Functional genomics of Staphylococcus aureus

Kyra Y.L. Chua, Timothy P. Stinear and Benjamin P. Howden

Advance Access publication date 21 February 2013

Abstract

Staphylococcus aureus remains a major opportunistic human pathogen, and while in many individuals it is associated with asymptomatic colonization, it is also capable of causing a range of clinical syndromes from minor skin infections to life-threatening septicemia. Staphylococcus aureus has also demonstrated a remarkable capacity to acquire antimicrobial resistance. Recent technological advances in genomics have led to an avalanche of studies providing deep insights into how S. aureus is evolving globally and within the human host. However, there are still significant experimental barriers in using these insights to try and better understand the biology of S. aureus. Here, we summarize recent advances in the understanding of S. aureus through the use of genomic approaches, and contemplate what the near future holds for truly functional genomics that will allow us to better understand the biology of this pathogen.

Keywords: Staphylococcus aureus; genomics; genome sequencing; pathogenesis; antibiotic resistance

INTRODUCTION

Staphylococcus aureus is a major cause of human infections ranging from relatively minor skin infections to more severe life-threatening infections, such as overwhelming sepsis and endocarditis in both the apparently healthy hosts and immuno-compromised patients [1]. In addition to its protean clinical manifestations, S. aureus is characterized by the ability to acquire resistance to almost any antibiotic, most importantly, methicillin. Methicillin-resistant S. aureus or MRSA, which was first reported in hospitalized patients in 1961, has now become a worldwide pandemic [2–4]. MRSA strains are resistant not only to methicillin but also to other penicillins and β-lactams. More recent emerging antimicrobial resistance issues include low-level vancomycin resistance and resistance to recently released agents, such as daptomycin and linezolid [5–7].

For many years, MRSA was a healthcare-associated problem. Although the incidence of healthcare-associated MRSA (HA-MRSA) has decreased in some regions, the phenomenon of community-associated MRSA (CA-MRSA) is being increasingly reported worldwide [8,9]. The CA-MRSA phenomenon was first described in patients from the remote Kimberley region in Western Australia in the early 1990s, but there have been subsequently multiple reports of heterogenous clones in widely dispersed geographical locations [9,10]. These clones are different to the HA-MRSA clones and cause infection and colonization in people with no healthcare contact. They also appear to cause more severe infections than HA-MRSA clones with early reports of fulminant sepsis in young children [11]. Staphylococcus aureus has a vast battery of defenses against the host immune system and is able to evade it on many levels [12]. It is thus a combination of virulence, immune evasion and rapid acquisition of resistance that makes S. aureus a highly efficient pathogen.

Kyra Y.L. Chua is undertaking her PhD in the area of Staphylococcus aureus pathogenesis where she has utilized whole-genome sequencing and genetic manipulation of clinical staphylococci. She is also a trainee Medical Microbiologist in Melbourne, Australia.

Timothy P. Stinear leads a research group at the University of Melbourne in Australia, and uses genomic approaches to understand how certain bacteria evolve and cause disease in humans.

Benjamin P. Howden is a practicing Infectious Diseases Physician and Medical Microbiologist in Melbourne, Australia. He leads a number of research projects in the areas of staphylococcal resistance and pathogenesis.

© The Author 2013. Published by Oxford University Press. All rights reserved. For permissions, please email: journals.permissions@oup.com
In the last decade, there have been significant advances in understanding the pathogenesis of staphylococcal infections, aided by the advent of high-throughput DNA sequencing and other new technologies. The key questions that have been addressed include understanding the organization and plasticity of the \textit{S. aureus} genome, defining the contributors to virulence in recently emerged highly virulent clones of CA-MRSA and uncovering novel mechanisms of antibiotic resistance. In this review, we will discuss lessons learnt using these genomic approaches, in particular, attempts to gain functional insights around questions of \textit{S. aureus} pathogenesis, using whole-genome sequencing of \textit{S. aureus}, transcriptome studies and the genetic manipulation of \textit{S. aureus}.

