583 Malachowa, N., Sabat, A., Gniadkowski, M., Krzyszton-Russjan, J., Empel, J.,
584 Miedzobrodzki, J., Kosowska-Shick, K., Appelbaum, P.C., Hryniewicz, W., 2005.

585 Comparison of multiple-locus variable-number tandem-repeat analysis with pulsed-field gel
586 electrophoresis, spa typing, and multilocus sequence typing for clonal characterization of
587 Staphylococcus aureus isolates. J Clin Microbiol. 43, 3095 - 3100.
588 Miragaia, M., Thomas, J.C., Couto, I., Enright, M.C., de Lencastre, H., 2007. Inferring a
589 Population Structure for Staphylococcus epidermidis from Multilocus Sequence Typing Data.
590 J. Bacteriol. 189, 2540-2552.
591 Murchan, S., Kaufmann, M.E., Deplano, A., de Ryck, R., Struelens, M., Zinn, C.E., Fussing,
592 V., Salmenlinna, S., Vuopio-Varkila, J., El Solh, N., Cuny, C., Witte, W., Tassios, P.T.,
593 Legakis, N., van Leeuwen, W., van Belkum, A., Vindel, A., Laconcha, I., Garaizar, J.,
594 Haeggman, S., Olsson-Liljequist, B., Ransjo, U., Coombes, G., Cookson, B., 2003.
595 Harmonization of Pulsed-Field Gel Electrophoresis Protocols for Epidemiological Typing of
596 Strains of Methicillin-Resistant Staphylococcus aureus: a Single Approach Developed by
597 Consensus in 10 European Laboratories and Its Application for Tracing the Spread of
598 Related Strains. J. Clin. Microbiol. 41, 1574-1585.
599 Noller, A.C., McEllistrem, M.C., Pacheco, A.G.F., Boxrud, D.J., Harrison, L.H., 2003.
600 Multilocus Variable-Number Tandem Repeat Analysis Distinguishes Outbreak and Sporadic
601 Escherichia coli O157:H7 Isolates. J. Clin. Microbiol. 41, 5389-5397.
602 Noller, A.C., McEllistrem, M.C., Stine, O.C., Morris, J.G., Jr., Boxrud, D.J., Dixon, B.,
603 Harrison, L.H., 2003. Multilocus Sequence Typing Reveals a Lack of Diversity among
604 Escherichia coli O157:H7 Isolates That Are Distinct by Pulsed-Field Gel Electrophoresis. J.
605 Clin. Microbiol. 41, 675-679.
606 Nouri, L.B.Z., Caitriona, M.G., Fitzgerald, J.R., 2008. Pathogenomics of the staphylococci:
607 insights into niche adaptation and the emergence of new virulent strains. FEMS
608 Microbiology Letters. 289, 1-12.
609 Octavia, S., Lan, R., 2009. Multiple-Locus Variable-Number Tandem-Repeat Analysis of
610 Salmonella enterica Serovar Typhi. J. Clin. Microbiol. 47, 2369-2376.
611 Pérez-Losada, M., Browne, E.B., Madsen, A., Wirth, T., Viscidi, R.P., Crandall, K.A., 2006.
612 Population genetics of microbial pathogens estimated from multilocus sequence typing
613 (MLST) data. Infection, Genetics and Evolution. 6, 97-112.
614 Pettersson, B., Andersson, A., Leitner, T., Olsvik, O., Uhlen, M., Storey, C., Black, C., 1997.
615 Evolutionary relationships among members of the genus Chlamydia based on 16S ribosomal
616 DNA analysis. J. Bacteriol. 179, 4195-4205.
617 Pinto, F.R., Melo-Cristino, J., Ramirez, M., 2008. A Confidence Interval for the Wallace
618 Coefficient of Concordance and Its Application to Microbial Typing Methods. PLoS ONE. 3,
619 e3696.
620 Pourcel, C., Hormigos, K., Onteniente, L., Sakwinska, O., Deurenberg, R.H., Vergnaud, G.,
621 2009. Improved MLVA assay for Staphylococcus aureus providing a highly informative
622 genotyping technique together with strong phylogenetic value. J. Clin. Microbiol.,
623 JCM.3121-3128.
624 Sabat, A., Krzyszton-Russjan, J., Strzalka, W., Filipek, R., Kosowska, K., Hryniewicz, W.,
625 Travis, J., Potempa, J., 2003. New Method for Typing Staphylococcus aureus Strains:
626 Multiple-Locus Variable-Number Tandem Repeat Analysis of Polymorphism and Genetic
627 Relationships of Clinical Isolates. J. Clin. Microbiol. 41, 1801-1804.
628 Schwalbe, R.S., Stapleton, J.T., Gilligan, P., 1987. Emergence of vancomycin-resistance in
629 coagulase-negative staphylococci. N Engl J Med. 316, 927 - 931.

