

Rapid Molecular Genotyping and Clonal Complex Assignment of *Staphylococcus aureus* Isolates by PCR Coupled to Electrospray Ionization-Mass Spectrometry[∇]

Thomas A. Hall,¹ Rangarajan Sampath,¹ Lawrence B. Blyn,¹ Raymond Ranken,¹ Cristina Ivy,¹ Rachael Melton,¹ Heather Matthews,¹ Neill White,¹ Feng Li,¹ Vanessa Harpin,¹ David J. Ecker,¹ Linda K. McDougal,² Brandi Limbago,² Tracy Ross,⁴ Donna M. Wolk,³ Vicki Wysocki,³ and Karen C. Carroll^{4*}

Ibis Biosciences, Inc., 1896 Rutherford Rd., Carlsbad, California 92008¹; Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, 1600 Clifton Rd., Atlanta, Georgia 30333²; University of Arizona, 1501 N. Campbell Ave., Tucson, Arizona³; and Johns Hopkins University, 600 N. Wolfe Street, Baltimore, Maryland 21287⁴

Received 12 November 2008/Returned for modification 6 January 2009/Accepted 10 March 2009

We describe a high-throughput assay using PCR coupled to electrospray ionization-mass spectrometry (PCR/ESI-MS) to determine the genotypes of *Staphylococcus aureus* isolates. The primer sets used in the PCR/ESI-MS assay were designed to amplify the same genes analyzed in multilocus sequence typing (MLST). The method was used to identify the clonal complex and USA type of each isolate and is suitable for use in a clinical or public-health setting. The method was validated using a panel of diverse isolates from the Centers for Disease Control and Prevention that were previously characterized by MLST and pulsed-field gel electrophoresis (PFGE). Clinical isolates from two geographically distinct hospitals were characterized, and the clustering results were in agreement with those for repetitive-element PCR and PFGE. The PCR/ESI-MS method enables genotyping of over 180 samples of *S. aureus* per day in an automated fashion.

Microbial genotyping methods are used to study global evolution of pathogens and to reveal genetic clonality, which is used to determine point sources during epidemiological investigations (1, 2). There are a variety of molecular genotyping methods available, each with certain strengths and weaknesses (4). The methods that are most valuable for understanding pathogen evolution over the long term rely on techniques such as multilocus sequence typing (MLST). This genotyping method is suitable for tracking global pathogen evolution over decades because it focuses on stable housekeeping genes and provides information that places isolates into genetic categories that provide perspective on a global scale (10, 12). For real-time outbreak investigations in health care settings, the molecular genotyping method used must focus on identification of recent genetic changes. Generally, techniques such as pulsed-field gel electrophoresis (PFGE) (27), repetitive-element PCR (rep-PCR) (32), or *Staphylococcus* protein A typing (19, 20) are used to support these investigations. These methods provide genetic signatures indicative of a microbe's recent history. However, electrophoretic methods are based on gel mobility and therefore provide analog rather than digital signatures. This means that it may be difficult to compare data between laboratories, and the techniques do not reliably reveal where an isolate fits with respect to the global evolutionary history of a pathogen.

Yet, a third category of molecular genotyping is direct PCR analysis of an isolate. Direct PCR is aimed at determining the

presence and specific types of certain acquired genetic elements in the microbial genome. The regions targeted for amplification are those portions of the genome known to change over very short periods of time. Examples of such elements include virulence factors, genes that mediate antimicrobial resistance, and housekeeping genes that encode proteins that are the target of antibiotics. This category of molecular genotyping is pathogen specific, as acquired genetic elements tend to be unique to each pathogen (28).

Here, we describe the application of PCR coupled with electrospray ionization-mass spectrometry (PCR/ESI-MS) (7, 16, 24) to *Staphylococcus aureus* genotyping. The primer sets used in the PCR/ESI-MS assay were designed to amplify the same genes analyzed in MLST. This assay allows us to determine the molecular genotype, as is done with MLST, and to directly characterize specific methicillin (meticillin)-resistant *S. aureus* (MRSA) genetic elements.

MATERIALS AND METHODS

Bacterial isolates. Thirty isolates of *Staphylococcus aureus* were selected to represent the diversity of isolates from a Centers for Disease Control and Prevention (CDC) national database, including a variety of health care-associated and community-acquired strain types (21, 28). The isolates obtained from Johns Hopkins Hospital were a convenience sample of 240 clinical isolates of *S. aureus* from the following sources: blood samples ($n = 31$), wounds/skin cultures ($n = 43$), nasal surveillance sites ($n = 57$), respiratory nonsurveillance sites ($n = 78$), urine samples ($n = 10$), sterile body sites ($n = 10$), and medical devices, such as catheter tips ($n = 11$). The 47 clinical isolates from the University of Arizona Medical Center were obtained from various sources, including wounds, abscesses, drainages, sputum samples, cardiovascular catheter tips, and tissues.

PCR/ESI-MS. General methods for genotypic characterization of isolates by multilocus PCR/ESI-MS using the commercially available Ibis T5000 instrument (Ibis Biosciences) have previously been described (9, 13). For strain differentiation and genotyping of *S. aureus*, a set of unique sequences from the seven genes commonly used for MLST of *S. aureus* was assembled (10, 12). These reference

* Corresponding author. Mailing address: Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21287. Phone: (410) 955-5077. Fax: (410) 614-8087. E-mail: kcarroll7@jhmi.edu.

[∇] Published ahead of print on 18 March 2009.

TABLE 1. Primer pair sets for PCR/ESI-MS analysis of *S. aureus*

Primer pair code	Gene target	Coordinates	Primer sequence
BCT3025	<i>arcC</i>	2724791 to 2724920	5'-TGAATAGTGATAGAAGCTGTAGGCACAATCGT-3' 5'-TGCCTAATCTTCAACTTCTTCTTCGT-3'
BCT2149	<i>aroE</i>	1674546 to 1674697	5'-TGGGGCTTTAAATATTCCAATTGAAGATTTTCA-3' 5'-TACCTGCATTAATCGCTTGTTTCATCAA-3'
BCT2150	<i>aroE</i>	1674392 to 1674523	5'-TGATGGCAAGTGGATAGGGTATAATACAG-3' 5'-TAAGCAATACCTTTACTTGCACCACCTG-3'
BCT2156	<i>gmk</i>	1191206 to 1191337	5'-TCACCTCCAAGTTTAGATCACTTGAGAGA-3' 5'-TGGGACGTAATCGTATAAAATTCATCATTTTC-3'
BCT2157	<i>pta</i>	629121 to 629229	5'-TCTTGTATTATGCTGGTAAAGCAGATGG-3' 5'-TGGTACACCTGGTTTCGTTTTGATGATTTGTA-3'
BCT2161	<i>tpi</i>	830671 to 830799	5'-TCCCACGAAACAGATGAAGAAATTAACAAAAAG-3' 5'-TGGTACAACATCGTTAGCTTTACCACTTTCACG-3'
BCT2163	<i>yqi</i>	379057 to 379199	5'-TGAATTGCTGCTATGAAAGGTGGCTT-3' 5'-TCGCCAGCTAGCACGATGTCATTTTC-3'
BCT2166	<i>yqi</i>	379190 to 379311	5'-TAGCTGGCGGTATGGAGAATATGTCT-3' 5'-TCCATCTGTAAACCATCATATACCATGCTATC-3'

