Monitoring meticillin resistant *Staphylococcus aureus* and its spread in Copenhagen, Denmark, 2013, through routine whole genome sequencing

M D Bartels (mette.damkjaer@dadlnet.dk)[1,2], H Larner-Svensson[2,3], H Meiniche[3], K Kristoffersen[3], K Schønning[1,4], J B Nielsen[1,3], S M Rohde[3], L B Christensen[3], A W Skibsted[3], J O Jarløv[5], H K Johansen[6], L P Andersen[7], I S Petersen[5], D W Crook[8], R Bowden[9], K Boye[1], P Worning[3], H Westh[1,3,4]

1. Department of Clinical Microbiology, Hvidovre University Hospital, Denmark
2. These authors contributed equally to this work
3. MRSA Knowledge Center, Hvidovre University Hospital, Denmark
4. Institute of Clinical Medicine, University of Copenhagen, Denmark
5. Department of Clinical Microbiology, Herlev University Hospital, Denmark
6. Department of Clinical Microbiology, Rigshospitalet, Denmark
7. Department of Infection control, Rigshospitalet, Denmark
8. Infectious Diseases and Microbiology Unit John Radcliffe Hospital, Oxford, United Kingdom
9. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

Typing of meticillin resistant *Staphylococcus aureus* (MRSA) by whole genome sequencing (WGS) is performed routinely in Copenhagen since January 2013. We describe the relatedness, based on WGS data and epidemiological data, of 341 MRSA isolates. These comprised all MRSA (n = 300) identified in Copenhagen in the first five months of 2013. Moreover, because MRSA of staphylococcal protein A (*spa*)-type 304 (t304), sequence type (ST) 6 had been associated with a continuous neonatal ward outbreak in Copenhagen starting in 2011, 41 t304 isolates collected in the city between 2010 and 2012 were also included. Isolates from 2013 found to be of t304, ST6 (n=14) were compared to the 41 earlier isolates. In the study, isolates of clonal complex (CC) 22 were examined in detail, as this CC has been shown to include the hospital-acquired epidemic MRSA (EMRSA-15) clone. Finally, all MRSA ST80 were also further analysed, as representatives of an important community-acquired MRSA in Europe. Overall the analysis identified 85 *spa*-types and 35 STs from 17 CCs. WGS confirmed the relatedness of epidemiologically linked t304 neonatal outbreak isolates. Several non-outbreak related patients had isolates closely related to the neonatal isolates suggesting unrecognised community chains of transmission and insufficient epidemiological data. Only four CC22 isolates were related to EMRSA-15. No community spread was observed among the 13 ST80 isolates. WGS successfully replaced conventional typing and added information to epidemiological surveillance. Creation of a MRSA database allows clustering of isolates based on single nucleotide polymorphism (SNP) calling and has improved our understanding of MRSA transmission.

Introduction

Whole genome sequencing (WGS) is expected to transform the practice of clinical microbiology and infection control [1,2]. Advances in WGS technology and extended multiplexing on desktop-based WGS machines has reduced sequencing cost to approximately EUR 100 per genome. While detailed bioinformatics analysis remains a challenge, commercial software and web-based solutions are now available for performing multilocus sequence typing and screening for gene-based antibiotic resistance on the WGS data [3].

The sequence-based typing method staphylococcal protein A (*spa*)-typing of meticillin-resistant *Staphylococcus aureus* (MRSA) has been performed since 2003 at the MRSA Knowledge Center, Department of Clinical Microbiology, at Hvidovre Hospital [4]. In January 2013, this Sanger sequencing method was replaced by WGS of all MRSA isolates and we routinely produce 24 full genomes twice a week. From WGS data, script-based bioinformatics programmes are used to identify *mec*A, *mec*C, *nuc, ccr* and Panton–Valentine leukocidin (PVL) genes as well as the arginine catabolic mobile element (ACME) genes (*arc*A to *arc*D). Bioinformatics is also used to determine direct repeat units *(dru)* types and multilocus sequence types (MLST). Furthermore, single nucleotide polymorphism (SNP) analysis is used routinely to compare relatedness of MRSA isolates.

Clonal complex	Isolates N	Sequence types (N)	*Spa*-types t (N)	Panton–Valentine leukocidin N (%)	Obtained by screening N (%)
1	20	ST1 (12); ST1_SLV (1); ST772 (6); ST2689 (1)	t127 (10); t177 (1); t345 (1); t386 (1); t657(4); t5388 (1); t5608 (1); t7358 (1)	7 (35)	8 (40)
5	45	ST5 (30); ST5_SLV (1); ST105 (1); ST149 (6); ST149_SLV (2); ST225 (2); ST1457 (1); ST2626 (2)	t002 (33); t003 (2); t062 (3); t067 (1); t105 (1); t442 (1); t579 (2); t688 (1); t854 (1)	6 (13)	27 (60)
6	56[b]	ST6	t304 (55); t104 (1)	0 (0)	40 (71)
8	59	ST8	t008 (28); t024 (17); t118 (1); t121 (1); t304 (3); t801 (1); t596 (1); t1476 (1); t1774 (2); t2849 (1); t3240 (1); t6197 (1); t13119 (1)b	26 (44)	30 (51)
22	29	ST22 (20); ST22_SLV (1); ST1327 (6); ST2371 (2)	t005 (8); t022 (1); t032 (1)d; t223 (6); t670 (1); t852 (5); t1328 (1); t2834 (1); t4326 (2); t11808 (2); t12093 (1)	10 (34)	19 (66)
30	42	ST30 (36); ST36 (1); ST1456 (4)	t019 (29); t021 (4); t253 (2); t685 (1); t964 (1); t1133 (2); t1752 (2); t2872 (1)	38 (90)	18 (43)
45	13	ST45 (9); ST46 (2); ST508 (1)	t015 (4); t026 (1); t050 (1); t069 (1); t116 (1); t230 (1); t583 (1); t728 (1); t2988 (2)	0 (0)	7 (54)
59	9	ST59 (8); ST59_SLV (1)	t437 (8); t3513 (1)	6 (67)	5 (56)
72	18	ST72 (16); ST72_SLV (2)	t148 (5); t324 (8); t791 (3); t2409 (1); t11917 (1)	2 (11)	14 (78)
80	17	ST80 (13); ST80_SLV (4)	t044 (11); t131 (1); t376 (3); t1028 (2)	17 (100)	14 (82)
88	10	ST88 (9); ST859 (1)	t325 (1): t690 (2); t729 (1); t1855 (1); t5147 (5)	2 (20)	2 (20)
97	1	ST97	t267	0 (0)	0 (0)
130	1	ST130	t1048	0 (0)	0 (0)
152	2	ST152	t355 (1); t454 (1)	1 (50)	1 (50)
228	1	ST111	t041	0 (0)	0 (0)
239	1	ST239	t138	0 (0)	1 (100)
398	5	ST398	t011 (1); t034 (4)	0 (0)	2 (40)
Not known	6	ST140 (2); ST1245 (2); ST1633_SLV; ST1943 (1)	t296 (2); t355 (1); t843 (2); t978 (1)	1 (17)	4 (67)
MLST[c] failure	7	MLST failure[c]	t024 (2); t062 (2); t304 (1); t437 (1)	1 (14)	2 (29)
Total	341	–	–	117 (34)	194 (57)

