Characterisation of Staphylococci species from neonatal blood cultures in low- and middle-income countries

Kirsty Sands¹,²*, Maria J. Carvalho¹,³, Owen B. Spiller¹, Edward A. R. Portal¹, Kathryn Thomson¹,², William John Watkins¹, Jordan Mathias¹, Calie Dyer¹,⁴, Chinenyе Akpulu¹,²,⁵, Robert Andrews¹, Ana Ferreira¹, Thomas Hender¹, Rebecca Milton¹,⁴, Maria Nieto¹, Rabaab Zahra⁶, Haider Shirazi⁷, Adil Muhammad⁶, Shermeen Akif⁶, Muhammad Hilal Jan⁶, Kenneth Iregbu⁵, Fatima Modibbo⁸, Stella Uwaezuoke⁹, Grace J. Chan¹⁰,¹¹, Delayehu Bekele¹², Semaria Solomon¹³, Sulagna Basu¹⁴, Ranjan Kumar Nandy¹⁴, Sharmi Naha¹⁴, Jean-Baptiste Mazarati¹⁵, Aniceth Rucogoza¹⁵, Lucie Gaju¹⁵, Shaheen Mehtar¹⁶,¹⁷, Andre N. H. Bulabula¹⁷,¹⁸, Andrew Whitelaw¹⁹,²⁰, BARNARDS Group and Timothy R. Walsh¹,²

Abstract

Background: In low- and middle-income countries (LMIC) Staphylococcus aureus is regarded as one of the leading bacterial causes of neonatal sepsis, however there is limited knowledge on the species diversity and antimicrobial resistance caused by Gram-positive bacteria (GPB).

Methods: We characterised GPB isolates from neonatal blood cultures from LMICs in Africa (Ethiopia, Nigeria, Rwanda, and South Africa) and South-Asia (Bangladesh and Pakistan) between 2015–2017. We determined minimum inhibitory concentrations and performed whole genome sequencing (WGS) on Staphylococci isolates recovered and clinical data collected related to the onset of sepsis and the outcome of the neonate up to 60 days of age.

Results: From the isolates recovered from blood cultures, Staphylococci species were most frequently identified. Out of 100 S. aureus isolates sequenced, 18 different sequence types (ST) were found which unveiled two small epidemiological clusters caused by methicillin resistant S. aureus (MRSA) in Pakistan (ST8) and South Africa (ST5), both with high mortality (n = 6/17). One-third of S. aureus was MRSA, with methicillin resistance also detected in Staphylococcus epidermidis, Staphylococcus haemolyticus and Mammaliicoccus sciuri. Through additional WGS analysis we report a cluster of M. sciuri in Pakistan identified between July–November 2017.

Conclusions: In total we identified 14 different GPB bacterial species, however Staphylococci was dominant. These findings highlight the need of a prospective genomic epidemiology study to comprehensively assess the true burden of GPB neonatal sepsis focusing specifically on mechanisms of resistance and virulence across species and in relation to neonatal outcome.

Keywords: LMIC, Staphylococci, Neonatal sepsis, Genomics, Mammaliicocci, Mortality, Early onset, Late onset, Epidemiology

*Correspondence: kirsty.sands@zoo.ox.ac.uk

1 Department of Medical Microbiology, Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, UK

Full list of author information is available at the end of the article

© The Author(s) 2022. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Background
Infections are a major cause of neonatal mortality [1, 2] with over 200,000 deaths reported each year [3]. Globally there is a large variation in the incidence of neonatal sepsis, although the burden is greatest in low- and middle-income countries (LMICs) [4, 5]. Despite *Staphylococcus aureus* being frequently reported as one of the top three bacterial causes of neonatal sepsis in LMICs along with GNB including *Klebsiella pneumoniae* and *Escherichia coli* [2, 6–9], surveillance data particularly using whole genome sequencing (WGS) is scarce. Understanding the true burden of coagulase negative Staphylococci (CoNS) in neonatal sepsis is far more complex. Despite multiple studies over the past decade suggesting that *Staphyloccus epidermidis* and *Staphylococcus haemolyticus* are important causes of nosocomial and bloodstream infections in neonates [10–13], there is often little capacity to implement routine surveillance, and these species are in many cases still considered contaminants, as they commonly reside on skin. Studies have shown that neonates born at a very low birth weight (VLBW, <1,500 g) may have a greater risk of contracting CoNS [11, 14]. Over the past two decades the number of specialised neonatal units in LMICs has fortunately increased, further increasing survival of VLBW infants. However, the likelihood of neonates exposed to nosocomial acquired CoNS is therefore higher, due to increased hospital stay for VLBW infants. Microbial investigation of CoNS isolated from blood cultures in countries with limited resources is often problematic due to a lack of simple and cost-effective methods of identification [9, 15]. Even if laboratories are equipped with manual or automated biochemical tests, these are laborious leading to subjective and inaccurate results; the tests add to the laboratory costs, and this may be difficult to justify given the often uncertain significance of the isolates. Although studies have reported cases of *Staphylococcus epidermidis* and *Staphylococcus haemolyticus* neonatal sepsis [9, 12, 13] and a PubMed search of “Staphylococci neonatal sepsis” produced over 300 hits (the majority of studies conducted in high-income countries) there is still a lack of data revealing the genomic diversity and resistome and virulome of staphylococcal species in LMICs.

BARNARDS is a network of 12 clinical sites distributed among four African (Ethiopia, Nigeria, Rwanda, South Africa) and three South-Asian (Bangladesh, India, Pakistan) countries aiming to assess the burden of antimicrobial resistance (AMR) in neonates. Herein we report data on sepsis caused by Gram-positive bacteria (GPB). We focus on the reporting of genomic diversity, genomic traits including antimicrobial resistance and virulence genes, and the phenotypic susceptibility profiles against 14 antibiotics for *S. aureus* isolates. Where possible, we performed genomics analysis on CoNS isolates to disclose AMR genotypes and whether mobile genetic elements were present. Finally, using WGS we report on an epidemiological cluster of neonatal sepsis caused by *Mammallolococcus sciuri*, previously known as *Staphylococcus sciuri* in Islamabad, Pakistan between July-November 2017.

Methods

Study design
Ethical approval was obtained prior to the start of the study (Additional file 1: Table S1). We have used the term ‘neonate’ for all enrolments, this includes up to 60 days of age. Upon clinical signs of sepsis, a blood culture (using automated blood culture systems either BD BACTEC™ or BacT/ALERT® bioMérieux) was collected and processed using automated blood culture systems. Early onset sepsis (EOS) was defined as sepsis occurring <72 h and late onset sepsis (LOS) >72 h after birth. Neonates were followed up until 60 days. Clinical data collected included onset and outcome of sepsis. Due to high reporting of GPB sepsis and in particular *S. aureus* during a BARNARDS network meeting in 2017 (a program originally designed with an emphasis on carbapenem resistance in GNB neonatal sepsis) we retrospectively recovered GPB isolates for genomics characterisation in addition to GNB [16]. Unfortunately, at this point, 970/1266 GPB isolates, predominantly non-*S. aureus* were not stored or recoverable. Blood isolates were stored on charcoal swabs (Deltalab) for transport under UN3373 regulations to Cardiff University, UK.

