Title: Prospective genomic surveillance reveals cryptic MRSA outbreaks with local to international origins among NICU patients.

Authors: Jay N. Worley^{a,b}, Jessica W. Crothers^c, William J. Wolfgang^d, Sai Laxmi Gubbala Venkata^d, Maria Hoffmann^e, Victor Jayeola^e, Michael Klompas^{f,g}, Marc Allard^e, and Lynn Bry^{a,h,*}

- a) Massachusetts Host-Microbiome Center, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
- b) National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA
- c) Department of Pathology and Laboratory Medicine, University of Vermont Medical Center, Burlington VT USA. Translational Global Infectious Disease Research Center, Larner College of Medicine, University of Vermont Burlington VT USA.
- d) Wadsworth Center, Division of Infectious Diseases, New York State Department of Health, Albany, NY, USA
- e) Center for Food Safety and Nutrition, US Food and Drug Administration, College Park, MD, USA
- f) Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA
- g) Division of Infectious Diseases, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
- h) Clinical Microbiology Laboratory, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

*Corresponding author

Abstract:

MRSA infections cause significant morbidity and mortality in neonates. Clinical testing and routine surveillance screening identified an increase in neonates with MRSA colonization and infection which triggered prospective genomic surveillance. Here we show the complex transmission dynamics of MRSA in a NICU setting. Analyses revealed concurrent transmission chains affecting 16 of 22 MRSA-colonized patients (68%), and 3.1% of all NICU patients (n=517). Prematurity and longer lengths of stay increased risks for colonization. Intervals of up to 7 months occurred among some cluster-related isolates. 3 of 22 MRSA-colonized patients developed invasive infections with the colonizing strain. Comparisons with 21,521 isolates in the NCBI Pathogen Detection Resource revealed NICU strains to be distinct from MRSA seen locally and internationally. Integration of international strain datasets in analyses increased the resolution of strain clusters and helped rule-out suspected transmission events. Analyses also identified sequence type 1535 isolates, emergent in the Middle East, carrying a unique *SCCmec* with *fusC* and *aac(6')-Ie/aph(2'')-1a* that provided a multi-drug resistant phenotype. NICU genomic surveillance identified cryptic MRSA colonization events, including NICU-endemic strains not linked with local hospital or international clusters, and has rich potential to guide improvements in infection prevention for this vulnerable patient population.

Introduction

Staphylococcus aureus is a common cause of healthcare associated infections^{1,2}. The pathogen can cause sepsis, respiratory and skin and soft tissue infections and is a major cause of patient morbidity and mortality³. Methicillin-resistant *Staphylococcus aureus* (MRSA) colonization is a well-recognized risk factor for infection^{4,5}. Efforts to reduce infections have focused on decreasing MRSA transmission via institutional screening and decolonization^{6,7}.

Neonatal populations in general, and pre-term and low birthweight infants in particular, are at increased risk of MRSA colonization and infection⁸⁻¹¹ given their lack of protective colonizing microbiota and immature immune system^{4,12,13}. Nosocomial transmission of *S. aureus* in neonates can occur among patients¹⁴⁻¹⁷, as well as from colonized healthcare workers¹⁸⁻²⁰, parents^{21,22}, and environmental reservoirs²³. However, the relative contribution of these reservoirs remains controversial^{13,24}. Additionally, neonatal intensive care units (NICUs) have witnessed a shift over the past twenty years from hospital acquired MRSA strains to community-acquired MRSA strains, requiring modifications in infection control practices for this patient population^{3,25}.

Whole genome sequencing (WGS) enables epidemiologic outbreak investigations through high-resolution strain tracking²⁶⁻²⁹. While *S. aureus* outbreak investigations have historically used targeted molecular strain-typing, the discriminatory power of conventional techniques has been largely insufficient to establish definitive genomic and strain-specific linkages³⁰. In contrast, WGS enables definitive clonal assessments and associated temporal relationships when combined with epidemiologic data³¹. Applied to NICU-populations, WGS has improved the detection of patient-to-patient^{17,27} and healthcare worker transmission events³²⁻³⁴, environmental reservoirs of infection^{23,35}, and supported more effective MRSA screening

programs³⁶. Importantly, WGS can also rule-out nosocomial transmission events, providing reassurance to staff and appropriate utilization of hospital resources^{29,37}.

The staphylococcal cassette chromosome *mec* (SCC*mec*) confers broad-spectrum resistance to beta-lactam antibiotics and is a common genetic cause for the MRSA phenotype. SCC*mecs* contain a *mec* gene encoding PBP2A, a peptidoglycan transpeptidase homolog of PBP2 with reduced affinity for beta-lactam antibiotics, particularly cephalosporins³⁸. Many SCC*mec* types, as well as non-*mec* SCCs, encode other resistance genes such as the fusidic acid resistance gene *fusC*^{38,39}.

Observations of an increase in NICU MRSA colonizations and infections raised concerns for transmission events among neonates. Genomic analyses identified multiple cryptic outbreaks missed by clinical surveillance and diagnostic testing and also identified strain recurrences in distinct time periods that could not be attributed solely to patient-to-patient transmission. Expanded genomic analyses with >20,000 internationally deposited MRSA genomes in the NCBI Pathogen Detection resource significantly improved the phylogenetic resolution of local strain clusters and their dynamics of spread within the NICU, demonstrating the power of large-scale comparative genomic analyses for local hospital efforts to improve clinical and epidemiologic MRSA outbreak investigations.