630 Shittu, A., Lin, J., Morrison, D., Kolawole, D., 2004. Isolation and molecular
631 characterization of multiresistant Staphylococcus sciuri and Staphylococcus haemolyticus
632 associated with skin and soft-tissue infections. J Med Microbiol. 53, 51-55.
633 Spratt BG, H.W., Li B, Aanensen DM, Feil EJ, 2004. Displaying the relatedness among
634 isolates of bacterial species -- the eBURST approach. FEMS Microbiol Lett. Dec 15, 129-
635 134.
636 Spratt, B.G., Maiden, M.C.J., 1999. Bacterial population genetics, evolution and
637 epidemiology. Philosophical Transactions of the Royal Society of London. Series B:
638 Biological Sciences. 354, 701-710.
639 Tabe, Y., Nakamura, A., Oguri, T., Igari, J., 1998. Molecular characterization of epidemic
640 multiresistant Staphylococcus haemolyticus isolates. Diagnostic Microbiology and Infectious
641 Disease. 32, 177-183.
642 Takeuchi, F., Watanabe, S., Baba, T., Yuzawa, H., Ito, T., Morimoto, Y., Kuroda, M., Cui, L.,
643 Takahashi, M., Ankai, A., Baba, S., Fukui, S., Lee, J.C., Hiramatsu, K., 2005. Whole-
644 genome sequencing of staphylococcus haemolyticus uncovers the extreme plasticity of its
645 genome and the evolution of human-colonizing staphylococcal species. J Bacteriol. 187,
646 7292-7308.
647 Tamura K, D.J., Nei M & Kumar 2007. MEGA4: Molecular Evolutionary Genetics Analysis
648 (MEGA) software version 4.0. , Molecular Biology and Evolution 24:1596-1599,
649 (Publication PDF at http://www.kumarlab.net/publications)
650 te Witt, R., van Belkum, A., MacKay, W., Wallace, P., van Leeuwen, W., 2010. External
651 quality assessment of the molecular diagnostics and genotyping of meticillin-resistant
652 Staphylococcus aureus. European Journal of Clinical Microbiology & Infectious Diseases. 29,
653 295-300.
654 Tenover, F., Arbeit, R., Goering, R., Mickelsen, P., Murray, B., Persing, D., Swaminathan,
655 B., 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel
656 electrophoresis: criteria for bacterial strain typing. J Clin Microbiol. 33, 2233 - 2239.
657 Tenover, F.C., Vaughn, R.R., McDougal, L.K., Fosheim, G.E., McGowan, J.E., Jr., 2007.
658 Multiple-Locus Variable-Number Tandem-Repeat Assay Analysis of Methicillin-Resistant
659 Staphylococcus aureus Strains. J. Clin. Microbiol. 45, 2215-2219.
660 Thomas, J.C., Vargas, M.R., Miragaia, M., Peacock, S.J., Archer, G.L., Enright, M.C., 2007.
661 Improved Multilocus Sequence Typing Scheme for Staphylococcus epidermidis. J. Clin.
662 Microbiol. 45, 616-619.
663 van Belkum, A., 1999. Short sequence repeats in microbial pathogenesis and evolution.
664 Cellular and Molecular Life Sciences. 56, 729-734.
665 van Belkum, A., Scherer, S., van Leeuwen, W., Willemse, D., van Alphen, L., Verbrugh, H.,
666 1997. Variable number of tandem repeats in clinical strains of Haemophilus influenzae.
667 Infect. Immun. 65, 5017-5027.
668 Watanabe, S., Ito, T., Morimoto, Y., Takeuchi, F., Hiramatsu, K., 2007. Precise Excision and
669 Self-Integration of a Composite Transposon as a Model for Spontaneous Large-Scale
670 Chromosome Inversion/Deletion of the Staphylococcus haemolyticus Clinical Strain
671 JCSC1435. J. Bacteriol. 189, 2921-2925.
672 Yang, X.-M., Li, N., Chen, J.-M., Ou, Y.-Z., Jin, H., Lu, H.-J., Zhu, Y.-L., Qin, Z.-Q., Qu, D.,
673 Yang, P.-Y., 2006. Comparative proteomic analysis between the invasive and commensal
674 strains of Staphylococcus epidermidis. FEMS Microbiology Letters. 261, 32-40.
675