Antimicrobial susceptibility testing
In Cardiff, GPB were plated onto Columbia Blood Agar supplemented with 5% sterile sheep/horse blood and incubated aerobically at 37 °C for 24–48 h. Isolates were identified using Microflex LT MALDI-TOF MS (Bruker Daltonik, GmbH, UK) with α-Cyano-4-hydroxycinnamic acid (HCCA) matrix (Sigma Aldrich) and stored in TS/72 beads (Technical Service Consultants, UK) at -80 °C. Minimum inhibitory concentrations (MICs) were determined by in-house agar dilution for 14 antibiotics (Additional file 1: Table S2) using *Staphylococcus aureus* ATCC 29,213, *Escherichia coli* ATCC 25,922 and *Pseudomonas aeruginosa* ATCC 2785 for quality control and interpreted according to the EUCAST v11 guidelines [17]. The MIC₉₀ and MIC₅₀ for each antibiotic was determined for *S. aureus*. Each bacterial isolate was tested once per antibiotic concentration to generate a single data point following validation using appropriate control strains listed above.
Whole genome sequencing and bioinformatics analysis

Assembly and annotation

WGS, assembly and annotation were performed as previously described [16], and individual accession numbers are listed in Additional file 1: Table S3. On average, isolates were paired-end sequenced to > 40 × coverage using Nextera XT v2 on an Illumina Miseq generating 300 bp reads. A summary of the bioinformatics programs and parameters used in this study can be found in Additional file 1: Fig. S1. Genome coverage was assessed using Quail-map (v 2.2.1) [18]. ABRicate (v 0.9.7) [19] (> 98% coverage and identity) was used to detect antimicrobial resistance genes (ARG) [20] and virulence factors (VF) [21]. Species identification was confirmed using BLAST v2.7.1 [22]. In silico MLST was determined for S. aureus, S. epidermidis and S. haemolyticus using mlst (v 2.17.6) [23]. Novel alleles and profiles were submitted using PubMLST [24]. Staphyloccocus aureus specific staphylococcal protein A (spa) gene typing, IS256 screening, and SCCmec cassette determination was performed using CGE [25] in combination with the following databases; spa typer (v 1.0) [26], MGE (v 1.0.2) [27] and SCCmecFinder (1.2) [28].

Comparative genomics

S. aureus (n = 251) from a European collection by Aanensen et al., 2016 [29] were incorporated into a species wide analysis using Prokka (v 1.14.0) [30], Roary (v 3.12.0) [31] and FastTree (v 2.1.11) [32]. For all core genome SNP analyses performed (depicted in Additional file 1: Fig. S1), a representative reference of each dominant ST was selected from the genomes submitted to NCBI marked as ‘complete’, for variant calling using Snippy (v 4.0.5) [33], Gubbins (v 2.3.4) [34] snp-sites (v 2.5.1) [35] and Raxml-ng (v 0.9.0) [36] and IQ-tree (v 2.0) where specified [37]. SNP pairwise distances were generated using snp-dist (v 0.6) [38]. Phylogenetic trees were mid-rooted and annotated using iTOL (v 5.7) [39]. In total n = 2,327 S. epidermidis and n = 310 S. haemolyticus genomes were downloaded from NCBI on 28th October 2020. Strains of the same ST (assessed using mlst v 2.17.6 [23]) of those found within this study were retained for comparative analysis. Metadata on whether these were reported as clinical, commensal, or other was recorded (Additional file 1: Table S4). SNP phylogeny workflows were performed as described above.

WGS analysis of Mammaliicoccus sciuri

An isolate belonging to the cluster was selected for complementary long read sequencing to produce a high-quality reference. gDNA was extracted and quantified as described [16], and SPRI beads (Mag-Bind TotalPure, Omega) were used to concentrate and purify at a 1:1 ratio with a final elution volume of 15 µL to achieve an optimal range between 40-60 ng/µL. gDNA was quantified with the dsDNA BR assay kit using the qubit 4.0 fluorometer (Thermofisher). Genomic libraries were prepared using the Rapid Barcoding Kit (SQK-RBK004; Oxford Nanopore), sequenced on a R9.4 flow cell using a MinION (Oxford Nanopore) and basecalling was performed using Guppy within MinKnow. Long reads were demultiplexed using Porechop (v 0.2.4) [40] and assembled with corresponding short reads generated from the Illumina MiSeq using Unicycler (v 0.4.7) [41]. Mash dist (mash v 2.2) [42] was performed on all M. sciuri genomes available in NCBI to determine genomic distance estimation to that of the cluster strain; the ten isolates with the greatest genomic similarity were incorporated into core SNP analysis as outlined previously. Metadata from NCBI was annotated within the phylogeny (Additional file 1: Table S5).

Statistical analysis

Statistical associations between clinical data (onset; EOS v LOS and outcome of sepsis; reported deceased v non reported deceased), phenotypic (MRSA) and genomic traits (ARG/ST/VF) were explored for S. aureus using Chi-Square (χ²) tests, comparing means T test and the Log Rank test SPSS (version 25.0.0.1). Kaplan–Meier survival curves were produced using the outcome follow up data (up to 60 days, or the last follow up record was used) in SPSS (version 25.0.0.1). The presence/absence of S. aureus VFs were analysed according to clinical onset and outcome of sepsis. Firstly, VFs associated with sepsis [43] were counted to provide a tailored VF score and a Kaplan–Meier survival curve was produced using the presence of Panton-Valentine Leukocidin (PVL), chosen as a frequently used clinical marker of S. aureus virulence. In addition, the frequency of all VFs detected > 98% identity and coverage from VFDB was counted against the clinical onset and outcome of sepsis to look for trends between EOS/LOS and alive/deceased. Statistical significance was taken at P value ≤ 0.05, and the false discovery rate was controlled using the Benjamini–Hochberg procedure [44]. A presence/absence matrix of all VFs detected per isolate (output from ABRicate vfdb) was generated using splist2presabs in the fuzzysim package in R 3.6.2.