Results

Antibiotic resistance phenotypes of NICU MRSA trigger prospective genomic surveillance. MRSA NICU isolates were evaluated phenotypically for similar antibiotic resistance profiles during a 341-day period over 2018-2019. Two suspected outbreaks during this targeted surveillance period triggered prospective genomic surveillance of all MRSA isolates from days

342-558, in addition to sequencing of prior available isolates from days 1-341 (Figure 1, Supplemental Figure 1). Over the 558 day period, 58 MRSA isolates, 55 screening and 3 diagnostic, were collected from 22 NICU patients. Of these patients, 16 neonates of 517 admitted during the prospective genomic surveillance period had MRSA isolates identified, representing 3.1% of NICU patients.

Prematurity and increased NICU lengths of stay elevate risks for MRSA colonization.

MRSA colonization was significantly associated with prematurity as well as low-birth-weight compared to non-colonized patients (Table 1, $q=0.002$; Supplemental Table 1). In addition to prematurity and low birth weight, MRSA-colonized patients had longer average lengths of stay in the NICU than non-colonized patients at 52.5 days (± 34.8 days) versus 20.0 days (± 24.8 days) ($p < 0.001$). Prematurity was also associated with longer NICU lengths of stay than patients carried to term (defined as >36 weeks of gestation), at 29.7 days (± 29.3 days) versus 6.8 days (± 7.2 days), respectively (Supplemental Table 2).

Genomic analyses identify 4 outbreak SNP clusters and unique NICU patient isolates.

Genomic analyses using NCBI's Pathogen Detection Resource identified 6 SNP clusters that accounted for 91.2% of sequenced NICU MRSA isolates originating from 17 of 22 MRSA-colonized patients (Figure 1). The remaining 5 patients with non-clustered MRSA strains each had short NICU admissions that yielded only one isolate. These 5 patient-unique isolates showed no relationship to any *S. aureus* isolates deposited in NCBI and were not temporally linked to other NICU transmission events (Supplemental Figure 1).

SNP clusters 1, 3, 4, and 5, involved multiple patients (Figure 2). Clusters 1 and 3 were found across distinct time periods without patient overlap. The patient isolates in cluster 1 occurred in two separate intervals separated by 259 days, while strains in cluster 3 occurred across three time periods separated by 49 and 56 days (Figure 1). In contrast, clusters 4 and 5 occurred within single time periods. Among the two other NCBI-associated clusters, cluster 2 included one NICU isolate that clustered with an isolate reported from another institution, but was not clonally related (Supplemental File 2). Cluster 6 included two clonally-related isolates from the same patient that were separated by 5 SNPs. In total, 15 of 22 (68%) MRSA-colonized NICU patients, including all twins, and all cases of invasive disease occurred in clusters 1, 3, 4 and 5.

Comparisons of NICU strains in clusters 1, 3, 4 and 5 with adult nosocomial MRSA isolates seen at the hospital showed no clonal relationships between the populations, indicating that the outbreak clusters seen in the NICU were genomically distinct. Further, the NICU outbreak clusters did not include any of the 21,521 *S. aureus* isolate genomes deposited in NCBI's Pathogen Isolates browser. To further assess nearest neighbor strains, whole-genome MLST (wgMLST) of NICU and 19,670 isolates with assemblies (Figure 3, Supplemental File 1, Supplemental File 2) confirmed that the NICU SNP clusters remained distinct even within broader wgMLST groups.

International MRSA genomic datasets enable rule-out of putative transmission events.

Analyses of Patient 21's MRSA isolate showed an identical antibiogram, MLST, and SCCmec type to isolates found in NICU cluster 4. However, incorporating the additional external strain

genomic datasets showed clear placement of this patient's isolate in a genomically distinct cluster (Figure 3, Supplemental File 1).

A gain of the *SCCmec* type IVa (2B) is evident in the sub-clade of multi-locus sequence type (MLST) 5 including this isolate. Eight of the nine other closely related *SCCmec* type IVa (2B) strains were submitted by Mount Sinai Hospital in New York City, New York, and the 9th by the University of Pittsburgh Medical Center in Pittsburgh, Pennsylvania. Thus, while Patient 21's isolate and cluster 4's isolates appear to belong to the same regional clade, this patient strain was ruled out as belonging to NICU cluster 4.

Genomic identification of internationally-occurring MRSA in NICU patients

Cluster 5 isolates, from patients 15 and 16 in twin set 3, included two nasal-perirectal swab isolates from patient 15 and one epidermal abscess isolate from patient 16. The three isolates belonged to MLST 1535 (Figure 4A). The twin's parents had recently immigrated from the Middle East where this MLST is more prevalent⁴⁰. While ST1535 appears to be rare in the US, one ST1535 genome from a pediatric cystic fibrosis patient in Cincinnati (GCA_002124405.1) was included in analyses⁴¹. The MLST 1535 genomes share the presence of three antibiotic resistance genes; *aac(6')-Ie/aph(2'')-1a*, *fusC*, and *mecA* (Figure 4B). High-resolution closed genomes of the cluster 5 isolates revealed an *SCCmec* region containing these resistance genes flanked by multiple mobilizing insertion sequences and novel repetitive elements that had not been resolvable in the draft-level genomic data (Figure 4C). The *SSCmec* contains *ccrC* and *ccrAA* with *mecA* allele C2, *fusC* in the J3 region and an IS256 flanked integron in the J1 region that contains *aac(6')-Ie-aph(2'')-1a*³⁸. Mobilizing IS431*mec* and IS256 were each found multiple

156 times in this region, as well as a set of three direct repeats sharing a 13 base long core sequence
157 not found elsewhere in the chromosome. The cluster 5 isolates demonstrated fusidic acid
158 resistance at 8 µg/ml, oxacillin resistance at 2 or 0.5 µg/ml, and gentamicin resistance at ≥16
159 µg/ml (Supplemental File 1).

