Characterisation of neonatal *Staphylococcus capitis* NRCS-A isolates compared with non NRCS-A *Staphylococcus capitis* from neonates and adults.

Heather Felgate[1,2], Dheeraj Sethi[1,3], Kirstin Faust[4], Cemsid Kiy[4], Christoph Härtel[5], Jan Rupp[6], Rebecca Clifford[2], Rachel Dean[1,3], Catherine Tremlett[3], John Wain[1,2], Gemma Langridge[1,2], Paul Clarke[2,3], Andrew Page[1], Mark A Webber[1,2*].

[1] Quadram Institute Bioscience, Norwich Research Park, NR4 7UQ, Norwich, UK

[2] Norwich Medical School, University of East Anglia (UEA), Norwich, UK

[3] Norfolk and Norwich University Hospital (NNUH), NR4 7UY, Norwich, UK

[4] Department of Pediatrics, University of Lübeck, Lübeck, Germany

[5] Department of Pediatrics, University of Würzburg, Würzburg, Germany

[6] Department of Infectious Diseases and Microbiology, University of Lübeck, Lübeck, Germany

[*]mark.webber@quadram.ac.uk

01603255233

Keywords: microbiota, skin, infection control, CRISPR, Antimicrobial Resistance, neonates, Late Onset Infection

Abstract

Staphylococcus capitis is a frequent cause of Late-Onset Sepsis (LOS) in neonates admitted to Neonatal Intensive Care Units (NICU). The NRCS-A clone of *S. capitis* has been isolated from NICUs globally although the reasons for the global success of this clone are not understood.

We analysed a collection of *S. capitis* colonising babies admitted to two NICUs, one in the UK and one in Germany as well as corresponding pathological clinical isolates. Genome analysis identified 3 groups; non-NRCS-A isolates, NRCS-A isolates, and a group of 'proto NRCS-A' - isolates closely related to NRCS-A but not associated with neonatal infection. All bloodstream isolates belonged to the NRCS-A group and were indistinguishable from strains carried on the skin or in the gut. NRCS-A isolates showed increased tolerance to chlorhexidine and antibiotics relative to the other *S. capitis* as well as enhanced ability to grow at higher pH values. Analysis of 138 pangenomes of the clades identified characteristic *nsr* and *tarJ* genes in the NRCS-A and proto groups with a CRISPR-cas system only seen in NRCS-A isolates which also showed enrichment of genes for metal acquisition and transport.

We found evidence for transmission of *S. capitis* NRCS-A within NICU, with related isolates shared between babies and multiple acquisitions by some babies. Our data show NRCS-A strains commonly colonise uninfected babies in NICU representing a potential reservoir for potential infection. This work provides more evidence that adaptation to survive in the gut and on skin facilitates spread of NRCS-A, and that metal acquisition and tolerance may be important to the biology of NRCS-A. Understanding how NRCS-A survives in NICUs can help develop infection control procedures against this clone.

41 **Introduction**

42 Non-*aureus* Staphylococci are common commensal bacteria, however many have also been implicated in
43 nosocomial infections and are also found to cause Late Onset Sepsis (LOS). LOS is defined as sepsis occurring
44 ≥ 72 hrs after birth which can be caused by various bacterial species, the most common being the Non-
45 *aureus* Staphylococci (NAS) (Bentlin et al., 2010, Makhoul et al., 2005, Dong and Speer, 2014, Harvey et al.,
46 2022). LOS increases the length of hospital stay, invasive procedures, and provokes more and longer
47 antibiotic treatments all of which can have negative impacts on the long-term outcomes of new-born babies
48 (Bentlin et al., 2010, Makhoul et al., 2005, Dong and Speer, 2014, Rasigade et al., 2012).

49 *Staphylococcus capitis* is a common commensal of humans and is regularly cultured from skin or nasal swabs
50 (Becker et al., 2020), but a particular clone of *S. capitis* NRCS-A has been found to associated with LOS.
51 Originally identified in France (the archetypal strain is known as CR01) it has subsequently been found in
52 neonatal intensive care units (NICUs) in 17 countries (Lemriss et al., 2014, Lepainteur et al., 2013, Wirth et
53 al., 2020, Butin et al., 2017, Harvey et al., 2022). Infection with NRCS-A has been associated with a higher
54 morbidity and mortality than other NAS (Butin et al., 2019, Ben Said et al., 2016). The transmission routes of
55 NRCS-A are not well understood but it has been documented to colonise incubators and equipment and has
56 reduced susceptibility to antiseptics which has been proposed to aid its environmental survival. Unlike other
57 NAS, infection with NRCS-A is not associated with maternal-fetal transmission, and babies born by
58 Caesarean-section have been shown to be colonised more often than those delivered vaginally (Butin et al.,
59 2019).

60 NRCS-A isolates typically have a multidrug resistance profile including methicillin, aminoglycosides and
61 fosfomycin, hetero-resistance to vancomycin (a common treatment for LOS) and decreased susceptibility to
62 chlorhexidine (Wirth et al., 2020, Lepainteur et al., 2013, Rasigade et al., 2012, Harvey et al., 2022). Many
63 NRCS-A isolates have been found to contain a S*CCmec* mobile genetic element (*SCCmec-SCCcad-SCCars-*
64 *SCCcop*) responsible for the methicillin resistance (Carter et al., 2018, Du et al., 2021, Becker et al., 2014) and
65 decreased susceptibility to vancomycin has been associated with SNPs in *sarA* and *glnQ* (Wirth et al., 2020).
66 Analysis of genomes of a panel of NRCS-A isolates from various sites and sources suggested the clone
67 emerged in the 1960s before expanding in the 1980s with vancomycin use proposed as a major driver in the
68 selection and evolution of NRCS-A (Wirth et al., 2020).

69 *S. capitis* infection has only been studied sporadically outside of the neonatal setting but *S. capitis* has been
70 identified in bone and joint infections and, whilst rarely associated with adult disease, NRCS-A has been
71 identified in endocarditis, osteomyelitis and prosthetic joint infections (PJI) in adults (Tevell et al., 2020).

72 Here we report results from a longitudinal survey of NAS from skin and gut swabs taken from babies on
73 NICUs from two countries (the UK and Germany) where different antiseptic regimens were in place (Sethi et
74 al., 2021). Over 1000 NAS isolates were characterised for antimicrobial susceptibility and sequenced. From
75 this collection we have analysed the population structure of *S. capitis* and compared carriage isolates with
76 those from neonatal blood cultures as well as additional isolates from healthy adult skin swabs and PJI.

77 We found that NRCS-A strains are commonly carried by uninfected neonates representing a reservoir of
78 potential infection and identified probable transmission between babies on NICU. We also identified a clade
79 closely related to NRCS-A but without an association with disease. Comparison of the NRCS-A isolates with
80 this clade ('proto NRCS-A') identified genomic and phenotypic features which may contribute to the success
81 and prevalence of this epidemic clone.

82 **Methods**

83 **Collection of isolates of *S. capitis***

84 Infants admitted to the NICUs of the Norfolk and Norwich University Hospital (Norwich, UK) or University
85 Children's Hospital (Lubeck, German) over 10-week periods in 2017 or 2018 were enrolled in this study as
86 recently described (Sethi et al., 2021). Swabs are taken routinely from babies upon admission and
87 throughout their stays in both sites for MRSA surveillance; duplicate swabs were taken for this study and
88 staphylococci isolated. In addition, isolates from positive blood, cerebrospinal fluid, urine, and wound
89 cultures were included, if taken during the study period.

90 The UK unit enrolment was between November 2017 and January 2018, and the German enrolment was
91 from January to March 2018. Amies Charcoal Swabs (Thermo Fisher Scientific) were used to isolate bacteria
92 from all infants on admission and weekly until discharge. Swabs from the ear, nose, axilla, groin and gut
93 were taken and streaked on 5% horse blood agar (Thermo Fisher), prior to incubation at 37 °C for 24 hours
94 and final identification of coagulase-negative Staphylococci after sub-culture on mannitol-salt agar (Oxoid,
95 Thermo Fisher Scientific), coagulase testing (MERCK; 75832), and/or MALDI-TOF mass spectrometry (Bruker)
96 as previously described (Sethi et al., 2021). Isolates were stored on preservation beads at -80 °C (Protect,
97 Technical Service Consultants Ltd) and in 96 deep-well plates in 20% glycerol at -40 °C. Clinical isolates of *S.*
98 *capitis* from both units which were identified by local microbiology departments during the study periods as
99 part of usual practice were also included as well as further anonymised clinical isolates taken from routine
100 blood tests taken from neonates with suspected sepsis at the NNUH NICU collected in 2018 (n=7) and
101 between June and May 2022 (n=5). In addition to isolates from neonates, a further panel of 15 *S. capitis*
102 were taken from a pre-existing Staphylococci collection isolated from adult healthy skin swabs, and clinical
103 isolates taken from adult blood cultures (where infection was suspected) and 5 recovered from PJI (Felgate
104 et al., 2021).

105 **DNA extraction**

106 Isolates were grown in 1 mL Brain Heart Infusion (BHI, MERCK) broth overnight at 37 °C. Cultures were
107 pelleted and resuspended in 100 µl 0.5 mg/mL lysostaphin (from *Staphylococcus staphylolyticus,* MERCK)
108 and incubated at 37 °C for a minimum of 1 hour. For Illumina sequencing DNA was extracted from the lysate
109 with the Quick-DNA Fungal/Bacterial 96 kit (Cambridge Bioscience), in accordance with the manufacturer's
110 guidelines. DNA was quantified using the Quant-iT™ dsDNA HS assay (ThermoFisher), and fluorescence was
111 measured on a FLUOstar Optima plate reader at 480/530 nm (excitation/emission).

112 For high molecular weight DNA, the Quick-DNA HMW MagBead Kit (Zymo) was used with a modified
113 protocol. Starting culture volume was increased to 500 µl and 50 µl of 0.5 mg/mL lysostaphin (from

114 *Staphylococcus staphylolyticus,* MERCK) was added, instead of lysozyme. Isolates were incubated at 37 °C for
115 1 hour prior to adding ProteinaseK. DNA was quantified with dsDNA HS Qubit Assay (ThermoFisher).

116 **Sequencing**

117 For Illumina sequencing genomic DNA was normalised to 5ng/µl in 10mM Tris-HCl. 0.5 µl of TB1
118 tagmentation DNA Buffer was mixed with 0.5 µl BLT, Tagment DNA Enzyme (Illumina) and 4 µl PCR grade
119 water in a master mix and 5 µl added to a chilled 96 well plate. 2 µl of normalised DNA (10 ng total) was
120 pipette mixed with the 5 µl of the tagmentation mix and heated to 55 °C for 15 minutes in a PCR block. A
121 PCR master mix was made up using 4 ul kapa2G buffer, 0.4 µl dNTP's, 0.08 µl Polymerase and 4.52 µl PCR
122 grade water, contained in the Kap2G Robust PCR kit (MERCK) per sample and 9 µl added to each well need
123 to be used in a 96-well plate. 2 µl of each P7 and P5 of Nextera XT Index Kit v2 index primers (Illumina) were
124 added to each well. Finally, the 7 µl of Tagmentation mix was added and mixed. The PCR was run with 72°C
125 for 3 minutes, 95°C for 1 minute, 14 cycles of 95°C for 10s, 55°C for 20s and 72°C for 3 minutes. Following
126 the PCR reaction, the libraries were quantified using the Promega QuantiFluor® dsDNA System and run on a
127 GloMax® Discover Microplate Reader. Libraries were pooled following quantification in equal quantities. The
128 final pool was double-SPRI size selected between 0.5 and 0.7X bead volumes using KAPA Pure Beads (Roche).
129 The final pool was quantified on a Qubit 3.0 instrument and run on a D5000 ScreenTape (Agilent) using the
130 Agilent Tapestation 4200 to calculate the final library pool molarity.

131 The pool was run at a final concentration of 1.5 pM on an Illumina Nextseq500 instrument using a Mid
132 Output Flowcell (NSQ® 500 Mid Output KT v2(300 CYS) Illumina) following the Illumina recommended
133 denaturation and loading recommendations which included a 1% PhiX spike in (PhiX Control v3 Illumina).
134 Data was uploaded to Basespace (www. basespace.illumina.com) where the raw data was converted to 8
135 FASTQ files for each sample.