Results

Characteristics of Gram-positive blood cultures reported during BARNARDS

In total, 1,266 GPB sepsis isolates were recorded from 1,239 neonates compared to 1,038 GNB sepsis isolates recorded from 1,007 neonates (Additional file 1: Table S6, Additional file 1: Fig. S2 [16]). From 296 GPB
Fig. 1 Core genome phylogenetic analysis of Staphylococcus aureus including a global contextual analysis. a Detailed core genome characterisation of 100 Staphylococcus aureus isolates (BARNARDS) using Roary (v3.12.0) and Fasttree (v2.1.11). Isolate labels are coloured according to clinical site of origin. The in-silico sequence type (ST) is shown outside of the isolate code (leaf). Presence of meca, and whether the isolate was classified phenotypically as MRSA (as inferred from oxacillin MIC > 2 mg/l) is denoted by a filled triangle and/or circle respectively. Presence of IS256 is denoted by a filled rectangle. b Core genome characterisation of 351 Staphylococcus aureus isolates, incorporating a European collection [29] using Roary (v3.12.0) and Fasttree (v2.1.11). Coloured ranges in blue represent a S. aureus from the BARNARDS collection. Branch labels are coloured according to country of origin. Symbol represents source of isolate.
isolates available for identification, we identified 14 different species (Additional file 1: Fig. S3) and of these, 271 isolates were *Staphylococci* with 101 *Staphylococcus aureus* and 170 CoNS. Twenty-six isolates were identified as *Staphylococcus sciuri*, but this species has since been reclassified to *Mammaliicoccus sciuri*. Overall, the onset of sepsis was 36% (n = 62) EOS, 54% (n = 71) LOS and for 10% (n = 17) cases the onset was undetermined (Additional file 1: Fig. S4). For *S. aureus*, the onset was 40% (n = 40) EOS, 47% (n = 47) LOS and undetermined for 13% (n = 14, Additional file 1: Fig. S4). Both *S. epidermidis* and *S. haemolyticus* was more common in LOS whereas *M. sciuri* was more common in EOS (Additional file 1: Fig. S4), however this data is indicative only.

This study specifically reports the antimicrobial susceptibility profiles of 144 isolates (n = 100 *S. aureus* and n = 44 CoNS) and WGS of 130 *Staphylococci*.

---

**Fig. 2** Core genome phylogenetic analysis of coagulase negative *Staphylococcus* blood culture isolates. Core genome phylogeny of CoNS isolates displaying key genomic traits for comparison using Roary (v3.12.0) and Fasttree (v2.1.11). Isolate labels are coloured according to clinical site of origin. Clades are highlighted according to species. The in-silico sequence type (ST) is shown outside of the isolate code (leaf). Presence of mecA, and whether the isolate was classified phenotypically as MRSA (as inferred from oxacillin MIC > 2 mg/l) is denoted by a filled triangle and/or circle respectively. Presence of IS256 is denoted by a filled rectangle.
Mammaliicocci sepsis isolates, (n = 100 S. aureus and n = 30 CoNS) which we were able to recover for microbiology characterisation after storage at −80 °C.

**Staphylococcus aureus and CoNS phylogeny**

*S. aureus* for WGS characterisation included isolates predominantly from Pakistan (n = 42), Nigeria (n = 26), and South Africa (n = 16), along with nine from Bangladesh, two from Ethiopia and one from Rwanda, (n = 100 in total). In total, we detected 18 different sequence types (ST) (Fig. 1a) including the assignment of four novel STs; ST5426, ST5494, ST5945 and ST5496. We detected seven *S. aureus* clonal complex ST22 (CC22) from Bangladesh (n = 5), Pakistan (n = 1) and Rwanda (n = 1) and three ST5426 (within CC22) from Pakistan. Within BARNARDS, the commonest clonal complexes were CC5 (n = 36; 36%; consisting of ST16 [n = 18], ST5 [n = 17], ST650 [n = 1]) and CC8 (n = 13; 13%, all isolates belonged to ST8). Comparative genomic analyses including a previously known European collection, revealed ST5 isolates were distributed along 16 countries 45 (Fig. 1b). Interestingly, ST152 (n = 15) which is not assigned to a CC was the third most common ST identified. These isolates collectively represented 64% of the dataset and were isolated across the sampling period (Fig. 1a, Fig. 3a). All seventeen ST5 (Nigeria n = 6; South Africa, n = 11) were found carrying mecA (n = 12; 71%; Nigeria n = 3, South Africa n = 9; Fig. 2a) and the Nigerian and South African isolates grouped on different clades (Additional file 1: Fig. S5). Nine ST5 isolates collected from distinct South African neonates between July 2016 and July 2017 were within 10 pairwise SNPs (Supplementary Fig. 5), all were MRSA spa type t045, contained IS256 and virulence factors hlg, scn, spa and hly. Mortality following biological sepsis (MfBS) was reported for four out of these nine neonates. Additionally, 12 ST152 were isolated from African sites (Ethiopia n = 1, Nigeria n = 10 and South Africa n = 1, Fig. 1a, Additional file 1: Fig. S6) and were MSSA.

ST6 (CC5) and ST8 (CC8) were found in South Asia (ST6 Pakistan n = 17/18, Bangladesh n = 1/18 and ST8 Pakistan n = 13, Fig. 1b, Supplementary Figs. 7–8). In the species-wide phylogeny, ST8 was found in 11 European countries, most notably Belgium and France (Fig. 1b). Within BARNARDS we saw three major ST8 clades, of which two were predominantly MRSA.

(See figure on next page.)

**Fig. 3** Temporal frequency and survival curve data for *Staphylococcus aureus* blood culture isolates. a Stacked bar graph to show the temporal frequency of *S. aureus* isolates recovered from blood cultures during the BARNARDS sampling, per month. The bar graph is coloured according to the dominant STs, with all other STs being grouped as ‘Other’. b A Kaplan Meier survival plot comparing the four dominant ST groups (with all other STs detected grouped together as a single group called ‘other’) against the age at outcome for the neonate and up to 60 days. Findings are suggestive as the data presented is from a sample size (ST152 n = 15, ST other n = 37, ST5 n = 17, ST6 n = 18 and ST8 n = 13, overall comparison P 0.041)
Fig. 3  (See legend on previous page.)
isolates from blood cultures collected in Nigeria, both were collected from the same clinical site (NK), both were resistant to methicillin (Fig. 2) and had 707 pairwise SNPs.

We found that all 8 S. epidermidis isolates belonged to different STs, including the assignment of three novel STs (ST888-890), of which two (ST889 and ST890) were to different STs, including the assignment of three novel SNPs.

In Antimicrobial resistance
We found that all 8 S. aureus isolates belonged to different STs, including the assignment of three novel STs (ST888-890), of which two (ST889 and ST890) were to different STs, including the assignment of three novel SNPs.

| Antibiotic class | Antimicrobial | MIC(µg/mL) | Resistance breakpoint | % Resistant |
|------------------|---------------|------------|-----------------------|-------------|
| Aminoglycosides  | Amikacin      | 4–32       | >8                    | 40.0        |
|                  | Gentamicin    | 0.5–8      | >4                    | 18.0        |
|                  | Tobramycin    | 0.5–8      | >4                    | 29.0        |
| Fluoroquinolones | Ciprofloxacin | 0.5–>4     | >4                    | 20.0        |
|                  | Levofloxacin  | 0.5–>4     | >4                    | 15.0        |
| Glycopeptides    | Vancomycin    | 1–2        | >2                    | 0.0         |
|                  | MLS           | 1–>8       | >4                    | 3.0         |
|                  | Oxazolidinones| 2–8        | >4                    | 1.0         |
|                  | Rifampicin    | 0.03–0.03  | >0.5                  | 0.0         |
| Penicillins      | Ampicillin    | 2–128      | NA                    | NA          |
|                  | Flucloxacillin| 0.5–>8     | NA                    | NA          |
|                  | Oxacillin     | 1–>8       | >2                    | 33          |
| Tetracyclines    | Minocycline   | 0.25–>2    | >0.5                  | 7.0         |
|                  | Tigecycline   | 0.25–0.5   | >0.5                  | 0.0         |

Minimum inhibitory concentrations of antibiotics were determined by agar dilution. For S. aureus n = 100, except Azithromycin where n = 93. For CoNS and M. scuri n = 44. Results interpreted according to EUCAST v11 guidelines and documents. The MIC50 and MIC90 were calculated for S. aureus.

