



A genomic portrait of the emergence, evolution and global spread of a methicillin resistant *Staphylococcus aureus* pandemic

Matthew T. G. Holden, Li-Yang Hsu, Kevin Kurt, et al.

Genome Res. published online January 8, 2013

Access the most recent version at doi:[10.1101/gr.147710.112](https://doi.org/10.1101/gr.147710.112)

P<P Published online January 8, 2013 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Comprehensive immune receptor profiling.
Discover the **DriverMap™ AIR Assay** difference.

LEARN
MORE



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

© 2013, Published by Cold Spring Harbor Laboratory Press

A genomic portrait of the emergence, evolution and global spread of a methicillin resistant *Staphylococcus aureus* pandemic

Matthew T.G. Holden^{1*}, Li-Yang Hsu^{1,2}, Kevin Kurt³, Lucy A. Weinert^{4,†}, Alison E. Mather¹, Simon R. Harris¹, Birgit Strommenger³, Franziska Layer³, Wolfgang Witte³, Herminia de Lencastre^{5,6}, Robert Skov⁷, Henrik Westh^{8,9}, Helena Žemličková¹⁰, Geoffrey Coombs¹¹, Angela M. Kearns¹², Robert L.R. Hill¹², Jonathan Edgeworth¹³, Ian Gould¹⁴, Vanya Gant¹⁵, Jonathan Cooke¹⁶, Giles F. Edwards¹⁷, Paul R. McAdam¹⁸, Kate E. Templeton¹⁹, Angela McCann²⁰, Zhemin Zhou²⁰, Santiago Castillo-Ramírez²¹, Edward J. Feil²¹, Lyndsey O. Hudson⁴, Mark C. Enright^{4,‡}, Francois Balloux^{4,+}, David M. Aanensen⁴, Brian G. Spratt⁴, J. Ross Fitzgerald¹⁸, Julian Parkhill¹, Mark Achtman^{20*}, Stephen D. Bentley¹, Ulrich Nübel^{3*}.

¹ The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.

² National University Health System, 1E Kent Ridge Road, NUHS Tower Block Level 10, Singapore 119228, Singapore.

³ Robert Koch Institute, Wernigerode, Germany.

⁴ Department of Infectious Disease Epidemiology, Imperial College London, St. Mary's Hospital Campus, London, UK.

⁵ Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica (ITQB), Universidade Nova de Lisboa, Oeiras, Portugal.

⁶ Laboratory of Microbiology, The Rockefeller University, New York, USA.

⁷ Statens Serum Institut, Copenhagen, Denmark.

⁸ Hvidovre Hospital, Hvidovre, Denmark.

⁹ University of Copenhagen, Faculty of Health, Copenhagen, Denmark.

¹⁰ National Institute of Public Health, Prague, Czech Republic.

¹¹ Royal Perth Hospital, Perth, Western Australia.

¹² Microbiology Services, Health Protection Agency, 61 Colindale Avenue, London, UK.

¹³ Department of Nephrology & Transplantation, King's College London, Guy's, King's & St Thomas' Medical School, Guy's Hospital, London, UK.

¹⁴ Aberdeen Royal Infirmary, Aberdeen, UK.

¹⁵ University College London Hospital, London, UK.

¹⁶ Division of Infectious Diseases, Department of Medicine, Imperial College London, London, UK.

¹⁷ Scottish MRSA Reference Laboratory, NHS Greater Glasgow and Clyde, Stobhill Hospital, Glasgow, UK.

¹⁸ The Roslin Institute and Edinburgh Infectious Diseases, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, UK.

¹⁹ Microbiology, Royal Infirmary of Edinburgh, Edinburgh, UK.

²⁰ Environmental Research Institute, University College Cork, Lee Road, Cork, Ireland.

²¹ Department of Biology and Biochemistry, University of Bath, South Building, Claverton Down, Bath, UK.

† present address: University of Cambridge, UK.

‡ present address: University of Bath, UK.

+ present address: University College, London, UK.

Running title: Drug resistance and the genesis of a MRSA pandemic

Corresponding Authors: Ulrich Nübel, Robert Koch Institute, Burgstr. 37, 38855 Wernigerode, Germany. email: nuebelu@rki.de

Mark Achtman, Environmental Research Institute, University College Cork, Lee Road, Cork, Ireland. email: M.Achtman@ucc.ie

Matthew Holden, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. Email: mh3@sanger.ac.uk

Abstract

The widespread use of antibiotics in association with high-density clinical care has driven the emergence of drug-resistant bacteria that are adapted to thrive in hospitalised patients. Of particular concern are globally disseminated methicillin-resistant *Staphylococcus aureus* (MRSA) clones that cause outbreaks and epidemics associated with healthcare. The most rapidly spreading and tenacious healthcare-associated clone in Europe currently is EMRSA-15, a lineage that was first detected in the UK in the early 1990s and subsequently spread throughout Europe and beyond. To understand the genetic events that have accompanied the emergence of the EMRSA-15 pandemic, we obtained genome sequences for 193 isolates that were chosen for their geographical and temporal diversity, and belong to the same multilocus sequence type as EMRSA-15. Using phylogenomic methods, we were able to show that the current pandemic population of EMRSA-15 descends from a healthcare-associated MRSA epidemic that spread through England in the 1980s, which had itself previously emerged from a primarily community-associated methicillin-sensitive population. The emergence of fluoroquinolone resistance in this EMRSA-15 sub-clone in the English Midlands during the mid-1980s appears to have played a key role in triggering pandemic spread, and occurred shortly after the first clinical trials of this drug. Genome-based coalescence analysis estimated that the population of this sub-clone over the last twenty years has grown four times faster than its progenitor. Using comparative genomic analysis we were able to identify the molecular genetic basis of 99.8% of the antimicrobial resistance phenotypes of the isolates, highlighting the potential of pathogen genome sequencing as a diagnostic tool. We document the genetic changes associated with adaptation to the hospital

environment and with increasing drug resistance over time, and how MRSA evolution likely has been influenced by country-specific drug use regimens.

Introduction

Methicillin-resistant *Staphylococcus aureus* (MRSA) remains a major global cause of healthcare-associated infections since it was first described five decades ago following the introduction of penicillinase-stable β -lactam antibiotics into clinical practice (Klein et al. 2007). During this period, MRSA strains have emerged several times by acquiring variants of staphylococcal cassette chromosome *mec* (SCC*mec*) elements that carry the *mecA* methicillin-resistance determinant (Robinson and Enright 2003; Nübel et al. 2008). However, the vast majority of MRSA isolated worldwide belong to a limited number of clones, some of which are associated with global epidemics.

Epidemic MRSA-15 (EMRSA-15) has proven to be particularly successful over the last two decades, transmitting rapidly within and between hospitals, as well as to different countries. EMRSA-15 carries a type IV SCC*mec* element, and belongs to multilocus sequence type ST22. Initially isolated in the south-east of England in 1991 (Richardson and Reith 1993), the spread of EMRSA-15 in the UK healthcare setting was rapid; by 2000, EMRSA-15 accounted for over 60% of MRSA nosocomial bacteremias in England (Johnson et al. 2001). Over this period the proportion of MRSA among *S. aureus* bacteremia in the UK increased from less than 2% to 40% (Johnson et al. 2001).

Shortly after EMRSA-15 was reported in England, MRSA indistinguishable from EMRSA-15 were reported to cause outbreaks in other geographic regions, including Germany, the Czech Republic, Portugal, New Zealand, Australia, and Singapore (Pearman et al. 2001; Witte et al. 2001; Hsu et al. 2005; Melter et al. 2006; Amorim et al. 2007; Grundmann et al. 2010). These belong to multilocus sequence type ST22 or its close relatives. In a recent structured survey, ST22 was the most frequently found MRSA across Europe (Grundmann et al. 2010), and it was among the three most frequent MRSA clones in seven out of 15 countries for which quantitative data were available (Grundmann et al. 2010). Although reported in many countries, ST22-MRSA accounted for only 0.2% of MRSA in a surveillance study recently performed in the USA (Limbago et al. 2009) and has not, as yet, been reported from South America (Sola et al. 2008) and from large parts of Asia (Arakere et al. 2005; Ko et al. 2005). ST22-MRSA has repeatedly demonstrated its ability to supplant and replace other, formerly established MRSA strains (Hsu et al. 2005; Amorim et al. 2007; Aires-de-Sousa et al. 2008; Witte et al. 2008). Of further concern, ST22-MRSA containing the *lukS/F* genes (encoding the Panton Valentin leucocidin) has been reported to cause community-associated infections (Linde et al. 2005; Witte et al. 2007).

The mechanisms that lead to the emergence and spread of MRSA strains are of great interest. Despite intense research in this area, the spatiotemporal dynamics of the population structure and epidemic spread remain poorly understood, partially due to the limited resolution provided by contemporary genotyping tools. In this study we have investigated the origins and evolution of EMRSA15 by whole genome sequencing 193 ST22 isolates from 15 countries collected between 1990 and 2009.

Integrating evolutionary and spatial information, we have reconstructed the spatial and temporal dynamics underpinning the expansion of this clone, and ascertained the genetic changes correlating with its enhanced spreading success.

Results

Phylogenomic reconstruction of the ST22 lineage. We determined full genome sequences from 193 ST22 isolates (Supplemental Table S1), including isolates both from hospital and community settings, and also twelve isolates from the earliest description of EMRSA-15 (Richardson and Reith 1993). Once mobile genetic elements (MGEs) were excluded, a total of 8,095 single nucleotide polymorphisms (SNPs) in the 'core' genome were used to reconstruct the evolutionary history of *S. aureus* ST22 (Fig. 1). These data revealed that the core genome has accumulated variation at a rate of 1.3×10^{-6} substitutions per nucleotide site per year (95% Bayesian credible intervals, 1.2×10^{-6} to 1.4×10^{-6} ; Supplemental Table S2). Substitution rates varied slightly among different phylogenetic clades within ST22, but these differences were not significant (Supplemental Table S2). Our rate estimate was similar to those previously reported for other *Staphylococcus aureus* lineages (Lowder et al. 2009; Harris et al. 2010; Nübel et al. 2010). This estimate then provided a temporal calibration with which to date emergence and geographic spread of EMRSA-15, and evolutionary events that have contributed to its success.

The level of homoplasy in the SNP dataset was extremely low (homoplasy index, 0.009), reflecting the clonal propagation of the core genome. The majority of isolates (162 of 193, 84%), collected from multiple countries, clustered in a single clade (ST22-A; Fig. 1A), which encompasses the epidemic strain EMRSA-15. All but four of the ST22-A isolates carried a single nucleotide deletion within the *ureC* gene, resulting in a frame shift inactivation of the urease alpha subunit and the consequent non-production of urease (Supplemental Table S1), which is a phenotypic

characteristic of EMRSA-15 (Pearman et al. 2001). The four isolates lacking this deletion are the most basal isolates of the ST22-A clade (for a list of the genetic variation that distinguishes the ST22-A clade from the rest of the ST22 population see Supplemental Table S3).

