Bacterial fingerprints across Europe

Glasner, Corinna

CHAPTER 5

A combined genomic and proteomic time-lapse movie of sequential *Staphylococcus aureus* isolates from a patient and his successive female partners

Corinna Glasner, Annette Dreisbach, Wouter F.W. Bierman, Anne de Jong, Peder Worning, Jesper Boye Nielsen, Dörte Becher, Andreas Otto, Janina Dordel, Sandra Reuter, Jan P. Arends and Jan Maarten van Dijl

Manuscript in preparation

ABSTRACT

Whole-genome sequencing (WGS) has revolutionized our understanding of bacterial evolution and pathogenesis, but the combination of WGS with proteomics in clinical settings has so far been lacking. The Gram-positive bacterium *Staphylococcus aureus* is usually a harmless commensal carried by about 30% of the population chronically. However, it can change into a dangerous pathogen causing severe invasive diseases. In the present study, we combined WGS and proteomics to investigate sequential *S. aureus* isolates, both MRSA and MSSA, collected over a period of five years from a patient suffering from recurrent ear infections and his two successive partners. WGS revealed merely 59 single nucleotide variants (SNPs) between the seven MRSA isolates, including 28 intragenic and 31 intergenic SNPs; between the three MSSA isolates, only 27 SNPs were identified, including 15 intragenic and 12 intergenic SNPs. The smaller portion of these SNPs was found in the core genome, with only 12 different SNPs among the seven MRSA isolates and three SNPs among the three MSSA isolates. Specifically, SNPs were detected in the *rsbU*, *agrC* and *clfB* genes that could affect virulence and/or adhesion, while other SNPs in intergenic regions could influence gene expression and might alter the overall proteome repertoire. Mass spectrometric analyses of the exoproteome revealed that the majority of identified proteins are important virulence factors, such as protein A, Sbi, IsdA, IsdB, CHIPS and PVL. Importantly, in this study we applied for the first time a combination of WGS and proteomics for the investigation of sequential *S. aureus* isolates from an individual patient over a longer time period, taking into account the chronology of antibiotic treatments and the patient's travel destinations. In conclusion, this comprehensive analysis of the *S. aureus* genome over time and the proteins secreted by consecutive isogenic isolates sheds new light on the impact of genome evolution on the global production of virulence factors.

INTRODUCTION

The renowned opportunistic pathogen *Staphylococcus aureus* is a regular component of the human microbiota with its preferred niches in the nasopharynx and on the skin. In general, *S. aureus* displays a commensal lifestyle, accompanying about 30% of the human population either chronically or intermittently [1,2]. Clearly, the prevalence of asymptomatic carriage dwarfs the incidence of staphylococcal disease, which implies that the pathogenic state represents an infrequent lifestyle of *S. aureus*. The commensal colonization remains generally unnoticed by the human host, suggesting a balance between *S. aureus* and the host that most likely resulted from co-evolution [3]. However, upon disturbance of this balance, *S. aureus* can transform into a dangerous invasive pathogen. The underlying mechanisms causing this lifestyle switch are still unknown [3]. Once *S. aureus* becomes invasive it can cause an array of diseases ranging from relatively mild skin infections to life-threatening conditions. Host risk factors that can have an effect on this delicate balance have been identified and include age, sex or an underlying health condition (e.g. immune deficiency and age-related diseases) [4]. A genetic predisposition of the host has also been suggested to affect this balance. Several studies reported high inter-individual variability of the anti-staphylococcal host immune response, and genome-wide association studies in humans found a link between particular genetic traits and the development of *S. aureus* bacteremia in a healthcare setting [5-7]. Bacterial factors that have been suggested to cause a destabilization, include virulence factors and antibiotic resistance genes, although no consistent association between the presence of particular genes and outcome of disease has been demonstrated so far [8]. To no surprise, *S. aureus* nasal carriage is a well-established risk factor for the development of staphylococcal infections, which are in 80% of the cases caused by the endogenous strain [1,2]. Despite this, persistent carriers have a lower risk of death by bacteraemia compared to non-carriers [9].

The large and variable virulence factor repertoire of *S. aureus* underpins the view that this bacterium generally harbors all the equipment needed to cause disease facultatively [10,11]. Virulence factors described to play a key role in within-host survival via mechanisms like adhesion, invasion and evasion of tissues, or evasion of the immune defenses include amongst others fibronectin-binding proteins, clumping factors, various toxins and the immune evasion cluster (IEC) [10,12]. The development of antibiotic resistance provides *S. aureus* with a further advantage for survival within the human host. In view of the diverse and variable virulence and antibiotic resistance potential of *S. aureus*, the genome of this bacterium has been under intense investigation in the past decade. The new era of whole-genome sequencing (WGS) revolutionized our understanding of genome plasticity and evolution of *S. aureus*. The numerous applications of WGS have provided valuable insights into the overall population structure and evolution of *S. aureus*, transmission routes in outbreak situations and the within-host evolution, exploiting *S. aureus* isolates from the same individual [13-18]. The latter builds an important avenue for research into the aetiology and mechanisms of staphylococcal disease. At present, only a limited number of studies have investigated series of isogenic *S. aureus* patient isolates, with the majority solely focussing on the correlation between disease, antibiotic treatment and the subsequent emerging antibiotic resistance in the respective isolates [4,13,14,19-24]. These studies have clearly contributed to our understanding of the *in vivo* evolution of *S. aureus* in patients and have collectively identified a limited number of single nucleotide variants (SNPs) that cause the emergence of antibiotic resistance in these isolates following antibiotic treatment. Nevertheless, no study has so far addressed both the whole genome and the complex proteome of sequential *S. aureus* isolates from their human host. Since anti-staphylococcal immune responses in both carriers and patients have revealed a high inter- and intra- individual variability, the expression and secretion of *S. aureus* virulence factors within the human host seem to play an important role in the delicate balance between *S. aureus* and the host. Previous investigations of virulence factor expression *in vitro* revealed a remarkable heterogeneity in the examined clinical *S. aureus* isolates [25,26]. It was suggested from these studies that the heterogeneous expression patterns under identical *in vitro* conditions, could reflect a similar high degree of variability *in vivo* [25].

Our group has recently shown that the *S. aureus* lineage with the *spa*-type t437 resembles a genetically tight cluster across Europe. Notably, *S. aureus* t437 in association with the multilocus sequence type (MLST) clonal complex (CC) 59 is the predominant community-associated methicillin-resistant *S. aureus* (MRSA) clone in Asia. The discovery of this lineage within a male patient suffering from recurrent external otitis and otitis media, who visited our hospital at regular intervals, was the incentive to investigate the possible evolution of this lineage within one patient. For this purpose, seven MRSA isolates with the *spa*-type t437 and three methicillin-sensitive *S. aureus* (MSSA) isolates from a different lineage (ST30, t021) isolated within a time period of five years from this patient and his two female partners were examined. The aim of this study was to unite the fields of genomics and proteomics, with the application of basic geno- and phenotypic assays, WGS and gel-free proteomics, to create a 'time-lapse movie' of the carried *S. aureus* lineages within one patient.

MATERIALS AND METHODS

Bacterial isolates and patient history

The sequential MRSA isolates E52, E42, E53, E44, E45 and E48, followed by two MSSA isolates E54 and E57, collected between 2007 and 2012 from the ear or nose from a male patient suffering from recurrent external otitis and otitis media, were retrospectively selected for this study. Two additional *S. aureus* isolates E51 (MRSA) and E58 (MSSA) were also included, originating from his two successive female partners in 2007 and 2012. For an overview of the *S. aureus* isolation dates and their key molecular characteristics see Figure 1. The first MRSA eradication attempt of the patient in October 2007, a combination therapy with doxycycline, rifampicin and mupirocin, failed. In the subsequent two years, the patient continued to carry MRSA in his ear with two more failed eradication attempts in November 2008 and March 2009. Between the first and second eradication attempt the patient underwent a tympanoplasty in October 2008. Finally, the patient was successfully cleared of MRSA, possibly due to a ciprofloxacin therapy that was administered for gastroenteritis in July 2009 combined with a susceptible situation, because his chronic ear infection had finally resolved. In fact, since the patient was not tested between July 2009 and July 2011, the exact cause resulting in the loss of MRSA cannot be determined with certainty. During this time, the patient also received amoxicillin and doxycycline for other medical reasons, however only in July 2011 another MRSA test was performed, resulting in a negative outcome. The MRSA negative results led to the incentive to test the patient for MSSA, which was subsequently isolated from his nose in November 2011. The first MRSA eradication procedure of the female partner, utilizing the same combination therapy as the patient, was immediately successful in August 2007. In early 2012, the patient and his new female partner at the time provided the final two MSSA isolates. The traceable history of the patient's antibiotic treatment is displayed in Figure 1 and the documented *S. aureus* isolation history of the patient between 2007 and 2013 is shown in the Supplementary Table 1 (available upon request). Unfortunately, it cannot be determined with certainty since when the patient was MRSA positive prior to the first identification in July 2007, since only his short hospital visit in France prior to his outpatient visit at the University Medical Center Groningen (UMCG) led to the active MRSA screening in accordance with the protocol for MRSA screening implemented by the UMCG.

