ROBERT KOCH INSTITUT

Originally published as:

Nübel, U., Strommenger, B., Layer, F., Witte, W.
From types to trees: Reconstructing the spatial spread of Staphylococcus aureus based on DNA variation
(2011) International Journal of Medical Microbiology, 301 (8), pp. 614-618.

DOI: 10.1016/j.ijmm.2011.09.007

This is an author manuscript.
The definitive version is available at: http://www.sciencedirect.com

From types to trees: reconstructing the spatial spread of *Staphylococcus aureus* based on DNA variation.

Ulrich Nübel*, Birgit Strommenger, Franziska Layer, Wolfgang Witte.
Robert Koch-Institut, Wernigerode, Germany.

Corresponding author:
Ulrich Nübel
Robert Koch-Institut
Burgstr. 37
38855 Wernigerode
Germany

phone: +49-3943-679-338
email: nuebelu@rki.de

Summary. Tracing the spatial spread of pathogens is a key objective of molecular infectious disease epidemiology. Accordingly, a wide range of genotyping approaches have been used to monitor the dissemination of *Staphylococcus aureus* strains, from localized outbreaks to global spread. We provide a critical review of available methods, revealing that molecular markers currently in use for typing *S. aureus* acquire changes so slowly that they monitor evolutionary change over timescales that are largely irrelevant to epidemiology. Moreover, the more variable markers frequently do not reflect the pathogen's evolutionary history and, hence, provide potentially misleading information about spread. More recent work has demonstrated that staphylococcal evolution proceeds sufficiently fast that the dynamics of *S. aureus* spatial spread can be elucidated at great detail on the basis of genome-wide single-nucleotide polymorphisms.

Key words. MRSA, MLST, PFGE, *spa* typing, MLVA, SNPs, phylogeography

Introduction. *Staphylococcus aureus* is both, a widespread commensal colonizer of the human anterior nose and a notorious pathogen causing a range of infectious diseases. Resistance against multiple classes of antibiotics is frequently observed, and the high prevalence of methicillin-resistant *S. aureus* (MRSA) in many places is particularly worrisome. Ever since the ability of *S. aureus* was demonstrated to cause outbreaks of disease and to disseminate within and among hospitals (Williams, 1959), a great interest has existed to monitor the spatial distribution and spread of this pathogen. Accordingly, considerable effort has been invested over the decades to apply and continuously refine a multitude of methods to differentiate and track peculiar strains of *S. aureus* over spatial scales ranging from localized outbreaks to global spread (Shopsin and Kreiswirth, 2001). Nevertheless, however, the extent and the preferential routes of spatial spread of *S. aureus* and the best options for curbing it are not well understood to date (Nübel et al., 2008). Recent work has demonstrated that staphylococcal evolution proceeds sufficiently fast that the dynamics of *S. aureus* spatial spread can be elucidated on the basis of DNA sequence variation, provided the sequences are long enough (Harris et al., 2010, Nübel et al., 2010). Here, we review molecular typing methods for *S. aureus* with respect to their utility

for tracing the pathogen's spatial spread. We discuss present limitations and provide a hopeful outlook in the light of increasingly affordable genome sequencing.

Multilocus sequence typing. For multilocus sequence typing (MLST) of *S. aureus*, 5 nucleotide sequences from seven metabolic housekeeping genes are determined. Strains with identical sequences at all seven genetic loci are assigned unique 'sequence types' (ST), and clusters of closely related STs are called 'clonal complexes' (Enright et al., 2002). MLST provides excellent inter-laboratory reproducibility and data portability. Sequences can be submitted to a central, curated 10 database (available at http://saureus.mlst.net/), which in turn enables the identification of alleles and STs via queries over the internet. This database currently contains 1,851 STs based on 3,655 entries (as of 28th January 2011).

MLST has provided an extremely useful, basic understanding of the population 15 structure of *S. aureus*. Used in conjunction with the characterization of SCC*mec* variability, it has revealed the evolutionary origins of major MRSA clones (Enright et al., 2002, Robinson and Enright, 2003). Even though a large number of STs has been discovered, a limited number of clonal complexes (CC1, 5, 8, 15, 22, 30, 45, 59, 80, 97, 121) appear to predominate the *S. aureus* population. These clonal 20 complexes display a worldwide distribution, as they have been found ubiquitously through local and national surveys in many countries on every continent. Interestingly, most clonal complexes prevail among both, local populations of methicillin-susceptible *S. aureus* from non-diseased, nasal carriers (Ruimy *et al.* 2009, investigated samples from France, Algeria, Moldova, Cambodia; Fan *et al.* 25 2009, China; Sakwinska *et al.* 2009, Switzerland; Ko *et al.* 2008, Korea; Ruimy *et al.* 2008, Mali; Feil *et al.* 2003, England; Grundmann *et al.* 2002, England; Kuehnert *et al.* 2006, USA) and MRSA, where most surveys have been targeted at human infections (examples: Ko *et al.* 2005, 12 countries in Asia; Holtfreter *et al.* 2007, Germany; Coombs *et al.* 2004, Australia; Limbago *et al.* 2009, USA). Notable 30 exceptions are CC15, 97, and 121, which have rarely been detected among MRSA. Commonly, these clonal complexes each are represented by single STs that are abundant, plus a number of rare variants (Feil et al., 2003, Ruimy et al., 2008, Sakwinska et al., 2009). It follows that a small number (not much larger than the number of major CCs listed above) of STs make up a large proportion of the *S.*

