

# Whole-genome sequencing to explore nosocomial transmission and virulence in neonatal methicillin-susceptible *Staphylococcus aureus* bacteremia

Bibi C.G.C. Slingerland, Margreet C. Vos, Willeke Bras, René F. Kornelisse, Dieter De Coninck, Alex van Belkum, Irwin K.M. Reiss, Wil H.F. Goessens, Corné H.W. Klaassen, Nelianne J. Verkaik

*Antimicrobial Resistance and Infection Control*, 2020

## ABSTRACT

### Background

Neonatal *Staphylococcus aureus* (*S. aureus*) bacteremia is an important cause of morbidity and mortality. In this study, we examined whether methicillin-susceptible *S. aureus* (MSSA) transmission and genetic makeup contribute to the occurrence of neonatal *S. aureus* bacteremia.

### Methods

A retrospective, single-centre study was performed. All patients were included who suffered from *S. aureus* bacteremia in the neonatal intensive care unit (NICU), Erasmus MC-Sophia, Rotterdam, the Netherlands, between January 2011 and November 2017. Whole-genome sequencing (WGS) was used to characterize the *S. aureus* isolates, as was also done in comparison to reference genomes. Transmission was considered likely in case of genetically indistinguishable *S. aureus* isolates.

### Results

Excluding coagulase-negative staphylococci (CoNS), *S. aureus* was the most common cause of neonatal bacteremia. Twelve percent ( $n = 112$ ) of all 926 positive blood cultures from neonates grew *S. aureus*. Based on core genome multilocus sequence typing (cgMLST), 12 clusters of genetically indistinguishable MSSA isolates were found, containing 33 isolates in total (2-4 isolates per cluster). In seven of these clusters, at least two of the identified MSSA isolates were collected within a time period of one month. Six virulence genes were present in 98-100% of all MSSA isolates. In comparison to *S. aureus* reference genomes, toxin genes encoding staphylococcal enterotoxin A (*sea*) and toxic shock syndrome toxin 1 (*tsst-1*) were present more often in the genomes of bacteremia isolates.

### Conclusion

Transmission of MSSA is a contributing factor to the occurrence of *S. aureus* bacteremia in neonates. *Sea* and *tsst-1* might play a role in neonatal *S. aureus* bacteremia.

## INTRODUCTION

*Staphylococcus aureus* (*S. aureus*) is a well-established nosocomial pathogen that causes multiple types of neonatal infections [1,2]. Invasive *S. aureus* infections in neonates (e.g. bacteremia) are common in very low birth weight (VLBW) infants, which makes this bacterial species one of the most important pathogens in neonatal intensive care units (NICU) [3-5]. A significant risk factor for *S. aureus* bacteremia in VLBW infants is the presence of intra-vascular catheters, which are frequently required [6-8]. In addition, *S. aureus* bacteremia can result in severe complications such as endocarditis and osteomyelitis [5,9,10]. All-cause mortality among neonates suffering from *S. aureus* bacteremia varies between 10 and 20% [7,11]. So there is an urgent need to prevent this infection. To prevent *S. aureus* bacteremia in neonates, it is important to know the factors contributing to the high frequency and severity of this infection.

Previously, the virulence factors *tsst-1* and *sea* were implicated to play a role in *S. aureus* bacteremia [12-14]. Furthermore, transmission of *S. aureus* might contribute to the high frequency of bacteremia. Outbreaks of methicillin-resistant *S. aureus* (MRSA) at the NICU are described and relatively easy to detect [15-18]. Meanwhile, the detection of methicillin-sensitive *S. aureus* (MSSA) outbreaks seems to be more difficult, excluding outbreaks in patients who suffer from a skin infection [19-22]. In this study, whole-genome sequencing (WGS), the typing method with the highest discriminatory power, was used to determine whether MSSA transmission and genetic makeup, contribute to the occurrence of neonatal *S. aureus* bacteremia.

## METHODS

### Population

The NICU of Erasmus MC-Sophia, Rotterdam, the Netherlands, is a level IV, 27-beds facility. It is divided into four units with six to eight beds each. Per year, about 750 neonates are admitted. Nearly 40% of them are below 32 weeks of gestation and were in majority born in this hospital.

### Screening

We included neonates with a presumed infection, of whom blood cultures were obtained between January 2011 and November 2017 that showed to be positive for *S. aureus*. Clinical data concerning gender, gestational age, birth weight and survival were obtained from patient records.