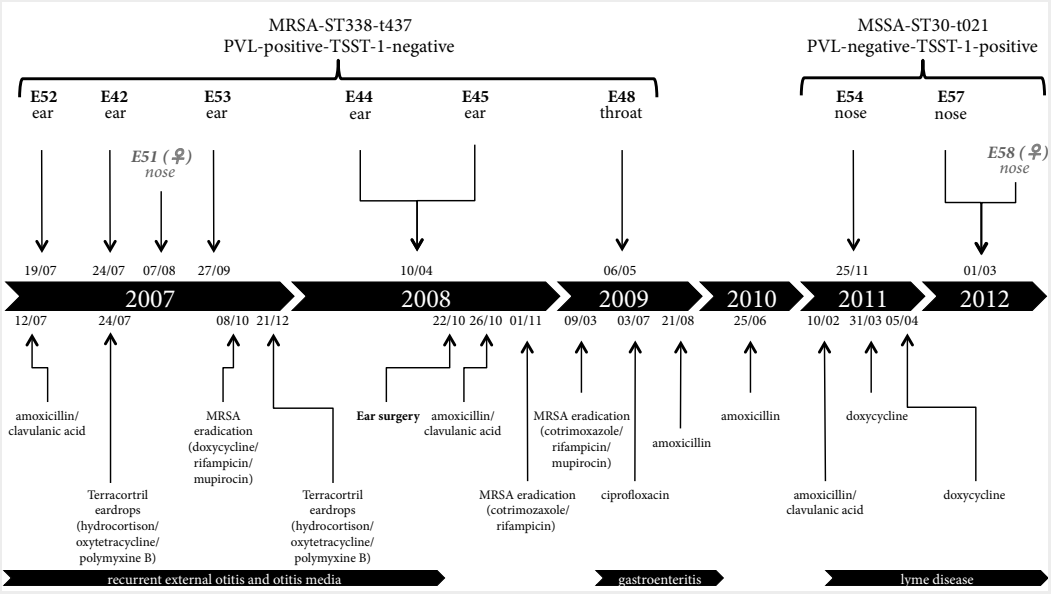

Figure 1. Timeline of the collection of seven MRSA and three MSSA isolates, their basic molecular characteristics and the antibiotic treatment of the patient between 2007 and 2012. Dates of *S. aureus* isolation and treatments are indicated. Isolates from the patient's female partners are indicated in italics and grey letters (marked ♀).

Antibiotic susceptibility testing

Antibiotic susceptibility was determined with the VITEK 2 system utilizing the AST P633 card for Gram-positive bacteria (bioMérieux, Marcy l'Etoil, France). The VITEK 2 minimum inhibitory concentration results were interpreted using the advanced expert system following EUCAST guidelines (http://www.eucast.org).

Molecular and phenotypic analyses

The 10 *S. aureus* study isolates were examined for the presence/absence of the following genes by PCR with the subsequent visualization of amplified fragments on a 1% agarose gel as described elsewhere: *pvl*, *tst-1, hla, hld, hlg, hlg-2* and *hlb* [27], *scn, chp* and the full-length *hlb* [28], *agr*-type [29], and SCC*mec* type [30,31]. Additionally, the *spa*-type of all 10 *S. aureus* study isolates was determined in accordance with published protocols [32,33]. The Ridom StaphType software v2.0 (Ridom GmbH, Münster, Germany) was used to assign the *spa*-type. The MRSA E42 isolate was as well typed by MLST according to the previously described protocol [34]stN/. *In vitro* production of the 3 haemolysins (i.e. α-haemolysin, β-haemolysin and δ-haemolysin) was determined with the haemolysis pattern assay as described by Adhikari *et al.* [35]. To assess the spreading motility of *S. aureus* and the associated functionality of the *agr* locus, a previously described soft agar plate assay was implemented [36]

Genome analyses

Bacterial growth for genomic DNA extraction

Total DNA of the *S. aureus* study isolates was isolated with the QiAmp DNA kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions with some modifications. In brief, bacterial cultures were grown overnight in a total volume of 5 mL under static conditions in tryptic soy broth (TSB, Oxoid, Hampshire, United Kingdom) at 37°C. Subsequently, 3 mL overnight culture was used to extract the genomic DNA. The addition of 0.1 mg lysostaphin per sample with an incubation step at 37°C for 90 min resulted in an efficient cell wall degradation of *S. aureus* that could be observed by

the clearing of the turbid cell suspension. The final DNA concentration and purity was measured and DNA samples with a concentration of 80-100 ng/μL were stored at -20°C until further use.

Genome sequencing and analysis

Whole-genome sequencing was performed using the Illumina HiSeq2000 platform (Illumina, San Diego, USA). All pipetting and incubation steps during the preparation workflow of the DNA for sequencing were performed with the Caliper Sciclone liquid handling station (PerkinElmer, Waltham, USA). Firstly, the DNA samples were randomly fragmented using the Covaris DNA shearing instrument (Covaris, Woburn, USA). Subsequently, barcoded adapters were ligated to both ends of the DNA fragments, according to the New England BioLabs NEBNext sample prep kit (BioLabs, Ipswitch, USA). Fragments with an insert size of 400 bp on average were collected using the LabChipXT system with the LabChip XT DNA 750 Assay Kit (PerkinElmer, Waltham, USA) and the extracted DNA was amplified by PCR. The quality of the product was verified on the BioRad Experion Automated Electrophoresis System (BioRad, Hercules, USA), prior to multiplexing the samples in an equimolar pool of 20 products. Finally, the pooled samples were sequenced on the HiSeq2000 with 100 bp paired-end reads. Image files were processed using standard Illumina base calling software with *in silico* de-multiplexing of the generated reads.

High quality reads of the ten *S. aureus* study isolates were assembled *de novo* using Velvet v1.2.03 [37] (https://www.ebi.ac.uk/~zerbino/velvet/). Velvetoptimiser was used to optimize the three primary parameter options (K, -exp_cov, -cov_cutoff) for the Velvet *de novo* sequence assembler to give the highest number of bases in contigs with lengths greater than 1 kb (http://www.bioinformatics.net.au/software.velvetoptimiser.shtml). Genome annotation was performed automatically on the rapid annotation using subsystem technology (RAST) server 4.0 with the RAST gene caller and FigFam release 69 (http://www.rast.nmpdr.org). An overview of the genome features is displayed in the Supplementary Table 2 (available upon request).

SNPs in the genomes of seven MRSA isolates were called by implementation of two different approaches. In the first approach, the *de novo* assembled genome sequences were analysed with the SNPsFinder algorithm (http://www.snpsfinder.lanl.gov/) [38]. As the outcome of this tool is fairly limited due to the fact that already assembled genomes were utilized, a second approach was followed in which the raw fastq read files were used. In this approach, the short reads of the 7 MRSA genomes were mapped to a reference genome (*S. aureus* strain SA957, GenBank accession No. CP003603.1) using the burrows-wheeler alignment (BWA) with subsequent visualization in Artemis [39]

Comparative genomics analyses between the 7 MRSA isolates and a reference genome (*S. aureus* SA957, GenBank accession CP003603.1), and between the 3 MSSA isolates, were performed on a pairwise basis and in the order of isolation to determine the overall genomic relatedness between the study isolates and the reference genome (data not shown). For this purpose, the ten contig files were first re-ordered to the respective reference genome using Abacas1.3.1.1 and, subsequently, a BLASTN comparison file between the newly re-ordered contigs and the query sequence was generated with WebACT (http://www.webact.org/WebACT/generate) and displayed in ACT [40].

Phylogenetic and statistical analyses

Phylogenetic analyses were performed using SNP calling against a reference genome USA300 TCH1516 (genome accession No. NC_010079.1). The pairwise SNP distance was calculated and a distance matrix was constructed. The tree was constructed as a neighbour-joining tree from the distance matrix. Bootstrapping was implemented in SplitsTree4 v4.11.3 [41]. Tree-drawing was performed with FigTree (v1.4.2; http://tree.bio.ed.ac.uk/software/figtree/).

SeqSphere⁺ data analysis

The *de novo* assembled contigs of the ten *S. aureus* genomes were imported into the SeqSphere⁺ software v2.0 (Ridom GmbH, Münster, Germany). Three different publicly available target schemes were predefined in the SeqSphere+ software: MLST, MLST⁺ and the accessory scheme [42]. The *in silico* MLST type was determined and additionally a gene-by-gene comparison with the defined MLST+ and Accessory scheme was performed that resulted in a SeqSphere+ specific allelic profile that was used for comparisons between the different *S. aureus* isolates.

Exoproteome analyses

Bacterial growth and protein extraction

For the preparation of extracellular protein extracts, bacteria were grown in triplicates overnight in 10 mL TSB under vigorous shaking (115 rpm) at 37°C in a water bath. The cultures were then diluted into 10 mL pre-warmed RPMI 1640 medium (GE Healthcare/PAA, Little Chalfont, United Kingdom) to an OD_{600} of 0.1 and cultivation was continued under the same conditions. Exponentially growing cells with an OD_{600} of ± 0.5 were again diluted into 30 mL fresh and pre-warmed RPMI medium to a final OD_{600} of 0.1 and their cultivation was continued until the cultures had reached 90 min within the stationary growth phase. At this time point, which usually corresponded to an OD_{600} of approx. 1.5, the extracellular proteins from 1.5 mL culture aliquots were collected, precipitated, washed, and dried until further use [25,43]. In brief, 1.5 mL culture aliquots were centrifuged for 10 min at 4°C and 8000×*g* with the subsequent application of a 0.22 µM filter step (GE Healthcare Systems, Little Chalfont, United Kingdom) to remove the remaining bacterial cells. The extracellular proteins in the supernatant were precipitated with 10% w/v TCA on ice at 4°C overnight. The precipitates were collected by centrifugation for 20 min at 4°C and 8000×*g*, washed with ice-cold acetone and dried at room temperature. The dried protein pellets were stored at -20°C until further use.

Sample preparation

Dried protein samples were processed as described before [44]. In brief, protein pellets were dissolved in 50 mM ammonium bicarbonate buffer (Fluka, Buchs, Switzerland), reduced with 10 mM dithiothreitol (Duchefa Biochemie, Haarlem, the Netherlands) for 30 min and alkylated with 10 mM iodoacetamide (Sigma-Aldrich, St. Louis, USA) for 30 min in the dark. For the digestion of the complex protein samples, 80 ng trypsin (Promega, Madison, USA) was added and the samples were incubated overnight at 37°C under static conditions. To stop the digestive reaction, the samples were acidified with a final concentration of 0.1% trifluoroacetic acid (TFA, Sigma-Aldrich, St. Louis, USA) and subsequently purified using ZipTips (Millipore, Billerica, USA). For this purpose, the tips were stepwise equilibrated with 30 µL acetonitrile (ACN, Fluka, Buchs, Switzerland),, 30 µL 80% ACN/0.1% TFA, 50% ACN/0.1% TFA, 30 µL 30% ACN/0.1% TFA and finally 30 µL 0.1% TFA. Peptides were bound to ZipTips by pipetting 10 times 10 µL of the sample. Impurities were removed by washing with 50 µL 0.1% TFA and finally peptides were eluted with 20 µL 50% ACN/0.1% TFA and 20 µL 80% ACN/0.1% TFA. The final eluates were concentrated using a vacuum centrifuge (Eppendorf, Hamburg, Germany) and stored at 4°C until further use.