aureus population, both locally and globally. As a consequence, the import of these widely distributed STs from external sources into a local setting cannot be identified. Further, the discriminatory power of MLST does not suffice to track spatial spread over local scales (Willems et al., 2011). Application of the recently estimated rate of (short-term) evolution for the MRSA core genome (Harris et al., 2010, Nübel et al., 2010) predicts that contemporary STs on average are many years old and that newly emerging and spreading strains will rarely be associated with novel STs. Striking differences between the frequencies of individual clonal complexes at different localities have been observed quite commonly, but it is usually impossible to infer whether these differences may reflect historic migration patterns of the pathogen or very recent expansions of local populations. The latter has been reported for a sample of Methicillin-susceptible isolates from Oxfordshire, UK (Fraser et al., 2005) and for carriage populations in individual kindergartens within Chengdu city, China (Fan et al., 2009).

To summarize, MLST in most cases will be of limited use for tracking the spread of individual *S. aureus* clones, due to the method's insufficient discriminatory power. It has to be noted, however, that most of the data supporting our conclusion represent the situation in industrialised countries, as large regions of the world remain undersampled. Exceptions have also been reported; for example, ST93 had long been considered unique to Australia, and ST93-MRSA-IV (the 'Queensland clone') was only recently discovered in the UK in association with long-distance travel. Additional examples include ST152, which has been reported from several countries in West Africa, but is rarely found elsewhere (Ghebremedhin et al., 2009, Ruimy et al., 2008), and hospital-associated ST22-MRSA, which to date is virtually non-existant in North America (Limbago et al., 2009) and got introduced into Portugal (Amorim et al., 2007) and Singapur (Hsu et al., 2005) only recently.

Also of note, the fact that the global diversity of *S. aureus* as resolved by MLST is not much greater than the diversity in any representative sample from a local population, is not reflected by the dataset currently available from the central MLST database (http://saureus.mlst.net/). Instead, rare STs are over-represented in the database, because it is mostly used for ST assignment, and strain information typically gets submitted only for novel STs. Moreover, there are only few studies where population

samples have been fully typed by MLST. Instead, isolates commonly are typed by DNA macrorestriction or *spa* typing (see below), and only selected isolates representing the respective, differentiated groups subsequently get MLST-typed. As a consequence, unfortunately, the MLST database is not a representative population sample.

DNA macrorestriction. DNA macrorestriction typing of *S. aureus* isolates is based on the comparison of banding patterns resulting from pulsed-field gel electrophoresis (PFGE) of *Sma*I-digested genomic DNA (Ichiyama et al., 1991, Prevost et al., 1991). PFGE was the first molecular typing tool for MRSA that was widely applied and considerable effort has been spent to standardize assays internationally, in order to overcome limitations of interlaboratory reproducibility (Chung et al., 2000, Murchan et al., 2003). Presently, PFGE is no longer considered the gold standard for typing *S. aureus* isolates (Cookson et al., 2007). It was replaced by MLST in this regard in the new millennium, because the unambiguous nature of DNA sequences is advantageous with respect to data reproducibility, interpretation, and portability. However, PFGE continues to be the most popular method for typing *S. aureus*, particularly outside Europe. Groupings based on PFGE banding patterns in most cases are concordant with MLST at the level of clonal complexes (Cookson et al., 2007, Grundmann et al., 2002, Strommenger et al., 2006). PFGE provides higher discriminatory power than MLST, as band patterns vary considerably within MLST sequence types (e. g., by up to 5 band differences in an early MLST study (Enright et al., 2000)). This variation of patterns within CCs is not always concordant with the phylogenetic relationships among isolates (Harris et al., 2010, Nübel et al., 2008), however, which is not surprising as changes in band patterns may result from a variety of genetic events, including the gain or loss of mobile genetic elements and intrachromosomal recombination (Tenover et al., 1995). Specific MLST sequence types frequently cannot be identified on the basis of PFGE band patterns (Grundmann et al., 2002, Nübel et al., 2010).

PFGE has been applied mostly for studies of local epidemiology and has proven very useful for identifying MRSA strains during outbreaks of disease (Tenover et al., 1995). One limitation is, however, that individual strains recognized by PFGE quite commonly dominate local populations, in which case the differentiation of short-term

outbreaks from endemic situations may be very difficult (Ghebremedhin et al., 2007, McDougal et al., 2003, Roberts et al., 1998, Tenover et al., 1995). At regional and national scales, PFGE frequently enabled the recognition of newly emerging MRSA clones (Deplano et al., 2000, Hookey et al., 1998, Witte et al., 1994).

5

When MRSA isolates from international sources were compared, PFGE grouped many of them according to geographic origin (Diekema et al., 2000). Accordingly, regional groupings commonly were assigned names such as 'Japan', 'New York', 'Iberian', 'South German' etc. (Murchan et al., 2003). Remarkably, however, some
10 identical band patterns have been found in multiple countries on several continents (Aires de Sousa et al., 1998, Diekema et al., 2000, Murchan et al., 2003), suggesting the widespread occurrence of a small number of MRSA strains (Oliveira et al., 2002). Indeed, the extremely close relatedness of full genome sequences from isolates affiliated to the 'Brazilian' MRSA clone (ST239) recently confirmed its intercontinental
15 spread within few decades (Harris et al., 2010). In contrast, MRSA have emerged multiple times within ST5 (encompassing PFGE types 'New York', 'EMRSA-3', 'Rheinhessen', and others). Individual MRSA clones within ST5 tend to have a more regional distribution, and PFGE analysis lumps strains that are more distantly related (Nübel et al., 2008).