alignments were then used for the design of 22 primer pairs for PCR/ESI-MS analysis. The resolution provided by PCR/ESI-MS analysis was calculated, starting with the primer pair providing the best sequence resolution, using theoretical amplicon base compositions determined for each one of 710 *S. aureus* sequence types (<http://saureus.mlst.net>). Comparison of these base composition signatures defined the number of sequence types (STs) that were incompatible with any particular type at this particular locus. The average proportion of sequence types excluded by PCR/ESI-MS, with the corresponding standard deviations, was compared against the number of loci used in the analysis. This process was repeated, using base composition signatures extended by one additional locus at a time. A final panel of eight primer pairs was chosen to maximize average resolution with a minimum number of primer pairs. The final choice of primer pairs (Table 1) did not include all original gene targets used in sequence-based MLST but was instead chosen on the basis of strain resolution ability, amplification specificity, and amplification sensitivity with laboratory strains and sample throughput considerations of the PCR/ESI-MS platform. DNA from isolated colonies was distributed into each of 8 wells of a 96-well plate that contained pairs of PCR primers targeted to eight regions of six different MLST alleles. Twelve colonies were analyzed on each microtiter plate. The resulting amplicons were desalted and injected into a mass spectrometer by using a robotic autosampler as described previously (6). Spectra were generated and analyzed using automated software to calculate the base composition (A, G, C, and T counts) in each strand of the amplicon. The data from each of the eight spectra were concatenated to produce a 32-position PCR/ESI-MS type. This experimentally derived base composition was then compared to the calculated base compositions derived from the sequences of the alleles in the MLST database to identify the STs and associated clonal complexes.

Correlation of base composition signatures with clonal complexes. The set of 853 distinct STs available from the *S. aureus* MLST isolate database (<http://saureus.mlst.net/portable/portable.xls>) as of 7 July 2008 was analyzed with the BURST algorithm (<http://eburst.mlst.net/>) to assign clonal complexes by using the default stringency method, where every member of a complex shares six or seven alleles with at least one other member. A total of 682 MLST types were assigned to clonal complexes, and 171 were not. With the exception of four distinct cases of base composition profiles that correlate with MLST types belonging to two different clonal complexes, theoretical base composition signatures were correctly grouped. The four pairs of sequence types that belong to two different clonal complexes that generated the same base composition signature are as follows: (i) sequence types 673 (clonal complex 8) and 717 (clonal complex 12), (ii) sequence types 514 and 550 (clonal complex 1) and 221 and 402 (clonal complex 5), (iii) sequence types 437 (clonal complex 25) and 117 (clonal complex 97), and (iv) sequence types 247 (clonal complex 5) and 615 (clonal complex 8).

There were 32 base composition signatures (out of a total of 368) that correlated with at least one sequence type not assigned to any clonal complex.

rep-PCR. The DiversiLab System (bioMérieux, Durham, NC) combines standardized DNA extraction methods, primers, and reagents with postamplification microfluidics-based electrophoresis to perform bacterial strain typing of MRSA. MRSA strains were cultured in trypticase soy broth incubated overnight at 35°C. DNA from each isolate was extracted using an UltraClean microbial DNA isolation kit (Mo Bio Laboratories, Solana Beach, CA). The extracted DNA (25 to 50 ng/ml) was amplified using a DiversiLab *Staphylococcus* fingerprinting kit, *Taq* polymerase, and 10× PCR buffer per the manufacturer's instructions. During PCR, the rep-PCR primers bind repetitive DNA sequences and generate amplicons of different sizes. Amplicons are size fractionated using disposable microfluidics chips that contain a copolymer matrix and an intercalating dye, which allows charge separation of the DNA fragments in a microseparation channel in the chip placed when electrophoresis occurs in an Agilent 2100 Bioanalyzer (Agilent Technologies, Foster City, CA). Analysis of amplicon fragments and subsequent data interpretation were accomplished via Web-based software which classifies unique fragment patterns into rep-PCR strain types (14, 15, 30). Sample similarities were calculated by two methods: the Pearson correlation coefficient and the Kullback-Leibler coefficient. Both methods compare all possible pairs of samples in a data set with respect to band presence or absence and position as well as band intensity. DiversiLab dendrograms are created from a distance matrix by using the unweighted-pair group method with arithmetic means. Known pulsed-field types (PFT) USA100 to USA1200 were inferred from the rep-PCR strain types by comparison with DiversiLab's PFT reference library.

PFGE methods. Isolates from Johns Hopkins were genotyped using PFGE by comparison of the pulsed-field patterns with those of reference strains obtained from the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA, Reston, VA), supported under NIAID, NIH, contract no. N01-AI-95359. The NARSA reference numbers corresponding to clones USA100 to USA800 are NRS 382, NRS 383, NRS 384, NRS 123, NRS 385, NRS 22, NRS 386, and NRS 387, respectively. PFGE was performed using a protocol developed at the Johns Hopkins Hospital Division of Medical Microbiology. Overnight broth cultures of clinical strains of *S. aureus* as well as two well-characterized nosocomial MRSA control strains were pelleted, and bacterial DNA was extracted from agarose plugs by using a solution containing lysostaphin and lysozyme. Restriction enzyme digestion was performed using SmaI. Restriction endonuclease fragments were analyzed by PFGE using a contour-clamped homogeneous electric field DR-II apparatus (Bio-Rad Laboratories, Inc., Hercules, CA) set at 14°C, with the following settings: initial switch, 5 s; final switch, 50 s; and time, 23.5 h. After electrophoresis, gels were stained with ethidium bromide. Macrorestriction DNA banding patterns were digitized and analyzed using Mo-