MLST: multilocus sequence type; SLV: single locus variant; *Spa*-type: staphylococcal protein A-type; ST: sequence type.

[a] Whole genome sequencing was conducted in 2013 on the 341 isolates. Of these, 300 had been obtained in 2013, while 41 were selected t304, ST6 isolates from the years 2010 to 2012.

[b] 41 of these isolates were from the years 2010 to 2012.

[c] MLST failure: In seven cases a ST could not be obtained due to a gene being in two contigs, the identification of a new allele or a new combination of alleles leading to a novel ST not listed in the MLST database.

[d] This isolate was included as a ST22 isolate having six of the alleles in ST22, but a missing *glpF* gene.

In this study, the overall relatedness of all consecutive MRSA isolates from the first five months of 2013 (n=300) in Copenhagen, Denmark, was assessed by SNP analysis. In addition, 41 isolates of *spa*-type (t) 304 (t304) from 2010 to 2012 were whole genome sequenced, and compared with 14 of the isolates from 2013, that were determined to be t304, ST6. A continuous outbreak has been caused by t304, ST6 in neonatal wards in Copenhagen since 2011, and the t304 isolates from prior to 2013 were included to study the t304, ST6 isolates both before and during the outbreak period in more detail. All clonal complex (CC) 22 isolates were selected for detailed analysis as this CC contains the globally important international hospital-acquired MRSA (HA-MRSA) clone epidemic MRSA-15 (EMRSA-15) [5]. Finally, all ST80 isolates (European clone) were described in more detail as representatives of an important community-acquired MRSA (CA-MRSA) in Europe [6].

Phylogenetic tree of all meticillin resistant *Staphylococcus aureus* isolates found in Copenhagen, Denmark, January–May 2013 (n=300)
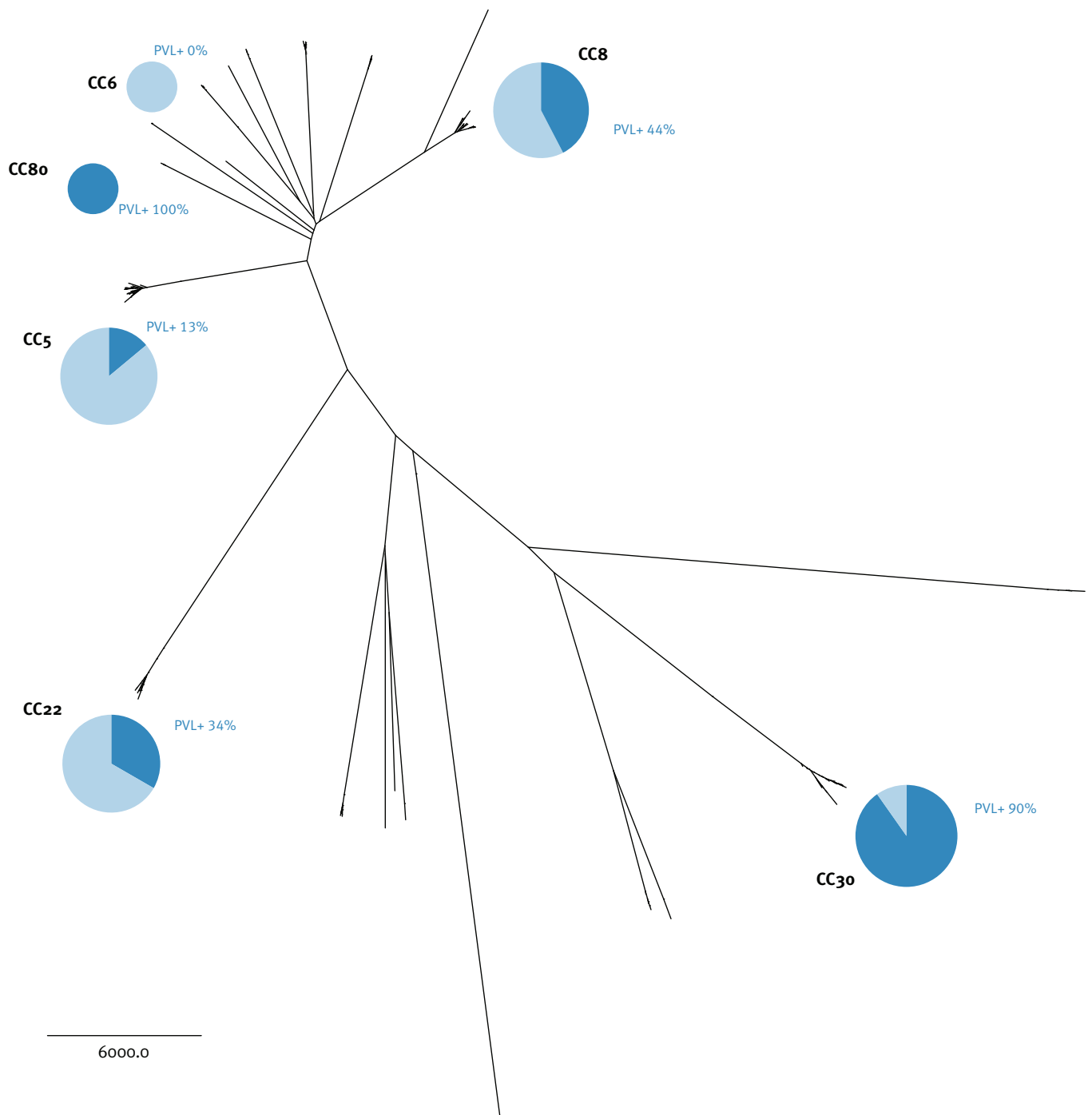
CC: clonal complex; PVL: Panton–Valentine leukocidin.