Mass spectrometry analyses

The peptides were separated by reversed phase liquid chromatography (LC) and coupled online to electrospray ionization mass spectrometry (ESI-MS) on an LTQ Orbitrap as described in Bonn *et al.* [45]. Database searching was done with Sorcerer-SEQUEST 4 (Sage-N Research, Milpitas, USA). After extraction from the raw files, *.dta files were searched with Sequest against a target-decoy database with a set of common laboratory contaminants. The databases for the respective peptide/protein search were created from the own genome sequences of the ten *S. aureus* study isolates. The RAST annotation file of the ten study isolates was used to create a non-redundant database comprising protein sequences of all isolates. Protein sequences that differed in only 1 amino acid were included

in this database. Finally, the gene names and uniprot identifiers were added. Validation of MS/MS based peptide and protein identifications was performed with Scaffold v4.3.4 (Proteome Software, Portland, USA). Peptide identifications were accepted if they exceeded specific database search engine thresholds. SEQUEST identifications required at least deltaCn scores of greater than 0.1 and XCorr scores of greater than 2.6, 3.5 and 3.5 for doubly, triply and quadruply charged peptides, respectively. With these filter parameters no false-positive hits were obtained. Protein identifications were accepted if they were identified in 2 out of 3 biological replicates. Protein data was exported from Scaffold and curated in Microsoft Excel before further analysis. For the in-depth analysis of the identified proteins, their predicted subcellular locations and thepresence of signal peptides the PSort (http://www.psort.org/psortb/index.html), Phobius (http://www.phobius.sbc.su.se/) and LipoP (http://www.cbs.dtu.dk/services/LipoP/) algorithms were used [46-48].

RESULTS AND DISCUSSION

Characterization of the *S. aureus* study isolates reveals a high relatedness among the seven MRSA and three MSSA isolates, respectively

A total of ten clinical *S. aureus* isolates, seven MRSA and three MSSA, were examined in the present study. Eight of these isolates were isolated from a patient who visited our hospital in July 2007 suffering from recurrent external otitis and otitis media, whereas the other two isolates were collected from his two successive female partners in 2007 and 2012. The primary geno– and phenotypic investigations revealed that the seven MRSA isolates were highly similar to each other, and that the same was true for the three MSSA isolates. However, these investigations also uncovered clear differences between the respective MRSA and MSSA isolate groups (Table 1). Specifically, all seven MRSA isolates featured the following basic molecular profile: ST338, *spa*-type t437, PVL-positive, TSST-1-negative, *agr*-type I and SCC*mec* type V (5C2&5). The three MSSA isolates represented a different lineage possessing the following basic molecular profile: ST30, *spa*-type t021, PVL-negative, TSST-1 positive and *agr*-type III. Based on the results of these initial investigations, it was clear that the patient had acquired a totally unrelated *S. aureus* MSSA lineage at some point between July 2009 and July 2011 and that the patient and his two successive female partners carried the same *S. aureus* lineage at the respective time points. The latter is indicative of transmission events between the partners, which has been described numerous times in different types of studies [18,49,50].

Intriguingly, the ST338 of the seven MRSA isolates is a single-locus variant of the ST59, both belonging to the clonal complex 59 (CC59), which is the predominant community-associated MRSA clone in Asia [51,52]. Moreover, the Asian CC59 is primarily linked with the *spa*-type t437, which was also determined for the seven MRSA isolates. Therefore we mapped the travel destinations of the patient from 2005 onwards. Although this showed that the patient had not travelled to Asia before the first MRSA isolation (July 2007), it clearly illustrated his 'globetrotting' behaviour (Figure 2). The patient reported to have only travelled to Belgium, France and Norway shortly before visiting the ear, nose and throat outpatient clinic of our hospital in July 2007. During his travel to France the patient visited a hospital for one day, but unfortunately, no information about a possible treatment in this hospital was available. Notably, the patient's newly acquired MSSA isolate from 2011 with the ST30 is, in its MRSA form, the most dominant community-associated MRSA type found in Oceania. It is therefore also referred to as the South West Pacific clone [53]. Since the combination of epidemiological and molecular studies have shown that certain lineages of *S. aureus* attain a geo-spatial predominance, it is very well conceivable that the patient acquired the two different *S. aureus* lineages during his travels. This supports the view that the global exchange, spread and transmission of certain *S. aureus* lineages through travel should not be under-estimated. Especially when clones represent a very tight genomic cluster irrespective of their epidemiological features, they could pose a clear threat to public health [16,17,54,55].

With regard to previously published *S. aureus in vivo* evolution studies, the antibiotic resistance

geno- and phenotypes of the ten *S. aureus* study isolates were examined (Table 2). All MRSA isolates displayed a multidrug-resistant phenotype, including resistance against penicillin, oxacillin, kanamycin, erythromycin, clindamycin and chloramphenicol. Resistance to tetracycline was detected in all but the very first MRSA isolate (E52). Only five days separate the first and second *S. aureus* isolation dates, of which the second date on which MRSA E42 was collected coincided with a treatment with terracortril eardrops containing oxytetracycline, hydrocortisone and polymyxin B (Figure 1). Although no evidence for a previous treatment with tetracycline could be found in the patient's medical records, he may have received antibiotics during his brief visit in the French hospital. Consistent with the outcomes of antibiotic susceptibility testing, the resistance genotypes of the seven MRSA isolates were highly similar, and only the first MRSA isolate E52 lacked the *tetK* and *cadD* genes (Table 2). Intriguingly, although the patient received mupirocin, rifampicin and co-trimoxazole on several occasion between July 2007 and July 2009, none of the collected MRSA isolates was resistant to these antibiotics. In this context it is noteworthy that the first five MRSA isolates originated from the continuously infected ear of the patient. Hence, it seems that the *S. aureus* residing in this niche was insufficiently exposed to mupirocin, rifampicin and co-trimoxazole. Alternatively, some *S. aureus* cells may have invaded host cells and/or entered a persistent/dormant state. Especially, internalized *S. aureus* may be protected against the administered antibiotics resulting in failure of the therapy. Accordingly, no clear connection between the antibiotic treatment and the observed antibiotic resistance profiles of the different isolates can be made. In contrast to the MRSA isolates, the resistance geno- and phenotypes of the three MSSA isolates were very meagre as the latter isolates only carried the *blaZ* gene causing penicillin-resistance (Table 2).

Table 1. Genotypic and phenotypic characterisation of the ten *S. aureus* study isolates.

Isolate	Sample date	Sample site	*spa*-type	MLST[a]	*mecA*	*pvl*	*tst-1*	*hla*	*hlb*	*hld*	*hlg*	*hlg-2*	*scn*	*chp*	*agr* type	SCCmec type	α-haemolysin[b]	β-haemolysin[b]	δ-haemolysin[b]	Spreading
E52	19/07/07	ear	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	+	+	+	+
E42	24/07/07	ear	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	+	+	+	+
E51	07/08/07	nose	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	-	-	-	-
E53	27/09/07	ear	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	-	+	-	+
E44	10/04/08	ear	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	-	+	-	+
E45	10/04/08	ear	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	-	+	-	-
E48	06/05/09	throat	t437	ST338	+	+	-	+	+	+	-	+	+	+	I	V (5C2&5)	-	-	-	-
E54	25/11/11	nose	t021	ST30	-	-	+	-	-	+	+	-	+	+	III	-	-	-	-	+
E57	01/03/12	nose	t021	ST30	-	-	+	-	-	+	+	-	+	+	III	-	-	-	-	+
E58	01/03/12	nose	t021	ST30	-	-	+	-	-	+	+	-	+	+	III	-	-	-	-	+

[a]MLST was determined with SeqSphere+ software (the ST of E42 was also assessed through the original MLST protocol); [b] phenotypic determination; '+', tested positive; '-', tested negative.

Table 2. Overview of the antibiotic resistance geno- and phenotypes of the ten *S. aureus* study isolates.

Isolate	Resistance phenotype[a]								Resistance genotype[b]									
	CXT	PEN	OXA	KAN	ERY	CLI	TET	CHLO	*mecA*	*blaZ*	*aadE-aphA3*	*ermB*	*tetK*	*cat*	*cadD*	*cadA*	*copA*	*mco*
E52	R	R	R	R	R	R	S	R	+	+	+	+	-	+	-	-	-	-
E42	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E51	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E53	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E44	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E45	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E48	R	R	R	R	R	R	R	R	+	+	+	+	+	+	+	-	-	-
E54	S	R	S	S	S	S	S	S	-	+	-	-	-	-	-	+	+	+
E57	S	R	S	S	S	S	S	S	-	+	-	-	-	-	-	+	+	+
E58	S	R	S	S	S	S	S	S	-	+	-	-	-	-	-	+	+	+

[a] Determined with the VITEK 2 system; [b] Determined by a BLAST search of the respective genes against the whole genome sequences. Abbreviations: CXT, cefoxitin; PEN, penicillin; OXA, oxacillin; KAN, kanamycin; ERY, erythromycin; CLI, clindamycin; TET, tetracycline; CHLO, chloramphenicol; 'R', resistant; 'S', sensitive; '+', tested positive; '-', tested negative.