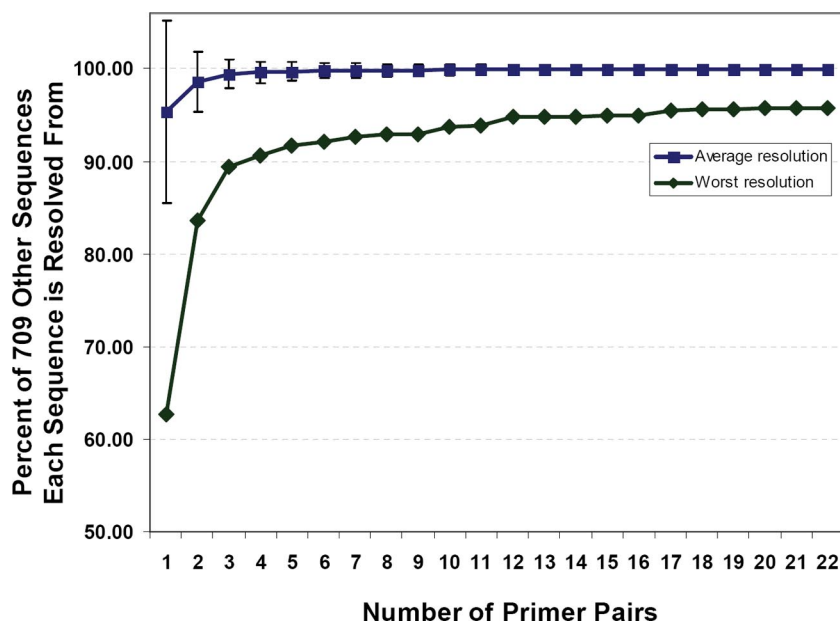


FIG. 1. Cumulative resolution of 710 concatenated *Staphylococcus aureus* MLST sequences from each other, using PCR product base compositions from 1 to 22 pairs of primers. The blue line shows the average resolution (\pm standard deviation) of each sequence from the 709 others as increasing numbers of primer pairs are added. The green line shows the worst resolution for each number of primer pairs. For example, with the best combined set of eight primer pairs, the 710 sequences are broken into 320 groups. Sequences on average are each different from 707.78 other sequences, or 99.83% of the other 709 sequences (\pm 0.69). With the same primer pair set, the largest unresolved group of sequences has 51 members. The worst resolution of any sequence from the other 709 sequences for primer pairs is therefore 92.95%.

lecular Analyst DNA fingerprinting software (Bio-Rad). Patterns were compared to the USA control strains, and an isolate with three or fewer band differences from the USA clone was presumed to be that strain (27).

RESULTS

Multilocus PCR/ESI-MS follows the logic of MLST. PCR amplicons were generated from the same genes used in MLST analysis. In the case of PCR/ESI-MS, amplicons were analyzed by mass spectrometry to determine amplicon base composition as described previously (7). The final set of eight primer pairs used for the assay was selected to maximize MLST discrimination with a minimum number of primer pairs. The primer pair set provided an average discrimination of $99.83\% \pm 0.69\%$ when assessed using a set of 710 distinct STs (Fig. 1). The eight primer pairs that make up the *Staphylococcus aureus* genotyping panel are shown in Table 1. The process of associating mass spectrometry-derived base composition signatures to STs and clonal complexes is illustrated in Fig. 2. The mass spectrometry measurements required approximately 45 s per PCR, so analysis of each *S. aureus* colony required about 6 min of mass spectrometry time. Thus, PCR/ESI-MS analysis can be carried out as a highly automated, high-throughput process. One instrument is capable of analysis of the nucleic acid extracts from 180 bacterial colonies (15 microtiter plates) per day.

Because base composition analysis is not as information rich as sequencing, a theoretical analysis of the *S. aureus* MLST database was conducted to determine the resolving power of base composition analysis compared with that of sequencing. Each of the STs in the database were analyzed by calculating the expected base compositions of each of the primer pair

target sites used in the PCR/ESI-MS assay and concatenating the base composition signatures from the alleles to create a theoretical PCR/ESI-MS type for each MLST sequence type. Figure 3A shows the comparison of sequence types and PCR/ESI-MS types of the *S. aureus* MLST database organized by clonal complexes.

The PCR/ESI-MS base composition analysis cannot distinguish the order of nucleotides in the sequence, and the entire region sequenced in MLST is not encompassed within the PCR/ESI-MS amplicon. Therefore, there was typically more than one sequence type consistent with each PCR/ESI-MS type. For example, clonal complex 5, the most abundantly populated clonal complex in the MLST database, comprises 124 sequence types that reduce to 50 PCR/ESI-MS types. Clonal complex 8, the next-most-abundant clonal complex, comprises 109 sequence types that produce 34 unique PCR/ESI-MS types. Although PCR/ESI-MS analysis does not unambiguously assign an isolate to a unique sequence type, 98.8% of 682 STs that were assigned to clonal complexes by eBurst analysis had base composition signatures that correlated with only the correct clonal complex. There were only four pairs of STs (of 682 assigned to clonal complexes) where the theoretical PCR/ESI-MS types were consistent with more than one clonal complex (Fig. 3A). Thus, while PCR/ESI-MS provides less information than sequence type, assignments of STs to clonal complexes upon the basis of theoretical base composition signatures were in agreement with MLST assignments for 99% of analyzed sequence types. The PCR/ESI-MS technique also provides a substantial degree of resolution within each complex (as shown by the pie slices in Fig. 3A).