20

Genetic lineages identified by PFGE that are prevalent among MRSA have also repeatedly been found among MSSA, both contemporary and historic (i. e., revived from strain archives pre-dating the era of MRSA), and these susceptible strains may be equally widespread geographically (Crisóstomo et al., 2001, Hallin et al., 2007,
25 Kuehnert et al., 2006, Witte et al., 1994). As a consequence, it may be difficult to ascertain whether a particular MRSA strain has evolved locally or been imported from another region in recent times, especially as the rate at which PFGE band patterns may change over time is unknown at present.

30 ***Spa* typing.** *Spa* typing is based on sequence analyses of the 'X-region' of the *spa* (staphylococcal protein A) gene, which consists of a highly polymorphic succession of short, sequence-variable tandem repeats (Koreen et al., 2004). Isolates with different *spa* sequences are assigned distinct *spa* types through comparison to a central database (http://spaserver.ridom.de). The discriminatory power of *spa* typing

is inferior to that of PFGE, but groupings resolved by *spa* typing and MLST correlate well at the level of clonal complexes, so that clonal designations can be made with sufficient confidence (Cookson et al., 2007, Strommenger et al., 2006). The application of *spa* typing has been simplified through the availability of a dedicated computer software, which performs half-automated quality control and data interpretation (Harmsen et al., 2003). Due to the ease of use, high interlaboratory reproducibility, and portability of DNA sequences, *spa* typing has replaced PFGE in many reference laboratories in Europe and has lead to a more de-centralized surveillance of MRSA (Cookson et al., 2007).

One limitation inherent to *spa* typing is that the high mutation rate of the *spa* locus can result in evolutionary convergence. When *spa* types were mapped onto high-resolution phylogenies based on genome-wide SNPs, several *spa* sequences were distributed in two or more distinct phylogenetic sub-lineages (Harris et al., 2010, Nübel et al., 2008). Sequence identity in these unrelated isolates most likely is the result of repeated evolution (convergence) of *spa* sequences, reflecting homoplasies. Homoplasies impair the usefulness of *spa* typing for the investigation of MRSA spatial spread, because s*pa* sequence identity may misleadingly suggest the geographic spread of individual clones. For example, identical *spa* sequences (type t138) were found in ST239 isolates from Greece and Brazil, even though genome-based phylogenetic analysis indicated they were unrelated (Harris et al., 2010) (of note, *spa* type t138 also occurs in unrelated ST30; http://spa.ridom.de).

A more serious limitation may be that -- despite the high mutation rate and the resulting extraordinary variability of the *spa* locus -- evidence grows that *spa* sequences may be too conserved to effectively monitor the spatial spread of MRSA. While methicillin-susceptible *S. aureus* display relatively greater *spa* variability (Grundmann et al., 2010, Strommenger et al., 2008), *spa* typing provides too little discriminatory power for MRSA in many situations, simply because single variants dominate the pathogen population. This limitation hampers investigations at all spatial scales, from local transmission to intercontinental spread.

Once a *spa* type has reached high abundance and become endemic in an institution or area, its dispersal can no longer be followed and local outbreaks or regional

epidemics caused by an indistinguishable strain become difficult to survey (Khandavilli et al., 2009, Shore et al., 2010, Strommenger et al., 2008). At a national scale, it is common that one or two *spa* types account for large proportions of the MRSA population. In Germany, for instance, *spa* types t032 (35%) and t003 (28%) currently dominate the MRSA population (Witte et al., 2008). In the UK, t032 (62%) is most frequent, and in France it is t008 (48%) (for information on predominant *spa* types in these and other European countries, see Grundmann et al. (2010)). In a recent study at a Europe-wide scale, only five *spa* types together accounted for 48% of MRSA isolates from 26 countries (Grundmann et al., 2010). Even in globally representative collections of MRSA, single *spa* types predominate. In the case of ST239 MRSA from five continents, *spa* type t037 accounted for 56% of isolates (Harris et al., 2010), and in the case of ST5 from six continents, t002 accounted for 54% (Nübel et al., 2008).

Sometimes, MRSA strains causing multiple infections in a spatiotemporal context are associated with novel or uncommon *spa* types. Such clusters may be recognized through routine surveillance, especially when a geographically broad perspective is supported by state-of-the-art statistical methods (Grundmann et al., 2010). However, a potentially large proportion of newly emerging strains may go undetected because they display an unmodified, ancestral *spa* sequence. For example, MRSA strain 'USA300', which causes an ongoing, large-scale epidemic of community-associated infections in the USA and is recognized by PFGE, clearly forms a distinct phylogenetic lineage (Kennedy et al., 2008). Since USA300 displays *spa* type t008, however, *spa* typing does not distinguish it from other t008/ST8 strains that occur the world over. Similarly, within ST239 and ST5, multiple geographically confined sub-clones were apparent that were not associated with specific *spa* types (Harris et al., 2010, Nübel et al., 2008). In summary, it is unclear, at present, what proportion of emerging MRSA strains are associated with unique *spa* sequences, and, hence, may actually be recognized on the basis of *spa* typing.