## S. aureus isolates

Blood from neonates was cultured in BACTEC plus PEDS aerobic bottles and incubated in the Bactec FX (BD, Heidelberg, Germany). In case of positive blood cultures, plates were inoculated and, after 16-24 hours of incubation at 37°C, screened for *S. aureus* based on colony morphology. Identification was performed by means of a latex agglutination test (Slidex Staph Plus, bioMérieux, Marcy-l'Etoile, France) and/or via matrix-assisted laser desorption/ionisation, time-of-flight, mass spectrometry (MALDI-TOF MS system, Bruker). *S. aureus* isolates were stored at – 20 °C or – 80 °C until use. The VITEK 2 system (bioMérieux) was used for antimicrobial susceptibility testing (AST).

## Whole-genome sequencing

### *Transmission*

*S. aureus* isolates were processed according to the bioMérieux EpiSeq<sup>cs</sup> V1 programme and sent to LGC Genomics GmbH (Berlin, Germany) for next-generation sequencing (NGS). We used Illumina chemistry, which generated paired end 2x150 bp reads. Sequences were assembled using the proprietary built-in assembler from CLC Genomics Workbench v11 software (Qiagen, Hilden, Germany) with default parameters. We analysed them by means of the available *S. aureus* core genome multilocus sequence typing scheme (cgMLST) [23] in BioNumerics 7.6.3 (bioMérieux, Sint-Martens-Latem, Belgium) which contains 1,861 loci. Allele calling was performed using two algorithms, one based on the assembly using a BLAST approach (assembly-based calling) and one based on the trimmed sequencing data using a kmer based approach (assembly-free calling). A consensus of both algorithms was used to assign final allele calls: when both algorithms were in agreement or when an allele call was made by only one of the algorithms, the allele call was considered in the consensus. However, when both algorithms were in disagreement, the allele call was not considered in the consensus. Both allele calling algorithms were executed using default parameters. Conventional MLST types were inferred *in silico* from the WGS data. To this end, the seven MLST loci were identified using the sequence extraction tool and the MLST plugin from BioNumerics 7.6.3 that is synchronized to the pubMLST.org public repository (accession date: April 5, 2019). For the visualisation of the genetic relatedness between the isolates, we used a minimum spanning tree for the cgMLST data. The MST was generated using default parameters, and no re-sampling was performed. Isolates containing less than 12 allelic differences in the *S. aureus* core genome were considered genetically indistinguishable [23]. We defined a cluster as more than two genetically indistinguishable isolates and, within a cluster, considered transmission of *S. aureus* likely. To further validate the results based on the cgMLST approach, as additional method, we evaluated transmission events using a SNP based approach (Additional file 1: Table S1).

### Virulence

The presence of virulence genes was assessed, using the sequence extraction tool in BioNumerics 7.6.3. Extraction parameters (percentage coverage and identity) were individualised to accommodate for the different levels of sequence diversity within and between the virulence genes. Anticipating problems upon assembling virulence genes containing repetitive motifs (*sdrA*, *-B* and *-C*, *clfA* and *-B*, *cna*, *sasG*) using the short read sequence data, only the largest non-repetitive part of these genes was used for querying. In order to obtain data from a general *S. aureus* population, the prevalence of virulence genes was also assessed by means of the available genomic sequences in the Refseq Genome Database, using the BLAST interface (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). This database contained 10,288 *S. aureus* genomes at the time of analysis. Virulence gene-specific search parameters were used as discussed above. Role and function of the *S. aureus* virulence genes were described in more detail earlier [12,24]. An overview of analysed virulence genes, their role, search parameters and query sequence are shown in Additional file 2: Table S2.