**Structure of the \textit{S. aureus} genome and comparative genomics**

The era of \textit{S. aureus} comparative genomics began with descriptions of the complete genomes of the MRSA strain N315, and the vancomycin-intermediate \textit{S. aureus} (VISA) strain Mu50 [13]. Initial observations noted the breadth of mobile DNA within \textit{S. aureus}. This included transposons, bacteriophages, pathogenicity islands and genomic islands. Many of these elements carried antibiotic resistance and virulence genes. In a subsequent study from the same group, the whole-genome sequence of a CA-MRSA strain, MW2, was analyzed [14]. Comparison of these three sequenced strains demonstrated overall a close relationship between the strains, although the N315 and Mu50 were more closely related to each other than to MW2. In comparison to the former two strains, MW2 had a novel repertoire of genes encoding putative virulence determinants including Panton–Valentine Leukocidin (PVL), staphylococcal enterotoxins, collagen adhesin protein (\textit{cna}) and staphylococcal superantigen-like proteins, which the authors postulated contributed to increased virulence of MW2. The staphylococcal collagen adhesin protein had previously demonstrated to be a virulence factor in infectious endocarditis [15]. MW2 also carried the bacteriocin gene cluster in one of its genomic islands, which may be a toxin to other bacteria, therefore assisting in its success in colonizing humans.

Since then there have been many other \textit{S. aureus} genomes completed and publically released (total number of completed genomes, \textit{n} = 28; Table 1), in addition to many other draft genomes. \textit{Staphylococcus aureus} genomes are approximately 2.8 Mbp in size and all sequenced strains have very similar chromosome architecture. Figure 1 illustrates a comparison of the completed \textit{S. aureus} genomes and highlights the variable or accessory regions.

Lindsay and Holden first discussed the concept of the ‘core’ \textit{S. aureus} genome [40]. This is the part of the genome that is present in all strains of \textit{S. aureus}. This is in contrast to the accessory genome that is variably present. The core \textit{S. aureus} genome is approximately 2.3 Mbp in size and contains housekeeping genes, genes required for growth and survival and many virulence genes including those encoding many of the microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), such as protein A and the fibronogen-binding proteins, as well as some exotoxins including \textit{α}-hemolysin and the phenol soluble modulins (PSMs) [41,42].

The accessory genome is composed of mobile genetic elements or elements that were once mobile. These include insertion sequences, plasmids, transposons, integrated bacteriophages and genomic or pathogenicity islands that contain genes which may encode proteins that contribute to the virulence and fitness in a particular environment.

**Role of the accessory genome in resistance and virulence**

An example of adaptation of the organism in the face of antibiotic pressure is acquisition of resistance genes. A key element in the armory of staphylococcal resistance is the SCC\textit{mec} cassette, which contains \textit{mecA} that encodes the altered penicillin-binding protein, PBP2a, allowing for reduced affinity for \textit{β}-lactam antibiotics [43,44]. Recently, a novel PSM gene, \textit{psm-mec}, was found in the SCC\textit{mecII}, III and VIII cassettes [45,46]. PSM-mec is a staphylococcal virulence factor and causes both neutrophil and red cell lysis [45]. In contrast to this direct cytotoxic effect, the PSM-mec ORF transcription product may also act as a regulatory factor to decrease the production of the \textit{α}-type PSMs, which have been demonstrated to be important virulence factors in septicemia and pneumonia animal models [47]. Therefore, the effect of this locus to increase or decrease virulence of a specific strain may, in fact vary depending on the degree of baseline expression of PSM-mec [46]. Unlike other \textit{psm} genes, which are located on the core staphylococcal genome, \textit{psm-mec} is an example of how virulence and resistance can
### Table 1: *Staphylococcus aureus* genome sequences