Six of the larger clonal complexes (CC) are indicated at the end of branches, as well as the percentage of PVL-positive isolates within each CC. The length of the scale bar corresponds to 6,000 single nucleotide polymorphisms (SNPs).

Unrooted neighbour joining tree of meticillin resistant *Staphylococcus aureus* isolates of staphylococcal protein A-type 304, sequence type 6, Copenhagen, Denmark, 2010–2013 (n=55)
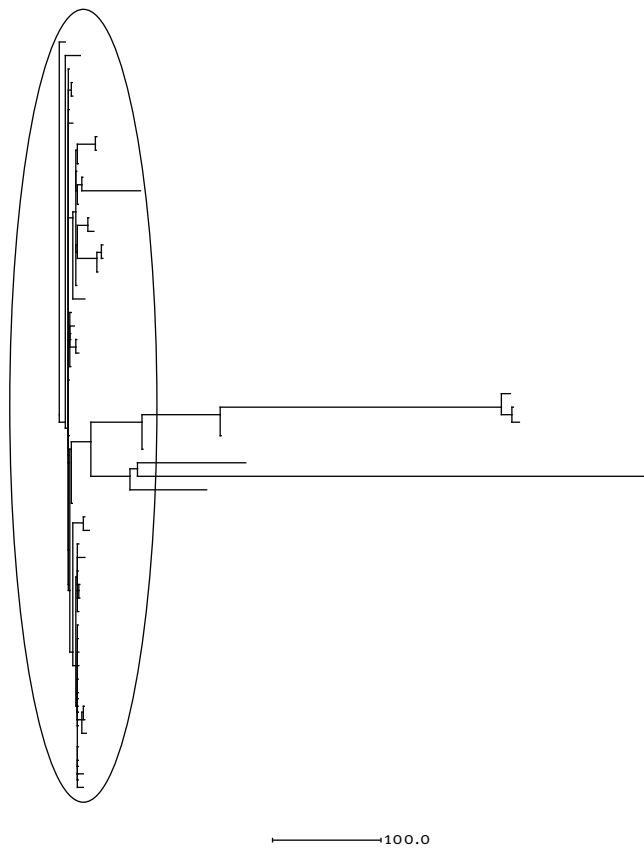
The circle marks 49 isolates that include the 34 neonatal outbreak isolates and 15 isolates with between four and 36 single nucleotide polymorphisms (SNPs) to the closest neonatal isolates. These 49 isolates are further presented in Figure 3. The length of the scale bar corresponds to 100 SNPs.

Methods

Setting

The Capital Region of Denmark manages 12 hospitals and the general practice healthcare services of the region. This study covers the region's 1.73 million inhabitants. MRSA isolates are from individual patients. They are identified in clinical or screening samples submitted to the three Departments of Clinical Microbiology in the Capital Region. WGS is only performed at Hvidovre Hospital and all MRSA isolates from the two other Departments of Clinical Microbiology are sent to Hvidovre Hospital for WGS.

Isolates and patient data

Each cultured isolate is immediately confirmed to be MRSA with an in-house multiplex real-time polymerase chain reaction (PCR) that can detect the presence of

*nuc, mec*A and *mec*C (data not shown). In this study, a total of 341 MRSA-confirmed isolates were whole genome sequenced. Three-hundred consecutive isolates were collected between 1 January 2013 and 31 May 2013, while 41 isolates of t304, ST6 were collected in the period from 2010 to 2012. The latter 41 isolates were originally *spa*-typed by Sanger sequencing and were whole genome sequenced in 2013.

According to the national MRSA guidelines all patients seen in a hospital are asked for MRSA risk factors. These include hospitalisation abroad within the last six months, previous MRSA positivity, contact to a MRSA positive person and contact with pigs within the last six months. In our internal microbiology database, a case report form is generated for each patient. The report contains information on date of sampling, submitting hospital unit or general practitioner, sample site, known MRSA contacts, household relations and antibiogram of the isolate. Sample types were categorised as being either from infection or screening. Screening samples were from hospitalised patients having risk factors as mentioned above, household contacts of MRSA patients and/or outbreak investigations in hospitals and nursing homes.

Whole genome sequencing procedures

Each confirmed MRSA was whole genome sequenced on a MiSeq (Illumina, United States (US)). The current workload is a four-day set-up. DNA concentrations were normalised using a Qubit (Invitrogen, United Kingdom (UK)). Libraries were made with Nextera XT DNA sample preparation kit (Illumina, US), genomes multiplexed to 24 isolates per run and sequenced with 2 x 150 bp paired-end reads. For the detection of single nucleotide variants relative to the reference, we used a reference-based mapping approach via Stampy (Lunter G). Reads were mapped to a USA300 reference sequence (US300_TCH1516) using Stampy v1.0.11 with no BWA pre-mapping and an expected substitution rate of 0.01 [7]. Variants were called using SAMtools v0.1.12 mpileup command with options −M0 -Q30 −q30 −o40 −e20 −h100 −m2 −D −S. The genome was assembled using Velvet v1.0.11 [8], with hash (kmer) size and coverage parameters optimised to give the highest number of bases in contigs with length greater than 1 kb. In cases were a *spa*-type could not be generated, WGS was repeated.

