MAJOR ARTICLE

# Roles of 34 Virulence Genes in the Evolution of Hospital- and Community-Associated Strains of Methicillin-Resistant *Staphylococcus aureus*

**Binh An Diep,[1,2] Heather A. Carleton,[1] Richard F. Chang,[3] George F. Sensabaugh,[2] and Françoise Perdreau-Remington[1]**

[1]Division of Infectious Diseases, Department of Medicine, University of California, San Francisco, [2]Program in Infectious Diseases and Immunity, School of Public Health, and [3]Undergraduate Program in Molecular Cell Biology, College of Letters and Science, University of California, Berkeley

***Background.*** The extent to which the horizontal transfer of virulence genes has contributed to the emergence of contemporary virulent strains of methicillin-resistant *Staphylococcus aureus* (MRSA) in hospital and community settings is poorly understood.

***Methods.*** Epidemiologically well-characterized MRSA isolates collected over 8.5 years were genotyped and tested for the presence of 34 virulence genes.

***Results.*** Six strain types accounted for 88.2% of all MRSA infections. The evolution of contemporary hospital and community phenotypes within the CC8 and CC30 lineages—2 background genomes that produced historical pandemic MRSA clones—were associated with multiple horizontal acquisitions of virulence genes. The epidemic community phenotype of a CC8 strain, designated ST8:USA300, was linked to the acquisition of staphylococcal cassette chromosome (SCC)*mec* type IV, the genes for Panton-Valentine leukocidin (PVL), and the enterotoxin Q and K genes. Similarly, the epidemic community phenotype of a CC30 strain, ST30:USA1100, was linked to the acquisition of SCC*mec* type IV and the *pvl* genes. In contrast, the epidemic hospital phenotype of another CC30 strain, ST36:USA200, was associated with the acquisition of SCC*mec* type II, the enterotoxin A gene, and the toxic shock syndrome toxin 1 gene. The *pvl* genes appear not to be essential for the evolution of other community-associated strains of MRSA, including ST8:USA500 and ST59:USA1000.

***Conclusions.*** The horizontal transfer of virulence genes, although infrequent, is epidemiologically associated with the emergence of new virulent strains of MRSA.

*Staphylococcus aureus* causes a variety of infections and toxinoses, ranging in severity from cutaneous abscesses to life-threatening necrotizing fasciitis and necrotizing pneumonia [1, 2]. The introduction of penicillin and β-lactamase–stable penicillins, although dramatically improving the management of staphylococcal infection, have also contributed to the emergence of methicillin-resistant *S. aureus* (MRSA) strains. The original MRSA strains have become established in hospitals worldwide and are classified by multilocus sequence typing (MLST) as belonging to 5 distinct clonal complexes, CC5, CC8, CC22, CC30, and CC45 [3, 4]. Recently, other MRSA strains have become established in community settings [5]. The 2 community-associated MRSA (CA-MRSA) lineages with the greatest medical impact belong to CC8 and CC30 [6, 7]—the 2 background genomes that produced the very first MRSA strain [8] and the second most common MRSA strain found in hospitals in the United Kingdom and the United States [9]. Other CA-MRSA strains belong to CC1, CC59, and CC80 [5]. It is not clear why certain MRSA strains predominate in hospitals and other MRSA strains predominate in community settings.

Also not clear are the epidemiological differences among different strains of CA-MRSA originating from the same clonal complex. For example, we have shown that 2 local CA-MRSA descendants of the CC8 lineage,

Received 28 September 2005; accepted 27 December 2005; electronically published 21 April 2006.

Presented in part: International Staphylococci and Staphylococcal Infections Meeting, 24–27 October 2004, Charleston, South Carolina.

Potential conflicts of interest: F.P.-R. is a consultant for Pfizer Corporation.

Financial support: Pfizer Corporation (unrestricted grant to F.P.-R.); Eugene Cota-Robles Predoctoral Fellowship and a Genetics Training Grant (5-T32-GM07127-29-30), University of California, Berkeley, to B.A.D.

Reprints or correspondence: Dr. Françoise Perdreau-Remington, University of California, San Francisco, Div. of Infectious Diseases, San Francisco General Hospital, 1001 Potrero Ave. UCSF 1372, Bldg. 100, Rm. 301, San Francisco, CA 94110 (fpr@epi-center.ucsf.edu).

**The Journal of Infectious Diseases 2006; 193:1495–1503**



0022-1899/2006/19311-0004$15.00

ST8:USA500 and ST8:USA300, cause very different types of disease in community settings [6]. Unlike ST8:USA500, which causes sporadic disease, ST8:USA300 is epidemic in community populations in at least 16 US states [6, 10–12] and has been linked to clusters of necrotizing fasciitis [1], necrotizing pneumonia [13], and other rapidly progressing and life-threatening staphylococcal infections [14]. The molecular genetic basis of the high transmissibility and hypervirulence of ST8:USA300, which far exceeds those of ST8:USA500 and archaic CC8 strains, is not known.

The objective of the present study was to gain new insight into the epidemiological and molecular factors that could explain the differences between hospital-associated MRSA (HA-MRSA) and CA-MRSA strains as well as the differences between CA-MRSA lineages exhibiting distinct epidemiological characteristics. We hypothesized that variations in virulence-gene content and allelic diversity contribute to MRSA lineage diversification and the evolution of distinct epidemiological characteristics. To test this hypothesis, we characterized the virulence-gene content of a representative sample of MRSA isolates collected from inpatients and outpatients over 8.5 years of laboratory-based surveillance in San Francisco. The study revealed the vital contribution that virulence determinants have made to the evolution of contemporary epidemic CA-MRSA strains.

## METHODS

***Bacterial strains.*** Laboratory-based surveillance for the Community Health Network of San Francisco identified 8382 *S. aureus* isolates from unique patient clinical specimens between January 1996 and June 2004. Of the 3713 MRSA isolates, we selected 715 (19.3%) for genotyping. For 1996–2002, 70 isolates were randomly selected each year [6]; for 2003, 150 isolates were randomly selected; and for January–June 2004, 75 isolates were randomly selected. From these isolates, we selected a subset of 132 (18.5%) for virulence-gene profiling and staphylococcal cassette chromosome (SCC)*mec* allotyping. The isolates were selected to represent a broad diversity of pulsed-field gel electrophoresis (PFGE) subtypes, sites of infection, and year of collection. CA-MRSA was defined as an isolate that was obtained either from an outpatient or from an inpatient ⩽72 h after hospital admission; HA-MRSA was defined as an isolate that was obtained from an inpatient >72 h after hospital admission [6].