<table>
<thead>
<tr>
<th>Genome</th>
<th>MLST (CC)</th>
<th>Features</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCTC8125</td>
<td>8 (8)</td>
<td>Laboratory strain, UK, before 1949.</td>
<td>[16]</td>
</tr>
<tr>
<td>FPR3757 USA300</td>
<td>8 (8)</td>
<td>USA300 CA-MRSA causing skin and musculoskeletal infection, USA, reported in 2006. ACME locus described.</td>
<td>[76]</td>
</tr>
<tr>
<td>TCH1516 USA300</td>
<td>8 (8)</td>
<td>USA300 CA-MRSA, adolescent with severe sepsis, USA, reported in 2005.</td>
<td>[17]</td>
</tr>
<tr>
<td>USA300 genomes, n = 10</td>
<td>8 (8)</td>
<td>Comparative genomics of USA300 isolates demonstrating recent clonal diversification and expansion. Decreased virulence in one strain possibly due to mutations in <em>agrA</em> and <em>vraF</em>.</td>
<td>[18]</td>
</tr>
<tr>
<td>COL</td>
<td>250 (8)</td>
<td>Early MRSA strain, UK, 1961.</td>
<td>[20]</td>
</tr>
<tr>
<td>TW20</td>
<td>239 (8)</td>
<td>HA-MRSA strain causing bacteremia, UK, 2003.</td>
<td>[21]</td>
</tr>
<tr>
<td>JKD6008</td>
<td>239 (8)</td>
<td>HA-MRSA VISA strain causing endocarditis, New Zealand, 2003.</td>
<td>[22]</td>
</tr>
<tr>
<td>T0131</td>
<td>239 (8)</td>
<td>HA-MRSA, China, 2006. Novel gene sasX described, later demonstrated to have a role in colonization and pathogenesis [23].</td>
<td>[23]</td>
</tr>
<tr>
<td>ST239 genomes, n = 63</td>
<td>239 (8)</td>
<td>High-resolution phylogeny and detailed intercontinental transmission of this clone described.</td>
<td>[25]</td>
</tr>
<tr>
<td>VC40</td>
<td>8 (8)</td>
<td>Vancomycin and daptomycin resistant laboratory generated strain derived from RN4220, Germany, reported in 2012.</td>
<td>[26]</td>
</tr>
<tr>
<td>N315</td>
<td>5 (5)</td>
<td>HA-MRSA, Japan, 1982.</td>
<td>[27]</td>
</tr>
<tr>
<td>Mu50</td>
<td>5 (5)</td>
<td>HA-MRSA VISA, Japan 1997.</td>
<td>[28]</td>
</tr>
<tr>
<td>Mu3</td>
<td>5 (5)</td>
<td>HA-MRSA IVISA, Japan 1997. Mutation in groR resulting in change in vancomycin resistance phenotype.</td>
<td>[29]</td>
</tr>
<tr>
<td>JH1</td>
<td>105 (5)</td>
<td>HA-MRSA VSSA, USA, 2000.</td>
<td>[30]</td>
</tr>
<tr>
<td>ED98</td>
<td>5 (5)</td>
<td>Poultry strain, Ireland, reported in 2009</td>
<td>[32]</td>
</tr>
<tr>
<td>04-0298I, ST225 genomes, n = 73</td>
<td>225 (5)</td>