Multilocus variable number tandem repeat analysis (MLVA). MLVA exploits length variation of repetitive DNA in bacterial genomes; loci to be investigated are amplified by PCR -- frequently in a multiplex fashion -- and subsequently analysed electrophoretically. Highly discriminatory MLVA typing schemes are available for

various bacterial pathogens, and several attempts have been made in the recent past to establish such a method for *S. aureus*. Depending on the protocol, the number of genetic loci investigated for MLVA typing of *S. aureus* varied from five (Sabat et al., 2003) to fourteen (Pourcel et al., 2009). While earlier attempts had relied on agarose gel electrophoresis (Hardy et al., 2006, Sabat et al., 2003), capillary electrophoresis allows for semi-automated fragment length determination and subsequent cluster analysis (Francois et al., 2005, Pourcel et al., 2009, Schouls et al., 2009).

For MRSA, MLVA provides discriminatory power comparable to PFGE (Sabat et al., 2003, Schouls et al., 2009) and, reportedly, may be usefully applied for identifying outbreaks in hospitals (Hardy et al., 2006, Holmes et al., 2010, Ikawaty et al., 2008). In a sample from diverse geographic sources on several continents, the majority of *S. aureus* isolates displayed unique profiles (81 MLVA types among 104 isolates), but in some cases identical profiles were found in multiple countries (Conceicao et al., 2009). However, genome re-sequencing of eight of those isolates previously investigated by Conceicao *et al.* (2009) revealed that groupings obtained by MLVA were not necessarily concordant with the isolates' phylogenetic relationships (Harris et al., 2010). Some isolates with similar MLVA profiles (differing at a single locus) turned out to be only distantly related to each other as they were affiliated to distinct major clades on the basis of genome-wide SNPs. Conversely, extremely closely related isolates had yielded MLVA profiles that differed at up to three loci and on that basis had not been grouped together (Conceicao et al., 2009, Harris et al., 2010).

We conclude that MLVA in the case of ST239 does not accurately reflect MRSA evolutionary history and, hence, may also mislead investigations of the history of MRSA spatial spread. The application of MLVA may be more appropriate for studies of short-term epidemiology within localized settings, where its high discriminatory power might be an advantage. Technically, MLVA provides for a low cost typing tool amenable to high-throughput, but the procedure will require standardization of multiplex-PCR amplification, electrophoresis, and data analysis to demonstrate interlaboratory reproducibility.

Genome-wide SNPs enable the detailed reconstruction of spatial spread. Recent investigations based on long DNA sequences from globally representative

isolate collections revealed that MRSA genomes had accumulated one point mutation per every six to eight weeks (Harris et al., 2010, Nübel et al., 2010, Smyth et al., 2010). This short-term evolutionary rate was virtually identical in two different clones of MRSA and much faster than previously acknowledged. Importantly, such a measurable accumulation of DNA variation over epidemiologically relevant timescales can be exploited to accurately infer the spatial and temporal dynamics of pathogen spread (Cottam et al., 2008, Hue et al., 2005, Jombart et al., 2011, Lemey et al., 2009). The genealogical relationships among isolates from a single outbreak or epidemic can be reconstructed on the basis of serially sampled DNA sequences, if these are sufficiently variable (Jombart et al., 2011). On this basis, and by considering sampling dates and locations, the time course and pathways of transmission and spatial spread can then be determined and visualized (Jombart et al., 2011, Lemey et al., 2009).

The application of this ancestry-based tracking approach to MRSA ST225 recently revealed that this clone got introduced into Europe in the mid 1990s and that its dispersal among hospitals across Central Europe was more rapid and complex than reflected by *spa* sequences (Nübel et al., 2010). An investigation of ST239 suggested it may even be possible to trace exact chains of person-to-person transmission, provided that DNA sequences of sufficient length (i. e., whole genomes) are analysed (Harris et al., 2010). Hence, *S. aureus* genome sequences carry information about the isolates' ancestral history ranging from weeks to many years, and about their geographical history at scales from localized outbreaks to global spread.

These recent findings provide very promising outlooks for the study of MRSA spatial distribution and spread. DNA sequencing recently became more cost-effective and more rapid by orders of magnitude, today enabling sequencing projects at scales that were deemed unapproachable just a few years ago (Harris et al., 2010, Parkhill, 2008). As sequencing costs continue to drop, it seems likely that genome sequencing will soon become the new gold standard for genotyping *S. aureus* (and other bacterial pathogens) and that it may develop into a first-line typing tool.

Obviously, whole-genome sequences provide ultimate discriminatory power for typing purposes. Beyond a mere quantitative advance, however, the sampling and comparative analyses of multiple long DNA sequences also mean a conceptual progress in staphylococcal molecular epidemiology. Traditionally, *S. aureus* isolates are classified into discrete 'types' by any of the methods discussed above and spreading routes are attempted to be inferred from the spatial distribution of these units (Grundmann et al., 2010). However, the respective molecular markers in use reflect evolutionary relationships among isolates to a varying extent and in most cases provide information about evolutionary events that precede the epidemiological process under study (see above). In fact, the invariability of specific types to be tracked down during outbreak analyses is considered a desirable feature of such typing data (van Belkum et al., 2007). In contrast, the comparative analysis of long DNA sequences which have accumulated variation during an outbreak may enable an ancestry-based reconstruction of spread, provide information even about unsampled ancestral sequences, and potentially unravel preferential routes of transmission. These inferences will be based on reconstructed trees that reflect genealogical relationships between sequences. Hence, this approach will imply a conceptual shift 'from types to trees'.

References.