Control isolates for polymerase chain reaction (PCR)–based assays were obtained from the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA) and included the 7 reference strains with fully sequenced genomes—N315, MU50, MW2, methicillin-susceptible *S. aureus* (MSSA) 476, MRSA252, COL, and NCTC8325. Other isolates from NARSA were used as positive controls for virulence genes not present in the reference genome strains, including NRS171 (*sed* and *sej*), NRS166 (*agr4, eta,* and *etb*), and NRS227 (*sel* and *sep*).

***Genotyping.*** PFGE with *Sma*I [15] and multilocus restriction-fragment typing (MLRFT) [16] were used to characterize all *S. aureus* isolates. A PFGE group was considered to consist of isolates with differences of no more than 6 bands from all other isolates in the group [17]. The PFGE genotype nomenclature used was in accordance with that of McDougal et al. [10]. MLRFT, which assesses restriction-site variation in 7 housekeeping gene loci, provided a low-cost method for grouping distantly related PFGE subtypes [16]. MLST was performed for representative isolates from each genotype as defined by PFGE and MLRFT [18]. To provide a simplified, uniform nomenclature for describing MRSA, strains are designated by both sequence type and PFGE genotype (e.g., ST8:USA300). Fisher's exact test was used to test for significant associations between strain types and sites of infection as well as between strain types and other categorical variables [19].

***PCR-based assays for virulence genes and SCC*mec *allotypes.*** PCR-based assays for *sea, seb, sec, sed, see, seg, seh, sei, sej, sen, seo, sem, tst, eta, etb, lukM, hla, hlb, hld,* and *hlg* were performed using the primers described by Jarraud et al. [20], and assays for the *pvl* genes were performed using the primers described by Lina et al. [21]. The primers used for assays for adhesin genes (*cna, clfA, clfB, fnbA,* and *icaA*) were those described by Peacock et al. [22], and the primers used for assays for the capsular polysaccharide type 5 and type 8 genes (*cap5* and *cap8*) were those described by Moore and Lindsay [23]. Multiplex PCR–based protocols for allotyping of *agr* classes I–IV [24] and SCC*mec* types I–IV [25] were performed as described elsewhere. Table 1 lists our in-house primers for detection of *bsaA, lukE,*

**Table 1. Oligonucleotide primers designed for the present study.**

| Gene (GenBank accession no.), primer | Primer sequence (5′→3′) |
|---|---|
| *bsaA* (BA000033) | |
| bsaA-F | ACAGAAGCTGTTAAAACTACCC |
| bsaA-R | GATTAATATGACAATTGAAGTGGGTC |
| *lukE* (BA000033) | |
| lukE-F | GCAACTTTGTCAGTAGGACTG |
| lukE-R | GTCTACTTCACTGACATAACTC |
| *sek* (AF410775) | |
| sek-F | GGTGTCTCTAATAGTGCCAG |
| sek-R | TCGTTAGTAGCTGTGACTCC |
| *sel* (BA000033) | |
| sel-F | ATCAATGGCAAGCATCAAACAG |
| sel-R | TGGAAGACCGTATCCTGTG |
| *sep* (BA000018) | |
| sep-F | GACCTTGGTTCAAAAGACACC |
| sep-R | TGTCTTGACTGAAGGTCTAGC |
| *seq* (AF410775) | |
| seq-F | TCTAGCATATGCTGATGTAGG |
| seq-R | CAATCTCTTGAGCAGTTAC(C/T)TC |

**NOTE.** F, forward; R, reverse.

*sek, sel, sep,* and *seq*. All 7 reference genome strains were included as controls in each PCR experiment. Where appropriate, other positive control isolates (described above) were included for virulence genes not present in the reference genome strains.

## RESULTS

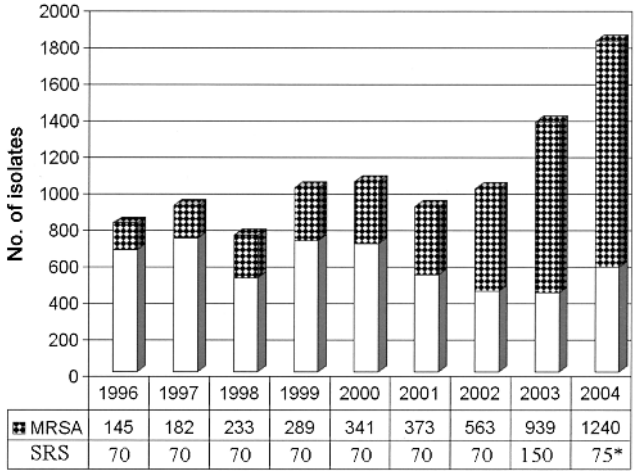
***Overview of epidemiological surveillance.*** The prevalence of MRSA strains from unique patients infected with *S. aureus* increased from 17.8% (145/815) in 1996 to 68.4% (1240/1813) in 2004 (figure 1). Overall, there were 3713 MRSA infections, with 42.0% of the cases occurring during the last 18 months of the collection period, ending in June 2004. To assess the population structure of HA-MRSA and CA-MRSA strains, we selected a nonproportional stratified random sample of 715 isolates for genotyping (figure 1). Six major strain types—as defined by MLST, MLRFT, and PFGE—accounted for 88.2% of the MRSA isolates (table 2). The remaining isolates belonged to 22 minor strain types, with each accounting for no more than 2% of the total. Two major strain types that had been previously identified as HA-MRSA, ST5:USA100 and ST36:USA200, were cultured predominantly (72.8% and 72.5%, respectively) from inpatients >72 h after hospital admission [6]. The remaining 4 major strain types—ST30:USA1100, ST59:USA1000, ST8:USA500, and ST8:USA300—have been classified as CA-MRSA and were cultured predominantly (range, 63.6%–87.9%) from outpatients or from inpatients ≤72 h after hospital admission. The CA-MRSA strains have been shown previously to differ in epidemiological character: ST8:USA500 and ST59:USA1000 exhibited a sporadic phenotype, ST30:USA1100 exhibited a localized outbreak phenotype, and ST8:USA300 exhibited an epidemic phenotype [12].