676
677

678 **Table 1.** Genes and primer sequences used in the MLST scheme.

Gene loci	Primer sequence 5`→ 3`	Amplicon size (bp)	Ref.
Arc [a]	F AGTGACTCAAGTTGAA R AATCTTACCATCTAGG	520	This study
SH 1200 [b]	F CGGTAATGTAACACACGCAGT R TCTTCCTAGTAGCTGACCAG	450	This study
hemH [c]	F CTGATCGTCAAGCTGAAGCAT R GTACCTGTGTGACCCTCAGA	450	This study
leuB [d]	F AGCCATAGATTCGCATGGTGT R CCTAATGAACCTGGAATGGTAG	450	This study
SH 1431[e]	F TCAGACCAATTCCCAACC R CTTTAGCGTCACGATGGTCG	450	This study
cfxE [f]	F GAAGCACAAATTGATGGTCTGC R TCTGCCCCATTATCAACACA	450	This study
Ribose ABC	F GAGACGATTCAGCTAAGCAA R CGCCTTTCATTAGGCCATTA	450	This study

679 **[a] *arc*, carbamate kinase; [b] SH 1200, Ser A; D-3 -phosphoglycerate dehydrogenase; [c]**

680 **hemH, ferrochelatase; [d] leuB, 3- isopropymalate dehydrogenase; [e] SH 1431, cell surface**

681 **elastin binding protein; [f] cfxE, ribulose 5- phosphate epimerase.**

682

683 **Table 2.** Primers and repeat sequences used in the MLVF scheme

Orf	Repeat position [a]	Primer sequences (5`→3`)
SH0999	406--624	**SH0999_F** [b]
		CATCAATCTGATACCCAAGATTCAACTGAATTAG
		SH0999_R [c]
		TCCAGTGTCTGGTTTACCTGAATCATTG
SH0324	251--809	**SH0324_F**
		GATGCTTTTCAGCATAGCCA
		SH0324_R
		GGTCAACCAATTACATCCCA
SH1184	46-235	**SH1184_F**
		ATATAATCGCGACGCATTTG
		SH1184_R
		CAGCTGAACCGATTAAAGCA
SH1645	300--357	**SH1645_F**
		ATAATAACAAAAATAATGCCAAAA
		SH1645_R
		AGCTGCCGGTTTGTTATTTT
SH0326	2221--2575	**SH0326_F**
		CAAGTGCAAGCACATCATTG
		SH0326_R
		CTTGCACTTGTTGAATCGCT

684 **[a] location of the tandem repeats on the chromosome of *S. haemolyticus* JCSC1435.**

685 **[b] F, Forward primer,[c] R, reverse primer.**

686

687

688 **Table 3.** Discriminatory power of three molecular typing methods evaluated with 45 *S.*

689 *haemolyticus* isolates.

