Genome and virulence determinants of high virulence community-acquired MRSA

Tadashi Baba, Fumihiko Takeuchi, Makoto Kuroda, Harumi Yuzawa, Ken-Ichi Aoki, Akio Oguchi, Yoshimi Nagai, Natsuko Iwama, Kazuyuki Asano, Timothy Naimi, Hiroko Kuroda, Longzhu Cui, Kenji Yamamoto, Keiichi Hiramatsu

Summary

Background A new type of meticillin-resistant Staphylococcus aureus (MRSA), designated community-acquired MRSA, is becoming increasingly noticeable in the community, some strains of which cause fatal infections in otherwise healthy individuals. By contrast with hospital-acquired MRSA, community-acquired MRSA is more susceptible to non β-lactam antibiotics. We investigated the high virulence potential of certain strains of this bacterium.

Methods We ascertained the whole genome sequence of MW2, a strain of community-acquired MRSA, by shotgun cloning and sequencing. MW2 caused fatal septicemia and septic arthritis in a 16-month-old girl in North Dakota, USA, in 1998. The genome of this strain was compared with those of hospital-acquired MRSA strains, including N315 and Mu50.

Findings Meticillin resistance gene (mecA) in MW2 was carried by a novel allelic form (type Iva) of staphylococcal cassette chromosome mec (SCCmec), by contrast with type II in N315 and Mu50. Type Iva SCCmec did not carry any of the multiple antibiotic resistance genes reported in type II SCCmec. By contrast, 19 additional virulence genes were recorded in the MW2 genome. All but two of these virulence genes were noted in four of the seven genomic islands of MW2.

Interpretation MW2 carried a range of virulence and resistance genes that was distinct from those displayed on the chromosomes of extant S aureus strains. Most genes were carried by specific allelic forms of genomic islands in the MW2 chromosome. The combination of allelic forms of genomic islands is the genetic basis that determines the pathogenicity of medically important phenotypes of S aureus, including those of community-acquired MRSA strains.

Lancet 2002; 359: 1819–27

See Commentary page 1791

Introduction

Meticillin-resistant Staphylococcus aureus (MRSA), which arose in the 1960s, typically causes infections in inpatients, who have risk factors associated with health care. In the past 5 years however, MRSA infections have been described in the general population. These infections arise in the community, and thus affected people lack traditional risk factors such as recent admittance, surgery, or long-term residence in care facilities. Community-acquired MRSA infections can cause serious and even fatal infections in otherwise healthy hosts.1,2

Community-acquired MRSA typically affects children and young adults, and it causes a range of infections similar to those caused by community-acquired meticillin-susceptible S aureus (MSSA).1 Furthermore, these infections are more likely to be susceptible to many antimicrobial classes and to have different pulsed-field gel electrophoresis subtypes than isolates that are associated with health care. There is an ominous sign that MRSA strains circulating beyond nosocomial settings are replacing community-acquired MSSA as the flora of healthy human beings in some communities.

Recently, we established the whole genome sequence of two MRSA strains, N315 and Mu50, both of which are strains associated with health care.1 Two further hospital-acquired MRSA strains have been sequenced by others (COL, E-MRSA-16 [strain 252]; see Methods). However, the genome of any community-acquired strain with high virulence has not been analysed. In this study, we sequenced the whole genome of the community-acquired MRSA strain MW2, and aimed to investigate the genetic basis for this strain’s high virulence by comparison of its genome with those of MRSA strains that are associated with health care.

Methods

MW2 is a typical community-acquired strain of MRSA, which was isolated in 1998 in North Dakota, USA.1 MW2 is so-called because strains of similar pulsed-field gel electrophoresis genotypes are reported widely in midwestern USA. MW2 caused fatal septicemia and septic arthritis in a 16-month-old American-Indian girl who had no risk factors associated with health care. She died within 2 h of arrival at hospital. The MW2 strain is susceptible to many antibiotic classes, apart from β-lactam antibiotics, to which the strain shows a typical heterogeneous-type resistance.