## RESULTS

### Patient characteristics

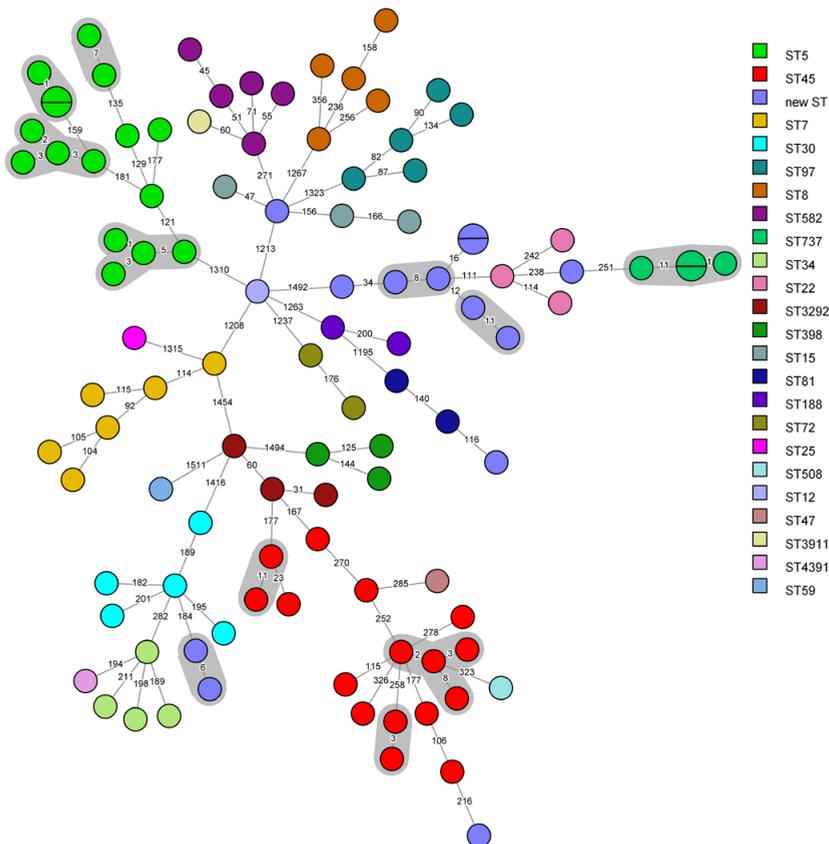
After coagulase-negative staphylococci (CoNS), MSSA was the most frequent causative pathogen of bacteremia in neonates. Several species of CoNS were isolated from neonatal blood, but they were considered to be one group. Twelve percent ( $n = 112$ ) of 926 positive blood cultures from neonates (one blood culture per episode per patient), taken in the period January 2011 – November 2017, were positive for MSSA. Fifty-nine of the 112 neonates (52.7%) with MSSA bacteremia were male. The median (interquartile range) for gestational age and birth weight were 26 3/7 (25 1/7 – 30) weeks and 880 (680 – 1150) grams, respectively. The onset of all episodes of MSSA bacteremia occurred 72 hours after birth, at a median postnatal age of 10 (7 – 19) days. The overall mortality among the included 112 patients was 20.5% while 11 of these 23 neonates died of MSSA septicaemia.

### Genetic relatedness

One hundred and four MSSA isolates from the total of 112 neonatal bloodstream isolates (93%) were available and therefore included for WGS (including only the first isolate per patient). Based on WGS, a total of 23 classical MLST types were identified. The most predominant MLST types were ST5 and ST45 (for both  $n = 16$ ). For 11 MSSA isolates a novel MLST type was found. To assess the genetic relatedness between the 104 isolates based on the more discriminatory cgMLST scheme, we visualised the number of allelic differences of the isolates in Fig. 1. Twelve cgMLST clusters of genetically indistinguishable isolates were observed, containing a total of 33 isolates (2 - 4 isolates per cluster). In seven of these

cgMLST clusters, at least two of the identified MSSA isolates were collected within a time period of one month. In two cgMLST clusters, all MSSA isolates were found within a time period of one year, but the shortest time interval between isolates of two neonates was forty days. In the other three cgMLST clusters, there was a time interval of more than one year between culturing the MSSA bloodstream isolates of two neonates. The SNP approach confirmed our results based on the cgMLST approach (Additional file 1: Table S1).

**Figure 1.** Minimum spanning tree, based on the core genome of 104 *S. aureus* isolates.



Colours indicate the classical MLST sequence types (ST). Twelve cgMLST clusters containing at least two isolates with a maximum of eleven allelic differences are indicated with a grey background.

***S. aureus* virulence genes**

An overview of virulence genes present in the 104 MSSA isolates is provided in Table 1. Of the immunomodulatory proteins, staphylococcal complement inhibitor (*scin*) was present in 100% of all bloodstream isolates. Alpha-hemolysin (*hla*) was present in 99% of the isolates.