160

Discussion

Preterm neonates are highly susceptible to colonization with pathogens, including MRSA, per their lack of colonization with a complex microbiota, immature immune system and frequent co-morbidities associated with preterm birth. Genomic analyses of surveillance and diagnostic MRSA cultures from NICU patients identified multiple outbreak clusters and transmission events. Many of these transmissions were cryptic events that fell below thresholds of detection using standard phenotypic assessments with clinical microbiologic data.

Four MRSA outbreaks associated with patient-to-patient transmission, clusters 1, 3, 4 and 5, occurred during the period of prospective genomic surveillance (Figures 1 and 2). These clusters accounted for isolates in 15 of 22 MRSA-colonized patients (68%) and all invasive MRSA infections. Genomic analyses provided high resolution information to resolve clusters and to also exclude putative transmission events in cases where strains co-occurring in time had identical microbiologic phenotypes. Geomic analyses resolved co-occurring MRSA strains among patients into three distinct clusters 1, 3, and 4. In addition, the sequence type 5 (ST5) isolate seen in surveillance cultures from patient 21 showed identical microbiologic phenotypes to strains in cluster 4 but a genomically distinct SNP profile allowing it to be ruled out as part of a transmission chain (Supplemental File 1). The higher resolution analyses enabling this distinction used 21,520 publicly-available genomic datasets from NCBI to resolve sub-clusters among ST5 strains (Figure 3). The external MRSA genomic data greatly improved the phylogenetic resolution of locally obtained strains and was thus able to rule out strain relatedness and further transmission events. Incorporation of large genomic datasets from geographically diverse regions enhances the accuracy of local genomic investigations of nosocomial outbreaks⁴².

We identified two potentially endemic strains in the NICU. Clusters 1 and 3 occurred over discontinuous time periods (Figure 1). Cluster 1 involved a NICU strain seen prior to initiating prospective genomic surveillance, with only 9 SNPs difference between the point of first detection of the cluster and its recurrence 259 days later (Figure 2). This finding concurs with prior literature showing that persistent strain carriage or presence on surfaces may accumulate low levels of SNP heterogeneity over time⁴²⁻⁴⁵. Similarly, cluster 3 isolates showed no consistent SNP differences while also appearing in three discontinuous time periods separated by 56 days and 49 days. Both clusters raise the possibility of sustained reservoirs that can seed intermittent outbreaks.

Two of four patients with isolates in cluster 4 developed invasive disease. Patient 12 had positive MRSA surveillance cultures 3 weeks in advance of infection while patient 14's infection was detected on the same day as positive surveillance cultures (Figure 2C). In all cases, longitudinal MRSA isolates from individual patients and among twin sets were clonally related. Surveillance cultures therefore have predictive value for patients at higher risk for infections. However, patients 11, 15 and 16 were also co-colonized by methicillin sensitive *S. aureus*, demonstrating that colonization by multiple *S. aureus* strains can occur in NICU patients and can shift over time.

The use of extensive and geographically diverse datasets improved NICU-focused assessments of strain reservoirs and potential transmission events. The ST1535 isolate seen in patients 15 and 16 from twin set 3 (Figure 4, Supplemental Table 1) occurred primarily in locations from the Middle East, including clinical and environmental reservoirs⁴⁶⁻⁴⁸. While the parents and visiting family had recent international travel from the Middle East, domestic sources cannot be ruled out, as evidenced by a similar strain found in Cincinnati⁴¹. Recent

observations of confirmed and suspected MDR SCCs encoding *mecA* with other antimicrobial resistance genes, including several with *fusC*, raises concerns globally for Staphylococcal infections requiring complicated antimicrobial therapy⁴⁹⁻⁵¹. *fusC* fusidic acid resistance has been seen in non-*mec* SCC and *SCCmec* elements^{49,52,53}. The *SCCmec* sequence from the closed isolate genomes were highly similar to draft *SCCmec* sequences reported by Senok, et al.⁴⁰. Closing of these isolate genomes identified clear *IS43I**mec* and *IS256* sequences and additional direct repeats within the *SCCmec* that may facilitate accrual of additional resistance genes via further transposition or inter-recombination events with homologous mobile elements⁵⁴.

All MRSA-colonized neonates had been admitted to the NICU for prematurity (Figure 2, Figure S1). Prematurity and associated long lengths of stay in the NICU elevate risks for MRSA colonization and infection^{4,5}. Importantly, NICU-specific outbreak strains were found in 68% of MRSA-colonized patients (15 of 22 neonates). Prospective genomic surveillance thus offers a tool to rapidly identify related strains, aid the discovery of risk factors for transmission, define potential environmental or staff reservoirs, and thus inform interventions to halt further spread.