We established the whole genome sequence as described previously,1 with the exception that we used genomic DNA fragments of 1·5–3·0 kb instead of 1·0–2·2 kb to make the plasmid library. We based the final genome sequence of MW2 on about 64 000 sequences. We have entered the whole genome sequence of MW2 primarily in the DNA Database of Japan, with accession numbers AP004822–AP004831 (chromosome), AP004832 (plasmid), and BA000033 (11 entries and contig).
GENOMIC ISLAND ALLOTYPING

Clinically important properties, drug resistance, and the pathogenic potential of S aureus strains are mostly established by the set of genomic islands that strain possesses. Each island is found in several alternative forms with different sets of virulence or resistance genes among S aureus strains—i.e., it exists in several allelic forms. Since different allelic forms of islands encode different properties, identifying the allelic set of islands that the strain carries gives information on the overall properties of the strain, including its pathogenic potential and the pattern of antibiotic susceptibilities.

MEC OPERON

Composed of mecA, mecR1, and mecI. The latter two are gene homologues of blaR1 and bla, respectively. MecR1 is the membrane-bound signal transducer recognising β-lactam antibiotics, and Bla is a repressor. Therefore, the same sequence of events arises as with bla operon. The only difference is that mecA encodes cell-wall synthesis enzyme PBP2* instead of penicillinase encoded by bla: β-lactams induce metillin resistance in this case.

PANTON-VALENTINE LEUKOCIDIN

Member of the leukocidin group of bacterial toxins that kill leucocytes by creation of pores in the cell membrane. Apart from S aureus Panton-Valentine leukocidin, the group includes Pasteurella haemolytica leukotoxin, Actinobacillus actinomycetemcomitans leukotoxin, Listeria monocytogenes listeriolysin, Escherichia coli haemolysin, and Fusobacterium necrophorum leukotoxin.

RESTRICTION-MODIFICATION SYSTEM

Bacteria protect themselves from invasion of foreign DNA (eg, bacteriophage DNA) by cleaving DNA (restriction) at specific recognition sites. The restriction enzyme HsdR (HSD is the acronym of host-specificity determinant) cleaves the specific DNA site that HsdS (S stands for specificity) recognises. To protect the bacteria’s own DNA from restriction, DNA methylase, HsdM, and HsdS together modify the specific site by methylation and methylated recognition sites are not cleaved by HsdR. When the hsdM gene encoding HsdM is present in the genomic island, the island is considered to be protected from spontaneous loss from the chromosome, because such cells may not survive the self-digestion of the chromosome that occurs when hsdM is gene is lost.

STAPHYLOCOCCAL CASSETTE CHROMOSOME (SCC)

A unique family of mobile genetic elements found on the chromosomes of Staphylococcus spp. In S aureus chromosome, it is always integrated at a fixed site near the origin of replication. If SCC carries metillin resistance gene (meca), it is called SCCmec, the integration of which in the S aureus chromosome generates metillin-resistant S aureus (MRSA).

SUPERANTIGENS

A variety of molecules that share the ability to activate large populations of T lymphocytes through co-ligation between major histocompatibility complex (MHC) class II molecules on antigen-presenting cells and the variable portion of the T-cell antigen receptor β chain; the pattern of Vβ activation is specific for each of these superantigens. They include staphylococcal enterotoxins, exotoxins, and toxic-shock-syndrome toxin.

SYNONYMOUS CODON USAGE

Each aminoacid of protein is coded by a set of three nucleotides (a codon) in the gene. Some codons, for example GCT, GCC, GCA, GCG, code the same aminoacid, alanine. Such codons are called synonymous codons. The proportion of usage of synonymous codons to encode each aminoacid is known to vary for different proteins and for different species.