Table 2 summarizes the estimated number of isolates exhibiting specific disease characteristics. CA-MRSA strains accounted for a disproportionate number (1687/2053 [82.2%]) of abscess/wound infections (table 2). Sterile site infections due to CA-MRSA strains were 2-fold more common than were sterile site infections due to HA-MRSA strains. Nonetheless, the proportion of CA-MRSA strains associated with sterile site infections was 11.4% (284/2486), compared with 13.3% (116/872) for HA-MRSA strains ($P = .14$), suggesting that there is no difference in the capacity to cause severe disease between CA-MRSA and HA-MRSA strains. Importantly, HA-MRSA strains were more likely to be multidrug resistant than were CA-MRSA strains ($P < .001$) (table 2).

***Diversity of chromosomal PFGE profiles.*** Table 3 summarizes the diversity of PFGE subtypes in the nonproportional stratified random sample and in the subset of isolates subjected to virulence-gene profiling. Each MRSA strain type consisted of isolates with closely related PFGE profiles that differed by no more than 6 bands. For 5 of the 6 major strain types, a single prevalent PFGE subtype accounted for at least 80% of the isolates. The ST30:USA1100 isolates, however, had 2 prevalent subtypes, each accounting for >45% of the isolates. The number of distinct PFGE subtypes varied for each strain type, ranging from 6 subtypes among the 38 ST36:USA200 isolates to 34 subtypes among the 161 ST5:USA100 isolates (table 3). Of particular note, 20 distinct PFGE subtypes were observed among the 188 isolates belonging to ST8:USA300. Between September 2001 and December 2002, only 1 PFGE subtype, identical to USA300-0114 [11], was recovered from patients infected with CA-MRSA. The other 19 minor PFGE subtypes of ST8:USA300 were identified subsequently in 2003–2004. These data suggest that a single prevalent PFGE subclone was introduced into the area, increased in frequency in the population, and, through many infectious cycles, diversified into a cluster of closely related subtypes, all of which had descended from the founding USA300-0114 subtype.

***Comparison of HA-MRSA and CA-MRSA gene content.*** To assess variation in virulence-gene content among HA-MRSA and CA-MRSA strain types, we screened for the presence of 34 virulence genes and SCC*mec* allotypes in 132 isolates from the 6 major strain types (table 3). The isolates were selected to represent a broad temporal distribution of PFGE subtypes and sites of infection. This analysis revealed several major findings.

Each strain type is characterized by a distinct consensus repertoire of virulence genes (figure 2). Although the isolates surveyed were selected to represent a broad diversity in PFGE profiles and sites of infection, relatively few deviated from the consensus profile. When isolates did differ from their strain-type consensus—for example, 4 of the 29 ST5:USA100 isolates carried the




| | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 |
|---|---|---|---|---|---|---|---|---|---|
| MRSA | 145 | 182 | 233 | 289 | 341 | 373 | 563 | 939 | 1240 |
| SRS | 70 | 70 | 70 | 70 | 70 | 70 | 70 | 150 | 75* |

**Figure 1.** Secular trends of methicillin-susceptible *(white bars)* and methicillin-resistant *(shaded bars) Staphylococcus aureus* (MRSA) infections in unique patients, identified during laboratory-based surveillance for the Community Health Network of San Francisco. The total no. of MRSA isolates is listed for each year; we selected a nonproportional stratified random sample (SRS) of MRSA isolates from each year for electronic medical record review and genotyping. For January–June 2004, we randomly selected 75 isolates (indicated by the asterisk).

**Table 2. Characteristics of community- and hospital-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA and HA-MRSA, respectively) clinical isolates from the Community Health Network of San Francisco obtained between January 1996 and June 2004.**

| Category | HA-MRSA strain types | | CA-MRSA strain types | | | | Low-frequency strains[a] |
|---|---|---|---|---|---|---|---|
| | ST5:USA100 | ST36:USA200 | ST30:USA1100 | ST59:USA1000 | ST8:USA500 | ST8:USA300 | |
| SFGH nomenclature[b] | ST5:D | ST36:A | ST30:Z | ST59:P | ST8:C | ST8:S | … |
| SCC*mec* type | II | II | IVa | IVa | IVa | IVa | … |
| Stratified random sample ($n$ = 715), no. of isolates[c] | 161 | 38 | 93 | 68 | 80 | 188 | 86 |
| Estimated no. (%) of isolates[d] ($n$ = 3713) | | | | | | | |
| Total | 734 | 138 | 438 | 360 | 352 | 1336 | 353 |
| Time of MRSA culture | | | | | | | |
| >72 h after hospital admission | 536 (73.0) | 100 (72.5) | 53 (12.1) | 106 (29.4) | 128 (36.4) | 161 (12.1) | 137 (38.8) |
| ⩽72 h after hospital admission[e] | 198 (27.0) | 38 (27.5) | 385 (87.9) | 254 (70.6) | 224 (63.6) | 1175 (87.9) | 216 (61.2) |
| Site of infection | | | | | | | |
| Abscess/wound | 183 (24.9) | 24 (17.4) | 350 (79.9) | 177 (49.2) | 137 (38.9) | 1023 (76.6) | 159 (45.0) |
| Sterile site[f] | 85 (11.6) | 31 (22.5) | 21 (4.8) | 57 (15.8) | 49 (13.9) | 157 (11.8) | 43 (12.2) |
| Respiratory tract | 226 (30.8) | 42 (30.4) | 23 (5.3) | 79 (21.9) | 70 (19.9) | 79 (5.9) | 84 (23.8) |
| Urinary tract | 207 (28.2) | 29 (21.0) | 13 (3.0) | 37 (10.3) | 56 (15.9) | 50 (3.7) | 42 (11.9) |
| Other[g] | 33 (4.5) | 12 (8.7) | 31 (7.1) | 10 (2.8) | 40 (11.4) | 27 (2.0) | 25 (7.1) |
| Antimicrobial resistance, % of isolates | | | | | | | |
| Gentamicin | 39.8 | 10.5 | 1.1 | 1.5 | 6.3 | 0.0 | 11.7 |
| Erythromycin | 96.3 | 97.4 | 7.5 | 72.1 | 78.8 | 89.4 | 65.0 |
| Clindamycin | 86.3 | 92.1 | 3.2 | 11.8 | 11.3 | 11.7 | 31.7 |
| Ciprofloxacin | 93.2 | 97.4 | 3.2 | 4.4 | 28.8 | 62.8 | 38.3 |
| Tetracycline | 4.3 | 2.6 | 3.2 | 7.4 | 5.0 | 24.5 | 16.7 |
| Trimethoprim-sulfamethoxazole | 8.1 | 0.0 | 2.2 | 0.0 | 11.3 | 0.0 | 11.7 |