The origin of 173 of the isolates was recorded. Whereas the vast majority (97%) of ST22-A isolates were collected in hospitals, this proportion was significantly smaller in non-ST22 isolates ($p < 0.001$; Supplemental Table S1). This suggests that members of ST22-A have adapted to transmission and survival in healthcare settings. In contrast, other clades within ST22 encompass methicillin-susceptible *S. aureus* isolates, MRSA with different types and subtypes of SCCmec elements, and isolates with genes for the *Panton-Valentine* leucocidin toxin. The phylogeny indicates these isolates have distinct evolutionary origins from the healthcare-associated clone ST22-A isolates (Supplemental Table S1) and represent the ST22 founding population from which ST22-A/EMRSA-15 recently emerged.

Included in the collection sequenced are 12 isolates from the original description of EMRSA-15 (Richardson and Reith 1993). At the time Richardson and Reith noted that the majority of EMRSA-15 in the UK were ciprofloxacin sensitive, however they identified a group of variant isolates from the English West Midlands (Birmingham region) that were ciprofloxacin resistant. All eight ciprofloxacin sensitive EMRSA-15 isolates from the Richardson and Reith study are found in the ST22-A1 clade (Fig. 1B; Supplemental Table S1), whereas the four ciprofloxacin resistant isolates form a basal group in clade ST22-A2. It therefore appears that ST22-A1, while spreading

among hospitals in the British Isles, formed a progenitor population from which a fluoroquinolone resistant pandemic clone (i.e. ST22-A2) emerged.

EMRSA-15 genotype. A combination of gene acquisition and core genome mutation shapes the genome of ST22. Therefore, in addition to SNP discovery based on mapping sequencing reads to a reference genome, we also determined the full genetic repertoire from each of the isolates, including MGEs, by *de novo* assemblies of sequencing reads. Using the phylogenetic reconstruction, we have been able to examine the evolution of ST22-A, and also pinpoint the genetic events that accompanied the emergence of EMRSA-15.

A significant component of the *S. aureus* genome is derived from MGEs that contribute to the accessory genome, and includes staphylococcal cassette chromosome *mec* elements (SCC*mec*), prophages, *S. aureus* Pathogenicity Islands (SaPI), IS elements, transposons and plasmids (Lindsay and Holden 2004). Comparative analysis reveals the composition and fluidity of the ST22 accessory genome (Fig. 2), and that this accounts for between 4%-14% of CDS within the ST22 genomes. Comparison of the ST22-A and non-ST22-A strains reveals that, typically, the accessory genome of ST22-A strains is larger and less variable than that of non-ST22-A strains (Fig. 2A).

Analysis of the origins and location of CDSs in the ST22 accessory genomes show MGEs belonging to a range of classes (Fig. 2B), with prophages being the largest component. The accessory genome variation within the ST22-A isolates results from both the loss of ancestrally acquired elements such as prophages and SCC*mec* (Fig.

2), and the gain on multiple occasions of plasmids and prophages (Supplemental Table S1).

The ST22 reference genome, HO 5096 0412, is a representative of EMRSA-15 and is found in the ST22-A2 clade (Supplemental Table S1). Amongst the MGEs it contains are: an *SCCmec* IVh element (Milheirico et al. 2007), which carries the *mecA* gene encoding methicillin resistance; a Tn554 like-transposon, encoding a β -lactamase; two prophages (ϕ Sa2_{HO 5096 0412}, which encodes a *BcgI*-like restriction enzyme, and ϕ Sa3_{HO 5096 0412}, which encodes staphylokinase (Sako and Tsuchida 1983) and the immune evasion proteins CHIPS and SCIN (van Wamel et al. 2006)); and a 2.4 kb *ermC* carrying plasmid (Köser et al. 2012). Comparative genomic analysis of the ST22 isolates showed that these elements were conserved in the ST22-A population, with few exceptions (Supplemental Table S1).

Using the phylogeny we showed that the *SCCmec* IVh element was acquired once, just prior to the emergence of ST22-A. Similarly, ϕ Sa2 appears to have been acquired prior to the emergence of ST22-A (Supplemental Table S3). Prophage ϕ Sa3 seems to have been acquired before the emergence of the ST22-A lineage because it is present in members of the non-ST22-A population. A SaPI element encoding enterotoxin C (Novick 2003) is absent from the reference genome but present in ~68% of the ST22-A isolates (Supplemental Table S1). This element is also found in non-ST22-A isolates at the same site (3' region of GMP synthase, SAEMRSA1503430), suggesting that like ϕ Sa3, its association with the ST22 lineage pre-dates the emergence of ST22-A. In addition, other MGEs are variably present in the pandemic clone, several of which are associated with the transfer of antibiotic resistance genes,

but, unlike the elements above, none of these have been stably maintained since ST22-A2 emerged.

In addition to the MGE-induced changes in the genomic architecture of ST22-A, all ST22-A genomes possess five indels in the core genome, including a 2268 bp deletion in the fibronectin-binding protein (FnBP) locus (Supplemental Table S3). This FnBP deletion was caused by homologous recombination between the C-terminal region of *fnbA* and the N-terminal region of *fnbB*, resulting in a gene fusion, *fnbA-fnbB*. FnBPA and FnBPB bind the host extracellular-matrix proteins fibronectin, fibrinogen and elastin (Greene et al. 1995; Roche et al. 2004), allowing the *S. aureus* to connect to cellular integrins and trigger internalization (Fowler et al. 2000). In the ST22 isolates containing *fnbA* and *fnbB* (non-ST22-A), these two homologues are approximately 71% identical (~67% identical at the amino acid level) and 2871 bp and 3045 bp in length. The predicted recombination site occurs ~1850 bp into *fnbA*, thereby generating a gene fusion that retains most of the sequence of *fnbA*. The fusion protein is the same length as FnBPA and is 98.5% identical, whereas it is only 68.6% identical to FnBPB. It is therefore likely that the encoded fusion protein retains most of binding properties of FnBPA (Roche et al. 2004).

From the phylogeny we defined the genetic changes that differentiate the ST22-A2 and ST22-A1 populations (Supplemental Table S4). That differentiation was not marked by any unique acquisitions of MGEs. Instead, the variation in the core genome that marks this split includes a single insertion and 30 SNPs, of which 18 result in non-synonymous substitutions, and one nonsense mutation. The nonsense mutation introduces a premature stop codon in the ECM-binding protein homologue,

Ebh. This very large surface protein binds fibronectin and contains a C-terminal transmembrane domain that anchors the intact protein in the cell envelope (Clarke et al. 2002). The truncated variant caused by the nonsense mutation lacks this transmembrane domain, and ST22-A2 isolates likely do not express this adhesin on their surface.

Demography and phylogeography. The high proportion of ST22-A isolates in the sample and their very low sequence diversity, points to ST22-A being a highly fit, recently emerged MRSA strain. This inference is further supported by coalescence analyses estimating that the population size of ST22-A has grown exponentially by 24% each year (Bayesian credibility intervals, 20% to 29%) since the 1990s, which is four times faster than the average growth of other ST22 strains (non-ST22-A) (Fig. 3). The rapid population expansion is accompanied by star-shaped branching pattern of the ST22-A clade (Fig. 1A).

A Bayesian phylogeographic reconstruction of ancestral nodes in the phylogenetic tree supports a root of the ST22-A clade in the UK. Within a few years of its emergence there multiple exports spread it from the UK to Australia as well as to various countries in Europe, Asia, and Africa (Fig. 1B; Supplemental Fig. S1). Early ST22-A isolates (clade ST22-A1 in Fig. 1B) were susceptible to fluoroquinolones and apparently did not spread beyond the British Isles. Bayesian modelling of the spread of fluoroquinolone-resistant ST22-A2 in the UK suggests that ST22-A2 originated within the English Midlands (Fig. 4), with it being restricted to this geographical region until the late 1980s. The same analysis further suggests that isolates from the Midlands seeded infection to the north, reaching Scotland by around 1993, but also to

the south, reaching London around 1998 (Fig. 4; movie S1). By 2000, the epidemic had spread to most locations within the UK via multiple routes and sources (e.g. Bristol, from the Southern England and Midlands; Manchester, from the Midlands and the North; Aberdeen from elsewhere in Scotland and North England). The UK epidemic then spread globally through multiple international transmission events (Fig. 1).

Estimates for the timings of first appearances of ST22-A in countries beyond the UK (Supplemental Fig. S1) are concordant with available epidemiological evidence for EMRSA-15, suggesting epidemic expansion in those countries soon after its introduction (Richardson and Reith 1993; Pearman et al. 2001; Witte et al. 2001; Hsu et al. 2005; Melter et al. 2006). Our data reveal multiple introductions from the UK into Germany, Denmark and Australia, possibly via the migration of colonised health-care workers (Pearman et al. 2001). Country-specific clades in the phylogenetic tree point to subsequent diversification within many recipient countries, and the data support additional spreading events, e.g. from Portugal to Germany (Fig. 1B). Because basal isolates outside the ST22-A clade are also geographically widespread, ST22 was almost certainly present in many countries prior to the emergence of ST22-A (Fig. 1A, Supplemental Table S1). Indeed, methicillin-susceptible ST22 is known to form a common component of the nasal carriage population in many places (Feil et al. 2003; Holtfreter et al. 2007; Ruimy et al. 2009).

The development of drug resistance in ST22. The broad range of resistance traits within ST22-A contrasts strikingly with the more susceptible and diverse population from which it emerged (Supplemental Table S1). For ST22-A, the number of

resistance traits per isolate increased with time ($p < 0.001$; Fig. 5), whereas the trend in the rest of the population (non-ST22-A) was not statistically significant ($p > 0.1$).

In order to investigate the evolutionary events responsible for the increase in resistance, we analyzed the whole genome sequences to elucidate the antibiotic resistance genotype of ST22. Molecular determinants within the genomes explained 99.8% of the measured phenotypic resistance traits (847 resistance traits in total). Only two phenotypes did not correlate with the predicted genotypes. HO 7258 0475 05 was erythromycin resistant but lacked *ermC* and all other erythromycin resistance mechanisms that have been documented in the literature. Strain 09-00678 contained *thyA* and *dfrA* but was not phenotypically co-trimoxazole resistant. For both isolates it is possible that the drug resistance determinants were lost during lab culturing prior to DNA preparation, or phenotypic testing.