PCR/ESI-MS was then used to analyze 30 blinded isolates of

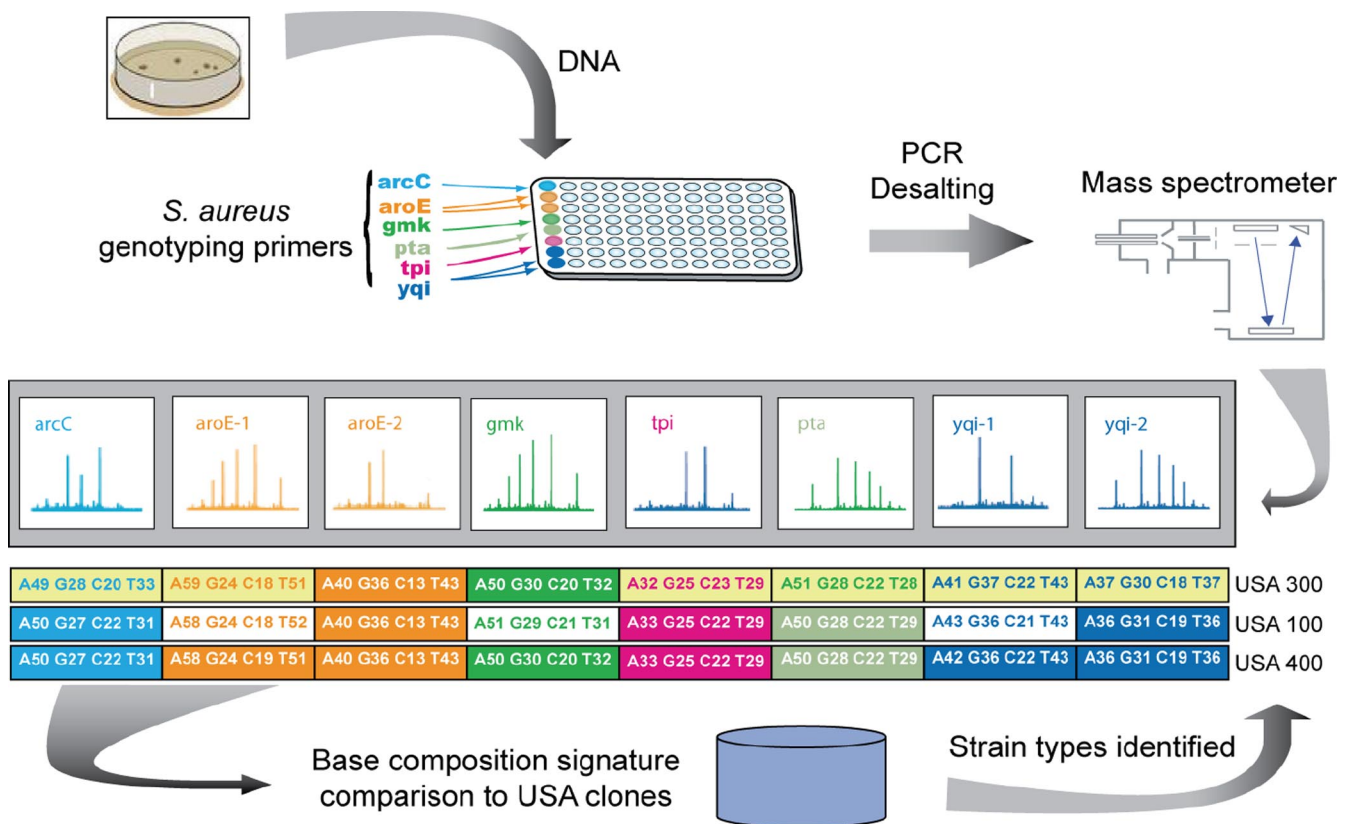


FIG. 2. Flow scheme for PCR/ESI-MS genotyping of *S. aureus*. Isolated DNA is distributed into eight wells of a microtiter plate, allowing 12 samples to be analyzed per plate. Each reaction mixture contains a pair of primers targeted to a gene used in MLST. Following PCR and desalting, amplicons are analyzed by ESI-MS. Amplicon masses are used to calculate base compositions—the A, G, C, and T counts—of the PCR products. Comparison to a database of calculated base compositions derived from the sequences in the MLST database allows assignment of clonal complexes and USA types.

S. aureus isolates from the CDC that had previously been characterized by PFGE and MLST (21). Twenty-eight of the 30 isolates had previously been associated with clonal complexes. One of the 30 isolates, a methicillin-susceptible *S. aureus* isolate had not been characterized by MLST. The distribution of isolates by clonal complexes is shown in Fig. 3B, and the detailed PCR/ESI-MS signatures are shown in Fig. 4. As anticipated from the theoretical analysis of the MLST database, when PCR/ESI-MS was used, all isolates were correctly assigned to clonal complexes and were sometimes parsed further within the complex. For example, clonal complex 8 had two PCR/ESI-MS types and eight PFGE types. PFGE is known to parse *S. aureus* isolates more finely than MLST. For example, USA300 and USA500 are both MLST type 8; these PFGE types were not resolved by PCR/ESI-MS and are also sometimes difficult to distinguish based only on PFGE (B. Limbago, unpublished results). Thus, PCR/ESI-MS provided resolution at the level of clonal complexes for the diverse CDC *S. aureus* panel, with somewhat finer subdividing capability within a clonal complex.

The PCR/ESI-MS approach using a diverse set of isolates from the CDC was successfully validated as described above. Subsequently, analysis was performed using 47 unique isolates cultured from clinical specimens that included wounds, abscesses, drainages, sputum samples, cardiovascular catheter

tips, and tissues collected at the University of Arizona Medical Center. The purpose of this sample set was to evaluate the correlation of PCR/ESI-MS with an independent typing method (rep-PCR in this case) in an outbreak setting. Another 240 *S. aureus* isolates obtained from Johns Hopkins Hospital in Maryland were also analyzed by PCR/ESI-MS. These isolates represented clinical strains recovered from patients from various sources, including blood cultures, upper- and lower-respiratory-tract specimens, wounds, sterile body fluids, tissues, and intravascular catheter tips, and were chosen to evaluate the PCR/ESI-MS method in a diverse set of samples acquired in the context of a local setting with high MRSA endemicity.

The distribution of clonal complexes and PCR/ESI-MS genotypes for the Arizona isolates are shown in Fig. 5A. The Arizona isolates were predominantly clonal complex 5 (36 isolates), which were subdivided into PCR/ESI-MS types 2, 14, and 10 (Fig. 5A), and clonal complex 8 (10 isolates), comprising PCR/ESI-MS types 1 and 7. The 240 Johns Hopkins Hospital isolates (Fig. 5B) were also predominantly of clonal complex 8 (82 isolates), subdivided into PCR/ESI-MS types 1 and 7, and clonal complex 5 (71 isolates), subdivided into PCR/ESI-MS types 2, 16, 10, 19, and 23. There were also a substantial number of representatives from clonal complexes 30, 15, 45, and 59 and smaller numbers of isolates from other clonal complexes (for minor genotypes, see the legend to Fig. 5B).