**Table 1.** Presence of virulence genes in neonatal *S. aureus* isolates compared to reference genomes.

genes	neonatal isolates (%)	refseq (%)	genes	neonatal isolates (%)	refseq (%)
<i>sea</i>	24.0	9.4	<i>cna</i>	51.0	34.6
<i>seb</i>	4.8	5.9	<i>eap-map</i>	64.4	93.7
<i>sec</i>	18.3	10.8	<i>ebp</i>	95.2	95.2
<i>sed</i>	10.6	8.8	<i>fnbpA</i>	66.3	75.2
<i>seg</i>	51.9	55.3	<i>fnbpB</i>	65.4	75.0
<i>seh</i>	5.8	1.7	<i>sdrC</i>	91.3	96.3
<i>sei</i>	55.8	55.2	<i>sdrD</i>	50.0	74.9
<i>sej</i>	10.6	11.1	<i>sdrE</i>	89.4	85.0
<i>sek</i>	3.8	19.0	<i>efb</i>	77.9	94.7
<i>sel</i>	18.3	10.8	<i>icaA</i>	100.0	97.8
<i>sem</i>	55.8	55.6	<i>icaB</i>	89.4	98.5
<i>sen</i>	45.2	55.2	<i>icaC</i>	99.0	97.8
<i>seo</i>	58.7	49.1	<i>icaD</i>	100.0	98.0
<i>sep</i>	4.8	19.1	<i>icaR</i>	95.2	97.7
<i>seq</i>	4.8	19.1	<i>isaA</i>	100.0	98.3
<i>ser</i>	10.6	10.9	<i>isdA</i>	100.0	97.6
<i>ses</i>	0.0	0.3	<i>isdH</i>	100.0	97.5
<i>set</i>	0.0	0.3	<i>sasG</i>	54.8	56.7
<i>seu</i>	51.9	54.8	<i>edn</i>	1.0	1.4
<i>sey</i>	1.0	5.8	<i>etA</i>	1.0	1.0
<i>aur</i>	100.0	98.2	<i>etB</i>	0.0	0.2
<i>coa</i>	70.2	83.0	<i>etD</i>	1.0	0.8
<i>geh</i>	87.5	97.5	<i>hla</i>	99.0	97.1
<i>hysA</i>	98.1	96.5	<i>hIb</i>	100.0	95.1
<i>sak</i>	73.1	78.5	<i>hId</i>	90.4	98.0
<i>sspA</i>	99.0	98.1	<i>hlgA</i>	96.2	97.6
<i>sspB</i>	91.3	98.1	<i>hlgB</i>	100.0	97.8
<i>sspC</i>	100.0	98.3	<i>hlgC</i>	98.1	97.5
<i>vWbp</i>	78.8	98.1	<i>lukD</i>	49.0	66.5
<i>adsA</i>	96.2	98.1	<i>lukE</i>	49.0	65.9
<i>chp</i>	68.3	67.2	<i>lukF</i>	0.0	18.8
<i>sbi</i>	100.0	97.7	<i>lukM</i>	0.0	0.7
<i>scn</i>	98.1	98.2	<i>lukS</i>	0.0	19.0
<i>clfA</i>	98.1	97.4	<i>tsst-1</i>	18.3	5.8
<i>clfB</i>	100.0	97.8			

We also found a 98-100% presence of the MSCRAMMs clumping factors A and B (*clfA*, *clfB*), immunodominant surface antigen A (*isaA*) and iron-responsive surface determinants A and H (*isdA*, *isdH*). When compared to a reference population of *S. aureus* genomes, a few observations stand out. Remarkably, staphylococcal enterotoxin A (*sea*) and toxic shock syndrome toxin 1 (*tsst-1*) were, respectively, 2.6 and 3.2 times more prevalent among the 104 neonatal bloodstream isolates, relative to the reference genomes. Likewise, staphylococcal enterotoxin h (*seh*) was 3.4 times more prevalent although, in absolute numbers, this involved only a few isolates (6/104 versus 173/10288). For the other virulence genes, no such increases were detected (Table 1).

## DISCUSSION

At our level IV neonatal intensive care unit, as in many centres [3-5], *S. aureus* is a frequent cause of neonatal bacteremia. In our study, we explored the role of MSSA transmission and the possible contribution of virulence genes. By using WGS, 12 different cgMLST clusters of MSSA isolates were found. Seven of these twelve cgMLST clusters included at least two MSSA isolates, cultured from blood of neonates within one month, indicative for transmission. Transmission should therefore be considered as a contributing factor for the frequent occurrence of neonatal *S. aureus* bacteremia, as was recently described by Rouard et al. [13]. Although it seems reasonable to assume that transmission, irrespective of the source, can only occur through the hands of healthcare workers (HCWs), we did not prove this, since we did not culture the environment, nor the HCWs or parents. Still, general measures such as improvement of the current (daily) cleaning, disinfection procedures as well as hand hygiene, will be likely to help. It was already proven that neonatal hospital-acquired infections could in part be prevented by strict infection control measures [8,25,26]. In addition, reinforcement of the implementation of central-line bundles has the potential to reduce the incidence of central line-associated bloodstream infections (CLABSIs); although these bundles are already implemented, compliance can still be improved and additional measures can be explored [27].