136 For long read sequencing up to 400 ng of DNA was incubated with Ultra II End-prep reaction buffer and Ultra
137 II End-prep enzyme mix for end repair. The samples then had native barcodes (NB01-24) added with the
138 addition of the NEB Blunt/TA Ligase Master Mix and then the samples were heated on thermal cycler at 20
139 °C for 20 m, 65 °C for 10 m. The libraries were pooled and cleaned using AMPure XP beads, DNA was
140 quantified using dsDNA HS Qubit (ThermoFisher) and Tapestation (Aligent). For the adaptor ligation step,
141 Adapter Mix II (AMX II), NEBNext Quick Ligation Reaction Buffer (5X) and Quick T4 DNA Ligase was added to
142 the pooled library and cleaned using AMPure XP beads. The library was eluted and loaded on to a primed
143 MinION flow cell (Oxford Nanopore Technologies). Base calling (June 2022) was carried out using Guppy
144 (v6.06) (Oxford Nanopore Technologies).

145 **Genomic and Phylogenetic analysis**

146 Once generated, sequence data was processed via a series of pipelines on a Galaxy instance hosted at the
147 Quadram Institute Bioscience. FASTQ files were used to assigned a microbial classification and check for
148 contamination for each sample using Centrifuge (v0.15) (Kim et al., 2016). *S. capitis* strains were then used
149 for further analysis, isolates were assembled using SPAdes (v3.12.0) (Bankevich et al., 2012), with default
150 parameters applied, and resulting assemblies analysed for quality with QUAST (v5.0.2) (Mikheenko et al.,
151 2018). Isolates with less than X10 genomic coverage were omitted from the final collection.

152 Long read data was assembled using Flye (v2.5) (Lin et al., 2016, Kolmogorov et al., 2019) the Illumina short
153 reads were then mapped to the Flye scaffolds using minimap2 (v2.17) (Li and Durbin, 2010, Li and Durbin,
154 2009) and the sequence files were then polished with 2 rounds of pilon (v1.20.1) (Walker et al., 2014).

155 The phylogeny of the *S. capitis* isolates was determined using gene presence/absence after creating a tree
156 based on the core gene alignment. Prokka (v1.14.5) (Seemann, 2014) was used to annotate the assembled
157 contigs along with nine available *S. capitis* reference genomes (https://www.ncbi.nlm.nih.gov,
158 Supplementary Data 1). The GFF3 files were then submitted to ROARY (v3.13.0) to determine the core and
159 accessory genomes (requiring that 75% percent of isolates should carry a gene to be considered to be in the
160 core) (Page et al., 2015). A phylogenetic tree was inferred using IQTREE (v 1.6.12) (Nguyen et al., 2015) from
161 the core gene alignment output from Roary, and then visualised and annotated in iTOL (Letunic I. and Bork,
162 2019).

163 The presence of antimicrobial resistance and virulence genes were identified using ABRicate (v2.13.2, with
164 75% minimum identity, and 90% minimum coverage filters applied) (Seeman, 2016) and ARIBA (Hunt et al.,
165 2017) using the CARD database (v3.1.1) (Alcock et al., 2020).

166 To identify SNPs, FASTQ files were processed by Snippy (v4.4.3), and Snippy-core (v4.4.3) used to generate a
167 core alignment FASTA which was then analysed by snp-dist (v0.6.3) (Seemann, 2015, Seeman, 2019). The
168 reference genome used for all Snippy tools was the CR05 genome (NZ_CTEO01000001.1). The Snippy-core
169 alignment was then used to create a phylogeny based on core SNPs and run through IQTREE (Nguyen et al.,
170 2015), the resulting maximum likelihood tree was visualised using iTOL (Letunic I. and Bork, 2019).

171 Homologues of specific sequences of interest were identified using BLASTn, sequence data for the *SCCmec-*
172 *SCCcad-SCCars-SCCcop* (KF049201.1) mobile element was used as a reference to identify other similar
173 sequences in the collection using BLASTn megablast (v2.10.1; default settings)) (Camacho et al., 2009, Cock
174 et al., 2015). The NCBI protein database was searched for homologues to Nsr (CDI72769.1) and TarJ
175 (CDI72761.1) with BLASTp (Cock et al., 2015).

176 Alignment of CRISPR-Cas Type-III-A genes from different strains was assessed by taking a reference sequence
177 from *S. capitis* CR05 and creating a Snippy-core full alignment which was then submitted to ClustalW (v1.91)
178 (Larkin et al., 2007).

179 Long read sequences were compared using BLASTn pairwise alignment (NCBI, 2022) and visualised using
180 Artemis Comparison Tool (Carver et al., 2008).

181 **Antimicrobial Susceptibility Testing**

182 The susceptibility of all *S. capitis* isolates to the biocides chlorhexidine gluconate (CHX) and octenidine
183 hydrochloride (OCT) and the antibiotics, benzylpenicillin (PEN), cefotaxime (CEF), daptomycin (DAP),
184 gentamicin (GEN), fusidic Acid (FUS) and vancomycin (VAN) was determined in Mueller-Hinton agar (for
185 daptomycin, Ca^{2+} was added to a final of concentration 50 mg/L) (Sigma-Aldrich) in accordance with EUCAST
186 guidelines. Plates were inoculated with ~10^4 cells diluted from overnight cultures grown in Mueller-Hinton
187 broth (Sigma-Aldrich) using a multi-point inoculator (Denley). *Staphylococcus aureus* controls were run on all
188 the plates, ST 239 ('TW20') and NCTC 8532 ('F77') were used on the biocide plates and ATCC 29213 for the
189 antibiotic plates. The plates were incubated at 37 °C for 24 hours before minimum inhibitory concentrations
190 (MICs) were determined. MIC_{50} and MIC_{90} were determined by identifying the MIC which inhibits 50% or 90%
191 of the collection. Clinical breakpoints from EUCAST (2021) were used where applicable and cut off points for
192 CHX and OCT were set at 4 µg/mL and 2 µg/mL respectively, as defined by Htun et al. (2019). Statistical
193 analysis was performed using Graph Pad Prism (v5.04). For analysis of more than 2 groups a 1-way ANOVA
194 Kruskal-Wallis, Non-Parametric test was used.

195 **Sensitivity to pH**

196 Two randomly selected isolates from each of the Non NRCS-A, proto NRCS-A and NRCS-A groups were used
197 to test for ability to grow at pH (pH range 3.5 -10). Overnight cultures of each were used to inoculate plates
198 in duplicate (PM10 MicroPlate™ pH, Biolog, U.K) in accordance with the manufacturer's protocol and
199 respiratory activity in each sample measured over 24 hours automatically using the OmniLog reader (Biolog,
200 U.K).

201

202 **Results**

203 ***Staphylococcus capitis* were isolated from various body sites of babies on NICUs**

204 Routine MRSA swabs were taken from a total of 159 babies across their stay in NICU, *S. capitis* was isolated
205 from 55 babies (34.5 %) admitted to NICU in the UK and Germany yielding a total of 102 *S. capitis* isolates
206 (UK n=90, German n=14) from 4 different body sites (ear n=47; gut n=13; nose n=17; skin n=27). A further 12
207 isolates were from neonates with infection, 11 from blood and one from skin. A total of 15 *S. capitis* were
208 collected from adults including from prosthetic joint infection (n=3), blood (n=5) healthy skin swabs (n=6)
209 and one from an abdominal swab. All these 129 strains were sequenced and gave an average genome size of
210 2.52 Mbp with an average GC content of 32.86 %. A further 9 previously published reference genomes were
211 also included in later analysis of genomic relationships.

212 **Antiseptic susceptibility differed between isolates from different geographical sites**

213 All the isolates from the NICU CoNS collection were tested for antiseptic and antibiotic susceptibility
214 (Supplementary Data 2, Sethi et al. (2021)). The *S. capitis* isolates demonstrated greater sensitivity to
215 octenidine than chlorhexidine (MIC^{50} of 2 µg/mL and 16 µg/mL respectively, and MIC^{90} being 4 µg/mL and 32
216 µg/mL) (Figure 1). There was no significant difference found in susceptibility of isolates from different body
217 sites, but the MIC^{50} values for both octenidine and chlorhexidine in the German isolates (OCT 1 µg/mL; CHX 8
218 µg/mL) were lower than those of the UK isolates (OCT 2 µg/mL; CHX 16 µg/mL). As well as the antiseptics,
219 UK isolates had higher MIC^{50} values for gentamicin, penicillin and fusidic acid. The largest discrepancy being
220 for fusidic acid where the MIC^{50} increased from 0.625 µg/mL for the German population to 4 µg/mL in UK
221 isolates. Seven adult isolates showed hetero-resistance to daptomycin (defined as an MIC ≥ 1 µg/mL, Becker
222 et al. (2014)) but this was not seen in any neonatal isolates. No vancomycin resistance was observed in the
223 isolates with the highest MIC being 2 µg/mL, however nearly a quarter of the isolates exhibited
224 heteroresistance to vancomycin (29 in total), with the majority of these (23/29) being from NICU.

225 **Phylogenetic characterisation of *S. capitis* identified two major clades**

226 A phylogenetic tree was built using the core gene alignment output by Roary of all isolates in the panel (as
227 well as the nine available NCBI reference *S. capitis* genomes in Supplementary Data 1) then aligned with
228 IQTREE (Figure 1). Clades including NRCS-A strains were determined based on similarity to known NRCS-A
229 reference strains (CR01, 03 and 05) and by the presence of the two genes *nsr* and *tarJ,* common identifiers
230 used for this clone (Martins Simões et al., 2013). The tree showed a diversity of non-NRCS-A clones and then
231 a clonal group which all carried *nsr* and *tarJ* and included the NRCS-A strains but could be divided into two
232 further groups based on carriage of a CRISPR system.

233 **The 'non NRCS-A'** group (Figure 1, green, n=27) contained no known NRCS-A strains but did include other *S.*
234 *capitis* NCBI reference genomes and isolates from adults as well as some NICU isolates. Isolates in this group
235 tended to show a low tolerance to OCT and CHX (OCT MIC_{50} = 2 µg/mL, CHX MIC_{50} = 8 µg/mL; Table 1 and
236 fusidic acid with *fusB* being absent in all but one strain (Supplementary Data 3). Of these non NRCS-A isolates
237 12 came from the NICU (4 German, 8 UK), nine came from healthy adult skin swabs (n=5) or blood cultures
238 (n=4) and this group also included six non NRCS-A reference strains (*S. capitis* strains AYP1020, C87, QN1,
239 SK14, VCU116 and NCTC11045 (Supplementary Data 1)).

240 **The 'Proto NRCS-A'** group (denoted as such as it appears ancestrally related to the NRCS-A group and
241 contain *nsr* and *tarJ* but no CRISPR system) (Figure 1, orange, n=16) carried fewer AMR genes than the
242 NRCS-A isolates and typically lacked *fusB,* however these strains were closely related to the NRCS-A group.
243 This proto-NRCS-A clade had OCT and CHX susceptibilities lower than NRCS-A (OCT MIC_{50} = 2 µg/mL, CHX
244 MIC_{50} = 8 µg/mL: Table 1). Along with ten isolates from NICU, the remaining isolates were sourced from
245 adult PJI (n=3), blood (n=1), an abdominal swab and a healthy adult skin swab.

246 **The 'NRCS-A'** group (Figure 1, red, n=95) was the largest and contained most isolates as well as NRCS-A
247 reference strains, *S. capitis* CR01, CR03 and CR05 along with the remaining *S. capitis* isolated from the NICU.
248 No isolates from adults were found in this clade. Tolerance to OCT and CHX was marginally higher than
249 isolates from proto NRCS-A (OCT MIC_{50} = 2 µg/mL, CHX MIC_{50} = 16 µg/mL: Table 1). NRCS-A was defined by
250 isolates carrying multiple AMR genes including *blaZ* (β-lactamase)*, fusB* (fusidic acid resistance)*, AAC(6')-Ia-*
251 *APH(2")-Ia* (aminoglycoside resistance) and *mecA* (penicillin/methicillin resistance) and contained the *nsr*
252 and *tarJ* genes as well as a CRISPR-Cas Type III-A system (Figure 1 and Supplementary Data 3).

253 Multiple isolates of *S. capitis* were isolated on different sampling dates from 24 separate babies (

254 Figure 2) with the remaining 40 babies only yielding a single isolate. Only 4/64 babies carried only non NRCS-
255 A *S. capitis* strains (babies 48, 58, 94 and 146, Figure 2). The remaining 60 babies carried NRCS-A isolates
256 (Supplementary Data 2). Three babies (baby 8, 59, and 137) yielded *S. capitis* isolates from two of the three
257 groups (Non NRCS-A 1 and NRCS-A), and one baby had two distinct isolates from the proto NRCS-A group
258 (baby 28), and another yielded one non NRCS-A and one proto NRCS-A isolate (Supplementary Data 2).