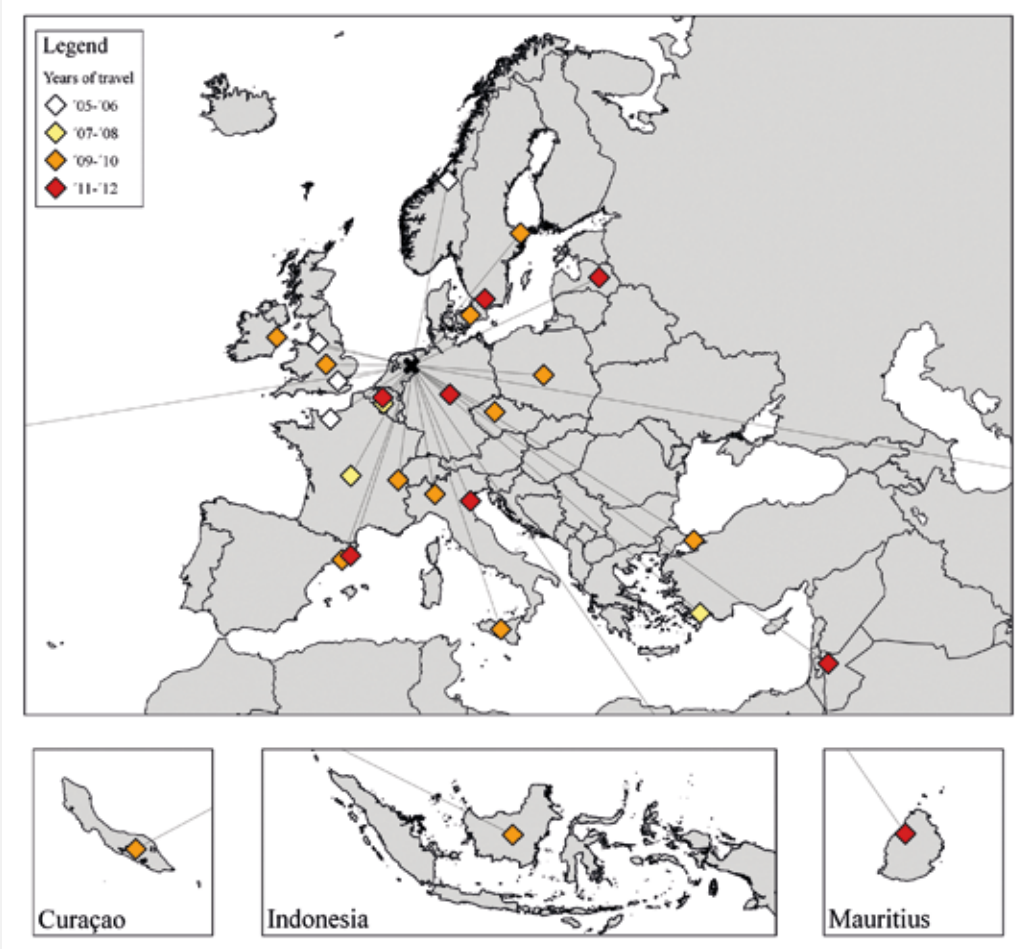

Figure 2. Map displaying all travel destination of the patient between 2005 and 2012. The black cross indicates the place of residence of the patient. His travel destinations where grouped per two years; white diamonds for 2005/2006, yellow diamonds for 2007/2008, orange diamonds for 2009/2010 and red diamonds for 2011/2012.

However, a revision of the available patient's medical history clearly revealed that the patient had a long-lasting problem with his upper respiratory tract suffering from recurrent ear infections since early childhood. At several occasion he received treatment with a diverse range of eardrops and steroids that were administered to his ear and nose, including Otosporin (containing polymyxin B, neomycin and hydrocortisone) and Sofradex (containing framycetin, dexamethasone and gramicidin) eardrops in the year prior to the first MRSA isolation. In addition to the ear infections, the patient suffered from gastroenteritis in July 2009 and Lyme disease in 2011 (Figure 1).

Intriguingly, a crude screen of the secreted protein fractions of the seven MRSA and three MSSA isolates by SDS-PAGE revealed substantial differences in the respective protein profiles, despite the fact that isolates belonging to either the MRSA or MSSA groups appeared geno- and phenotypically very similar. Differences between individual isolates were also clearly evident from their respective haemolysis patterns and spreading ability on soft agar plates (Table 1). Together, these observations formed the incentive to perform an in-depth genome and proteome analyses as described in the following paragraphs. However, with regard to the studied patient isolates, it has to be noted that the present analyses were performed in a retrospective manner using a rather narrow selection of isolates. This relates to the fact that all isolates were collected during routine diagnostic practice, where only one colony per sample is retained and stored at -80°C.

Close genomic relationship among the seven MRSA and three MSSA isolates

To identify potential genomic changes in the *S. aureus* carried by the patient over a prolonged time period, the whole genomes of his eight sequentially collected *S. aureus* isolates were sequenced. In addition, the genomes of two MSSA isolates from the patient's successive partners, who presumably acquired the respective isolates from the patient, were also sequenced. This WGS analysis revealed the close genomic relationships between the seven MRSA isolates and within the group of three MSSA isolates, which are in full agreement with the above-presented basic molecular and phenotypic characteristics of the respective isolates. Moreover, the application of WGS allowed a more in-depth analysis of the overall genome contents.

Phylogenetic analysis of the ten study isolates based on core genome SNPs clearly indicated that the two isolate groups, the seven MRSA isolates and the three MSSA isolates, each form a very narrow, distinct clade within the tree, representing two totally different lineages (Figure 3). Following this approach, only ten SNPs among the seven MRSA isolates and three SNPs among the three MSSA isolates were identified in their core genomes. The ten SNPs between the seven MRSA isolates were identified in intergenic regions, a hypothetical protein-encoding gene, the *sigB* gene for the σ factor B, and genes for proteins involved in basic cellular functions. The three SNPs among the three MSSA isolates were identified in an intergenic region and genes for proteins involved in basic cellular functions (data not shown).

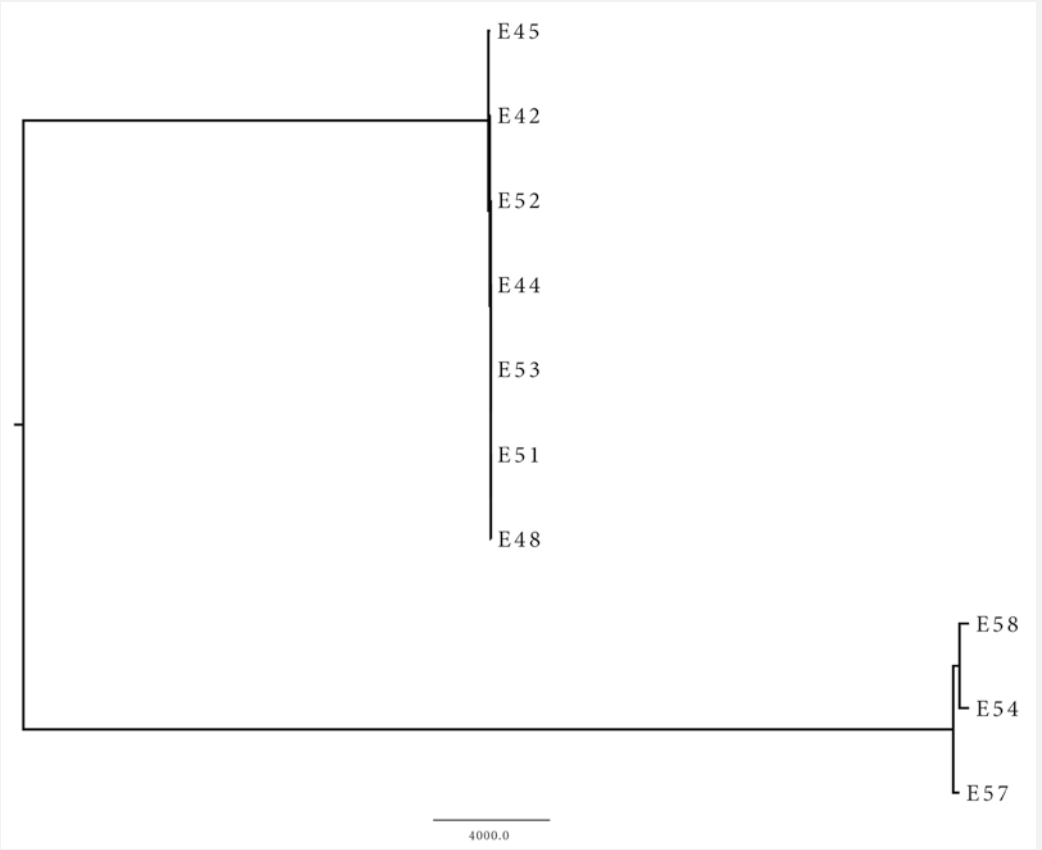

Figure 3. Phylogenetic relationship of the ten sequential *S. aureus* study isolates. Neighbor-joining phylogeny based on core genome point mutations (based on USA300_TCH1516) for the seven MRSA and three MSSA isolates. Scale represents substitutions per site.

For a more in-depth SNP analyses, including both the core and accessory genomes, two independent approaches were employed using different input files to call the SNPs among the seven MRSA and three MSSA isolates. A complete list of all identified SNPs among the MRSA and MSSA isolates with the SNPsFinder tool is displayed in the Supplementary Table 3 (available upon request). The application of the online tool SNPsFinder was very user-friendly and easy-to-use and resulted in the identification of 59 SNPs among the seven MRSA study isolates, including 28 intragenic and 31 intergenic SNPs. Crucially, compared to some other studies the total number of identified SNPs was relatively high [17,18,56]. This may have been caused by the specific *in vivo* pressure on *S. aureus* in the infected ear, relating either to the host immune defences, the different antibiotic therapies, or both. Other studies have reported similarly high numbers of SNPs in *S. aureus* isolates from an outbreak cluster and *S. aureus* isolates collected from a single patient during the progression from carriage to disease [21,57]. The latter study even identified numerous SNPs between different colonies from the same swab or culture, but also SNPs that occurred over time within the individual [21]. More specifically, a single, cohesive *S. aureus* population was discovered in one patient exhibiting a repertoire of 30 SNPs, which would be in line with the results of our present study. Notably, no large insertions or deletions were identified when comparing the seven MRSA isolates with each other and the reference genome. Surprisingly, over time no accumulation of particular SNPs could be observed among the seven consecutive MRSA study isolates. For example, the MRSA isolates E51and E48 contained a SNP within the *agr*C gene that was not identified in the other MRSA isolates, whereas the MRSA isolate

E45 contained a different SNP in the same gene. Most interestingly, the second MRSA isolate E42 did not harbour any of the SNPs that were found in the other MRSA isolates. Regarding the *rsbU* locus, only two out of the seven MRSA isolates (E53 and E44) possessed the same SNP (Supplementary Table 3, available upon request). The results of the SNPsFinder analysis were confirmed with the BWA alignment and variant calling, which allowed a visualization of the identified SNPs in relation to the query reference genome. Most interestingly, the seven MRSA isolates only contained a total number of 120 SNPs compared to the *S. aureus* SA957 reference isolate that originated from a blood culture from a young patient in Taiwan suffering from bacteraemia, arthritis and a cutaneous abscess [51,58]. This observation clearly indicated a high genomic relatedness between the seven MRSA isolates from the present study and this Taiwanese isolate, which was assigned as ST59. A total number of 27 SNPs, including 15 intragenic and 12 intergenic SNPs, were identified among the three MSSA isolates (Supplementary Table 3, available upon request). Important to note is that only four months lay between the isolation of the first MSSA isolate (E54) and the last two MSSA isolates (E57 and E58). Altogether, the identified SNPs in the MRSA isolates only overlap in a few instances with previously discovered SNPs in within-host evolution studies, including the *agrC* and *rsbU* genes. Surprising was the observation that none of the identified SNPs accumulated in the seven sequential MRSA isolates, which is in contrast to other within-host evolution studies of *S. aureus* [13,19,20]. This observed difference may have been caused by the distinct nature of the investigated sequential *S. aureus* isolates. Previous studies almost solely investigated bacteraemia samples from mostly extremely sick patients, whereas the present study utilized isolates from a relatively healthy young man who suffered from a common ear infection. Whether this different study set up might have caused the distinct SNP patterns still needs to be further investigated. An alternative explanation could be that the lack of overlap in identified SNPs relates to the fact that per time point of sampling only one *S. aureus* isolate was retained, while the patient may have carried at each time point various different *S. aureus* cells with a different range of only partially overlapping SNPs.

