Authors' Accepted Manuscript

1   **An emerging Panton-Valentine leukocidin (PVL)-positive CC5-meticillin-resistant**

2   ***Staphylococcus aureus*-IVc clone recovered from hospital and community settings over a**

3   **17-year period from 12 countries investigated by whole-genome sequencing**

4

5   Bisola K. Aloba[a], Peter M. Kinnevey[a], Stefan Monecke[b,c,d], Gráinne I. Brennan[e], Brian

6   O'Connell[f], Anita Blomfeldt[g], Brenda A. McManus[a], Wulf Schneider-Brachert[h], Jan Tkadlec[i],

7   Ralf Ehricht[b,d,j], Abiola Senok[k], Mette Damkjær Bartels[l,m], David C. Coleman[a],*

8
9   *[a]Microbiology Research Unit, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin,*
10 *Dublin, Ireland*
11
12 *[b]Leibniz Institute of Photonic Technology (IPHT), Jena, Germany*
13
14 *[c]Institut für Medizinische Mikrobiologie und Virologie, Uniklinikum Dresden*
15 *Fiedlerstrasse 42, D-01307 Dresden, Germany*
16
17 *[d]InfectoGnostics Research Campus, Jena, Germany*
18
19 *[e]National MRSA Reference Laboratory, St. James's Hospital, Dublin, Ireland*
20
21 *[f]Department of Clinical Microbiology, St. James's Hospital, Dublin, Ireland*
22
23 *[g]Department of Microbiology and Infection Control, Akershus University Hospital, Lørenskog, Norway*
24
25 *[h]Department of Infection Prevention and Infectious Diseases, University Hospital Regensburg, Regensburg,*
26 *Germany*
27
28 *[i]Department of Medical Microbiology, Charles University, 2nd Faculty of Medicine and Motol University*
29 *Hospital, Prague, Czech Republic*
30
31 *[j]Institute of Physical Chemistry, Friedrich-Schiller University, Jena, Germany*
32
33 *[k]College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab*
34 *Emirates*
35
36 *[l]Department of Clinical Microbiology, Amager and Hvidovre Hospital, Hvidovre, Denmark*
37
38 *[m]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark*

39
40 *Running Title:* An emerging PVL-positive CC5-MRSA-IVc clone

41

42 *Corresponding author. Address: Microbiology Research Unit, Division of Oral Biosciences,

43 Dublin Dental University Hospital, University of Dublin, Trinity College, Lincoln Place,

44 Dublin D02 F859, Ireland. Tel.: +353 1 6127276; fax: + 353 1 6127295.

45 *E-mail address:* david.coleman@dental.tcd.ie (D.C. Coleman).

46

47 Abbreviations: CC; clonal complex, ST; sequence type; WGS, whole-genome sequencing;

48 cgMLST, core-genome multilocus sequence typing; cgSNP, core genome single-nucleotide

49 polymorphism; RIGs, related isolate groups.

1

50 **ABSTRACT**

51

52 ***Background:*** A novel Panton-Valentine leukocidin (PVL)-positive meticillin-resistant
53 *Staphylococcus aureus* (MRSA) clonal complex (CC)5-MRSA-IVc ('Sri Lankan' clone) was
54 recently described from Sri Lanka. Similar isolates caused a recent Irish hospital outbreak.

55

56 ***Aims:*** To investigate the international dissemination and diversity of PVL-positive CC5-
57 MRSA-IVc isolates from hospital and community settings using whole-genome sequencing
58 (WGS).

59

60 ***Methods:*** Core-genome single nucleotide polymorphism (cgSNP) analysis, multilocus
61 sequence typing (cgMLST) and microarray-based detection of antimicrobial-resistance and
62 virulence genes were used to investigate PVL-positive CC5-MRSA-IVc (*N*=214 including 46
63 'Sri Lankan' clone) from hospital and community settings in 12 countries over 17 years.
64 Comparators included 29 PVL-positive and 23 PVL-negative CC5/ST5-MRSA-
65 I/II/IVa/IVc/IVg/V.

66 ***Findings:*** Maximum-likelihood cgSNP analysis grouped 209/214 (97.7%) CC5-MRSA-IVc
67 into Clade I; average of 110 cgSNPs between isolates. Clade III contained the five remaining
68 CC5-MRSA-IVc; average of 92 cgSNPs between isolates. Clade II contained seven PVL-
69 positive CC5-MRSA-IVa comparators, whereas the remaining 45 comparators formed an
70 outlier group. Minimum-spanning cgMLST analysis revealed a comparably low average of 57
71 allelic differences between all CC5/ST5-MRSA-IVc. All 214 CC5/ST5-MRSA-IVc were
72 identified as 'Sri Lankan' clone, predominantly *spa* type t002 (186/214) with low population
73 diversity and harboured a similar range of virulence genes and variable antimicrobial-
74 resistance genes. All 214 Sri Lankan clone isolates and Clade II comparators harboured a
75 9,616 bp chromosomal PVL-encoding phage remnant, suggesting both arose from a PVL-
76 positive meticillin-susceptible ancestor. Most Sri Lankan clone isolates were from infections
77 (142/214), and where detailed metadata were available (168/214), most were community-
78 associated (85/168).

79 ***Conclusions:*** Stable chromosomal retention of *pvl* may facilitate Sri-Lankan clone
80 dissemination.

81
82 **Keywords:** CC5-MRSA-IVc, dissemination, epidemiology, phylogenomics, PVL, Sri Lankan
83 clone
84

Introduction

Meticillin-resistant *Staphylococcus aureus* (MRSA) contributes significantly to the prevalence of infectious diseases worldwide. Expression of Panton-Valentine leukocidin (PVL), a bicomponent beta-barrel toxin that causes leukocyte lysis or apoptosis via pore formation[1], has been associated with increased MRSA virulence and transmissibility[2-4]. PVL is encoded by the *lukF-PV* and *lukS-PV* genes (known as *pvl*) harboured by lysogenic converting bacteriophages[5]. Carriage of *pvl* was traditionally associated with community-associated (CA)-MRSA, frequently responsible for skin and soft tissue infections (SSTIs), however CA-MRSA lineages are increasingly associated with hospital infection and outbreaks[6-9]. In some cases, highly virulent CA-MRSA with increased transmissibility and greater clonal diversity have surpassed healthcare-associated (HA)-MRSA as the dominant hospital MRSA lineages[7,10,11]. Over the last decade, PVL-positive MRSA clones causing mainly superficial SSTIs have emerged in Irish healthcare and community settings[12,13]. In 2020, the Irish National MRSA Reference Laboratory (NMRSARL) identified *pvl* in 25% of all non-bloodstream infection MRSA submitted for investigation, up from 20% in 2017 and attributed this to an increase in HA-outbreaks[14,15].

McTavish et al. recently described a dominant PVL-positive clonal complex (CC) 5 MRSA lineage harbouring a type IVc staphylococcal cassette chromosome *mec* (SCC*mec*) element in Sri Lanka, and also identified it in the United Kingdom and Australia, referred to hereafter as the 'Sri-Lankan clone'[16]. This lineage was also recently reported in the United Arab Emirates[17]. The emergence of a PVL-positive CC5-MRSA-IVc lineage in Irish hospitals was reported in 2021[13]. Multiple suspected PVL-positive CA-MRSA outbreaks were investigated and identified a PVL-positive CC5-MRSA-IVc lineage as responsible for a 15-month maternity unit outbreak involving 13 patients. The widespread dissemination of novel MRSA lineages with subsequent replacement of predominant clonal types is not uncommon and has been described previously in Ireland and internationally[18-21].

To date, only localised investigations focused on country-specific genomic characterisation of PVL-positive ST5-MRSA-IVc isolates, such as the Sri-Lankan clone, have been reported[13,16,17]. This study sought to further investigate the PVL-positive CC5-MRSA-IVc population from Irish hospital and community settings in comparison to a comprehensive collection of similar isolates from 12 countries spanning 17 years using whole-genome sequencing (WGS). WGS provides unrivalled sensitivity and precision for comparing and monitoring the development and spread of historical, current, and emerging clones, as well as tracing infection outbreaks with extremely high resolution.

120 **Methods**

121

122 *MRSA isolates*

123   MRSA isolates (N=266) recovered between 2003–2022 were investigated: (i) 214

124 PVL-positive CC5/sequence type (ST)5-MRSA-IVc isolates (2005–2022) from 12 countries

125 similar to and including 46 previously described Sri-Lankan clone isolates[16] and (ii) 52

126 comparator CC5/ST5-MRSA-SCC-I/II/IVa/IVc/IVg/V isolates (29 PVL-positive and 23

127 PVL-negative) recovered between 2003–2021. Isolates were cryogenically stored at -80°C.

128 Detailed isolate information and available metadata is provided in Tables I and SI.

129

130 *Irish MRSA*

131   All 47 Irish MRSA isolates investigated were submitted to the NMRSARL between

132 2013–2022. These included 14 previously described PVL-positive ST5/t002-MRSA-IVc

133 maternity unit outbreak isolates recovered between 2018–2020 that were similar to the Sri-

134 Lankan clone and had a median of 3 (average: 6; range: 0–27) core-genome multi-locus

135 sequence type (cgMLST) allelic differences between isolates[13]. The emergence of this

136 clonal type in Irish hospitals appeared to be recent but a search of the NMRSARL collection

137 for PVL-positive t002-MRSA-IVc and related *spa* types revealed 16 additional isolates

138 recovered between 2013-2022 (Tables I and SI). These included two more recent (2021)

139 patient isolates from the maternity unit and 14 patient isolates from nine other hospitals.

140 Seventeen comparator NMRSARL isolates (two PVL-positive ST5-MRSA-IVa and 15 PVL-

141 negative ST5-MRSA-I/IVc/IVg/V) recovered between 2013–2019 were also investigated

142 (Tables I and SI).

143

144 *International MRSA*

145   Additional PVL-positive ST5-MRSA-IVc isolates or WGS datasets from disparate

146 geographical locations were sought for comparison. Contact with international collaborators,

147 an extensive search of the National Center for Biotechnology Information (NCBI) Sequence

148 Read Archive (SRA)/GenBank, the European Nucleotide Archive (ENA) databases and a

149 literature search (Table SII) yielded 219 international isolates or WGS datasets (80 clinical

150 isolates and 139 WGS sequences; Tables I and SI). Of these, 184 were PVL-positive ST5-

151 MRSA-IVc isolates similar to the Sri-Lankan clone, whereas 35 were PVL-positive (N=27)

152 and PVL-negative (N=8) ST5-MRSA-II/IVa/IVc/V comparator isolates.

153

154 *(i) Clinical isolates*

155   Fifty-six PVL-positive ST5-MRSA-IVc isolates recovered between 2005–2021 in the
156   Czech Republic (*N*=6), Germany (*N*=4), Kuwait (*N*=1), Norway (*N*=24), Saudi Arabia
157   (*N*=4), Sweden (*N*=2) and the United Arab Emirates (UAE) (*N*=15) underwent WGS at the
158   Dublin Dental University Hospital Microbiology Research Unit (Ireland) (Tables I and SI).
159   Twenty-four ST5-MRSA-II/IVa/IVc/V comparator isolates (19 PVL-positive and five PVL-
160   negative) recovered in Algeria (*N*=1), the Czech Republic (*N*=2), Germany (*N*=1), Kuwait
161   (*N*=1), Norway (*N*=12), Saudi Arabia (*N*=1), Senegal (*N*=2), Slovakia (*N*=1) and the UAE
162   (*N*=3) between 2003–2021 were also sequenced (Tables I and SI).

163

164   *(ii) Whole-genome sequences*
165   WGS datasets for 46 previously described PVL-positive ST5-MRSA-IVc Sri-Lankan
166   clone isolates were downloaded from ENA (accession number PRJEB27049)[16]. These
167   patient isolates were recovered in a Sri-Lankan hospital over four months in 2014 (*N*=33), the
168   United Kingdom between 2005–2015 (*N*=12) and Australia in 2015 (*N*=1) (Table I). WGS
169   datasets for PVL-positive ST5-MRSA-IVc isolates from Denmark (2007–2021) (*N*=66) and
170   Germany (2011–2019) (*N*=16) were received. Comparator ST5-MRSA-IVa/IVc/V WGS
171   datasets from Denmark (seven PVL-positive and three PVL-negative; 2013–2015) and
172   Germany (one PVL-positive; 2017) were also included (Tables I and SI).