We demonstrate the application of longitudinal genomic surveillance in a NICU setting using international repositories of deposited strains to improve outbreak detection, including cryptic transmission events, and potential pathogen reservoirs. The addition of robust and genomically diverse information improves the resolution of analyses and better enables critical rule-out determinations. Analyses also provide global context for uncommon strains that may be seen in an increasingly connected world.

Methods

IRB study protocol and data collection

The study was conducted under IRB protocol 2011-P-002883 (PI: Bry, Partners Healthcare). The Crimson LIMS was used for sample retrieval, and patient data and microbiological test results were retrieved using the Partners Research Patient Data Registry^{55,56}. Data used for analyses were deidentified.

Study design and clinical interventions

From days 1-341 MRSA isolates flagged from prior clinical microbiologic testing or nasal-perirectal swab screening were saved for phenotypic analyses of susceptibility profiles to evaluate suspicious outbreak clusters. Findings triggered a prospective genomic surveillance period from days 342 to 558 during which MRSA isolates from all nasal-perirectal swab cultures and diagnostic cultures were sequenced, in addition to the previously collected strains. Medical records for NICU patients during the surveillance period were evaluated for clinical diagnoses, reasons for NICU admission, and additional culture-based analyses. In the context of the clinical genomic surveillance program, only patient surveillance and diagnostic strains were evaluated. Non-clinical samples were not collected from patient family members or NICU personnel. Neonates colonized or infected with MRSA were placed in single rooms and managed with Contact Precautions.

Sample collection

BD BBL CultureSwabs (BD, San Jose, CA) were used for nasal-perirectal (NP) screening for *S. aureus*. Swabs were cultured to colistin nalidixic acid (CNA) blood agar for 18-24 hours at 37°C.

S. aureus colonies were identified by colony morphology, Gram stain and positive catalase and coagulase assays (Thermo). Diagnostic cultures with putative *S. aureus* colonies were evaluated similarly. One MRSA isolate from a patient that tested positive on a single nasal-perirectal screening swab was missed during the surveillance period and excluded from the study. Collected isolates were preserved at -80° C in 2ml Cryovials (Corning, Corning, NY). In the context of the clinical genomic surveillance program, and the IRB protocol under which it operates, only patient surveillance and diagnostic strains were evaluated. Non-clinical samples were not collected from patient family members or NICU personnel.

Antibiotic resistance testing

S. aureus isolates were tested clinically for antibiotic resistance by VITEK 2 with Susceptibility Card AST-GP78 (Biomérieux, Durham, NC). Microdilution testing for fusidic acid resistance was done by broth microdilution according to CLSI guidelines⁵⁷. MRSA isolates were called by resistance to ceftioxin.

Genome sequencing

All isolate genomes were sequenced using Illumina NextSeq as described (Illumina, San Diego, CA)⁵⁶. Oxford nanopore sequencing (Oxford Nanopore, Cambridge, MA) and subsequent assembly using a combination of nanopore and Illumina reads was performed as previously described for MLST 1535 isolates⁵⁸. Datasets were deposited under NCBI Bioproject PRJNA278886.

Bioinformatic analyses

Patient data was visualized and analyzed using Python with the packages Matplotlib and SciPy^{59,60}. Isolate genome sequencing reads were uploaded to the NCBI Sequence Read Archive for inclusion in the Pathogen Detection program (Supplemental table 1), which includes SNP clustering, virulence gene detection, antimicrobial resistance gene detection, and isolate metadata⁶¹. Phylogenetic results were visualized using the Pathogen Detection Isolates Browser.

Whole-genome MLST was performed using the program tblastn and coding sequences from RefSeq genome NC_007795.1^{62,63}. The threshold for inclusion was an e-value ≤ 0.001 and a length between 0.5x and 1.5x the reference sequence coding length. Gene matches were retained if there was at most one high quality hit in each genome for all genomes in a downsampled set of all *S. aureus* isolate genomes available in the NCBI Pathogen Detection program, limited to 5 from each SNP cluster. Gene sequences were only kept if they matched the reference sequence length to allow for efficient SNP calculation across thousands of genomes. Genes were excluded if present in <90% of genomes, resulting in a database of 2021 genes. Genomes with <90% of these genes represented within the criteria above were excluded, resulting in a final database representing 15,477 genomes. From this set of genomic information, allelic differences were calculated to determine the closest set of 25 genomes to NICU MRSA isolates, keeping all tied for 25th closest. SNP matrices were then created using the gene encoding sequences among these ≥ 25 genomes plus the NICU MRSA isolates. Phylogenies from these matrices were calculated with RAxML 8.2.11 using 1000 bootstraps and the GTR substitution matrix⁶⁴.

Each strain's MLST was calculated using the reference sequences and definitions from pubMLST.org⁶⁵. SCCmec types were determined using the tools and guidelines from SCCmec Finder⁶⁶. Virulence gene profiles and antibiotic resistance gene profiles were taken from the analysis at the NCBI Pathogen Detection Program⁶¹.

297 Annotation of the MLST 1535 *SCCmec* region was performed as follows. First, the
298 genome was annotated by the PATRIC RASTtk-enabled Genome Annotation Service⁶⁷. Open
299 reading frames within this sequence were confirmed or identified using blastp searches against
300 the non-redundant protein sequences database at NCBI⁶³. Cassette chromosome recombinase
301 (*ccr*) gene identity was investigated using blastp and representative alleles using hits above 80%
302 identity and length^{40,68}.