Our genomic analyses showed that evolution of resistance in ST22 has occurred in a stepwise fashion encompassing both the core and accessory genomes. Acquisition of a transposon-encoded β -lactamase BlaZ confers penicillin-resistance, and the widespread distribution of this transposon in ST22 is consistent with a single, primordial development of β -lactam resistance, and probably reflects usage of penicillin since the 1940s. Only six ST22 isolates in our collection are sensitive to penicillin, four of which have lost the Tn552-like transposon carrying the *blaZ* gene, and one of which carries a *blaZ* gene inactivated through an IS element insertion in its promoter region (Supplemental Table S1). The penicillin-resistant ST22 population then gave rise to several methicillin-resistant clones, which emerged through imports of SCCmec elements into their chromosomes. The phylogenomic distribution of

different *SCCmec* variants in the phylogenomic context provides evidence of at least seven independent imports into ST22, one replacement of *SCCmec* type IV by *SCCmec* type V, and two secondary losses of *SCCmec* through precise excision by homologous recombination between flanking direct repeat sequences (Fig. 1; Supplemental Table S1). As described above, ST22-A emerged after the acquisition of *SCCmec* IVh, which is predicted to have occurred over a decade before the first report of EMRSA-15 (Fig. 1) (Richardson and Reith 1993).

Around 1986 (95% Bayesian credible intervals 1984-1988) ST22-A developed fluoroquinolone resistance through two point mutations generating amino acid substitutions Ser80Phe in topoisomerase IV (GrlA) and Ser84Leu in gyrase A (GyrA). This genotype is found in the ST22-A2 clade (Fig. 1B) that spread throughout the UK, and globally. We note that our dating estimate for acquisition of fluoroquinolone resistance coincides closely with the introduction of fluoroquinolone drugs into clinical medicine in the UK, ciprofloxacin being first licensed in February 1987. Further, phylogeographic modelling suggested that fluoroquinolone resistance emerged in the English Midlands in the early 1990s (Fig. 4). Notably, clinical studies with ciprofloxacin occurred in the Midlands prior to the licensing of ciprofloxacin in 1987 (Crump et al. 1983; Finch et al. 1986; Silverman et al. 1986).

In the subsequent decades the pandemic has continued along a trajectory of increasing antibiotic resistance, developing resistance to a range of antibiotics of different classes (tetracyclines, aminoglycosides, co-trimoxazole, fusidic acid, mupirocin, and oxazolidinones) on multiple occasions (Supplemental Fig. S2, Supplemental Table S1). The development of these resistances has been rather sporadic, and also less

stably maintained than those against β -lactams and fluoroquinolones, but nevertheless demonstrates clear evidence of the strong selective pressure exerted by antibiotic usage and clinical practice. For example, our dataset documents multiple independent events of erythromycin resistance acquisition based on the uptake of plasmids encoding the ErmC erythromycin ribosomal methylase (Fig. 1B). Phylogenetic analysis also indicates that the *ermC* plasmid was lost multiple times, and in several cases these losses affect entire clades, suggesting they represent genuine events in the natural population rather than laboratory artefacts (Fig. 1B). ErmC additionally confers resistance to clindamycin, when mutations in the leader peptide region of the *ermC* gene cause its constitutive expression (Levin et al. 2005). All clindamycin-resistant isolates possess genetic rearrangements that disrupt the promoter region upstream of *ermC*, including deletions of part, or all of the *ermC* leader peptide region, or insertion of an IS element in *ermC* leader peptide region. At least eight independent, structurally different, mutations have generated clindamycin resistance in the population examined (Supplemental Table S1), and five of these were found in ST22-A (denoted 'c-*ermC*' in Fig. 1B). The data also suggest that country-specific variation in antibiotic usage may be shaping the geographic distribution of resistance. The prescription level of clindamycin is much higher in Germany than in the UK (Coenen et al. 2006; Meyer et al. 2006; de With et al. 2011), causing selective pressure for a high prevalence of clindamycin resistance. Among the 34 erythromycin-resistant isolates from Germany, 32 (94%) are also clindamycin-resistant, whereas none of the 45 UK isolates are resistant to clindamycin (significant difference, $p < 0.001$, Fisher's Exact Test).

Discussion

EMRSA-15 represents a 'measurably evolving population' (Drummond et al. 2003), accumulating genetic variation over epidemiological timescales at an average rate similar to previous estimates for other *S. aureus* strains (Lowder et al. 2009; Harris et al. 2010; Nübel et al. 2010). One consequence of this fast nucleotide substitution rate is that evolutionary events could be dated, including the stepwise acquisition, refinement, and loss of multiple drug resistance traits. Furthermore, past temporal changes in population size and geographic expansion could be analysed retrospectively based on the molecular traces that these processes had left in the genomes. These conclusions were possible because our genome sequences span a wide range of time and space, which is a prerequisite for such coalescence-based analyses. Our data revealed much more detail about the pathogen's population history than would have been possible with a snapshot survey from a single time point.

In principle, changes in the frequency of a genetic variant may be due to stochastic events or genetic adaptation and selection (Keim and Wagner 2009). For the highly successful MRSA strains described here, it seems plausible that they are more specialized to the restricted ecological niches in the hospital environment, where exposure to antibiotics is common, and survival depends on efficient transmission between patients in close contact, than is the ancestral population from which they emerged. This scenario is supported by the observation that hospital-associated MRSA rarely establish themselves in the community. Aside from the development of antibiotic resistance, little is known about other adaptive genetic changes that promote the success of *S. aureus* (Gomes et al. 2005; Chambers and Deleo 2009). A recent microarray-based investigation did not reveal any concordant differences between

gene contents of highly successful ('epidemic') versus sporadic *S. aureus* strains (Kuhn et al. 2010). Recently, a phage-encoded virulence determinant (which is not present in ST22) has been shown to be associated with enhanced spreading of MRSA in Asia (Li et al. 2012). Phenotypic changes may be caused by genetic changes that are subtler than the gain of entire genes through horizontal gene transfer, or their loss (Kennedy et al. 2008; Olsen et al. 2010). However, our analysis of the genetic changes that accompanied the emergence of EMRSA-15 failed to pinpoint a single determinant, but by contrast, identified a series of gene acquisition events and mutations that potentially promote the success of this clone in the hospital setting (Supplemental Tables S3 & S4). Prominent amongst these are those associated with antibiotic resistance.

Evidence of the selective pressure exerted by the widespread use of broad-spectrum antibiotics in hospitals and beyond is manifest in the ST22 isolates' genomes. In order to unravel the evolutionary events surrounding the development of resistance, we examined the genotypic basis of the antibiotic resistance phenotypes. These analyses permit an evaluation of utility of genome sequencing as a diagnostic tool for the prediction of antibiotic resistance. Our *in silico* analysis of genome sequences identified the molecular correlates for >99% of the phenotypic antimicrobial resistance traits. This strong correlation documents the impressive success of past research into resistance against established antibiotics, leaving few undiscovered resistance mechanisms in MRSA. Of note, several mutations in the *ermC* promoter region were structurally related but not identical to previously reported mutations (Levin et al. 2005), and oxazilidinone resistance in our isolate collection was due to a mechanism that had been recently described from *in vitro* experiments (Locke et al.

2009) but had not previously been reported from clinical isolates (Jones et al. 2009). These observations highlight the potential of large-scale sequence analyses for discovering novel mechanisms. While various molecular methods were used in the past to detect resistance determinants in clinical isolates (Bergeron and Ouellette 1998), whole-genome sequences enabled a screen for all known genes and mutations in a single experiment. For the 12 antibiotics, we screened for 116 different alleles that were known to confer antibiotic resistance by searching antibiotic resistance genes in the *de novo* assemblies or SNPs in the core genome (Supplemental Table S5). Accordingly, we were able to demonstrate that it should be possible to predict with high accuracy the full resistance profile from the sequence data alone. Such predictions may be more challenging for other organisms whose mechanisms of resistances are poorly studied, or more complex, such as those associated with efflux and permeability.

The reduction in costs and time to generate genomic sequences has recently led to genome sequencing being developed as a clinical diagnostic tool which can be used to investigate and stop outbreaks in hospitals (Eyre et al. 2012; Harris et al. 2012; Köser et al. 2012; Snitkin et al. 2012). The observations from this study herald an important development in the progress of transitioning whole genome sequencing from a research tool to a diagnostic application. Whole-genome sequencing can provide both high discriminatory power for molecular epidemiology and a prediction of antibiotic resistance, which is valuable for “near real-time” infection control and clinical management, and also for surveillance, horizon scanning of the clinical population for the evolution of drug resistance and the identification of emerging threats.

With the benefit of hindsight it is clear that the emergence of the ST22-A2 variant of the EMRSA-15 population in the UK in the 1980s marked one such threat. Originating in the English Midlands, it spread rapidly during the 1990s, becoming endemic in UK hospitals, and was associated with an upsurge in the number of MRSA infections (Johnson et al. 2001).

Of the genetic events that define the emergence of ST22-A2, the two non-synonymous SNPs associated with fluoroquinolone resistance would appear to be good candidates to explain the success of the clone. Fluoroquinolones are readily excreted in sweat and can suppress the growth of the normal microbiota, which includes *S. aureus* that colonizes skin, nose and throat (Hawkey 1997). Fluoroquinolone resistant ST22-A2 therefore may have a competitive edge, promoting colonization and survival in the hospital environment where this class of antibiotic is frequently used. Indeed, widespread fluoroquinolone usage has previously been reported to promote the spread of fluoroquinolone-resistant MRSA, even though this class of antibiotics is not normally used to treat *S. aureus* infections (Weber et al. 2003; Salangsang et al. 2010). A recent study examining MRSA rates in a tertiary hospital over a 10-year period, identified a significant increase in MRSA infection rate after an increase in fluoroquinolone prescription levels following a period of restricted usage (Parienti et al. 2011). Conversely, significant falls in MRSA rates have been linked to a decrease in the usage of fluoroquinolones over a 4-year period (Lafaurie et al. 2012). In a prevalence study looking at elderly patients admitted to hospitals in the UK, a strong association was found between MRSA colonization and prior ciprofloxacin exposure (Hori et al. 2002), with EMRSA-15 being the dominant lineage detected in this study (75.9% of MRSA isolates).

The biological and ecological advantage of fluoroquinolone resistance would therefore appear to be clear. However, to understand the remarkable success of ST22-A2 requires insights into the historic clinical context in which it emerged. Our analysis suggests that development of fluoroquinolone resistance coincided closely with the introduction of fluoroquinolone drugs into routine clinical medicine in 1987. Initially, ST22-A2 isolates were restricted to the English Midlands until 1990, after which point the clone spread rapidly. During this time the use of fluoroquinolones in the hospital and community settings increased rapidly, doubling from 1990 to 1993 (Livermore et al. 2002). Throughout the rest of the 1990s, usage in the community stabilized, but the usage in hospitals continued to grow steadily both in absolute terms and also as a proportion of all antibiotic usage, constituting 31.5% of total usage in 1999 compared with 18.9% in 1992 (Livermore et al. 2002). The emergence and rapid spread of ST22-A2 in the UK 1990s therefore perhaps reflects changes in clinical prescription regimens. In this regard it is worth noting that we also uncovered evidence of the recent influence that country-specific antibiotic regimens (clindamycin usage in Germany vs. the UK) have had in shaping the MRSA population within Europe, illustrating the link between pathogen population structure and national antibiotic prescription policymaking.