173

174   *Genomic DNA extraction and whole-genome sequencing*
175   For short-read sequencing, genomic DNA was extracted and sequencing libraries
176   prepared using the Illumina® DNA Prep Kit (Illumina, Eindhoven, The Netherlands) as
177   described previously[19]. Libraries were scaled to exhibit ≥50× coverage and sequenced
178   using a 600-cycle MiSeq paired-end Reagent Kit v3 (Illumina) on an Illumina MiSeq
179   sequencer according to the manufacturer's instructions. Short- and long-read datasets for
180   isolates sequenced in Dublin were submitted to the NCBI SRA database under BioProject
181   Nos. PRJNA896922 and PRJNA638834). Short-read datasets for Danish isolates were
182   submitted to the NCBI SRA database under BioProject Nos. PRJNA839593, PRJNA865897,
183   PRJNA869909 and PRJNA898141.

184   For long-read sequencing, genomic DNA extractions and library preparations were
185   performed as described previously[22]. Sequencing was performed on the MinION platform
186   using a R9.4.1 Flow Cell with the MinKNOW software v20.10 (Oxford Nanopore
187   Technologies, United Kingdom) as per manufacturer's instructions.

188   Hybrid assemblies were performed by genome scaffolding using paired-end short-read
189   and   long-read   sequences   using   the   Unicycler   v0.5.0   pipeline

190 (https://github.com/rrwick/Unicycler). Assembled genomes were annotated using the web-

191 based RAST v2.0 server (https://rast.nmpdr.org) and visualised using Bandage v0.8.1

192 (https://rrwick.github.io/Bandage/) and SnapGene v6.0.6 (GSL Biotech LLC;

193 https://www.snapgene.com).

194

195 *Whole-genome sequence analysis*

196 Short-read FASTQ files were assembled, quality assessed and analysed using

197 BioNumerics software (BioNumerics v8.0; Applied Maths, Sint-Martens-Latem, Belgium),

198 Ridom SeqSphere+ software v7.0.4 (Ridom GmbH, Münster, Germany) and web-based

199 SCC*mec*Finder tool (https://cge.cbs.dtu.dk/services/SCCmecFinder/) as described

200 previously[13].

201

202 *Molecular characterisation*

203 DNA microarray profiling was undertaken using the *S. aureus* Genotyping Kit 2.0

204 (Abbott [Alere Technologies GmbH], Jena, Germany) or WGS analysis. The DNA

205 microarray chip harbours 333 target sequences for approximately 170 antimicrobial-resistance

206 and virulence-associated genes and other genes and sequences that can assign *S. aureus* to

207 CCs and/or STs as described previously[23,24]. WGS-based DNA microarray profiling was

208 undertaken using *in silico* probes of the *S. aureus* Genotyping Kit 2.0. Probe sequences map

209 onto assembled genomes to predict DNA array hybridisation patterns[25] and these patterns

210 were compared to *in vitro* array results. Additional investigations into alleles of interest were

211 performed using the Clustal Omega multiple sequence alignment tool,

212 (https://www.ebi.ac.uk/Tools/msa/clustalo/) and NCBI BLAST search engine

213 (https://blast.ncbi.nlm.nih.gov/Blast.cgi). *S. aureus* immune evasion cluster (IEC) types were

214 assigned using microarray profiling. β-haemolysin converting bacteriophages encode

215 combinations of the IEC genes (*sea/sak/scn/sep/chp*), permitting isolates to be clustered into

216 one of eight IEC types based on the combination of genes carried[26].

217

218 *Phylogenetic analysis*

219 Relatedness between MRSA recovered over an extended time period (2003–2022) was

220 investigated using cgMLST. A minimum spanning tree (MST) based on 1,861 core-gene loci

221 was generated for all isolates using the SeqSphere+ (Ridom) cgMLST scheme as previously

222 described[27,28]. Core-genome alignment and variant calling based on single-nucleotide

223 polymorphisms (SNPs) was performed on all isolates mapped against a study-specific

224 reference genome from MRSA isolate 141087 (2005; earliest year of recovery for Sri-Lankan

clone isolates investigated), using Snippy v4.6.0 (https://github.com/tseemann/snippy).

Recombinant SNPs were removed using Gubbins v3.2.1 (https://github.com/nickjcroucher/gubbins) and a pairwise cgSNP distance matrix generated using snp-dists v3 (https://github.com/tseemann/snp-dists). A cgSNP-based maximum-likelihood tree (MLT) was constructed through IQ-TREE v2.2.0 (http://www.iqtree.org) using recommended IQ-TREE guidelines. The phylogenetic tree was visualised and annotated through Interactive Tree of Life v6.5.8 (https://itol.embl.de).

Results

MRSA

MRSA isolates (*N*=266) recovered between 2003–2022 were investigated. These included 214 PVL-positive CC5/ST5-MRSA-IVc isolates from 12 countries similar to and including 46 'Sri-Lankan clone' isolates[16] and 29 PVL-positive and 23 PVL-negative ST5-MRSA-I/II/IVa/IVc/IVg/V comparator isolates (Tables I and SI). The PVL-positive ST5-MRSA-IVc isolates belonged to eight closely related *spa* types, with t002 predominating (186/214; 86.9%) (Tables I and SI). Six closely related *spa* types were identified among the comparator isolates, half of which were t002 (26/52; 50%) (Tables I and SI).

Phylogenetic analysis of Sri-Lankan clone isolates

To investigate the population structure of Irish PVL-positive CC5/ST5-MRSA-IVc relative to international isolates, all isolates and comparators were subjected to WGS-based phylogenetic analyses. The construction of SNP-based MLT and cgMLST-based MST trees yielded comparable findings regarding isolate relatedness and clustering (Figures 1a and 1b).

(i) Core-genome SNP analysis

cgSNP analysis based on 12,245 SNPs showed that all 214 PVL-positive CC5/ST5-MRSA-IVc isolates exhibited a pairwise SNP-distance median of 107 (average: 116; range: 0–410) (Table SIII). The SNP-based MLT grouped the vast majority of isolates (209/214, 97.7%) including all 46 'Sri-Lankan clone' isolates[16] into one clade, termed Clade I (Figure 1a). Clade I isolates had a median of 106 (average: 110; range: 0–287) SNPs between isolates (Table SIII). The remaining five PVL-positive ST5-MRSA-IVc grouped into Clade III exhibiting a median of 86 (average: 92; range: 69–127) SNPs (Figure 1a). Clade III differed from Clade I by a median of 232 (average: 237; range: 159–410) SNPs. Most comparator isolates (44/52) formed an outgroup at the base of the MLT (Figure 1a). A single PVL-

260 positive CC5/ST5-MRSA-IVa comparator (isolate Z4294) branched out next to Clade III

261 (thick black branch in Figure 1a). The remaining seven comparators (all PVL-positive

262 ST5/t002-MRSA-IVa) grouped into Clade II, forming the closest neighbour to Clade I (Figure

263 1a). Clade II differed from Clades I and III by a median of 176 (average: 178; range: 123–

264 331) and 230 (average: 234; range: 207–277) SNPs, respectively (Table SIII). This tree

265 topology confirmed the identity of all 214 CC5/ST5-MRSA-IVc as 'Sri-Lankan clone'.

266 Hereafter Clades I and III isolates are referred to as 'Sri-Lankan clone'. In general, Sri-

267 Lankan clone isolates did not group according to their country of origin or year of recovery;

268 however, 24 small country-specific clusters of closely related isolates that differed by ≤10

269 cgSNPs were evident (Figure 1a).

270

271 *(ii) Core-genome MLST analysis*

272 As cgSNP analysis revealed low genotypic diversity among Sri-Lankan clone isolates

273 recovered over 17 years, the previously recommended threshold of ≤24 cgMLST allelic

274 differences for defining closely related isolates[28] was lowered to ≤20. The 214 Sri-Lankan

275 clone isolates exhibited a median of 55 allelic differences from one another (average: 59;

276 range: 0–200) (Table SIV). The 209 Clade I and five Clade III Sri-Lankan clone isolates

277 exhibited a median of 54 (average: 57; range: 0–153) and 42 (average: 43; range: 28–56)

278 allelic differences, respectively. Clade III isolates differed from Clade I by a median of 116

279 (average: 117; range: 70–200) allelic differences. Clade II comparator isolates differed from

280 Sri-Lankan clone Clade I and Clade III by a median of 87 (average: 86; range: 36–174) and

281 106 (average: 93; range: 28–134) allelic differences, respectively (Table SIV). These findings

282 confirmed limited diversity within the Sri-Lankan clone population.

283 Thirty-six related isolate groups (RIGs) comprising 123/266 study isolates were

284 evident in the cgMLST-based MST (Figure 1b, Table SV). Isolates within each RIG exhibited

285 ≤20 allelic differences to the closest neighbouring isolate in the RIG. Most RIGs included Sri-

286 Lankan clone isolates only (30/36 RIGs) and the remaining six (RIGs 31–36) included

287 comparator isolates only (Figure 1b and Table SV). There were 32 country-specific RIGs (27

288 Sri-Lankan clone isolates only (RIGs 1–2, 4–10, 12–13, 15–30) and five comparator isolates

289 only (RIGs 31, 33–36)) as follows: Denmark (*N*=16), Norway (*N*=4), Ireland (*N*=4),

290 Germany (*N*=3), UAE (N=3), Sri Lanka (*N*=1) and Czech Republic (*N*=1). The remaining

291 four RIGs comprised isolates from two or more countries. RIG-3 comprised eight Sri-Lankan

292 clone isolates from Denmark (*N*=3), Sri Lanka (*N*=2), Germany (*N*=1), Kuwait (*N*=1) and

293 the UAE (*N*=1), with an allelic difference range of 6–20 between neighbouring isolates in the

294 RIG and a range of 6-36 allelic differences for the entire RIG (Figure 1b and Table SV). Sri-

295   Lankan clone isolates from Norway (*N*=2) formed two separate RIGs (RIG-11 and RIG-14)

296   with Sri-Lankan clone isolates from Germany (*N*=1) and Denmark (*N*=1) with allelic

297   difference ranges of 1–21 and 18-20, respectively. One comparator isolate from Norway

298   formed a third RIG (RIG-32) with two comparator isolates from Ireland.

299

300   *(iii) Irish sub-clade*

301   Potential sub-clades (RIGs 1–36) observed on the cgSNP-based MLT and cgMLST-

302   based MST phylogenetic trees were further investigated using *in silico* DNA microarray

303   profiling and WGS data to identify possible RIG/country-specific characteristics. A genotypic

304   difference was observed between the overall Sri-Lankan clone population and 15/18 Irish Sri-

305   Lankan clone isolates in RIG-1. These 15 isolates formed a distinct Irish sub-clade within the

306   cgSNP-based MLT (Figure 1a) and lacked the *bbp* gene (also known as *sdrE*) encoding a

307   surface-associated, bone sialoprotein-binding protein. The absence of *bbp* in these 15 isolates

308   was confirmed by analysing hybrid assembled genomes (Figure S1a).

309

310   *Genotypic profiling of the Sri-Lankan clone*

311

312   *(i) Strain assignment and antimicrobial resistance genes*

313   DNA microarray profiling, SCC*mec*Finder and Ridom Seqsphere+ template tools for

314   detection of antimicrobial-resistance and virulence-associated genes revealed that most genes

315   in Sri-Lankan clone isolates (*N*=214) were homogenously distributed (Table SI). Microarray

316   analysis grouped Sri-Lankan clone isolates into two categories including "CC5-MRSA-IVc

317   (*sed/sej/ser+*)" (200/214; 93.5%) and "CC5-MRSA-IVc (*sed/sej/ser*-)" (14/214; 6.5%). The

318   two groups differed by the presence/absence of the *sed/sej/ser* enterotoxin genes, which were

319   located on a plasmid of approximately 27 kb. Of the five Sri-Lankan clone Clade III isolates

320   (Figure 1a), four isolates were "CC5-MRSA-IVc (*sed/sej/ser*-)" (Table SI). The *β-lactamase*

321   gene *blaZ* and the multidrug transporter encoding gene *lmrP* mediating resistance to

322   macrolides, lincosamides, streptogramins and tetracycline was harboured by the majority of

323   Sri-Lankan clone isolates (202/214; 94.4% and 212/214; 99.1%, respectively). Antimicrobial

324   genes detected are shown in Tables I and SI.