Staphylococcal protein A-types, sequence types and Panton–Valentine leukocidin genes

An additional script has been developed in-house to analyse MRSA genomes for *mec*A, *mec*C, *nuc, ccr* genes, *spa*-type, MLST, *dru* types and PVL. *Spa*-types and ST were called from the assembled contigs by comparison to the published types on the SpaServer [9] and the *S. aureus* MLST.net database [10]. A CC was assigned based on the MLST.net webpage. Each isolate was reported in the Laboratory Information System with confirmation of being an MRSA, *spa*-type, ST and presence or absence of PVL genes.

Unrooted neighbour joining tree of meticillin resistant *Staphylococcus aureus* isolates of staphylococcal protein A-type 304, sequence type 6, that were implicated in a neonatal outbreak (n=34) as well as closely related isolates (n=15), Copenhagen, Denmark, 2010–2013

5 SNPs
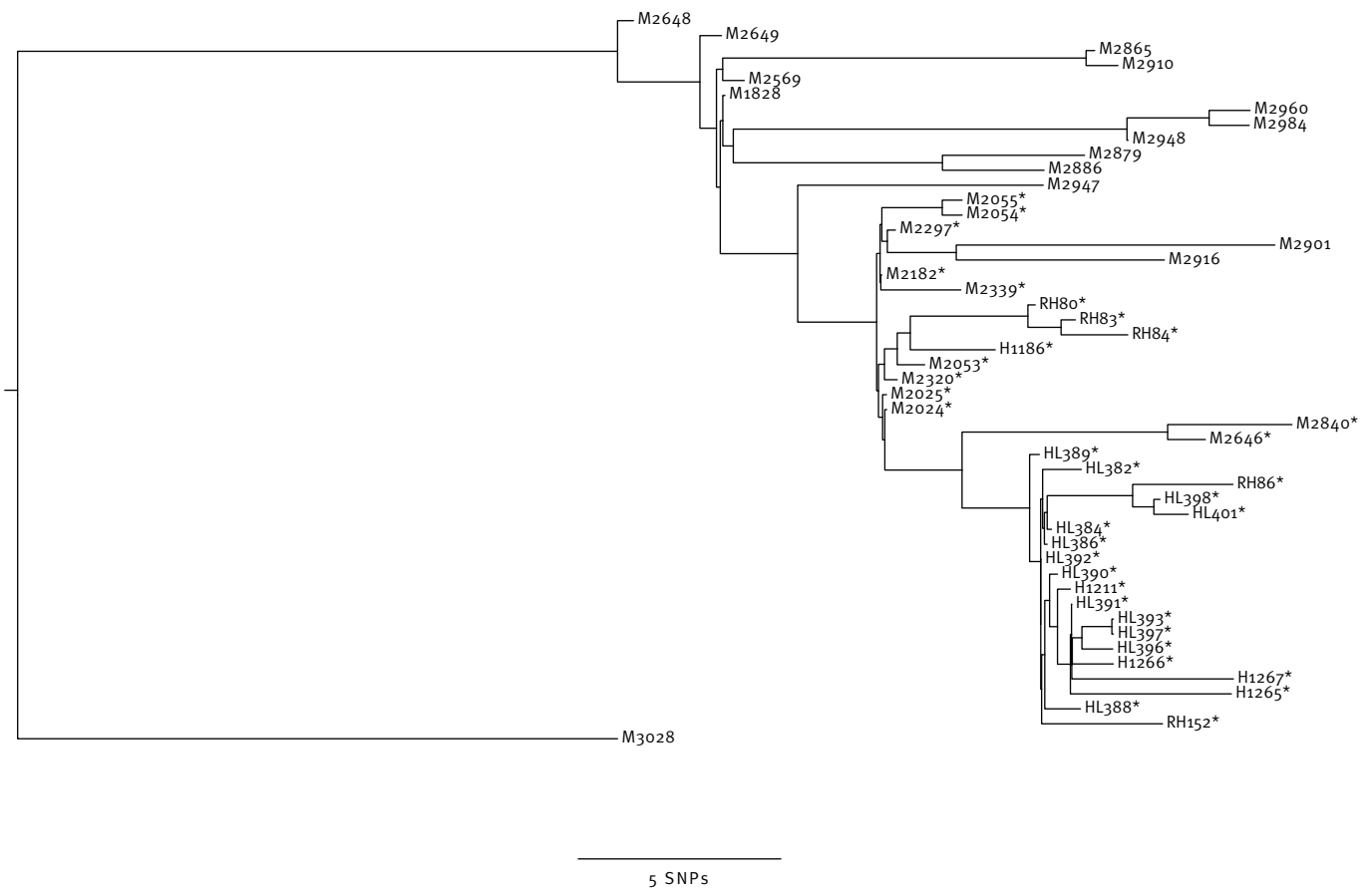
The isolates depicted in the tree are those marked in Figure 2. The numbers at the end of branches are the internal isolate numbers. The 34 isolates epidemiologically related to the neonatal outbreak are marked with asterisks (*). The length of the scale bar corresponds to five single nucleotide polymorphisms (SNPs).