The pandemic success of EMRSA15 cannot however be simply be characterized as a case of the right mutation at the right time. Fluoroquinolone resistance is not unique to ST22-A2, but is also found among non-ST22-A isolates in the collection as the result of homoplastic substitutions (Supplemental Table S1), as well as in other epidemic MRSA clones (Fluit et al. 2001; Woodford and Livermore 2009; Tenover et

al. 2011). Hence, other facets of the ST22-A2 must have contributed to its outstanding success. Fluoroquinolone resistant ST22-A2 emerged out of a hospital-adapted genetic background (ST22-A1; Fig. 1) that was already successful, and had spread through English hospitals by the start of the 1990s (Richardson and Reith 1993).

An important class of *S. aureus* proteins involved in host colonization are surface-expressed binding-proteins (Edwards et al. 2012). Analysis of the genetic events that define the emergence of ST22-A identified mutations in two surface-expressed binding-proteins, FnBP and Ehb, both of which bind fibronectin, a glycoprotein component of the host's extracellular matrix. Ehb is one of the largest bacterial proteins (~1.1MDa) (Clarke et al. 2002). The pandemic clone of EMRSA-15 has a nonsense mutation that ablates expression of Ehb at the cell envelope. Previous studies demonstrated that mutation of *ebh* had no effect on the growth rate, or pathogenicity in a murine skin abscess model of infection (Clarke et al. 2002). Notably this protein is intact in some successful MRSA lineages (Gill et al. 2005; Diep et al. 2006; Holden et al. 2010) and mutated in others (Kuroda et al. 2001; Holden et al. 2004). Therefore it is unlikely that mutation of Ehb is the sole determinant of increased transmissibility of ST22-A. Its effect may be more pronounced due to other mutations, for example the deletion in the fibronectin-binding protein locus.

Recombination between the C-terminal regions of *fnbA* and *fnbB* results in an *fnbA-fnbB* gene fusion and an effective reduction in *fnb* gene copy number at this locus due to the deletion of *fnbB*. Studies with clinical isolates suggested that strains associated with invasive disease are significantly more likely to have two *fnb* genes (Peacock et

al. 2000). Molecular studies have also shown that both FnBPA and FnBPB are required for the establishment of septic infection (Shinji et al. 2011). However in *S. aureus* that have both homologues, deletion of either gene does not abolish fibronectin-binding capacity of the cell (Greene et al. 1995). The functional consequences of the deletion and resulting gene fusion on the host-cell interaction and transmissibility are unclear, especially when considering that there is also a mutation in *ebh*.

Interestingly, although most other representatives of the main MRSA lineages contain two copies of the *fnb* genes (Kuroda et al. 2001; Gill et al. 2005; Diep et al. 2006; Holden et al. 2010), EMRSA-16, an unrelated epidemic MRSA lineage that originated in the UK around the same time as EMRSA-15, also contains a deletion-mediated fibronectin-binding protein gene fusion (Holden et al. 2004). Like EMRSA-15, EMRSA-16 also has evolved fluoroquinolone resistance via *gyrA* and *grlA* mutations just prior to its emergence (McAdam et al. 2012). Both of these epidemic clones spread throughout UK hospitals since the 1990s (Johnson et al. 2001), albeit EMRSA-15 has been the more successful of the two clones (Ellington et al. 2010). It is therefore intriguing that they share the fluoroquinolone resistance and *fnbA*-*fnbB* genotypes that affect resistance and host cell interaction. However pinpointing the genetic basis of their success is more complex. In the EMRSA-16 background, mutations in the accessory gene regulator C and α -hemolysin (DeLeo et al. 2011), and squalene desaturase (McAdam et al. 2012) have been identified and hypothesized to contribute to its epidemic success. However these are not found in EMRSA-15. In order to fully elucidate the genetic basis for the relative success of EMRSA-15,

extensive experiments investigating the colonization and transmission properties of the hospital adapted clone are required.

Methods

Bacterial isolates

Sources and properties of 193 *S. aureus* isolates are listed in Supplemental Table S1. In assembling the collection in this study, we included isolates that encompass wide geographical and temporal ranges, but also included isolates that reflect the documented history and prevalence of EMRSA-15 and ST22.

Susceptibility to antibiotics was tested by using the standardized broth microdilution method according to DIN58940 instructions (Anonymous 2004), and phenotypic resistance was defined by applying minimum inhibitory concentration breakpoints from the European Committee on Antimicrobial Susceptibility Testing (EUCAST; <http://www.eucast.org/>). In total 18 antibiotics were tested (penicillin, PEN; oxacillin, OXA; gentamycin, GEN; linezolid, LNZ; erythromycin, ERY; clindamycin, CLI; ciprofloxacin, CIP; fusidic acid, FUS; mupirocin, MUP; moxifloxacin, MFL; co-trimoxazole, SXT; tetracycline, TET; vancomycin, VAN; teicoplanin, TPL; rifampicin, RAM; phosphomycin, PHO; tigecycline, TGC; daptomycin, DAP), of which, 12 had resistance in at least 1 isolates in the collection (Supplemental Table S1). Bacterial typing was performed as described previously (Strommenger et al. 2008).

Genomic library creation and multiplex sequencing

Unique index-tagged libraries for each sample were created, and up to 12 separate libraries were sequenced in each of eight channels in Illumina Genome Analyser GAII cells with 54-base paired-end reads of Illumina HiSeq with 75 base paired-end reads. The index-tag sequence information was used for downstream processing to assign reads to the individual samples.

Detection of SNPs in the core genome

The paired-end reads were mapped against the chromosome of *S. aureus* HO 5096 0412 (accession number HE681097) (Köser et al. 2012), a representative of EMRSA-15 that was isolated from a fatal neonatal infection in Suffolk, UK in February 2005. SNPs were identified as described in (Croucher et al. 2011). Indels were identified using Dindel (Albers et al. 2011). Unmapped reads and sequences that were not present in all genomes were not considered as part of the core genome, and therefore SNPs from these regions were not included in the analysis. SNPs falling within MGEs regions were also excluded from the core genome, as well those falling in high-density SNPs regions, which could have arisen by recombination. The core genome was curated manually to ensure a high quality data set for subsequent phylogenetic analysis and comprised of 2,643,131 bp.

Phylogenetic analysis

A phylogeny was drawn for ST22 using RAxML v0.7.4 (Stamatakis et al. 2005) to estimate the trees for all SNPs called from the core genome. The general time-reversible model with gamma correction was used for among-site rate variation for ten initial trees.

Comparative analysis

Raw Illumina data were split to generate paired-end reads, and assembled using a *de novo* genome-assembly program, Velvet v0.7.03 (Zerbino and Birney 2008), to generate a multi-contig draft genome for each of 193 *S. aureus* isolates. Gene prediction was performed using Prodigal (Hyatt et al. 2010). Predicted CDSs were compared using BlastP (Ausubel et al. 1995), and the proteins were clustered into homology groups using TribeMCL (Centre for Mathematics and Computer Science and EMBL-EBI) (Enright et al. 2002) with a cut-off of $1e^{-50}$.

Included in the clustering dataset were the complete genome sequences of *S. aureus* available in the public sequence databases. MGEs in these sequences were identified using comparative genomic analysis and manual curation. Homology groups (HG) were categorised according to their origins based on matches to CDSs from the MGE reference set that clustered within the group. For the purposes of the preliminary accessory genome analysis, HG origins were subdivided into: Prophage, Conjugative transposon, Plasmid, *S. aureus* Pathogenicity Island (SaPI), SCC*mec* element, Transposon, Hypothetical proteins and Other. Conjugative transposon/Plasmid reference sequences included Tn916-like (Clewett et al. 1985) and Tn5801-like elements (Kuroda et al. 2001), as well as integrated conjugative elements such as ICE6013 (Smyth and Robinson 2009), whereas Transposon included Tn554 (Phillips and Novick 1979) and Tn552 (Rowland and Dyke 1989), and IS elements. Hypothetical proteins HGs did not have any matches with either core, or MGE, CDSs from the reference set. HGs defined as ‘Other’ indicates that there were no matches with MGE CDSs, but there were matches with CDSs from the core regions of the reference genomes. In some cases, there were examples of clustered proteins from

more than one origin being present in the single HG, for example Prophage and SaPI, and Conjugative transposons and Plasmid. This reflects the shared evolutionary origins and mechanisms of transfer MGEs. In these cases, HGs were assigned arbitrarily.

Detailed comparisons of genome sequences was conducted on the *de novo* assemblies using BlastN (Ausubel et al. 1995), and was facilitated by using the Artemis Comparison Tool (ACT) (Carver et al. 2005).

Determination of antibiotic resistance genotype.

The literature was extensively mined for the genetic basis of antibiotic resistance in *S. aureus*. A database was compiled which contained horizontally acquired genes and core genes mutations that conferred antibiotic resistance. Antibiotic resistance conferred by SNPs in components of the core chromosome were identified by manual curation; substitutions were compared to antibiotic resistance mutations in *S. aureus* proteins detailed in the literature. The presence of antibiotic resistance genes of MGEs was investigated using BlastN (Altschul et al. 1990) to compare *de novo* assemblies against a database of *S. aureus* resistance genes compiled from literature and public sequence database searches, and also by mapping sequence reads to a pseudo-molecule consisting of concatenated antibiotic resistance gene as detailed by Köser et al. (Köser et al. 2012).

The genotype of each isolate was searched for the resistance genes and/or mutations for the antibiotics tested; for the 12 antibiotics for which resistance was detected, 847 incidences in total, 116 different genes or alleles were searched for. The presence of

the following known resistance genes and mutations/deletions of determinants conferring resistance were identified: *mecA* (Utsui and Yokota 1985), *ermC* (Catchpole et al. 1988), *blaZ* (Voladri and Kemodle 1998), *tetK* (Khan and Novick 1983), *fusC* (O'Neill et al. 2007), *fusA* (mutation) (Chen et al. 2010), *dfrA* (Rouch et al. 1989), *thyA* (deletion) (Leelaporn et al. 1994), *ileS-2* (Hodgson et al. 1994), *aacA-aphD* (Rouch et al. 1987), *griA* (mutation) (Griggs et al. 2003), *gyrA* (mutation) (Griggs et al. 2003), *rplC* (mutation) (Locke et al. 2009), and *ermC* plasmid (deletion) (Weisblum 1995) (see Supplemental Table S1).