Antimicrobial resistance
In S. aureus, resistance rates were 53% (n=49) to azithromycin, 32% (n=32 to methicillin (oxacillin), 29% (n=29) to tobramycin and 18% (n=18) for gentamicin, (n=100 isolates, except azithromycin n=93; Table 1).

Vancomycin and linezolid were the most active agents (0% resistance). The MIC50 of amikacin, tigecycline, minocycline, rifampicin, vancomycin and linezolid were all below the respective susceptible breakpoints (Table 1). Thirty-three were methicillin resistant S. aureus (MRSA), however the meca gene was detected in 43% (n=43). We also identified the reverse discrepancy, with three isolates negative for mecA (and mecC), but were classified MRSA by phenotypic methods. Typing of the staphylococcal cassette chromosome mec (SCCmec), revealed SCC-mec_type_1Va(2B) was the most frequent (n=30/43). All MRSA from Bangladesh carried SCCmec_type_1Va(2B) whereas all MRSA from South Africa carried SCCmec_type_1Va(1B).

Although a 20% reduction in patient survival at 60 days was seen for neonates with MRSA (chi-squared test, P 0.061), which may be due to the small sample size (n=33). We found no significant difference between the onset of sepsis and MRSA (χ², P 0.935 for phenotypic method and P 0.135 when mecA was used to denote MRSA).

Although the MIC values indicate almost all S. aureus isolates were sensitive to both minocycline (Table 1; 93% [n=93] sensitive) & tigecycline (100% [n=100] sensitive), 25 isolates carried at least one tetracycline ARG (Fig. 2; n=18 tet(K), n=7 tet(M) and n=3 tet(L). Additionally, aminoglycoside resistance genes were found in all S. aureus, and up to seven genes, including aac(3), aadD, sat4a and str, were found in a cluster of isolates.
Fig. 4  Neonatal survival curve data for Staphylococcus aureus and pathogenicity markers: MRSA and PVL. Kaplan Meier survival plot comparing a MRSA v MSSA and time to neonatal outcome, and b The presence of virulence factor PVL and time to neonatal outcome censored at the last available follow up appointment.
from South Africa (Fig. 1). The insertion element IS256 is associated with the aminoglycoside resistance-mediating transposon Tn4001, and IS256 was found in 13/100 isolates. Macrolide ARG (ermA, ermC, InuA, mphC and msrA) were predominantly found in South Asian isolates, including ST1, ST6, ST8, ST22 and ST121 S. aureus. We only detected macrolide ARG in ST5 isolates from Africa, where msrA genes were detected in isolates from Nigeria, and ermA genes were detected in ST5 from South Africa (Fig. 1a).

For CoNS, isolates’ resistance rates were 50% (n = 22) to tobramycin and 41% (n = 18) to gentamicin (Table 2). Fourteen isolates carried b1aZ, 10 of which also carried at least one aminoglycoside ARG (aadD and aph(3) genes; Fig. 2). We also detected IS256 in seven isolates (six S. haemolyticus and one S. epidermidis). Resistance to methicillin in CoNS (MR-CoNS) was seen in n = 11/18 isolates of S. epidermidis, n = 9/11 of S. haemolyticus and n = 2/13 of M. sciuri (Table 1, Fig. 2). mecA was detected in an additional four isolates, however all were sensitive to oxacillin. All M. sciuri isolates had the mecA1 gene, however this did not confer resistance to β-lactam antibiotics, as reported previously 46. We detected two copies of mecA (mecA1_2 and mecA9) in one isolate from Bangladesh, ultimately rendering this isolate resistant to oxacillin with an MIC > 8 mg/l (Fig. 2, Table 1). Furthermore, a single M. sciuri isolate from Pakistan was resistant to methicillin, tigecycline and vancomycin. Aside from these two isolates, the remaining 13 M. sciuri tested were sensitive to all other antibiotics. We also found a single vgaALC and catPC gene in two S. haemolyticus isolates, both of which also carried mecA, and aminoglycoside resistance genes (aadD and aph(3)-III).

**Virulence factors**

From 100 S. aureus isolates, 5494 genes putatively related to virulence (>98% coverage and identity) were detected in total. All VF hits were summarised according to the gene presence/absence against sepsis onset (EOS vs LOS) and outcome (alive v deceased) extending on analysis reported in [46]. For example, for the gene geh, we saw 32/40 (80%) presence in EOS as opposed to 24/46 (52.2%) in LOS. For the six genes that were significant before FDR correction (geh, cap8H, sak, cap8I, cap8K and scn), all had 13% or greater presence for EOS than for LOS (Additional file 1: Table S7). Similarly, lukD was present in 16/18 (88.9%) of S. aureus isolates recovered from deceased neonates, but only in 50/82 (61%) of neonates recorded as alive at the latest follow up appointment/60 days. For the eight genes that were significant before FDR correction, (lukD, sbi, aur, cap8B, scn, esaE, essC, hlgB), all bar one had at least 19% greater gene presence in the deceased than the survivors (Additional file 1: Table S7).

| Function | VF | Onset EOS | LOS | ND | Total | Outcome Alive | Deceased | ND | Total |
|----------|----|-----------|-----|----|-------|---------------|-----------|----|-------|
| Inhibiting host complement systems | chp | 16 | 15 | 5 | 36 | 19 | 8 | 9 | 36 |
| | spa | 18 | 22 | 8 | 48 | 27 | 10 | 11 | 48 |
| Death of B cells | coa | 3 | 3 | 2 | 8 | 8 | 0 | 0 | 8 |
| | fnbB | 2 | 2 | 1 | 4 | 4 | 0 | 0 | 4 |
| Formation of aggregates in blood | clfA | 10 | 5 | 2 | 17 | 12 | 3 | 1 | 16 |
| | fnbA | 5 | 8 | 0 | 13 | 12 | 0 | 1 | 13 |

**Table 2** In silico *Staphylococcus aureus* virulence factors analysed for both onset and outcome of sepsis.
We also grouped VF related to bloodstream invasion/infection (n = 656) according to their biological functions as described by Powers et al., 2014 43 (Table 2, Additional file 1: Fig. S13). *hlgA/B/C* and *lukD* genes were found in 100% (n = 100) and 66% (n = 66) of *S. aureus* respectively, and PVL was found in 30% (n = 30). In neonates reported as deceased (n = 17), we found that the presence of PVL was found to significantly reduce the time to death (comparing means test, P 0.028), with the mean time to mortality at eight days compared to 31 days for PVL negative *S. aureus*. Similarly, the Kaplan Meier survival curve suggests the presence of PVL may result in a rapid MfBS, and this was more pronounced in the first 30 days of life, however, was not statistically significant (Log Rank test, P 0.251, Fig. 4b). In addition to virulence associated genes targeting leukocytes, we detected genes that contribute to inhibiting host complement systems (*chp*, and *scn*) and genes that contribute to *S. aureus* aggregation.
(Table 2, coa n = 8 isolates, cflB n = 4, clfA n = 16 and νWbp n = 17).