Phylogenetic analyses

Phylogenetic analyses were performed using a distance method, based on pairwise nucleotide sequence alignments for the *S. aureus* core genome, as defined by USA300, among all strains. Phylogeny was inferred by neighbour-joining analysis as implemented in SplitsTree v4.11.3 [11]. Tree drawing was managed with FigTree [12].
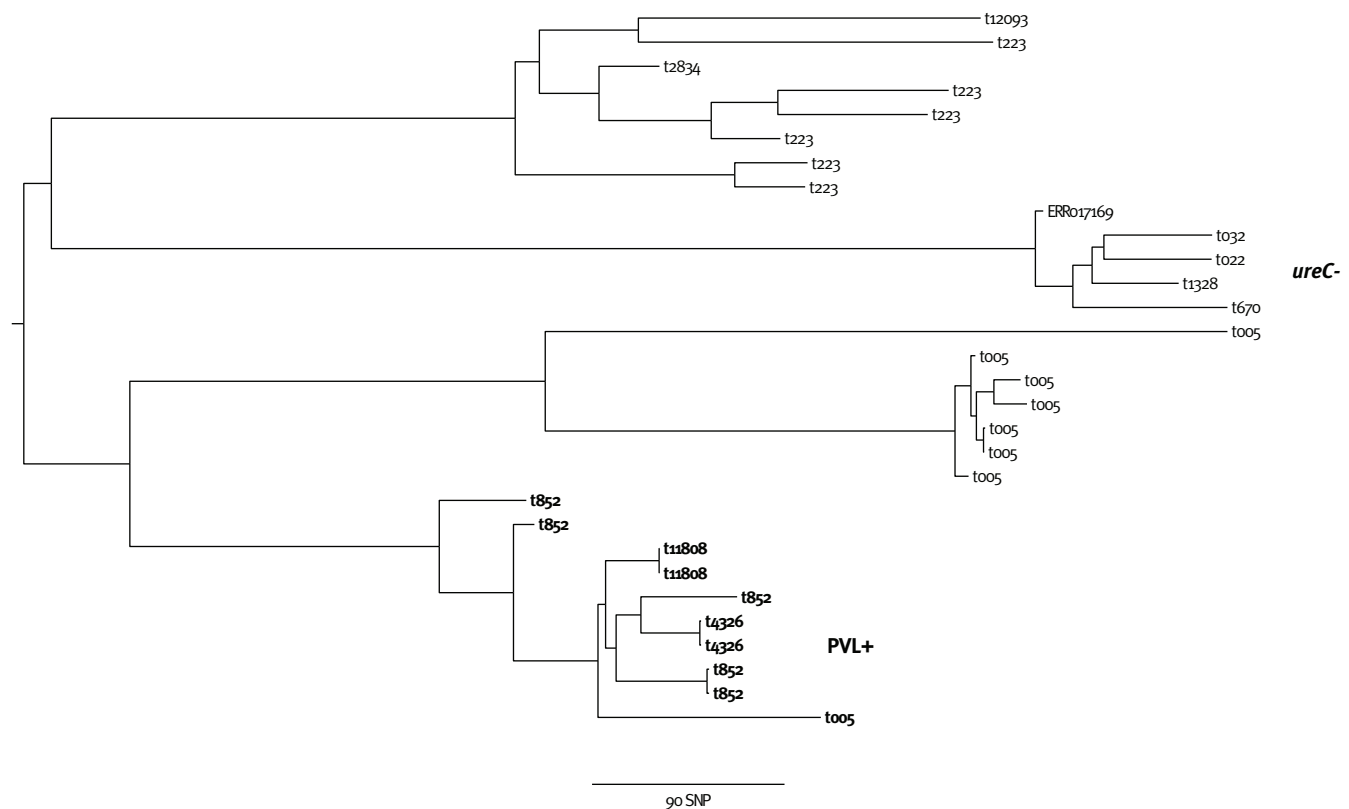
Staphylococcal protein A-type 304 analysis

To illustrate the value of WGS in outbreak investigation 14 t304 isolates obtained in 2013 and 41 t304 isolates from 2010 to 2012 were sequenced. t304 has been associated with a neonatal intensive care unit outbreak spreading to several hospitals in Copenhagen mainly in 2011 and 2012. Both isolates from persons with and without known contact to the neonatal units were included. Epidemiological data for all persons were registered and compared with the findings in the SNP analysis.

Analysis of sequence type 80 and clonal complex 22 isolates

To study the import and spread of international well known MRSA clones, we analysed in depth CC22 isolates and their relationship to a representative of the international, HA-MRSA clone, EMRSA-15. The reference isolate ERR017169 of ST22-A2 was included as a quality control (QC) isolate for EMRSA-15 [5]. All CC22 isolates were also screened for a 1bp deletion in the *ure*C gene, an identifier of EMRSA-15 according to a recently published paper [5]. The European clone (ST80, PVL positive) was studied more in detail as a representative of a typical CA-MRSA. This clone is known to be very homogeneous and therefore a QC isolate was not included [13]. Epidemiological data for all persons were registered and the results of the analysis of these data were compared with the findings in the SNP analysis.

Unrooted neighbour joining tree with the ERR017169 strain representative of epidemic meticillin resistant *Staphylococcus aureus* (EMRSA)-15 and 29 MRSA isolates of clonal complex 22, Copenhagen, Denmark, January–May 2013

PVL: Panton–Valentine leukocidin; SNP: single nucleotide polymorphisms.

For each isolate the staphylococcal protein A-type (t) is displayed at the end of branches, and the 10 PVL-positive isolates are in bold. Isolates with a 1 bp deletion in the *ureC* gene characteristic of EMRSA-15 are also indicated. The length of the scale bar corresponds to 90 SNPs.

Results

Overall relatedness of the isolates

A total of 194 (57%) of the 341 isolates were obtained through MRSA screening (Table). *mec*A was found in 337 isolates while four isolates contained *mec*C. Based on WGS, a *spa*-type could be assigned for all 341 isolates. Of these, 15 obtained a *spa*-type after repeating WGS. A ST was generated by our script in 334 of 341 isolates (98%). A lacking ST (7 isolates) was caused by a gene being in two contigs, the identification of a new allele, or a new combination of alleles leading to a novel ST not listed in the MLST database. Eighty-five *spa*-types and 35 STs as well as eight single locus variants were identified (Table). Of the 334 isolates with a ST, 328 fell into 17 CCs. The CC for the remaining six isolates with an identified ST could not be determined using the MLST database.