<td>HA-MRSA, Germany, 2004. High-resolution phylogeny of strains from geographically different backgrounds.</td>
<td>[33]</td>
</tr>
<tr>
<td>MW2</td>
<td>1 (1)</td>
<td>USA400 CA-MRSA causing fatal childhood infection, USA, 1998.</td>
<td>[14]</td>
</tr>
<tr>
<td>MSSA476</td>
<td>1 (1)</td>
<td>Community-acquired MSSA, UK, 1998.</td>
<td>[31]</td>
</tr>
<tr>
<td>MRSA252</td>
<td>36 (30)</td>
<td>EMRSA-16 HA-MRSA epidemic strain, UK, 1997.</td>
<td>[31]</td>
</tr>
<tr>
<td>CC30 genomes, n = 8</td>
<td>30, 36 (30)</td>
<td>Comparison of contemporary hospital-associated strains, CA-MRSA and phage 80/81 strains. Hospital-associated strains less virulent due to mutations in <em>agr</em> gene regulatory system and <em>hla</em>.</td>
<td>[33]</td>
</tr>
<tr>
<td>CC30 genomes, n = 87</td>
<td>30, 36, 500 (30)</td>
<td>High-resolution phylogeny of global CC30 strains, in particular EMRSA-16.</td>
<td>[34]</td>
</tr>
<tr>
<td>RF22</td>
<td>151</td>
<td>Cattle mastitis strain, Ireland, 1993.</td>
<td>[32]</td>
</tr>
<tr>
<td>S0385</td>
<td>398</td>
<td>‘LA-MRSA’ strain causing human endocarditis, Netherlands 2006. Many novel mobile genetic elements identified.</td>
<td>[33]</td>
</tr>
<tr>
<td>08BA02i76</td>
<td>398</td>
<td>‘LA-MRSA’ causing post-operative site infection. No known animal livestock contact. Canada, 2008.</td>
<td>[34]</td>
</tr>
<tr>
<td>ST398 genomes, n = 88</td>
<td>398</td>
<td>Human-associated strains found to be the ancestral clade, and containing human IEC, carried on integrated prophage.</td>
<td>[35]</td>
</tr>
<tr>
<td>ST398 genomes, n = 2</td>
<td>398</td>
<td>Human-associated strains contain IEC.</td>
<td>[36]</td>
</tr>
<tr>
<td>JKD6159</td>
<td>93</td>
<td>CA-MRSA causing severe pneumonia and musculoskeletal infection, Australia, 2004.</td>
<td>[37]</td>
</tr>
<tr>
<td>MSHRI132</td>
<td>CC75</td>
<td>CA-MRSA causing necrotizing fasciitis, Australia, 2006.</td>
<td>[38]</td>
</tr>
<tr>
<td>II189-97</td>
<td>80 (80)</td>
<td>CA-MRSA, skin infection, Denmark, reported in 2012.</td>
<td>[39]</td>
</tr>
<tr>
<td>MO3</td>
<td>59</td>
<td>CA-MRSA, pediatric musculoskeletal infection, Taiwan, 2002.</td>
<td>[40]</td>
</tr>
<tr>
<td>EDI33</td>
<td>133</td>
<td>Ovine mastitis strain, reported in 2010.</td>
<td>[41]</td>
</tr>
</tbody>
</table>