We initially identified open reading frames (orf)—sections of the genome likely to encode proteins—with GLIMMER and rbsfinder software; we ignored orfs comprising 30 or fewer codons. We reviewed predicted orfs individually with the GAMBLER program (Xanagen, Kawasaki, Japan), and mutually overlapping orfs were identified and removed in some cases. We searched a non-redundant protein database with ORFs—peptides encoded by orfs—with BLAST software. We identified transfer RNA and tmRNA* genes by tRNAscan-SE and with a procedure available online.

The best-hit entries identified by BLASTP searches with e-values of 1·0×10−5 or less were classified into taxonomic groups in accordance with the NCBI (National Center for Biotechnology Information) taxonomy database. We excluded hits against entries for Staphylococcus genus from this analysis.

We calculated bias in SYNONYMOUS CODON USAGE against averages of all orfs longer than 150 bp (B.) and of ribosomal proteins (B.) and GC content at the third codon position (GC3), as described previously. With these values, we defined putative highly expressed genes and alien genes as orfs with B.<0·35 and B.>0·52 and B.>0·345, respectively. We also regarded orfs as possible alien genes when their GC3 values differed more than 1·5 SD from the average of all orfs longer than 150 bp, which corresponds to GC3<0·16 or GC3>0·29.

To compare the MW2 genome with those of other S aureus strains, we first processed homologous regions of the two chromosomes with bl2seq (blast for two sequences), and we visualised them for gross comparison. To do this, we divided a 2·8-Mb genome into three 1-Mb fragments, and blast homology was calculated for the nine pairs of fragments. Then, we processed blast results by a script written in Perl (http://www.perl.com) and visualised with Mathematica (Wolfram Research, Champaign, IL, USA). Similarity between the two virtual chromosomes—which are obtained by removal of all genomic islands and the mobile genetic elements IS1181, Tn554, and Tn5801 from the chromosomes—was assessed by AVID whole genome alignment as the proportion of matched nucleotides in the entire aligned nucleotides including gaps. To verify some of the ideas derived from the comparison, we compared the genome sequence of MW2 with those of NCTC8325, E-MRSA-16 (strain 252), and MRSA strain COL.

To assess spontaneous excision of genomic islands, LightCycler PCR was done with the LightCycler-FastStart DNA Master SYBR Green I kit (Roche Diagnostics, Mannheim, Germany) with specific primers designed in and around the genomic islands. To calculate frequency of the chromosome from which the island in question was precisely cut out, and to ascertain copy number of closed circular DNA formed by the excised island, we did quantitative PCR amplification, and then established the relative copy numbers of each PCR product with LightCycler software.

Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Strains used for comparison

N315 is a MRSA strain named after Nagasaki University Hospital, Japan, isolate number 315. Mu50 is a vancomycin-resistant MRSA (VRSAs) isolate from the pus of a postoperative wound. NCTC8325 is a registered strain of the National Collection of Type Cultures and has been widely used as laboratory S aureus strains. E-MRSA-16 (strain 252) is an epidemic MRSA from an outbreak in 1991–92 in the UK. The classification of number 16 was
established by phage-typing. MRSA COL was isolated in the UK in 1965.

**MW2 chromosome compared with N315**

Figure 1 shows overall nucleotide sequence similarity of MW2 (abscissa) and N315 (ordinate) chromosomes. The discontinuous regions of the central diagonal line of homology correspond with seven staphylococcal genomic islands that discriminate the main functional properties of the two genomes (see below; figures 1 and 2). Other minor discontinuities correspond with insertion of mobile elements in the N315 genome. The term genomic island is defined as a chromosomal region whose exogenous derivation is inferable. STAPHYLOCOCCAL CASSETTE CHROMOSOME meC (SCCMec; meC stands for meticillin resistance) and prophage are included, together with regions previously called pathogenicity islands. These islands carried genes associated with pathogenicity of S aureus. The reason for adoption of the term genomic island instead of pathogenicity island is that we are now finding new islands that are very similar to the previous ones, but no genes associated with pathogenicity are found on them (eg, type I vSa3 and type II vSa4). Well-established transposons and insertion sequences are excluded from genomic island terminology.