With the numerous WGS applications for microbiological research in recent years an arsenal of different bioinformatic tools for data analysis has been developed, with the ultimate goal to simplify and accelerate the downstream analyses of WGS and to make it accessible for bacteriological routine work in hospitals and laboratories. One of these novel tools is the SeqSphere$^+$ software for easy microbial typing [42]. The application of this software underscored the above-mentioned high genomic homology within the respective groups of MRSA and MSSA isolates. With the utilization of the publically available MLST+ scheme, which covers approximately 2740 coding regions defined by the *S. aureus* reference strain COL, a gene-by-gene comparison revealed that only 13 loci were different between the seven MRSA isolates and only five loci were different between the three MSSA isolates. These differences were either caused by a different known allele, a different new allele, a missing target or a failed target (data not shown). In addition, the implemented MLST+ scheme confirmed that the seven MRSA isolates were ST338 and a preliminary script for *spa*-typing confirmed the experimentally identified *spa*-type t437, whereas the three MSSA were assigned to the ST30 with *spa*-type t021. Applying the SeqSphere+ accessory genome scheme, only 11 targets were different between the seven MRSA isolates and only six targets were different between the three MSSA isolates. In summary, the application of the SeqSphere+ software was simple, user-friendly and resulted in rapid results. The possible analysis steps that are provided within the SeqSphere+ software are certainly sufficient enough to inform clinicians and microbiologists on a daily basis when implemented as a regular screening method for effective infection control (e.g. in outbreak situations). For research purposes, this software is very limited and restricts the user to predefined or self-made schemes. As an example, SNP calling can be performed, but only considers SNPs within the included targets.

Exoproteome analyses of the ten *S. aureus* study isolates revealed high heterogeneity

Proteomics, especially gel-free proteomics, is a very powerful tool to investigate the virulence potential of *S. aureus*. In view of the infectious isolation site of most of the study isolates, and since the exoproteome of *S. aureus* has been repeatedly shown to represent a key reservoir of virulence factors,

the present study focused solely on this particular subproteome. Moreover, the stationary phase of *S. aureus* during *in vitro* growth was chosen for the time point of investigation since previous studies have revealed that the majority of virulence factors are secreted during this particular growth phase [25,26]. As suggested by regular SDS-PAGE analyses (data not shown), a high degree of heterogeneity existed in the extracellular proteins of the seven MRSA and three MSSA isolates, the most prominent differences being observed between the MRSA and MSSA isolates, which belong to two totally unrelated *S. aureus* lineages as shown above.

Altogether 275 distinct proteins were identified by LC-MS/MS in the growth media of the seven MRSA isolates (Supplementary Table 4, available upon request). The total numbers of identified proteins between the MRSA isolates differed largely: E52 (131 proteins), E42 (172 proteins), E51 (110 proteins), E53 (132 proteins), E44 (246 proteins), E45 (118 proteins) and E48 (150 proteins). For 87 (31%) of all proteins a signal peptide could be predicted with SignalP 3.0 and other algorithms [47]. Using PSORTb version 3.0 (http://www.psor-t.org/psortb) on the complete protein dataset, 37 (13%) of the proteins were predicted to be extracellular, 15 (5.5%) were predicted to be cell wall-bound proteins and the localization of another 53 (20%) proteins is currently unknown [46] (Figure 4A). Moreover, 140 (51%) cytoplasmic proteins and 30 (11%) membrane proteins were also identified in the exoproteomes of the seven MRSA isolates, which is not unusual for bacterial exoproteomes and has been previously reported (Figure 4A) [25,26]. Among the 275 identified proteins, 27 were annotated as hypothetical proteins. These proteins are putative lipoproteins as judged by BLAST searches (http://www.uniprot.org/). Only 68 proteins were identified in all seven MRSA isolates, including protein A, Sbi, IsdA, IsdH, CHIPS, SEB, SEK, PVL, ClfB, SasH, Eap, LytM and a large number of cytoplasmic proteins (Supplementary Table 4, available on request). Proteins identified in the exoproteomes of all seven MRSA isolates clearly belong to the group of well-known virulence factors with a diverse range of functions. To what extent the presence of these proteins might have had an impact on the continuous ear infection of the patient would have to be investigated in an appropriate *in vivo* animal model. Strikingly, 110 proteins were found to be unique for one or two MRSA isolates under the tested conditions. Figure 4B shows the different predicted proteins per MRSA isolate. The two isolates E42 and E44 produced the majority of proteins that were not identified in any of the other isolates with the larger part of these proteins predicted to be located in the cytoplasm. Since a detailed protein comparison of all seven MRSA isolates with each other would be too complex, a selection of two different combinations is presented in the following. The first selection of isolates deals with the first three MRSA isolates of the patient (E52, E42 and E53) and the MRSA isolate of his first female partner (E51) (Figure 4C).

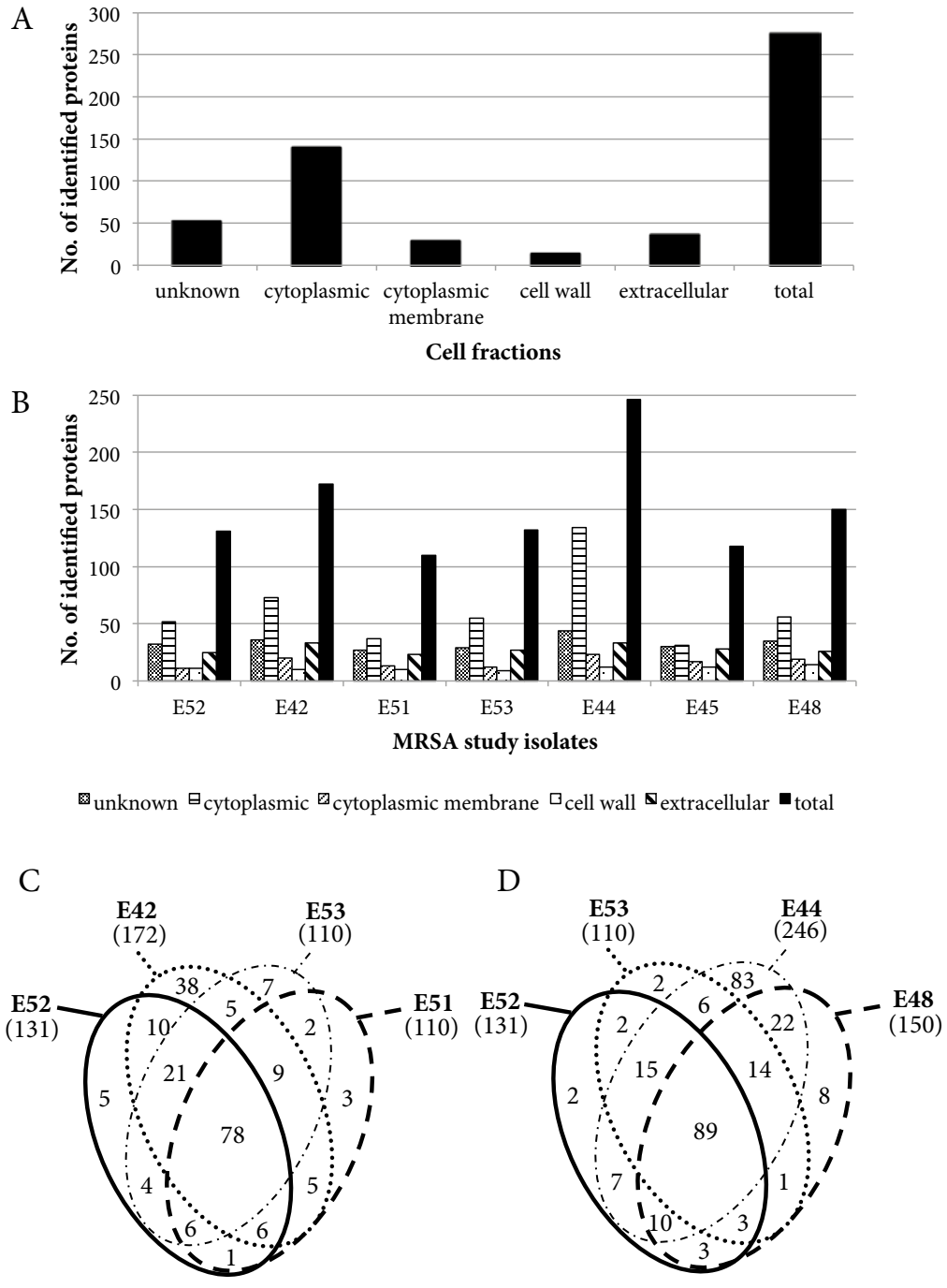

Figure 4. Overview of the extracellular protein identifications for the seven MRSA study isolates. (A) Total overview of all identified proteins in the growth media of the seven MRSA isolates and their predicted subcellular localization. **(B)** Identified proteins and their predicted subcellular localization indicated separately for each of the seven MRSA isolates. **(C and D)** Venn diagrams giving an overview of the numbers of consistently or uniquely identified extracellular proteins of the four MRSA isolates E52, E42, E53 and E51 (C) and the four MRSA isolates E52, E53, E44 and E48 (D).

