

# Epidemic community-associated methicillin-resistant *Staphylococcus aureus*: Recent clonal expansion and diversification

Adam D. Kennedy\*, Michael Otto\*, Kevin R. Braughton\*, Adeline R. Whitney\*, Liang Chen†, Barun Mathema†, Jose R. Mediavilla†, Kelly A. Byrne\*, Larye D. Parkins\*, Fred C. Tenover‡, Barry N. Kreiswirth†, James M. Musser§, and Frank R. DeLeo\*¶

\*Laboratory of Human Bacterial Pathogenesis, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT 59840; †International Center for Public Health, Public Health Research Institute, 225 Warren Street, Newark, NJ 07103; ‡Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, GA 30333; and §Center for Molecular and Translational Human Infectious Diseases Research, Methodist Hospital Research Institute and Department of Pathology, 6565 Fannin Street, B490, Houston, TX 77030

Edited by Richard M. Krause, National Institutes of Health, Bethesda, MD, and approved December 10, 2007 (received for review October 26, 2007)

Emerging and re-emerging infectious diseases, especially those caused by drug-resistant bacteria, are a major problem worldwide. Community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) appeared rapidly and unexpectedly in the United States, resulting in an epidemic caused primarily by isolates classified as USA300. The evolutionary and molecular underpinnings of this epidemic are poorly understood. Specifically, it is unclear whether there has been clonal emergence of USA300 isolates or evolutionary convergence toward a hypervirulent phenotype resulting in the independent appearance of similar organisms. To definitively resolve this issue and understand the phylogeny of USA300 isolates, we used comparative whole-genome sequencing to analyze 10 USA300 patient isolates from eight states in diverse geographic regions of the United States and multiple types of human infection. Eight of 10 isolates analyzed had very few single nucleotide polymorphisms (SNPs) and thus were closely related, indicating recent diversification rather than convergence. Unexpectedly, 2 of the clonal isolates had significantly reduced mortality in a mouse sepsis model compared with the reference isolate ( $P = 0.0002$ ), providing strong support to the idea that minimal genetic change in the bacterial genome can have profound effects on virulence. Taken together, our results demonstrate that there has been recent clonal expansion and diversification of a subset of isolates classified as USA300. The findings add an evolutionary dimension to the epidemiology and emergence of USA300 and suggest a similar mechanism for the pandemic occurrence and spread of penicillin-resistant *S. aureus* (known as phage-type 80/81 *S. aureus*) in the 1950s.

genomics | phylogeny | virulence | single-nucleotide polymorphism

Study of bacterial genome diversification and pathogen emergence has historically relied on strategies that index genomic variation at a relatively small number of nucleotides. Although these strategies are powerful, the fact that members of a bacterial species can differ dramatically in gene content and allelic variation means that analysis of strains at the whole genome level is required to understand important biomedically relevant processes such as epidemics. Whole genome sequencing approaches represent a next generation of studies to address bacterial genome diversification.

One of the most unexpected events in bacterial infectious diseases in recent years has been the rapid emergence of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) infections, now a leading cause of disease in otherwise healthy individuals (1–3). For example, the majority of all skin and soft tissue infections in patients presenting to emergency departments in the United States are caused by CA-MRSA, primarily by isolates classified as pulsed-field gel electrophoresis (PFGE) type USA300 (4–6). In addition, a

recent study by Klevens *et al.* (7) indicates that USA300 isolates account for the majority (67%) of invasive community MRSA infections, some of which result in death. These and other studies clearly indicate there is an ongoing epidemic of CA-MRSA in the United States (8). The origin of isolates classified as USA300 is not known nor do we understand the extent of genetic diversity or the basis of enhanced virulence in these organisms.

To address this deficiency in knowledge and gain understanding about molecular evolutionary processes contributing to temporal variation in disease frequency, we compared the genome sequence of 10 geographically and clinically diverse isolates classified as USA300, and evaluated strain virulence in a mouse infection model.

## Results and Discussion

**Analysis of Geographically Diverse *S. aureus* Isolates Classified as USA300 by PFGE.** As a first step toward understanding the extent of genetic diversity of epidemic community-MRSA, we analyzed 11 isolates classified as USA300 that were isolated from patients with diverse infections and hospitalized at several regions of the U.S. between 2002 and 2005 (Table 1). These strains included community- and healthcare-associated MRSA recovered from patients with diverse clinical manifestations, including skin and soft tissue infections, bacteremia, endocarditis, and necrotizing pneumonia. Organisms were selected largely from a repository of 864 MRSA isolates collected over a 1-year period by the Active Bacterial Core Surveillance system (ABCs) at the Centers for Disease Control and Prevention (CDC) and are representative of major geographic regions of the United States (7). The strain sampling strategy is described in detail in Materials and Methods. Five of the 11 isolates had PFGE profiles that were indistinguishable from one another, suggesting that these organisms are related (Fig. 1). The other 6 isolates differed from the common PFGE type by 1–4 bands (Fig. 1).