Figure 1 also shows many dots forming lattice-like patterns that cover the entire chromosomes of both MW2 and N315. This pattern indicates presence of multiple repeats that are well conserved between the two chromosomes. The core sequence 5'-TGTTGGG-GCCC-3' designated staphylococcus uniform repeats, is present in most lattice cross-sections. There are 60–85 copies of these repeats in the six S aureus chromosomes and 18 copies in a S epidermidis chromosome (based on the draft sequence obtained from The Institute for Genomic Research), but are not reported in virus, archaea, or eubacteria other than Staphylococcus spp.

Table 1 lists general information about the MW2 genome compared with those of N315 and Mu50. MW2 has six ribosomal RNA operons compared with five in N315 and Mu50 (figure 2). Transposons and insertion sequences are scarcely recorded in the MW2 genome (see also figure 2). A truncated copy of insertion sequence IS1272 is reported within SCCmec at a site downstream of...

---

**Figure 1:** Seven genomic islands as major differences between N315 and MW2 chromosomes

Genomic islands in MW2 chromosome=open circles. Yellow=SCCmec; pink=prophages; red=elements (genomic islands apart from SCC and prophages). Insertion of mobile elements in N315 genome=closed circle. Name of the allelic form of each genomic island in N315 and MW2 is given above and below the diagonal line, respectively. No allelic genomic island is recorded in N315 for vSa2 and vSa3.
Mu50 was 99.7%, whereas that between MW2 and N315 and Mu50 was 94.8% and 94.7%, respectively. Data suggest that N315 and Mu50 are closely related to each other, whereas MW2 is relatively distant from N315 and Mu50.

Genetic basis for non-multiresistance of MW2

Absence of multiresistance to non-/H9252-lactam antibiotics is a characteristic of community-acquired MRSA strains, a feature also shared by MW2. Consistent with this, a truncated copy of the mec regulator gene, mecr (part of the MEC OPERON), and the remains of a Tn552-related transposon carrying BLA OPERON (encoding penicillinase and its regulator function) were noted on plasmid pMW2. IS431 mec is the only apparently functional insertion sequence in the MW2 genome that is always present linked to mecA in SCCmec.

Evolutionary association of the three strains was assessed by alignment of their virtual chromosomes. Percentage of matched nucleotides between N315 and

Mu50 was 99.7%, whereas between MW2 and N315 and Mu50 was 94.8% and 94.7%, respectively. Data suggest that N315 and Mu50 are closely related to each other, whereas MW2 is relatively distant from N315 and Mu50.

Genetic basis for non-multiresistance of MW2

Absence of multiresistance to non β-lactam antibiotics is a characteristic of community-acquired MRSA strains, a feature also shared by MW2. Consistent with this
phenotype, the MW2 genome contains no antibiotic resistance genes apart from blaZ (encoding penicillinase) on plasmid pMW2 and mecA encoding penicillin-binding protein 2 (PBPs) found within SCCmec.

The structure of SCCmec in MW2 is quite distinct from that of hospital-acquired MRSA strains. Figure 3 shows the structure of SCCmec in MW2 (designated type IVa) compared with type I SCCmec (reported in hospital-acquired MRSA strains discovered in the 1960s, such as NCTC10442 and COL),17 and type II SCCmec (carried by hospital-acquired MRSA strains such as N315 and Mu50).4 The type IVa allelic form of SCCmec of MW2 is much smaller in size than the previously identified types in hospital-acquired MRSA strains. The boundary of SCCmec is defined by two functional direct repeats, at which excision of the element takes place. Type IVa SCCmec of MW2 comprises two allelic elements—class-B mec-gene complex (mecA and its regulatory genes) and type-2 ccrA and B genes.19 All SCCmec elements are integrated at exactly the same site, and yet each has distinct genetic organisation, presumably as a result of complex series of mutual crossovers. No other resistance genes, apart from mecA, were recorded on MW2 SCCmec.