325

326   *(iii) IEC types*

327   IEC-type G was predominant amongst Sri-Lankan clone isolates (196/214; 91.6%).

328   The remaining isolates harboured IEC-type F (7/214; 3.3%), IEC-type E (2/214; 0.9%), IEC-

329   type D (1/214; 0.5%) or were non-typeable IEC variants harbouring *sep* only (2/214; 0.9%),

330 *sak* and *sep* (1/214; 0.5%) or *sak, scn* and *sea-sep* (4/214; 1.9%). One Sri-Lankan clone

331 isolate (M130242; Table SI) lacked lysogenic β-haemolysin converting bacteriophages and

332 carried no IEC genes.

333

334 *Epidemiological data*

335 Where detailed metadata were available (168/214; 78.5%), the majority of Sri-Lankan

336 clone isolates were CA-MRSA (85/168; 50.6%), while the remainder (50/168; 29.8%) were

337 HA-MRSA or were from hospitalised patients (33/168; 19.6%) (Table SI). Most isolates were

338 from infection sites (142/214; 66.3%), with the remainder from carriage (50/214; 23.4%) or

339 unknown sites (22/214; 10.3%). The majority of infection isolates were from SSTIs (83/142,

340 58.5%), other infection types (9/142) or were unknown (50/142).

341 Epidemiological information available for some Sri-Lankan clone isolates from

342 Denmark (*N=9*), Ireland (*N=2*) and the UAE (*N=1*) revealed that these isolates were from

343 patients with international links. The Irish isolates were recovered from patients with a history

344 of travel to Sri Lanka and Turkey, respectively. Within the Danish subset, one patient had

345 been hospitalised in Vietnam, four were from Sri-Lanka and four had travelled to Sri-Lanka.

346 The isolate from the UAE was from a patient from Bangladesh (Table SI).

347

348 *Pvl-encoding bacteriophage regions*

349 Clade I and III Sri-Lankan clone (*N=214*) and Clade II (*N=7*) comparator isolate

350 short-read assembled genomes were investigated for *pvl*-associated bacteriophage DNA. All

351 isolates harboured the *lukF/S*-PV genes, the phage lysis genes encoding amidase and holin

352 and remnants of phage structural genes encoding the tail fiber and major teichoic acid

353 biosynthesis protein. Genes associated with lysogeny, DNA replication/transcriptional

354 regulation and packaging/structure were not detected[29].

355 Twenty-six representative Sri-Lankan clone isolates (24 MLT Clade I and two MLT

356 Clade III isolates) available for long- and short-read sequencing underwent hybrid-assembly

357 to further investigate chromosomal regions surrounding *pvl*. These 26/214 isolates (2005–

358 2021) were from Czech Republic (*N=1*), Denmark (*N=6*), Germany (*N=2*), Ireland (*N=8*),

359 Kuwait (*N=1*), Norway (*N=3*), Saudi Arabia (*N=2*), Sweden (*N=1*) and the UAE (*N=2*).

360 Additionally, eight comparator isolates underwent hybrid-assembly (five outgroup and three

361 MLT Clade II comparators, Table SI). All 26 Sri-Lankan clone isolates lacked an intact

362 lysogenized *pvl*-encoding phage genome, but harboured a chromosomal remnant encoding the

363 *lukF/S-PV* genes as well as remnants of phage structural and lysis genes (Figure S2). In each

364 case, the phage remnant was 9,616 bp, with an intact upstream attachment site (*attL*), but no

365   downstream attachment site (*attR*) (Figure S2b). An identical phage remnant was observed in

366   the three Clade II comparator isolates and the single PVL-positive ST5-MRSA-IVa

367   comparator isolate (Z4294) next to Clade III (Figure 1a). The four remaining outgroup

368   comparator isolates all harboured a complete bacteriophage genome of ~45,000 bp which

369   shared 99.99% sequence homology with the well-characterised PVL-encoding phage

370   phiSa2wa (accession no. ON989481.1) (Figure S2a)[29]. The phage remnant exhibited 100%

371   sequence homology with the 3' junction of phage phiSa2wa (Figure S2b). Chromosomal

372   sequences adjacent the *pvl*-phage remnant were identical in all isolates investigated by hybrid

373   assembly.

374
375
376   **Discussion**
377

378   The emergence of PVL-positive MRSA is a public health concern globally. These

379   organisms were originally associated with community-onset infections, especially SSTIs but

380   also including necrotizing pneumonia, necrotizing fasciitis, and sepsis[2-4,30,31]. Patients

381   with community onset SSTIs often seek treatment in hospital emergency departments,

382   providing entry routes for CA-MRSA clones into hospitals[32,33]. The spread of PVL-

383   positive CA-MRSA clones into hospitals and resistance to a wide range of antimicrobials is

384   well documented[8,12,34,35].

385   The increasing prevalence of PVL-positive MRSA isolates from non-bloodstream

386   infections and hospital outbreaks both in Ireland and internationally is

387   concerning[8,13,34,35]. In 2019, McTavish et al. characterised a dominant PVL-positive

388   CC5-MRSA-IVc lineage in a Sri Lankan hospital and also identified it in the United Kingdom

389   and Australia[16]. In 2021, similar isolates from 13 patients during a protracted Irish

390   maternity unit hospital outbreak were described[13]. Consequently, our investigation sought

391   to compare Irish PVL-positive CC5-MRSA-IVc with the previously reported Sri-Lankan

392   clone and similar international isolates to determine the clone's global distribution, diversity

393   and population structure for the first time.

394   Phylogenetic analysis of 214 Sri-Lankan clone and 52 comparator isolates revealed

395   that the Sri-Lankan clone is relatively homogenous compared to other PVL-positive CA-

396   MRSA clones that have diverged more significantly over time[35]. Greater diversity maybe

397   revealed in future studies with more disparately recovered isolates. The vast majority of Sri-

398   Lankan clone isolates (209/214, 97.7%) recovered over 17-years grouped into Clade I

399   (Figures 1a) by cgSNP analysis with an average of 110 cgSNPs (57 cgMLST allelic

400   differences) between isolates (Tables SIII and SIV, respectively). The five remaining Sri-

401 Lankan clone isolates formed Clade III that differed from Clade I by an average of 237 SNPs
402 (117 cgMLST allelic differences). Seven PVL-positive ST5/t002-MRSA-IVa comparator
403 isolates in Clade II formed the closest neighbour to Sri-Lankan Clade I. Segregation of Sri-
404 Lankan clone isolates into country-specific RIGs (27/30 RIGs) by cgMLST probably reflects
405 local transmission and clonal evolution (Table SV). Some Danish isolates (N=27) in country-
406 specific RIGs also formed household-specific clusters (RIGs 4, 9–10, 13, 18–21, 24 and 29)
407 (Table SI). In some cases, different members of the same household presented with either
408 carriage or infection. Additionally, two isolates recovered from separate patients in a Danish
409 hospital clustered in RIG-30, with 16 cgSNPs (9 cgMLST allelic differences) between
410 isolates (Table SI). These findings highlight the significance of CA-MRSA transmission in
411 both community and hospital settings. Only limited inter-country dissemination of closely
412 related Sri-Lankan clone isolates was detected (RIG-3, RIG-11 and RIG-14), although this
413 possibly reflects the limited collection of isolates available for investigation.

414   Sri-Lankan clone isolates investigated were ST5, predominantly *spa* type t002 or
415 closely related *spa* types and harboured a relatively small number of antimicrobial-resistance
416 genes (Tables I and SI). DNA microarray and WGS data analyses revealed variable IEC gene
417 cluster (*sea/sak/scn/sep/chp*) and plasmid-encoded enterotoxin genes (*sed/sej/ser*) detection,
418 while the majority of other molecular characteristics were highly conserved (Table SI).
419 Although IEC-type G (91.6%) was predominant among Sri-Lankan clone isolates, six other
420 IEC types were detected (Table SI). Additionally, *sed/sej/ser* enterotoxin genes were absent in
421 only a small number of isolates (6.5%). The absence of the *bbp* gene within Irish maternity
422 unit hospital outbreak-associated isolates (Figure S1a) very likely reflects local loss of the
423 gene as other Irish isolates harboured the gene. Variation in IEC types and enterotoxin genes
424 probably reflects loss/gain of converting bacteriophages encoding IEC genes and *sed/sej/ser*-
425 encoding plasmids[36,37]. The prevalence of the multi-drug resistant PVL-negative European
426 CC1-MRSA-IV clone in Ireland[18,19] exemplifies the importance of mobile genetic
427 elements in the successful dissemination of emerging MRSA clones[31,35,38]. Earls *et al.*
428 described the emergence of European CC1-MRSA-IV from a South-Eastern European
429 meticillin-susceptible *S. aureus* (MSSA) CC1 lineage, and its subsequent rapid expansion
430 across Europe in the late 1990s[18,19]. European CC1-MRSA-IV is now the predominant
431 endemic CC1-MRSA clone in Ireland, associated with community transmissions and multi-
432 hospital outbreaks[19]. Periodic replacement of predominant MRSA clones in Irish hospitals
433 is well-documented[39], thus the recovery of the Sri-Lankan clone in 10 Irish hospitals over a
434 nine-year period (2013–2022) is concerning (Tables I and SI).

435   The Sri-Lankan clone chromosomally integrated defective 9.6 kb *pvl*-encoding phage
436   remnant (Figure S2b) may be a useful genetic marker, as the earliest Sri-Lankan clone study
437   isolate (141087, 2005) harboured this remnant. The remnant probably arose by imprecise
438   excision of a lysogenised *pvl*-phage genome, possibly as a result of a fitness cost imposed on
439   the bacteria through carriage of the entire prophage genome[40]. Stable chromosomal *pvl*-
440   retention without phage mobility-associated genes may provide a survival advantage.
441   Defective *pvl*-encoding bacteriophages with truncated tail formation genes have been
442   described in MRSA[41,42]. Furthermore, a defective *pvl*-phage has been reported in the
443   successful CA-MRSA clone USA300[43]. In the Sri-Lankan clone, approximately 80% of the
444   phage genome has been deleted leaving the *pvl*-encoding remnant.

445   The seven PVL-positive ST5/t002-MRSA-IVa comparator isolates in Clade II (the
446   closest neighbour to Sri-Lankan Clade I) and the single ST5-MRSA-IVa comparator isolate
447   Z4294 located adjacent to Clade III in the cgSNP MLT also harboured the 9.6 kb phage
448   remnant. These findings suggest that Clade II isolates, isolate Z4294 and Sri-Lankan clone
449   Clade I and III isolates emerged from a PVL-positive common ancestor harbouring the 9.6 kb
450   phage remnant, very likely a PVL-positive ST5-MSSA[19]. Sri-Lankan clone Clade I then
451   went on to disseminate widely. Interestingly, a PVL-positive CC5-MSSA isolate identified in
452   the puBMLST database by *in silico* PCR that also harboured the 9.6 kb phage remnant
453   clustered beside comparator isolate Z4294 and adjacent to Sri Lankan Clade III (Figure 1 and
454   Supplementary Figure S2).

455   This study had some limitations. Limited Sri-Lankan clone isolates/WGS datasets
456   were recovered following comprehensive literature and WGS database searches making it
457   difficult to assess its true prevalence (Table SII). Historical and contemporary data on MSSA
458   progenitor populations is limited in most MRSA lineages[44], including the Sri-Lankan clone.
459   Future investigations require a more comprehensive isolate collection with good quality
460   metadata, including potential progenitor MSSAs from more numerous and disparate regions.