259 The number of SNPs distinguishing all 138 isolates was determined using snippy with UK NRCS-A CR05 as a
260 reference. The mean (standard deviation) number of SNPs seen between isolates in the NRCS-A, and proto
261 NRCS-A groups were 115.47 and 265 respectively (Supplementary Data 4). On average the non NRCS-A
262 isolates contained 23084.5 (3263.2) SNPs, a much higher number confirming that this clade is genetically
263 distant from the NRCS-A isolates and more diverse. The 92 NRCS-A isolates from NICU had a maximum of 31
264 snps between each other (mean aligned core genome: 2401507 bp), these isolates were only isolates from

265 the UK (Supplementary Data 4). This high degree of relatedness of isolates in the UK NICUs suggests a recent
266 common ancestor with transition across the unit and further confirms the nature of NRCS-A.

267 **Evidence for transmission within the NICU**

268 We analysed relationships between isolates taken longitudinally from babies (

269 Figure 2), this revealed a cluster of 8 isolates (209, 878, 33, 97, 191, 1048, 355, 489) differentiated by only 3
270 SNPS, (Supplementary Data 5) recovered between over 70 days, from six different babies (

271 Figure 2) suggesting transmission within the NICU environment and persistence of isolates over time. The
272 total number of SNPs between all NRCS-A isolates was only 37. This indicates that very similar closely related
273 isolates were present in the UK NICUs and they remain present for years with very little genomic change.

274 **Differences distinguishing clades**

275 The collection was analysed by Roary, (Page et al., 2015) to identify genes that were either present or absent
276 in the clades, 'Scoary' was then used to calculate statistical associations between the presence and absence
277 of genes, clade membership and a series of phenotypes (Brynildsrud et al., 2016). Notably the genes being
278 most significantly associated with the NRCS-A clade were the CRISPR genes, all NRCS-A isolates harboured a
279 complete CRISPR-Cas Type-III-A system. None of the closely related proto NRCS-A isolates harboured any
280 CRISPR Cas Type-III-A genes. These Cas Type-III-A genes are inserted in the *SCCmec-SCCcad/ars/cop* element
281 (Figure 3). Two virulence genes *paaZ* and *ugpQ,*along with *dut* are known to be closely located to *mecA,* and
282 were also strongly associated with being present in the NRCS-A isolates (Datta et al., 2021, Westberg et al.,
283 2022).

284 Scoary analysis also identified groups of other genes significantly associated with the NRCS-A isolates. Many
285 of these genes were annotated as being of unknown function and clustered together in the genomes.
286 Further analysis of these genes, using Blastn identified several as being putative phage terminases which are
287 inserted in-between *yfnK* and *murI,* immediately downstream from *frdA, uvrC* and *mutS.* Scoary also
288 identified *mcrC* in NRCS-A isolates which is involved in regulation of degrading methylated bacteriophage
289 DNA (Nirwan et al., 2019) (Table 2, Supplementary Data 6).

290 Previous research by Martins Simões et al. (2013) suggested the presence of *nsr* and *tarJ* as a defining
291 characteristic of the NRCS-A clone, and both these genes are present in both NRCS-A and proto NRCS-A
292 isolates. The *tarLIJ* locus encodes a Ribulose-5-phosphate reductase which are involved in teichoic acid
293 biosynthesis (Tevell et al., 2020, Martins Simões et al., 2013). These were significantly associated with the
294 NRCS-A group, parts of the operon were found in Non NRCS-A but were usually incomplete (Figure 1, Table
295 2). All NRCS-A and proto NRCS-A isolates had *tarJ* compared to only 3.7% of non NRCS-A isolates (Figure 1).
296 The *nsr* gene associated with nisin resistance was identified in both NRCS-A and proto NRCS-A isolates, and

297 *kdpDC* was also significantly over-represented in these groups which encodes a potassium pump encoded
298 close to *nsr* in the genome (Martins Simões et al., 2013).

299 There was a significant increase in the presence of genes involved in metal transport in the NRCS-A group
300 compared to the others (Table 3, Supplementary Data 6). ABC permeases *cnt ABCDEFKL, nikC* and *oppDF*
301 were all identified and are involved in the acquisition of Zn and Ni (Grim et al., 2017). Further genes involved
302 in metal detoxification were also seen in NRCS-A, the *copB, copAZ, arsBC* and *cadC* systems were over-
303 represented in this clade. These have all been previously identified as part of an SCC*mec* element (*SCCmec-*
304 *SCCcad/ars/cop*) also containing the CRISPR-Cas Type III-A system in NCRS-A isolates. As well as genes
305 associated with this element, the NRCS-A isolates could also be distinguished from the proto-NRCS-A clade
306 by the presence of *mntH* (Table 3, Supplementary Data 6).

307 Previously it has been suggested that NRCS-A isolates are adapted to survive in both the gut and on the skin,
308 two environments where pH differs. We observed that growth of both proto NRCS-A and NRCS-A isolates at
309 lower pH was higher compared to non NRCS-A isolates (

310 Figure 4). Growth at pH 6 and pH 5.5 was also marginally higher compared to pH 7 in both the NRCS-A and
311 proto NRCS-A isolates (

312 Figure 4).

313 To compare conservation of the *SCCmec* element seen in all the isolates, the S*CCmec-SCC-cad/ars/cop*
314 sequence of CR01 (GenBank KF049201.1) was compared to all the *S. capitis* isolates using snippy-core and
315 snp-dist (Seemann, 2015, Seeman, 2019). All isolates and references except for VCUII6 aligned to the
316 S*CCmec-SCC-cad/ars/cop* reference. The NRCS-A isolates, on average, had 12.3 ± 0.5 SNPs, which increased
317 to 64.4 ± 11.1 in non NRCS-A and 190.9 ± 22.5 in proto NRCS-A isolates. The higher divergence in the proto
318 NRCS-A group is surprising given its closer relationship to the NRCS-A group according to the core genome
319 phylogeny.

320 Further analysis of the CRISPR-Cas Type III-A system showed that the genes were highly conserved when
321 aligned. In all isolates it was possible to identify direct repeat (DR) sequences (beginning with GATAACT), and
322 there were three different DRs present in the isolates as previously observed in CR01 Cao et al. (2016). In a
323 full SNP ClustalW alignment DR5 (4 times) and DR6 (3 times) were found before the Cas genes, and then two
324 repeats of DR8 towards the end of the alignment after the CRISPR genes (Cao et al., 2016).

325 **Carriage of AMR genes varied between non NRCS-A and NRCS-A**

326 All isolates contained *lmrS, norA, arlR* and *dfrC*, and all but one isolate contained *mgrA*. NRCS-A was defined
327 by the additional presence of *blaZ, mecA, ACC(6')-le(APH)(2')-la* and *fusB* in the majority of isolates whereas

328 proto NRCS-A had a smaller number of AMR genes which were interspersed across the clade without a
329 consistent pattern (Figure 1, Supplementary Data 3). Analysis of the pangenomes with Scoary was used to
330 identify genes correlating with susceptibility to different antimicrobials. Unsurprisingly *mecA* was strongly
331 associated with high cefotaxime MICs (4 µg/mL, Supplementary Data 6). Carriage of *blaZ* was also associated
332 with penicillin resistance (≥1 µg/mL, P= 0.005), however CRISPR genes (P= 0.002) and other genes commonly
333 found on the S*CCmec-SCC-cad/ars/cop* MGE were more significant (Table 4)

334 . The *qac* family of genes has previously been implicated in tolerance to antiseptics, *qacA* was detected but
335 only present in 22/138 (16 %) isolates and there was no correlation with the presence of *qacA* and CHX
336 tolerance. Carriage of *qacA* was only observed in 16.25% of isolates with high chlorhexidine MICs ≥16 µg/mL
337 (Figure 1). For isolates that had an MIC to chlorhexidine of ≥16 µg/mL the genes mostly associated with this
338 tolerance were the CRISPR-Cas Type III-A, Supplementary Data 6). This is a marker of NRCS-A isolates which
339 had significantly higher chlorhexidine MICs (

340 Figure 5) although there was no obvious genetic basis for this. Scoary was unable to identify any significant
341 genes associated with octenidine susceptibility.

342 The classic antibiogram for NRCS-A isolates includes being resistant to fusidic acid (PHE, 2021), with the
343 clinical breakpoint being ≥0.5 µg/mL for all CoNS (EUCAST, 2021). Based on this criterion 79% of our
344 collection tested were clinically resistant and carriage of *fusB* was seen in 78.75 % of these resistant isolates,
345 leaving 17 isolates with no known mechanism of fusidic acid resistance. The presence of *fusB* was mainly
346 observed in NRCS-A, where there was a small increase in tolerance to fusidic acid, with the mean MIC being
347 2.95 (SD: 0.155) µg/mL, higher than non NRCS-A (1.40, SD: 0.301 µg/mL). The UK isolates had a higher
348 median FUS MIC (4 µg/mL) than the German isolates (1.5 µg/mL) (P = <0.01). Penicillin resistance encoded
349 by *mecA* was seen in 86% of the isolates with an MIC ≥ 2 µg/mL (Supplementary Data 2 and 3), and carriage
350 of *ACC(6')-Ia-APH(2")-Ia* was seen in 73% of isolates with an MIC ≥1 µg/mL of gentamicin. No genes were
351 associated with ciprofloxacin resistance although the main mechanism of resistance to this agent is target
352 site mutation within *gyrA* rather than the presence of a gene.

353 *S. capitis* NRCS-A are often referred to as demonstrating hetero-resistance to vancomycin, indicated by a
354 MIC of 2 µg/mL (Satola et al., 2011). Despite 23.5% of the isolates having an MIC of 2 µg/mL there was no
355 presence of *vanA* apart from in one isolate (with an MIC of 1 µg/mL) or any other significant genes
356 associated with vancomycin resistance.

357 **Discussion**

358 Even with the increased focus on their prevention, nosocomial infections remain a major cause of disease
359 and death in hospitals. Transmission of bacterial pathogens can occur via multiple routes and vectors with
360 both health care workers and the environment sources of infection for many pathogens such as
361 *Staphylococcus aureus, Pseudomonas aeruginosa* and Enterobacteriaceae (Risso et al., 2019, Götting et al.,
362 2020). The effective adaption of the NRCS-A clone to the NICU environment is well established, though the
363 reasons for this remain unclear. One suggestion is that NRCS-A isolates are adapted to survival in the gut
364 which provides an additional reservoir to the skin and may allow persistent colonisation, even when skin
365 organisms are lost during antisepsis (Butin et al., 2017). Carriage of the *nsr* gene responsible for nisin
366 resistance has been proposed to aid survival in the gut where this antimicrobial peptide is produced (Martins
367 Simões et al., 2013). In this study we documented carriage of NRCS-A on both the skin of neonates and in the
368 gut, and this was much more common than for other groups of *S. capitis* (Figure 1). We did observe that the
369 *nsr* gene is present in NRCS-A isolates but not in more distantly related strains. Interestingly we identified a
370 group of isolates closely related to NCRS-A, the 'proto NRCS-A' strains which did carry *nsr* but which were
371 isolated from the gut less often (Figure 1). It was also observed that the proto NRCS-A and NRCS-A isolates,
372 which possess *nsr* and *tarJ,* had increased growth in pH 6-5.5 cultures (

373 Figure 4), analogous to their niche in the gastrointestinal tract.

374 *Metal acquisition and survival*

375 Analysis of the pangenomes of each group identified other genes that may aid the survival of the NRCS-A in
376 the NICU. In the NRCS-A isolates there was a marked increase in genes associated with metal chelation
377 (Table 2), availability of many metals is limited *in vivo* and in the gut metals are precious commodities. Metal
378 chelators can also influence virulence, for example, the nickel/cobalt transporter *nixA* in *S. aureus* has been
379 implicated in an increase in virulence in UTI and kidney infection and is required for urease activity in mice
380 (Remy et al., 2013). Other metal related genes found in NRCS-A were *mntH*, a metal transporter found to
381 minimise impact of Mn and Zn starvation (Kehl-Fie et al., 2013), and *cntABCDEF* which functions with
382 Staphylopine, StP (CntKLM) a staphylococcal zincophore involved in Zn acquisition which can also bind to Cu,
383 Co and Ni (Grim et al., 2017, Song et al., 2018). These high affinity chelators and zincophores have been
384 shown to be able to outcompete human calprotectin (a Zn chelator excreted by neutrophils) to evade
385 nutritional immunity. The presence of *cntABCDEF cntKLM, nixA* and *mntH* may indicate that even in
386 nutritionally deprived environments, the NRCS-A isolates are able to acquire essential metal ions and survive
387 in a gut environment. Similarly, the presence of *copA, cad,* and *ars,* encoded within the *SCCmec-*
388 *SCCcad/ars/cop* element in NRCS-A provide defence against toxic metals by encoding efflux systems.
389 Together these data suggest that an enhanced ability to scavenge and survive metal exposure may be crucial

390 for the success of NRCS-A and supports the theory that gut colonisation is a key strategy exploited by this
391 clone.