Besides transmission, it was determined whether the presence of certain virulence factors is associated with neonatal *S. aureus* bacteremia. Since it was difficult to define a suitable control population of neonates, we chose to compare neonatal *S. aureus* bacteremia isolates to all available *S. aureus* genomes from the Refseq Genome Database ( $N = 10,288$  at the time of analysis). Remarkably, the genes *sea* and *tsst-1* were found a factor 2.6 and 3.2 times more often in the MSSA bloodstream isolates, compared to the reference genomes in the Refseq Genome Database. The overrepresentation of *tsst-1* could not be explained by the frequent presence of MLST ST5 and ST45 in our isolates collection, since *tsst-1* was not associated with these sequence types. On the other hand, 11 of the 25 isolates carrying *sea*

were found in ST5 isolates. Still, this cannot be the full explanation for finding an association between *sea* and neonatal MSSA bacteremia. Many studies have been executed on *S. aureus* toxins and their pathogenic roles, particularly on *sea* and *tsst-1*. Previously, it was described that antibody responses to these two specific toxins were higher in patients with *S. aureus* bacteremia, compared to control patients [12]. In addition, in a recent publication about a NICU MSSA outbreak, *tsst-1* and especially *sea* were found in bloodstream isolates, compared to colonisation isolates [13]. Another review article describes the association of these toxins with bacteremia [14]. Therefore, this may suggest that *sea* and *tsst-1* might play a role in the pathogenesis of *S. aureus* bacteremia. The other virulence genes were present in virtually all study isolates, but in virtually all reference genomes as well (Table 1).

Our study has its limitations. It was performed retrospectively, in a single centre. We considered less than 12 allelic differences in the *S. aureus* core genome as indistinguishable, as described by *Leopold et al.* for MRSA outbreaks [23]. Still, it is a matter of debate which cut-off should be used to define MSSA isolates as indistinguishable. If the cut-off would have been set at 20 alleles [28], this would have led to 10 larger instead of 12 indistinguishable MSSA clusters, which would not have changed the conclusion regarding transmission. Additional studies are needed to define a clear cut-off. Finally, we compared neonatal isolates to a large number of reference genomes, but these originate from several countries, several clinical sites and patients of all ages. It would have been ideal if the reference genomes had originated from colonized but not infected neonates, admitted to the same NICU, in the same time period.

## CONCLUSIONS

In conclusion, transmission of MSSA seems a contributing factor to the occurrence of *S. aureus* bacteremia in neonates. The possibility of MSSA transmission in neonatal intensive care should be explored to prevent this invasive and serious infection. The exact role of *sea* and *tsst-1* warrants further investigation.

## ACKNOWLEDGEMENTS

We thank Marian ten Kate and Heleen van de Spek for their help in collecting the *S. aureus* isolates. We thank Marius Vogel for his help in data collection. We thank Bruno Lacroix and Deborah Braun (both bioMérieux France) for their help.

## **AUTHOR'S CONTRIBUTIONS**

BS contributed to the study design, collected, analysed and interpreted the data, and drafted the initial manuscript. NV conceptualised and designed the study, collected, analysed, interpreted and supervised the data collection and critically revised and reviewed the manuscript for intellectual content. MV contributed to the study design, while MV WB collected and interpreted the data, and critically revised and reviewed the manuscript for intellectual content. RK IR contributed to the study design, collected data, and critically revised and reviewed the manuscript for intellectual content. DDC CK contributed to the study design, collected, analysed and interpreted the data, and critically revised and reviewed the manuscript for intellectual content. AVB WG conceptualised and designed the study, interpreted the data and critically revised and reviewed the manuscript for intellectual content.

## **FUNDING**

bioMérieux funded this study. bioMérieux designs, develops and sells diagnostic tests in the domain of infectious diseases. The company had no influence on the design and execution of the current study.

## **AVAILABILITY OF DATA AND MATERIALS**

The data generated and analysed in this study are included in the current article.

## **ETHICS APPROVAL AND CONSENT TO PARTICIPATE**

Because this was a retrospective observational study in which anonymised patient data were used, collected during routine clinical practice, informed consent was not mandatory according to the Dutch Medical Research Involving Human Subjects Act (WMO). The Institutional Ethics Review Board of the Erasmus MC reviewed the study protocol and provided an exemption from formal ethical assessment (MEC-2015-306), based on the non-interventional design. The study was carried out in accordance with the current ethical guidelines for epidemiological research.

## **CONSENT FOR PUBLICATION**

Not applicable.

## **COMPETING INTERESTS**

All other authors declare that there are no competing personal or institutional interests.