**Notes:** MLST, multi-locus sequence type; CC, clonal complex; IVISA, hetero-VISA; HA-MRSA, healthcare-associated MRSA; LA-MRSA, life-stock associated MRSA; IEC, immune evasion complex.
co-evolve—selection of methicillin resistance in a staphylococcal population also allows for selection of this virulence determinant.

Comparative genomics of a collection of clonal complex 30 strains, including those that were the etiological agent for menstrual toxic shock syndrome epidemic, demonstrated a high sequence identity of the staphylococcal pathogenicity island, SaPI2 amongst the 
\textit{tst}} positive strains [48]. This suggests a single horizontal gene transfer event in a common ancestor, which occurred 10–140 years before the epidemic in the 1970s and 80s. The epidemic was therefore caused by strains carrying \textit{tst} which were already widely disseminated rather than rapid clonal expansion of a single strain. This finding confirmed what had previously been found using microarray analysis of these \textit{tst} positive strains [49].

Other recent comparative genomics studies that have lent insight into the pathogenesis of \textit{S. aureus} include the examination of the ST398 clone. MRSA strains of this type have emerged as a significant colonizer in livestock, in particular, pigs [50]. There have been many CC398 draft genomes sequenced and published in the last 12 months [51,52]. Among these are isolates obtained from both livestock and humans and across many geographical areas, and have permitted a reconstruction of the evolution of CC398. This is a feature of comparative genomic studies that have a dataset that is sufficiently representative of the \textit{S. aureus} population being studied both temporally and spatially.

Price \textit{et al.} analyzed 88 CC398 strains from 19 countries and found a significant difference between the human-associated and animal-associated strains [51]. Comparative genomics demonstrated that the human-associated strains were the ancestral clade. Human-associated MSSA strains also contained phages encoding human innate immune modulators [immune evasion complex (IEC)] allowing for infection/colonization in humans. Lifestock-associated CC398 was also marked by the acquisition of resistance genes encoding for tetracycline and methicillin resistance.

Uhlemann \textit{et al.} also independently identified this lack of prophage carrying the IEC genes [52]. These authors provided further evidence for the role of this gene complex in the adaptation to different hosts. They found that human-associated CC398 had increased \textit{in vivo} binding to human epidermal keratinocytes compared with livestock-associated CC398.

Using whole-genome comparisons, the evolution of \textit{S. aureus} during transition from colonizer to invasive pathogen has been recently explored. This study demonstrated that mutations accumulate at a steady rate during colonization; however, a number of mutations predicted to lead to non-functional proteins were detected during the transition to invasive pathogen, suggesting a possible role for these proteins in colonization and a selective disadvantage for their presence during the transition to invasive phenotype [53].

\textbf{Moving from the genome to the transcriptome, with an emphasis on the use of RNA-seq}

Methods used to study the expression of genes on a genome-wide scale under varying environmental conditions include DNA microarrays, and more recently RNA sequencing (RNA-seq). Over the past decade, DNA microarrays have been the staple method to define the transcriptional reach of several staphylococcal regulators including the \textit{agr} system, as well as SarA [54], Rot [55] and SigB [56]. These studies compare the global transcriptional profile of selected mutant to wild-type strains, often defining a large ‘regulon’, in some cases many hundreds of genes that are differentially regulated in the mutant strains. While such studies have been important for defining the global effects of selected regulators, in many cases it has not been possible to define whether the transcriptional changes are due to the direct or indirect impact of these proteins.

Another powerful application of DNA microarrays has been the examination of antibiotic resistant strains of \textit{S. aureus}. Using DNA microarrays to compare the transcriptomes of pairs of vancomycin–susceptible \textit{S. aureus} (VSSA) and VISA, researchers have shown that diverse global transcriptional changes can occur between different pairs of strains, with some consistent changes that are central to the resistance mechanism [57]. Subsequently, based on the comparisons of the whole-genome sequences of clinical, paired VSSA and VISA strains, mutations were identified, such as in the essential regulator \textit{walKR}, and the mutants (or wild type) strains were each re-generated by making single base substitutions. DNA microarrays were then employed and helped define a clear role of \textit{WalKR} as a regulator of genes involved in several aspects of staphylococcal metabolism [58]. Transcriptome studies have also examined the role of \textit{WalKR} as a positive regulator of
MSCRAMMs genes (efb, emp, fnbA, fnbB), genes involved in cytolysis (hlgACB, hla, hlb) and innate immune defense evasion (scn, chp, sbi), through activation of the SaeSR two-component system [59]. Strains producing constitutively active WalRc were found to be less virulent due to stimulation of the host inflammatory response by higher levels of released peptidoglycan fragments, presumably as a result of deranged peptidoglycan degradation [59].