Acknowledgements: The authors wish to thank Mary Delaney for clinical microbiologic support, the Wadsworth Center Advanced Genomic Technologies Center (AGTC) for sequencing, and Beth Flanigan, MD for helpful review of results. The work of JW was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health, and the work of JC by the BWH Clinical Microbiology Fellowship. Studies were supported by P30 DK034854 (Bry). Sequencing at the AGTC was supported by FDA Cooperative Agreements 5U18FD006229 and 1U18FD006763 (Wolfgang).

References

- 1 Sievert, D. M. *et al.* Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infection Control and Hospital Epidemiology* **34**, 1-14, doi:10.1086/668770 (2013).
- 2 Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L. & Fowler, V. G. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clinical Microbiology Reviews* **28**, 603-661, doi:10.1128/CMR.00134-14 (2015).
- 3 Nelson, M. U. & Gallagher, P. G. Methicillin-resistant Staphylococcus aureus in the neonatal intensive care unit. *Seminars in Perinatology* **36**, 424-430, doi:10.1053/j.semperi.2012.06.004 (2012).
- 4 Huang, Y.-C., Chou, Y.-H., Su, L.-H., Lien, R.-I. & Lin, T.-Y. Methicillin-resistant Staphylococcus aureus colonization and its association with infection among infants hospitalized in neonatal intensive care units. *Pediatrics* **118**, 469-474, doi:10.1542/peds.2006-0254 (2006).
- 5 Zervou, F. N., Zacharioudakis, I. M., Ziakas, P. D. & Mylonakis, E. MRSA colonization and risk of infection in the neonatal and pediatric ICU: a meta-analysis. *Pediatrics* **133**, e1015-1023, doi:10.1542/peds.2013-3413 (2014).
- 6 Gerber, S. I. *et al.* Management of outbreaks of methicillin-resistant Staphylococcus aureus infection in the neonatal intensive care unit: a consensus statement. *Infection Control and Hospital Epidemiology* **27**, 139-145, doi:10.1086/501216 (2006).
- 7 Milstone, A. M. *et al.* Effect of treating parents colonized with Staphylococcus aureus on transmission to neonates in the intensive care unit: a randomized clinical trial. *JAMA* **323**, 319-328, doi:10.1001/jama.2019.20785 (2020).
- 8 Gastmeier, P. *et al.* Outbreaks in neonatal intensive care units - they are not like others. *American Journal of Infection Control* **35**, 172-176, doi:10.1016/j.ajic.2006.07.007 (2007).
- 9 Gastmeier, P., Sohr, D., Geffers, C., Behnke, M. & Rüden, H. Risk factors for death due to nosocomial infection in intensive care unit patients: findings from the Krankenhaus Infektions Surveillance System. *Infection Control and Hospital Epidemiology* **28**, 466-472, doi:10.1086/510810 (2007).
- 10 Conceição, T. *et al.* Staphylococcus aureus reservoirs and transmission routes in a Portuguese neonatal intensive care unit: a 30-month surveillance study. *Microbial Drug Resistance (Larchmont, N.Y.)* **18**, 116-124, doi:10.1089/mdr.2011.0182 (2012).
- 11 Dong, Y., Glaser, K. & Speer, C. P. New threats from an old foe: methicillin-resistant Staphylococcus aureus infections in neonates. *Neonatology* **114**, 127-134, doi:10.1159/000488582 (2018).
- 12 Sakaki, H., Nishioka, M., Kanda, K. & Takahashi, Y. An investigation of the risk factors for infection with methicillin-resistant Staphylococcus aureus among patients in a neonatal intensive care unit. *American Journal of Infection Control* **37**, 580-586, doi:10.1016/j.ajic.2009.02.008 (2009).

- 13 Washam, M., Woltmann, J., Haberman, B., Haslam, D. & Staat, M. A. Risk factors for methicillin-resistant *Staphylococcus aureus* colonization in the neonatal intensive care unit: a systematic review and meta-analysis. *American Journal of Infection Control* **45**, 1388-1393, doi:10.1016/j.ajic.2017.06.021 (2017).
- 14 Raboud, J. *et al.* Modeling transmission of methicillin-resistant *Staphylococcus aureus* among patients admitted to a hospital. *Infection Control and Hospital Epidemiology* **26**, 607-615, doi:10.1086/502589 (2005).
- 15 Laing, I. A., Gibb, A. P. & McCallum, A. Controlling an outbreak of MRSA in the neonatal unit: a steep learning curve. *Archives of Disease in Childhood. Fetal and Neonatal Edition* **94**, F307-310, doi:10.1136/ad.2008.151290 (2009).
- 16 Geva, A. *et al.* Spread of methicillin-resistant *Staphylococcus aureus* in a large tertiary NICU: network analysis. *Pediatrics* **128**, e1173-1180, doi:10.1542/peds.2010-2562 (2011).
- 17 Price, J. R. *et al.* Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of *Staphylococcus aureus* in an intensive care unit. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **58**, 609-618, doi:10.1093/cid/cit807 (2014).
- 18 Bertin, M. L. *et al.* Outbreak of methicillin-resistant *Staphylococcus aureus* colonization and infection in a neonatal intensive care unit epidemiologically linked to a healthcare worker with chronic otitis. *Infection Control and Hospital Epidemiology* **27**, 581-585, doi:10.1086/504933 (2006).
- 19 Méan, M. *et al.* A neonatal specialist with recurrent methicillin-resistant *Staphylococcus aureus* (MRSA) carriage implicated in the transmission of MRSA to newborns. *Infection Control and Hospital Epidemiology* **28**, 625-628, doi:10.1086/513616 (2007).
- 20 Price, J. R. *et al.* Transmission of *Staphylococcus aureus* between health-care workers, the environment, and patients in an intensive care unit: a longitudinal cohort study based on whole-genome sequencing. *The Lancet. Infectious Diseases* **17**, 207-214, doi:10.1016/S1473-3099(16)30413-3 (2017).
- 21 Morel, A.-S. *et al.* Nosocomial transmission of methicillin-resistant *Staphylococcus aureus* from a mother to her preterm quadruplet infants. *American Journal of Infection Control* **30**, 170-173, doi:10.1067/mic.2002.119819 (2002).
- 22 Al-Tawfiq, J. A. Father-to-infant transmission of community-acquired methicillin-resistant *Staphylococcus aureus* in a neonatal intensive care unit. *Infection Control and Hospital Epidemiology* **27**, 636-637, doi:10.1086/505097 (2006).
- 23 Cheng, V. C. C. *et al.* Whole-genome sequencing data-based modeling for the investigation of an outbreak of community-associated methicillin-resistant *Staphylococcus aureus* in a neonatal intensive care unit in Hong Kong. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology* **38**, 563-573, doi:10.1007/s10096-018-03458-y (2019).
- 24 Albrich, W. C. & Harbarth, S. Health-care workers: source, vector, or victim of MRSA? *The Lancet. Infectious Diseases* **8**, 289-301, doi:10.1016/S1473-3099(08)70097-5 (2008).
- 25 Lessa, F. C. *et al.* Trends in incidence of late-onset methicillin-resistant *Staphylococcus aureus* infection in neonatal intensive care units: data from the National Nosocomial