Assessing the association of number of genetic determinants of resistance with time

The total number of resistance genes and mutations/deletions conferring resistance were summed for each strain. Linear regression was used to assess the association of the number of genetic determinants of resistance per strain with time. The dependent variable was the number of genetic determinants of resistance per strain, and the independent variable was time (year of isolation). Significance was assessed at $p < 0.05$. This analysis was applied to all 162 ST22-A isolates, from all countries, and separately to the 77 ST22-A isolates from the UK. This latter analysis was performed to evaluate the association of time with the number of genetic resistance determinants in a single country. Linear regression with the non-ST22-A strains demonstrated that the number of genetic determinants of resistance per strain was not significantly associated with time ($p > 0.10$).

Assessing the association of ST22-A strains with hospital origin

Out of the 162 global ST22-A strains, there were 147 where it was known whether or not they were hospital-associated: 142 (96%) were hospital-associated, and five were not. Of the 31 non-ST22-A strains, there were 26 where it was known whether or not they were hospital-associated: 13 (50%) were hospital-associated, and 13 were not. These proportions were significantly different according to a Fisher's Exact Test ($p < 0.05$).

Bayesian analyses

We used the Bayesian software package, BEAST (v1.7.1) (Drummond et al. 2012) to investigate the temporal, spatial and demographic evolution of ST22. The core genome was retained for the analyses after discarding all ambiguous bases and strains that showed admixture, resulting in a final alignment of 193 taxa and 2,154,246 bps. The different clades of ST22 (A and non-A) have different demographic histories, as inferred from the shape of their coalescent trees. Non-ST22-A individuals coalesce further back in time and at a slower rate than ST22-A individuals, and the latter have a more star-like genealogy. For this reason, we analysed these two clades separately (25 taxa for non-ST22-A and 162 taxa for ST22-A). We used an uncorrelated lognormal distribution (variable clock rate) to model the rate of evolution and constrained the tips of the phylogeny to their dates of isolation to calibrate that rate, as well as giving it a default uniform prior of between 0 and 1. We used the HKY85 model of nucleotide substitution (Hasegawa et al. 1985) with four discrete gamma-distributed rate categories and to model the relative node ages, ran the two datasets under both the constant size and exponential size coalescent tree prior. All other default parameters in the programme BEAUti v1.6.2 were used. To model the global

geographical spread, we calculated the probabilities of transfer between each discrete location (i.e. country) via a continuous-time Markov chain with a non-reversible infinitesimal rate matrix, using this to inform location at ancestral nodes (Lemey et al. 2009; Edwards et al. 2011). To reduce the number of parameters to estimate, we employed Bayesian Stochastic Search Variable Selection (BSSVS), where at any point in the Monte Carlo Markov Chain (MCMC), the transition probabilities of some states can be set to zero, with the prior designed to minimise the number of location changes as described in Lemey *et al.* (2009). To infer the geographic spread within the UK and to predict the root state of the fluoroquinolone-resistant isolates, we modelled diffusion over a continuous two-dimensional space assuming a Brownian motion (or Weiner process), with the rate of diffusion assumed to be constant over the entire tree and estimated from the data (Lemey et al. 2010). For this analysis, we used ST22-A strains and discarded strains that were fluoroquinolone-sensitive and those that did not have exact latitude and longitude information, leaving a total of 60 taxa. The uncertainty in root location is represented as 80% of the Highest Posterior Density of the continuous traits: latitude and longitude. We calculated a posterior distribution of all parameters using two separate MCMCs and terminated the runs after checking for efficient mixing, convergence and that 90% of the distribution was post ‘burn-in’. To estimate the timing of first introduction to each country (Supplemental Fig. S1), we took the date of the youngest node for any individual location in the posterior distribution of trees.

FIGURES.

Fig. 1. Phylogeny of ST22 and the emergence of MRSA clones (A) Maximum likelihood phylogenetic tree of ST22 isolates. The tree was rooted by using the distantly related *S. aureus* isolate MSSA476 as an outgroup. Colors indicate the isolates' countries of origin. Roman numerals indicate acquisitions of structurally different *SCCmec* elements, which cause methicillin resistance. (B) Maximum clade credibility tree of the ST22-A clade based on BEAST analysis using a variable clock rate (uncorrelated lognormal) model. Tips of the tree are constrained by isolation dates, the time scale is shown at the bottom. Gains and losses (Δ) of genetic determinants for resistance to methicillin (*SCCmec* IVh), fluoroquinolones (point mutations in *griA* and *gyrA*), erythromycin (plasmid-encoded *ermC*), and clindamycin (mutations in *ermC* leader peptide region, *c-ermC*) have been mapped on the tree by applying the parsimony criterion.

Fig. 2. Accessory genome diversity in ST22. CDSs in the accessory genome of each isolate were clustered and assigned to homology groups (HGs). (A.) Conservation of accessory genome HGs across the ST22 population. The isolates are ordered according to the phylogeny displayed along the left and top of the figure. The conserved HGs in the pairwise comparison are displayed as a heat map matrix, and colored according to the number of conserved HGs (range 64 to 235 HGs) per pair; see figure for scale. (B.) Size and composition of each isolate's accessory genome based on the number of HG CDSs. The HGs have been subdivided into different mobile genetic element types based on matches to an annotated MGE reference set included in the clustering (see figure for legend).

Fig. 3. Exponential growth rates (λ) of ST22-A and other ST22 isolates (posterior probability density functions from Bayesian analysis). ST22-A is estimated to have grown significantly faster than non-ST22-A. The population size N at time t was modelled by $N(t) = N(0) \cdot \exp^{\lambda \cdot t}$, where λ is the exponential growth rate and t is measured in years. Accordingly, the population size change per year is $\exp^{\lambda \cdot \text{year}}$.

Fig. 4. Bayesian reconstruction of the spread of ST22-A2 in the UK. A continuous spatial diffusion model was used to reconstruct the finer-scale geographical dispersal of ST22-A2 within the UK and to predict the origin of fluoroquinolone resistance. Lines indicate the inferred routes of spread with confidence displayed as green ovals representing 80% of the Highest Posterior Density (HPD) for latitude and longitude. The timing of transmission events are represented by red (old) to black (recent) lines and light- to dark-green oval shading (for the animation of the reconstruction of the spread, see File S2). The maps are based on satellite pictures made available in Google Earth (<http://earth.google.com>).

Fig. 5. Increase in antibiotic resistance traits within ST22-A population over time. Number of genotypic resistance determinants per ST22-A isolate (red, n. 162) and non-ST22-A isolates (blue, n. 31) from all countries, 1990-2008, including regression lines with 95% confidence interval for each group. Size of circle corresponds to number of isolates.

Data access

Data has been deposited in the European Nucleotide Archive under study numbers ERP000150 and ERP000633.

Acknowledgments:

We thank the core sequencing and informatics teams at the Sanger Institute for their assistance and The Wellcome Trust for its support of the Sanger Institute Pathogen Genomics and Biology groups, MTGH, SH, AEM, SDB and JP were supported by Wellcome Trust grant 098051, BGS and DMA were funded by Wellcome Trust grant 089472, AM, ZZ and MA were supported by grant 05/FE1/B882 from the Science Foundation of Ireland. LAW and FB were supported by grant 260801-BIG-IDEA from the European Research Council. KK, BS, FL, UN, and WW were supported by the Robert Koch Institute, the Bundesministerium für Bildung und Forschung (Germany; Pathogenomik Plus Grant 101570326667) and the European Commission, Framework Programme 7 (project Trocar). HZ was supported by grant IGA 9642-4 from the Internal Grant Agency, Ministry of Health (CZ). JRF, GFE, and KET were supported by a grant from the Chief Scientists Office Scotland and by a Medical Research Council doctoral training grant (for PRM), and by the BBSRC (JRF). HdL was supported by grants from US Public Health Service 2 RO1 AI457838-12 to A. Tomasz, and grant # PEst-OE/EQB/LA0004/2011 from Fundação para a Ciência e a Tecnologia, Portugal.

We are grateful to M. Henkel, A. Weller, C. Cuny I, H. Illiger, and the sequencing team at the Robert Koch Institute for technical assistance, and we thank M. Aires-de-Sousa, C. Milheiriço and N. Faria for helping to select the Portuguese isolates used in this study, and S. Tong for his comments regarding the manuscript. We also thank the

Genomics core facility (Ark Genomics) at The Roslin Institute, University of Edinburgh for sequencing services.

Supplementary Materials:

Figure S1. Bayesian reconstruction of the timings of first ST22-A introduction per country (posterior probability density functions from Bayesian analysis). For an animation of the Bayesian reconstruction of the global spread of ST22-A see File S1.

Figure S2. Phylogenetic distribution of antibiotic resistance in ST22. A Maximum Likelihood tree (RAxML) based on SNPs in the core genome is displayed on the left. The matrix displays the resistance genetic determinants. For each of 193 isolates, a colored shaded box indicates a positive genotypic trait conferring resistance, the identity of which is indicated by the text at the top of the column. Phenotypic resistance against the following antibiotics was observed: penicillin, PEN; oxacillin, OXA; gentamicin, GEN; linezolid, LIN; erythromycin, ERY; clindamycin, CLI; ciprofloxacin, CIP; fusidic acid, FUS; mupirocin, MUP; moxifloxacin, MFL; co-trimoxazole, SXT; tetracycline, TET. A list of resistance genotypes from extensive mining of the literature was also used to query the data. Only genotypes conferring antibiotic resistance are indicated and these include the presence of horizontally acquired antibiotic resistance genes (*blaZ*, *mecA*, *aacA-aphD*, *ermC*, *fusC*, *ileS-2*, *thyA* and *dfrA*, and *tetK*); deletions that ablated function (*rplC* Δ S145 and *ermC* Δ lp), and non-synonymous substitutions in core genes that alter protein functions (GrlA and GyrA, and FusA).

Table S1.

ST22 isolates used in this study and accompanying metadata.

Table S2.

Substitution rates and time to most recent common ancestor (tMRCA) for the ST22-A clade and for non-ST22-A isolates.

Table S3

Genetic changes that differentiate the ancestral ST22-A and non-ST22-A1 populations.

Table S4

Genetic changes that differentiate the ancestral ST22-A2 and ST22-A1 populations

Table S5

Antibiotic resistance alleles examined for the ST22 antibiotic resistance genotype. The presence of accessory genes, and core genome mutations were searched for in all 193 isolates. The table details antibiotic resistance determinants for the 12 antibiotics described in this study.