**NOTE.** Of the 3713 clinical isolates recovered from the Community Health Network during the 8.5-year period, a nonproportional stratified random sample ($n$ = 715) was selected for further characterization. For 1996–2002, we randomly selected 70 isolates per year; for 2003, we randomly selected 150 isolates; and for January–June 2004, we randomly selected 75 isolates. MLST, multilocus sequence typing; PFGE, pulsed-field gel electrophoresis; SCC, staphylococcal cassette chromosome.

[a] All other MRSA strains, with each type accounting for no more than 2% of the total.

[b] PFGE strain type nomenclature used in previous studies conducted at San Francisco General Hospital (SFGH) [6, 12, 26–28] corresponding to the MLST:PFGE nomenclature used in the present study. Corresponding *spa* types associated with the 6 major MRSA strain types are as follows: TJMGMK or TJMBMDMBMK for ST5:USA100; WGKAKAOMQQQ for ST36:USA200 and ST30:USA1100; ZDGDGDEB for ST59:USA1000; YHGFMBQBLO or YHGCMBQBLO for ST8:USA500; and YHGFMBQBLO or YGFMBQBLO for ST8:USA300 [16].

[c] Shown are the actual no. of isolates of the stratified random sample for each of the 6 major MRSA strain types.

[d] Shown are the estimated no. of isolates obtained after appropriate weighting to restore the original proportions, to account for the nonproportional stratified random sampling design [6]. Percentages shown are for the total no. of isolates belonging to each strain type.

[e] Or from outpatients.

[f] Blood, cerebrospinal fluid, and bone.

[g] All other clinical sites, such as fluids of unspecified origin, ear, gallbladder, etc.

**Table 3. Diversity of pulsed-field gel electrophoresis (PFGE) subtypes within each major community- and hospital-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA and HA-MRSA, respectively) strain type.**

| | HA-MRSA strain types | | CA-MRSA strain types | | | |
|---|---|---|---|---|---|---|
| Category | ST5:USA100 | ST36:USA200 | ST30:USA1100 | ST59:USA1000 | ST8:USA500 | ST8:USA300 |
| PFGE diversity | | | | | | |
| Distinct subtypes | 34 | 6 | 12 | 13 | 10 | 20 |
| Prevalent subtypes[a] | 1 | 1 | 2 | 1 | 1 | 1 |
| Representative PFGE subtypes surveyed for virulence genes | | | | | | |
| Isolates | 29 | 16 | 32 | 16 | 20 | 19 |
| Unique subtypes | 12 | 6 | 9 | 8 | 7 | 7 |
| Subtypes demonstrating horizontal acquisition of virulence genes | 1 | 0 | 3 | 1 | 1 | 1 |

**NOTE.** Data are no. of isolates or subtypes.

[a] The prevalent PFGE subtype accounted for at least 80% of the isolates for each major MRSA strain type, except for ST30:USA1100, in which there were 2 prevalent subtypes, each accounting for >45% of the isolates.

enterotoxin D, J, and P genes—they were found to belong to the same PFGE subtype. Strain types belonging to the same clonal complexes—for example, ST8:USA500 and ST8:USA300 in CC8 and ST36:USA200 and ST30:USA1100 in CC30—possessed very similar virulence-gene repertoires. This suggests the possibility that these virulence-gene repertoires are coadapted to the background genome and, thus, contribute to the fundamental biological fitness of the lineages.

The only marker that differentiated the HA-MRSA and CA-MRSA strain types was the allotype of the SCC*mec* element; strains of the 2 HA-MRSA types carried the SCC*mec* type II element, whereas strains of the 4 CA-MRSA types carried the SCC*mec* type IV element. Among the CA-MRSA types, strains belonging to 2 of them, ST8:USA300 and ST30:USA1100, carried the genes for Panton-Valentine leukocidin (PVL), a bicomponent cytolytic toxin. Although the recent emergence of CA-MRSA has been associated with the carriage of the *pvl* genes in SCC*mec* type IV–bearing MRSA strains [5, 29], the fact that none of the ST8:USA500 and ST59:USA1000 CA-MRSA strains tested in the present study, and very few in other surveys [26], carried the *pvl* genes suggests that PVL is not essential for a community lifestyle. Nonetheless, this toxin may bestow an epidemic phenotype in the appropriate background genome—for example, *pvl*-harboring ST30:USA1100 has been implicated in an infection peak during 1998–2000, and, likewise, the epidemic expansion of *pvl*-harboring ST8:USA300 has been observed since 2002 [6].

None of the other virulence markers analyzed in this survey appeared to differentiate HA-MRSA from CA-MRSA or the epidemiologically distinct CA-MRSA from each other. For example, there was no obvious pattern to the distribution of the accessory gene regulator *(agr)* allotype among the different MRSA lineages; such a link might have been expected, given that this locus controls the expression of most of the virulence genes involved in pathogenesis [30]. Similarly, *cna* (which encodes collagen-binding protein) was present only in the 2 CC30 strains, HA-MRSA ST36:USA200 and CA-MRSA ST30:USA1100; this suggests that the *cna* gene itself contributes little to the differentiation of hospital and community phenotypes, although the possibility of differences in regulatory control cannot be discounted.