Sample ID	PCR/ESI-MS Types	Clonal complex	USA Type by PFGE	MLST Sequence Type	BCT3025 <i>aroC</i>	BCT2149 <i>aroE</i>	BCT2150 <i>aroE</i>	BCT2156 <i>gmk</i>	BCT2157 <i>pta</i>	BCT2161 <i>tpi</i>	BCT2163 <i>yqi</i>	BCT2166 <i>yqi</i>			
CDC0010	1	CC8	Archaic	250	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC0026			EMRSA10	254	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC0030			Iberian	247	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC0015			USA300	ND	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC0014				8	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC004			USA500	8	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC0019				8	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			
CDC008				USA300/500	ND	49 28 20 33	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37		
CDC0023			4	EMRSA9	Brazilian	239	50 28 20 32	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37	
CDC0025					240	50 28 20 32	59 24 18 51	40 36 13 43	50 30 20 32	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37		
CDC001	2	CC5	USA100	5	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDC0022				225	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDC006				ND	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDCVRS-A				371	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDCVRS-B				ND	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDC007				USA800	5	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36		
CDC0027				10	USA100	231	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	33 25 22 29	50 28 22 29	43 36 21 43	36 31 19 36	
CDC0011				3	CC30	USA200	ND	50 28 20 32	58 24 19 51	41 36 12 43	51 29 21 31	32 25 23 29	50 28 22 29	42 36 22 43	37 30 18 37
CDC0021						36	50 28 20 32	58 24 19 51	41 36 12 43	51 29 21 31	32 25 23 29	50 28 22 29	42 36 22 43	37 30 18 37	
CDC0012						USA1100	30	50 28 20 32	58 24 19 51	41 36 12 43	51 29 21 31	32 25 23 29	50 28 22 29	42 36 22 43	37 30 18 37
CDC002	6	CC45	USA600	ND	51 27 22 30	58 24 19 51	42 36 12 42	51 29 20 32	33 25 23 28	50 28 22 29	42 36 22 43	37 30 18 37			
CDC0028			45	51 27 22 30	58 24 19 51	42 36 12 42	51 29 20 32	33 25 23 28	50 28 22 29	42 36 22 43	37 30 18 37				
CDC005	5	CC59	USA1000	59	49 28 20 33	59 23 19 51	40 36 14 42	51 29 21 31	32 25 24 28	51 27 21 30	42 36 22 43	37 30 18 37			
CDC0018			59	49 28 20 33	59 23 19 51	40 36 14 42	51 29 21 31	32 25 24 28	51 27 21 30	42 36 22 43	37 30 18 37				
CDC0029	11	CC1	USA400	1	50 27 22 31	58 24 19 51	40 36 13 43	50 30 20 32	33 25 22 29	50 28 22 29	42 36 22 43	36 31 19 36			
CDC0020	12	CC15	USA900	15	49 27 23 31	59 23 18 52	40 36 13 43	50 30 20 32	33 25 22 29	50 28 21 30	42 36 22 43	36 31 18 37			
CDC0013	8	CC80	undefined	80	50 27 22 31	59 24 18 51	40 36 13 43	51 29 21 31	32 25 23 29	51 28 22 28	42 36 22 43	37 30 18 37			
CDC0024	13	CC22	EMRSA15	22	50 27 22 31	57 25 19 51	40 36 13 43	51 29 22 30	33 25 22 29	51 28 22 28	42 36 22 43	37 30 18 37			
CDC0016	9	None	USA1200	ND	50 27 20 33	58 24 19 51	40 37 13 42	51 29 21 31	32 25 24 28	50 28 21 30	42 36 22 43	37 30 18 37			
CDC003	7	None	USA700	72	50 27 22 31	58 24 18 52	40 36 13 43	51 29 21 31	32 25 23 29	51 28 22 28	41 37 22 43	37 30 18 37			

FIG. 4. PCR/ESI-MS signatures of the CDC panel of *S. aureus* isolates. The eight columns on the right show base compositions (given as the numbers of A, G, C, and T counts) of amplicons from the MLST target genes (Table 1). The results shown in the clonal complex, PFGE type, and MLST sequence type columns were determined previously. Base compositions that are the same within a column are colored the same. ND, not determined.

analysis of a large collection of isolates from around the world and established the evolutionary history of MRSA (23). They determined that there are 11 major MRSA clones that fall into five major groups of related genotypes, designated clonal complexes (11). MLST is particularly useful for molecular evolution studies because it provides unambiguous digital readouts that are portable over the Internet and amenable to database population.

However, MLST requires multiple PCRs, sequencing, and multiple sequence analysis; these techniques are difficult to perform and validate in a clinical laboratory setting.

L. K. McDougal and colleagues from the CDC conducted an extensive PFGE analysis of *S. aureus* clones from around the United States and determined that there were eight major lineages, designated PFT USA100 through USA800, categori-

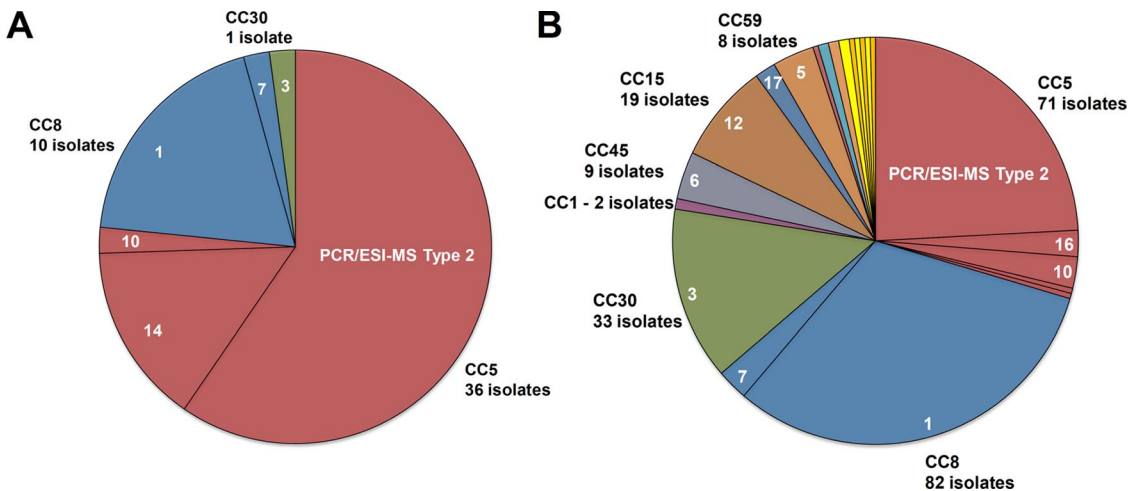


FIG. 5. Genotypes and assignment into clonal complexes by PCR/ESI-MS of Arizona and Johns Hopkins Hospital isolates. (A) Distribution of clonal complexes (colors) and PCR/ESI-MS types for 47 isolates from the University of Arizona. (B) Distribution of clonal complexes (CCs) and PCR/ESI-MS types for 240 isolates from Johns Hopkins Hospital. The unlabeled minor isolates were CC97/ESI-MS type 17 (four isolates); ESI-MS type 9 (two isolates); two isolates each of CC12/ESI-MS type 18, CC1/ESI-MS type 21, and CC7/ESI-MS type 24; and one isolate each of ESI-MS types 11, 20, 22, 25, 27, and 28 (not assigned to clonal complexes).