392 *Phage defence*

393 If NRCS-A has developed a lifestyle which includes ability to colonise the gut then it also needs to adapt to
394 attack from phages, which are prevalent in this environment. In *Escherichia coli* the *mcrBC* genes provide
395 protection from phages and McrC activates McrB to degrade methylated bacteriophage DNA (Nirwan et al.,
396 2019). MrcC was seen in our collection but *mcrB* was not detected and it is unclear what impact *mcrC* alone
397 may have. Other mechanisms of phage defence observed in the NRCS-A isolates were the CRISPR-associated
398 genes (Cas), used in defence against MGEs, viruses and plasmids (Tamulaitis et al., 2017, Rossi et al., 2017).
399 CRISPR-Cas systems show low abundance in most Non-aureus Staphylococci although they have been
400 documented in the genomes of *S. aureus* strains as well as in *S. epidermidis, S. schleiferi, S. haemolyticus, S.*
401 *lugdunensis*. Carriage of a CRISPR-Cas system is a known feature of NRCS-A *S capitis* and CRISPR-Type III-A
402 Cas genes were identified in all NRCS-A isolates. This was the main defining feature differentiating the proto
403 NRCS-A and the NRCS-A clades, however when this system was acquired by NCRS-A remains unclear. The
404 direct repeats identified are similar to that as previously seen in CR01 and CR03 which both contain the
405 conserved CCCC and GGGG motifs which are thought to form hairpin structures (Rossi et al., 2017). The
406 highly conserved regions suggest that a common ancestor is responsible for the acquisition of the CRISPR-
407 Cas Type-III-A genes.

408 *Antiseptic tolerance*

409 One defining feature of the NRCS-A clade was increased tolerance of chlorhexidine (Figure 5), this was not
410 associated with carriage of *qac* genes, and the mechanism remains unclear. This increased ability to survive
411 chlorhexidine exposure may allow both survival of skin antisepsis but also environmental decontamination
412 where chlorhexidine is a common part of cleaning and disinfection routines for incubators and medical
413 equipment. There was no elevation in tolerance to octenidine, which is the main alternative for antisepsis in
414 neonates (Sethi et al., 2021).

415 *Stability of NRCS-A*

416 The NRCS-A isolates compared here were very similar including those separated by geography (UK and
417 German isolates) and time (reference genomes spanning years in isolation) (

418 Figure 2), and there were no more than 37 SNPs separating strains in the core genome. We provide evidence
419 for transfer between babies on the NICU (

420 Figure 2), studies identifying NRCS-A in the general environment and on different surfaces or medical
421 implements following clusters of LOS NRCS-A positive cases have suggested a role for the environment in
422 transmission The isolates from carriage and the 10 from neonatal blood cultures were highly similar (Figure
423 1). This supports the idea that carriage acts as a reservoir for infection but clearly carriage by neonates in a
424 nosocomial setting does not usually result in infection.

425 **Conclusions**

426 We have identified that carriage of the NRCS-A clone is abundant on the skin and gut of uninfected neonates
427 and that they are likely to be transferred within NICUs. NRCS-A isolates separated in time and space showed
428 little genetic variation and carriage isolates were indistinguishable to those from blood culture, suggesting
429 carriage can be a precursor to infection. We identified a closely related proto-NRCS-A clade which was less
430 associated with gut carriage, did not possess the CRISPR-Cas III-A system, demonstrated lower antimicrobial
431 resistance and chlorhexidine tolerance and carried fewer metal acquisition or detoxification genes. Our data
432 supports the idea that gut colonisation is a key survival strategy of NRCS-A and expands our understanding
433 of the likely mechanisms employed by this clone to both survive on skin, in the environment and in the gut.
434 This opens the possibility to develop a probiotic approach to hinder colonisation of NCRS-A based upon
435 nutritional immunity where strains that sequester essential metals may prove beneficial to the infant gut.
436 Recent deployment of a probiotic product into the NNUH NICU has been used to reduce incidence of
437 necrotising enterocolitis with success and was also associated with reduced episodes of infection with CoNS
438 ((Robertson et al., 2020). More work can help develop strategies to identify reservoirs and transmission
439 routes of NRCS-A and to eradicate carriage from neonates to prevent infection which can have serious
440 consequences.

441 **References**

442 ALCOCK, B., RAPHENYA, A., LAU, T., TSANG, K., BOUCHARD, M., EDALATMAND, A., HUYNH, W., NGUYEN, A.,
443 CHENG, A., LIU, S., MIN, S., MIROSHNICHENKO, A., TRAN, H., WERFALLI, R., NASIR, J., OLONI, M.,
444 SPEICHER, D., FLORESCU, A., SINGH, B., FALTYN, M., HERNANDEZ-KOUTOUCHEVA, A., SHARMA, A.,
445 BORDELEAU, E., PAWLOWSKI, A., ZUBYK, H., DOOLEY, D., GRIFFITHS, E., MAGUIRE, F., WINSOR, G.,
446 BEIKO, R., BRINKMAN, F., HSIAO, W., DOMSELAAR, G. & MCARTHUR, A. 2020. CARD 2020: antibiotic
447 resistome surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids*
448 *Research,* 48**,** D517-D525.
449 BANKEVICH, A., NURK, S., ANTIPOV, D., GUREVICH, A., DVORKIN, M., KULIKOV, A., LESIN, V., NIKOLENKO, S.,
450 PHAM, S., PRJIBELSKI, A., PYSHKIN, A., SIROTKIN, A., VYAHHI, N., TESLER, G., ALEKSEYEV, M. & PA., P.
451 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.
452 *Journal of Computational Biology,* 19.
453 BECKER, K., BOTH, A., WEIßELBERG, S., HEILMANN, C. & ROHDE, H. 2020. Emergence of coagulase-negative
454 staphylococci. *Expert Review of Anti-infective Therapy,* 18**,** 349-366.
455 BECKER, K., HEILMANN, C. & PETERS, G. 2014. Coagulase-Negative Staphylococci. *Clinical Microbiology*
456 *Reviews,* 27**,** 870-926.
457 BEN SAID, M., HAYS, S., BONFILS, M., JOURDES, E., RASIGADE, J. P., LAURENT, F. & PICAUD, J. C. 2016. Late-
458 onset sepsis due to *Staphylococcus capitis* 'neonatalis' in low-birthweight infants: a new entity?
459 *Journal of Hospital Infection,* 94**,** 95-98.
460 BENTLIN, M., DE SOUZA, R. & MS., L. 2010. Late-onset Sepsis: Epidemiology, Evaluation, and Outcome.
461 *NeoReviews,* 11**,** e426-e435.
462 BRYNILDSRUD, O., BOHLIN, J., SCHEFFER, L. & ELDHOLM, V. 2016. Rapid scoring of genes in microbial pan-
463 genome-wide association studies with Scoary. *Genome Biol.,* 17.
464 BUTIN, M., DUMONT, Y., MONTEIX, A., RAPHARD, A., ROQUES, C., MARTINS SIMOES, P., PICAUD, J.-C. &
465 LAURENT, F. 2019. Sources and reservoirs of Staphylococcus capitis NRCS-A inside a NICU.
466 *Antimicrobial Resistance & Infection Control,* 8**,** 157.
467 BUTIN, M., MARTINS-SIMÕES, P., RASIGADE, J.-P., PICAUD, J.-C. & LAURENT, F. 2017. Worldwide Endemicity
468 of a Multidrug-Resistant *Staphylococcus capitis* Clone Involved in Neonatal Sepsis. *Emerging*
469 *infectious diseases,* 23**,** 538-539.
470 CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009.
471 BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.
472 CAO, L., GAO, C.-H., ZHU, J., ZHAO, L., WU, Q., LI, M. & SUN, B. 2016. Identification and functional study of
473 type III-A CRISPR-Cas systems in clinical isolates of Staphylococcus aureus. *International Journal of*
474 *Medical Microbiology,* 306**,** 686-696.
475 CARTER, G., USSHER, J., DA SILVA, A., BAINES, S., HEFFERNAN, H., RILEY, T., BROADBENT, R., VAN DER
476 LINDEN, A., LEE, J., MONK, I., STINEAR, T., HOWDEN, B. & WILLIAMSON, D. 2018. Genomic Analysis
477 of Multiresistant *Staphylococcus capitis* Associated with Neonatal Sepsis. *Antimicrobial Agents and*
478 *Chemotherapy,* 62**,** e00898-18.
479 CARVER, T., BERRIMAN, M., TIVEY, A., PATEL, C., BÖHME, U., BARRELL, B. G., PARKHILL, J. & RAJANDREAM,
480 M.-A. 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational
481 database. *Bioinformatics,* 24**,** 2672-2676.
482 COCK, P. J. A., CHILTON, J. M., GRÜNING, B., JOHNSON, B. E. & SORANZO, N. 2015. NCBI BLAST+ integrated
483 into Galaxy. *GigaScience,* 4.
484 DATTA, M. S., YELIN, I., HOCHWALD, O., KASSIS, I., BORENSTEIN-LEVIN, L., KUGELMAN, A. & KISHONY, R.
485 2021. Rapid methicillin resistance diversification in Staphylococcus epidermidis colonizing human
486 neonates. *Nature Communications,* 12**,** 6062.
487 DONG, Y. & SPEER, C. 2014. The role of Staphylococcus epidermidis in neonatal sepsis: Guarding angel or
488 pathogenic devil? . *Int J Med Microbiol,* 304**,** 513-20.
489 DU, X., LARSEN, J., LI, M., WALTER, A., SLAVETINSKY, C., BOTH, A., SANCHEZ CARBALLO, P., STEGGER, M.,
490 LEHMANN, E., LIU, Y., LIU, J., SLAVETINSKY, J., DUDA, K., KRISMER, B., HEILBRONNER, S.,

491 WEIDENMAIER, C., MAYER, C., ROHDE, H., WINSTEL, V. & PESCHEL, A. 2021. Staphylococcus
492 epidermidis clones express Staphylococcus aureus-type wall teichoic acid to shift from a commensal
493 to pathogen lifestyle. *Nature Microbiology*.
494 EUCAST 2021. European Committee on Antimicrobial Susceptibility Testing.
495 FELGATE, H., CROSSMAN, L. C., GRAY, E., CLIFFORD, R., CORREIA, A., DEAN, R., WAIN, J. & LANGRIDGE, G. C.
496 2021. Known mechanisms account for less than half of antimicrobial resistance in a diverse
497 collection of non-aureus staphylococci. *bioRxiv*, 2021.06.22.449369.
498 GÖTTING, T., REUTER, S., JONAS, D., HENTSCHEL, R., HENNEKE, P., KLOTZ, D., HOCK, S., WOLKEWITZ, M.,
499 BLÜMEL, B., HÄCKER, G., GRUNDMANN, H. & MUTTERS, N. 2020. Protocol for a prospective cohort
500 study: Prevention of Transmissions by Effective Colonisation Tracking in Neonates (PROTECT-Neo).
501 *BMJ open,* 10, e034068-e034068.
502 GRIM, K. P., SAN FRANCISCO, B., RADIN, J. N., BRAZEL, E. B., KELLIHER, J. L., PÁRRAGA SOLÓRZANO, P. K.,
503 KIM, P. C., MCDEVITT, C. A. & KEHL-FIE, T. E. 2017. The Metallophore Staphylopine Enables
504 Staphylococcus aureus To Compete with the Host for Zinc and Overcome Nutritional Immunity.
505 *mBio,* 8.
506 HARVEY, E. J., ASHIRU-OREDOPE, D., HILL, L. F., DEMIRJIAN, A. & TEAM., U. K. H. S. A. S. C. I. M. 2022. Need
507 for standardized vancomycin dosing for coagulase-negative staphylococci in hospitalized infants. *Clin*
508 *Microbiol Infect*.
509 HTUN, H. L., HON, P. Y., HOLDEN, M. T. G., ANG, B. & CHOW, A. 2019. Chlorhexidine and octenidine use,
510 carriage of *qac* genes, and reduced antiseptic susceptibility in methicillin-resistant *Staphylococcus*
511 *aureus* isolates from a healthcare network. *Clinical Microbiology and Infection,* 25, 1154.e1-1154.e7.
512 HUNT, M., MATHER, A., SÁNCHEZ-BUSÓ, L., PAGE, A., PARKHILL, J., KEANE, J. & HARRIS, S. 2017. ARIBA: rapid
513 antimicrobial resistance genotyping directly from sequencing reads. *Microbial Genomics,* 3.
514 KEHL-FIE, T. E., ZHANG, Y., MOORE, J. L., FARRAND, A. J., HOOD, M. I., RATHI, S., CHAZIN, W. J., CAPRIOLI, R.
515 M. & SKAAR, E. P. 2013. MntABC and MntH contribute to systemic Staphylococcus aureus infection
516 by competing with calprotectin for nutrient manganese. *Infection and immunity,* 81, 3395-3405.
517 KIM, D., SONG, L., BREITWIESER, F. & SALZBERG, S. 2016. Centrifuge: rapid and sensitive classification of
518 metagenomic sequences. *Genome research,* 26, 1721-1729.
519 KOLMOGOROV, M., YUAN, J., LIN, Y. & PEVZNER, P. 2019. Assembly of long, error-prone reads using repeat
520 graphs. *Nature Biotechnology,* 37, 540-546.
521 LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H.,
522 VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G.
523 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23, 2947-2948.
524 LEMRISS, H., MARTINS SIMÕES, P., LEMRISS, S., BUTIN, M., IBRAHIMI, A., EL KABBAJ, S., RASIGADE, J. &
525 LAURENT, F. 2014. Non-contiguous finished genome sequence of Staphylococcus capitis CR01
526 (pulsetype NRCS-A). *Standards in genomic sciences,* 9, 1118-1127.
527 LEPAINTEUR, M., ROYER, G., BOURREL, A. S., ROMAIN, O., DUPORT, C., DOUCET-POPULAIRE, F. &
528 DECOUSSER, J. W. 2013. Prevalence of resistance to antiseptics and mupirocin among invasive
529 coagulase-negative staphylococci from very preterm neonates in NICU: the creeping threat? *J Hosp*
530 *Infect,* 83, 333-6.
531 LETUNIC I. & BORK, P. 2019. *Tree Of Life v4.4.2* [Online]. Available: https://itol.embl.de/ [Accessed 2019].
532 LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform.
533 *Bioinformatics,* 25, 1754-1760.
534 LI, H. & DURBIN, R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform.
535 *Bioinformatics,* 26, 589-595.
536 LIN, Y., YUAN, J., KOLMOGOROV, M., SHEN, M., CHAISSON, M. & PEVZNER, P. 2016. Assembly of long error-
537 prone reads using de Bruijn graphs. *Proceedings of the National Academy of Sciences,* 113, E8396-
538 E8405.
539 MAKHOUL, I., SUJOV, P., SMOLKIN, T., LUSKY, A. & REICHMAN, B. 2005. Pathogen-Specific Early Mortality in
540 Very Low Birth Weight Infants with Late-Onset Sepsis: A National Survey. . *Clin Infect Dis,* 40, 218-24.