Method	No. of types	SID (95% CI)[a]	
PFGE	38	0.991	(0.983-0.999)
MLST	17	0.877	(0.813-0.940)
MLVF	14	0.831	(0.749-0.914)

690 **[a] Simpson's index of diversity (SID); CI, confidence interval**

691

692 **Table 4.** Concordance of PFGE, MLST and MLVF for the 45 *S. haemolyticus* isolates.

	Adjusted Rand			Wallace coefficient		
Methods	**PFGE**	**MLST**	**MLVF**	**PFGE**	**MLST**	**MLVF**
PFGE					0.333	0.444
MLST	0.029			0.025		0.254
MLVF	0.029	0.084		0.024	0.186	

693

694

695 **Figure legends, Fig. 1-2:**

696

697 **Fig. 1.**

698 Isolate information and type assignment made by PFGE, MLST and MLVF.

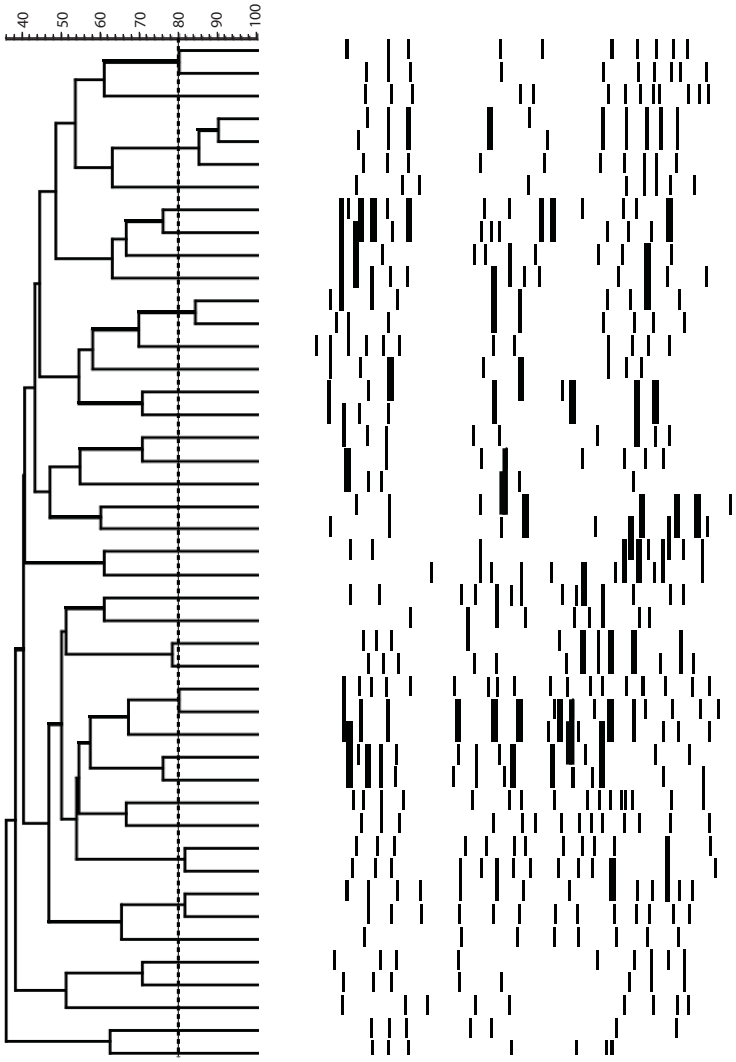
699

700 **Fig. 2.** ML dendrogram from the concatenated sequences of six MLST genes (SH 1200,

701 hemH, leuB, SH 1431, cfxE and Ribose ABC) for the 45 isolates included in the study

Figure(s)

Figure 1; Strain information and type assignment made by PFGE ,MLST and MLVA.