461   In conclusion, international and local surveillance of emerging MRSA clones is
462   important for monitoring transmission. The association of Sri-Lankan clone isolates with
463   SSTIs in both community and hospital settings in 12 countries spanning 17 years reflects its
464   emergence internationally. The stable chromosomal integration of *pvl* in the Sri-Lankan clone
465   potentially contributes to its dissemination.

466

470 for investigation as well as Elke Müller and Annett Reissig (Jena) for performing array

471 experiments identifying suspect isolates. We thank Edet Udo (Kuwait), Bo Söderquist

472 (Sweden) Anne Tristran (Lyon) for providing some isolates for investigation.

473

Conflict of interest statement

475 None of the authors have any conflicts of interest to declare.

476

482

References

[1] Kaneko J, Kamio Y. Bacterial two-component and hetero-heptameric pore-forming cytolytic toxins: structures, pore-forming mechanism, and organization of the genes. Biosci Biotechnol Biochem 2004;68:981-1003. https://doi.org/10.1271/bbb.68.981

[2] Bhatta DR, Cavaco LM, Nath G, Kumar K, Gaur A, Gokhale S, et al. Association of Panton Valentine leukocidin (PVL) genes with methicillin resistant *Staphylococcus aureus* (MRSA) in Western Nepal: a matter of concern for community infections (a hospital based prospective study). BMC Infect Dis 2016;16:199. https://doi.org/10.1186/s12879-016-1531-1

[3] Jaiswal R, Garg A, Tripathi P, Venkatesh V. Epidemiology of Panton Valentine leukocidin in clinical *Staphylococcus aureus* isolates - A prospective study at a tertiary care centre in North India. Clin Epidemiol Glob Health 2022;15:101006. https://doi.org/10.1016/j.cegh.2022.101006

[4] Hussain K, Bandyopadhyay A, Roberts N, Mughal N, Moore L, Fuller LC. Panton-Valentine leucocidin-producing *Staphylococcus aureus*: a clinical review. Clin Exp Dermatol 2022;10 https://doi.org/10.1111/ced.15392

[5] Boakes E, Kearns AM, Ganner M, Perry C, Hill RL, Ellington MJ. Distinct bacteriophages encoding Panton-Valentine leukocidin (PVL) among international methicillin-resistant *Staphylococcus aureus* clones harboring PVL. J Clin Microbiol 2011;49:684-92. https://doi.org/10.1128/JCM.01917-10

[6] Otter JA, French GL. Community-associated meticillin-resistant *Staphylococcus aureus*: the case for a genotypic definition. J Hosp Infect 2012;81:143-8. https://doi.org/10.1016/j.jhin.2012.04.009

[7] Choo EJ. Community-associated methicillin-resistant *Staphylococcus aureus* in nosocomial infections. Infect Chemother 2017;49:158-9. https://doi.org/10.3947/ic.2017.49.2.158

[8] Steinig EJ, Duchene S, Robinson DA, Monecke S, Yokoyama M, Laabei M, et al. Evolution and global transmission of a multidrug-resistant, community-associated methicillin-resistant *Staphylococcus aureus* lineage from the Indian subcontinent. mBio 2019;10:01105-19. https://doi.org/10.1128/mBio.01105-19

[9] Hu Q, Cheng H, Yuan W, Zeng F, Shang W, Tang D, et al. Panton-Valentine leukocidin (PVL)-positive health care-associated methicillin-resistant *Staphylococcus aureus* isolates are associated with skin and soft tissue infections and colonized mainly by infective PVL-encoding bacteriophages. J Clin Microbiol 2015;53:67-72. https://doi.org/10.1128/JCM.01722-14

[10] Lakhundi S, Zhang K. Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. Clin Microbiol Rev 2018;31:00020-18. https://doi.org/10.1128/CMR.00020-18

[11] Peng H, Liu D, Ma Y, Gao W. Comparison of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus* isolates at a Chinese tertiary hospital, 2012–2017. Sci Rep 2018;8:17916 https://doi.org/10.1038/s41598-018-36206-5

[12] Shore AC, Tecklenborg SC, Brennan GI, Ehricht R, Monecke S, Coleman DC. Panton-Valentine leukocidin-positive *Staphylococcus aureus* in Ireland from 2002 to 2011: 21 clones, frequent importation of clones, temporal shifts of predominant methicillin-resistant *S. aureus* clones, and increasing multiresistance. J Clin Microbiol 2014;52:859-70. https://doi.org/10.1128/JCM.02799-13

[13] McManus BA, Aloba BK, Earls MR, Brennan GI, O'Connell B, Monecke S, et al. Multiple distinct outbreaks of Panton–Valentine leucocidin-positive community-associated meticillin-resistant *Staphylococcus aureus* in Ireland investigated by whole-genome sequencing. J Hosp Infect 2021;108:72-80. https://doi.org/10.1016/j.jhin.2020.11.021

[14] National methicillin-resistant *Staphylococcus aureus* reference laboratory. 2017. Annual report. Available at: https://www.stjames.ie/media/NMRSARLAnnualReport2017.pdf [last accessed October 2022].

[15] National methicillin-resistant *Staphylococcus aureus* reference laboratory. 2020. Annual report. Available at: https://www.stjames.ie/media/AnnRpt2020.pdf [last accessed October 2022].

[16] McTavish SM, Snow SJ, Cook EC, Pichon B, Coleman S, Coombs GW, et al. Genomic and epidemiological evidence of a dominant Panton-Valentine leucocidin-positive methicillin resistant *Staphylococcus aureus* lineage in Sri Lanka and presence among isolates from the United Kingdom and Australia. Front Cell Infect Microbiol 2019;9:123. https://doi.org/10.3389/fcimb.2019.00123

[17] Senok A, Nassar R, Celiloglu H, Nabi A, Alfaresi M, Weber S, et al. Genotyping of methicillin resistant *Staphylococcus aureus* from the United Arab Emirates. Sci Rep 2020;10:18551. https://doi.org/10.1038/s41598-020-75565-w

566 [18] Earls MR, Shore AC, Brennan GI, Simbeck A, Schneider-Brachert W, Vremerǎ T, et
567   al. A novel multidrug-resistant PVL-negative CC1-MRSA-IV clone emerging in
568   Ireland and Germany likely originated in South-Eastern Europe. Infect Genet Evol
569   2019;69:117-26. https://doi.org/10.1016/j.meegid.2019.01.021
570
571 [19] Earls MR, Steinig EJ, Monecke S, Samaniego Castruita JA, Simbeck A, Schneider-
572   Brachert W, et al. Exploring the evolution and epidemiology of European CC1-
573   MRSA-IV: tracking a multidrug-resistant community-associated meticillin-resistant
574   *Staphylococcus aureus* clone. Microb Genom 2021;7:000601.
575   https://doi.org/10.1099/mgen.0.000601
576
577 [20] Albrecht N, Jatzwauk L, Slickers P, Ehricht R, Monecke S. Clonal replacement of
578   epidemic methicillin-resistant *Staphylococcus aureus* strains in a German university
579   hospital over a period of eleven years. PLoS One 2011;6:28189.
580   https://doi.org/10.1371/journal.pone.0028189
581
582 [21] Das S, Anderson CJ, Grayes A, Mendoza K, Harazin M, Schora DM, et al. Nasal
583   carriage of epidemic methicillin-resistant *Staphylococcus aureus* 15 (EMRSA-15)
584   clone observed in three Chicago-area long-term care facilities. Antimicrob Agents
585   Chemother 2013;57[9]:4551-3. https://doi.org/10.1128/AAC.00528-13
586
587 [22] Egan SA, Kavanagh NL, Shore AC, Mollerup S, Samaniego Castruita JA, O'Connell
588   B, et al. Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals
589   a high prevalence of ST80 and spread of similar *vanA* regions via IS*1216E* and
590   plasmid transfer in diverse genetic lineages in Ireland. J. Antimicrob. Chemother
591   2021;77:320-30. https://doi.org/10.1093/jac/dkab393
592
593 [23] Monecke S, Jatzwauk L, Weber S, Slickers P, Ehricht R. DNA microarray-based
594   genotyping of methicillin-resistant *Staphylococcus aureus* strains from Eastern
595   Saxony. Clin Microbiol Infect 2008;14:534-45 https://doi.org/10.1111/j.1469-
596   0691.2008.01986.x
597
598 [24] Monecke S, Slickers P, Ehricht R. Assignment of *Staphylococcus aureus* isolates to
599   clonal complexes based on microarray analysis and pattern recognition. FEMS
600   Immunol Med Microbiol 2008;53:237-51 https://doi.org/10.1111/j.1574-
601   695X.2008.00426.x
602
603 [25] Monecke S, Jatzwauk L, Müller E, Nitschke H, Pfohl K, Slickers P, et al. Diversity of
604   SCC*mec* elements in *Staphylococcus aureus* as observed in South-Eastern Germany.
605   PLoS One 2016;11:0162654. https://doi.org/10.1371/journal.pone.0162654
606
607 [26] Hau SJ, Sun J, Davies PR, Frana TS, Nicholson TL. Comparative prevalence of
608   immune evasion complex genes associated with β-hemolysin converting
609   bacteriophages in MRSA ST5 isolates from swine, swine facilities, humans with
610   swine contact, and humans with no swine contact. PLoS One 2015;10:0142832.
611   https://doi.org/10.1371/journal.pone.0142832
612
613 [27] Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A. Bacterial whole-
614   genome sequencing revisited: portable, scalable, and standardized analysis for typing
615   and detection of virulence and antibiotic resistance genes. J Clin Microbiol
616   2014;52:2365-70. https://doi.org/10.1128/JCM.00262-14
617

[28] Earls MR, Coleman DC, Brennan GI, Fleming T, Monecke S, Slickers P, et al. Intra-hospital, inter-hospital and intercontinental spread of ST78 MRSA from two neonatal intensive care unit outbreaks established using whole-genome sequencing. Front Microbiol 2018;9:1485. https://doi.org/10.3389/fmicb.2018.01485

[29] Coombs GW, Baines SL, Howden BP, Swenson KM, O'Brien FG. Diversity of bacteriophages encoding Panton-Valentine leukocidin in temporally and geographically related *Staphylococcus aureus*. PLoS One 2020;15[2]:0228676. https://doi.org/10.1371/journal.pone.0228676

[30] Lina G, Piémont Y, Godail-Gamot F, Bes M, Peter M-O, Gauduchon V, et al. Involvement of Panton-Valentine leukocidin—producing *Staphylococcus aureus* in primary skin infections and pneumonia. Clin Infect Dis 1999;29:1128-32. https://doi.org/10.1086/313461

[31] Watkins RR, David MZ, Salata RA. Current concepts on the virulence mechanisms of meticillin-resistant *Staphylococcus aureus*. J Med Microbiol 2012;61:1179-93. https://doi.org/10.1099/jmm.0.043513-0

[32] Bouchiat C, Curtis S, Spiliopoulou I, Bes M, Cocuzza C, Codita I, et al. MRSA infections among patients in the emergency department: a European multicentre study. J. Antimicrob. Chemother 2016;72:372-5. https://doi.org/10.1093/jac/dkw431

[33] Kossow A, Stühmer B, Schaumburg F, Becker K, Glatz B, Möllers M, et al. High prevalence of MRSA and multi-resistant Gram-negative bacteria in refugees admitted to the hospital—But no hint of transmission. PLoS One 2018;13:0198103. https://doi.org/10.1371/journal.pone.0198103

[34] Blomfeldt A, Larssen KW, Moghen A, Gabrielsen C, Elstrøm P, Aamot HV, et al. Emerging multidrug-resistant Bengal Bay clone ST772-MRSA-V in Norway: molecular epidemiology 2004-2014. Eur J Clin Microbiol Infect Dis 2017;36:1911-21. https://doi.org/10.1007/s10096-017-3014-8

[35] Challagundla L, Luo X, Tickler IA, Didelot X, Coleman DC, Shore AC, et al. Range expansion and the origin of USA300 North American epidemic methicillin-resistant *Staphylococcus aureus*. mBio 2018;9:02016-17. https://doi.org/10.1128/mBio.02016-17