It is noteworthy that seh, encoding staphylococcal enterotoxin H, and a truncated copy of see, encoding enterotoxin O, are present next to SCCmec (figure 3). The seh gene exists beyond the right boundary of SCCmec, linked to the second copy of the left extremity region of SCC (containing the unique gene mecB encoding a putative restriction enzyme). These genes are only present in MW2: they were not recorded in any of the other five complete genomes. Work in animals and clinical observation suggest that Staphylococcal enterotoxin H is involved in an acute toxic-shock-like syndrome.21 This enterotoxin is also reported to be produced in disproportionately high amounts compared with other superantigens.21 The seh gene could be associated with a strong promoter, or its location near the origin of replication, at which more gene copy numbers per cell are expected because of the uncoupled nature of cell division and chromosome replication, could be important.

### Table 1: General features of NM2 genome compared with those of N315 and Mu50

<table>
<thead>
<tr>
<th>Phenotypes</th>
<th>MW2</th>
<th>N315*</th>
<th>Mu50*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td>28,862</td>
<td>28,816</td>
<td>28,804</td>
</tr>
<tr>
<td>G+C content</td>
<td>28.4%</td>
<td>28.7%</td>
<td>28.9%</td>
</tr>
<tr>
<td>Length of sequence (bp)</td>
<td>20,654</td>
<td>20,653</td>
<td>20,107</td>
</tr>
<tr>
<td>Protein coding regions</td>
<td>28.3%</td>
<td>28.7%</td>
<td>28.9%</td>
</tr>
<tr>
<td>Non-coding region</td>
<td>75.6%</td>
<td>78.2%</td>
<td>80.1%</td>
</tr>
<tr>
<td>Transfer RNAs</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Protein coding regions</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ribosomal RNAs</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Transposons</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Domains</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Functional direct repeats</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Protein coding regions</td>
<td>27</td>
<td>29</td>
<td>34</td>
</tr>
</tbody>
</table>

Provenance of features in bold. *Previously designated pathogenicity islands.

### Figure 3: Comparison of three allelic forms of SCCmec, in MW2, N315, and NCTC10442 strains

Arrows with gaps indicate frameshifted orfs.
formally negates involvement of TSST-1 (toxic shock syndrome toxin-1) in high virulence of MW2, which is reported to be the case with an exceptionally virulent hospital-acquired MRSA strain.\(^{26}\)

Islands \(\psi SA3\) and \(\psi SA4\) are spontaneously excised from chromosomes of all the strains tested and form extrachromosomal closed circular DNA. \(\psi SA3mw\) is spontaneously excised from MW2 chromosome at a frequency of 0-0044, but it exists at a 17 times greater copy number than the chromosome from which it was excised, indicating the replication capability of the excised island. Frequency of spontaneous excision of \(\psi SA4mw\) from MW2 chromosome is much lower, which could be because the integrase gene of the \(\psi SA4mw\) is framshifted by a point mutation (table 2).

The two other islands, \(\psi SAo\) and \(\psi SAB\), are reported in all six \(S\ aureus\) strains and all clinical strains so far tested (K Hiramatsu, unpublished). \(\psi SAo\) is distinctive because it carries many putative staphylococcal exotoxin (set) genes (figure 4).\(^{27,28}\) However, none of 11 set genes of MW2 encodes protein of identical aminoacid sequence to that encoded by 10 set genes of N315 (figure 4, A). Maximum aminoacid identity noted among these 21 putative exotoxins was 98%. Pairwise comparison of orthologous set gene products indicated that the difference comprised many non-synonymous mutations accumulated in the T-cell receptor-binding domain of the exotoxins.\(^{29}\)