Isolates				PFGE		MLST			MLVA				
	Origin	Year	Source	Type	Group	ST	Alleles	CC	RT	Patterns	CC	Biofilm	Resistant
CN 1197	UK	2005	H1	1	A	7	2454433		2	21111	1	+[4]	S[6]
CN 1138	UK	2005	H	1	A	8	1511214	2	2	21111	1	+	S
6249	Germany	2007	H	2		1	2111114	1	2	21111	1	+	R[7]
115609	Switzerland	2004	H	3	B	4	1111222	3	14	11113	1	+	R
133319	Switzerland	2007	H	3	B	10	1511111	1	1	11111	1	-	R
103709	Switzerland	2002	H	3	B	3	1111114	1	2	21111	1	-	R
CN 1219	UK	2005	H	4		1	2111114	1	2	21111	1	+	R
6660	Germany	2008	H	5		1	2111114	1	7	21211	1	+	R
6035	Germany	2007	H	6		1	2111114	1	7	21211	1	-	R
2111	Germany	2008	H	7		15	2111154	1	1	11111	1	+	R
TUH 51-55	Norway	1989	H	8		2	1111111	1	1	11111	1	+	R
CN 1134	UK	2005	H	9	C	3	1111114	1	2	21111	1	+	R
CN 1175	UK	2005	H	9	C	3	1111114	1	2	21111	1	+	R
CN 1167	UK	2005	H	10		9	2511214	2	8	21511	1	+	R
T 621	Japan	NA[8]	H	11		14	1521214	2	2	21111	1	-	R
21116	Japan	NA	H	12		8	1511214	2	1	11111	1	-	R
CN 1011	UK	2005	H	13		1	2111114	1	2	21111	1	+	R
W 114	Japan	NA	CA[2]	14		8	1511214	2	5	11311	1	-	R
643	Germany	2008	H	15		1	2111114	1	2	21111	1	+	R
W 139	Japan	NA	CA	16		8	1511214	2	2	21111	1	-	R
2263-3461	Germany	2004	V[3]	17		6	4332335		4	31312		+	S
CN 1048	UK	2005	H	18		13	1111212	3	1	11111	1	-	R
08074328	Spain	2008	H	19		16	2111122		10	21811	1	+	R
097208	Switzerland	2001	H	20		13	1111212	3	1	11111	1	-	R
5MB 278-10	Norway	2003	V	21		5	544344x		5	11311	1	+	S
W 75	Japan	NA	CA	22		8	1511214	2	2	21111	1	-	R
JCSC 1435	Japan	2000	H	23		2	1111111	1	1	11111	1	+	R
M-176	USA	1988	H	24		11	3221211		6	21411	1	+	R
2139	Germany	2008	H	25	D	12	1525614		13	41111	1	+	R
7532	Germany	2007	H	25	D	1	2111114	1	2	21111	1	+	R
7589	Germany	2008	H	26		1	2111114	1	3	31111	1	+	R
137772	Switzerland	2008	H	27		1	2111114	1	9	21611	1	+	R
123477	Switzerland	2005	H	28		1	2111114	1	9	21611	1	+	R
TUH 51-34	Norway	1991	H	29		4	1111222	3	1	11111	1	+	R
07080750	Spain	2007	H	30		1	2111114	1	3	31111	1	+	R
TUH 51-33	Norway	2005	H	31	E	3	1111114	1	2	21111	1	+	R
TUH 51-57	Greece	2000	H	31	E	3	1111114	1	2	21111	1	+	R
S5	Belgium	2010	V	32	F	1	2111114	1	12	31611	1	-	R
AB	Belgium	2010	V	32	F	1	2111114	1	11	31211	1	-	R
TUH 51-50	Norway	1989	H	33		17	3561614		2	21111	1	+	S
F06	Japan	1995	H	34		9	2511214	2	8	21511	1	-	R
F045	Japan	1995	H	35		3	1111114	1	9	21611	1	-	R
51-72	Norway	2010	CA	36		3	1111114	1	11	31211	1	-	S
114564	Switzerland	2004	H	37		4	1111222	3	1	11111	1	+	R
104626	Switzerland	2002	H	38		3	1111114	1	9	21611	1	+	R

[1]human clinical isolates, [2]community acquired, [3]Veterinary isolates, [4]biofilm positive, [5]biofilm negative, [6]= sensitive to ≥ 4 antimicrobial agents, [7]= resistant to ≥ 3 more antimicrobial agents, [8]not available.