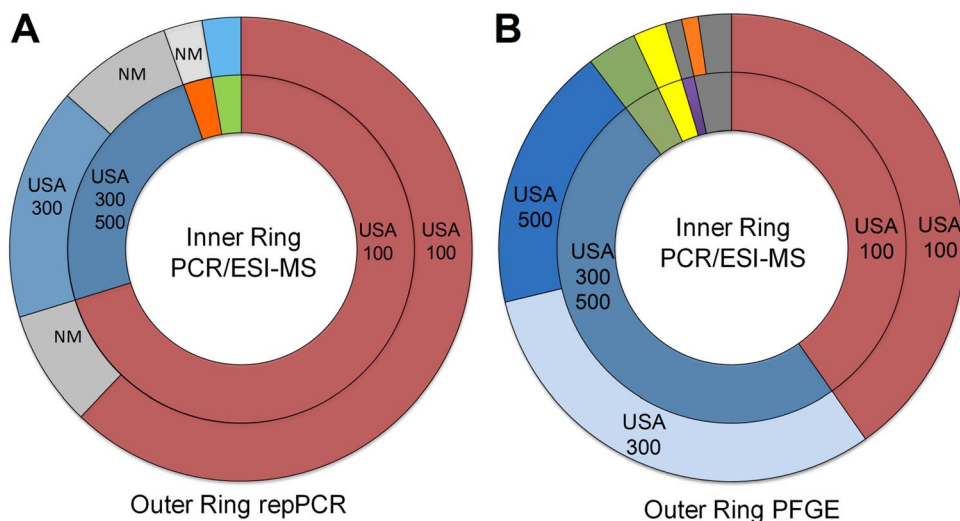


FIG. 6. Comparison of PCR/ESI-MS with rep-PCR and PFGE. (A) Comparison of PCR/ESI-MS and rep-PCR (Arizona isolates). The inner ring represents PCR/ESI-MS results, and the outer ring represents the corresponding isolates identified by rep-PCR. The unlabeled isolates in the inner ring (clockwise) were one each of PCR/ESI-MS 3 (USA700, orange) and PCR/ESI-MS type 7 (USA200/1100, green). The gray isolates in the outer ring (labeled NM) had no match in the DiversiLab database. The unlabeled blue isolate in the outer ring was USA200-like by rep-PCR and corresponded to USA200/1100 by PCR/ESI-MS. (B) Comparison of PCR/ESI-MS and PFGE (Johns Hopkins Hospital isolates). Red slices represent USA100, medium blue (inner ring) represents USA300/500, and light blue and dark blue (outer ring) represent USA300 and USA500, respectively. The isolates unlabeled by PCR/ESI-MS (inner ring) were USA200 USA1100 (green, three isolates), USA700 (yellow, two isolates), USA1000 (purple, one isolate), and undetermined isolates (gray, three isolates). The corresponding data for isolates unlabeled by PFGE indicated, in clockwise order, USA200 (green, three isolates), USA700 (yellow, two isolates), USA1000 (purple, one isolate), USA800 (orange, one isolate), and two unique PFGE types (gray) corresponding to types undetermined by PCR/ESI-MS.

zations that have now been extended to the USA1200 designation (29). By codetermination of the MLST and PFGE types of these isolates, correlation with the previously established MLST nomenclature was established (21). More recently, CDC scientists directly characterized a set of specific genetic elements found in MRSA strains that cause most community-acquired MRSA infections. These strains are designated USA300. The relevant genetic elements are found in MLST type 8 isolates and contain *SCCmec* type IV and the Panton-Valentine leukocidin toxin gene (28). Thus, the CDC approach to characterization of strains combines general molecular genotyping methods (i.e., PFGE) with direct molecular detection and typing of genetic elements that are characteristic of MRSA strains. Coupling of general genotyping methods with direct analysis of specific genetic elements is necessary for an accurate view of the strain genealogy because gene transfer between different strains of microbes and acquisition of mobile genetic elements occur over relatively short time intervals.

The clinical utility of a molecular genotyping or characterization method depends upon the value of the information that it provides as well as practical issues, such as technical difficulty, labor requirements, speed, throughput, and portability of data. Knowing the detailed molecular characteristics of an isolate can guide intervention decisions during an outbreak and enable prudent use of infection control resources. Despite some debate, pending health care legislation aimed at health care-associated infections may necessitate the ability to prove or disprove hospital-based acquisition in order to obtain insurance reimbursement (31).

Although base composition analysis as described here does not achieve the strain resolution of PFGE analysis, the PCR/

ESI-MS approach described here is simple, straightforward, and automated and can correctly assign isolates to clonal complexes with a high degree of agreement (~99%) with MLST analysis. It could be argued that both PCR/ESI-MS and MLST analyses lack discriminatory ability relative to PFGE. However, this does not make a strong statement, given that, for both MLST and PCR/ESI-MS, multiple loci, each displaying multiple alleles, are combined for inference of types. The probability of two unrelated genotypes having the same sequence type is low (11). It may even happen that the rapid accumulation of genetic variation that can be indexed by PFGE could lead to differences in PFGE patterns among the descendants of ancestral genotypes, resulting in some *S. aureus* clones being unnecessarily subdivided.