Structure of \(\psi SAB\) also differs greatly from strain to strain. The superantigen gene cluster (composed of five enterotoxin genes) carried by islands of N315 and Mu50 (type I) are missing from \(\psi SA3m\) and \(\psi SA4m\) (table 2). Instead, \(\psi SA3mw\) has a novel gene cluster, designated \(bsa\) (bacteriocin of \(S\ aureus\)), which encodes a putative bacteriocin (toxin or antibiotic to other bacteria) and its modification factors, which are similar to those reported in \(S\ epidermidis\)\(^{29}\) (figure 4, B). The type II allelic form of the island is probably important for MW2 to compete with other natural flora for successful colonisation in man.

### Genomic Islands of \(S\ aureus\)

The two island families, \(\psi SAo\) and \(\psi SAB\), both have a broken transposase gene and a \text{RESTRICION-MODIFICATION SYSTEM} composed of \(hsdS\) and \(hsdM\) (\(hsd\) stands for host specificity determinant.) The broken transposase gene feature is consistent with the fact that these two islands are not spontaneously cut out from the chromosome compared with \(\psi SA3\) and \(\psi SA4\). The common restriction-modification system seems to contribute to stabilisation of islands in the \(S\ aureus\) genome.\(^{4}\) In support of this view, \(HsdM\) methylases encoded by these islands are very well conserved across the six \(S\ aureus\) genomes, the aminoacid identity of which range from 99% to 100%. However, the \(HsdS\) protein that establishes target specificity of methylation, which is encoded by \(hsd\) located next to \(hSD\) on each island, exists in three allelic forms whose mutual aminoacid identities are smaller than 66% (figure 4, A and B; table 2). Their aminoacid differences are mainly localised in the putative nucleotide-sequence-recognition regions.\(^{26}\) For each island family, it is noted that mutual association of the three \(hSD\) alleles is well correlated with the mutual structural similarity of the three islands on which they are located. Therefore, \(\psi SAo\) and \(\psi SAB\) themselves exist in three allelic forms with their \(hSD\) genes as landmarks. \(\psi SAo\) and \(\psi SAB\) arise in different allelic combinations in each of the six \(S\ aureus\) chromosomes. N315 and Mu50 carry type I \(\psi SAo\) and type I \(\psi SAB\), NCTC8325 and COL carry type I \(\psi SA0\) and type II \(\psi SAB\), whereas MW2 carries type II \(\psi SA0\) and type II \(\psi SAB\) (table 2). E-MRSA-16 (strain 252) has different allelic forms for both islands (type III). This distribution of allelic forms of island is consistent with the view that the two islands were once mobile genetic elements, and were independently acquired by intraspecies genetic transfer between \(S\ aureus\) strains. In support of this hypothesis is the fact that every island of this family, apart from type II \(\psi SA0\) of MW2, has a remnant of transposase gene (figure 4, A and B), and notably, the allelotype of this gene also correlates with the allelotype of the island on which it is located.

### Virulence Genes in Prophages

All the sequenced \(S\ aureus\) strains, except for COL, have prophage \(\psi SA3\). Two new enterotoxin gene alleles, \(seg2\) and \(sek2\), are reported in \(\psi SA3mw\), which encodes putative enterotoxin G and K homologues. These toxin genes are
not carried by any \( \phi Sa3 \) prophages found in comparator genomes. Alignment of some of these prophages highlights their mosaic structures that have been generated by multiple crossovers, which is reminiscent of the case of SCCmec (figure 4, C). This finding shows clearly that prophages also exist in allelic forms.

The other prophage on MW2 chromosome, \( \phi Sa2mw \), carries genes \( \text{lukF-PV} \) and \( \text{lukS-PV} \), which encode PANTON-VALENTINE LEUKOCIDIN components, which have a potent toxic effect on human white-blood cells. The toxin is strongly associated with severe forms of pneumonia (necrotic pneumonia) caused by community-acquired \( S \) \( \text{aureus} \) strains. The \( \phi Sa2 \) attachment site is present in all six chromosomes, and three strains MW2, NCTC8325, and E-MRSA-16 (strain 252) do have integrated prophages. However, again they exist in
different allelic forms, and only \(\delta S\alpha 2m\) has Panton-Valentine leukocidin genes.