We found eight virulence associated genes from seven CoNS isolates. Two hits of the antiphagocytosis genes cap8G and cap8E within the same contig were detected in a S. haemolyticus isolate, whereas the remaining genes were detected within S. epidermidis. VF from the biofilm related ica operon (icaA) were detected in three S. epidermidis isolates, of which one contained IS256 and was MRSA (from Nigeria National Hospital, Abuja (NN)), whereas essA, a gene associated with bacterial survival (type VII secretion system), were found in three S. epidermidis isolates (NK n = 2, NN n = 1). VFs present in VFDB were not detected within M. sciuri isolates.

**M. sciuri cluster**

Twenty-two M. sciuri were collected from Islamabad, Pakistan, between July-November 2017 (Fig. 5a). The blood culture date was not available for eight and we inferred approximate dates based on the consecutive enrolment numbers for the study (Fig. 5a). Where the onset of sepsis was known (n = 13), all were EOS suggesting the same strain may have spread within the hospital environment, perhaps related to the delivery area/delivery practices as, commonly reported, hospital contamination is more often linked to LOS [5, 47]. WGS data was available for 11 isolates which were all within four pairwise SNPs and were assigned to ST82, suggesting these represented the same/phylogenetically very close strains. The hybrid assembly of short and long read sequences produced a single contig for the bacterial chromosome at 2,932,999 bp (Additional file 1: Fig. S14). From genome estimation screening using mash (v2.2), the 10 genetically closest genomes to the Pakistani M. sciuri were > 1,400 SNPs away. The isolate that shares the same branch point in the collective phylogeny (Fig. 5b), was an M. sciuri blood culture isolate collected from Bangladesh during this study and was assigned to ST81. Similarly, the single M. sciuri isolate available for WGS characterisation from Nigeria was assigned to ST80. The MLST scheme for M. sciuri was first made available within pubmlst.org on the 1st of February 2021. Of the M. sciuri genomes that had the greatest Jaccard similarity index and were included into the phylogenetic analysis, three were clinical samples from Europe (Denmark and Czech Republic) and these were distinct throughout the tree suggesting different lineages causing human infection.

**Discussion**

When a neonate is presenting with clinical signs of sepsis, empirical antibiotic treatment will be prescribed. Often, in LMICs hospitals there are not sufficient resources to perform necessary laboratory diagnostics, or a confirmatory blood culture. There is currently little to no data extending beyond an initial species identification (often via Gram staining and biochemical tests) and an antibiogram when Staphylococci/GPB are reported from neonatal blood cultures. We performed this retrospective study to gain an insight into the diversity of GPB isolated from blood cultures, and we highlight the need for future genomic and epidemiological studies to uncover the true burden of GPB neonatal sepsis in LMICs. A recent review by Garcia et al., 2022 describe a “Paucity of [care] bundles and elements aimed at HCAI detection (4% of 295 elements)” and “barriers to strengthening laboratory capacity” [48] when addressing matters of infection control in neonatal care in LMICs, emphasising the need for surveillance studies also mirrored elsewhere [49].

*S. aureus* is often the most frequently identified Gram-positive cause of neonatal sepsis in LMICs [2, 7, 50]. One of the most common STs found in this study was ST152, and although ST152 is not associated to a particular CC, scanning the list on the PubMLST database isolate collection [24], we noticed that of the 40 isolates with this ST, 12 were collected from blood cultures and all were from Africa (n = 9 The Gambia and n = 3 Mozambique). Herein, we found ST152 *S. aureus* neonatal sepsis in Ethiopia, Nigeria, and South Africa. Additionally, we detected a cluster of ST5 (CC5) in South Africa and previous published work from the same hospital suggests that ST5 is a successful and problematic lineage [51]. Many countries have reported CC5 MRSA from routine hospital molecular surveillance causing nosocomial infections further suggesting its global importance [51–56]. Frequently, however, nosocomial epidemiological clusters or outbreaks go undetected perhaps due to limited resources and/or protocols in place to detect outbreaks, outlining the importance of genomic surveillance studies to tackle nosocomial infection and provide insight into transmission dynamics of bacteria within neonatal units. Similar to all ST152 *S. aureus* being found only in Africa, during this study, all ST8 isolates were from neonates in South Asia. The numerous clades within the *S. aureus* ST8 group (as inferred from SNP analysis of the ST8 reads) suggest multiple introductions of ST8 causing neonatal sepsis in Pakistan. CC8 is well established and has commonly been reported from clinical infections worldwide, including South Asia [57, 58]. Importantly, we found that infection with ST8 and ST5 *S. aureus* had lower survival in the Kaplan Meir curve compared to other ST groups suggesting that certain lineages of *S. aureus* may be more pathogenic and problematic in neonatal sepsis in LMICs. Previous work on *S. aureus* bacteraemia also suggests that clonal differences may be a contributing indicator of infection outcome [59], and through preliminary data
analysis we indicate that certain STs, ARGs, MGEs and VFs may be used as indicators of neonatal outcome.

Although *S. aureus* neonatal sepsis isolates in this study remained susceptible to vancomycin, similar to many previous reports [8, 60], we report an MRSA rate of 33% slightly higher than reported in a single site study in Uganda [61], and similar to a study in Nigeria [62], emphasising MRSA as an important source of infection and further highlighting the burden of AMR in neonatal sepsis in LMICs. Importantly, CoNS isolates often carry methicillin resistant genes [9, 13, 15], and through WGS and MIC determination, we identified methicillin resistance in one-third of isolates analysed, highest in *S. epidermidis* (n=6/8) and *S. haemolyticus* (n=7/8).

Meric et al. [10] showed the importance of horizontal gene transfer events to increase the likelihood of divergent lineages causing infection due to the acquisition of genetic traits (virulence genes and ARG, e.g. SCCmecA) potentially conferring niche adaptations associated with pathogenicity. Kozitskaya et al. [63] also report a significant association between the presence of IS256 in clinical *S. epidermidis*, perhaps suggesting that isolates with increased antimicrobial resistance are successfully colonising hospital environments. Although our WGS dataset for CoNS is too small for such comparisons, and found relatively few IS256 in *S. epidermidis*, we have shown concordance with previous studies and published genomics data, indicating that further surveillance studies are needed to characterise GPB isolates recovered from neonatal blood cultures.