541 MARTINS SIMÕES, P., RASIGADE, J. P., LEMRISS, H., BUTIN, M., GINEVRA, C., LEMRISS, S., GOERING, R. V.,
542 IBRAHIMI, A., PICAUD, J. C., EL KABBAJ, S., VANDENESCH, F. & LAURENT, F. 2013. Characterization of
543 a novel composite staphylococcal cassette chromosome mec (SCCmec-SCCcad/ars/cop) in the
544 neonatal sepsis-associated *Staphylococcus capitis* pulsotype NRCS-A. *Antimicrobial agents and*
545 *chemotherapy,* 57**,** 6354-6357.
546 MIKHEENKO, A., PRJIBELSKI, A., SAVELIEV, V., ANTIPOV, D. & GUREVICH, A. 2018. Versatile genome assembly
547 evaluation with QUAST-LG. *Bioinformatics.* 5.0.2 ed.
548 NCBI, N. C. F. B. I. 2022. *The Basic Local Alignment Search Tool (BLAST)* [Online]. Available:
549 https://blast.ncbi.nlm.nih.gov/Blast.cgi [Accessed 2022].
550 NGUYEN, L.-T., SCHMIDT, H., VON HAESELER, A. & MINH, B. 2015. IQ-TREE: a fast and effective stochastic
551 algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution,* 32**,** 268-
552 274.
553 NIRWAN, N., ITOH, Y., SINGH, P., BANDYOPADHYAY, S., VINOTHKUMAR, K. R., AMUNTS, A. & SAIKRISHNAN,
554 K. 2019. Structure-based mechanism for activation of the AAA+ GTPase McrB by the endonuclease
555 McrC. *Nature Communications,* 10**,** 3058.
556 PAGE, A. J., CUMMINS, C., HUNT, M., WONG, V., REUTER, S., HOLDEN, M., FOOKES, M., FALUSH, D., KEANE, J.
557 & PARKHILL, J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics,* 31**,**
558 3691-3.
559 PHE. 2021. *Briefing Note 0 2021/005 Increased incidence of detections of Staphylococcus capitis in neonatal*
560 *clinical samples* [Online]. Public Health England. Available: https://hubble-live-
561 assets.s3.amazonaws.com/bapm/attachment/file/411/Increased_incidence_of_detections_of_Stap
562 hylococcus_capitis_in_neonates.pdf [Accessed April 2021].
563 RASIGADE, J. P., RAULIN, O., PICAUD, J. C., TELLINI, C., BES, M., GRANDO, J., BEN SAÏD, M., CLARIS, O.,
564 ETIENNE, J., TIGAUD, S. & LAURENT, F. 2012. Methicillin-resistant Staphylococcus capitis with
565 reduced vancomycin susceptibility causes late-onset sepsis in intensive care neonates. *PLoS One,* 7**,**
566 e31548.
567 REMY, L., CARRIÈRE, M., DERRÉ-BOBILLOT, A., MARTINI, C., SANGUINETTI, M. & BOREZÉE-DURANT, E. 2013.
568 The Staphylococcus aureus Opp1 ABC transporter imports nickel and cobalt in zinc-depleted
569 conditions and contributes to virulence. *Molecular Microbiology,* 87**,** 730-743.
570 RISSO, F., MINGHETTI, D., MARIANI, M., PARODI, A., LA MASA, D., CASTAGNOLA, E. & RAMENGHI, L. A. 2019.
571 Behaviours monitoring and infection control in neonatal intensive care unit: how to improve
572 ourselves? *J. Prev Med. Hyg.,* 60**,** E226-E228.
573 ROBERTSON, C., SAVVA, G., CLAPUCI, R., JONES, J., MAIMOUNI, H., BROWN, E., MINOCHA, A., HALL, L. &
574 CLARKE, P. 2020. Incidence of necrotising enterocolitis before and after introducing routine
575 prophylac *Lactobacillus* and *Bifidobacterium* probiotics. *Archives of Disease in Childhood - Fetal and*
576 *Neonatal Edition,* 105**,** 380.
577 ROSSI, C. C., SOUZA-SILVA, T., ARAÚJO-ALVES, A. & GIAMBIAGI-DEMARVAL, M. 2017. CRISPR-Cas Systems
578 Features and the Gene-Reservoir Role of Coagulase-Negative Staphylococci. *Frontiers in*
579 *Microbiology,* 8.
580 SATOLA, S. W., FARLEY, M. M., ANDERSON, K. F. & PATEL, J. B. 2011. Comparison of detection methods for
581 heteroresistant vancomycin-intermediate Staphylococcus aureus, with the population analysis
582 profile method as the reference method. *J Clin Microbiol,* 49**,** 177-83.
583 SEEMAN, T. 2016. ABRicate: mass screening of contigs for antiobiotic resistance genes. Github.
584 SEEMAN, T. 2019. snp-dists. GitHub repository: GitHub.
585 SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* 15**,** 2068-9.
586 SEEMANN, T. 2015. snippy: fast bacterial variant calling from NGS reads.
587 SETHI, D. K., FELGATE, H., DIAZ, M., FAUST, K., KIY, C., CLARKE, P., HÄRTEL, C., RUPP, J. & WEBBER, M. A.
588 2021. Chlorhexidine gluconate usage is associated with antiseptic tolerance in staphylococci from
589 the neonatal intensive care unit. *JAC Antimicrob Resist,* 3**,** dlab173.
590 SONG, L., ZHANG, Y., CHEN, W., GU, T., ZHANG, S.-Y. & JI, Q. 2018. Mechanistic insights into staphylopine-
591 mediated metal acquisition. *Proceedings of the National Academy of Sciences,* 115**,** 3942.

592 TAMULAITIS, G., VENCLOVAS, Č. & SIKSNYS, V. 2017. Type III CRISPR-Cas Immunity: Major Differences
593 Brushed Aside. *Trends in Microbiology,* 25**,** 49-61.
594 TEVELL, S., BAIG, S., HELLMARK, B., MARTINS SIMOES, P., WIRTH, T., BUTIN, M., NILSDOTTER-
595 AUGUSTINSSON, Å., SÖDERQUIST, B. & STEGGER, M. 2020. Presence of the neonatal Staphylococcus
596 capitis outbreak clone (NRCS-A) in prosthetic joint infections. *Scientific Reports,* 10**,** 22389.
597 WALKER, B., ABEEL, T., SHEA, T., PRIEST, M., ABOUELLIEL, A., SAKTHIKUMAR, S., CUOMO, C., ZENG, Q.,
598 WORTMAN, J., YOUNG, S. & EARL, A. 2014. Pilon: An Integrated Tool for Comprehensive Microbial
599 Variant Detection and Genome Assembly Improvement. *PLOS ONE,* 9**,** e112963.
600 WESTBERG, R., STEGGER, M. & SÖDERQUIST, B. 2022. Molecular Epidemiology of Neonatal-Associated
601 *Staphylococcus haemolyticus* Reveals Endemic Outbreak. *Microbiology Spectrum***,** e02452-22.
602 WIRTH, T., BERGOT, M., RASIGADE, J.-P., PICHON, B., BARBIER, M., MARTINS-SIMOES, P., JACOB, L., PIKE, R.,
603 TISSIERES, P., PICAUD, J.-C., KEARNS, A., SUPPLY, P., BUTIN, M., LAURENT, F., ADAMKOVA, V.,
604 BARKHAM, T., BECKER, K., BENNETT, D., CLARIS, O., CREECH, C., DE LENCASTRE, H., DEIGHTON, M.,
605 DENIS, O., FERGUSON, J., HUANG, Y.-C., KLINGENBERG, C., INGEBRETSEN, A., LAFERRIÈRE, C., DOS
606 SANTOS, K., SCHRENZEL, J., SPILIOPOULOU, I., STEFANI, S., TAEKSOO, K., TARKKA, E., FRIEDRICH, A.,
607 VANDENBROUCKE-GRAULS, C., USSHER, J., VANDENESCH, F., WESTBLADE, L., LINDSAY, J.,
608 VANDENESCH, F., LARSEN, A., ZANGER, P., KAHL, B., AYMERICH, C., THE INTERNATIONAL
609 CONSORTIUM FOR STAPHYLOCOCCUS CAPITIS NEONATAL, S. & THE, E. S. G. O. E. 2020. Niche
610 specialization and spread of Staphylococcus capitis involved in neonatal sepsis. *Nature Microbiology,*
611 5**,** 735-745.

612

613 **Acknowledgements**

614 We would like to acknowledge the nurses within NICUs in Lubeck and Norwich who helped with isolation of
615 S. *capitis* from neonates and the microbiology lab staff at the NNUH and Lubeck hospital who also helped
616 process and save isolates.

617 This work was supported by an award from the Biotechnology and Biological Sciences Research Council
618 (BB/T014644/1). For the purpose of open access, the author has applied a Creative Commons Attribution (CC
619 BY) licence to any Author Accepted Manuscript version arising from this submission.

620

621 **Data accessibility**

622 Sequence data for all isolates have been uploaded to the SRA under Project number PRJNA751027

623 **Ethics**

624 The collection of weekly neonatal skin swabs in 2017-18 received prior approval as a surveillance study by
625 the Research and Development Manager of the Norfolk and Norwich University Hospitals NHS Foundation
626 Trust and did not require formal research ethics review. In Germany, the study was reviewed and approved
627 as a surveillance study by the University of Lubeck Hospital ethics committee (reference AZ 15-034,
628 amendment 01/2018). Clinical isolates were provided under NHS Research Ethics Committee approval to the
629 Norwich Biorepository which banks blood, solid tissue and bacterial isolates from the NNUH and research

institutes on the Norwich Research Park, including the University of East Anglia, and makes these available to the research community.