- Infections Surveillance System, 1995-2004. *The Pediatric Infectious Disease Journal* **28**, 577-581, doi:10.1097/INF.0b013e31819988bf (2009).
- 26 Köser, C. U. *et al.* Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *The New England Journal of Medicine* **366**, 2267-2275, doi:10.1056/NEJMoa1109910 (2012).
- 27 Price, J. R., Didelot, X., Crook, D. W., Llewelyn, M. J. & Paul, J. Whole genome sequencing in the prevention and control of *Staphylococcus aureus* infection. *The Journal of Hospital Infection* **83**, 14-21, doi:10.1016/j.jhin.2012.10.003 (2013).
- 28 Azarian, T. *et al.* Whole-genome sequencing for outbreak investigations of methicillin-resistant *Staphylococcus aureus* in the neonatal intensive care unit: time for routine practice? *Infection Control and Hospital Epidemiology* **36**, 777-785, doi:10.1017/ice.2015.73 (2015).
- 29 Roisin, S. *et al.* Pan-genome multilocus sequence typing and outbreak-specific reference-based single nucleotide polymorphism analysis to resolve two concurrent *Staphylococcus aureus* outbreaks in neonatal services. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* **22**, 520-526, doi:10.1016/j.cmi.2016.01.024 (2016).
- 30 Lakhundi, S. & Zhang, K. Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. *Clinical Microbiology Reviews* **31**, doi:10.1128/CMR.00020-18 (2018).
- 31 SenGupta, D. J. *et al.* Whole-genome sequencing for high-resolution investigation of methicillin-resistant *Staphylococcus aureus* epidemiology and genome plasticity. *Journal of Clinical Microbiology* **52**, 2787-2796, doi:10.1128/JCM.00759-14 (2014).
- 32 Harris, S. R. *et al.* Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *The Lancet. Infectious Diseases* **13**, 130-136, doi:10.1016/S1473-3099(12)70268-2 (2013).
- 33 Layer, F. *et al.* Molecular typing of toxic shock syndrome toxin-1- and Enterotoxin A-producing methicillin-sensitive *Staphylococcus aureus* isolates from an outbreak in a neonatal intensive care unit. *International journal of medical microbiology: IJMM* **305**, 790-798, doi:10.1016/j.ijmm.2015.08.033 (2015).
- 34 Earls, M. R. *et al.* Intra-hospital, inter-hospital and intercontinental spread of ST78 MRSA from two neonatal intensive care unit outbreaks established using whole-genome sequencing. *Frontiers in Microbiology* **9**, 1485, doi:10.3389/fmicb.2018.01485 (2018).
- 35 Sullivan, M. J. *et al.* A complete genome screening program of clinical methicillin-resistant *Staphylococcus aureus* isolates identifies the origin and progression of a neonatal intensive care unit outbreak. *Journal of Clinical Microbiology* **57**, doi:10.1128/JCM.01261-19 (2019).
- 36 Nübel, U. *et al.* MRSA transmission on a neonatal intensive care unit: epidemiological and genome-based phylogenetic analyses. *PloS One* **8**, e54898, doi:10.1371/journal.pone.0054898 (2013).
- 37 Prosperi, M. *et al.* Molecular epidemiology of community-associated methicillin-resistant *Staphylococcus aureus* in the genomic era: a cross-sectional study. *Scientific Reports* **3**, 1902, doi:10.1038/srep01902 (2013).