Aires de Sousa, M., Sanches, I.S., Ferro, M.L., Vaz, M.J., Saraiva, Z., Tendeiro, T., Serra, J., de Lencastre, H., 1998. Intercontinental spread of a multidrug-resistant methicillin-resistant Staphylococcus aureus clone. J Clin Microbiol. 36, 2590-2596.

Amorim, M.L., Faria, N.A., Oliveira, D.C., Vasconcelos, C., Cabeda, J.C., Mendes, A.C., Calado, E., Castro, A.P., Ramos, M.H., Amorim, J.M., de Lencastre, H., 2007. Changes in the clonal nature and antibiotic resistance profiles of methicillin-resistant *Staphylococcus aureus* isolates associated with spread of the EMRSA-15 clone in a tertiary care Portuguese hospital. J Clin Microbiol. 45, 2881-2888.

Chung, M., de Lencastre, H., Matthews, P., Tomasz, A., Adamsson, I., Aires de Sousa, M., Camou, T., Cocuzza, C., Corso, A., Couto, I., Dominguez, A., Gniadkowski, M., Goering, R., Gomes, A., Kikuchi, K., Marchese, A., Mato, R., Melter, O., Oliveira, D., Palacio, R., Sa-Leao, R., Santos Sanches, I., Song, J.H., Tassios, P.T., Villari, P., 2000. Molecular typing of methicillin-resistant Staphylococcus aureus by pulsed-field gel electrophoresis: comparison of results obtained in a multilaboratory effort using identical protocols and MRSA strains. Microbial drug resistance. 6, 189-198.

Conceicao, T., Aires de Sousa, M., de Lencastre, H., 2009. Staphylococcal interspersed repeat unit typing of Staphylococcus aureus: evaluation of a new multilocus variable-number tandem-repeat analysis typing method. J Clin Microbiol. 47, 1300-1308.

Cookson, B.D., Robinson, D.A., Monk, A.B., Murchan, S., Deplano, A., de Ryck, R., Struelens, M.J., Scheel, C., Fussing, V., Salmenlinna, S., Vuopio-Varkila, J., Cuny, C., Witte, W., Tassios, P.T., Legakis, N.J., van Leeuwen, W., van Belkum, A., Vindel, A., Garaizar, J., Haeggman, S., Olsson-Liljequist, B., Ransjo, U., Muller-Premru, M., Hryniewicz, W., Rossney, A., O'Connell, B., Short, B.D., Thomas, J., O'Hanlon, S., Enright, M.C.,

2007. Evaluation of molecular typing methods in characterizing a European collection of epidemic methicillin-resistant *Staphylococcus aureus* strains: the HARMONY collection. J Clin Microbiol. 45, 1830-1837.

Coombs, G.W., Nimmo, G.R., Bell, J.M., Huygens, F., O'Brien, F.G., Malkowski, M.J., Pearson, J.C., Stephens, A.J., Giffard, P.M., 2004. Genetic diversity among community methicillin-resistant Staphylococcus aureus strains causing outpatient infections in Australia. J Clin Microbiol. 42, 4735-4743.

Cottam, E.M., Wadsworth, J., Shaw, A.E., Rowlands, R.J., Goatley, L., Maan, S., Maan, N.S., Mertens, P.P., Ebert, K., Li, Y., Ryan, E.D., Juleff, N., Ferris, N.P., Wilesmith, J.W., Haydon, D.T., King, D.P., Paton, D.J., Knowles, N.J., 2008. Transmission pathways of foot-and-mouth disease virus in the United Kingdom in 2007. PLoS Pathog. 4, e1000050.

Crisóstomo, M.I., Westh, H., Tomasz, A., Chung, M., Oliveira, D.C., de Lencastre, H., 2001. The evolution of methicillin resistance in *Staphylococcus aureus*: similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. Proc Natl Acad Sci U S A. 98, 9865-9870.

Deplano, A., Witte, W., van Leeuwen, W.J., Brun, Y., Struelens, M.J., 2000. Clonal dissemination of epidemic methicillin-resistant Staphylococcus aureus in Belgium and neighboring countries. Clin Microbiol Infect. 6, 239-245.

Diekema, D.J., Pfaller, M.A., Turnidge, J., Verhoef, J., Bell, J., Fluit, A.C., Doern, G.V., Jones, R.N., 2000. Genetic relatedness of multidrug-resistant, methicillin (oxacillin)-resistant Staphylococcus aureus bloodstream isolates from SENTRY Antimicrobial Resistance Surveillance Centers worldwide, 1998. Microbial drug resistance. 6, 213-221.

Enright, M.C., Day, N.P., Davies, C.E., Peacock, S.J., Spratt, B.G., 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol. 38, 1008-1015.

Enright, M.C., Robinson, D.A., Randle, G., Feil, E.J., Grundmann, H., Spratt, B.G., 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci U S A. 99, 7687-7692.

Fan, J., Shu, M., Zhang, G., Zhou, W., Jiang, Y., Zhu, Y., Chen, G., Peacock, S.J., Wan, C., Pan, W., Feil, E.J., 2009. Biogeography and virulence of Staphylococcus aureus. PLoS ONE. 4, e6216.

Feil, E.J., Cooper, J.E., Grundmann, H., Robinson, D.A., Enright, M.C., Berendt, T., Peacock, S.J., Smith, J.M., Murphy, M., Spratt, B.G., Moore, C.E., Day, N.P., 2003. How clonal is *Staphylococcus aureus*? J Bacteriol. 185, 3307-3316.