Because microarrays are inherently limited by the specificity of their fluorescent signals to allow detection of binding to the oligonucleotide or cDNA probes, there has been significant interest in using RNA-seq, which employs high-throughput DNA sequencing technology. RNA-seq also allows for determination of the staphylococcal non-coding genome including regulatory RNA. However, few studies have been published using this technology with S. aureus to date. The first staphylococcal transcriptome analysis using high-throughput sequencing technology was performed on S. aureus N315 with the aim of defining non-coding regulatory RNAs in the strain [60]. This study uncovered approximately 200 transcripts and antisense RNAs that may act as regulatory RNAs. The expression of antisense RNAs varied significantly depending on stress conditions imposed on the organism and also on the growth phase. In a recent study using RNA-seq, the extensive role of RNase III in post-transcriptional regulation, was defined, with novel RNase III substrates also identified [61].

**Moving from hypothesis to fact—integrating findings from genomics and the genetic manipulation of S. aureus**

Although the information gleaned from various omic approaches are powerful hypothesis generators, carefully performed mutagenesis or other genetic manipulation experiments are still necessary to confirm functions at the level of individual genes,
operons or even pathways. For example, comparison of allelic mutants or gene knockouts generated by genetic manipulation of strains in both \textit{in vitro} experiments and animal virulence studies.

Currently, the most common method for genetic manipulation of \textit{S. aureus} is by transformation by electroporation. In this technique, electrocompetent cells in the exponential phase of growth are exposed to a high-voltage electrical field that creates transient pores in the cell membrane, in the presence of plasmid DNA, thus facilitating plasmid uptake by the staphylococcal strain of interest [62,63]. The large number of plasmids available for the genetic manipulation of \textit{S. aureus} has been comprehensively reviewed by McNamara [64] and Monk \textit{et al.} [65]. These plasmids usually have both \textit{Escherichia coli} and \textit{S. aureus} origins of replication, antibiotic resistance genes to allow for selection, and a multiple cloning site for targeted insertion of DNA sequences. Bae and Schneewind described an allelic replacement method using the plasmid pKOR1, which allows for the generation of unmarked mutants that has been very useful to many groups [58,66,67].

An elegant study by DeLeo \textit{et al.} [67] is an example of integrating information gleaned from genomic comparisons of a large collection of strains and then assessing the impact of identified mutations in wet laboratory experiments. The authors were able to use these approaches to address complex and important questions regarding the evolution of \textit{S. aureus}. In this study, clonal complex 30 \textit{S. aureus} isolates that included contemporary hospital-associated strains, CA-MRSA clones and the epidemic penicillin-resistant phage type 80/81 clone were compared. The contemporary hospital-associated strains were found to contain non-synonymous SNP in \textit{hla}, the gene encoding \( \alpha \)-hemolysin, which was not present in the CA-MRSA and phage type 80/81 strains. This pore-forming toxin found in almost all strains of \textit{S. aureus} has been demonstrated to be an important virulence factor, in staphylococcal skin infection and pneumonia [68,69]. This SNP in \textit{hla} resulted in a premature stop codon that resulted in an absence of expression of Hla. In addition, the contemporary hospital-associated strains also contained a non-synonymous SNP in \textit{agrC}. The \textit{agr} system has been established as a key staphylococcal regulator in many different \textit{S. aureus} genetic backgrounds and its importance in virulence has been clearly demonstrated in many different animal models [70–72]. Using allelic replacement, these changes were introduced into a virulent phage type 80/81 strain, and this significantly attenuated virulence. This study is illustrative of how strains with the same genetic ‘backbone’ arise at different time points to give quite different phenotypes. The less virulent phenotype observed with the contemporary hospital-associated strains may have evolved to allow for long-term colonization and persistence, as well as restricting these strains to the hospitalized human host [67].