- 38 Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrobial Agents and Chemotherapy* **53**, 4961-4967, doi:10.1128/AAC.00579-09 (2009).
- 39 Lin, Y.-T. *et al.* A novel staphylococcal cassette chromosomal element, SCCfusC, carrying fusC and speG in fusidic acid-resistant methicillin-resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* **58**, 1224-1227, doi:10.1128/AAC.01772-13 (2014).
- 40 Senok, A. C. *et al.* Investigating a rare methicillin-resistant *Staphylococcus aureus* strain: first description of genome sequencing and molecular characterization of CC15-MRSA. *Infection and Drug Resistance* **10**, 307-315, doi:10.2147/IDR.S145394 (2017).
- 41 Ankrum, A. & Hall, B. G. Population dynamics of *Staphylococcus aureus* in cystic fibrosis patients to determine transmission events by use of whole-genome sequencing. *Journal of Clinical Microbiology* **55**, 2143-2152, doi:10.1128/JCM.00164-17 (2017).
- 42 Worley, J. *et al.* Genomic determination of relative risks for *Clostridioides difficile* infection from asymptomatic carriage in ICU patients. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, doi:10.1093/cid/ciaa894 (2020).
- 43 Orsi, R. H. *et al.* Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. *BMC Genomics* **9**, 539, doi:10.1186/1471-2164-9-539 (2008).
- 44 Narui, K. *et al.* Change in environmental bacterial flora in a new hospital building. *The Journal of Hospital Infection* **73**, 24-33, doi:10.1016/j.jhin.2009.05.019 (2009).
- 45 Stasiewicz, M. J., Oliver, H. F., Wiedmann, M. & den Bakker, H. C. Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Applied and Environmental Microbiology* **81**, 6024-6037, doi:10.1128/AEM.01049-15 (2015).
- 46 Williamson, D. A. *et al.* High usage of topical fusidic acid and rapid clonal expansion of fusidic acid-resistant *Staphylococcus aureus*: a cautionary tale. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **59**, 1451-1454, doi:10.1093/cid/ciu658 (2014).
- 47 Cabrera, R. *et al.* Molecular characterization of methicillin-resistant *Staphylococcus aureus* clinical strains from the endotracheal tubes of patients with nosocomial pneumonia. *Antimicrobial Resistance and Infection Control* **9**, 43, doi:10.1186/s13756-020-0679-z (2020).
- 48 Senok, A. *et al.* Genotyping of methicillin resistant *Staphylococcus aureus* from the United Arab Emirates. *Scientific Reports* **10**, 18551, doi:10.1038/s41598-020-75565-w (2020).
- 49 Ellington, M. J. *et al.* Emergent and evolving antimicrobial resistance cassettes in community-associated fusidic acid and methicillin-resistant *Staphylococcus aureus*. *International Journal of Antimicrobial Agents* **45**, 477-484, doi:10.1016/j.ijantimicag.2015.01.009 (2015).
- 50 Baines, S. L. *et al.* Rapid emergence and evolution of *Staphylococcus aureus* clones harboring fusC-containing staphylococcal cassette chromosome elements. *Antimicrobial Agents and Chemotherapy* **60**, 2359-2365, doi:10.1128/AAC.03020-15 (2016).

481 51 Senok, A. *et al.* Characterisation of a novel SCCmec VI element harbouring fusC in an
482 emerging *Staphylococcus aureus* strain from the Arabian Gulf region. *PLoS ONE* **14**,
483 doi:10.1371/journal.pone.0223985 (2019).

484 52 Monecke, S. *et al.* Characterisation of MRSA strains isolated from patients in a hospital
485 in Riyadh, Kingdom of Saudi Arabia. *BMC Microbiology* **12**, 146, doi:10.1186/1471-2180-
486 12-146 (2012).

487 53 Udo, E. E. *et al.* Emergence of methicillin-resistant *Staphylococcus aureus* belonging to
488 clonal complex 15 (CC15-MRSA) in Kuwait hospitals. *Infection and Drug Resistance* **13**,
489 617-626, doi:10.2147/IDR.S237319 (2020).

490 54 Larsen, J., Andersen, P. S., Winstel, V. & Peschel, A. *Staphylococcus aureus* CC395
491 harbours a novel composite staphylococcal cassette chromosome mec element. *The*
492 *Journal of Antimicrobial Chemotherapy* **72**, 1002-1005, doi:10.1093/jac/dkw544 (2017).

493 55 Murphy, S. *et al.* Instrumenting the health care enterprise for discovery research in the
494 genomic era. *Genome Research* **19**, 1675-1681, doi:10.1101/gr.094615.109 (2009).

495 56 Nudel, K. *et al.* Genomics of *Corynebacterium striatum*, an emerging multi-drug
496 resistant pathogen of immunocompromised patients. *Clinical microbiology and infection*
497 : the official publication of the European Society of Clinical Microbiology and Infectious
498 Diseases **24**, 1016.e1017-1016.e1013, doi:10.1016/j.cmi.2017.12.024 (2018).

499 57 CLSI. *Methods for dilution antimicrobial susceptibility tests for bacteria that grow*
500 *aerobically*. 11 edn, Vol. M07 (Clinical and Laboratory Standards Institute, 2018).

501 58 Gonzalez-Escalona, N. & Sharma, S. K. Closing *Clostridium botulinum* group I genomes
502 using a combination of short- and long-reads. *Frontiers in Microbiology* **11**,
503 doi:10.3389/fmicb.2020.00239 (2020).

504 59 Hunter, J. D. Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*
505 **9**, 90-95, doi:10.1109/MCSE.2007.55 (2007).