A total number of 200 proteins were identified in all four isolates that were collected within three months. The four isolates had 78 of these proteins in common. Interestingly, proteins that were solely found in E52, E53 or E51 were all predicted to have a cytoplasmic or unknown localization. Crucially, 38 proteins that were identified in E42 were not found in the other three MRSA isolates, of which the majority is as well predicted to be cytoplasmic. Within this comparison it is also clear that the two isolates E52 and E42, which were collected only five days apart, share the most identified proteins with each other. The second isolate selection includes four other MRSA isolates of the patient (E52, E53, E44 and E48) covering the complete time period of two years in which the patient was MRSA-positive and suffered from the ear infection (Figure 4D). These four isolates share a total number of 89 proteins (33%) with each other. And as already revealed for the first comparison, proteins that were solely identified in a single isolate were either predicted to have a cytoplasmic or an unknown localization. Most interesting are the exoproteome results obtained for the MRSA isolate E44, with the highest number of identified proteins, and therefore also with the highest number of proteins solely found for this isolate (83 proteins). Most of the latter proteins are predicted to have a cytoplasmic, membrane or unknown localization, suggesting that isolate E44 may be more prone to lysis than the other investigated MRSA isolates. Altogether, it can be concluded that the *in vitro* exoproteome profiles of the seven MRSA isolates revealed extensive heterogeneity under the tested conditions, although WGS demonstrated their high genomic similarity. Most interestingly are the few virulence factors that were identified among all seven MRSA isolates, but also the large number of cytoplasmic and membrane proteins.

Altogether only 48 distinct proteins were identified by LC-MS/MS in the growth media of the three MSSA isolates (Supplementary Table 5, available on request). The total numbers of identified proteins in the different MSSA isolates differed: E54 (46 proteins), E57 (29 proteins) and E58 (26 proteins). For 30 (62%) of the total identified proteins a signal peptide could be predicted with SignalP 3.0 and other algorithms [47]. Using PSORTb version 3.0 (http://www.psor-t.org/psortb) on the complete exoproteome dataset, 10 (21%) of these proteins were predicted to be extracellular, 8 (17%) were predicted to be cell wall-bound proteins and the predicted localization of another 15 (31%) proteins is currently unknown [46] (Figure 5A). Moreover, 10 (21%) cytoplasmic proteins and 5 (10%) membrane proteins were as well identified in the exoproteomes of these three MSSA isolates (Figure 5A). Among the 48 identified proteins, nine were annotated as hypothetical proteins but BLAST searches suggest that they are actually potential lipoproteins (http://www.uniprot.org/). Only 20 proteins were identified in all three MSSA isolates, including protein A, IsdA, Sbi IsdH, IsdD and a large number of cytoplasmic proteins. Accordingly, 28 proteins were uniquely identified for one or two strains (Supplementary Table 5, available on request). For an overview of the predicted localization of the different proteins identified for each of the three MSSA isolates refer to Figure 5B. Crucially, there is some overlap in the most abundantly identified proteins in the MRSA and MSSA groups, as protein A, Sbi, IsdA and IsdH were identified for both groups, illustrating their dominant expression under the tested conditions. The Venn diagram in Figure 5C displays a comparison of all identified proteins in the growth media of the three MSSA isolates. Altogether, the very low number of identified extracellular proteins and the large difference in numbers between the first and the last two MSSA isolates were unexpected. However, compared to the MRSA isolates, the number of cytoplasmic proteins observed in the media of the MSSA isolates was relatively low, suggesting that the investigated MSSA isolates are less prone to cell lysis than the MRSA isolates. Whether this is a specific trait of the investigated MSSA isolates or a more general trait of MSSA is unknown due to the current lack of proteomics studies on clinical MSSA isolates. It would therefore be of relevance to analyse the exoproteomes of a larger collection of MSSA isolates in comparison to the presently studied MSSA isolates.

Only five reports are so far available investigating the complex exoproteome of clinical *S. aureus* isolates under *in vitro* conditions [26,59-62]. However, comparing the results of the present study with these previous reports is not feasible, because of the following constraints: (i) in contrast to our experimental set-up growing all *S. aureus* isolates in RPMI until 90 min within stationary phase, all five previously

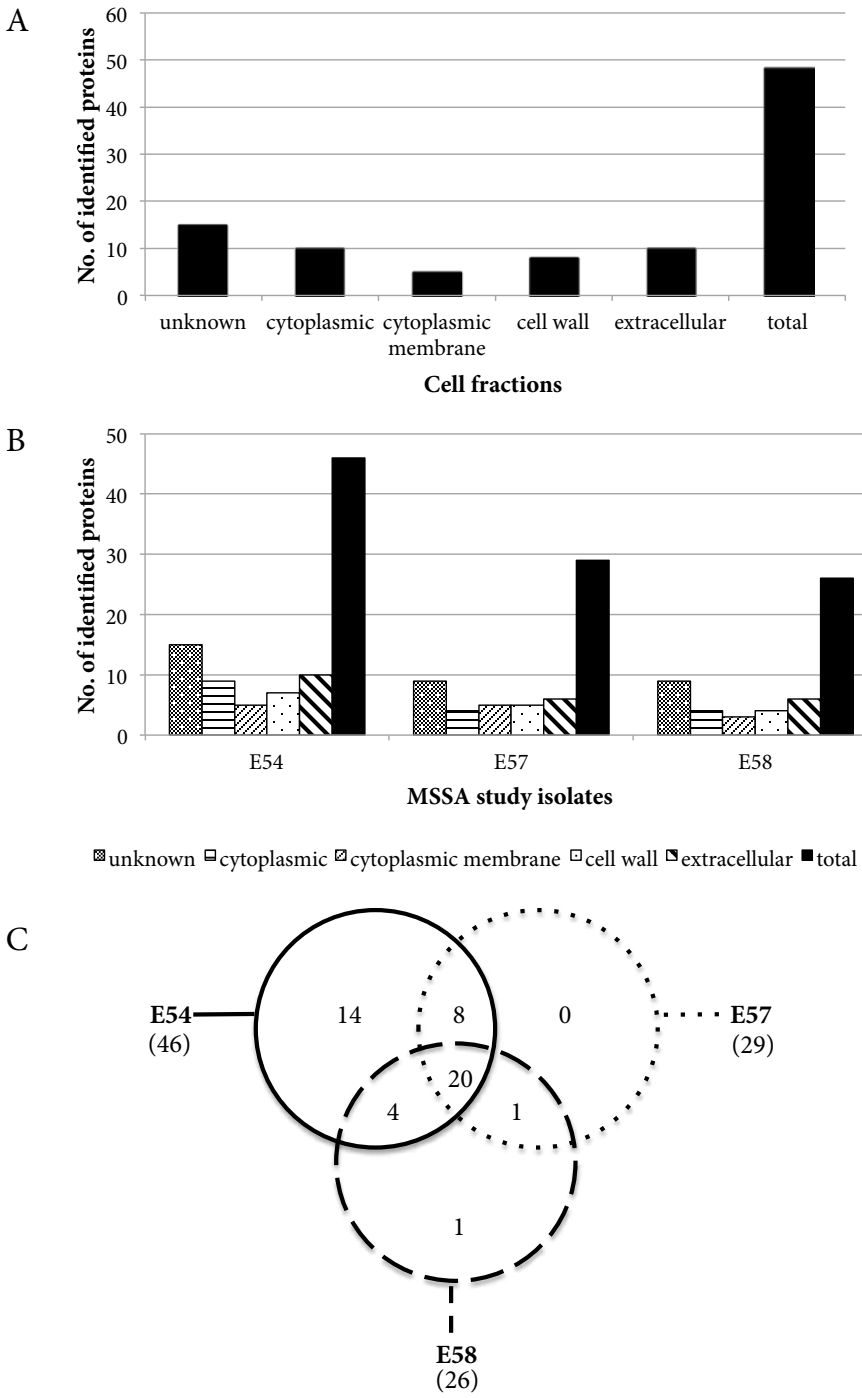

Figure 5. Overview of the proteins identified in the growth media of three MSSA study isolates. (A) Total overview of all identified proteins in the growth media of the three MSSA isolates and their predicted subcellular localization. **(B)** Identified proteins and their predicted subcellular localization indicated separately for each of the three MSSA isolates. **(C)** Venn diagram giving an overview of the numbers of consistently or uniquely identified extracellular proteins of the three MSSA isolates.

published studies used complex media and collected the proteins at a very high OD and at mostly undefined time points within the stationary phase; (ii) the present study used a gel-free approach, whereas all five previously published studies were based on two-dimensional gel electrophoresis, which is a gel-based method; and (iii) none of the previous studies performed WGS on their study isolates in combination with proteome analyses. Nevertheless, two studies investigated the transcriptome levels of RNAIII and a selection of other genes, and Wolf *et al.* also performed DNA microarray experiments [25,26]. Moreover, all these previous exoproteome analyses employed a protein database containing all publically available genome sequences at the time of investigation. This approach sets a limit to the identification of proteins, since clinical *S. aureus* isolates most probably differ in their genome repertoire compared to the utilized sequenced reference strains. To overcome this limitation, we used the study isolates' own genome sequences to create a protein database for searching the peptides identified by MS. This novel approach was explored in the present study by the creation of the protein database based on the RAST genome annotation. Though this approach led to the proper identification of a large number of proteins in all ten study isolates, we cannot exclude the possibility that some genes were not or not correctly annotated. It is thus conceivable that still some proteins produced by the investigated isolates were not included in the protein database and therefore overlooked in our analysis. However, this potential problem may also apply to protein databases based on published genome sequences, which are known to include sequencing errors and misannotations. One alternative approach would be to perform a 6-frame open reading frame translation of the genome sequences and to include the translated sequences in the final protein database. This approach is currently under investigation.