***Horizontal transfer of virulence genes and genomic diversity.*** The subset of 132 isolates selected for virulence-gene profiling included 49 distinct PFGE subtypes (table 3). Assuming 1 ancestral PFGE subtype for each strain type, a minimum of 43 genetic events are required to account for the generation of the 49 subtypes; possible genetic events giving rise to PFGE subtype variation include mutations resulting in the gain or loss of *Sma*I restriction sites and intergenomic transfers of large, mobile genetic elements—for example, prophages, pathogenicity islands, and plasmids [31]. The present study suggests, however, that whatever processes account for the generation of the PFGE variants observed here, the horizontal transfer of virulence genes is relatively infrequent; only 14.3% (7/49) of the PFGE subtypes included in our study exhibited a gain or loss of virulence genes (figure 2 and table 3). Not surprisingly, horizontal transfer of virulence genes involved only toxin genes, including the *pvl* genes, *tst,* and most of the enterotoxin genes (figure 2), all of which are known to be carried on mobile genetic elements [31]. Most of the toxin genes that did not demonstrate horizontal mobility belonged to the enterotoxin gene cluster (*seg, sei, sem, sen,* and *seo*), which is restricted to the CC5 and CC30 genomic background, and also included *seh,* which is restricted to the CC1 genomic background.

On the basis of comparisons of virulence-gene content, SCC*mec* allotypes, and MLST and PFGE markers, the lines of ancestry within CC8 and CC30 can be inferred as shown in figure 3, allowing some latitude for sequential arrows lacking intermediate isolates to show the exact order of events. The genetic events epidemiologically associated with the evolution

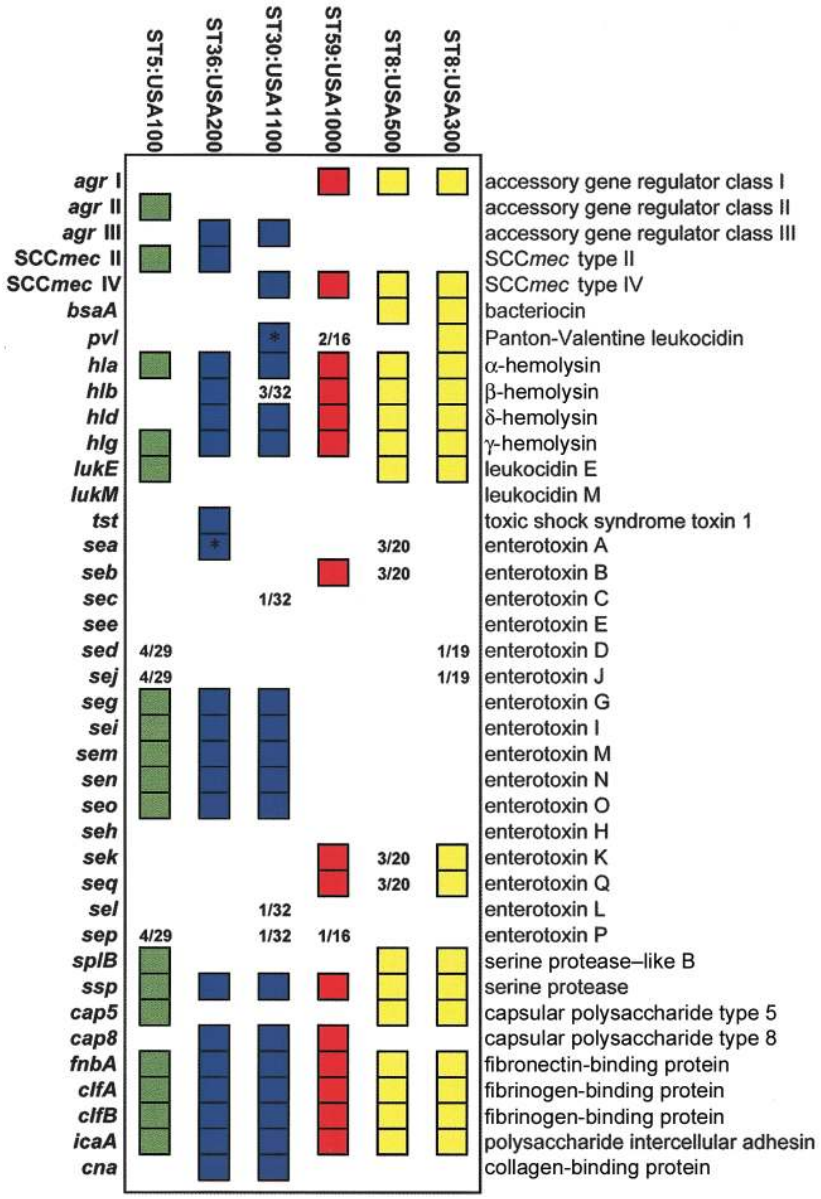



**Figure 2.** Distribution of 34 virulence determinants and staphylococcal cassette chromosome (SCC)*mec* allotypes in hospital-associated ($n = 45$) and community-associated ($n = 87$) methicillin-resistant *Staphylococcus aureus*. A square without an asterisk indicates that all isolates belonging to a given strain type tested positive for the virulence locus, and a square with an asterisk indicates that most isolates belonging to a given strain type tested positive for the virulence locus (15 of 16 ST36:USA200 isolates carried *sea,* and 29 of 32 ST30:USA1100 isolates carried *pvl*). The squares are colored according to the background genome, to illustrate the similarity in the consensus virulence-gene repertoire shared between ST36:USA200 and ST30:USA1100 (CC30 background) as well as between ST8:USA500 and ST8:USA300 (CC8 background).

of hospital or community phenotypes within these 2 clonal clusters are the horizontal acquisition of SCC*mec* type IV, the *pvl* genes, *seq*, and *sek* in the CC8 lineage and SCC*mec* type IV, the *pvl* genes, *sea*, and *tst* in the CC30 lineage.

## DISCUSSION

Detailed knowledge of gene content is a prerequisite to developing an understanding of the genetic basis underlying differences between HA-MRSA strains and CA-MRSA strains and the differences between CA-MRSA lineages exhibiting distinct epidemiological characteristics. Our characterization of virulence-gene profiles and SCC*mec* allotypes in epidemiologically well-characterized strain populations sheds light on the diversification of MRSA lineages and the emergence of new virulent CA-MRSA strains.