## REFERENCES

- 1 Jeong IS, Jeong JS, Choi EO: Nosocomial infection in a newborn intensive care unit (NICU), South Korea. *BMC Infect Dis* 2006;6:103.
- 2 Reichert F, Piening B, Geffers C, Gastmeier P, Buhner C, Schwab F: Pathogen-Specific Clustering of Nosocomial Blood Stream Infections in Very Preterm Infants. *Pediatrics* 2016;137
- 3 Ericson JE, Popoola VO, Smith PB, Benjamin DK, Fowler VG, Benjamin DK, Jr., Clark RH, Milstone AM: Burden of Invasive *Staphylococcus aureus* Infections in Hospitalized Infants. *JAMA Pediatr* 2015;169:1105-1111.
- 4 Carey AJ, Duchon J, Della-Latta P, Saiman L: The epidemiology of methicillin-susceptible and methicillin-resistant *Staphylococcus aureus* in a neonatal intensive care unit, 2000-2007. *J Perinatol* 2010;30:135-139.
- 5 Dolapo O, Dhanireddy R, Talati AJ: Trends of *Staphylococcus aureus* bloodstream infections in a neonatal intensive care unit from 2000-2009. *BMC Pediatr* 2014;14:121.
- 6 Ekkelenkamp MB, van der Bruggen T, van de Vijver DA, Wolfs TF, Bonten MJ: Bacteremic complications of intravascular catheters colonized with *Staphylococcus aureus*. *Clin Infect Dis* 2008;46:114-118.
- 7 Hakim H, Mylotte JM, Faden H: Morbidity and mortality of Staphylococcal bacteremia in children. *Am J Infect Control* 2007;35:102-105.
- 8 Murdoch F, Danial J, Morris AK, Czarniak E, Bishop JL, Glass E, Imrie LJ: The Scottish enhanced *Staphylococcus aureus* bacteraemia surveillance programme: the first 18 months of data in children. *J Hosp Infect* 2017;97:127-132.
- 9 Le J, Dam Q, Tran T, Nguyen A, Adler-Shohet FC, Kim S, Schmidt K, Lieberman JM, Bradley JS: Epidemiology and hospital readmission associated with complications of *Staphylococcus aureus* bacteremia in pediatrics over a 25-year period. *Epidemiol Infect* 2017;145:2631-2639.
- 10 Valente AM, Jain R, Scheurer M, Fowler VG, Jr., Corey GR, Bengur AR, Sanders S, Li JS: Frequency of infective endocarditis among infants and children with *Staphylococcus aureus* bacteremia. *Pediatrics* 2005;115:e15-19.
- 11 Kaufman D, Fairchild KD: Clinical microbiology of bacterial and fungal sepsis in very-low-birth-weight infants. *Clin Microbiol Rev* 2004;17:638-680, table of contents.
- 12 Verkaik NJ, Dauwalder O, Antri K, Boubekri I, de Vogel CP, Badiou C, Bes M, Vandenesch F, Tazir M, Hooijkaas H, Verbrugh HA, van Belkum A, Etienne J, Lina G, Ramdani-Bouguessa N, van Wamel WJ: Immunogenicity of toxins during *Staphylococcus aureus* infection. *Clin Infect Dis* 2010;50:61-68.
- 13 Rouard C, Bourgeois-Nicolaos N, Rahajamanana L, Romain O, Pouga L, Derouin V, De Luca D, Doucet-Populaire F: Evaluation of an 'all-in-one' seven-day whole-genome sequencing solution in the investigation of a *Staphylococcus aureus* outbreak in a neonatal intensive care unit. *J Hosp Infect* 2019;102:297-303.
- 14 Xu S, McCormick J: Staphylococcal superantigens in colonization and disease. *Front Cell Infect Microbiol* 2012;2:52.
- 15 Regev-Yochay G, Rubinstein E, Barzilai A, Carmeli Y, Kuint J, Etienne J, Blech M, Smollen G, Maayan-Metzger A, Leavitt A, Rahav G, Keller N: Methicillin-resistant *Staphylococcus aureus* in neonatal intensive care unit. *Emerg Infect Dis* 2005;11:453-456.
- 16 Reich PJ, Boyle MG, Hogan PG, Johnson AJ, Wallace MA, Elward AM, Warner BB, Burnham CA, Fritz SA: Emergence of community-associated methicillin-resistant *Staphylococcus aureus* strains in the neonatal intensive care unit: an infection prevention and patient safety challenge. *Clin Microbiol Infect* 2016;22:645 e641-648.