The biggest question raised upon the isolation of CoNS from blood cultures is whether this confirms biological sepsis or represents contamination from transient colonisation of the skin. The ANISA study classified all CoNS as “definite contaminants” [64] and CoNS were likewise excluded from Okomo et al.’s. [2] meta-analysis, many neonatal studies report CoNS as a leading cause of neonatal sepsis [65–67]. Our results must be interpreted with this in mind. A second blood culture is common practice following the identification of CoNS, however this remains challenging in neonates, as often a small volume of blood was collected (volume not recorded), and it was not possible for sites to consistently perform duplicate blood cultures following the local species identification of a CoNS. Speciation of CoNS can be difficult in laboratories where biochemical and conventional microbiology techniques are routinely employed [9], and long-term improvements in access to microbiology diagnostics in LMICs will be vital to understand the true burden of *Staphylococci* neonatal sepsis in these sites. Without this capacity, the cluster of *M. sciuri* described here could not be recognised and managed, as one clear example of the value of diagnostic capacity. *M. sciuri* have a broad range of habitats including animals, humans and the environment [45]. During a recent phylogenomic comparison of *Staphylococci*, several CoNS species were reclassified including *Staphylococcus sciuri* [68]. Although infection caused by *M. sciuri* has largely been reported in animals, there are case reports of *M. sciuri* causing peritonitis [69] and two distinct cases of catheter associated sepsis in Turkey [70] and Japan [71]. The clinical and microbiology team in Pakistan were routinely reporting and storing all GPB isolates from blood cultures, in contrast to other sites in South Asia which is likely to account for the steep difference in GPB neonatal sepsis, and the majority (n=21, 81%) of *M. sciuri* isolates from Pakistan (Fig. 5).

We did however identify five *M. sciuri* isolates from other clinical sites in Bangladesh, Ethiopia, and Nigeria. From these, two isolates (from Bangladesh and Nigeria) were shown to be sitting on separate phylogenetic branches through genomics analysis, (Fig. 5b). Furthermore, the identification of *M. sciuri* from 5/12 clinical sites highlights the potential for uncommon pathogens to be present in clusters or cause outbreaks.

There are limitations to this study. BARNARDS was established to determine the aetiology of GNB sepsis, therefore there was heterogeneity in the reporting and storage of GPB isolates among the clinical sites. Although many clinical sites that participated in BARNARDS identified GPB species from blood cultures as part of their routine standard operating procedures and guided by local practices, not all sites had the capacity and resources to identify bacteria to the species level. The retrospective recovery of Gram-positive isolates at local sites was greatest in Nigeria and Pakistan, therefore caution must be taken in any interpretation of results. There was discrepancy between the degree of reporting GPB neonatal sepsis within the clinical sites therefore we make no reference to prevalence or specific country analyses within this dataset. All statistical analyses between clinical data and microbiology findings are exploratory and hypothesis forming. Group B *Streptococci* (GBS), considered the dominant cause of sepsis in high-income countries [47, 72], was not found/reported, and this has also been observed in other studies [1, 5, 6, 8, 60, 73]. With limited microbiology resources in many LMICs however (GBS require additional media and growth requirements), septic neonates may be deceased prior to a blood culture and Vera et al., suggest GBS are perhaps underreported due to these infections presenting very early into life [47]. Additionally, due to the nature of our retrospective recovery, the loss of fastidious organisms at the local sites and/or transportation to the UK might have hindered the recovery of GBS and further studies are needed to determine the incidence of *Staphylococci* and GBS neonatal sepsis.
Conclusion

*S. aureus* is a leading cause of neonatal sepsis in LMICs, and we found one third of *S. aureus* to be MRSA. Future studies should assess the use of genomic characterisation to predict whether certain lineages of Staphylococci, and GPB species have greater pathogenic potential in neonatal sepsis. Unveiling a cluster of *M. sciuri* in Pakistan further demonstrates the importance of GPB in neonatal sepsis. There is an urgent need for microbiological and genomic surveillance studies with strict infection control practices to categorically evaluate the true burden of GPB in neonatal sepsis in LMICs, laying emphasis on whether associations exist between certain species, lineages and virulence factors and the clinical onset or outcome of sepsis.

Abbreviations

AMR: Antimicrobial resistance; ARG: Antibiotic resistance genes; ATCC: American Type Culture Collection; BARNARDS: Burden of Antibiotic Resistance in Neonates from Developing Societies; CC: Clonal complex; CoNS: Coagulase negative Staphylococci; EOS: Early onset sepsis; EUCAST: European Committee on Antimicrobial Susceptibility Testing; GNB: Gram negative bacteria; GPB: Gram positive bacteria; IS: Insertion sequence; LMIC: Low- and- middle income country; LOS: Late onset sepsis; MRSA: Methicillin resistant *Staphylococcus aureus*; MfBs: Mortality following biological sepsis; MGE: Mobile genetic element; MIC: Minimum inhibitory concentration; MLST: Multi locus sequence typing; Mm: M. morganii; M. sciuri: *M. sciuri*; M. tuberculosis; M. truncatula: *M. truncatula*; M. xenopi: *M. xenopi*; MRIA: Multi resistant *Staphylococcus aureus*; MRSA: Methicillin resistant *Staphylococcus aureus*; PVL: Panton-Valentine Leukocidin; SNP: Single nucleotide polymorphism; ST: Sequence type; VF: Virulence factors; VFDB: Virulence factor database; VLBW: Very low birth weight; WGS: Whole genome sequencing.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12879-022-07541-w.

Additional file 1. Additional tables and figures.

Acknowledgements

We would firstly like to acknowledge all the mothers’ and neonates that enrolled into the BARNARDS study, and all hospital and research staff at each of the participating clinical sites. We thank and acknowledge the hard work of the team at Liofilchem®, Roseto, Italy, specifically F. Brocco and M. Collett. We would like to acknowledge Specialist Antimicrobial Chemotherapy Unit (SACU) at Public Health Wales for support with bacterial isolate identification using MALDI-TOF MS. We thank Maksuda Islam for providing access to the clinical sites and epidemiology data in Bangladesh. We would like to acknowledge Rubina Kamran the Microbiologist from PIMS who sadly passed away in 2018. We would like to thank Julian Parkhill for advice on SNP and phylogenetic analysis. We would like to thank the team at Bill and Melinda Gates Foundation namely Padmini Srikanth, Rasa Izadnezhadgar, Keith Grammer and Sara Vernam. We would also like to acknowledge Wales Gene Park and ARCCA for their continued bioinformatics support. Bioinformatics analysis was undertaken using the supercomputing facilities at Cardiff University operated by Advanced Research Computing at Cardiff (ARCCA) on behalf of the Cardiff Supercomputing Facility and the HPC Wales and Supercomputing Wales (SCW) projects. We acknowledge the support of the latter, which is part-funded by the European Regional Development Fund (ERDF) via the Welsh Government. The Centre for Trials Research is funded by Health and Care Research Wales and Cancer Research UK. We thank the team of curators for the databases hosted on PubMedMLST https://pubmlst.org/databases/.