[36] Xia G, Wolz C. Phages of *Staphylococcus aureus* and their impact on host evolution. Infect Genet Evol 2014;21:593-601. https://doi.org/10.1016/j.meegid.2013.04.022

[37] Varshney AK, Mediavilla JR, Robiou N, Guh A, Wang X, Gialanella P, et al. Diverse enterotoxin gene profiles among clonal complexes of *Staphylococcus aureus* isolates from the Bronx, New York. Appl Environ Microb 2009;75:6839-49. https://doi.org/10.1128/AEM.00272-09

[38] Lindsay JA, Knight GM, Budd EL, McCarthy AJ. Shuffling of mobile genetic elements (MGEs) in successful healthcare-associated MRSA [HA-MRSA]. Mob Genet Elements 2012;2:239-43. https://doi.org/10.4161/mge.22085

[39] Kinnevey PM, Shore AC, Brennan GI, Sullivan DJ, Ehricht R, Monecke S, et al. Extensive genetic diversity identified among sporadic methicillin-resistant

670 *Staphylococcus aureus* isolates recovered in Irish hospitals between 2000 and 2012.
671 Antimicrob Agents Chemother 2014;58:1907-17. https://doi.org/10.1128/AAC.02653-
672 13
673

674 [40] Rohmer C, Wolz C. The role of *hlb*-converting bacteriophages in *Staphylococcus*
675 *aureus* host adaption. Microb Physiol 2021;31:109–122.
676 https://doi.org/10.1159/000516645
677

678 [41] Ma XX, Ito T, Chongtrakool P, Hiramatsu K. Predominance of clones carrying
679 Panton-Valentine leukocidin genes among methicillin-resistant *Staphylococcus aureus*
680 strains isolated in Japanese hospitals from 1979 to 1985. J Clin Microbiol
681 2006;44:4515-27. https://doi.org/10.1128/JCM.00985-06
682

683 [42] Kaneko J, Kimura T, Narita S, Tomita T, Kamio Y. Complete nucleotide sequence
684 and molecular characterization of the temperate staphylococcal bacteriophage phiPVL
685 carrying Panton-Valentine leukocidin genes. Gene 1998;215:57-67.
686 https://doi.org/10.1016/s0378-1119(98)00278-9
687

688 [43] Wirtz C, Witte W, Wolz C, Goerke C. Transcription of the phage-encoded Panton–
689 Valentine leukocidin of *Staphylococcus aureus* is dependent on the phage life-cycle
690 and on the host background. Microbiology 2009;155:3491-9.
691 https://doi.org/10.1099/mic.0.032466-0
692

693 [44] Steinig E, Aglua I, Duchêne S, Meehan MT, Yoannes M, Firth C, et al. Phylodynamic
694 signatures in the emergence of community-associated MRSA. Proc Natl Acad Sci
695 USA 2022;119:e2204993119. https://www.pnas.org/doi/10.1073/pnas.2204993119

696
697
698
699
700
701
702
703
704
705
706

Table 1. Antimicrobial resistance and virulence-associated gene profiles of 214 PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone isolates and 52 additional PVL-positive (*N*=29) and PVL-negative (*N*=23) CC5/ST5-MRSA-I/II/IVa/IVc/IVg/V comparator isolates investigated

Country	Isolates (*N*)	Year(s) of isolation	*spa*-ST-SCC*mec* (*N*)	Antimicrobial resistance genes (*N*)	PVL (+/-)	IEC Type (*N*)	Reference
Algeria	Comparator (1)	2003	t450-ST5-IVa	*aadD*, *erm*(C), *fosB*, *lmrP*, *mprF*, *tet*(M), *vga*(A), *sdrM*	+	B	This study
Australia	Sri-Lankan clone (1)	2015	t002-ST5-IVc	*blaZ*, *fosB*, *lmrP*, *sdrM*	+	G	McTavish *et al.* [16]
Czech Republic	Sri-Lankan clone (6)	2018–2021	t002-ST5-IVc	*blaZ* (4), *fosB* (6), *lmrP* (5), *sdrM* (6), *mprF* (6), *erm*(C) (1)	+	G (4) E(1) Novel type 3 (*sak, sep)* (1)	This study
	Comparators (2)	2019–2021	t002-ST5-IVa (1) t002-ST5-II (1)	*aadD* (1), *blaZ* (2), *erm*(A) (1), *fosB* (2), *kdpA/B/C/D/E* (1), *lmrP* (2), *mprF* (2), *sdrM* (2), *xylR* (1)	-	G (1) B (1)	
Denmark	Sri-Lankan clone (66)	2007–2021	t002-ST5-IVc	*blaZ* (61), *fosB* (65), *lmrP* (65), *mprF* (65), *sdrM* (66), *erm*(C) (31), *tet*(K) (3), *mupA* (3), *qacA* (3), *qacC* (1), *cat* (1)	+	G (59) F (7)	This study
	Comparators (10)	2013–2015	t002-ST5-IVa (2) t002-ST5-V (8)	*blaZ* (9), *fosB* (10), *lmrP* (10), *mprF* (10), *sdrM* (10), *aacA-aphD* (7), *erm*(C) (2), *tet*(K) (1)	+ (7) - (3)	G (8) F (1) B (1)	
Germany	Sri-Lankan clone (20)	2011–2019	t002-ST5-IVc (15) t535-ST5-IVc (2) t579-ST5-IVc (1) ND-ST5-IVc (2)	*blaZ* (18), *erm*(C) (11), *fosB* (20), *lmrP* (20), *mprF* (20), *sdrM* (19), *tet*(K) (1), *qacA* (1), *msr*(A) (1)	+	G (16) E (1) Novel type 1(*sep* only) (1) Novel type 2(*sak, scn, sea, sep*) (1)	This study
	Comparators (2)	2014–2017	t105-ST5-IVc	*blaZ* (1), *fosB* (1), *lmrP* (1), *mprF* (2), *sdrM* (1)	+	B	
Ireland	Sri-Lankan clone (30)	2013–2022	t002-ST5-IVc	*blaZ* (29), *fosB* (30), *lmrP* (30), *mprF* (30), *sdrM* (30), *erm*(C) (5)	+	G (29) None (1)	This study, McManus *et al.* [13]
	Comparator (17)	2013–2019	t002-ST5-I (1) t002-ST5-IVa (2) t002-ST5-IVc (3) t002-ST5-IVg (2) t311-ST5-V (9)	*blaZ* (13), *erm*(C) (10), *fosB* (17), *lmrP* (17), *mprF* (17), *sdrM* (17), *fusC* (10), *fexA* (1), *aadD* (1), *qacA* (1), *merA* (1)	+ (2) - (15)	G (2) F (2) B (3) E (9) Novel type 1(*sep* only) (1)	
Kuwait	Sri-Lankan clone (1)	2013	t002-ST5-IVc	*blaZ*, *erm*(C), *fosB*, *lmrP*, *mprF*, *sdrM*	+	G	This study

	Comparator (1)	2013	t002-ST5-IVa	*blaZ, erm*(C)*, fosB, lmrP, mprF, sdrM*	+	G	
Norway	Sri-Lankan clone (24)	2007–2021	t002-ST5-IVc (23) t1062-ST5-IVc (1)	*blaZ* (23)*, fosB* (24)*, lmrP* (24)*, mprF* (24)*, sdrM* (24)*, erm*(C) (9)*, tet*(K) (2)*, vga*(A) (1)	+	G (23) Novel Type 2 *(sak, scn, sea, sep)* (1)	This study
	Comparators (12)	2003–2020	t311-ST5-IVa (4) t311-ST5-IVc (2) t002-ST5-IVa (3) t105-ST5-IVc (1) t3089-ST5-IVa (1) t442-ST5-V (1)	*blaZ* (12)*, fosB* (12)*, lmrP* (12)*, mprF* (12)*, sdrM* (12)*, erm*(C) (1)*, tet*(K) (1)*, aacA-aphD* (2)*, dfrA* (2)*, tet*(M) (2)*, aphA3* (3)*, mph*(C) (1)*, msr*(A) (1)*, sat* (3)*, qacC* (1)	+ (11) - (1)	G (3) B (5) A (4)	
Saudi Arabia	Sri-Lankan clone (4)	2010–2017	t002-ST5-IVc	*blaZ* (3)*, fosB* (4)*, lmrP* (4)*, mprF* (4)*, sdrM* (4)*, erm*(C) (2)*, aphA3* (1)*, sat* (1)	+	G (4)	This study
	Comparator (1)	2010	t311-ST5-IVa	*blaZ, fosB, lmrP, mprF, sdrM*	-	B	
Senegal	Comparators (2)	2007	t311-ST5-IVa	*aacA-aphD* (1)*, aadD* (1)*, blaZ* (1)*, dfrA* (2)*, fosB* (2)*, lmrP* (2)*, mprF* (2)*, qacC* (1)*, tet*(M) (2)*, sdrM* (2)	+	B (2)	This study
Slovakia	Comparator (1)	2020	t002-ST5-IVc	*blaZ, fosB, lmrP, mprF, sdrM*	-	B	This study
Sri-Lanka	Sri-Lankan clone (33)	2014	t002-ST5-IVc (21) t010-ST5-IVc (1) t045-ST5-IVc (2) t062-ST5-IVc (4) t1062-ST5-IVc (1) ND-ST5-IVc (4)	*blaZ* (33)*, fosB* (33)*, lmrP* (33)*, mprF* (33)*, sdrM* (33)*, erm*(C) (14)*, tet*(K) (4)	+	G (31) Novel type 1 *(sep* only) (1) Novel Type 2 *(sak, scn, sea, sep)* (1)	McTavish *et al.* [16]
Sweden	Sri-Lankan clone (2)	2005–2009	t002-ST5-IVc	*blaZ* (2)*, erm*(C) (1)*, fosB* (2)*, lmrP* (2)*, mprF* (2)*, sdrM* (2)*, tet*(K) (1)	+	G (2)	This study
United Arab Emirates	Sri-Lankan clone (15)	2017–2019	t002-ST5-IVc (12) t010-ST5-IVc (1) t045-ST5-IVc (1) t306-ST5-IVc (1)	*blaZ* (15)*, erm*(C) (9)*, fosB* (15)*, lmrP* (17)*, mprF* (17)*, sdrM* (17)	+	G (15)	This study
	Comparator (3)	2018	t105-ST5-IVc (2) t002-ST5-IVa (1)	*blaZ* (1)*, erm*(C) (1)*, fosB* (3)*, lmrP* (3)*, mprF* (3)*, sdrM* (3)	+	G (1) B (2)	

| United Kingdom | Sri-Lankan clone (12) | 2005–2015 | t002-ST5-IVc (5)
t062-ST5-IVc (2)
ND-ST5-IVc (5) | *blaZ* (12), *fosB* (12), *lmrP* (12), *mprF* (12), *sdrM* (12), *erm*(C) (4), *tet*(K) (2), *sat* (1) | + | G (11)
D (1) | McTavish *et al.* [16] |

09 Genotypic information for 266 study isolates (Sri-Lankan clone and comparator isolates) recovered from 15 different countries between 2003–2022. Sri-Lankan clone
10 isolates were recovered from 12 countries across Europe, Asia, Australia, and the Middle East between 2005–2022. 214 PVL+ CC5/ST5-MRSA-IVc Sri-Lankan clone
11 isolates and 52 additional PVL-positive (*N*=29) and PVL-negative (*N*=23) CC5/ST5-MRSA-I/II/IVa/IVc/IVg/V comparator isolates were investigated. The isolates were
12 subjected to whole-genome sequencing and subsequent analyses and profiling to determine antimicrobial resistance gene patterns and virulence gene profiles. Genotypic
13 information was extracted from whole-genome data using Ridom SeqSphere+ v7.0.4 (Ridom GmbH, Münster, Germany) genotyping & *S. aureus* Genotyping Kit 2.0
14 (Abbott) microarray technology[19, 23–25]. The isolates also underwent core-genome multi-locus sequence typing and single nucleotide polymorphic analyses.
15
16 Abbreviations: ND, not determined – isolates not available and *spa* types could not be determined using *in-silico* techniques on the available genomic sequence data; ST,
17 sequence type; SCC*mec*, staphylococcal chromosomal cassette harbouring *mecA*; PVL, Panton-Valentine leukocidin; +, positive; -, negative; IEC, immune evasion cluster;
18 *scn*, staphylococcal complement inhibitor, *sea*; staphylococcal enterotoxin a gene, s*ep*, staphylococcal enterotoxin p gene; *sak*, staphylokinase gene.