**Virulence genes of MW2 found outside genomic islands**

Collagen-adhesin protein (CNA)\(^2\) is uniquely encoded in the MW2 genome near the origin of replication (orf MW2612, indicated by green arrow in figure 2). This protein is not implicated in pathogenesis of osteomyelitis and septic arthritis.\(^2\) Small direct repeat sequences are found around cna, but there is no evidence that the gene is a part of a mobile genetic element.

**Discussion**

Transposons and insertion sequences can integrate themselves into any chromosome loci by illegitimate recombination.\(^5\) Thus, they tend to shuffle genome structure and are thought to contribute much to adaptability of *S aureus* to the adverse environment. From this viewpoint, it is curious to note that MW2 has scarcely any transposons or insertion sequences, whereas N315 and Mu50 have many (table 1).

It would be reasonable to assume that hospitals are a severe environment for microorganisms to survive in, because they are constantly exposed to various antiseptics and new antibiotics. Multiple insertions of transposons and insertion sequences in hospital-acquired MRSA genomes might be testament to the evolutionary ordeal they have gone through. In support of this view, we have previously shown that the copy number of insertion sequences of an MRSA strain increases in time during its propagation within a hospital.\(^4\) By contrast, MW2 is thought to represent a successful clone of *S aureus* in the community. The bacteriocin operon in MW2 might be indispensable for community-acquired *S aureus* strains; in this environment, competition with many other bacterial species for colonisation serves as a big selective pressure for survival.

Growth rate is also important. In fact, MW2 grows much faster than hospital-acquired MRSA strains, as doubling time shows, which—in drug-free Mueller-Hinton broth—is \(23\cdot 5, 34\cdot 8, \text{ and } 46\cdot 8\) min for MW2, N315, and Mu50, respectively. As an already successful part of the natural flora of healthy human beings, MW2 seems to establish an additional advantage by acquisition of a simple and less encumbering allelic version of SCC\textit{mec} (type IVa) to survive sporadic exposure to limited varieties of antibiotics prescribed in the outpatients’ clinic.

MW2 has a total of 18 toxins, which are not recorded in any of the five comparative *S aureus* genomes (N315, Mu50, E-MRSA-16 [strain 252], NCTC8325, and COL). The Panton-Valentine leukocidin components recorded on \(\delta S\alpha 2m\) should at least partly contribute to virulence of MW2, since results of an epidemiological study showed a close association of Panton-Valentine leukocidin with nectocic pneumonia,\(^3\) a symptom that is seen in patients infected with MW2 and related strains.\(^1\)

The staphylococcal enterotoxin H, which is reported to have the highest binding affinity to major histocompatibility complex class II molecules among staphylococcal enterotoxins, is also uniquely reported in MW2 genome.\(^1\) Besides this superantigen, MW2 has 15 unique superantigen genes: 11 exotoxin (\(se\)) genes on \(\sigma S\alpha\) island and four enterotoxin genes on \(\sigma S\alpha\) and \(\delta S\alpha3\) genomic islands. These putative superantigens have at least \(2\%\) aminoacid differences compared with their extant homologues, and together with staphylococcal enterotoxin H, constitute an unique allelic superantigen range in MW2. These superantigens are expected to trigger human T cells in a quite distinct fashion from those of other *S aureus* strains. The culture supernant of MW2 contained ten times stronger human T cell proliferation activity than N315, and its activity is directed towards distinct T-cell subsets compared with those activated by N315 (T Uchiyama, unpublished observation). This strong stimulation of distinct T-cell subsets might be correlated with high virulence of MW2.