Author contributions

KS guided the analysis, wrote the manuscript and produced the Figures. KS performed the bioinformatics analysis supported by RA, MJC, JP and BS. KS, EP, and JM performed whole genome sequencing experiments. KT performed the minimum inhibitory concentration (MIC) experiments. KT and MJC analysed the MIC dataset. WJ and MJC performed the statistical analysis. KS, MJC, EP, KT, CD, CA, AF, TH, JM, and MN performed microbiology culture and sample processing in Cardiff University. CD and RM provided the managed epidemiological dataset. GC, DB, SS, SB, RNK, SN, KI, FM, SU, RZ, HS, AM, JBM, AR, LG, SM, ANHB and AW collected and processed blood cultures and collected clinical data at the clinical sites. BS and JM guided the analysis and writing of the manuscript. TRW, KS, and MJC. RM, GC, SB, KL, RZ, JBM and SM designed the BARNARDS study. All authors reviewed and commented on the manuscript.

Funding

BARNARDS was funded by two awards S4.28M (OPP1119772) and S849K (OPP191522) from the Bill and Melinda Gates Foundation at 500 Fifth Avenue North-Seattle, WA 98109. The funders had no role in the design of the study, data collection, analysis, interpretation of data or manuscript preparation.

Availability of data and materials

Sequences have been deposited in ENS with the project number: PRJEB40908, with individual accessions reported in the supplementary dataset (Additional file 1: Table S3). The *M. sciuri* genome used as a reference for SNP calling has been deposited in Genbank with the accession number CP071138.

Declarations

Ethics approval and consent to participate

Ethical approval was obtained at each of the seven participating countries (Additional file 1: Table S1). Bangladesh: Ethical Review Committee, Bangladesh Institute of Child Health (BICH-ERC-4/3/2015), Ethiopia: Boston Children’s Hospital (188–P00025058), India: Institutional Ethics Committee, National Institute of Cholera and Enteric Diseases and Institute of Post Graduate Medical Education and Research, JIPMER-BRA Research Oversight Committee (A/1-2016-IEC and Inst/IEC/2016/508), Nigeria: Kano State Hospitals Management Board (8/10/1437AH), Health Research Ethics Committee (HREC) and National Hospital, Abuja (NHA/EC/017/2015), Pakistan: Shaheed Zulfiqar Ali Bhutto Medical University, Pakistan Institute of Medical Sciences (PIMS) Islamabad (Ref No NA, signed letter from Prof Tabish Hazir). Rwanda: Republic of Rwanda, National Ethics Committee (No342/RNEC/2015), South Africa: Stellenbosch University and Tygerberg Hospital, Research projects, Western Cape Government (N15/07/063). In local languages, research nurses provided mothers with study information and collected consent for mother and/or neonatal enrolment. Informed consent was obtained in writing unless this was not possible (due to literacy barriers), and oral consent was collected from the mothers by trained researchers. Verbal consent was approved for use as informed consent if written consent was not possible by each committee and this was documented by the participant signing/marking the consent form.

Consent for publication

Not applicable.

Competing interests

The authors have no competing interests to declare.

Author details

1Department of Medical Microbiology, Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, UK. 2Department of Zoology, Ineos Oxford Institute of Antimicrobial Research, Oxford, UK. 3Institute of Biomedical Sciences, Department of Medical Sciences, University of Aveiro, Aveiro, Portugal. 4Centre for Trials Research, Cardiff University, Cardiff, UK. 5National Hospital Abuja, Abuja, Nigeria. 6Department of Microbiology, Quaid-I-Azam University, Islamabad, Pakistan. 7Pakistan Institute of Medical Sciences, Islamabad, Pakistan. 8Murtala Muhammad Specialist Hospital, Kano, Nigeria. 9Federal Medical Centre Jabi, Abuja, Nigeria. 10Division of Medical Critical Care,
References

1. Hamer DH, Darmstadt GL, Carlin JB, Zaidi AKM, Bs MB. Etiology of bacteremia in young infants in six countries. Pediatr Infect Dis J. 2015;34(1):1–8.

2. Okomu U, Akpulu EN, Le do Keco, Roca A, Coussens S, Jarde A, et al. Articles Aetiology of invasive bacterial infection and antimicrobial resistance in neonates in sub-Saharan Africa: a systematic review and meta-analysis in line with the STROBE-NI reporting guidelines. Lancet Infect Dis. 2019;3099(19):1–16. https://doi.org/10.1016/S1478-7274(19)30414-1.

3. Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10108):1151–210.

4. Fleischmann C, Reichert F, Casiini A, Ferber J, Harder F, Markwart R, et al. Global incidence and mortality of neonatal sepsis: a systematic review and meta-analysis. Arch Dis Child. 2021;106:1–8.

5. Popescu CR, Cavanagh MMM, Tembo B, Lufesi N, Goldfarb DM, Kissoon N, et al. Review of clinically and epidemiologically relevant coagulase-negative staphylococci in Africa. Microb Drug Resist. 2020;26(8):951–70.

6. Drmagowski A, Madide A, Bekker A. Neonatal nosocomial bloodstream infections at a referral hospital in a middle-income country: burden, pathogens, antimicrobial resistance and mortality. Paediatr Int Child Health. 2015;35(3):265–72.

7. Medugu N, Iregbu K, Tam Pyong I, Obaro S. Aetiology of neonatal sepsis in Nigeria, and relevance of Group b streptococci in neonates. J Med Microbiol. 2016;65(14):2068–9.

8. Asante J, Armaoog DG, Abia AM, Ogbunike I, Bester LA, et al. Review of clinically and epidemiologically relevant coagulase-negative staphylococci in Africa. J Infect Chemother. 2020;26(8):951–70.

9. Méric G, Mageiros L, Pensar J, Klingenberg C, Flægstad T, et al. Neutonal sepsis in low-income countries: epidemiology, diagnosis and prevention. Expert Rev Anti Infect Ther. 2020. https://doi.org/10.1080/14787210.2020.1732818.

10. Méric G, Mageiros L, Pensar J, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. J Antimicrob Chemother. 2021;76(1):101–9.

11. EUCAST. EUCAST V11: The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 11.0, 2021.

12. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2. Advanced and refined database for big data analysis—10 years on. Nucleic Acids Res. 2016;44(D1):D694–7.

13. Popescu CR, Cavanagh MMM, Tembo B, Lufesi N, Goldfarb DM, Kissoon N, et al. BMC Infectious Diseases (2022) 22:593

14. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2. Advanced and refined database for big data analysis—10 years on. Nucleic Acids Res. 2016;44(D1):D694–7.

15. McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 2004;32:20–5.

16. Seemann T. mlst, Github [Internet]. 2019. https://github.com/tseemann/mlst.

17. Jolley KA, Bray JE, Maiden MC. Open-access bacterial population genomics: BigSBd software, the PubMLST.org website and their applications [version 1; referees: 2 approved]. Wellcome Open Res. 2018;3:1–20.