***MRSA strain diversification.*** The multiplicity of PFGE subtypes within each of the 6 MRSA lineages characterized in the present study indicates that strain diversification is an on-

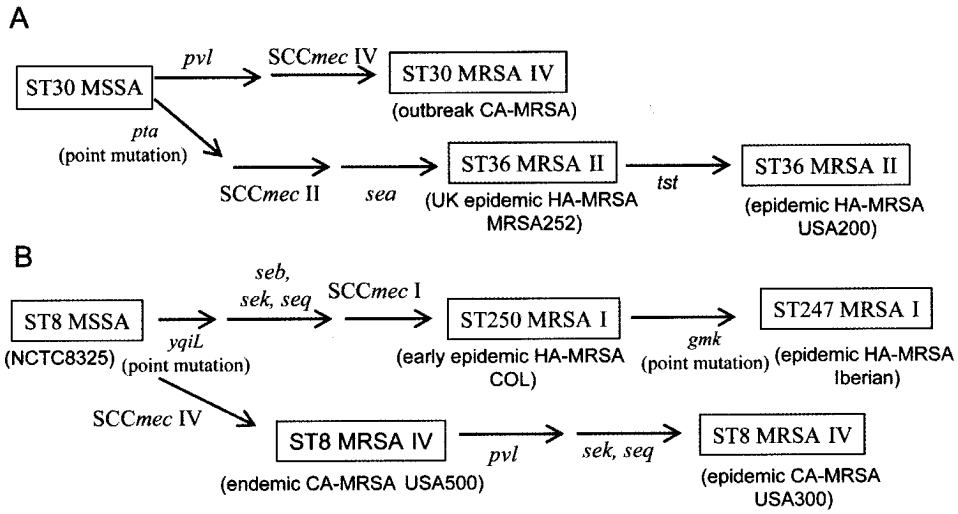



**Figure 3.** Proposed genetic events leading to the emergence of epidemic strains of community- and hospital-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA and HA-MRSA, respectively) in the CC30 *(A)* and CC8 *(B)* lineages. Strain diversification through the horizontal transfer of virulence genes and staphylococcal cassette chromosome (SCC)*mec* allotypes are indicated above the arrows. Virulence genes common across all CC8 and CC30 strains are illustrated in figure 2.

going process. Our study suggests, however, that there is infrequent concomitant change in virulence-gene profiles. Each of the 6 strain types was found to possess a characteristic consensus virulence-gene profile, with relatively few variants deviating from the consensus; moreover, the consensus profiles were very similar for closely related genomes—for example, ST30:USA1100 and ST36:USA200 in CC30 and the 2 strain types in ST8. Only 14.3% (7/49) of the PFGE subtypes chosen to represent the 6 MRSA lineages acquired virulence genes. That all acquired genes were toxin genes known to be carried on mobile genetic elements implicates horizontal gene transfer (figure 2 and table 3) [31]. The relatively low rate of horizontal transfer of virulence genes in our strain population stands in contrast to the results of other studies that have reported significant variation in virulence-gene content among genetically related isolates [20, 22, 32, 33]. This may be a consequence of the local epidemic population structure, which is characterized by nonoverlapping outbreaks of successful MRSA strain types. For example, during the 1998–2000 infection peak, 2 successful outbreak-associated PFGE subtypes increased in prevalence, predominated for a time period [26, 27], then decreased rapidly in number after 2002 as a result of the rapid ascendency of the next outbreak strain, ST8:USA300 [6, 12]. Such rapid expansion and contraction in population size likely obscured the background level of horizontal transfer of virulence genes observed in our strain population.

The pattern of gene transfer observed in the present study, coupled with existing historical and epidemiological data, allows a plausible reconstruction of the evolutionary lines of descent within CC30 and CC8 that accounts for some of the distinctive epidemiological properties exhibited by the 4 strain types characterized in this study as belonging to the 2 clonal clusters (figure 3).

***Evolutionary descent within the CC30 background genome.*** The CC30 lineage includes 2 contemporary strain types differing significantly in epidemiological character and sites of infection. The SCC*mec* type II–bearing ST36:USA200 strain was endemic in our health care facilities during 1996–2000, where it was the second most common cause of HA-MRSA infection. In contrast, the SCC*mec* type IV–bearing ST30:USA1100 strain was epidemic in community populations during 1998–2001, causing skin and soft-tissue infections among the homeless and injection drug users [34, 35]. Evolutionary patterns of descent of these 2 strains have been inferred by Enright et al. [4] and Robinson et al. [7], using data obtained from MLST, SCC*mec* allotyping, and *pvl* gene detection. Although the 2 strain types possess very similar virulence-gene profiles, the differences provide additional data for inferring evolutionary pathways (figure 3*A*).

The inferred progenitor of the CC30 lineage is an MSSA ST30 strain bearing the same repertoire of virulence genes common to both ST30:USA1100 and ST36:USA200 (figure 2). It has been proposed previously by Robinson et al. [7] that the progenitor strain acquired, through horizontal gene transfer, the phage-encoded *pvl* genes and, more recently, SCC*mec* type IV. ST30:USA1100 as well as the southwest Pacific/Oceania clone are descended from this branch.

A second branch in CC30 leads to ST36 strains. Two alternative models of ST36 evolution have been postulated by Enright et al. [4] and Robinson et al. [7]. The first model postulates that ST36 descended from a SCC*mec* type IV–bearing ST30 strain through replacement of the SCC*mec* type IV element with the SCC*mec* type II element [4, 7], whereas the second has the ST30

MSSA progenitor acquiring a point mutation in *pta* (phosphate acetyltransferase) followed by a subsequent direct acquisition of SCC*mec* type II [4, 7]. We favor the second, more parsimonious model (figure 3*A*). We further hypothesize that the next step involved the horizontal acquisition of the phage-encoded enterotoxin gene *sea,* which resulted in the E-MRSA16 strain, the second most common MRSA strain in hospitals in the United Kingdom; this strain is represented by the fully sequenced isolate MRSA252. Subsequent acquisition by this strain of the toxic-shock syndrome toxin 1 gene *(tst)* leads to ST36:USA200, the second most common HA-MRSA strain in hospitals in the United States [10].

The acquisition by ST36:USA200 of 2 pyrogenic toxin superantigen genes, *sea* and *tst,* may explain its capacity to cause significantly more sterile site infections, respiratory tract infections, and urinary tract infections than does ST30:USA1100 ($P = .001$) (table 2). In contrast, ST30:USA1100 causes significantly more abscess/wound infections than does ST36:USA200 ($P = .001$) (table 2), likely because of its acquisition of the *pvl* genes, a virulence determinant with well-established linkage to skin and soft-tissue infections.