Some of the virulence genes on MW2 probably contribute to its pathogenicity in healthy human hosts, but it is not clear which of these virulence factors is most important. For example, although work in animals shows that MW2 has an LD\(_50\), that is five times less than N315, results of gene knockout studies show that no single gene, such as \(lukS-PV, lukF-PV, cna, \text{ or seh}\), accounts for this difference (T Baba, unpublished). Therefore, it is likely that the total pathogenic potential of MW2 is a result of synthetic contribution of many virulence genes listed in this study (table 2, figure 2). Moreover, although some genes (eg, cna) arise singly in the chromosome, most virulence genes tend to exist as a set of allelic genes carried by genomic islands. This fact was first noticed from scrutiny of \(\sigma S\alpha\), in which the \(se\) gene cluster encoded a totally different range of superantigens between N315 and MW2. This finding strongly suggests that the pathogenic potential of *S aureus* is established first, by the genomic island families it is part of, and second, by whatever allelic form of each island family it takes. Since most of the genomic islands are potentially mobile, pathogenic potential of each strain of *S aureus* is not an intrinsic trait and is likely to change as a result of interspecies exchange of virulence genes via lateral gene transfer.

On the basis of our findings, we propose that GEMENIC ISLAND ALLOTYPING is a novel approach to *S aureus* genome typing. This process allows us to predict the pathogenic capability of an *S aureus* clinical strain, and would even help us predict the symptom, severity, and prognosis of the illness that it causes. Genomic island allotyping, done either by PCR, microarray hybridisation, or nucleotide sequence determination of certain regions of genomic islands, would add medically relevant information to extant genotypings, such as pulsed-field gel electrophoresis or multilocus sequence typing,\(^1\) which indicate only the physical genomic information of *S aureus* chromosome.

**Contributors**

All investigators contributed to design of the study. K Hiramatsu was responsible for overall experimental design and interpretation of data. T Baba did comparative genome analysis and interpretation of structures of the pathogenicity islands and prophages of MW2. F Takeuchi and K Yamanoto did computer-directed analysis on MW2 genome and its comparison with other strains. M Kuroda constructed the genomic library and annotated the MW2 genome. H Yuzawa analysed structural differences of SCC\textit{mec} between MW2 and other strains. A Oguchi, Y Nagai, and N Iwama established the MW2 genomic sequence. K Aoki and K Asano assembled contigs and did the BLAST search. H Kuroda-Murakami studied mobile capability of genomic islands by PCR with a Light Cycler. T Nami did clinical and epidemiological characterisation of MW2 and associated mid-western community-acquired MRSA strains. L Cui analysed allelic genes of MW2 and comparator strains.

**Conflict of interest statement**

None declared.

**Acknowledgments**

We thank Hisashi Kikuchi, chairman of National Institute of Technology and Evaluation, for his support; Chisana Temastorn, Rika Nishiko and Satomi Suzuki for technical assistance; Susan Johnson, Kathy LeDell, Ruth Lynfield, Richard Danila, and Michael Osterholm (Minnesota Department of Health), Larry Shirley (North Dakota Department of Health), and Fred C Tenover (Centers for Disease Control and Prevention, Atlanta, USA) for clinical and epidemiological characterisation of MW2 strain; Hiroyuki Takase and Ken-ichi Sato (Daichi Pharmaceutical Co, Japan) for virulence evaluation of MW2 in animal infection model; and Yuh Morimoto for help in preparation of computer graphics.
This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (13226141) from the Ministry of Education, Science, Sports, Culture, and Technology of Japan; and the Core University System Exchange Programme under the Japan Society for the Promotion of Science, coordinated by the University of Tokyo Graduate School of Medicine and Mahidol University. The study was also partly supported by a grant for the National Health Cooperation Research (11G-4) from the Ministry of Health and Welfare, Tokyo, Japan, and by the Research for the Future Program of the Japan Society for the Promotion of Science.

References