633 **Figure 1** Population structure of *S. capitis*. A maximum likelihood (ML) tree based on core gene alignment
634 from Roary. Three groups are indicated; Non NRCS-A isolates are highlighted in green, proto NRCS-A in
635 orange, and NRCS-A, in pink. Country of isolation is indicated by the colour of tip branches; Germany (grey
636 circles), UK (black circles). Body site of isolation are shown by coloured boxes; orange – neonatal skin,
637 brown - neontal gut, red – neonatal bloodstream, green – healthy adult skin, blue – adult prosthetic joint,
638 pink – adult bloodsteam. Antimicrobial resistance (AMR) genes are represented by red boxes (blank boxes
639 represent absence of the corresponding gene). The percentage identity of *nsr* and *tarJ* in each strain
640 (compared to the reference sequences present in CR01 and CR05 respectively using BLASTp) are shown as
641 pie charts. MIC data for chlorhexidine (CHX, purple bars) and octenidine (OCT, yellow bars) are shown with
642 the dotted lines representing 4 µg/mL of OCT or CHX.

643

644 **Figure 2.** Detailed phylogeny of NRCS-A based on core SNP alignment. The week of isolation is indicated by
645 purple squares for all UK isolates and sources of each isolate are indicated by coloured squares (neonatal
646 skin = orange; neonatal gut = brown, neonatal bloodstream = red).

647

648 **Figure 3** Insertion of the CRISPR-Cas Type III-A system in NRCS-A isolates, compared to the related Proto
649 NRCS-A isolates.

650

651 **Figure 4** Growth of *S. capitis* in different pH (pH7, 6 and 5.5), showing growth at final time point of 48 hours
652 from PM10 MicroPlate ™. Points indicate mean values and bars the standard error.

653

654 **Figure 5** Chlorhexidine susceptibility of the three groups of *S. capitis*. Points represent values from individual
655 strains, horizontal lines indicate the means with bars showing the standard error of the mean.

656 **Table 1** MIC^{50}- and MIC^{90} values of antimicrobials against the *S. capitis* groups.

	OCT		CHX		VAN		GEN		PEN		CEF		CIP		DAP		FUS	
	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}	MIC^{50}	MIC^{90}
Non NRCS-A	2	2	8	16	1	2	0.064	32	0.5	8	1	4	0.5	4	0.5	2	2	2
Proto NRCS-A	2	4	8	16	1	2	1	32	2	32	1	4	0.25	0.5	0.5	1	0.064	2
NRCS-A	2	4	16	32	1	2	16	32	8	126	4	4	0.5	4	0.25	1	4	4
Overall	2	4	16	16	1	2	8	32	4	126	4	4	0.5	4	0.5	1	2	4
UK	2	4	16	32	1	2	8	32	4	126	4	4	0.5	4	0.5	1	4	4
German	1	2	8	16	1	1	8	32	4	126	4	4	0.5	4	0.5	1	2	2

657

658 **Table 2** Selected genes where presence/absence was strongly statistically associated with the NRCS-A
659 isolates compared other *S. capitis* clades.

Gene Presence	Annotation	Sensitivity	Specificity	Adjusted P value
cas 1, 2, 6, 10	CRISPR-associated endonucleases (subtype III-A)	100.00	100.00	9.42E-34
csm 2, 3, 4, 5, 6	CRISPR system Cms proteins	100.00	100.00	9.42E-34
ynfM	Inner membrane transport protein YnfM	97.94	100.00	2.97E-31
betI	HTH-type transcriptional regulator BetI	94.85	95.35	6.54E-25
paaZ	Bifunctional protein PaaZ	98.97	86.05	4.83E-24
ugpQ	Glycerophosphodiester phosphodiesterase cytoplasmic	98.97	86.05	4.83E-24
mecA	PBP2a family beta-lactam-resistant peptidoglycan transpeptidase MecA	98.97	83.72	6.37E-23
repB	Replication protein RepB	86.60	100.00	1.56E-22
trpS2	Tryptophan--tRNA ligase 2	82.47	100.00	2.97E-20
mntH	Divalent metal cation transporter MntH	91.75	90.70	7.06E-20
bcr	Bicyclomycin resistance protein	80.41	100.00	3.12E-19
copA	Copper-exporting P-type ATPase	96.91	74.42	1.10E-16
mcrC	Protein McrC	100.00	62.79	6.70E-16
copZ	Copper chaperone CopZ	95.88	74.42	7.84E-16
csoR	Copper-sensing transcriptional repressor CsoR	95.88	74.42	7.84E-16
czcD	Cadmium cobalt and zinc/H(+)-K(+) antiporter	95.88	74.42	7.84E-16
tarIJL	Ribitol-5-phosphate cytidylyltransferase,Ribulose-5-phosphate reductase, Teichoic acid ribitol-phosphate polymerase	100.00	60.47	4.08E-15
xerD	Tyrosine recombinase XerD	100.00	60.47	4.08E-15
nixA	Nickel transporter NixA	100.00	58.14	2.47E-14
lacE	PTS system lactose-specific EIICB component	97.94	62.79	1.66E-13
aacA-aphD	Bifunctional AAC/APH	76.29	90.70	4.90E-12
hchA	Protein/nucleic acid deglycase HchA	95.88	62.79	8.59E-12
yhfK	putative sugar epimerase YhfK	95.88	62.79	8.59E-12
appA	Oligopeptide-binding protein AppA	100.00	48.84	2.23E-11
yfmJ	Putative NADP-dependent oxidoreductase YfmJ	96.91	55.81	2.06E-10
dppB	Dipeptide transport system permease protein DppB	98.97	48.84	3.71E-10
hsdM	Type I restriction enzyme EcoKI M protein	87.63	69.77	2.40E-09
qorB	Quinone oxidoreductase 2	96.91	46.51	8.53E-08
gltC	HTH-type transcriptional regulator GltC	100.00	32.56	8.45E-07
gtfA	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase GtfA subunit	100.00	32.56	8.45E-07
aes	Acetyl esterase	100.00	30.23	3.26E-06
cimH	Citrate/malate transporter	100.00	30.23	3.26E-06
cntABCDEF	Metal-staphylopine-binding proteins, permases and Staphylophine export/import proteins	100.00	30.23	3.26E-06
fabHB	3-oxoacyl-(acyl-carrier-protein) synthase 3 protein 2	100.00	30.23	3.26E-06
mleS	Malolactic enzyme	100.00	30.23	3.26E-06

mshD	Mycothiol acetyltransferase	100.00	27.91	8.12E-06
nmoA	Nitronate monooxygenase	100.00	27.91	8.12E-06
rl5	Polyphenol oxidase	100.00	27.91	8.12E-06
rutB	Peroxyureidoacrylate/ureidoacrylate amidohydrolase RutB	100.00	27.91	8.12E-06
cntKLM	Histidine racemase and Staphylopine synthase	100.00	25.58	3.10E-05
dppC	Dipeptide transport system permease protein DppC	87.63	53.49	3.22E-05
copB	Copper-exporting P-type ATPase B	91.75	41.86	0.0003
tarF	hypothetical protein	98.97	25.58	0.0003
gtfB	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase stabilizing protein GtfB	92.78	39.53	0.0003
kdpD	Sensor protein KdpD	100.00	20.93	0.0004
cspA	Cold shock protein CspA	97.94	25.58	0.0012
efeM	putative iron uptake system component EfeM	100.00	16.28	0.0052
paiA	Spermidine/spermine N(1)-acetyltransferase	100.00	16.28	0.0052
hsdR	Type-1 restriction enzyme R protein	98.97	18.60	0.0094
dtd	D-aminoacyl-tRNA deacylase	100.00	13.95	0.0175
npr	NADH peroxidase	100.00	13.95	0.0175
clpB	Chaperone protein ClpB	93.81	27.91	0.0224

Gene Absence	**Annotation**	**Sensitivity**	**Specificity**	**Adjusted P value**
aaxC	Arginine/agmatine antiporter	0.00	62.79	4.85E-08
hdcA	Histidine decarboxylase proenzyme	0.00	62.79	4.85E-08
metNPQ	Methionine import ATP-binding protein, permease and putative D-methionine-binding lipoprotein	0.00	62.79	4.85E-08
gltR	HTH-type transcriptional regulator GltR	0.00	65.12	2.05E-07
dapL	LL-diaminopimelate aminotransferase	0.00	67.44	8.45E-07
pyrB	Aspartate carbamoyltransferase catalytic subunit	0.00	69.77	3.26E-06
argR	Arginine repressor	2.06	62.79	3.81E-06
pap	Polyphosphate:AMP phosphotransferase	2.06	62.79	3.81E-06
arsD	Arsenical resistance operon trans-acting repressor ArsD	0.00	72.09	8.12E-06
panS	Pantothenate precursors transporter PanS	0.00	72.09	8.12E-06
pfbA	Plasmin and fibronectin-binding protein A	0.00	72.09	8.12E-06
pncA	Nicotinamidase/pyrazinamidase	0.00	72.09	8.12E-06
wecD	dTDP-fucosamine acetyltransferase	0.00	72.09	8.12E-06
yclM	Aspartokinase 3	0.00	72.09	8.12E-06
arcA	Arginine deiminase	3.09	62.79	1.15E-05
arcR	HTH-type transcriptional regulator ArcR	0.00	74.42	3.10E-05
malP	PTS system maltose specific EIICB component	0.00	74.42	3.10E-05
tarM	Poly(ribitol-phosphate) alpha-N-acetylglucosaminyltransferase	0.00	74.42	3.10E-05
ydfJ	Membrane protein YdfJ	0.00	74.42	3.10E-05
arcC1	Carbamate kinase 1	4.12	62.79	4.28E-05
arsR	Arsenical resistance operon repressor	1.03	72.09	7.40E-05
nikE	Nickel import system ATP-binding protein NikE	0.00	76.74	0.0001
pnoA	Nitronate monooxygenase	0.00	76.74	0.0001
arcD	Arginine/ornithine antiporter	5.15	62.79	0.0001

relJ	Antitoxin RelJ	3.09	67.44	0.0001
arsA	Arsenical pump-driving ATPase	11.34	51.16	0.0001
kimA	Potassium transporter KimA	1.03	74.42	0.0003
argF	Ornithine carbamoyltransferase	6.19	62.79	0.0004
yoeB	Toxin YoeB	2.06	72.09	0.0004
dppE	Dipeptide-binding protein DppE	0.00	79.07	0.0004
bcrA	Bacitracin transport ATP-binding protein BcrA	4.12	67.44	0.0004
ktrB	Ktr system potassium uptake protein B	7.22	62.79	0.0009
merB	Alkylmercury lyase	0.00	81.40	0.0015
oppC	Oligopeptide transport system permease protein OppC	0.00	81.40	0.0015
sraP	Serine-rich adhesin for platelets	0.00	81.40	0.0015

Sensitivity and specificity of linkage with phenotypes calculated by Scoary based on Roary results. Adjusted P value determined using Benjamini-Hochberg's method (Brynildsrud et al., 2016).

664 **Table 3** Selected genes where presence was strongly statistically associated with the NRCS-A isolates
665 compared to the proto-NRCS-A isolates.

Gene	Annotation	Sensitivity	Specificity	P value
cas1, 2, 6, 10	CRISPR-associated endonucleases (type III-A)	100.00	100.00	1.81E-16
csm2,3,4,5,6	CRISPR system Cms proteins	100.00	100.00	1.81E-16
bceA	Bacitracin export ATP-binding protein BceA	100.00	93.75	6.26E-15
ynfM	Inner membrane transport protein YnfM	97.94	100.00	6.26E-15
smc	Chromosome partition protein Smc	96.91	100.00	3.07E-14
npr	NADH peroxidase	95.88	100.00	1.27E-13
lytO	putative autolysin LytO	95.88	100.00	1.27E-13
xerD	Tyrosine recombinase XerD	96.91	93.75	1.60E-12
araB	Ribulokinase	92.78	100.00	4.57E-12
cmoM	tRNA 5-carboxymethoxyuridine methyltransferase	90.72	100.00	3.45E-11
lexA	LexA repressor	90.72	100.00	3.45E-11
dut	Deoxyuridine 5'-triphosphate nucleotidohydrolase	88.66	100.00	1.87E-10
betI	HTH-type transcriptional regulator BetI	94.85	87.50	7.21E-10
repB	Replication protein RepB	86.60	100.00	9.40E-10
trpS2	Tryptophan--tRNA ligase 2	82.47	100.00	1.52E-08
mntH	Divalent metal cation transporter MntH	91.75	87.50	1.71E-08
hisC	Histidinol-phosphate aminotransferase	81.44	100.00	2.71E-08
bcr	Bicyclomycin resistance protein	80.41	100.00	4.93E-08
blaR1-2	beta-lactam sensor/signal transducer BlaR1	95.88	75.00	7.08E-08
hin	DNA-invertase hin	95.88	75.00	7.08E-08
ldh1	L-lactate dehydrogenase 1	78.35	100.00	1.41E-07
mecA	PBP2a family beta-lactam-resistant peptidoglycan transpeptidase MecA	98.97	62.50	1.59E-07
paaZ	Bifunctional protein PaaZ	98.97	62.50	1.59E-07
ugpQ	Glycerophosphodiester phosphodiesterase cytoplasmic	98.97	62.50	1.59E-07
blaI	Penicillinase repressor	93.81	75.00	6.17E-07
rhaS	HTH-type transcriptional activator RhaS	79.38	93.75	2.38E-06
blaZ	Beta-lactamase	84.54	87.50	2.55E-06
copA	Copper-exporting P-type ATPase	96.91	62.50	3.14E-06
copZ	Copper chaperone CopZ	95.88	62.50	9.94E-06
dtd	D-aminoacyl-tRNA deacylase	83.51	81.25	5.74E-05
menH	2-succinyl-6-hydroxy-2 4-cyclohexadiene-1-carboxylate synthase	93.81	62.50	6.91E-05
ydhP	Inner membrane transport protein YdhP	90.72	68.75	7.62E-05
csoR	Copper-sensing transcriptional repressor CsoR	95.88	56.25	9.74E-05
czcD	Cadmium cobalt and zinc/H(+)-K(+) antiporter	95.88	56.25	9.74E-05

666 Sensitivity and specificity of linkage with phenotypes calculated by scoary based on Roary results. Adjusted P
667 value (<0.001) determined using Benjamini-Hochberg's method (Brynildsrud et al., 2016).