Here, we describe a new, high-throughput method for genotyping of *S. aureus* that is similar to traditional MLST strain typing but provides analysis of PCR amplicons via ESI-MS rather than standard DNA sequencing. The base compositions of the PCR amplicons, determined from PCR product masses, provide a high-resolution digital signature that is directly related to sequences of known genotypes. Digital genomic signatures, such as a nucleotide sequence or base composition of a selected region of a microbe's genome, are precise and comparable among laboratories (5, 8, 12). Another advantage of the PCR/ESI-MS method is automation and instrument throughput: a specimen is analyzed in 6 min of mass spectrometry time, and 180 samples can be analyzed per day. The instrument used in the PCR/ESI-MS studies is a prototype and not the version expected to be placed in clinical microbiology laboratories. This paper reports only on the utility of the genotyping component of the assay. The assay will also have the

capability of identifying staphylococci directly from clinical samples while simultaneously allowing for strain characterization, using a multiplexed assay. The ability of the assay to identify staphylococci to the species level while also determining resistance genes such as *mecA* and *mupA* and virulence factors such as Panton-Valentine leukocidin will be reported in a separate paper (D. Wolk, L. Blyn, T. Hall, R. Sampath, R. Renken, C. Ivy, R. Melton, H. Matthews, N. White, F. Li, V. Harpin, D. Ecker, B. Limbago, L. M. McDougal, V. Wysock, M. Cai, and K. Carroll, submitted for publication).

In this report, we first analyzed a diverse panel of MRSA isolates obtained from a national surveillance study containing representative clonal complexes and PFGE groups (21, 28). The PCR/ESI-MS method correctly grouped all the isolates into clonal complexes previously assigned by MLST and provided additional resolution within the clonal complexes. When used to analyze two sets of MRSA isolates from hospitals in different geographic locations, PCR/ESI-MS grouped the isolates in excellent agreement with complexes determined by PFGE (26) or rep-PCR. Each of these methods measures different genetic variations in *S. aureus*. However, since all three methods measure signals grounded in the DNA sequences of the microbes, related isolates should generally group together, as was observed. Here, we focused on genotyping, but the PCR/ESI-MS method is also amenable to more-targeted genomic analysis of *S. aureus*, such as determination of the presence or absence of virulence factors or acquired genes or mutations that mediate antibiotic resistance (Wolk et al., submitted).

In summary, our PCR/ESI-MS approach to genotyping of *S. aureus* follows the logic of MLST but uses mass spectrometry to analyze the PCR products instead of sequencing. The method correctly assigned 100% of the isolates from a CDC validation panel and two collections of hospital isolates into the clonal complexes. The PCR/ESI-MS approach provides a rapid, high-throughput, and automated method of genotyping *S. aureus* isolates suitable for use in a clinical or public-health setting.

ACKNOWLEDGMENTS

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the CDC.

The work at the University of Arizona was partially supported by grant AI065359.