Francois, P., Huyghe, A., Charbonnier, Y., Bento, M., Herzig, S., Topolski, I., Fleury, B., Lew, D., Vaudaux, P., Harbarth, S., van Leeuwen, W., van Belkum, A., Blanc, D.S., Pittet, D., Schrenzel, J., 2005. Use of an automated multiple-locus, variable-number tandem repeat-based method for rapid and high-throughput genotyping of Staphylococcus aureus isolates. J Clin Microbiol. 43, 3346-3355.

Fraser, C., Hanage, W.P., Spratt, B.G., 2005. Neutral microepidemic evolution of bacterial pathogens. Proc Natl Acad Sci U S A. 102, 1968-1973.

Ghebremedhin, B., König, W., Witte, W., Hardy, K.J., Hawkey, P.M., König, B., 2007. Subtyping of ST22-MRSA-IV (Barnim epidemic MRSA strain) at a university clinic in Germany from 2002 to 2005. J Med Microbiol. 56, 365-375.

Ghebremedhin, B., Olugbosi, M.O., Raji, A.M., Layer, F., Bakare, R.A., Konig, B., Konig, W., 2009. Emergence of a community-associated methicillin-resistant Staphylococcus aureus strain with a unique resistance profile in Southwest Nigeria. J Clin Microbiol. 47, 2975-2980.

Grundmann, H., Aanensen, D.M., van den Wijngaard, C.C., Spratt, B.G., Harmsen, D., Friedrich, A.W., 2010. Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: a molecular-epidemiological analysis. PLoS Med. 7, e1000215.

Grundmann, H., Hori, S., Enright, M.C., Webster, C., Tami, A., Feil, E.J., Pitt, T., 2002. Determining the genetic structure of the natural population of Staphylococcus aureus: a comparison of multilocus sequence typing with

pulsed-field gel electrophoresis, randomly amplified polymorphic DNA analysis, and phage typing. J Clin Microbiol. 40, 4544-4546.

Hallin, M., Denis, O., Deplano, A., De Mendonca, R., De Ryck, R., Rottiers, S., Struelens, M.J., 2007. Genetic
5 relatedness between methicillin-susceptible and methicillin-resistant *Staphylococcus aureus*: results of a national survey. J Antimicrob Chemother. 59, 465-472.

Hardy, K.J., Oppenheim, B.A., Gossain, S., Gao, F., Hawkey, P.M., 2006. Use of variations in staphylococcal interspersed repeat units for molecular typing of methicillin-resistant Staphylococcus aureus strains. J Clin
10 Microbiol. 44, 271-273.

Harmsen, D., Claus, H., Witte, W., Rothgänger, J., Claus, H., Turnwald, D., Vogel, U., 2003. 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. 41, 5442-5448.

15

Harris, S.R., Feil, E.J., Holden, M.T., Quail, M.A., Nickerson, E.K., Chantratita, N., Gardete, S., Tavares, A., Day, N., Lindsay, J.A., Edgeworth, J.D., de Lencastre, H., Parkhill, J., Peacock, S.J., Bentley, S.D., 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science. 327, 469-474.

20 Holmes, A., Edwards, G.F., Girvan, E.K., Hannant, W., Danial, J., Fitzgerald, J.R., Templeton, K.E., 2010. Comparison of two multilocus variable-number tandem-repeat methods and pulsed-field gel electrophoresis for differentiating highly clonal methicillin-resistant Staphylococcus aureus isolates. J Clin Microbiol. 48, 3600-3607.

25 Holtfreter, S., Grumann, D., Schmudde, M., Nguyen, H.T., Eichler, P., Strommenger, B., Kopron, K., Kolata, J., Giedrys-Kalemba, S., Steinmetz, I., Witte, W., Bröker, B.M., 2007. Clonal distribution of superantigen genes in clinical *Staphylococcus aureus* isolates. J Clin Microbiol. 45, 2669-2680.

Hookey, J.V., Richardson, J.F., Cookson, B.D., 1998. Molecular typing of Staphylococcus aureus based on PCR
30 restriction fragment length polymorphism and DNA sequence analysis of the coagulase gene. J Clin Microbiol. 36, 1083-1089.

Hsu, L.Y., Koh, T.H., Singh, K., Kang, M.L., Kurup, A., Tan, B.H., 2005. Dissemination of multisusceptible methicillin-resistant *Staphylococcus aureus* in Singapore. J Clin Microbiol. 43, 2923-2925.

35

Hue, S., Pillay, D., Clewley, J.P., Pybus, O.G., 2005. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. Proc Natl Acad Sci U S A. 102, 4425-4429.

Ichiyama, S., Ohta, M., Shimokata, K., Kato, N., Takeuchi, J., 1991. Genomic DNA fingerprinting by pulsed-
40 field gel electrophoresis as an epidemiological marker for study of nosocomial infections caused by methicillin-resistant Staphylococcus aureus. J Clin Microbiol. 29, 2690-2695.

Ikawaty, R., Willems, R.J., Box, A.T., Verhoef, J., Fluit, A.C., 2008. Novel multiple-locus variable-number tandem-repeat analysis method for rapid molecular typing of human Staphylococcus aureus. J Clin Microbiol.
45 46, 3147-3151.

Jombart, T., Eggo, R.M., Dodd, P.J., Balloux, F., 2011. Reconstructing disease outbreaks from genetic data: a graph approach. Heredity. 106, 383-390.