File S1. Animation of the Bayesian reconstruction of the global spread of ST22-A. KML file of the BEAST phylogeographic analysis using the ST22-A dataset. A continuous spatial diffusion model was used. Lines indicate the inferred routes of spread. The timing of transmission events are represented by blue (old) to mauve (recent) lines. For the posterior probabilities of the introductions see Supplemental

Fig. S2. To view the animation, load in Google Earth (<http://earth.google.com>) and click the time slider animation button.

File S2. Animation of the Bayesian reconstruction of the spread of ST22-A2 in the UK. KML file of the BEAST phylogeographic analysis using the ST22-A2 dataset illustrated in Fig. 4. A continuous spatial diffusion model was used to reconstruct the finer-scale geographical dispersal of ST22-A2 within the UK and to predict the origin of fluoroquinolone resistance. Lines indicate the inferred routes of spread with confidence displayed as green ovals representing 80% of the Highest Posterior Density (HPD) for latitude and longitude. The timing of transmission events are represented by red (old) to black (recent) lines and light- to dark-green oval shading. To view the animation, load in Google Earth (<http://earth.google.com>) and click the time slider animation button.

References

- Aires-de-Sousa M, Correia B, de Lencastre H. 2008. Changing patterns in frequency of recovery of five methicillin-resistant *Staphylococcus aureus* clones in Portuguese hospitals: surveillance over a 16-year period. *J Clin Microbiol* **46**(9): 2912-2917.
- Albers CA, Lunter G, MacArthur DG, McVean G, Ouwehand WH, Durbin R. 2011. Dindel: accurate indel calls from short-read data. *Genome research* **21**(6): 961-973.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of molecular biology* **215**(3): 403-410.
- Amorim ML, Faria NA, Oliveira DC, Vasconcelos C, Cabeda JC, Mendes AC, Calado E, Castro AP, Ramos MH, Amorim JM et al. 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**(9): 2881-2888.
- Anonymous. 2004. DIN 58940. Medical microbiology -- susceptibility testing of pathogens to antimicrobial agents. Part 8 -- microdilution. In *DIN-Taschenbuch 222: medizinische Mikrobiologie und Immunologie -- diagnostische Verfahren*. Deutsches Institut für Normung, Beuth-Verlag, Berlin, Germany.
- Arakere G, Nadig S, Swedberg G, Macaden R, Amarnath SK, Raghunath D. 2005. Genotyping of methicillin-resistant *Staphylococcus aureus* strains from two hospitals in Bangalore, South India. *J Clin Microbiol* **43**(7): 3198-3202.
- Ausubel FM, Brent R, Kingston RB, Moore DD, Seidman JG, Smith JA, Struhl K. 1995. In *Current protocols in molecular biology*. John Wiley & Sons, New York.

- Bergeron MG, Ouellette M. 1998. Preventing antibiotic resistance through rapid genotypic identification of bacteria and of their antibiotic resistance genes in the clinical microbiology laboratory. *J Clin Microbiol* **36**(8): 2169-2172.
- Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. 2005. ACT: the Artemis Comparison Tool. *Bioinformatics* **21**: 3422-3423.
- Catchpole I, Thomas C, Davies A, Dyke KG. 1988. The nucleotide sequence of *Staphylococcus aureus* plasmid pT48 conferring inducible macrolide-lincosamide-streptogramin B resistance and comparison with similar plasmids expressing constitutive resistance. *Journal of general microbiology* **134**(3): 697-709.
- Chambers HF, Deleo FR. 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* **7**(9): 629-641.
- Chen HJ, Hung WC, Tseng SP, Tsai JC, Hsueh PR, Teng LJ. 2010. Fusidic acid resistance determinants in *Staphylococcus aureus* clinical isolates. *Antimicrob Agents Chemother* **54**(12): 4985-4991.
- Clarke SR, Harris LG, Richards RG, Foster SJ. 2002. Analysis of Ebh, a 1.1-megadalton cell wall-associated fibronectin-binding protein of *Staphylococcus aureus*. *Infection and Immunity* **70**(12): 6680-6687.
- Clewell DB, An FY, White BA, Gawronburke C. 1985. *Streptococcus faecalis* sex-pheromone (cAM373) also produced by *Staphylococcus aureus* and identification of a conjugative transposon (Tn918). *Journal of bacteriology* **162**(3): 1212-1220.
- Coenen S, Ferech M, Malhotra-Kumar S, Hendrickx E, Suetens C, Goossens H. 2006. European Surveillance of Antimicrobial Consumption (ESAC): outpatient macrolide, lincosamide and streptogramin (MLS) use in Europe. *J Antimicrob Chemother* **58**(2): 418-422.
- Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JH, Ko KS et al. 2011. Rapid pneumococcal evolution in response to clinical interventions. *Science* **331**(6016): 430-434.
- Crump B, Wise R, Dent J. 1983. Pharmacokinetics and tissue penetration of ciprofloxacin. *Antimicrob Agents Chemother* **24**(5): 784-786.
- de With K, Kern WV, Meyer E. 2011. Antibiotikaverbrauch im Krankenhaus. In *GERMAP 2010 Antibiotika-Resistenz und -Verbrauch*, pp. 17-21, Düsseldorf, Germany.
- DeLeo FR, Kennedy AD, Chen L, Bubeck Wardenburg J, Kobayashi SD, Mathema B, Braughton KR, Whitney AR, Villaruz AE, Martens CA et al. 2011. Molecular differentiation of historic phage-type 80/81 and contemporary epidemic *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **108**(44): 18091-18096.
- Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF et al. 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet* **367**(9512): 731-739.
- Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG. 2003. Measurably evolving populations. *Trends in Ecology and Evolution* **18**(9): 481-488.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution* **29**(8): 1969-1973.
- Edwards AM, Massey RC, Clarke SR. 2012. Molecular mechanisms of *Staphylococcus aureus* nasopharyngeal colonization. *Molecular oral microbiology* **27**(1): 1-10.
- Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, Barnett R, O'Connell TC, Coxon P, Monaghan N et al. 2011. Ancient hybridization and an Irish origin for the modern polar bear matriline. *Current biology : CB* **21**(15): 1251-1258.
- Ellington MJ, Hope R, Livermore DM, Kearns AM, Henderson K, Cookson BD, Pearson A, Johnson AP. 2010. Decline of EMRSA-16 amongst methicillin-resistant *Staphylococcus aureus* causing bacteraemias in the UK between 2001 and 2007. *J Antimicrob Chemother* **65**(3): 446-448.
- Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic acids research* **30**(7): 1575-1584.
- Eyre DW, Golubchik T, Gordon NC, Bowden R, Piazza P, Batty EM, Ip CL, Wilson DJ, Didelot X, O'Connor L et al. 2012. A pilot study of rapid benchtop sequencing of *Staphylococcus aureus* and *Clostridium difficile* for outbreak detection and surveillance. *BMJ Open* **2**(3).
- Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, Peacock SJ, Smith JM, Murphy M, Spratt BG et al. 2003. How clonal is *Staphylococcus aureus*? *Journal of bacteriology* **185**(11): 3307-3316.

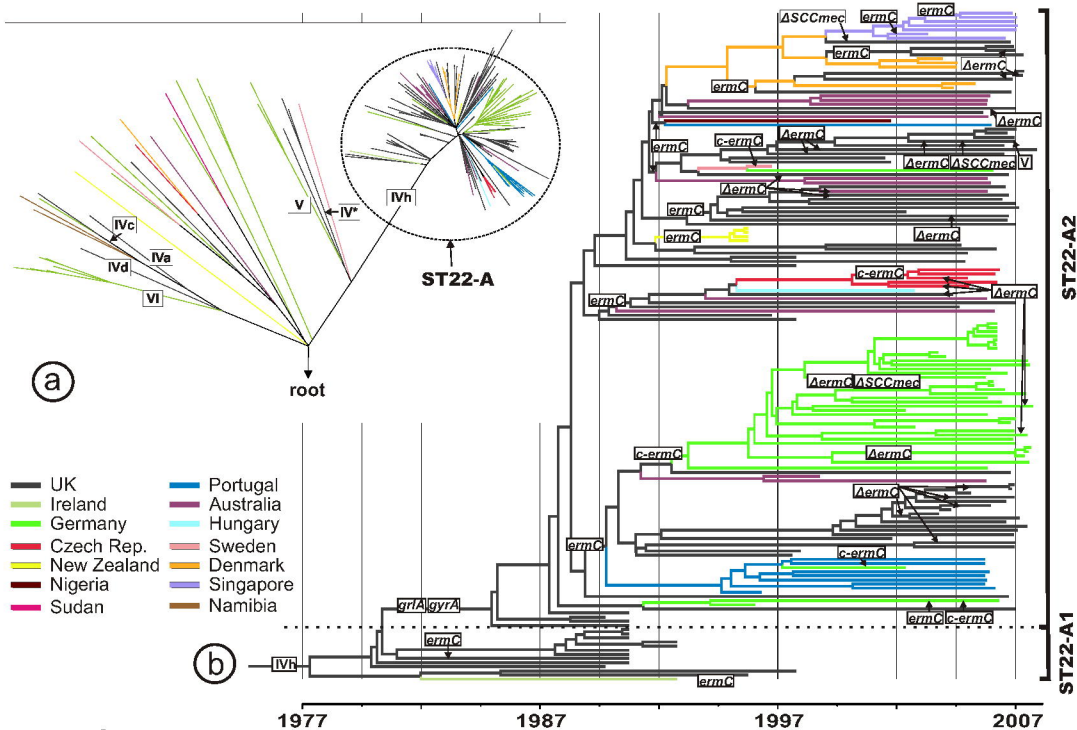
- Finch R, Whitby M, Craddock C, Holliday A, Martin J, Pilkington R. 1986. Clinical evaluation of treatment with Ciprofloxacin. *European Journal of Clinical Microbiology & Infectious Diseases* **5**(2): 257-259.
- Fluit AC, Wiolders CL, Verhoef J, Schmitz FJ. 2001. Epidemiology and susceptibility of 3,051 *Staphylococcus aureus* isolates from 25 university hospitals participating in the European SENTRY study. *J Clin Microbiol* **39**(10): 3727-3732.
- Fowler T, Wann ER, Joh D, Johansson S, Foster TJ, Hook M. 2000. Cellular invasion by *Staphylococcus aureus* involves a fibronectin bridge between the bacterial fibronectin-binding MSCRAMMs and host cell beta1 integrins. *European journal of cell biology* **79**(10): 672-679.
- Gill SR, Fouts DE, Archer GL, Mongodin EF, Deboy RT, Ravel J, Paulsen IT, Kolonay JF, Brinkac L, Beanan M et al. 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *Journal of bacteriology* **187**(7): 2426-2438.
- Gomes AR, Vinga S, Zavolan M, de Lencastre H. 2005. Analysis of the genetic variability of virulence-related loci in epidemic clones of methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* **49**(1): 366-379.
- Greene C, McDevitt D, Francois P, Vaudaux PE, Lew DP, Foster TJ. 1995. Adhesion properties of mutants of *Staphylococcus aureus* defective in fibronectin-binding proteins and studies on the expression of *fnb* genes. *Molecular microbiology* **17**(6): 1143-1152.
- Griggs DJ, Marona H, Piddock LJ. 2003. Selection of moxifloxacin-resistant *Staphylococcus aureus* compared with five other fluoroquinolones. *J Antimicrob Chemother* **51**(6): 1403-1407.
- Grundmann H, Aanensen DM, van den Wijngaard CC, Spratt BG, Harmsen D, Friedrich AW. 2010. Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: a molecular-epidemiological analysis. *PLoS Med* **7**(1): e1000215.
- Harris SR, Cartwright EJ, Torok ME, Holden MT, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J et al. 2012. Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis*.
- Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA et al. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327**(5964): 469-474.
- Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* **22**(2): 160-174.
- Hawkey PM. 1997. Quinolones in sweat and quinolone resistance. *Lancet* **349**(9046): 148-149.
- Hodgson JE, Cumock SP, Dyke KG, Morris R, Sylvester DR, Gross MS. 1994. Molecular characterization of the gene encoding high-level mupirocin resistance in *Staphylococcus aureus* J2870. *Antimicrob Agents Chemother* **38**(5): 1205-1208.
- Holden MT, Lindsay JA, Corton C, Quail MA, Cockfield JD, Pathak S, Batra R, Parkhill J, Bentley SD, Edgeworth JD. 2010. Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methicillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). *Journal of bacteriology* **192**(3): 888-892.
- Holden MTG, Feil EJ, Lindsay JA, Peacock SJ, Day NPJ, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R et al. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance. *Proceedings of the National Academy of Sciences of the United States of America* **101**(26): 9786-9791.
- Holtfreter S, Grumann D, Schmudde M, Nguyen HT, Eichler P, Strommenger B, Kopron K, Kolata J, Giedrys-Kalemba S, Steinmetz I et al. 2007. Clonal distribution of superantigen genes in clinical *Staphylococcus aureus* isolates. *J Clin Microbiol* **45**(8): 2669-2680.
- Hori S, Sunley R, Tami A, Grundmann H. 2002. The Nottingham *Staphylococcus aureus* population study: prevalence of MRSA among the elderly in a university hospital. *The Journal of hospital infection* **50**(1): 25-29.
- Hsu LY, Koh TH, Singh K, Kang ML, Kurup A, Tan BH. 2005. Dissemination of multisusceptible methicillin-resistant *Staphylococcus aureus* in Singapore. *J Clin Microbiol* **43**(6): 2923-2925.
- Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics* **11**: 119.
- Johnson AP, Aucken HM, Cavendish S, Ganner M, Wale MC, Warner M, Livermore DM, Cookson BD. 2001. Dominance of EMRSA-15 and -16 among MRSA causing nosocomial bacteraemia in the UK: analysis of isolates from the European Antimicrobial Resistance Surveillance System (EARSS). *J Antimicrob Chemother* **48**(1): 143-144.