REFERENCES

- Blanc, D. S. 2004. The use of molecular typing for epidemiological surveillance and investigation of endemic nosocomial infections. *Infect. Genet. Evol.* **4**:193–197.
- Blanc, D. S., C. Petignat, A. Wenger, G. Kuhn, Y. Vallet, D. Fracheboud, S. Trachsel, M. Reymond, N. Troillet, H. H. Siegrist, S. Ouevray, M. Bes, J. Etienne, J. Bille, P. Francioli, and G. Zanetti. 2007. Changing molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in a small geographic area over an eight-year period. *J. Clin. Microbiol.* **45**:3729–3736.
- Boucher, H. W., and G. R. Corey. 2008. Epidemiology of methicillin-resistant *Staphylococcus aureus*. *Clin. Infect. Dis.* **46**(Suppl. 5):S344–S349.
- Caugant, D. A. (ed.) Molecular epidemiology of microorganisms, in press. *Methods in molecular biology*, vol. 551. Humana Press, Totowa, NJ.
- Ecker, D. J., and K. C. Carroll. 2005. Investments in high-payoff technologies could reduce toll of infections. *ASM News* **71**:576–581.
- Ecker, D. J., J. Drader, J. Gutierrez, A. Gutierrez, J. Hannis, A. Schink, R. Sampath, J. A. Ecker, L. B. Blyn, M. W. Eshoo, T. A. Hall, M. Tobarmosquera, Y. Jiang, K. Sannes-Lowery, L. Cummins, B. Libby, D. J. Walcott, C. Massire, R. Ranken, S. M. Manalili, C. Ivy, R. Melton, H. Levene, V. Harpin, F. Li, N. White, M. Pear, V. Samant, D. Knize, D. Robbins, K. Rudnick, F. Hajjar, and S. A. Hofstadler. 2006. The Ibis T5000 universal biosensor—an automated platform for pathogen identification and strain typing. *JALA* **11**:341–351.
- Ecker, D. J., R. Sampath, L. B. Blyn, M. W. Eshoo, C. Ivy, J. A. Ecker, B. Libby, V. Samant, K. Sannes-Lowery, R. E. Melton, K. Russell, N. Freed, C. Barrozo, J. Wu, K. Rudnick, A. Desai, E. Moradi, D. J. Knize, D. W. Robbins, J. C. Hannis, P. M. Harrell, C. Massire, T. A. Hall, Y. Jiang, R. Ranken, J. J. Drader, N. White, J. A. McNeil, S. T. Crooke, and S. A. Hofstadler. 2005. Rapid identification and strain-typing of respiratory pathogens for epidemic surveillance. *Proc. Natl. Acad. Sci. USA* **102**:8012–8017.
- Ecker, D. J., R. Sampath, C. Massire, L. B. Blyn, T. A. Hall, M. W. Eshoo, and S. A. Hofstadler. 2008. Ibis T5000: a universal biosensor approach for microbiology. *Nat. Rev. Microbiol.* **6**:553–558.
- Ecker, J. A., C. Massire, T. A. Hall, R. Ranken, T. T. Pennella, C. Agasino Ivy, L. B. Blyn, S. A. Hofstadler, T. P. Endy, P. T. Scott, L. Lindler, T. Hamilton, C. Gaddy, K. Snow, M. Pe, J. Fishbain, D. Craft, G. Deye, S. Riddell, E. Milstrey, B. Petruccioli, S. Brisse, V. Harpin, A. Schink, D. J. Ecker, R. Sampath, and M. W. Eshoo. 2006. Identification of *Acinetobacter* species and genotyping of *Acinetobacter baumannii* by multilocus PCR and mass spectrometry. *J. Clin. Microbiol.* **44**:2921–2932.
- Enright, M. C., N. P. Day, C. E. Davies, S. J. Peacock, and B. G. Spratt. 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J. Clin. Microbiol.* **38**:1008–1015.
- Enright, M. C., D. A. Robinson, G. Randle, E. J. Feil, H. Grundmann, and B. G. Spratt. 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc. Natl. Acad. Sci. USA* **99**:7687–7692.
- Enright, M. C., and B. G. Spratt. 1999. Multilocus sequence typing. *Trends Microbiol.* **7**:482–487.
- Hannis, J. C., S. M. Manalili, T. A. Hall, R. Ranken, N. White, R. Sampath, L. B. Blyn, D. J. Ecker, R. E. Mandrell, C. K. Fagerquist, A. H. Bates, W. G. Miller, and S. A. Hofstadler. 2008. High-resolution genotyping of campylobacter species by use of PCR and high-throughput mass spectrometry. *J. Clin. Microbiol.* **46**:1220–1225.
- Healy, M., J. Huang, T. Bittner, M. Lising, S. Frye, S. Raza, R. Schrock, J. Manry, A. Renwick, R. Nieto, C. Woods, J. Versalovic, and J. R. Lupski. 2005. Microbial DNA typing by automated repetitive-sequence-based PCR. *J. Clin. Microbiol.* **43**:199–207.
- Healy, M., K. Reece, D. Walton, J. Huang, S. Frye, I. I. Raad, and D. P. Kontoyiannis. 2005. Use of the DiversiLab System for species and strain differentiation of *Fusarium* species isolates. *J. Clin. Microbiol.* **43**:5278–5280.
- Hofstadler, S. A., R. Sampath, L. B. Blyn, M. W. Eshoo, T. A. Hall, Y. Jiang, J. J. Drader, J. C. Hannis, K. A. Sannes-Lowery, L. L. Cummins, B. Libby, D. J. Walcott, A. Schink, C. Massire, R. Ranken, N. White, V. Samant, J. A. McNeil, D. Knize, D. Robbins, K. Rudnik, A. Desai, E. Moradi, and D. J. Ecker. 2005. TIGER: the universal biosensor. *Int. J. Mass Spectrom.* **242**:23–41.
- King, M. D., B. J. Humphrey, Y. F. Wang, E. V. Kourbatova, S. M. Ray, and H. M. Blumberg. 2006. Emergence of community-acquired methicillin-resistant *Staphylococcus aureus* USA 300 clone as the predominant cause of skin and soft-tissue infections. *Ann. Intern. Med.* **144**:309–317.
- Klevens, R. M., M. A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L. H. Harrison, R. Lynfield, G. Dumyati, J. M. Townes, A. S. Craig, E. R. Zell, G. E. Fosheim, L. K. McDougal, R. B. Carey, and S. K. Fridkin. 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA* **298**:1763–1771.
- Koreen, L., S. V. Ramaswamy, E. A. Graviss, S. Naidich, J. M. Musser, and B. N. Kreiswirth. 2004. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. *J. Clin. Microbiol.* **42**:792–799.
- Koreen, L., S. V. Ramaswamy, S. Naidich, I. V. Koreen, G. R. Graff, E. A. Graviss, and B. N. Kreiswirth. 2005. Comparative sequencing of the serine-aspartate repeat-encoding region of the clumping factor B gene (*clfB*) for resolution within clonal groups of *Staphylococcus aureus*. *J. Clin. Microbiol.* **43**:3985–3994.
- McDougal, L. K., C. D. Steward, G. E. Killgore, J. M. Chaitram, S. K. McAllister, and F. C. Tenover. 2003. Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. *J. Clin. Microbiol.* **41**:5113–5120.
- Moran, G. J., A. Krishnadasan, R. J. Gorwitz, G. E. Fosheim, L. K. McDougal, R. B. Carey, and D. A. Talan. 2006. Methicillin-resistant *S. aureus* infections among patients in the emergency department. *N. Engl. J. Med.* **355**:666–674.
- Robinson, D. A., and M. C. Enright. 2004. Multilocus sequence typing and the evolution of methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* **10**:92–97.
- Sampath, R., S. A. Hofstadler, L. Blyn, M. Eshoo, T. Hall, C. Massire, H. Levene, J. Hannis, P. M. Harrell, B. Neuman, M. J. Buchmeier, Y. Jiang, R. Ranken, J. Drader, V. Samant, R. H. Griffey, J. A. McNeil, S. T. Crooke, and D. J. Ecker. 2005. Rapid identification of emerging pathogens: coronavirus. *Emerg. Infect. Dis.* **11**:373–379.
- Spellberg, B., R. Guidos, D. Gilbert, J. Bradley, H. W. Boucher, W. M. Scheld, J. G. Bartlett, and J. Edwards, Jr. 2008. The epidemic of antibiotic-

- resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **46**:155–164.
26. Stevens, D. L., A. L. Bisno, H. F. Chambers, E. D. Everett, P. Dellinger, E. J. Goldstein, S. L. Gorbach, J. V. Hirschmann, E. L. Kaplan, J. G. Montoya, and J. C. Wade. 2005. Practice guidelines for the diagnosis and management of skin and soft-tissue infections. *Clin. Infect. Dis.* **41**:1373–1406.
 27. Tenover, F. C., R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. E. Murray, D. H. Persing, and B. Swaminathan. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J. Clin. Microbiol.* **33**:2233–2239.
 28. Tenover, F. C., L. K. McDougal, R. V. Goering, G. Killgore, S. J. Projan, J. B. Patel, and P. M. Dunman. 2006. Characterization of a strain of community-associated methicillin-resistant *Staphylococcus aureus* widely disseminated in the United States. *J. Clin. Microbiol.* **44**:108–118.
 29. Tenover, F. C., R. R. Vaughn, L. K. McDougal, G. E. Fosheim, and J. E. McGowan, Jr. 2007. Multiple-locus variable-number tandem-repeat assay analysis of methicillin-resistant *Staphylococcus aureus* strains. *J. Clin. Microbiol.* **45**:2215–2219.
 30. Versalovic, J., T. Koeuth, and J. R. Lupski. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res.* **19**:6823–6831.
 31. Weber, S. G., S. S. Huang, S. Oriola, W. C. Huskins, G. A. Noskin, K. Harriman, R. N. Olmsted, M. Bonten, T. Lundstrom, M. W. Climo, M. C. Roghmann, C. L. Murphy, and T. B. Karchmer. 2007. Legislative mandates for use of active surveillance cultures to screen for methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci: position statement from the Joint SHEA and APIC Task Force. *Infect. Control Hosp. Epidemiol.* **28**:249–260.
 32. Woods, C. R., Jr., J. Versalovic, T. Koeuth, and J. R. Lupski. 1992. Analysis of relationships among isolates of *Citrobacter diversus* by using DNA fingerprints generated by repetitive sequence-based primers in the polymerase chain reaction. *J. Clin. Microbiol.* **30**:2921–2929.