Similarly, using glycopeptide resistance as an example, we, and others, have used comparative genomics to identify mutations that differentiate vancomycin-susceptible and intermediate resistant (VISA) isolate pairs isolated from the same patient before and after vancomycin treatment failure [58,73–75]. Subsequently, we were able to use the allelic replacement vector, pKOR1 [66], to delineate the impact of selected single base substitutions in \textit{walKR} and \textit{gatRS}, using allelic exchange and functional assessment of antibiotic resistance and transcriptomics [58,73]. Interestingly, mutations in \textit{walKR} were found to mediate vancomycin and daptomycin resistance as well as reduced \textit{agr} activity, which resulted in reduced virulence in an invertebrate virulence model [58]. These combined comparative genomic and genetic manipulation studies highlight the potential global effects of subtle genetic changes in \textit{S. aureus}, where single base substitutions can impact resistance to multiple antibiotics, as well as affecting virulence.

The importance of wet laboratory experiments to test hypotheses generated by comparative genomics is illustrated in the following two examples.

The arginine catabolic mobile element (ACME) is a novel genetic locus that was uncovered during the whole-genome sequence of FPR3757 USA300, the epidemic North American CA-MRSA clone [76]. Although not all USA300 isolates carry the complete ACME locus, it was hypothesized to contribute to the increased virulence of USA300 and aid in human colonization, and thus explain the success of the USA300 clone [76,77]. However, direct examination of the role of ACME in virulence by deletion of ACME only demonstrated a modest reduction in virulence [78,79]. The exact contribution of ACME to persistence and transmission of \textit{S. aureus} is yet to be defined.

Recently, by examining the genomes of seven CA-MRSA isolates, we demonstrated that there were three elements in the accessory genome, including the PVL carrying phage, which were common to the majority of these isolates. The
other two elements were a plasmid, pMW2 and the SCCmecIV cassette [80]. While some of these strains were distantly related to each other based on SNP comparisons of the S. aureus core genome, the same analysis based on SNP comparisons of these accessory elements identified significantly less sequence differences. This indicated very recent acquisition of these elements from a single source. We inferred that, not only were there common selective pressures acting on genetically and geographically diverse S. aureus populations, there was a possible role for these elements in the fitness-for-spread and virulence of these strains in the community. However, direct examination of the role of PVL in the pathogenesis of staphylococcal infection using genetic mutants in various animal models has been fraught with controversy [81,82]. Some, but not all these differences in findings can be attributed to the differential susceptibility of host neutrophils to PVL. Murine and monkey neutrophils are significantly more resistant to PVL in comparison with human and rabbit neutrophils [83]. In addition, the importance of PVL is also likely to be dependent on site of infection. In the rabbit pneumonia model, PVL has been demonstrated to mediate severe lung necrosis and inflammation [84]. In contrast, in skin infection, even in the rabbit model, its role remains somewhat controversial [85,86].

Although these examples have employed plasmid-mediated allelic exchange methods to generate mutants, recently a targetron system that are mobile group II introns, has been described for site-directed gene disruption in S. aureus [87,88].

More large scale and less targeted generation of mutants may also be employed using transposon mutagenesis [89] and signature tagged mutagenesis [90]. Recently, a collection of transposon (bursa aurealis) insertion mutants of a S. aureus USA300 LAC strain has been generated (Nebraska Transposon Mutant Library, http://www.narsa.net, accessed 6 October 2012). This collection is available in collaboration with the Network on Antimicrobial Resistance in S. aureus. It is anticipated that this defined mutant library, which includes selected mutants for all non-essential S. aureus genes in the USA300 background, will allow for genome-scale screening of virulence factors or antibiotic resistance.

**Barriers to progress**

Whilst the staphylococcal research community has made significant progress, there remain some hurdles to rapid and exponential gains in the field. While the cost of genome sequencing continues to fall, a rate-limiting factor in many institutions is analyzing data generated from these analyses quickly, primarily due to a paucity of bioinformatics support. It is anticipated that as software and user interfaces to interrogate ‘omic’ data improve, this obstacle will be overcome [91].