- 17 Huang YC, Lien RI, Su LH, Chou YH, Lin TY: Successful control of methicillin-resistant *Staphylococcus aureus* in endemic neonatal intensive care units--a 7-year campaign. *PLoS One* 2011;6:e23001.
- 18 Ramsing BG, Arpi M, Andersen EA, Knabe N, Mogensen D, Buhl D, Westh H, Ostergaard C: First outbreak with MRSA in a Danish neonatal intensive care unit: risk factors and control procedures. *PLoS One* 2013;8:e66904.
- 19 Achermann Y, Seidl K, Kuster SP, Leimer N, Durisch N, Ajdler-Schaffler E, Karrer S, Senn G, Holzmann-Burgel A, Wolfensberger A, Leone A, Arlettaz R, Zinkernagel AS, Sax H: Epidemiology of Methicillin-Susceptible *Staphylococcus aureus* in a Neonatology Ward. *Infect Control Hosp Epidemiol* 2015;36:1305-1312.
- 20 Lin MF, Huang ML, Lai SH: Investigation of a pyoderma outbreak caused by methicillin-susceptible *Staphylococcus aureus* in a nursery for newborns. *J Hosp Infect* 2004;57:38-43.
- 21 Gomez-Gonzalez C, Alba C, Otero JR, Sanz F, Chaves F: Long persistence of methicillin-susceptible strains of *Staphylococcus aureus* causing sepsis in a neonatal intensive care unit. *J Clin Microbiol* 2007;45:2301-2304.
- 22 Koningstein M, Groen L, Geraats-Peters K, Lutgens S, Rietveld A, Jira P, Kluytmans J, de Greeff SC, Hermans M, Schneeberger PM: The use of typing methods and infection prevention measures to control a bullous impetigo outbreak on a neonatal ward. *Antimicrob Resist Infect Control* 2012;1:37.
- 23 Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A: Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J Clin Microbiol* 2014;52:2365-2370.
- 24 Verkaik NJ, van Wamel WJ, van Belkum A: Immunotherapeutic approaches against *Staphylococcus aureus*. *Immunotherapy* 2011;3:1063-1073.
- 25 Borghesi A, Stronati M: Strategies for the prevention of hospital-acquired infections in the neonatal intensive care unit. *J Hosp Infect* 2008;68:293-300.
- 26 Trevisanuto D, Arnolda G, Chien TD, Xuan NM, Thu le TA, Kumara D, Lincetto O, Moccia L: Reducing neonatal infections in south and south central Vietnam: the views of healthcare providers. *BMC Pediatr* 2013;13:51.
- 27 Ista E, van der Hoven B, Kornelisse RF, van der Starre C, Vos MC, Boersma E, Helder OK: Effectiveness of insertion and maintenance bundles to prevent central-line-associated bloodstream infections in critically ill patients of all ages: a systematic review and meta-analysis. *Lancet Infect Dis* 2016;16:724-734.
- 28 Goyal M, Javerliat F, Palmieri M, Mirande C, van Wamel W, Tavakol M, Verkaik NJ, van Belkum A: Genomic Evolution of *Staphylococcus aureus* During Artificial and Natural Colonization of the Human Nose. *Front Microbiol* 2019;10:1525.

## ADDITIONAL FILE 1

### Methods

#### *Whole-genome Single-Nucleotide Polymorphism analysis*

A whole-genome Single-Nucleotide Polymorphism (wgSNP) analysis was performed on the samples within a potential transmission cluster identified by wgMLST analysis (BioNumerics 7.6.3 (Applied Maths, Sint-Martens-Latem, Belgium)). For each potential transmission cluster a reference genome for mapping with Bowtie 2 using default parameters was selected based on the N50 of the assembly and number of loci which were assigned an allele during wgMLST analysis. The strict SNP filtering template was applied.

### Results

#### *SNP matrices of the possible transmission clusters detected by cgMLST.*

Cluster numbers correspond to the clusters mentioned in the manuscript and Figure 1 involving samples with a maximum of 11 alleles difference using the cgMLST approach.