Most of the 300 isolates from 2013 belonged to CC 8 (59 isolates), followed by CC5 (45 isolates) and CC30

(42 isolates). In the first five months of 2013, PVL genes were present in 117 of the 300 isolates (39%) and in 33 *spa*-types and 16 STs. Among the international well known MRSA clones, we found isolates that, based on *spa*-type, MLST/CC and PVL-positivity were similar to US300 (t008, ST8, PVL+; n=28), South-West Pacific clone (t019, ST30, PVL+; n=29), European clone (t044, ST80, PVL+; n=7), ST239 and EMRSA-15 (ST22; n=20). Only five of the 300 isolates belonged to the livestock associated ST398.

The neighbour joining tree of the 300 MRSA isolates from 2013 (Figure 1) reveals a picture of high genetic diversity of MRSA in Copenhagen. Major clones (clades) are easily seen as clusters at the end of branches (Figure 1). The larger CCs are marked, as is the amount of PVL-positive isolates within these CCs. Within each CC further discrimination can be obtained by comparing the isolates in an unrooted neighbour joining tree through SNP calling.

Unrooted neighbour joining tree of meticillin resistant *Staphylococcus aureus* isolates of sequence type 80 found in Copenhagen, Denmark, January–May 2013 (n=13)
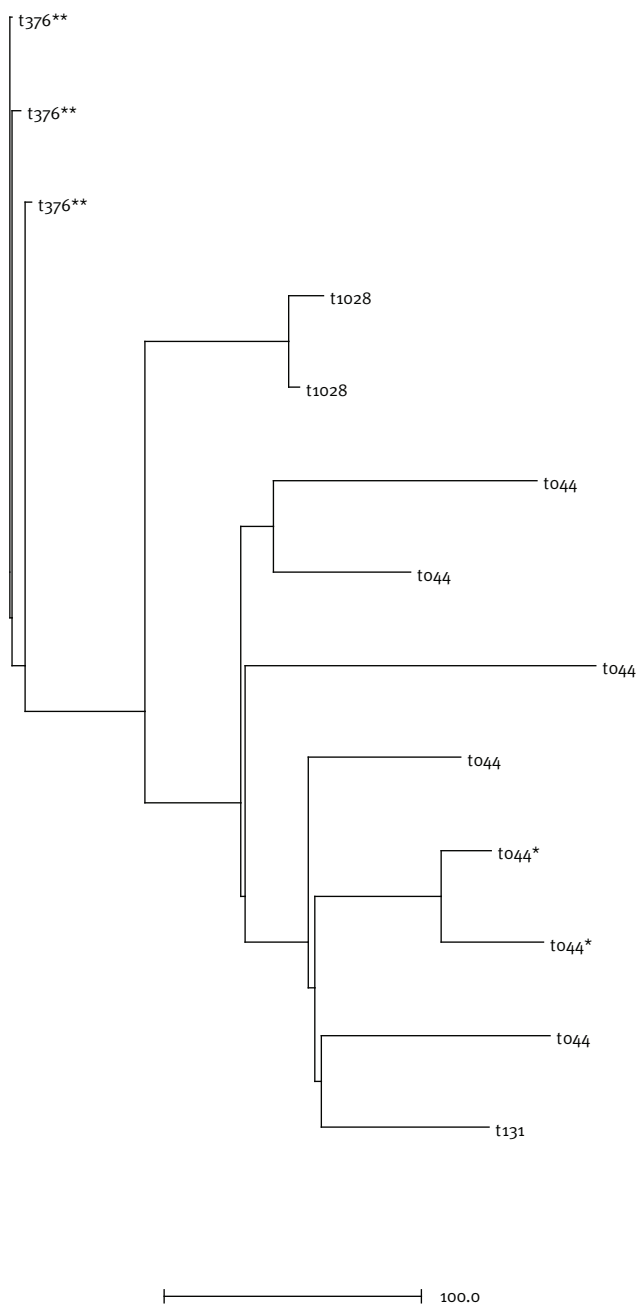
*Same household, 59 SNPs between the two isolates.
**Father and two daughters.

SNPs: single nucleotide polymorphisms.

The staphylococcal protein A-type (t) of each isolate is displayed at the end of branches. The length of the scale bar corresponds to 100 SNPs.

Staphylococcal protein A-type 304 analysis

Healthcare-associated MRSA outbreaks are rare in Denmark, and the only clusters identified as HA-MRSA were t304, ST6, PVL-negative, associated with a neonatal outbreak spreading in neonatal wards in the Capital Region (Figure 2 and Figure 3) and t024, ST8, PVL-negative, a clone that has been healthcare associated since 2003 [14]. t304, ST6 in neonates and their relatives were, apart from one case in February 2013, all identified either in the last seven months of 2011 (n = 7) or in 2012 (n = 26) . Comparing the 55 t304, ST6 isolates from Copenhagen in the period from 2010 to 2013, the epidemiologically related outbreak isolates (n = 34) were genetically closely related, having from 1 to 10 SNPs compared with the closest related outbreak isolates. t304 isolates from the same neonatal ward were in most cases even closer related with a maximum of five SNPs separating them. The t304 patients who were not epidemiologically related to the outbreak (n = 21), were in six cases clearly different from the outbreak strains with between 67 and more than 200 SNPs separating them from the outbreak strains. Seven isolates had between 11 to 36 SNPs compared with the closest related of the outbreak isolates, while eight isolates only had from four to ten SNPs (Figure 2 and Figure 3).