719 **Figure Legends**

720

721 **Figure 1a. Maximum likelihood tree (MLT) based on phylogenetic analysis of 12,245**
722 **core-genome single nucleotide polymorphisms (cgSNPs) for 214 PVL-positive CC5/ST5-**
723 **MRSA-IVc Sri-Lankan clone isolates and 52 additional PVL-positive (*N*=29) and PVL-**
724 **negative (*N*=23) CC5/ST5-MRSA-I/II/IVa/IVc/IVg/V comparator isolates.**

725 Sri-Lankan clone isolates were recovered from 12 countries across Europe, Asia, Australia
726 and the Middle East between 2005–2022. Separate node colours/shapes represent country of
727 recovery, sample types and SCC*mec* types as indicated in the legend. Blue branches represent
728 PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone isolates (*N*=209/214) forming Clade I.
729 Green branches represent PVL-positive CC5/ST5-MRSA-IVa comparator isolates (*N*=7/52)
730 forming Clade II. Red branches represent the remaining PVL-positive CC5/ST5-MRSA-IVc
731 Sri-Lankan clone isolates (*N*=5/214) forming Clade III. The thick black branch represents a
732 PVL-positive CC5/ST5-MRSA-IVa comparator isolate (*n*=1) branching out next to Clade III.
733 The thin black branches represent the comparator outgroup isolates (*N*=44) which separate
734 away from Clades I–III. Labels for the 52 comparator isolates are highlighted in orange.
735 Isolate names, year of recovery and *spa* types are all indicated in the branch labelling.
736 Country-specific isolate pairs or clusters containing closely related isolates that differed by
737 ≤10 cgSNPs are shaded in grey. The divergent subgroup of 15 Irish isolates (lacking the *bbp*
738 gene) within the large Sri-Lankan clade (Clade I) is indicated by an asterisk and the isolate
739 names are highlighted in green. The epidemiological and genotypic information for each
740 isolate investigated is provided in Table I and Supplementary Table SI. Corresponding SNP
741 distance matrix data is provided in Supplementary Table SIII.

742
743

744 **Figure 1b. Minimum spanning tree (MST) based on core genome multi-locus sequence**
745 **type (cgMLST) analysis of 1,861 target genes for 214 PVL-positive CC5/ST5-MRSA-IVc**
746 **Sri-Lankan clone isolates and 52 additional PVL-positive (*N*=29) & PVL-negative**
747 **(*N*=23) CC5/ST5-MRSA-I/II/IVa/IVc/IVg/V comparator isolates.**

748 Separate node colours represent country of isolation as indicated in the legend. Partitions
749 within nodes represent the presence of ≥2 isolates per node. Comparator isolates are indicated
750 by red squares. Closely related clusters of isolates (≤20 allelic differences to the closest
751 neighbouring isolate within the RIG) are outlined within grey shadowing. Branch numbers
752 indicate the number of allelic differences between neighbouring isolates. Node numbers
753 indicate the 36 related isolate groups (RIGs) in the population (Table SI). The subgroup of

754 closely related Irish isolates consisting of 15 isolates with a distinct genotypic profile to other
755 Sri-Lankan clone isolates is outlined in green. The corresponding cgMLST pairwise isolate
756 distance matrix is provided in Supplementary Table SIV. The cgMLST-based MST was
757 constructed using Ridom SeqSphere+ v7.0.4 (Ridom GmbH, Münster, Germany).
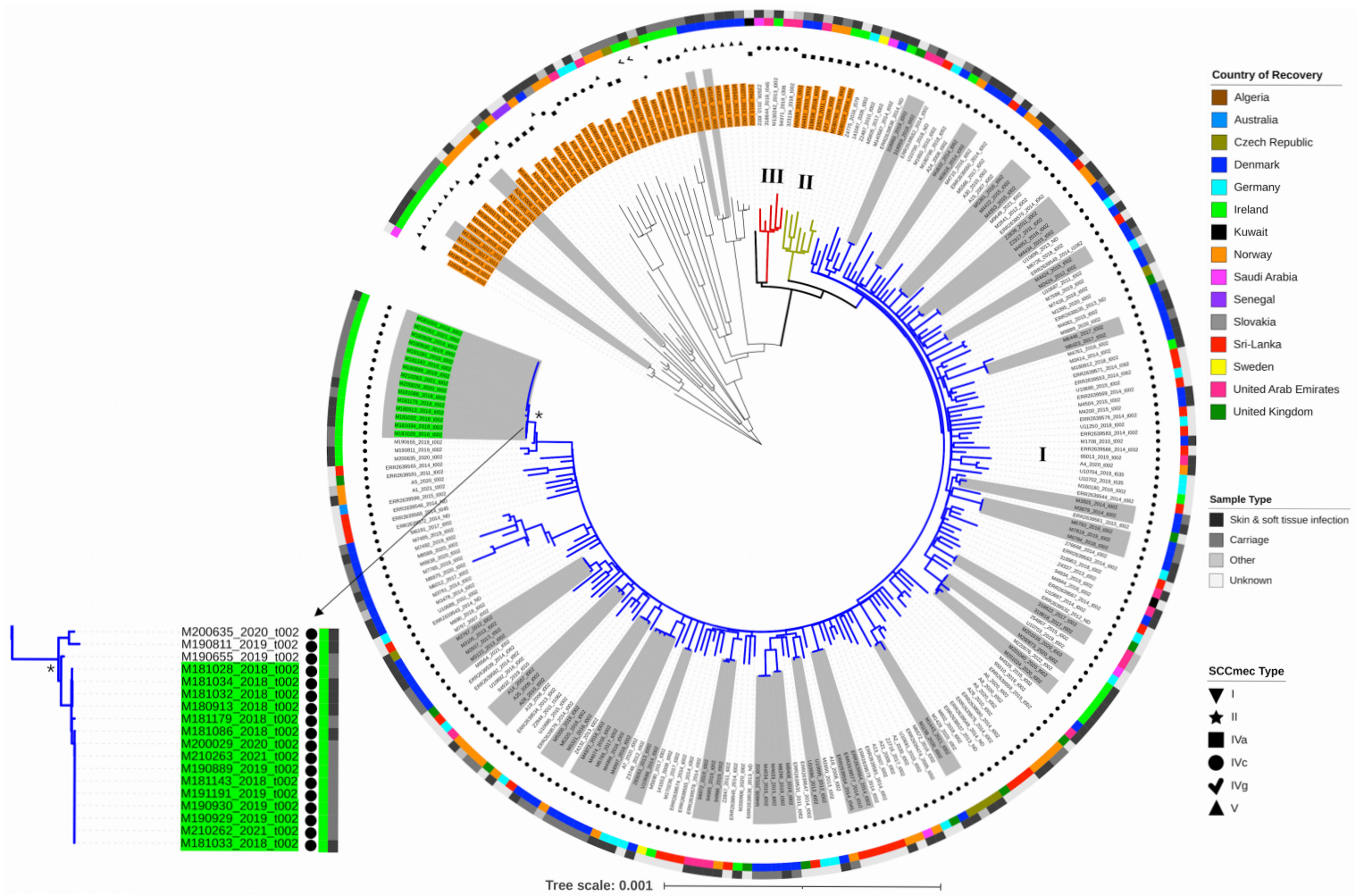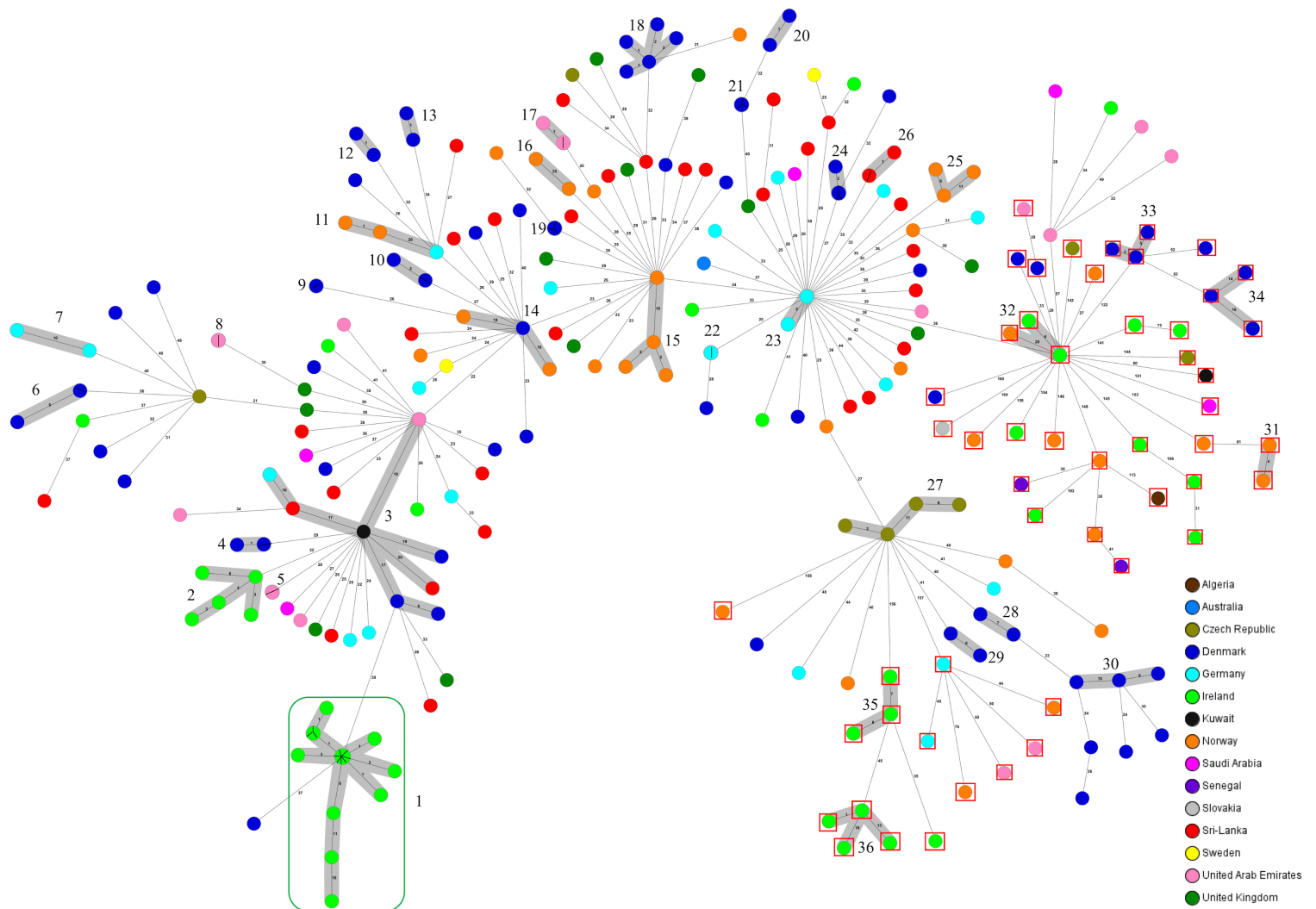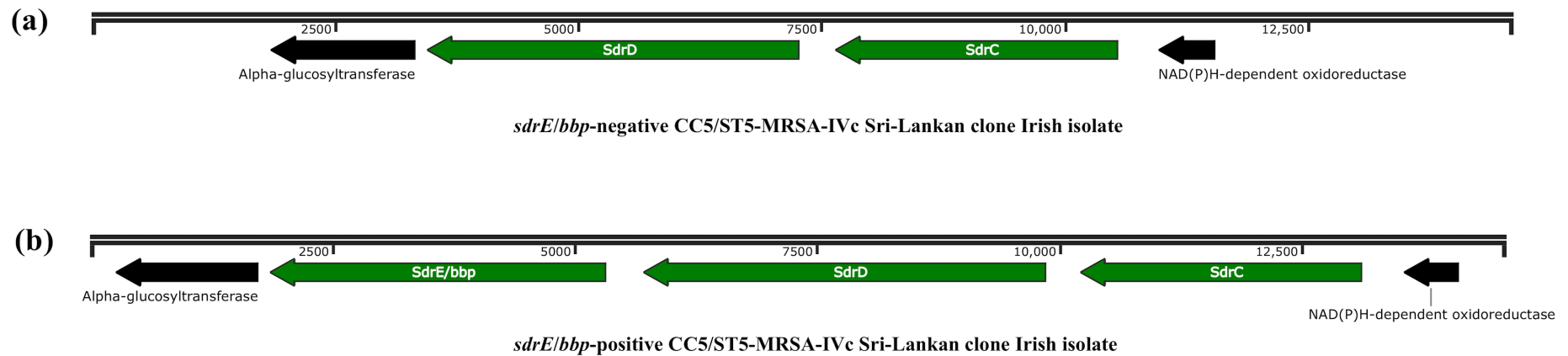758
759
760
761