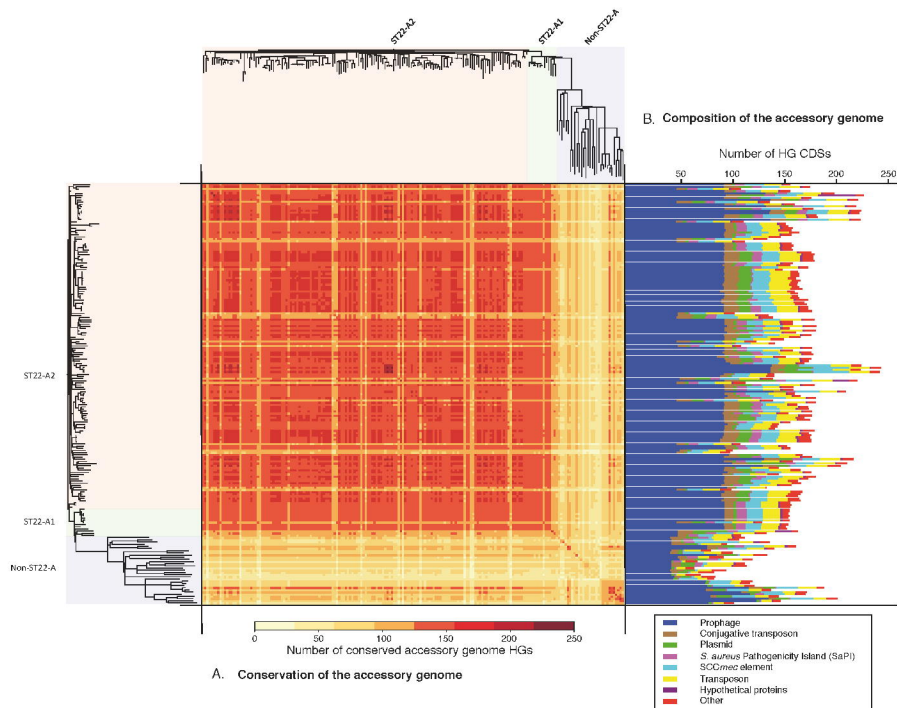
- Jones RN, Kohno S, Ono Y, Ross JE, Yanagihara K. 2009. ZAAPS International Surveillance Program (2007) for linezolid resistance: results from 5591 Gram-positive clinical isolates in 23 countries. *Diagn Microbiol Infect Dis* **64**(2): 191-201.
- Keim PS, Wagner DM. 2009. Humans and evolutionary and ecological forces shaped the phylogeography of recently emerged diseases. *Nat Rev Microbiol* **7**(11): 813-821.
- Kennedy AD, Otto M, Braughton KR, Whitney AR, Chen L, Mathema B, Mediavilla JR, Byrne KA, Parkins LD, Tenover FC et al. 2008. Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: recent clonal expansion and diversification. *Proc Natl Acad Sci U S A* **105**(4): 1327-1332.
- Khan SA, Novick RP. 1983. Complete nucleotide sequence of pT181, a tetracycline-resistance plasmid from *Staphylococcus aureus*. *Plasmid* **10**(3): 251-259.
- Klein E, Smith DL, Laxminarayan R. 2007. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999-2005. *Emerg Infect Dis* **13**(12): 1840-1846.
- Ko KS, Lee JY, Suh JY, Oh WS, Peck KR, Lee NY, Song JH. 2005. Distribution of major genotypes among methicillin-resistant *Staphylococcus aureus* clones in Asian countries. *J Clin Microbiol* **43**(1): 421-426.
- Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, Brown NM, Ogilvy-Stuart AL, Hsu LY, Chewapreecha C, Croucher NJ, Harris SR et al. 2012. A Neonatal MRSA Outbreak Investigation Using Rapid Whole Genome Sequencing. *New England Journal of Medicine* **366**(24): 2267-2275.
- Kuhn G, Koessler T, Melles DC, Francois P, Huyghe A, Dunman P, Vos MC, Zanetti G, Schrenzel J, van Belkum A et al. 2010. Comparative genomics of epidemic versus sporadic *Staphylococcus aureus* strains does not reveal molecular markers for epidemicity. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **10**(1): 89-96.
- Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui LZ, Oguchi A, Aoki K, Nagai Y et al. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**(9264): 1225-1240.
- Lafaurie M, Porcher R, Donay JL, Touratier S, Molina JM. 2012. Reduction of fluoroquinolone use is associated with a decrease in methicillin-resistant *Staphylococcus aureus* and fluoroquinolone-resistant *Pseudomonas aeruginosa* isolation rates: a 10 year study. *J Antimicrob Chemother* **67**(4): 1010-1015.
- Leelaporn A, Firth N, Byrne ME, Roper E, Skurray RA. 1994. Possible role of insertion sequence IS257 in dissemination and expression of high- and low-level trimethoprim resistance in staphylococci. *Antimicrob Agents Chemother* **38**(10): 2238-2244.
- Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLoS Comput Biol* **5**(9): e1000520.
- Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular biology and evolution* **27**(8): 1877-1885.
- Levin TP, Suh B, Axelrod P, Truant AL, Fekete T. 2005. Potential clindamycin resistance in clindamycin-susceptible, erythromycin-resistant *Staphylococcus aureus*: report of a clinical failure. *Antimicrob Agents Chemother* **49**(3): 1222-1224.
- Li M, Du X, Villaruz AE, Diep BA, Wang D, Song Y, Tian Y, Hu J, Yu F, Lu Y et al. 2012. MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med* **18**(5): 816-819.
- Limbago B, Fosheim GE, Schoonover V, Crane CE, Nadle J, Petit S, Heltzel D, Ray SM, Harrison LH, Lynfield R et al. 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**(5): 1344-1351.
- Linde H, Wagenlehner F, Strommenger B, Drubel I, Tanzer J, Reischl U, Raab U, Holler C, Naber KG, Witte W et al. 2005. Healthcare-associated outbreaks and community-acquired infections due to MRSA carrying the Panton-Valentine leucocidin gene in southeastern Germany. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* **24**(6): 419-422.
- Lindsay JA, Holden MTG. 2004. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol* **12**(8): 378-385.
- Livermore DM, James D, Reacher M, Graham C, Nichols T, Stephens P, Johnson AP, George RC. 2002. Trends in fluoroquinolone (ciprofloxacin) resistance in enterobacteriaceae from bacteremias, England and Wales, 1990-1999. *Emerg Infect Dis* **8**(5): 473-478.