Table S1

**CLUSTER 1**

isolate	1	2
1	0	32
2	32	0

**CLUSTER 2**

isolate	1	2	3
1	0	5	5
2	5	0	0
3	5	0	0

**CLUSTER 3**

isolate	1	2	3	4
1	0	10	11	7
2	10	0	12	8
3	11	12	0	10
4	7	8	10	0

**CLUSTER 4**

isolate	1	2	3	4
1	0	1	1	4
2	1	0	2	5
3	1	2	0	4
4	4	5	4	0

**CLUSTER 5**

isolate	1	2	3	4
1	0	2	2	37
2	2	0	0	37
3	2	0	0	37
4	37	37	37	0

**CLUSTER 6**

isolate	1	2
1	0	17
2	17	0

Table S1 (continued)

**CLUSTER 7**

isolate		1	2
1		0	21
2		21	0

**CLUSTER 8**

isolate		1	2
1		0	59
2		59	0

**CLUSTER 9**

isolate		1	2
1		0	62
2		62	0

**CLUSTER 10**

isolate		1	2
1		0	8
2		8	0

**CLUSTER 11**

isolate		1	2	3	4
1		0	1	4	16
2		1	0	3	17
3		4	3	0	20
4		16	17	20	0

**CLUSTER 12**

isolate		1	2
1		0	13
2		13	0

---



































Table S2. *S. aureus* virulence genes (continued)

Category	gene symbol	description	BLAST search parameters			Query sequence
			minimum percent-identity	minimum alignment length	minimum alignment length	
Toxin	PVL	Panton-Valentin leukocidin	90	1555	atgattttatgctcaaaaagaagactatgacigcaacattgctgtaggaataatcaactctattgctactctgctttctgaaicctaaagctgatacaataatgagaatattgagtgatggcgcctgag- gtagtcaaaaagacagaagagatacaagtagcgtatagtgggggctcacaanaatattcagagtgatttggtaaaagataaaaagatatacaaaagacgcttggatttaaaaagcaaggtttat- caattcaaaagactactatatacaattacaaaacacagatcatalabaagcgaatgagggccttccaataacataatgctcctcaaaacaaatgaccccaatgagattaaataatctactcta- aataaataatgattccagtaaaagtgtagcaaacataaagttataacatagtggttaatttttaagtggtgctcaataacagaggagaaagtggtctcaattaatgactcaaaaacataagttatataca- caaacatcatcagtgtaggaacgctcaaaatcacaagaaggttcaatggggaataaaagctaaattcattcaatcactataggttaaaagctgacatgactcaaatatttggtagataaaac- catatagtcanaaalccgagagactatgttccagacaatgaaatcccccattagacacagtggtttcaactcttattatgcaaacctgctctcagaaaaggctcagagagatacaagtgaaattt- gaataagatgtagcaaaataatggaatggttcaatgctactcaagcaaacacactatggaagaaacacactatggaagactctggaatacaaacgcaattgtaaacagaaaactcacagta- aatagaagtgactgaaaactcagaataaagtgaaaggacataattgatagaaaataatgtaaaatcagttcaatcagttgtaacatcagttgcttctcaatacagctgtagcagct- caacatcaccaccigtgaagtagaaaaggtgagabaaaatctttgtacaacaaactgcaacactgcaacactgcaagctcgaataagttaaaatctcagatttaactttatataagataaaag- tatgataagatcactaactcaaaagctctgggaacattatctggctatacaagcaaacactcaaaagcaactatagcttcaatttataggggctcagaacacacttcaataatctta- gattcaatgactcagtaaacctgttagattgcaacttaaaatcaaatgaaagattcagtaacaacacgtaggatttcttaagggagataataatcttaacggcttaacaggtgagag- gtaatggttcaaaactttttcagaagaactaaactataaaacagaagactatagaactgcttagataaagaactaaatgaaagaaatggtgaggtgtagcaataaataatgaaataag- gtagctcaatccagaattattgctgctctatctcgaaaacaaaacgctgcaaaaatgcaaaaactgtaacttaaaagaaagtagagatacaaaactttggatacaactctcaggttag- gtaataatataaagatgaaaatagaagcaactcaactcaattgaaaggtgattgggaataatacaggttaataatagaatactcaactaaggaaaataatcctatgagctaa	
Toxin	tsst-1	toxic shock syndrome toxin-1	90	566	atgaataaaaataactaaatgaataatgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtag- atataaaggatttgcagaactggtagtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtagtgtag- gtagctcttataagccctactctttcaaaagggaaaaagttgtagcttaaaacaaaaggactaaaaaagcccaactacgcaaaaggaaactggagatcattcccaataagtgtagtacaatac- gaaaattcctcctcatalagaactcacttaaaagtttaggttag- cagtagtaactcaatacagtagattatctgtcaagcgataaaacgaggtggttagtggaaaataacaatgaaagacgagatccaataatcaaaagtttactcaaaaagtttactaaactgaaaac- cacctataatgtagaataaaaactatagaagcagaaattatga	

Sequences marked with an asterisk contain repetitive elements. Only the largest non-repetitive part was used for querying.