The application of RNA-seq to analyze the transcriptome from in vivo infection, while potentially providing important new insights regarding gene regulation during the infection process, remains problematic, primarily because of the contamination of host RNA with staphylococcal RNA and the recovery of isolation of high quality input material that is accurately reflective and unbiased of the relative abundance of RNA molecules [92]. This is perhaps why there are still very few publications reporting on the use of this technique.

*Staphylococcus aureus* contains a number of restriction-modification (RM) systems that prevent easy genetic manipulation of strains. Indeed, these barriers are why certain staphylococcal strains have been, to date, not genetically tractable. The RM systems are classified into type I, II and IV RM systems [93]. The most widely distributed RM system present and functional in the majority of strains of S. aureus is the SauI type I RM system [94] which protects self-DNA from cleavage by catalyzing methylation of hemi-methylated residues within a target sequence whilst also recognizing foreign unmethylated DNA at the same target sequence and cleaving DNA at a site distant to the target sequence [94,95]. The laboratory S. aureus strain RN4220 lacks restriction barriers, allowing it to act as a recipient for foreign DNA such as plasmid DNA from *E. coli* [96]. After passaging through RN4220, plasmid DNA is then appropriately methylated and therefore ready to be used for the transformation of the target S. aureus strain. However, it should be noted that not all S. aureus strains, including those that belong to phage type 94/96 (ST25), will readily accept foreign DNA, even after the RN4220 intermediary step [97]. Recently, Corvaglia and colleagues identified a type IV RM system in S. aureus that contributed to the resistance against foreign DNA [88,98]. RN4220 contains mutations in both the type I and IV RM system [88,94].

Our experience with a unique S. aureus clone ST93, which is an emerging CA-MRSA clone in our region, demonstrated the difficulties with
genetically manipulating certain strains of \textit{S. aureus}. It has been very difficult to introduce DNA in ST93 (Kyra Chua, personal communication). In the ST93 genome, in addition to the usual complement of two type I RM systems, we identified an RM system that contains a distinct \textit{hsdS} gene \cite{80}. We predict that this previously undiscovered RM system explains, at least in part, the resistance of ST93 to genetic modification.

Recently, a high efficiency \textit{E. coli} cloning strain containing a DNA cytosine methyltransferase (\textit{dcm}) mutation was described and plasmid obtained from this strain was used to transform many strains of previously untransformable \textit{S. aureus} isolates \cite{65}. Tools such as these will greatly assist new advances in understanding the pathogenesis and resistance in the staphylococcal field.

**Future perspectives**

Much progress has been made in understanding the biology of \textit{S. aureus}. Despite this, \textit{S. aureus} infections remain very common, an effective vaccine is not available and new antibiotic resistance issues continue to evolve. Establishment of simple genetic manipulation tools for all \textit{S. aureus} strains will enhance progress toward a more comprehensive understanding of \textit{S. aureus} biology. It is critical that tools are developed for the effective analysis of \textit{S. aureus} during active infection, moving away from the standard \textit{in silico} and \textit{in vitro} analyses, where the important characteristics of the \textit{in vivo} milieu may be overlooked. The availability of high throughput, low cost whole-genome sequencing provides an opportunity to explore the genomics predictors of virulence and treatment outcome; however, this requires an ‘integrated omics’ or systematic systems biology approach incorporating large amounts of sequence data with standardized virulence assessment, transcriptional analysis and proteomics studies, as recently proposed \cite{99}. These types of analyses will provide key insights into the genetic predictors of virulence that are much more sophisticated than simple presence or absence of virulence genes.

- New tools for genetic manipulation of \textit{S. aureus} are urgently needed. In particular, we need to better understand the different \textit{S. aureus} restriction-modification systems so as to overcome barriers to genetically modify staphylococcal isolates and test hypotheses as to the contribution of the many targets identified by comparative genomics to virulence and antibiotic resistance.

**FUNDING**

This work was supported by the National Health and Medical Research Council (NHMRC), Australia.

**References**

Functional genomics of S. aureus