18. Thomsen MCT, Ahrenfeldt J, Cisneros JLB, Jutze V, Lar森 MV, Hasmann H, et al. A bacterial analysis platform: an integrated system for analysing bacterial whole genome sequencing data for clinical diagnostics and surveillance. PLoS ONE. 2016;11(16):1–14.

19. Bartels MD, Petersen A, Wörm M, Nielsen JB, Larners Svensson H, Johansons HK, et al. Whole-genome sequencing with sanger sequencing for spa typing of methicillin-resistant Staphylococcus aureus. J Clin Microbiol. 2014;52(21):4305–8.

20. Johansons M, Torgersen T, Tansirirchaiya S, Aarestrup FM, Roberts AP, Petersen TN. Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. J Antimicrob Chemother. 2021;76(1):101–9.

21. Seemann T. snippy: fast bacterial variant calling from NGS reads. [snippy] Bioinformatics. 2016;32(2):292–4.

22. Mcginnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 2004;32:20–5.

23. Seemann T. mlst, Github [Internet]. 2019. https://github.com/tseemann/mlst.

24. Aanensen DM, Feil EJ, Holden MTG, Dordel J, Yeats CA, Fedosejev A. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequencing data using Gubbins. Nucleic Acids Res. 2015;43(3):e15.

25. Aanensen DM, Feil EJ, Holden MTG, Dordel J, Yeats CA, Fedosejev A. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive Staphylococcus aureus in Europe. MBio. 2016;7(3):1–15.

26. Seemann T, Pointka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.

27. Page AJ, Cummings CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Sequence analysis Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(July):3691–3.

28. Price MN, Dehal PS, Arkin AP FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010. https://doi.org/10.1371/journal.pone.0009490.

29. Seemann T. snippy: fast bacterial variant calling from NGS reads. [snippy] Bioinformatics. 2015;31(14):2068–9.

30. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15.

31. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb Genom. 2016;2(4):e00056.

32. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NL: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35(21):4453–5.

33. Minh BQ, Schmidt HA, Forster J, Chenoune O, Schertel M, Dewhams MD, Von Haeseler A. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530–4.
38. Seemann T. snp-dist [Internet]. https://github.com/tseemann/snp-dist. Accessed Dec 2018.

39. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5.

40. Wick RR. Porechop v0.2.4 [Internet]. https://github.com/rnwick/Porechop. Accessed Nov 2018.

41. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(5):1–22.

42. Ondov BD, Treangen TJ, Melsted P, Mannon EE, Bergman NH, Koren S, et al. MASH: fast genome and metagenome distance estimation using Mink squash. Genome Biol. 2016. https://doi.org/10.1186/s13059-016-0997-x.

43. Powers ME, Wardenburg JB. Igniting the fire: Staphylococcus aureus virulence factors in the pathogenesis of sepsis. PLoS Pathog. 2014;10(2):10–3.

44. Url S, Society RS. Controlling the false discovery rate: a practical and powerful approach to multiple testing Yoav Benjamini; Yosef Hochberg Journal of the Royal Statistical Society. Series B (Methodological), Vol. 57, No. 1. (1995). JR Statist Soc B 1995;57(1):289–300.

45. Nemeghaire S, Argudín MA, Feßler AT, Hauschild T, Schwarz S, Butaye S, et al. Ori-Nection: Infection prevention and care bundles addressing health-care-associated infections in neonatal care in low-middle income countries: a scoping review. Ecological Medicine. 2022;44:101259.

46. Thomson K, Dyer C, Liu F, Sands K, Carvalho MJ, Barrell M, et al. Evaluating the roles of antibiotic resistance, drug target attainment, bacterial pathogenicity, virulence and antibiotic access and affordability in affecting outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective study. Lancet Infect Dis. 2021. https://doi.org/10.1016/S1473-3099(21)00050-5.

47. Zea-Vera A, Ochoa TJ. Challenges in the diagnosis and management of neonatal sepsis. J Trop Pediatr. 2015;61(6):11–13.

48. Molina García A, Cross JH, Fitchett EJA, Kawaza K, Okomo U, Spotswood J, et al. Challenges in the diagnosis and management of neonatal sepsis in Peshawar, Pakistan. Arch Iran Med. 2016;19(12):669–9.

49. Jansen van Rensburg MJ, Eliya Madikane V, Whitelaw A, Chachage M, Figueiredo AMS, et al. Emergence of clonal complex 5 (CC5) methicillin-resistant Staphylococcus aureus (MRSA) strains and ACME-positive methicillin-resistant Staphylococcus aureus (MRSA) strains and ACME-positive CC5 (ST8/SCCmec-IVa mecI-1) strains. J Clin Microbiol. 2013;52(1):649–56.

50. Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K, Ziebuhr W. The bacterial insertion sequence element is256 occurs preferentially in nosocomial Staphylococcus epidermidis isolates: association with biofilm formation and resistance to aminoglycosides. Infect Immun. 2004;72(2):1210–5.

51. Hossain B, Weber MW, Hamer DH, Hibbrell PL, Ahmed A, Marzan M, et al. Classification of blood culture isolates into contaminants and pathogens on the basis of clinical and laboratory data. Pediatr Infect Dis J. 2016;35(5):552–4.

52. Kotala S, Ghiya D, Kale L, Desai R, Logue A, Marques S, et al. Genetic diversity of emerging Panton-Valentine leukocidin/virulence-positive Staphylococcus aureus strains in neonates in a middle-income country. Pediatr Res. 2016;44(W1):W242–5.

53. Recker M, Laabei M, Toleman MS, Reuter S, Saunderson RB, Blane B, et al. Genetic diversity of emerging Panton-Valentine leukocidin/virulence-positive Staphylococcus aureus strains in neonates in sub-Saharan Africa: a systematic review and meta-analysis in line with the STROBE-NI reporting guidelines. Lancet Infect Dis. 2019;20(1):19–31.

54. Worah M, Achari S, Teo SK, Watanabe N, et al. Characterization of PVL/ACME-positive methicillin-resistant Staphylococcus aureus (MRSA) strains and ACME-positive CC5 (ST8/SCCmec-IVa mecI-1) strains. J Clin Microbiol. 2013;52(1):649–56.

55. Gorwitz RJ. Community-associated meticillin-resistant Staphylococcus aureus: epidemiology and update. Pediatric Infect Dis J. 2008;27(10):925–6.

56. Strauß L, Stegger M, Akpaka PE, Alabi A, Breurec S, Coombs G, et al. Origin, evolution, and global transmission of community-acquired Staphylococcus aureus ST31. Proc Natl Acad Sci U S A. 2017;114(49):E10596–604.

57. Recker M, Laabei M, Toleman MS, Reuter S, Saunderson RB, Blane B, et al. Clonal differences in Staphylococcus aureus bacteriaemia-associated mortality. Nat Microbiol. 2017;2(10):1381–8.