668 **Table 4** Selected genes where presence was statistically associated with antibiotics and chlorhexidine
669 susceptibility.

		Adjusted P value		
Gene	**Annotation**	**CHX 16 µg/mL**	**PEN 1 µg/mL**	**CEF 4 µg/mL**
aacA-aphD	Bifunctional AAC/APH	0.0001		
araB	Ribulokinase	2.47E-06		
bcr	Bicyclomycin resistance protein	7.49E-06		
betI	HTH-type transcriptional regulator BetI	1.57E-06		
bin3	Putative transposon Tn552 DNA-invertase bin3	2.79E-06	0.002	0.0126
blaI	Penicillinase repressor	3.50E-05		
blaR1-2	beta-lactam sensor/signal transducer BlaR1	0.000414331		
cas1,2,6,10	CRISPR-associated endonuclease Cas1	1.25E-07	0.002	0.0126
cls	Cardiolipin synthase	7.97E-05		
copA	Copper-exporting P-type ATPase	3.44E-06	0.002	0.0153
copZ	Copper chaperone CopZ	3.44E-06	0.002	0.0153
csm2,3,4,5,6	CRISPR system Cms protein Csm2	1.25E-07	0.002	0.0126
csoR	Copper-sensing transcriptional repressor CsoR	3.44E-06		
czcD	Cadmium cobalt and zinc/H(+)-K(+) antiporter	1.35E-06	0.003	
dut	Deoxyuridine 5'-triphosphate nucleotidohydrolase	7.83E-07	0.002	
hin	DNA-invertase hin	2.79E-06	0.002	0.0126
hisC	Histidinol-phosphate aminotransferase	7.49E-06		
ldh1	L-lactate dehydrogenase 1	3.06E-05		
lexA	LexA repressor	2.47E-06	0.002	
lytO	putative autolysin LytO	4.90E-05		
mboIR	hypothetical protein	1.25E-07		
mecA	PBP2a family beta-lactam-resistant peptidoglycan transpeptidase	1.20E-06		0.0126
mntH	Divalent metal cation transporter MntH	2.53E-06		0.0126
npr	NADH peroxidase	5.22E-06		
paaZ	Bifunctional protein PaaZ	3.34E-07		0.0126
repB	Replication protein RepB	0.0001		
rhaS	HTH-type transcriptional activator RhaS	1.02E-05		
rihC	Non-specific ribonucleoside hydrolase RihC		0.005	
smc	Chromosome partition protein Smc	1.25E-07	0.002	0.0126
trpS2	Tryptophan--tRNA ligase 2	2.79E-06		
ugpQ	Glycerophosphodiester phosphodiesterase cytoplasmic	3.34E-07		0.0126
xerC	Tyrosine recombinase XerC	1.25E-07	0.002	
xerD	Tyrosine recombinase XerD	1.24E-06		
ynfM	Inner membrane transport protein YnfM	1.25E-07	0.002	0.0126

670 A cut off MIC for chlorhexidine (CHX), penicillin (PEN) and cefotaxime (CEF) were used to distinguish
671 between strains with high and low tolerance (shown in Table 1). Adjusted P value determined using
672 Benjamini-Hochberg's method (Brynildsrud et al., 2016).

673 **Supplementary material**

674 **Supplementary Data 1** List of reference genomes of *S. capitis* strains used.

675 **Supplementary Data 2** Meta data for *Staphylococcus capitis* isolates. MICs OCT (octenidine range 1 – 32
676 µg/mL), CHX (chlorhexidine range 4 - 64 µg/mL), VAN (vancomycin range 0.25 - 4 µg/mL), GEN (gentamicin
677 range 0.016- 16 µg/mL), PEN (benzylpenicillin range 0.016 - 128 µg/mL), CEF (cefotaxime range 0.125 - 2
678 µg/mL), CIP (ciprofloxacin range 0.064 - 2µg/mL), DAP (daptomycin range 0.25 - 2 µg/mL), FUS (fusidic acid
679 range 0.064 -2 µg/mL). X indicates MIC data was not obtained.

680 **Supplementary Data 3** ABRICATE summary in Binary format, where 1 = ≥ 90% gene identity

681 **Supplementary Data 4** Snippy-core SNP distance matrix on all *S. capitis* collection using NRCS-A CR05 as the
682 reference genome.

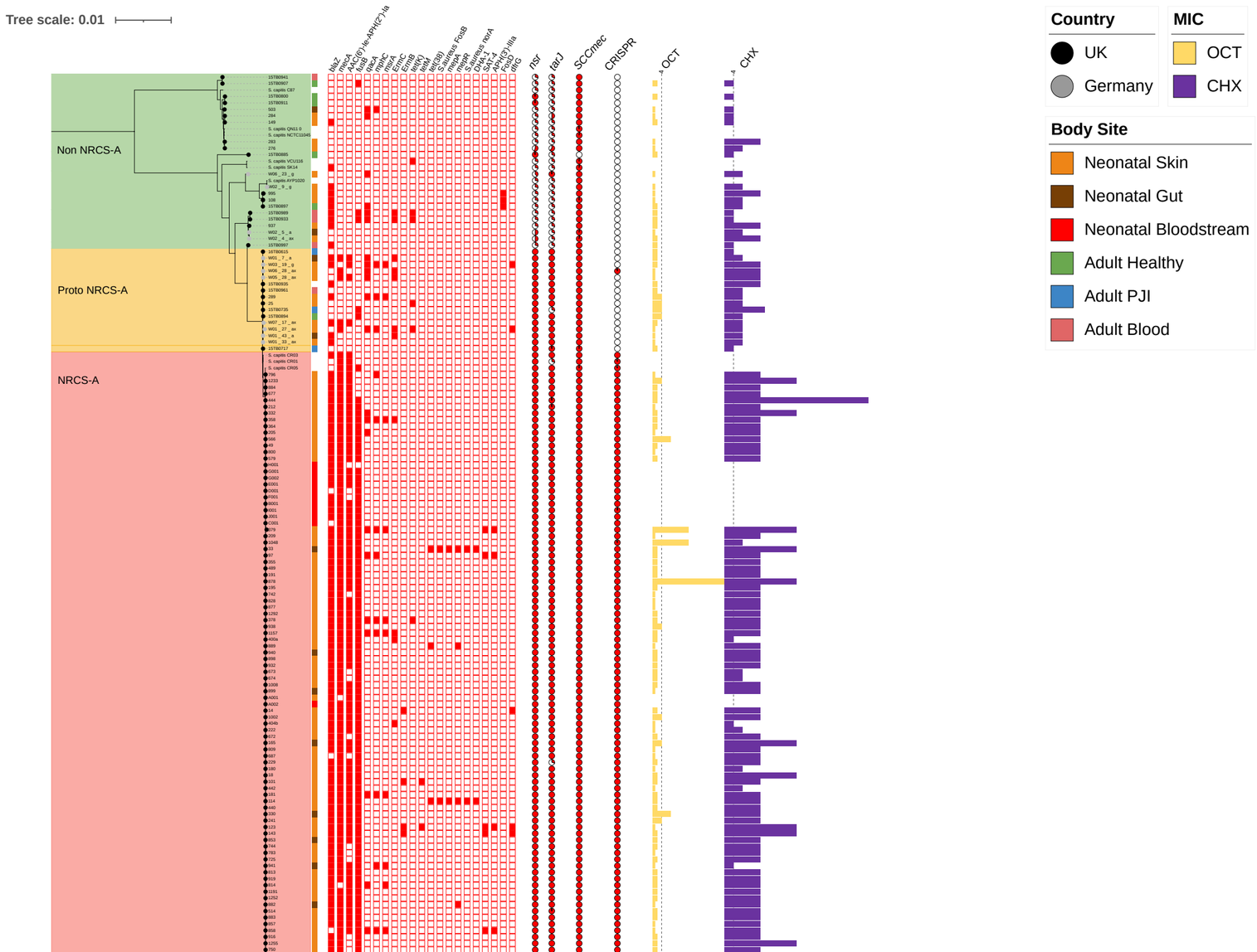
683 **Supplementary Data 5** Snippy-core SNP distance matrix on NRCS-A isolates using CR05 as the reference
684 genome.

685 **Supplementary Data 6** Scoary report on NRCS-A traits compared to proto NRCS-A, Country (UK), cefoxitin ≥
686 4 µg/mL (CEF4), chlorhexidine ≥ 16 µg/mL (CHX 16), ciprofloxacin ≥ 2 µg/mL (CIP2), fusidic acid ≥ 2 µg/mL
687 (FUS2), gentamicin ≥ 8 µg/mL