Correlation between the genome and exoproteome - building the bridge

The present combinatorial approach of WGS and proteomics was aimed at both identifying genomic and proteomic changes over time, but more importantly at enabling a potential correlation between the observed changes at these two 'omics' levels. As our present study shows, it remains challenging to pinpoint potential correlations between the genome and proteome over time. Most of *S. aureus*' virulence factors are regulated by global pleiotropic regulatory loci, such as the *agr* system, that form part of an interactive network ensuring the temporally coordinated expression of genes [63-65]. Intriguingly, to what extent these regulatory systems are responsible for the temporal expression of virulence factors *in vivo* has been debated largely in the past [65-67]. A search of the presently determined genome sequences showed that the key regulatory systems, including *sae*, *sar*, *walKR* and *vraFG* are present and, as judged by their sequences, fully functional in all ten study isolates. The most studied regulatory system, *agr*, on the contrary contained a non-synonymous SNP in three of the seven MRSA isolates. Although the SNPs occurred at two different locations, all three resulted in a premature stop codon, which probably led to a non– or malfunctioning *agr*C protein (Table 3 and Supplementary Table 3, available on request). This in turn will result in a non-functioning *agr* system leading to expression inhibition of RNAIII. Since RNAIII acts as the effector of the *agr* system and has previously been identified as a main regulator of virulence gene expression in *S. aureus*, the loss of RNAIII dramatically affects the extracellular protein pattern of *S. aureus* [43,68]. Although no transcriptome studies were performed in the present study, compelling evidence for the detrimental effect of the SNPs in *agr*C was obtained through our *in vitro* analyses. Firstly, in comparison to the other four MRSA isolates, the three MRSA isolates harbouring a premature stop codon in the *agr*C gene possessed a non-spreader phenotype. Secondly, MRSA isolates E48 and E51 did neither produce any of the three tested haemolysins in the *in vitro* assay, nor were any of these haemolysins identified in their exoproteomes. The MRSA isolate E45 also lacked these proteins in the exoproteome analyses, but showed a slight β-haemolysin-positive phenotype in the *in vitro* haemolysin assay. This observation is in line with the different locations of the SNPs within *agr*C as determined for the MRSA isolates E51 and E48 on the one hand, and the MRSA isolate E45 on the other hand. Since the SNP in MRSA isolate E45 occurred closer to the 3'-end within *agr*C it probably still permitted the production of a C-terminally truncated AgrC protein that could facilitate the weak β-haemolysin-positive phenotype. The remaining five MRSA isolates all possessed a functioning *agr* system as shown by the respective identification of the haemolysins both in the *in vitro* haemolysin

assay and the exoproteome analyses. Accordingly, these 5 MRSA isolates also showed a colony spreading phenotype on soft agar plates, which is dependent on an active *agr* system [36].

Notably, the examination of the overall gene repertoire of all MRSA study isolates revealed that they all possessed a basic virulence potential. Numerous well-known and characterized virulence genes were encoded in their genomes, including genes for enterotoxins, adhesion factors, parts of the immune evasion cluster, fibronectin-binding proteins and multiple virulence-associated cell-wall-anchored proteins (e.g. SasX, SasF). At least five possible explanations for the apparent absence of these proteins from the exoproteomes of our study isolates can be proposed: (i) these genes may have not been expressed under the experimental conditions; (ii) these genes may have been expressed at very low amounts below the detection limits of the applied methods; (iii) the encoded proteins may have been unstable under the experimental conditions; (iv) the encoded proteins may have been effectively retained in the cells; and/or (v) it was not possible to identify the encoded proteins by MS for technical reasons. The possible explanations (i) and (ii) could relate to mutations in intergenic regions or to unidentified regulators that inhibit the expression of the respective genes.

CONCLUDING REMARKS

In the post-genomic era, the availability of vast numbers of *S. aureus* genome sequences provides a strong basis for a better understanding of asymptomatic *S. aureus* carriage and the switch to staphylococcal infection. The here presented combined genomic and proteomic approach on a set of sequential *S. aureus* isolates from a patient and his two successive female partners over a 5-year time period is the first study to date implementing this type of approach. Moreover, it is the first study applying this approach to unveil potential genomic or proteomic changes in sequential isolates from one patient suffering from a common ear infection. Besides many well-known virulence factors that were identified in all ten study isolates, a large set of cytoplasmic and membrane proteins were also identified by analysis of the exoproteomes of these isolates. The latter might be of high interest for further pathogenesis studies especially in view of potential host-pathogen interactions, as these proteins could potentially play a more crucial role than previously believed. Overall, the combinatorial approach of genome and proteome analyses generates a first global view of their potential associations in different clinical *S. aureus* isolates. Notably, correlations between genomic and proteomic changes over time, based on genome-wide identifications of SNPs affecting the global regulation of virulence factors could have a profound impact on our understanding of the ecology of *S. aureus* in its commensal or pathogenic lifestyles. However, there are several potential limitations and pitfalls as exemplified in the present study. In particular, the retrospective study design sets limits to the possible interpretation of the obtained data. For future investigations, it will be preferable to implement a prospective study design with larger numbers of isolates collected at each time point of sampling. If clinically feasible, also a better-reviewed treatment regime should be implemented. Lastly, culture-independent methods for genome and proteome analysis should be applied to pinpoint the true *in vivo* situation within the patient. This would certainly provide more detailed insights into the *in vivo* lifestyles of *S. aureus*. Notwithstanding the fact that several improvements of the present study design are conceivable, we conclude that the here presented state-of-art approach paves the way on the route towards real-time investigations within the human host, with the ultimate goal to determine the lifestyles of *S. aureus* during carriage and disease.

ACKNOWLEDGEMENTS

We thank the anonymous patient who agreed to participate in this study. We thank Dr. Tjibbe Donker for generating the map displayed in Figure 2.

REFERENCES

1. Wertheim HF, Vos MC, Ott A, van Belkum A, Voss A, Kluytmans JA, et al. Risk and outcome of nosocomial *Staphylococcus aureus* bacteraemia in nasal carriers versus non-carriers. Lancet. 2004;364(9435):703–5.
2. Wertheim H, Melles D, Vos M, van Leeuwen W, van Belkum A, Verbrugh H, et al. The role of nasal carriage in *Staphylococcus aureus* infections. Lancet Infect Dis. 2005;5(12):751–62.
3. van Belkum A, Melles DC, Nouwen J, van Leeuwen WB, van Wamel W, Vos MC, et al. Co-evolutionary aspects of human colonisation and infection by *Staphylococcus aureus*. Infect Genet Evol. 2009;9(1):32–47.
4. Golubchik T, Batty EM, Miller RR, Farr H, Young BC, Larner-Svensson H, et al. Within-host evolution of *Staphylococcus aureus* during asymptomatic carriage. PLoS ONE. 2013;8(5):e61319.
5. Verkaik NJ, de Vogel CP, Boelens HA, Grumann D, Hoogenboezem T, Vink C, et al. Anti-staphylococcal humoral immune response in persistent nasal carriers and noncarriers of *Staphylococcus aureus*. J Infect Dis. 2009;199(5):625–32.
6. Verkaik NJ, Boelens HA, P VC, Tavakol M, Bode LGM, Verbrugh HA, et al. Heterogeneity of the humoral immune response following *Staphylococcus aureus* bacteremia. Eur J Clin Microbiol Infect Dis. 2010;29(5):509–18.
7. Nelson CL, Pelak K, Podgoreanu MV, Ahn SH, Scott WK, Allen AS, et al. A genome-wide association study of variants associated with acquisition of *Staphylococcus aureus* bacteremia in a healthcare setting. BMC Infectious Diseases. BMC Infectious Diseases; 2014;14(1):1–8.
8. Cameron DR, Howden BP, Peleg AY. The interface between antibiotic resistance and virulence in *Staphylococcus aureus* and its impact upon clinical outcomes. Clin Infect Dis. 2011;53(6):576–82.
9. Eiff von C, KK B, Machka K, Stammer H, Peters G. Nasal carriage as a source of *Staphylococcus aureus* bacteremia. N Engl J Med. European Centre for Disease Prevention and Control (ECDC); 2001;(344):11–6.
10. Diep BA, Carleton HA, Chang RF, Sensabaugh GF, Perdreau-Remington F. Roles of 34 virulence genes in the evolution of hospital- and Community-associated strains of methicillin-resistant *Staphylococcus aureus*. J Infect Dis. 2006 Jun 1;193(11):1495–503.
11. Otto M. Basis of virulence in community-associated methicillin-resistant *Staphylococcus aureus*. Annu Rev Microbiol. 2010;64(1):143–62.
12. Sibbald MJJB, Ziebandt AK, Engelmann S, Hecker M, de Jong A, Harmsen HJM, et al. Mapping the pathways to staphylococcal pathogenesis by comparative secretomics. Microbiol MolBiol Rev. 2006;70(3):755–88.
13. Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, et al. Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. Proc Natl Acad Sci USA. 2007;104(22):9451–6.
14. McAdam PR, Holmes A, Templeton KE, Fitzgerald JR. Adaptive evolution of *Staphylococcus aureus* during chronic endobronchial infection of a cystic fibrosis patient. PLoS ONE. 2011;6(9):e24301.
15. Strommenger B, Bartels MD, Kurt K, Layer F, Rohde SM, Boye K, et al. Evolution of methicillin-resistant *Staphylococcus aureus* towards increasing resistance. J Antimicrob Chemother. 2014;69(3):616–22.
16. Harris SR, Feil EJ, Holden MTG, Quail MA, Nickerson EK, Chantratita N, et al. Evolution of MRSA During Hospital Transmission and Intercontinental Spread. Science. 2010;327(5964):469–74.
17. Holden MTG, 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.
18. Uhlemann AC, Dordel J, Knox JR, Raven KE, Parkhill J, Holden MTG, et al. Molecular tracing of the emergence, diversification, and transmission of *S. aureus* sequence type 8 in a New York community. Proc Natl Acad Sci USA. National Acad Sciences; 2014;:201401006.
19. van Hal SJ, Steen JA, Espedido BA, Grimmond SM, Cooper MA, Holden MTG, et al. In vivo evolution of antimicrobial resistance in a series of Staphylococcus aureus patient isolates: the entire picture or a cautionary tale? J Antimicrob Chemother. 2014;69(2):363–7.
20. Peleg AY, Miyakis S, Ward DV, Earl AM, Rubio A, Cameron DR, et al. Whole genome characterization of the mechanisms of daptomycin resistance in clinical and laboratory derived isolates of *Staphylococcus aureus*. PLoS ONE. 2012;7(1):e28316.
21. Young BC, Golubchik T, Batty EM, Fung R, Larner-Svensson H, Votintseva AA, et al. Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to disease. Proc Natl Acad Sci USA. 2012;:1–6.
22. Schijffelen M, Konstantinov SR, Lina GR, Spiliopoulou I, Van Duijkeren E, Brouwer EC, et al. Whole genome analysis of epidemiologically closely related *Staphylococcus aureus* isolates. Horsburgh MJ, editor. PLoS ONE. Public Library of Science; 2013;8(10):e78340.
23. Gao W, Chua K, Davies JK, Newton HJ, Seemann T, Harrison PF, et al. Two novel point mutations in clinical *Staphylococcus aureus* reduce linezolid susceptibility and switch on the stringent response to promote persistent infection. PLoS Pathog. Public Library of Science; 2010;6(6):e1000944.
24. Howden BP, McEvoy CRE, Allen DL, Chua K, Gao W, Harrison PF, et al. Evolution of multidrug resistance during *Staphylococcus aureus* infection involves mutation of the essential two component regulator walkr. DeLeo FR, editor. PLoS Pathog. 2011;7(11):e1002359.
25. Ziebandt A-K, Kusch H, Degner M, Jaglitz S, Sibbald MJJB, Arends JP, et al. Proteomics uncovers extreme heterogeneity in the *Staphylococcus aureus* exoproteome due to genomic plasticity and variant gene regulation. Proteomics. 2010;10(8):1634–44.
26. Wolf C, Kusch H, Monecke S, Albrecht D, Holtfreter S, Eiff von C, et al. Genomic and proteomic characterization of *Staphylococcus aureus* mastitis isolates of bovine origin. Proteomics. 2011;11(12):2491–502.
27. Jarraud S, Mougel C, Thiolouse J, Lina G, Meugnier H, Forey F, et al. Relationships between *Staphylococcus aureus* genetic background, virulence factors, *agr* groups (alleles), and human disease. Infect Immun. 2002;70(2):631–41.
28. van Wamel WJB, Rooijakkers SHM, Ruyken M, van Kessel KPM, van Strijp JAG. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of staphylococcus aureus are located on -hemolysin-converting bacteriophages. J Bacteriol. American Society for Microbiology (ASM); 2006;188(4):1310–5.
29. Gilot P, Lina G, Cochard T, Poutrel B. Analysis of the genetic variability of genes encoding the rna iii-activating components agr and trap in a population of *Staphylococcus aureus* strains isolated from cows with mastitis. J Clin Microbiol. 2002;40(11):4060–7.
30. Kondo Y, Ito T, Ma XX, Watanabe S, Kreiswirth BN, Etienne J, et al. Combination of multiplex PCRs for staphylococcal cassette chromosome *mec* type assignment: rapid identification system for *mec*, *ccr*, and major differences in junkyard regions. Antimicrob Agents Chemother. 2006;51(1):264–74.
31. Chlebowicz MA, Nganou K, Kozytska S, Arends JP, Engelmann S, Grundmann H, et al. Recombination between *ccr*C genes in a type V (5C2) Staphylococcal Cassette Chromosome *mec* (SCC*mec*) of *Staphylococcus aureus* ST398 leads to conversion from methicillin resistance to methicillin susceptibility *in vivo*. Antimicrob Agents Chemother. 2010;54(2):783–91.