Figure 1a

Figure 1b

Algeria
Australia
Czech Republic
Denmark
Germany
Ireland
Kuwait
Norway
Saudi Arabia
Senegal
Slovakia
Sri-Lanka
Sweden
United Arab Emirates
United Kingdom

(a)

sdrE/bbp-negative CC5/ST5-MRSA-IVc Sri-Lankan clone Irish isolate

(b)

*sdrE/bbp*-positive CC5/ST5-MRSA-IVc Sri-Lankan clone Irish isolate

Supplementary Figure S1. Comparative structural organization of the serine-aspartate repeat protein-encoding (*sdr*) locus of (a) the *SdrE/bbp*-negative PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone Irish isolate M181179 and (b) the *SdrE/bbp*-positive PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone Irish isolate M180912. The tandemly arranged *sdr* genes *sdrC* (~2904 bp), *sdrD* (~4140 bp) and *sdrE* (~3462 bp) encode Sdr surface proteins (members of the Microbial Surface Components Recognising Adhesive Matrix Molecules [MSCRAMM] family). A divergent subgroup of 15 PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone Irish isolates recovered between 2018–2021 lacked this *sdrE/bbp* gene. Isolate sequences underwent hybrid assembly following long-read and short-read sequencing. The hybrid assembled genomes were annotated using RAST v2.0 (https://rast.nmpdr.org) and visualised using SnapGene v6.0.6 (https://www.snapgene.com).

(a)
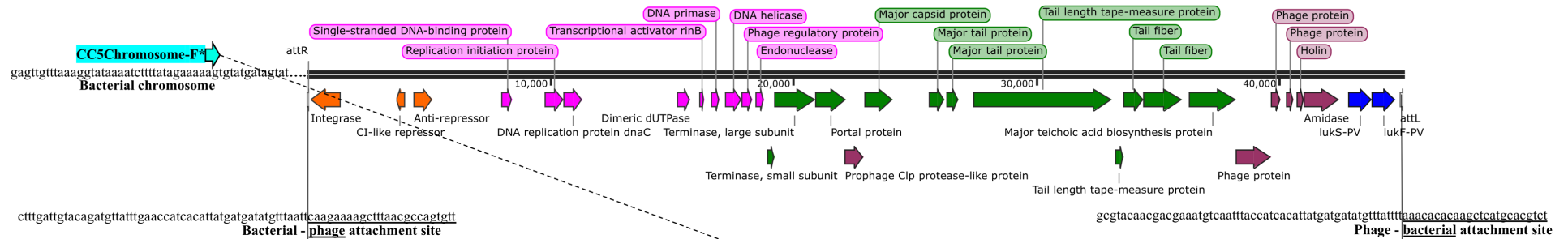
(b)
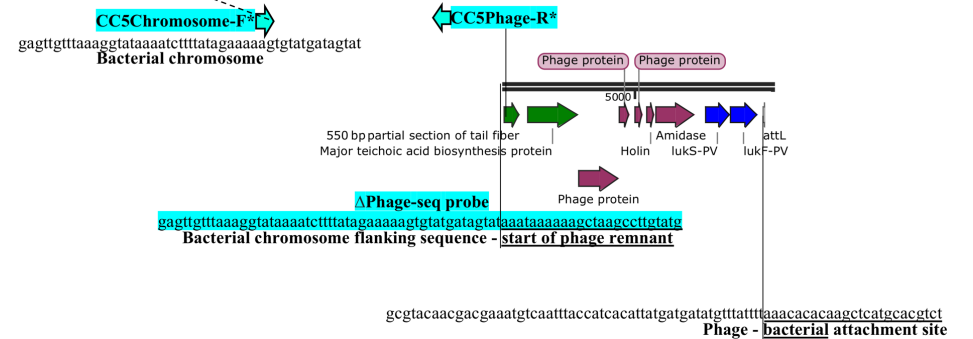
Supplemental Figure S2. Comparison of the structural organisation of the Panton-Valentine leukocidin (*pvl*) *lukF/S*-PV genes and associated phage DNA sequences within the chromosome of a PVL-positive CC5/ST5-MRSA-IVc comparator isolate harbouring an intact lysogenised *pvl*-encoding bacteriophage genome, and a PVL-positive CC5/ST5-MRSA-IVc Sri Lankan clone isolate carrying a *pvl*-encoding phage remnant. The organisation of phage DNA regions was determined by hybrid assembly. (a) Comparator isolate 323179 (United Arab Emirates, 2018) harbouring a fully intact lysogenised *pvl*-encoding bacteriophage genome (~45,000 bp) with 99.99% sequence homology to the *pvl*-encoding phage phiSa2wa (accession no. ON989481.1). (b) Sri Lankan clone isolate M181179 (Ireland, 2018) harbouring a 9,616 bp *pvl*-encoding phage remnant with 100% sequence homology to the 3' end junction of the phiSa2wa phage. The typical components of the bacteriophage genome are highlighted by colour; lysogeny module (orange), DNA replication and transcription genes (pink), head and tail packaging genes (green), lysis module (dark purple) and the

lukS/F-PV region (blue). Attachment sites (*attR* and *attL*) are typically present at the proximal and distal junction ends of lysogenized prophage genomes integrated in the bacterial chromosome. The phage remnant probably arose from an imprecise excision event that resulted in loss of most of the phage genome and a 614 bp chromosomal DNA sequence upstream and including the *attR* site. In isolate M181179, the annotated bacterial chromosome flanking sequence and phage remnant' sequence are indicated. The hybrid assembled genomes were annotated using RAST v2.0 (https://rast.nmpdr.org) and visualised using SnapGene v6.0.6 (https://www.snapgene.com).

All available 29,504 *S. aureus* assembled genomes in the pubMLST database (https://pubmlst.org) (accessed 2nd Nov. 2022) were screened by *in silico* PCR using primers CC5Chromosome-F* (5'-ATTCGATTGCACGTTCTG-3') and CC5Phage-R* (5'-ACTTAACAGACGAGTTATTGCAC-3') indicated in (b) which yield an 856 bp amplimer with assembled genomes harbouring the 9.6 kb phage remnant (https://pubmlst.org/bigsdb?db=pubmlst_saureus_isolates&page=query&genomes=1). Ten PVL-positive CC5/ST5-MRSA-IVc (8/10, t002) (Accession Nos. ERR211934, ERR211966, ERR212816, ERR212871, ERR540754, ERR540938, ERR541062, ERR541068, ERR714842 and SRR917592), seven PVL-positive CC5/ST5-MRSA-IVa (6/7, t002) (Accession Nos. ERR204190, ERR212783, ERR212986, ERR527305, ERR702114, ERR715326 and ERR737419) and one PVL-positive CC5/ST5/t002-MSSA (Accession No. ERR109505) were identified that harboured the 9.6 kb phage remnant. A 75 bp sequence "ΔPhage-seq probe" (highlighted in light blue in (b)) that spanned the chromosomal/5'-phage junction was used as an *in silico* probe against the assembled genomes to confirm *in silico* PCR results.

Investigation of these 18 isolates and all other study isolates by cgMLST SNP-based maximum likelihood analysis grouped all 10 PVL-positive CC5/ST5-MRSA-IVc into Sri Lankan clone Clade I and all 7 PVL-positive CC5/ST5-MRSA-IVa into Clade II alongside the other CC5/ST5-MRSA-IVa comparators. The PVL-positive CC5/ST5-MSSA isolate branched out next to Sri Lankan Clade III isolates. The complete absence of SCC*mec* element sequences in this MSSA isolate and its close phylogenetic proximity to Sri Lankan clone isolates further supports the probability of a PVL-positive MSSA progenitor giving rise to the Sri Lankan clone. Due to the lack of publicly available metadata, these 18 additional isolates were not included in the primary phylogenetic analysis (Figures 1a and 1b).

Supplementary Table SII. Literature search to identify previously investigated PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone isolates

Part (a): Literature search results; closest hits (isolates included in study are highlighted in red text)

Title	Country	Year	DOI	No. of isolates in study	ST	PVL+/-	SCC*mec* type
Genomic and epidemiological evidence of a dominant Panton-Valentine leucocidin-positive methicillin resistant *Staphylococcus aureus* lineage in Sri Lanka and presence among isolates from the United Kingdom and Australia	United Kingdom, Sri-Lanka, Australia	2019	https://doi.org/10.3389/fcimb.2019.00123	56 (46 included in study)	ST5	PVL+	IVc
Genotyping of methicillin resistant *Staphylococcus aureus* from the United Arab Emirates	United Arab Emirates	2020	https://doi.org/10.1038/s41598-020-75565-w	14	ST5	PVL+	IVc
Genomics of *Staphylococcus aureus* ocular isolates	United States of America	2021	https://doi.org/10.1371/journal.pone.0250975	2	ST5	PVL+	IVc
The molecular epidemiology of methicillin-resistant *Staphylococcus aureus* (MRSA) in the Czech Republic	Czech Republic	2021	https://doi.org/10.1093/jac/dkaa404	11	ST5	PVL+	IVc

Prevalence and molecular characterization of *Staphylococcus aureus* in commercially available meat over a one-year period in Iowa, USA	United States of America	2017	http://dx.doi.org/10.1016/j.fm.2017.01.015	3	ST5	PVL+	Not mentioned
Epidemiological and molecular characterization of community and hospital acquired *Staphylococcus aureus* strains prevailing in Shenyang, North-eastern China	China	2013	http://dx.doi.org/10.1016/j.bjid.2013.02.007	6	ST5	PVL+	IV
Epidemiology of MRSA in southern Sweden: strong relation to foreign country of origin, health care abroad and foreign travel	Sweden	2013	https://doi.org/10.1007/s10096-013-1929-2	25	ST5	PVL+	Not mentioned
Molecular characterization of methicillin-resistant *Staphylococcus aureus* (MRSA) nasal colonization and infection isolates in a veterans affairs hospital	United States of America	2015	https://doi.org/10.1186/s13756-015-0048-5	6	ST5	PVL+	Not mentioned

Dissemination of the Methicillin-resistant *Staphylococcus aureus* paediatric clone (ST5-t002-IV-PVL+) as a major cause of community-associated staphylococcal infections in Bedouin children, southern Israel	Israel	2017	https://doi.org/10.1093/ofid/ofx163.1709	28	ST5	PVL+	IV
Genomic surveillance of methicillin-resistant *Staphylococcus aureus* in the Philippines, 2013–2014	The Phillipines	2021	https://doi.org/10.5365/wpsar.2020.11.1.004	2	ST5	PVL+	IV
Entry of Panton–Valentine leukocidin-positive methicillin-resistant *Staphylococcus aureus* into the hospital: prevalence and population structure in Heidelberg, Germany 2015–2018	Germany	2020	https://doi.org/10.1038/s41598-020-70112-z	4	ST5	PVL+	IV
Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of *Staphylococcus aureus* strains in a paediatric hospital	Italy	2018	https://doi.org/10.1186/s13073-018-0593-7	10	ST5	PVL+	IV