688

689

690



Country

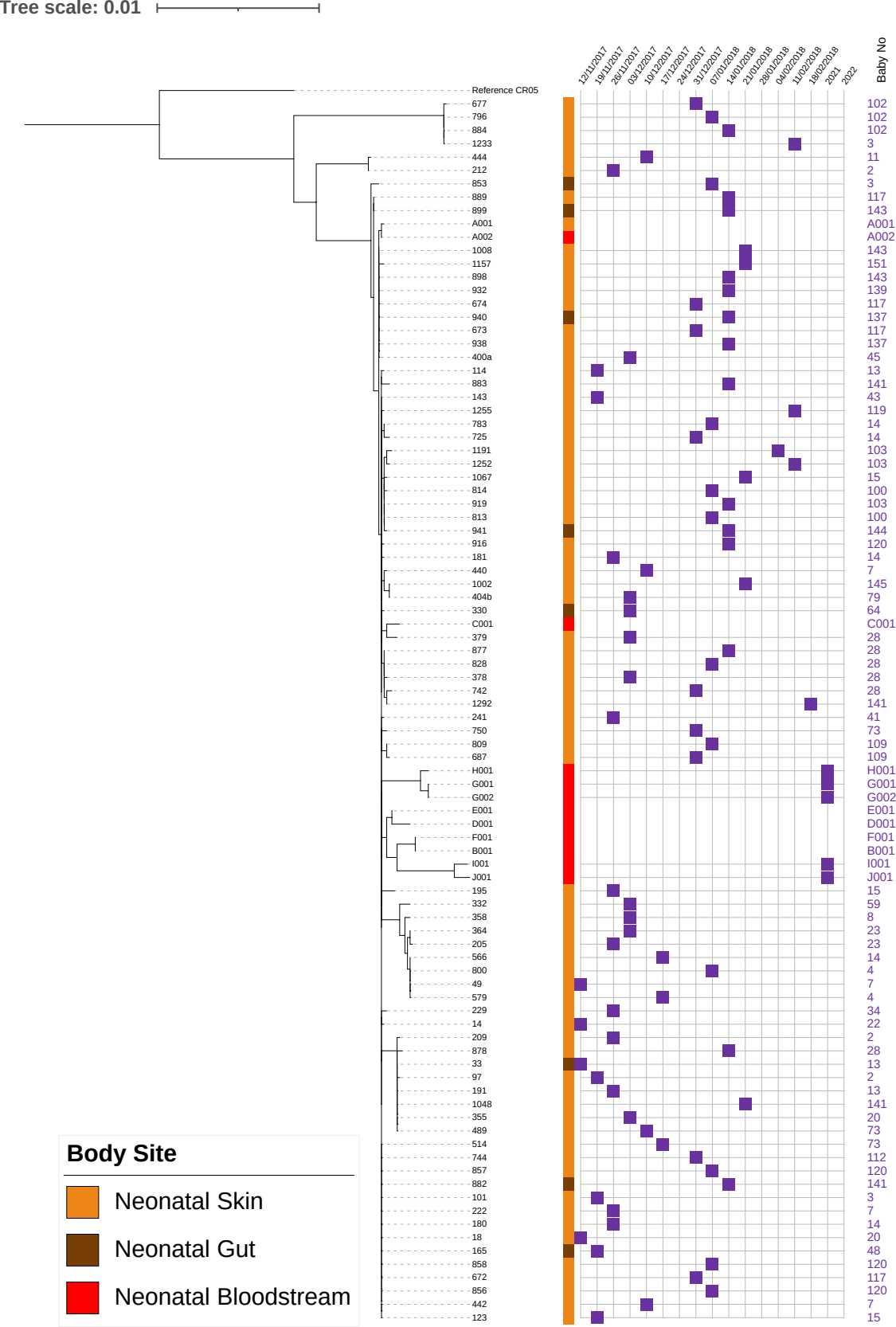
- UK
- Germany

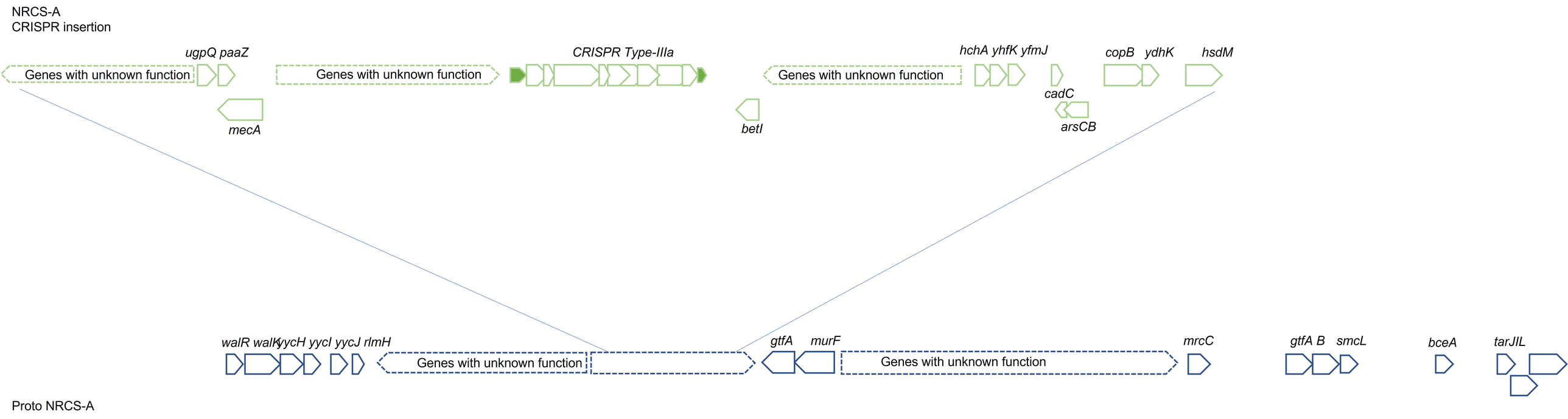
MIC

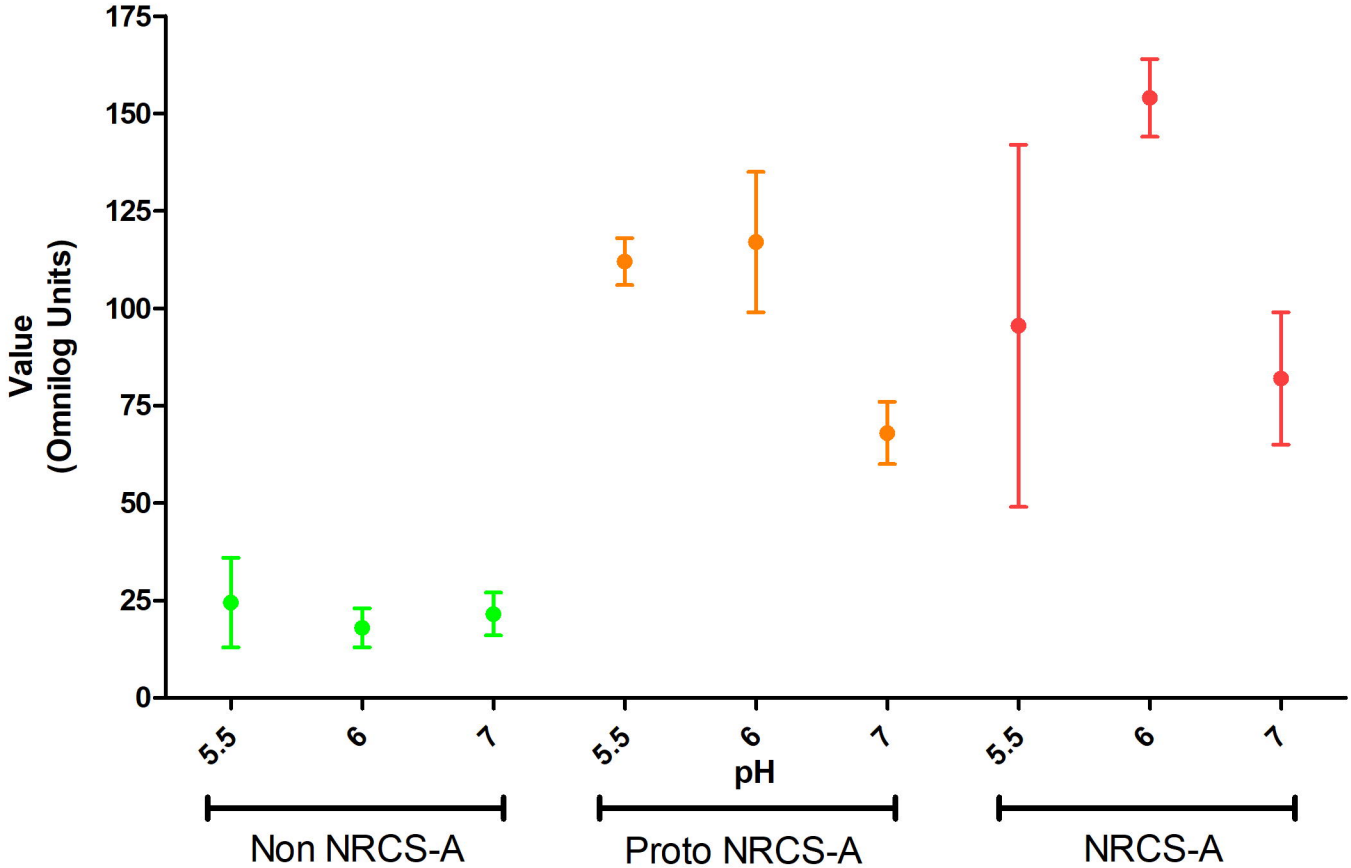
- OCT
- CHX

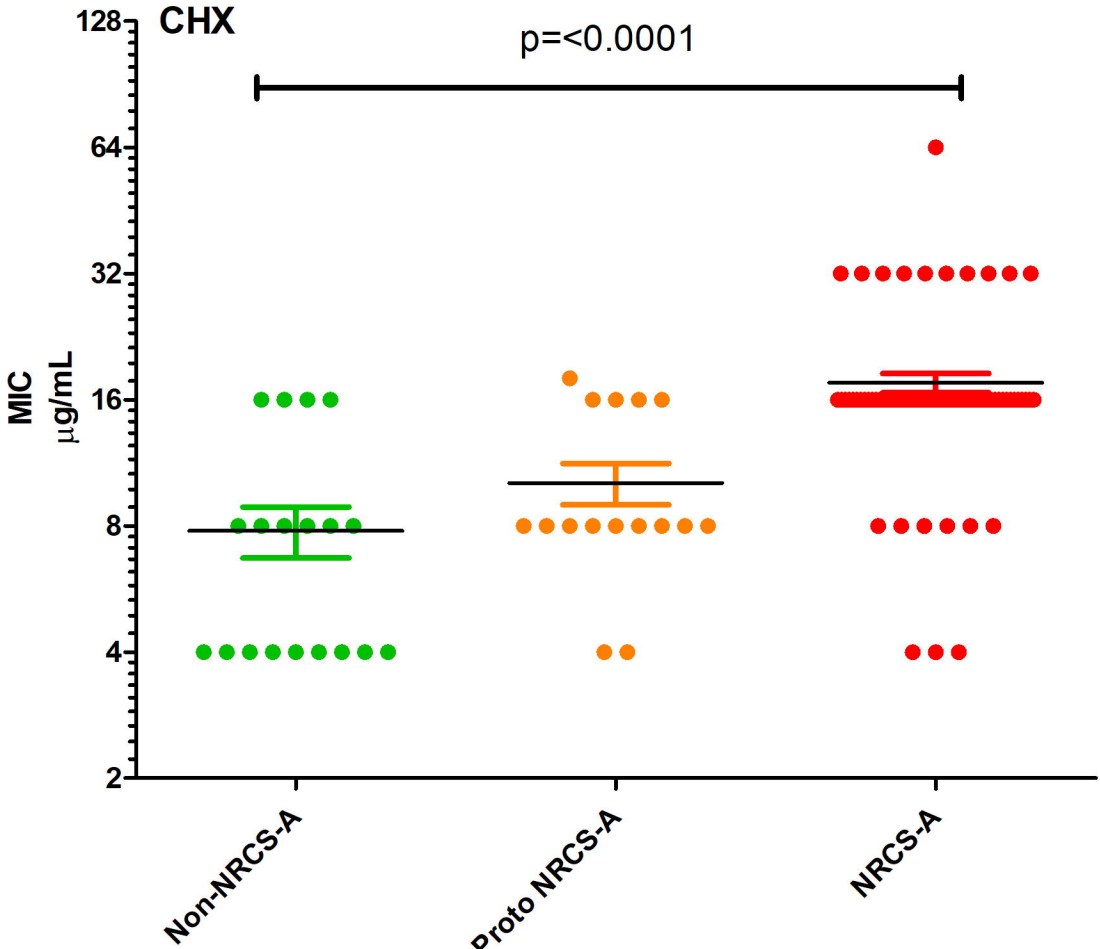
Body Site

- Neonatal Skin
- Neonatal Gut
- Neonatal Bloodstream
- Adult Healthy
- Adult PJI
- Adult Blood


Tree scale: 0.01
12/11/2017
19/11/2017
26/11/2017
03/12/2017
10/12/2017
17/12/2017
24/12/2017
31/12/2017
07/01/2018
14/01/2018
21/01/2018
28/01/2018
04/02/2018
11/02/2018
18/02/2018
2021
2022
Baby No
Reference CR05
677 102
796 102
884 102
1233 3
444 11
212 2
853 3
889 117
899 143
A001 A001
A002 A002
1008 143
1157 151
898 143
932 139
674 117
940 137
673 117
938 137
400a 45
114 13
883 141
143 43
1255 119
783 14
725 14
1191 103
1252 103
1067 15
814 100
919 103
813 100
941 144
916 120
181 14
440 7
1002 145
404b 79
330 64
C001 C001
379 28
877 28
828 28
378 28
742 28
1292 141
241 41
750 73
809 109
687 109
H001 H001
G001 G001
G002 G002
E001 E001
D001 D001
F001 F001
B001 B001
I001 I001
J001 J001
195 15
332 59
358 8
364 23
205 23
566 14
800 4
49 7
579 4
229 34
14 22
209 2
878 28
33 13
97 2
191 13
1048 141
355 20
489 73
514 73
744 112
857 120
882 141
101 3
222 7
180 14
18 20
165 48
858 120
672 117
856 120
442 7
123 15
Body Site
Neonatal Skin
Neonatal Gut
Neonatal Bloodstream


NRCS-A
CRISPR insertion
ugpQ paaZ
Genes with unknown function
Genes with unknown function
CRISPR Type-IIIa
Genes with unknown function
hchA yhfK yfmJ
copB ydhK
hsdM
cadC
mecA
betI
arsCB
walR walK yycH yycI yycJ rlmH
Genes with unknown function
gtfA murF
Genes with unknown function
mrcC
gtfA B smcL
bceA
tarJIL
Proto NRCS-A


175
150
125
100
75
50
25
0
Value
(Omnilog Units)
5.5
6
7
5.5
6
7
5.5
6
7
pH
Non NRCS-A
Proto NRCS-A
NRCS-A


CHX
p=<0.0001
MIC
µg/mL
128
64
32
16
8
4
2
Non-NRCS-A
Proto NRCS-A
NRCS-A