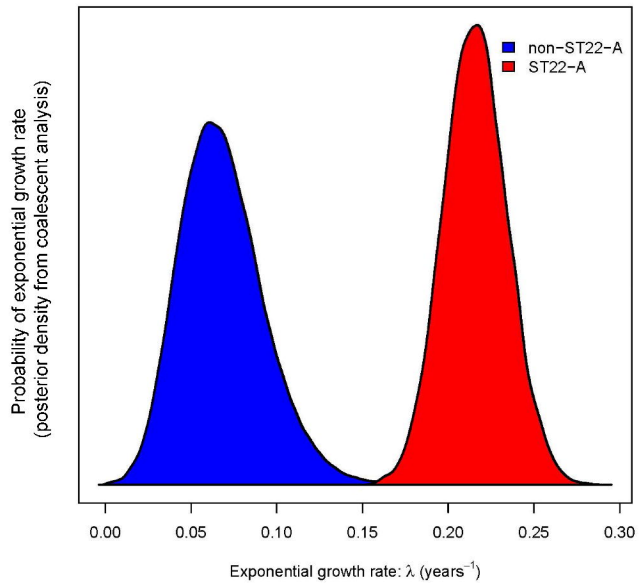
- Locke JB, Hilgers M, Shaw KJ. 2009. Novel ribosomal mutations in *Staphylococcus aureus* strains identified through selection with the oxazolidinones linezolid and terezolid (TR-700). *Antimicrob Agents Chemother* **53**(12): 5265-5274.
- Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, Cartwright RA, Simpson AJ, Rambaut A, Nübel U, Fitzgerald JR. 2009. Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **106**(46): 19545-19550.
- McAdam P, Templeton K, Edwards GF, Holden MTG, Feil EJ, Aanensen DM, Bargawi H, Spratt BG, Bentley SD, Parkhill J et al. 2012. Molecular Tracing of the Emergence, Adaptation and Transmission of Hospital-Associated MRSA. *Proc Natl Acad Sci U S A* **109**(23): 9107-9112.
- Melter O, Urbaskova P, Jakubu V, Mackova B, Zemlickova H. 2006. Emergence of EMRSA-15 clone in hospitals throughout the Czech Republic. *Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin* **11**(8): E060803 060806.
- Meyer E, Schwab F, Gastmeier P, Rueden H, Daschner FD. 2006. Surveillance of antimicrobial use and antimicrobial resistance in German intensive care units (SARI): a summary of the data from 2001 through 2004. *Infection* **34**(6): 303-309.
- Milheirico C, Oliveira DC, de Lencastre H. 2007. Multiplex PCR strategy for subtyping the staphylococcal cassette chromosome *mec* type IV in methicillin-resistant *Staphylococcus aureus*: 'SCCmec IV multiplex'. *J Antimicrob Chemother* **60**(1): 42-48.
- Novick RP. 2003. Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. *Plasmid* **49**(2): 93-105.
- Nübel U, Dordel J, Kurt K, Strommenger B, Westh H, Shukla SK, Zemlickova H, Leblois R, Wirth T, Jombart T et al. 2010. A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant *Staphylococcus aureus*. *PLoS Pathogens* **6**(4): e1000855.
- Nübel U, Roumagnac P, Feldkamp M, Song JH, Ko KS, Huang YC, Coombs G, Ip M, Westh H, Skov R et al. 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **105**(37): 14130-14135.
- O'Neill AJ, McLaws F, Kahlmeter G, Henriksen AS, Chopra I. 2007. Genetic basis of resistance to fusidic acid in staphylococci. *Antimicrob Agents Chemother* **51**(5): 1737-1740.
- Olsen RJ, Sitkiewicz I, Ayeras AA, Gonulal VE, Cantu C, Beres SB, Green NM, Lei B, Humbird T, Greaver J et al. 2010. Decreased necrotizing fasciitis capacity caused by a single nucleotide mutation that alters a multiple gene virulence axis. *Proc Natl Acad Sci U S A* **107**(2): 888-893.
- Parietti JJ, Cattoir V, Lebouvier G, Verdon R, Daubin C, du Cheyron D, Leclercq R, Charbonneau P. 2011. Hospital-wide modification of fluoroquinolone policy and methicillin-resistant *Staphylococcus aureus* rates: a 10-year interrupted time-series analysis. *The Journal of hospital infection* **78**(2): 118-122.
- Peacock SJ, Day NP, Thomas MG, Berendt AR, Foster TJ. 2000. Clinical isolates of *Staphylococcus aureus* exhibit diversity in *fnb* genes and adhesion to human fibronectin. *J Infect* **41**(1): 23-31.
- Pearman JW, Coombs GW, Grubb WB, O'Brien F. 2001. A British epidemic strain of methicillin-resistant *Staphylococcus aureus* (UK EMRSA-15) in Western Australia. *Med J Aust* **174**(12): 662.
- Phillips S, Novick RP. 1979. Tn554 - site-specific repressor-controlled transposon in *Staphylococcus aureus*. *Nature* **278**(5703): 476-478.
- Richardson JF, Reith S. 1993. Characterization of a strain of methicillin-resistant *Staphylococcus aureus* (EMRSA-15) by conventional and molecular methods. *The Journal of hospital infection* **25**(1): 45-52.
- Robinson DA, Enright MC. 2003. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* **47**(12): 3926-3934.
- Roche FM, Downer R, Keane F, Speziale P, Park PW, Foster TJ. 2004. The N-terminal A domain of fibronectin-binding proteins A and B promotes adhesion of *Staphylococcus aureus* to elastin. *The Journal of biological chemistry* **279**(37): 38433-38440.
- Rouch DA, Byrne ME, Kong YC, Skurray RA. 1987. The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and nucleotide sequence analysis. *Journal of general microbiology* **133**(11): 3039-3052.
- Rouch DA, Messerotti LJ, Loo LS, Jackson CA, Skurray RA. 1989. Trimethoprim resistance transposon Tn4003 from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS257. *Molecular microbiology* **3**(2): 161-175.

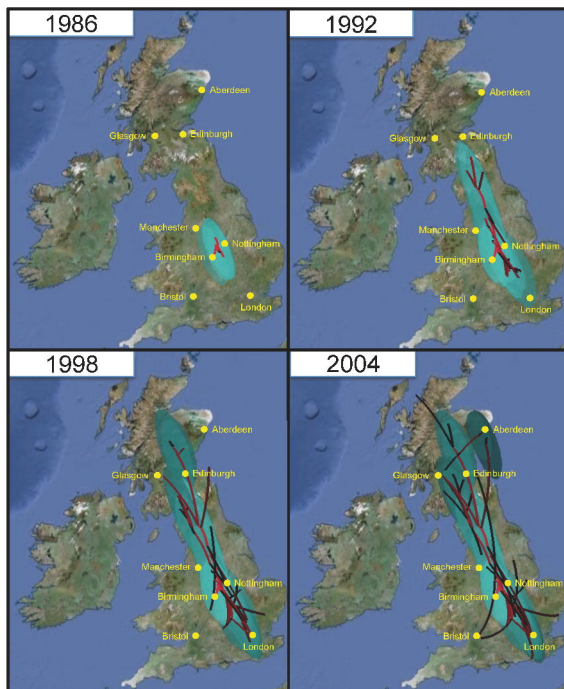
- Rowland SJ, Dyke KGH. 1989. Characterization of the Staphylococcal beta-lactamase transposon Tn552. *Embo J* **8**(9): 2761-2773.
- Ruimy R, Armand-Lefevre L, Barbier F, Ruppe E, Cocojaru R, Mesli Y, Maiga A, Benkalfat M, Benchouk S, Hassaine H et al. 2009. Comparisons between geographically diverse samples of carried *Staphylococcus aureus*. *Journal of bacteriology* **191**(18): 5577-5583.
- Sako T, Tsuchida N. 1983. Nucleotide sequence of the staphylokinase gene from *Staphylococcus aureus*. *Nucleic acids research* **11**(22): 7679-7693.
- Salangsang JA, Harrison LH, Brooks MM, Shutt KA, Saul MI, Muto CA. 2010. Patient-associated risk factors for acquisition of methicillin-resistant *Staphylococcus aureus* in a tertiary care hospital. *Infect Control Hosp Epidemiol* **31**(11): 1139-1147.
- Shinji H, Yosizawa Y, Tajima A, Iwase T, Sugimoto S, Seki K, Mizunoe Y. 2011. Role of fibronectin-binding proteins A and B in in vitro cellular infections and in vivo septic infections by *Staphylococcus aureus*. *Infect Immun* **79**(6): 2215-2223.
- Silverman SH, Johnson M, Burdon DW, Keighley MRB. 1986. Pharmacokinetics of single dose intravenous Ciprofloxacin in patients undergoing gastrointestinal surgery. *Journal of Antimicrobial Chemotherapy* **18**(1): 107-112.
- Smyth DS, Robinson DA. 2009. Integrative and sequence characteristics of a novel genetic element, ICE6013, in *Staphylococcus aureus*. *Journal of bacteriology* **191**(19): 5964-5975.
- Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Henderson DK, Palmore TN, Segre JA. 2012. Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Science translational medicine* **4**(148): 148ra116.
- Sola C, Saka HA, Vindel A, Bocco JL. 2008. Emergence and dissemination of a community-associated methicillin-resistant Panton-Valentine leucocidin-positive *Staphylococcus aureus* clone sharing the sequence type 5 lineage with the most prevalent nosocomial clone in the same region of Argentina. *J Clin Microbiol* **46**(5): 1826-1831.
- Stamatakis A, Ludwig T, Meier H. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**(4): 456-463.
- Strommenger B, Bräulke 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.
- Tenover FC, Tickler IA, Goering RV, Kreiswirth BN, Mediavilla JR, Persing DH. 2011. Characterization of nasal and blood culture isolates of methicillin-resistant *Staphylococcus aureus* from patients in United States Hospitals. *Antimicrob Agents Chemother* **56**(3): 1324-1330.
- Utsui Y, Yokota T. 1985. Role of an altered penicillin-binding protein in methicillin- and cephem-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* **28**(3): 397-403.
- van Wamel WJ, Rooijakkers SH, Ruyken M, van Kessel KP, van Strijp JA. 2006. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *Staphylococcus aureus* are located on beta-hemolysin-converting bacteriophages. *Journal of bacteriology* **188**(4): 1310-1315.
- Voladri RK, Kernodle DS. 1998. Characterization of a chromosomal gene encoding type B beta-lactamase in phage group II isolates of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **42**(12): 3163-3168.
- Weber SG, Gold HS, Hooper DC, Karchmer AW, Carmeli Y. 2003. Fluoroquinolones and the risk for methicillin-resistant *Staphylococcus aureus* in hospitalized patients. *Emerg Infect Dis* **9**(11): 1415-1422.
- Weisblum B. 1995. Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrob Agents Chemother* **39**(4): 797-805.
- Witte W, Cuny C, Klare I, Nübel U, Strommenger B, Werner G. 2008. Emergence and spread of antibiotic-resistant Gram-positive bacterial pathogens. *Int J Med Microbiol* **298**(5-6): 365-377.
- Witte W, Enright M, Schmitz FJ, Cuny C, Bräulke C, Heuck D. 2001. Characteristics of a new epidemic MRSA in Germany ancestral to United Kingdom EMRSA 15. *Int J Med Microbiol* **290**(8): 677-682.
- Witte W, Strommenger B, Cuny C, Heuck D, Nübel U. 2007. Methicillin-resistant *Staphylococcus aureus* containing the Panton-Valentine leucocidin gene in Germany in 2005 and 2006. *J Antimicrob Chemother* **60**(6): 1258-1263.
- Woodford N, Livermore DM. 2009. Infections caused by Gram-positive bacteria: a review of the global challenge. *J Infect* **59** Suppl 1: S4-16.

Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome research* **18**(5): 821-829.









Number of genotypic resistance determinants per isolate