Figure(s)

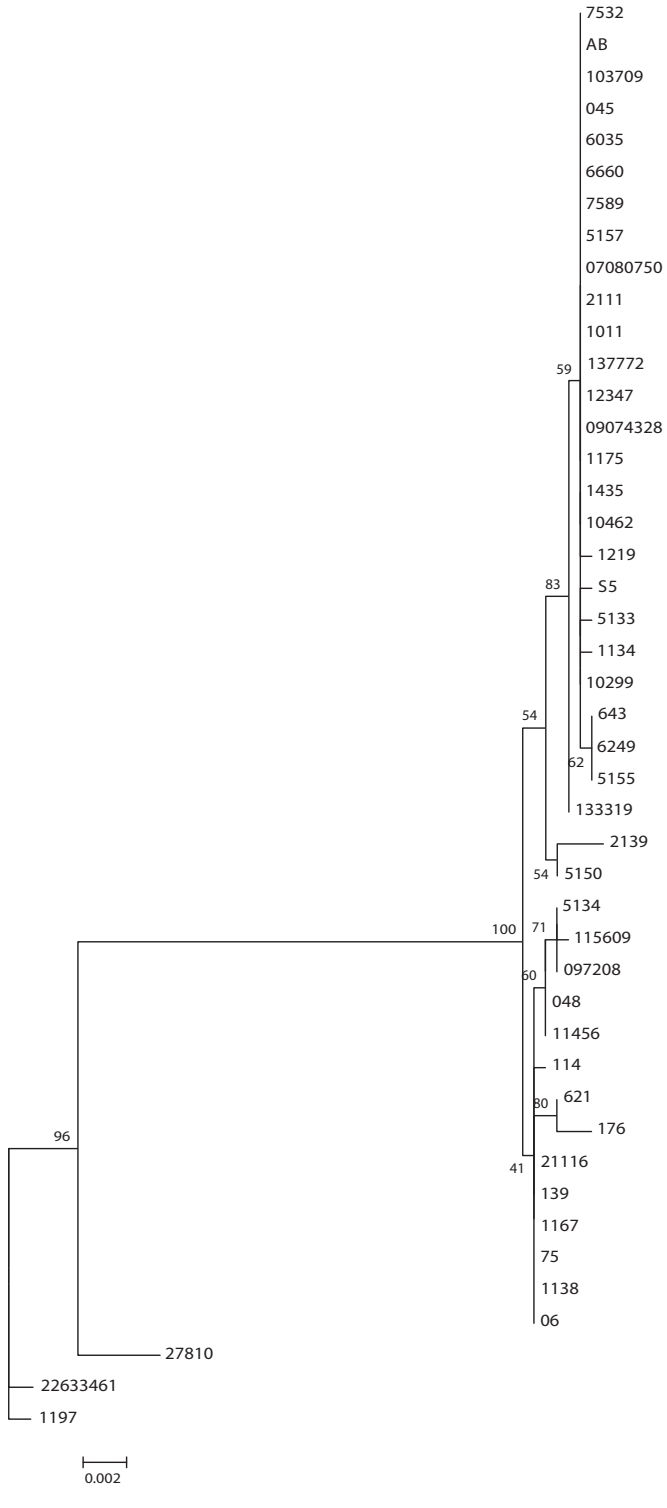


FIG. 2. ML dendrogram from the concatenated sequences of six MLST genes (SH 1200, hemH, leuB,

SH 1431, cfxE and Ribose ABC) for the 45 isolates included in the study.