***Evolutionary descent within the CC8 background genome.*** The CC8 lineage includes 2 fully sequenced genomes, NCTC8325 and COL [8], and the information contained in these genomes has been incorporated into the following analysis (figure 3*B*). NCTC8325 was isolated before 1949 and is the earliest known MSSA isolate belonging to the CC8 lineage. The introduction of methicillin in 1960 resulted in the emergence of COL and other early MRSA strains. Crisostomo et al. have shown that the early MRSA strains were descended from a prevalent ST250 MSSA strain, identified in hospitals in the United Kingdom and Denmark, through acquisition of the SCC*mec* type I element [3]. ST250 is a single-locus MLST variant of ST8, which deviated in only 1 bp in the *yqiL* (acetyl coenzyme A acetyltransferase) housekeeping gene.

The molecular event that is likely responsible for the evolution of the epidemic phenotype among early MSSA strains and early MRSA strains is the acquisition of the staphylococcal pathogenicity island (SaPI) 3, which contains the genes for 3 potent enterotoxins, *seb, sek,* and *seq* [36]. SaPI3 is present in COL but not in NCTC8325, a presumed progenitor of the CC8 lineage. The other major genomic differences between NCTC8325 and COL include the integration of 4 prophages ($\phi$11, $\phi$12, and $\phi$13 in NCTC8325 [37] and $\phi$COL in COL [8]), which do not appear to carry potent virulence genes aside from the $\phi$13-encoded staphylokinase. Taken together, these results suggest that SaPI3 bestows the epidemic phenotype and was first acquired horizontally by early MSSA strains; subsequently, it was inherited vertically by the early MRSA progeny. A descendant of SaPI3-bearing ST250 is the pandemic multidrug-resistant Iberian clone, a single-locus variant of ST250 that contains a linearized plasmid pUB110 inserted in the SCC*mec* type I element downstream of *mecA* [38].

The other branch of the CC8 lineage involves the evolution of 2 contemporary CC8 strains, ST8:USA500 and ST8:USA300. ST8:USA500, already present in our collection in 1996, probably evolved from a NCTC8325-like progenitor through the acquisition of SCC*mec* type IV. It has been almost entirely displaced by ST8:USA300 [12, 27], a new CC8 strain that is now epidemic in community populations in at least 16 US states [1, 6, 10, 11]. The 2 events epidemiologically associated with this occurrence are the acquisition of the *pvl* genes and the *seq-sek*–encoding pathogenicity island.

PVL, a bicomponent leukotoxin virulence factor, has been linked to severe necrotizing fasciitis and necrotizing pneumonia [1, 2]. It has been postulated that this toxin is the principal virulence factor responsible for the epidemic spread of many CA-MRSA strains [39], including ST1 (MW2) in the Midwest of the United States [40], ST30 in Australia [7] and the United States [27], and ST80 in Europe [5]. Our data on the virulence-gene content of ST8:USA300 and other archaic and contemporary CC8 isolates further identified *pvl* as the single virulence determinant associated with the evolution of this new virulent strain. It is important to note, however, that *pvl* is not essential in the evolution of other CA-MRSA strains, including ST8:USA500 and ST59:USA1000 (figure 2).

Two enterotoxin genes, *sek* and *seq,* encode pyrogenic toxin superantigens that lack emetic activity in monkeys, a defining biological feature of classical staphylococcal enterotoxins [41, 42]. In addition to ST8:USA300, these enterotoxins are carried by COL and by 2 recognized CA-MRSA strains, ST59:USA1000 and ST1 (MW2) [40]. Local production of enterotoxins on host mucosal membranes has been postulated to produce sufficient host immunosuppression for promotion of *S. aureus* colonization and transmission [43]. The presence of *sek* and *seq* in 3 CA-MRSA strains suggests that it may play a role in the evolution of community lifestyle.

## Acknowledgment

Reference isolates were obtained through the Network on Antimicrobial Resistance in *Staphylococcus aureus* Program, which is supported under National Institute of Allergy and Infectious Diseases/National Institutes of Health contract N01-AI-95359.

## References

1. Miller LG, Perdreau-Remington F, Rieg G, et al. Necrotizing fasciitis caused by community-associated methicillin-resistant *Staphylococcus aureus* in Los Angeles. N Engl J Med **2005**; 352:1445–53.
2. Gillet Y, Issartel B, Vanhems P, et al. Association between *Staphylococcus aureus* strains carrying gene for Panton-Valentine leukocidin and highly lethal necrotising pneumonia in young immunocompetent patients. Lancet **2002**; 359:753–9.
3. Crisostomo MI, Westh H, Tomasz A, Chung M, Oliveira DC, de Len-