Clonal complex 22 analysis

An expansion of the CC22 isolate tree is shown in Figure 4 and compared with a ST22 isolate representative of EMRSA-15 [5]. The 29 Danish CC22 isolates are diverse and include ten different *spa*-types. Ten of these isolates were from patients living in four different households. Based on SNP calling, the clusters of these isolates within households differed between 0 to 14 SNPs, while epidemiologically unrelated isolates differed at between 69 to 1,207 SNPs. PVL was found in ten of the 29 isolates (32%) including all t852, t4326, t11808 and one of eight t005 isolates. Four of the 29 isolates had the 1 bp deletion in *ure*C described as typical for EMRSA-15 and were the isolates closest related to the English prototype ST22 isolate ERR017169 based on SNP calling. These four patients had been hospitalised before the finding of MRSA, two of them in Italy and Spain and two in Denmark. Only one of the remaining 25 CC22 patients had been hospitalised before the finding of MRSA. Nineteen of the 29 patients were not of ethnic Danish origin, and one of the five ethnic Danes had recently travelled in India. We assume that most of these CC22 isolates were acquired through international travel.

Sequence type 80 analysis

A total of 13 isolates belonging to the European clone are presented in Figure 5. They were all PVL-positive and represented four different *spa*-types of which t044 accounted for seven isolates. Two of the t044 isolates differing with 59 SNPs were from two persons living in a three-generation household. SNP analysis revealed that the 59 SNPs were dispersed throughout the genome, and could therefore not be explained by a single recombination event. The rest of the t044 isolates were from patients not epidemiologically related, and they differed with 130 to 365 SNPs. There were three t376 isolates, which were from a father and his

two daughters and which had a maximum of 10 SNPs between them. Two isolates of t1028 had only 17 SNPs between them, but we have not identified an epidemiological link between these patients.

Discussion

The emergence of predominantly CA-MRSA was first recognised in Denmark in 2003 [14]. Since 2003 the number of new MRSA patients in Copenhagen has increased 20-fold from 33 to 674 cases in 2012 [4]. The genetic diversity that was already seen in the first years of the MRSA epidemic has increased over the years, and in 2012, 142 different *spa*-types were found in Copenhagen (data not shown). This diversity is mainly related to travel and tourism e.g. cases where Danish families visit relatives in countries where MRSA is prevalent in the community, and acquisition in persons working with healthcare abroad [15]. The high diversity based on *spa*-types indicates a low level of endemic transmission, and could be further differentiated in this study, based on WGS. Cases imported through hospitalisation abroad are routinely found through screening of these patients when they are transferred to Danish hospitals. In general, MRSA is first isolated in clinical samples and subsequent ring screening usually identifies carriers in their households. However, there is an under-identification of MRSA as a number of patients have had recurrent skin and soft tissue infections before they are sampled. This means that some patients who only have a single infection or short-term carriage state and then clear MRSA are not identified.

Livestock associated MRSA, ST398, accounted for less than 2% of MRSA in Copenhagen, a capital city surrounded by a large suburban area with little livestock production. However, ST398 was the most common clone in the whole of Denmark in 2012 accounting for 232 of 1,556 isolates (15%), predominantly seen in regions with a large pig production [4].

Two UK studies on neonatal MRSA outbreaks have confirmed the usefulness of WGS in outbreak investigations, and showed that the number of SNPs between outbreak related isolates were less than 15 [16,17], except in one case where a mutation in the *mutS* gene was explained as the reason for more divergence. Furthermore, Harris et al. [16] found additionally ten isolates that were initially not considered part of a particular outbreak, but were closely related to outbreak strains by WGS, and retrospective epidemiological links were found in most cases. In our study, WGS could confirm the relatedness between 34 t304 isolates that were epidemiologically related to the neonatal outbreak in Copenhagen. Six isolates were clearly different from the outbreak strains, and this was also consistent with patient data. Fifteen isolates, that were not considered outbreak-related based on epidemiological data, differed in eight cases with only four to 10 SNPs to the closest related outbreak strains, and in seven cases with 11 to 36 SNPs (Figure 3). Since 11 of these 15 isolates were obtained in 2013, we hypothesise that a community transmission chain might exist, but that the epidemiological link between patients is missing. WGS seems promising in outbreak investigations, and might add information on the spread of outbreak clones outside the hospital setting, where the connection between patients is frequently not clear.

The CC22 isolates found were diverse and only four isolates were similar to the English prototype CC22, EMRSA-15. Two of these four isolates were imported from Spain and Italy. In 2012, t032, CC22 accounted for 89 isolates and was the fifth most common MRSA type in Denmark representing almost 6% of new MRSA cases [4], but only seven of these 89 isolates were from the Capital Region [18].

The typical CA-MRSA ST80 was only found in 13 patients in the five month period. There is evidence of repeated introduction of this clone into Denmark with only one case of possible community spread, t1028, based on the number of SNPs (Figure 5). Although there is no consensus on how many SNPs define whether isolates are related or unrelated [19], it was surprising that two household isolates of ST80 had as many as 59 SNPs between them. A closer look at this household revealed that this family of three generations immigrated to Denmark five years prior, and presumably have been MRSA carriers for many years. According to Holden et al. [5], SNPs are estimated to occur in the core genome at a rate of 1.3×10^{-6} substitutions per nucleotide per year or three to four SNPs per year. Based on the relatively few closely related isolates in our study, we wonder if we will have a congruent SNP evolution for all MRSA clones.

Our policy until the end of 2012 has been to *spa*-type all MRSA using differences in *spa*-types to delineate different clones e.g. t024 different from t008 [20]. However, when less common *spa*-types were observed the link to international clones was often not seen. For example, the rare t1133, ST1456, PVL-positive, with only 11 isolates in the SpaServer database, was not easily associated to the South-West Pacific clone (t019, ST30, PVL+) before the neighbour joining tree clustered them together. Currently, the *spa*-type, ST and absence/presence of PVL are routinely found from the laboratory information system. The cost of WGS is steadily going down and today it would be more expensive for us to perform traditional Sanger sequencing for *spa*-typing and MLST. Furthermore, the WGS platform is universally applicable and is now routinely used in other healthcare-associated outbreak investigations.