50 Kennedy, A.D., Otto, M., Braughton, K.R., Whitney, A.R., Chen, L., Mathema, B., Mediavilla, J.R., Byrne, K.A., Parkins, L.D., Tenover, F.C., Kreiswirth, B.N., Musser, J.M., DeLeo, F.R., 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. Proc Natl Acad Sci U S A. 105, 1327-1332.

55 Khandavilli, S., Wilson, P., Cookson, B., Cepeda, J., Bellingan, G., Brown, J., 2009. Utility of spa typing for investigating the local epidemiology of MRSA on a UK intensive care ward. The Journal of hospital infection. 71, 29-35.

Ko, K.S., Lee, J.Y., Baek, J.Y., Peck, K.R., Rhee, J.Y., Kwon, K.T., Heo, S.T., Ahn, K.M., Song, J.H., 2008.
60 Characterization of Staphylococcus aureus nasal carriage from children attending an outpatient clinic in Seoul, Korea. Microbial drug resistance. 14, 37-44.

Ko, K.S., Lee, J.Y., Suh, J.Y., Oh, W.S., Peck, K.R., Lee, N.Y., Song, J.H., 2005. Distribution of major genotypes among methicillin-resistant *Staphylococcus aureus* clones in Asian countries. J Clin Microbiol. 43, 421-426.

5 Koreen, L., Ramaswamy, S.V., Graviss, E.A., Naidich, S., Musser, J.M., Kreiswirth, B.N., 2004. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. J Clin Microbiol. 42, 792-799.

Kuehnert, M.J., Kruszon-Moran, D., Hill, H.A., McQuillan, G., McAllister, S.K., Fosheim, G., McDougal, L.K.,
10 Chaitram, J., Jensen, B., Fridkin, S.K., Killgore, G., Tenover, F.C., 2006. Prevalence of *Staphylococcus aureus* nasal colonization in the United States, 2001-2002. J Infect Dis. 193, 172-179.

Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput Biol. 5, e1000520.

15

Limbago, B., Fosheim, G.E., Schoonover, V., Crane, C.E., Nadle, J., Petit, S., Heltzel, D., Ray, S.M., Harrison, L.H., Lynfield, R., Dumyati, G., Townes, J.M., Schaffner, W., Mu, Y., Fridkin, S.K., 2009. Characterization of methicillin-resistant *Staphylococcus aureus* isolates collected in 2005 and 2006 from patients with invasive disease: a population-based analysis. J Clin Microbiol. 47, 1344-1351.

20

McDougal, L.K., Steward, C.D., Killgore, G.E., Chaitram, J.M., McAllister, S.K., Tenover, F.C., 2003. Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. J Clin Microbiol. 41, 5113-5120.

25 Murchan, S., Kaufmann, M.E., Deplano, A., de Ryck, R., Struelens, M., Zinn, C.E., Fussing, V., Salmenlinna, S., Vuopio-Varkila, J., El Solh, N., Cuny, C., Witte, W., Tassios, P.T., Legakis, N., van Leeuwen, W., van Belkum, A., Vindel, A., Laconcha, I., Garaizar, J., Haeggman, S., Olsson-Liljequist, B., Ransjo, U., Coombs, G., Cookson, B., 2003. Harmonization of pulsed-field gel electrophoresis protocols for epidemiological typing of strains of methicillin-resistant *Staphylococcus aureus*: a single approach developed by consensus in 10 European
30 laboratories and its application for tracing the spread of related strains. J Clin Microbiol. 41, 1574-1585.

Nübel, U., Dordel, J., Kurt, K., Strommenger, B., Westh, H., Shukla, S.K., Zemlickova, H., Leblois, R., Wirth, T., Jombart, T., Balloux, F., Witte, W., 2010. A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. PLoS Pathogens 6, e1000855.

35

Nübel, U., Roumagnac, P., Feldkamp, M., Song, J.H., Ko, K.S., Huang, Y.C., Coombs, G., Ip, M., Westh, H., Skov, R., Struelens, M.J., Goering, R.V., Strommenger, B., Weller, A., Witte, W., Achtman, M., 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus* Proc Natl Acad Sci U S A. 105, 14130-14135.

40

Oliveira, D.C., Tomasz, A., de Lencastre, H., 2002. Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant Staphylococcus aureus. The Lancet infectious diseases. 2, 180-189.

Parkhill, J., 2008. Time to remove the model organism blinkers. Trends in microbiology. 16, 510-511.

45

Pourcel, C., Hormigos, K., Onteniente, L., Sakwinska, O., Deurenberg, R.H., Vergnaud, G., 2009. Improved multiple-locus variable-number tandem-repeat assay for Staphylococcus aureus genotyping, providing a highly informative technique together with strong phylogenetic value. J Clin Microbiol. 47, 3121-3128.

50 Prevost, G., Pottecher, B., Dahlet, M., Bientz, M., Mantz, J.M., Piemont, Y., 1991. Pulsed field gel electrophoresis as a new epidemiological tool for monitoring methicillin-resistant Staphylococcus aureus in an intensive care unit. The Journal of hospital infection. 17, 255-269.

Roberts, R.B., de Lencastre, A., Eisner, W., Severina, E.P., Shopsin, B., Kreiswirth, B.N., Tomasz, A., 1998.
55 Molecular epidemiology of methicillin-resistant Staphylococcus aureus in 12 New York hospitals. MRSA Collaborative Study Group. J Infect Dis. 178, 164-171.

Robinson, D.A., Enright, M.C., 2003. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother. 47, 3926-3934.