castre H. The evolution of methicillin resistance in *Staphylococcus aureus:* similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. Proc Natl Acad Sci USA **2001**; 98:9865–70.
4. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci USA **2002**; 99:7687–92.
5. Vandenesch F, Naimi T, Enright MC, et al. Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. Emerg Infect Dis **2003**; 9:978–84.
6. Carleton HA, Diep BA, Charlebois ED, Sensabaugh GF, Perdreau-Remington F. Community-adapted methicillin-resistant *Staphylococcus aureus* (MRSA): population dynamics of an expanding community reservoir of MRSA. J Infect Dis **2004**; 190:1730–8.
7. Robinson DA, Kearns AM, Holmes A, et al. Re-emergence of early pandemic *Staphylococcus aureus* as a community-acquired meticillin-resistant clone. Lancet **2005**; 365:1256–8.
8. Gill SR, Fouts DE, Archer GL, et al. 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. J Bacteriol **2005**; 187: 2426–38.
9. Holden MT, Feil EJ, Lindsay JA, et al. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci USA **2004**; 101:9786–91.
10. McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC. Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. J Clin Microbiol **2003**; 41:5113–20.
11. Kazakova SV, Hageman JC, Matava M, et al. A clone of methicillin-resistant *Staphylococcus aureus* among professional football players. N Engl J Med **2005**; 352:468–75.
12. Pan ES, Diep BA, Charlebois ED, et al. Population dynamics of nasal strains of methicillin-resistant *Staphylococcus aureus*—and their relation to community-associated disease activity. J Infect Dis **2005**; 192: 811–8.
13. Francis JS, Doherty MC, Lopatin U, et al. Severe community-onset pneumonia in healthy adults caused by methicillin-resistant *Staphylococcus aureus* carrying the Panton-Valentine leukocidin genes. Clin Infect Dis **2005**; 40:100–7.
14. Gonzalez BE, Martinez-Aguilar G, Hulten KG, et al. Severe staphylococcal sepsis in adolescents in the era of community-acquired methicillin-resistant *Staphylococcus aureus.* Pediatrics **2005**; 115:642–8.
15. Charlebois ED, Bangsberg DR, Moss NJ, et al. Population-based community prevalence of methicillin-resistant *Staphylococcus aureus* in the urban poor of San Francisco. Clin Infect Dis **2002**; 34:425–33.
16. Diep BA, Perdreau-Remington F, Sensabaugh GF. Clonal characterization of *Staphylococcus aureus* by multilocus restriction fragment typing, a rapid screening approach for molecular epidemiology. J Clin Microbiol **2003**; 41:4559–64.
17. Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol **1995**; 33:2233–9.
18. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus.* J Clin Microbiol **2000**; 38:1008–15.
19. Agresti A. Exact inference for categorical data: recent advances and continuing controversies. Stat Med **2001**; 20:2709–22.
20. Jarraud S, Mougel C, Thioulouse J, et al. Relationships between *Staphylococcus aureus* genetic background, virulence factors, *agr* groups (alleles), and human disease. Infect Immun **2002**; 70:631–41.
21. Lina G, Piemont Y, Godail-Gamot F, et al. Involvement of Panton-Valentine leukocidin–producing *Staphylococcus aureus* in primary skin infections and pneumonia. Clin Infect Dis **1999**; 29:1128–32.
22. Peacock SJ, Moore CE, Justice A, et al. Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus.* Infect Immun **2002**; 70:4987–96.
23. Moore PC, Lindsay JA. Genetic variation among hospital isolates of methicillin-sensitive *Staphylococcus aureus:* evidence for horizontal transfer of virulence genes. J Clin Microbiol **2001**; 39:2760–7.
24. Lina G, Boutite F, Tristan A, Bes M, Etienne J, Vandenesch F. Bacterial competition for human nasal cavity colonization: role of staphylococcal *agr* alleles. Appl Environ Microbiol **2003**; 69:18–23.
25. Oliveira DC, de Lencastre H. Multiplex PCR strategy for rapid identification of structural types and variants of the *mec* element in methicillin-resistant *Staphylococcus aureus.* Antimicrob Agents Chemother **2002**; 46:2155–61.
26. Diep BA, Sensabaugh GF, Somboona NS, Carleton HA, Perdreau-Remington F. Widespread skin and soft-tissue infections due to two methicillin-resistant *Staphylococcus aureus* strains harboring the genes for Panton-Valentine leukocidin. J Clin Microbiol **2004**; 42:2080–4.
27. Pan ES, Diep BA, Carleton HA, et al. Increasing prevalence of methicillin-resistant *Staphylococcus aureus* infection in California jails. Clin Infect Dis **2003**; 37:1384–8.
28. Charlebois ED, Perdreau-Remington F, Kreiswirth B, et al. Origins of community strains of methicillin-resistant *Staphylococcus aureus.* Clin Infect Dis **2004**; 39:47–54.
29. Naimi TS, LeDell KH, Como-Sabetti K, et al. Comparison of community- and health care-associated methicillin-resistant *Staphylococcus aureus* infection. JAMA **2003**; 290:2976–84.
30. Wright JS 3rd, Traber KE, Corrigan R, Benson SA, Musser JM, Novick RP. The *agr* radiation: an early event in the evolution of staphylococci. J Bacteriol **2005**; 187:5585–94.
31. Novick RP. Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus.* Plasmid **2003**; 49:93–105.
32. Booth MC, Pence LM, Mahasreshti P, Callegan MC, Gilmore MS. Clonal associations among *Staphylococcus aureus* isolates from various sites of infection. Infect Immun **2001**; 69:345–52.
33. van Leeuwen WB, Melles DC, Alaidan A, et al. Host- and tissue-specific pathogenic traits of *Staphylococcus aureus.* J Bacteriol **2005**; 187:4584–91.
34. Centers for Disease Control and Prevention. Soft tissue infections among injection drug users—San Francisco, California, 1996–2000. JAMA **2001**; 285:2707–9.
35. Binswanger IA, Kral AH, Bluthenthal RN, Rybold DJ, Edlin BR. High prevalence of abscesses and cellulitis among community-recruited injection drug users in San Francisco. Clin Infect Dis **2000**; 30:579–81.
36. Yarwood JM, McCormick JK, Paustian ML, Orwin PM, Kapur V, Schlievert PM. Characterization and expression analysis of *Staphylococcus aureus* pathogenicity island 3: implications for the evolution of staphylococcal pathogenicity islands. J Biol Chem **2002**; 277:13138–47.
37. Iandolo JJ, Worrell V, Groicher KH, et al. Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of *Staphylococcus aureus* 8325. Gene **2002**; 289:109–18.
38. Oliveira DC, Wu SW, de Lencastre H. Genetic organization of the downstream region of the *mecA* element in methicillin-resistant *Staphylococcus aureus* isolates carrying different polymorphisms of this region. Antimicrob Agents Chemother **2000**; 44:1906–10.
39. Chambers HF. Community-associated MRSA—resistance and virulence converge. N Engl J Med **2005**; 352:1485–7.
40. Baba T, Takeuchi F, Kuroda M, et al. Genome and virulence determinants of high virulence community-acquired MRSA. Lancet **2002**; 359:1819–27.
41. Orwin PM, Leung DY, Tripp TJ, et al. Characterization of a novel staphylococcal enterotoxin-like superantigen, a member of the group V subfamily of pyrogenic toxins. Biochemistry **2002**; 41:14033–40.
42. Orwin PM, Leung DY, Donahue HL, Novick RP, Schlievert PM. Biochemical and biological properties of staphylococcal enterotoxin K. Infect Immun **2001**; 69:360–6.
43. Ferens WA, Bohach GA. Persistence of *Staphylococcus aureus* on mucosal membranes: superantigens and internalization by host cells. J Lab Clin Med **2000**; 135:225–30.