A disadvantage is the slower turnaround time in our WGS-setup before typing results can be delivered. WGS is run in batches of 24 isolates, typically twice weekly. This means that from the date of sample to a WGS typing result can be up to 12 days or 4–8 days after the isolate is received for WGS. The current MLST *S. aureus* database contains fewer isolates and types than the *spa* server database and this leads to a number of

isolates only having six known alleles and one novel allele. In these cases our software calls all the STs with six matches, and we manually assign the ST as a single locus variant (SLV) of the most prevalent ST.

The data presented in this study show that after building up a local MRSA whole genome database, real time WGS gives additional information to the evolution and spread of MRSA. The method can be used to investigate outbreaks, but epidemiological data still need to be included. More insight into how to interpret the relatedness between isolates based on SNP analysis is needed as well as improved software designed for persons without bioinformatics skills.

Conflict of interest

None declared.

Authors' contributions

MDB, HLS and HW: Contributed substantially to the design, the analysis and the interpretation of data. Drafted the work and revised it for important intellectual content. HM, KK, KS, JBN, SMR, LBC, AJWS, JOJ, HKJ, LPA, ISP, DWC, RB, KB, PW: Contributed substantially to the interpretation of data. Revised the work critically for important intellectual content.

References

1. Le VT, Diep BA. Selected insights from application of whole-genome sequencing for outbreak investigations. Curr Opin Crit Care. 2013;19(5):432-9. http://dx.doi.org/10.1097/MCC.0b013e3283636b8c PMID:23856896

2. Reuter S, Ellington MJ, Cartwright EJ, Köser CU, Török ME, Gouliouris T, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. JAMA Intern Med. 2013;173(15):1397-404. http://dx.doi.org/10.1001/jamainternmed.2013.7734 PMID:23857503

3. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640-4. http://dx.doi.org/10.1093/jac/dks261 PMID:22782487

4. The Danish Integrated Antimicrobial Resistance Monitoring and Research Programme (DANMAP) Reports. [Accessed 5 Dec 2013]. Available from: http: www.danmap.org/Downloads/Reports.aspx

5. Holden MT, Hsu LY, Kurt K, Weinert LA, Mather AE, Harris SR, et al. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic. Genome Res. 2013;23(4):653-64. http://dx.doi.org/10.1101/gr.147710.112 PMID:23299977

6. Otter JA, French GL. Molecular epidemiology of community-associated meticillin-resistant Staphylococcus aureus in Europe. Lancet Infect Dis. 2010;10(4):227-39. http://dx.doi.org/10.1016/S1473-3099(10)70053-0 PMID:20334846

7. Didelot X, Bowden R, Wilson DJ, Peto TE, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. Nat Rev Genet. 2012;13(9):601-12. http://dx.doi.org/10.1038/nrg3226 PMID:22868263

8. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18(5):821-9. http://dx.doi.org/10.1101/gr.074492.107 PMID:18349386

9. Ridom SpaServer. [Accessed 5 Dec 2013]. Available from: http: www.spaserver.ridom.de/

10. Multi Locus Sequence Typing (MLST). [Accessed 5 Dec 2013]. Available from: http: www.mlst.net/

11. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006;23(2):254-67. http://dx.doi.org/10.1093/molbev/msj030 PMID:16221896

12. FigTree. [Accessed 5 Dec 2013]. Available from: http: tree.bio.ed.ac.uk/software/figtree/

13. Monecke S, Coombs G, Shore AC, Coleman DC, Akpaka P, Borg M, et al. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant Staphylococcus aureus. PLoS ONE. 2011;6(4):e17936. http://dx.doi.org/10.1371/journal.pone.0017936 PMID:21494333

14. Bartels MD, Boye K, Rhod Larsen A, Skov R, Westh H. Rapid increase of genetically diverse methicillin-resistant Staphylococcus aureus, Copenhagen, Denmark. Emerg Infect Dis. 2007;13(10):1533-40. http://dx.doi.org/10.3201/eid1310.070503 PMID:18258003

15. Böcher S, Gervelmeyer A, Monnet DL, Mølbak K, Skov RL; Danish CA-MRSA Study Group. Methicillin-resistant Staphylococcus aureus: risk factors associated with community-onset infections in Denmark. Clin Microbiol Infect. 2008;14(10):942-8. http://dx.doi.org/10.1111/j.1469-0691.2008.02055.x PMID:18752595

16. Harris SR, Cartwright EJ, Török ME, Holden MT, Brown NM, Ogilvy-Stuart AL, et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis. 2013;13(2):130-6. http://dx.doi.org/10.1016/S1473-3099(12)70268-2 PMID:23158674

17. Köser CU, Holden MT, Ellington MJ, Cartwright EJ, Brown NM, Ogilvy-Stuart AL, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med. 2012;366(24):2267-75. http://dx.doi.org/10.1056/NEJMoa1109910 PMID:22693998

18. Böcher S, Skov RL, Knudsen MA, Guardabassi L, Mølbak K, Schouenborg P, et al. The search and destroy strategy prevents spread and long-term carriage of methicillin-resistant Staphylococcus aureus: results from the follow-up screening of a large ST22 (E-MRSA 15) outbreak in Denmark. Clin Microbiol Infect. 2010;16(9):1427-34. http://dx.doi.org/10.1111/j.1469-0691.2010.03137.x PMID:20041904

19. Lindsay JA. Evolution of Staphylococcus aureus and MRSA during outbreaks. Infect Genet Evol. 2014;21:548-53. http://dx.doi.org/10.1016/j.meegid.2013.04.017 PMID:23665384

20. Larsen AR, Goering R, Stegger M, Lindsay JA, Gould KA, Hinds J, et al. Two distinct clones of methicillin-resistant Staphylococcus aureus (MRSA) with the same USA300 pulsed-field gel electrophoresis profile: a potential pitfall for identification of USA300 community-associated MRSA. J Clin Microbiol. 2009;47(11):3765-8. http://dx.doi.org/10.1128/JCM.00934-09 PMID:19759225