506 60 Virtanen, P. *et al.* SciPy 1.0 – fundamental algorithms for scientific computing in Python.
507 *Nature Methods* **17**, 261-272 (2020).

508 61 (National Library of Medicine (US), National Center for Biotechnology Information,
509 2016).

510 62 Gillaspay, A. F. *et al.* in *Gram-Positive Pathogens* 381-412 (John Wiley & Sons, Ltd,
511 2006).

512 63 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC bioinformatics* **10**, 421,
513 doi:10.1186/1471-2105-10-421 (2009).

514 64 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
515 large phylogenies. *Bioinformatics (Oxford, England)* **30**, 1312-1313,
516 doi:10.1093/bioinformatics/btu033 (2014).

517 65 Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics:
518 BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open*
519 *Research* **3**, doi:10.12688/wellcomeopenres.14826.1 (2018).

520 66 Kaya, H. *et al.* SCCmecFinder, a web-based tool for typing of staphylococcal cassette
521 chromosome mec in *Staphylococcus aureus* using whole-genome sequence data.
522 *mSphere* **3**, doi:10.1128/mSphere.00612-17 (2018).

523 67 Brettin, T. *et al.* RASTtk: A modular and extensible implementation of the RAST
524 algorithm for building custom annotation pipelines and annotating batches of genomes.
525 *Scientific Reports* **5**, doi:10.1038/srep08365 (2015).
526 68 Monecke, S. *et al.* Diversity of SCCmec elements in *Staphylococcus aureus* as observed
527 in South-Eastern Germany. *PloS One* **11**, e0162654, doi:10.1371/journal.pone.0162654
528 (2016).
529

Figure 1. NICU outbreak cluster occurrence: X-axis indicates day during the analysis period. "Targeted Outbreak Sampling" refers to microbiologic-only assessments of NICU strains from days 0-341. Triangles and crosses indicate all occurring strains among patients that were also included in later genomic analyses. "Prospective surveillance" over days 342-558 indicates the period of prospective genomic analyses of new strains. Y-axis indicates NICU MRSA strains occurring in six NCBI, Pathogen Detection SNP clusters; numbers in parentheses indicate number of NICU strains from this study that occurred in the broader NCBI cluster, as follows: SNP Clusters (MLST, SCCmec type): **1**, PDS000067849.1 (8, IVa(2B)); **2**, PDS000069849.1 (72, IVa(2B)); **3**, PDS000069850.1 (5, II(2A)); **4**, PDS000069851.1 (5, IVa(2B)); **5**, PDS000069882.1 (1535, unclassified SCCmec); **6**, PDS000069883.1 (8, IVa(2B)); Unc. = unclustered isolates with unrelated MLST and SCCmec types. SNP cluster 2 (NCBI outbreak cluster PDS000069849.1) contained only one isolate from this study (Supplemental File 2). Red dashed lines show a suspected outbreak during the, and a black dashed line marks when the Prospective Surveillance began.

Figure 2. SNP cluster analyses reveal patient outbreaks and occurrence over time. Panels A-D show NICU MRSA strain phylogenetic trees from the NCBI Pathogen Genomes isolates browser with addition of the anonymized patientID and day of isolation. Graphs below each tree show the associated patients and strain occurrences relative to surveillance or diagnostic cultures. Though patient isolates fell within the broader NCBI SNP clusters, no other closely related strains from other institutions fell within these regions of the cluster SNP tree. Bar in each graph indicates 5 SNPs distance.

Figure 3. International MRSA strain datasets support effective rule out of local NICU suspected transmission events. Phylogenetic tree of NCBI ST5 SNP cluster isolates showing the NCBI SNP cluster, MLST and genomic SCCmec. The ST5 SNP clusters were downsampled to 5 representative isolates and wgMLST distances were used to define the 25 closest related strains in the NCBI Pathogen Detection program. SNPs within the wgMLST genes were used to calculate phylogenetic relationships (Supplemental File 2). The outgroup strain (green) was most distantly related to the BWH strains by wgMLST SNPs. Bootstrap values are shown at phylogenetic nodes. NICU patient strains are shown in red, including 5 representative isolates from the 4 NICU patients in Cluster 4, and the additional MRSA isolate from patient 21. Bootstrap support values show placement of patient 21's isolate in a separate sub-cluster distinct from Cluster 4.

A

■ = BWH Submission (Cluster E)
■ = Outgroup

0.00001

B

□ = Absent genes
■ = present non-SCCmec genes
■ = Present SCCmec genes

SNP Cluster

C

Figure 4. Genomic analyses identify NICU MLST 1535 strains with international clusters similar resistance gene profiles and variant SCCmecs. **A.** SNP based phylogeny of a wgMLST defined group of related strains reveals a clade of MLST 1535 strains from NICU Cluster 5 (red) that occurred in two patients. Strains were most similar to others identified in Saudi Arabia and one from Cincinnati. Strains outside of the MLST 1535 box are MLST 15. **B.** Antibiotic resistance gene profiles show the stepwise acquisition of resistance genes. The three NICU patient strains have lost *lnuA* and *aac(6')-le/aph(2'')-la*. **C.** A high-resolution genomic map of the SCCmec region from NCBI cluster 5 isolates closed by long-read genome sequencing. A 13 base AAAC TAAAAAAT core sequence (DR1-DR3) with variable adjacent homologous sequence is directly repeated (DR) three times. This sequence does not appear elsewhere in the genome. Gradients in the IS element keys show directionality of the similar sequences.