Title	Country	Year	DOI	No. of isolates in study	ST	PVL+/-	SCC*mec* type
Prevalence of oxacillin-susceptible methicillin-resistant *Staphylococcus aureus* nasal carriage and their clonal diversity among patients attending public health-care facilities	Brazil	2020	https://doi.org/10.4103/ijmm.IJMM_20_157	2	ST5	PVL+	IV
Distribution of sasX, pvl, and qacA/B genes in epidemic methicillin-resistant *Staphylococcus aureus* strains isolated from East China	China	2018	https://doi.org/10.2147/IDR.S153399	1	ST5	PVL+	IV

Part (b): Literature search results; partial hits (potential comparator isolates)

Title	Country	Year	DOI	No. of isolates in study	ST	PVL+/-	SCC*mec* type
Multidrug-resistant methicillin-resistant *Staphylococcus aureus* associated with bacteraemia and monocyte evasion, Rio de Janeiro, Brazil	Brazil	2021	https://doi.org/10.3201/eid2711.210097	167	ST5	Not mentioned	IV

Title	Country	Year	DOI	N	ST	PVL	SCCmec
Comparing pulsed-field gel electrophoresis with multi-locus sequence typing, *spa* typing, Staphylococcal Cassette Chromosome *mec* (SCCmec) typing, and PCR for Panton-Valentine leukocidin, arcA, and opp3 in methicillin-resistant *Staphylococcus aureus* isolates at a U.S. medical center	USA	2013	https://doi.org/10.1128/JCM.02429-12	1	ST5	PVL+	II
Molecular epidemiology of hospital-onset methicillin-resistant *Staphylococcus aureus* infections in Southern Chile	Chile	2013	https://doi.org/10.1007/s10096-013-1907-8	2	ST5	PVL-	IVc, IV non-typeable
Antimicrobial susceptibility and molecular typing of MRSA in Cystic Fibrosis	USA	2014	https://doi.org/10.1002/ppul.22815	42	ST5	PVL-	II
Suspected transmission and subsequent spread of MRSA from farmer to T dairy cows	Sweden	2018	https://doi.org/10.1016/j.vetmic.2018.09.017	12	ST2659 (differs from ST5 by the *yqil* gene)	PVL+	Not mentioned

Molecular characteristics of antimicrobial resistance and virulence determinants of *Staphylococcus aureus* isolates derived from clinical infection and food	China	2018	https://doi.org/10.1002/jcla.22456	29	ST5	Not clarified	III
Spread of invasive Spanish *Staphylococcus aureus spa-* type t067 associated with a high prevalence of the aminoglycoside-modifying enzyme gene *ant(4')-Ia* and the efflux pump genes msrA/msrB	Spain	2009	https://doi.org/10.1093/jac/dkn430	15	ST5	PVL-	IV

Result of an extensive literature search for PVL-positive CC5/ST5-MRSA-IVc Sri-Lankan clone isolates: (a) closest hits from search (b) partial hits from search. Search of the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA)/GenBank, the European Nucleotide Archive (ENA) databases and publicly available literature using the following search terms: 'ST5', 'CC5', 'spa type t002', 'PVL-positive', 'lukF-pv', 'lukS-pv', 'MRSA' and 'SCCmec IVc'. Table provides all available information extracted from papers. Most papers did not confirm SCCmec subtypes.

Abbreviations: ST, sequence type; PVL, Panton Valentine Leukocidin; +, positive; -, negative; SCCmec type, Staphylococcal Cassette Chromosome *mec*

Supplementary Table SV. Genotypic details for 36 related isolate groups (RIGs). Of the 266 (214 CC5/ST5-MRSA-IVc Sri-Lankan clone isolates, 29 PVL-positive and 23 PVL-negative CC5/ST5-MRSA-I/II/IVa/IVc/IVg/V comparator strains) study isolates, 123 (105 Sri-Lankan clone isolates & 18 comparator isolates) formed RIGs

RIG	Isolate (n)	Country (n)	Years of isolation	*spa*-ST-SCC*mec* (n)	Antimicrobial resistance genes[1] (n)	PVL(+/-)	IEC Type (n)
1	18	Ireland	2018–2021	t002-ST5-IVc	*blaZ* (17), *fosB* (18), *lmrP* (18), *mprF* (18), *sdrM* (18)	+	G
2	5	Ireland	2020–2022	t002-ST5-IVc	*blaZ* (5), *fosB* (5), *lmrP* (5), *mprF* (5), *sdrM* (5), *ermC* (4)	+	G
3	8	Denmark (3) Germany (1) Kuwait (1) Sri-Lanka (2) United Arab Emirates (1)	2013–2019	t002-ST5-IVc	*blaZ* (8), *fosB* (8), *lmrP* (8), *mprF* (8), *sdrM* (8), *ermC* (6), *tetK* (1)	+	G
4	3	Denmark	2018–2019	t002-ST5-IVc (2) t002-ST5-IVa (1)	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (2)	+	G
5	2	United Arab Emirates	2017	t002-ST5-IVc	*blaZ* (2), *ermC* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	G
6	3	Denmark	2017–2020	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (2)	+	G
7	2	Germany	2019	t535-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	G
8	2	United Arab Emirates	2018	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	G
9	2	Denmark	2014	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	F
10	3	Denmark	2015–2016	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3)	+	G
11	3	Norway (2) Germany (1)	2011–2013	t002-ST5-IVc (2) ND-ST5-IVc (1)	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (1), *tetK* (1)	+	G

12	2	Denmark	2012–2015	t002-ST5-IVc	*blaZ* (2), *fosB* (1), *lmrP* (1), *mprF* (2), *sdrM* (2), *ermC* (2)	+	G
13	2	Denmark	2015–2016	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2), *ermC* (2)	+	G
14	3	Norway (2) Denmark (1)	2008–2010	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (2), *tetK* (1)	+	G
15	4	Norway	2007–2010	t002-ST5-IVc	*blaZ* (4), *fosB* (4), *lmrP* (4), *mprF* (4), *sdrM* (4)	+	G
16	2	Norway	2007–2008	t002-ST5-IVc	*blaZ* (1), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2), *ermC* (1)	+	G
17	3	United Arab Emirates	2019	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3)	+	G
18	5	Denmark	2016–2017	t002-ST5-IVc	*blaZ* (5), *fosB* (5), *lmrP* (5), *mprF* (5), *sdrM* (5), *ermC* (5)	+	G
19	3	Denmark	2016	t002-ST5-IVc	*fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (2), *mupA* (3), *qacA* (3)	+	F
20	2	Denmark	2019	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	G
21	3	Denmark	2015	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3)	+	G
22	2	Germany	2012	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	Novel Type 2 (*sak, scn, sea, sep*)
23	2	Germany	2013–2014	t002-ST5-IVc	*blaZ* (1), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (1), *ermC* (2)	+	G
24	4	Denmark	2012–2013	t002-ST5-IVc	*blaZ* (4), *fosB* (4), *lmrP* (4), *mprF* (4), *sdrM* (4), *ermC* (3), *tetK* (2)	+	G
25	3	Norway	2020	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (3)	+	G

26	3	Sri-Lanka	2014	t002-ST5-IVc t010-ST5-IVc t045-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3)	+	G
27	4	Czech Republic	2018–2021	t002-ST5-IVc	*blaZ* (4), *fosB* (4), *lmrP* (4), *sdrM* (4), *mprF* (4)	+	G (3) Novel Type 3 (*sak, sep*)
28	2	Denmark	2018–2019	t002-ST5-IVc t002-ST5-IVa	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2)	+	G
29	2	Denmark	2018–2019	t002-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2), *ermC* (1)	+	G
30	3	Denmark	2017–2019	t002-ST5-IVc	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3)	+	G
*31	2	Norway	2005–2008	t311-ST5-IVc	*blaZ* (2), *fosB* (2), *lmrP* (2), *mprF* (2), *sdrM* (2), *aphA3* (2), *sat* (2)	+	A
*32	3	Ireland (2) Norway (1)	2009–2018	t002-ST5-IVa (2) t002-ST5-IVa (1)	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *ermC* (3)	+	G
*33	3	Denmark	2013–2015	t002-ST5-V	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *aac-aphD* (2)	+ (2) - (1)	G
*34	3	Denmark	2013–2014	t002-ST5-V	*blaZ* (3), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *aacA-aphD* (3)	+ (2) - (1)	G
*35	3	Ireland	2015–2019	t311-ST5-V	*blaZ* (2), *fosB* (3), *lmrP* (3), *mprF* (3), *sdrM* (3), *fusC* (3), *ermC* (2)	-	E
*36	4	Ireland	2017–2019	t311-ST5-V	*blaZ* (4), *fosB* (4), *lmrP* (4), *mprF* (4), *sdrM* (4), *fusC* (4), *ermC* (3)	-	E

266 study isolates (Sri-Lankan clone and comparator isolates) were recovered from 15 different countries between 2003–2022. Sri-Lankan clone isolates were recovered from 12 countries across Europe, Asia, Australia & the Middle East between 2005–2022. 123/266 isolates clustered with closely related neighbouring isolates to form related isolate groups (RIGs). RIGs were defined as groups of isolates exhibiting ≤20 cgMLST allelic differences to their

closest neighbouring isolate within the RIG. The remaining 143 study isolates exhibited ≥20 cgMLST allelic differences and were therefore considered to not be as closely related to one another. Genotypic information including sequence types (STs), staphylococcal protein A (*spa*) types, Staphylococcal Chromosome Cassette *mec* element (SCC*mec*) subtypes & antimicrobial resistance profiles were extracted from whole genome data using Ridom SeqSphere+ v7.0.4 (Ridom GmbH, Münster, Germany) and *S. aureus* Genotyping Kit 2.0 (Abbott) microarray profiling[1–4]. RIGs 1–36 were identified following core-genome multi-locus sequence type minimum spanning tree (cgMLST MST) analysis using SeqSphere+ (Fig. 2)[5,6]. *Asterisk indicates RIGs containing only comparator isolates. Abbreviations: RIG, Related isolate group; n, number; ND, not determined – isolates not available & *spa* types could not be determined using in-silico techniques on the available genomic sequence data; ST, sequence type; SCC*mec*, staphylococcal chromosomal cassette harbouring *mecA*; PVL, Pantone-Valentine leukocidin; +, positive; -, negative; IEC, immune evasion cluster; s*ep*, staphylococcal enterotoxin p gene.

References:

[1] Monecke S, Jatzwauk L, Weber S, Slickers P, Ehricht R. DNA microarray-based genotyping of methicillin-resistant *Staphylococcus aureus* strains from Eastern Saxony. Clin Microbiol Infect 2008;14:534-45 https://doi.org/10.1111/j.1469-0691.2008.01986.x

[2] Monecke S, Slickers P, Ehricht R. Assignment of *Staphylococcus aureus* isolates to clonal complexes based on microarray analysis and pattern recognition. FEMS Immunol Med Microbiol 2008;53:237-51 https://doi.org/10.1111/j.1574-695X.2008.00426.x

[3] Monecke S, Jatzwauk L, Müller E, Nitschke H, Pfohl K, Slickers P, et al. Diversity of SCC*mec* elements in *Staphylococcus aureus* as observed in South-Eastern Germany. PLoS One 2016;11:0162654. https://doi.org/10.1371/journal.pone.0162654

[4] Earls MR, Steinig EJ, Monecke S, Samaniego Castruita JA, Simbeck A, Schneider-Brachert W, et al. Exploring the evolution and epidemiology of European CC1-MRSA-IV: tracking a multidrug-resistant community-associated meticillin-resistant *Staphylococcus aureus* clone. Microb Genom 2021;7:000601. https://doi.org/10.1099/mgen.0.000601

[5] Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A. Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. J Clin Microbiol 2014;52:2365-70. https://doi.org/10.1128/JCM.00262-14

[6] Earls MR, Coleman DC, Brennan GI, Fleming T, Monecke S, Slickers P, et al. Intra-hospital, inter-hospital and intercontinental spread of ST78 MRSA from two neonatal intensive care unit outbreaks established using whole-genome sequencing. Front Microbiol 2018;9:1485. https://doi.org/10.3389/fmicb.2018.01485