32. Mellmann A, Weniger T, Berssenbrugge C, Keckevoet U, Friedrich AW, Harmsen D, et al. Characterization of clonal relatedness among the natural population of *Staphylococcus aureus* strains by using *spa* sequence typing and the BURP (based upon repeat patterns) algorithm. J Clin Microbiol. 2008;46(8):2805–8.
33. Harmsen D, Claus H, Witte W, Rothgänger J, Claus H, Turnwald D, et al. Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. J Clin Microbiol. 2003;41(12):5442–8.
34. Enright MC, Day NPJ, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol. 2000;38(3):1008–15.
35. Adhikari RP, Arvidson S, Novick RP. A nonsense mutation in *agrA* accounts for the defect in *agr* expression and the avirulence of *Staphylococcus aureus* 8325-4 traP::kan. Infect Immun. 2007;75(9):4534–40.
36. Tsompanidou E, Sibbald MJJB, Chlebowicz MA, Dreisbach A, Back JW, van Dijl JM, et al. Requirement of the *agr* locus for colony spreading of *Staphylococcus aureus*. J Bacteriol. 2011;193(5):1267–72.
37. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18(5):821–9.
38. Song J, Xu Y, White S, Miller KWP, Wolinsky M. SNPsFinder--a web-based application for genome-wide discovery of single nucleotide polymorphisms in microbial genomes. Bioinformatics. 2005;21(9):2083–4.
39. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. Oxford University Press; 2012;28(4):464–9.
40. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21(16):3422–3.
41. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006;23(2):254–67.
42. 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. American Society for Microbiology; 2014;52(7):2365–70.
43. Ziebandt A-K, Becher D, Ohlsen K, Hacker J, Hecker M, Engelmann S. The influence of *agr* and σB in growth phase dependent regulation of virulence factors in *Staphylococcus aureus*. Proteomics. 2004;4(10):3034–47.
44. Dreisbach A, Hempel K, Buist G, Hecker M, Becher D, van Dijl JM. Profiling the surfacome of *Staphylococcus aureus*. Proteomics. 2010;10(17):3082–96.
45. Bonn F, Bartel J, Büttner K, Hecker M, Otto A, Becher D. Picking vanished proteins from the void: how to collect and ship/share extremely dilute proteins in a reproducible and highly efficient manner. Anal Chem. 2014;86(15):7421–7.
46. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010;26(13):1608–15.
47. Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol. 2004;338(5):1027–36.
48. Rahman O, Cummings SP, Harrington DJ, Sutcliffe IC. Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. World J Microbiol Biotechnol. Springer Netherlands; 2008;24(11):2377–82.
49. Ho P-L, Cheung C, Mak GC, Tse CWS, Ng T-K, Cheung CHY, et al. Molecular epidemiology and household transmission of community-associated methicillin-resistant *Staphylococcus aureus* in Hong Kong. Diagn Microbiol Infect Dis. 2007;57(2):145–51.
50. Knox J, van Rijen M, Uhlemann AC, Miller M, Hafer C, Vavagiakis P, et al. Community-associated methicillin-resistant Staphylococcus aureus transmission in households of infected cases: a pooled analysis of primary data from three studies across international settings. Epidemiol Infect. Cambridge University Press; 2014;:1–12.
51. Chen C-J, Unger C, Hoffmann W, Lindsay JA, Huang Y-C, Götz F. Characterization and comparison of 2 distinct epidemic community-associated methicillin-resistant *Staphylococcus aureus* Clones of ST59 lineage. PLoS ONE. 2013;8(9):e63210.
52. Song JH, Hsueh PR, Chung DR, Ko KS, Kang CI, Peck KR, et al. Spread of methicillin-resistant *Staphylococcus aureus* between the community and the hospitals in Asian countries: an ANSORP study. J Antimicrob Chemother. 2011;66(5):1061–9.
53. DeLeo FR, Otto M, Kreiswirth BN. Community-associated meticillin-resistant *Staphylococcus aureus*. Lancet. 2009;375(9725):1557–68.
54. Zanger P, Nurjadi D, Schleucher R, Scherbaum H, Wolz C, Kremsner PG, et al. Import and spread of Panton-Valentine Leukocidin-positive *Staphylococcus aureus* through nasal carriage and skin infections in travelers returning from the tropics and subtropics. Clin Infect Dis. 2012;54(4):483–92.
55. Tatem AJ, Rogers DJ, Hay SI. Global transport networks and infectious disease spread. Adv Parasitol. 2006;62:293–343.
56. McAdam PR, Templeton KE, Edwards GF, Holden MTG, Feil EJ, Aanensen DM, et al. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci USA. National Acad Sciences; 2012;109(23):9107–12.
57. Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, 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.
58. Chen C-J, Su L-H, Chiu C-H, Lin T-Y, Wong K-S, Chen Y-YM, et al. Clinical features and molecular characteristics of invasive community-acquired methicillin-resistant *Staphylococcus aureus* infections in Taiwanese children. Diagn Microbiol Infect Dis. 2007;59(3):287–93.
59. Ziebandt A-K, Weber H, Rudolph J, Schmid R, Höper D, Engelmann S, et al. Extracellular proteins of *Staphylococcus aureus* and the role of SarA and sigmaB. Proteomics. 2001;14(4):480–93.
60. Muthukrishnan G, Quinn GA, Lamers RP, Diaz C, Cole AL, Chen S, et al. Exoproteome of *Staphylococcus aureus* reveals putative determinants of nasal carriage. J Proteome Res. 2011;10(4):2064–78.
61. Pocsfalvi G, Cacace G, Cuccurullo M, Serluca G, Sorrentino A, Schlosser G, et al. Proteomic analysis of exoproteins expressed by enterotoxigenic *Staphylococcus aureus* strains. Proteomics. 2008;8(12):2462–76.
62. Nandakumar R, Nandakumar MP, Marten MR, Ross JM. Proteome analysis of membrane and cell wall associated proteins from *Staphylococcus aureus*. J Proteome Res. 2005 Apr;4(2):250–7.
63. Blevins JS, Beenken KE, Elasri MO, Hurlburt BK, Smeltzer MS. Strain-dependent differences in the regulatory roles of *sarA* and *agr* in *Staphylococcus aureus*. Infect Immun. 2002;70(2):470–80.
64. Goerke C, Gressinger M, Endler K, Breitkopf C, Wardecki K, Stern M, et al. High phenotypic diversity in infecting but not in colonizing *Staphylococcus aureus* populations. Environ Microbiol. 2007;9(12):3134–42.
65. Pragman AA, Schlievert PM. Virulence regulation in *Staphylococcus aureus*: the need for *in vivo* analysis of virulence factor regulation. FEMS Immunol Med Mic. 2004;42(2):147–54.

66. Yarwood JM, McCormick JK, Paustian ML, Kapur V, Schlievert PM. Repression of the *Staphylococcus aureus* accessory gene regulator in serum and *in vivo*. J Bacteriol. 2002;184(4):1095–101.
67. Xiong Y-Q, Bayer AS, Yeaman MR, van Wamel W, Manna AC, Cheung AL. Impacts of *sar*A and *agr* in *Staphylococcus aureus* strain Newman on fibronectin-binding protein a gene expression and fibronectin adherence capacity *in vitro* and in experimental infective endocarditis. Infect Immun. 2004;72(3):1832–6.
68. Novick RP. Autoinduction and signal transduction in the regulation of staphylococcal virulence. Mol Microbiol. 2003;48(6):1429–49.

‘The greatest glory in living lies not in never falling, but in rising every time we fall.’
Nelson Mandela (1918-2013)