60

Ruimy, R., Armand-Lefevre, L., Barbier, F., Ruppe, E., Cocojaru, R., Mesli, Y., Maiga, A., Benkalfat, M., Benchouk, S., Hassaine, H., Dufourcq, J.B., Nareth, C., Sarthou, J.L., Andremont, A., Feil, E.J., 2009.

Comparisons between geographically diverse samples of carried *Staphylococcus aureus*. J Bacteriol. 191, 5577-5583.

Ruimy, R., Maiga, A., Armand-Lefevre, L., Maiga, I., Diallo, A., Koumare, A.K., Ouattara, K., Soumare, S., Gaillard, K., Lucet, J.C., Andremont, A., Feil, E.J., 2008. The carriage population of *Staphylococcus aureus* from Mali is composed of a combination of pandemic clones and the divergent Panton-Valentine leukocidin-positive genotype ST152. J Bacteriol. 190, 3962-3968.

Sabat, A., Krzyszton-Russjan, J., Strzalka, W., Filipek, R., Kosowska, K., Hryniewicz, W., Travis, J., Potempa, J., 2003. New method for typing Staphylococcus aureus strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates. J Clin Microbiol. 41, 1801-1804.

Sakwinska, O., Kuhn, G., Balmelli, C., Francioli, P., Giddey, M., Perreten, V., Riesen, A., Zysset, F., Blanc, D.S., Moreillon, P., 2009. Genetic diversity and ecological success of *Staphylococcus aureus* strains colonizing humans. Appl Environ Microbiol. 75, 175-183.

Schouls, L.M., Spalburg, E.C., van Luit, M., Huijsdens, X.W., Pluister, G.N., van Santen-Verheuvel, M.G., van der Heide, H.G., Grundmann, H., Heck, M.E., de Neeling, A.J., 2009. Multiple-locus variable number tandem repeat analysis of Staphylococcus aureus: comparison with pulsed-field gel electrophoresis and spa-typing. PLoS ONE. 4, e5082.

Shopsin, B., Kreiswirth, B.N., 2001. Molecular epidemiology of methicillin-resistant Staphylococcus aureus. Emerg Infect Dis. 7, 323-326.

Shore, A.C., Rossney, A.S., Kinnevey, P.M., Brennan, O.M., Creamer, E., Sherlock, O., Dolan, A., Cunney, R., Sullivan, D.J., Goering, R.V., Humphreys, H., Coleman, D.C., 2010. Enhanced Discrimination of Highly-Clonal ST22-Methicillin Resistant *Staphylococcus aureus* (MRSA)-IV Isolates Achieved by Combining spa, dru and Pulsed Field Gel Electrophoresis (PFGE) Typing Data. J Clin Microbiol.

Smyth, D.S., McDougal, L.K., Gran, F.W., Manoharan, A., Enright, M.C., Song, J.H., de Lencastre, H., Robinson, D.A., 2010. Population structure of a hybrid clonal group of methicillin-resistant *Staphylococcus aureus*, ST239-MRSA-III. PLoS ONE. 5, e8582.

Strommenger, B., Braulke, C., Heuck, D., Schmidt, C., Pasemann, B., Nübel, U., Witte, W., 2008. *Spa*-typing of *Staphylococcus aureus* as a frontline tool in epidemiological typing. J. Clin. Microbiol. 46, 574-581.

Strommenger, B., Kettlitz, C., Weniger, T., Harmsen, D., Friedrich, A., Witte, W., 2006. Assignment of *Staphylococcus* isolates to groups by *spa* typing, *Sma*I macrorestriction analysis, and multilocus sequence typing. J. Clin. Microbiol. . 44, 2533-2540.

Tenover, F.C., Arbeit, R.D., Goering, R.V., Mickelsen, P.A., Murray, B.E., Persing, D.H., Swaminathan, B., 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol. 33, 2233-2239.

van Belkum, A., Tassios, P.T., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N.K., Fussing, V., Green, J., Feil, E., Gerner-Smidt, P., Brisse, S., Struelens, M., 2007. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 13 Suppl 3, 1-46.

Willems, R.J., Hanage, W.P., Bessen, D.E., Feil, E.J., 2011. Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. FEMS Microbiol Rev.

Williams, R.E.O., 1959. Epidemic staphylococci. Lancet. 1, 190-195.

Witte, W., Cuny, C., Braulke, C., Heuck, D., 1994. Clonal dissemination of two MRSA strains in Germany. Epidemiol. Infect. . 113, 67-73.

Witte, W., Cuny, C., Halle, E., Mauch, H., Wagner, J., 1994. Methicillin resistance in an epidemic *Staphylococcus aureus* strain with genomic fingerprints corresponding to those of a sensitive strain in the community. Med. Microbiol. Lett. . 3, 388-395.

Witte, W., Klare, I., Nübel, U., Strommenger, B., Werner, G., 2008. Emergence and spread of antibiotic resistant Gram positive bacterial pathogens. International Journal of Medical Microbiology. 298, 365-377.

(Ruimy et al., 2009)
(Fan et al., 2009)
(Sakwinska et al., 2009)
(Ko et al., 2008)
5 (Ruimy et al., 2008)
(Feil et al., 2003)
(Grundmann et al., 2002)
(Kuehnert et al., 2006)
(Ko et al., 2005)
10 (Coombs et al., 2004)
(Holtfreter et al., 2007)
(Limbago et al., 2009)

15