**Comparative Genome Sequencing of USA300 Isolates Indicates Recent Clonal Descent of Epidemic CA-MRSA.** PFGE profiling indexes variation at a relatively small number of base pairs and thus does

Author Contributions: A.D.K., F.C.T., B.N.K., J.M.M., and F.R.D. designed research; A.D.K., M.O., K.R.B., A.R.W., L.C., and J.R.M. performed research; F.C.T. contributed new reagents/analytic tools; A.D.K., B.M., K.A.B., L.D.P., B.N.K., J.M.M., and F.R.D. analyzed data; and A.D.K., J.M.M., and F.R.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¶To whom correspondence should be addressed at: 903 South 4th Street, Hamilton, MT 59840. E-mail: fdeleo@niaid.nih.gov.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0710217105/DC1](http://www.pnas.org/cgi/content/full/0710217105/DC1).

© 2008 by The National Academy of Sciences of the USA

**Table 1. USA300 isolates analyzed by comparative genome sequencing**

Isolate	Spa	PFGE*	SCCmec	ACME†	State	Syndrome	Origin‡	Date	Age§
FPR3757¶	1	0114	IVa	+	California	Abscess	CA	2003	36
LAC	1	0114	IVa	+	California	Abscess	CA	2002	n/a
18805	168	0114	IVa	+	Oregon	Endocarditis	CA	2005	54
18806	1	0180	IVa	–	Georgia	Bacteremia	HA	2005	49
18807	1	0114	IVa	+	Colorado	Bacteremia	CA	2005	39
18808	1	0047	IVa	+	New York	Endocarditis	CA	2005	43
18809	1	0114	IVa	+	Minnesota	Bacteremia	CA	2005	63
18810	1	0114	IVa	+	Pennsylvania	Necrotizing pneumonia	CA	2005	21
18811	1	0247	IVa	+	California	Osteomyelitis	HA	2005	94
18813	1	0045	IVb	–	Tennessee	Bacteremia	HA	2005	3
19321	1	0120	IVb	–	Oregon	Pneumonia / bacteremia	HA	2005	50

\* PFGE subtype.

†ACME, arginine catabolic mobile element.

‡CA, community-associated; HA, healthcare-associated.

§n/a, not available.

¶Strain FPR3757 is the reference strain to which all other isolates were compared.

||Los Angeles County clone, LAC.

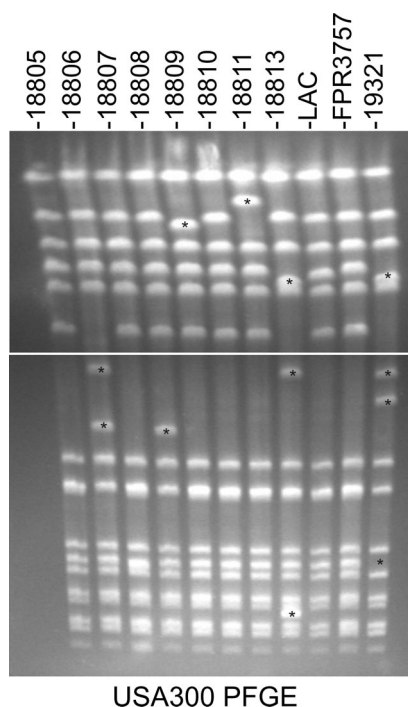
not determine the extent of single nucleotide polymorphisms (SNPs) and many other molecular events contributing to genetic variation. In principle, similarity in PFGE profile can be caused by different molecular processes, including evolutionary convergence or recent clonal descent from a common ancestor. To differentiate between these two possibilities, we performed comparative genome sequencing (CGS), a process that identifies the full extent of SNPs and non-SNP regions of difference (RDs) in genomic and plasmid DNA. [Fig. 2, and [supporting information \(SI\) Tables 2–7](#)]. Strain FPR3757 (USA300-0114) was used as the reference organism because its genome was recently sequenced by conventional methods (9). Non-SNP insertions,

deletions, and transpositions were defined as RDs. All putative SNPs and RDs identified by CGS were evaluated by conventional capillary DNA sequencing, and 91% of the SNPs (526 of 578 SNPs) were confirmed to be true polymorphisms. Compared with the reference strain FPR3757, there was a total of 578 unique SNPs in the core genome ( $\approx 2,753$  kbp excluding foreign elements) of the 10 USA300 isolates. This frequency (1 SNP in every 4,763 bp) is similar to the SNP frequency of 1 in 6,707 bp for the core metagenome of clonally related serotype M3 strains of group A *Streptococcus* reported recently (10). Eighteen percent ( $n = 106$ ) of SNPs in these USA300 isolates were located in intergenic regions and 29% ( $n = 165$ ) and 53% ( $n = 307$ ) were synonymous and nonsynonymous nucleotide changes, respectively (SI Table 2 and SI Fig. 6).

Compared with the reference genome, the number of SNPs in each of the 10 clinical isolates ranged from 11 to 408 SNPs, although 8 of these organisms had a more limited range of SNPs (11 to 48 SNPs) (Fig. 3A). The 8 most closely related isolates differed on average from the reference strain by 32 SNPs and by 50 SNPs from one another. The 2 most divergent isolates (18813 and 19321, Table 1) differed from each other by 477 SNPs, and differed from the reference strain by 408 and 139 SNPs, respectively (Fig. 3A).

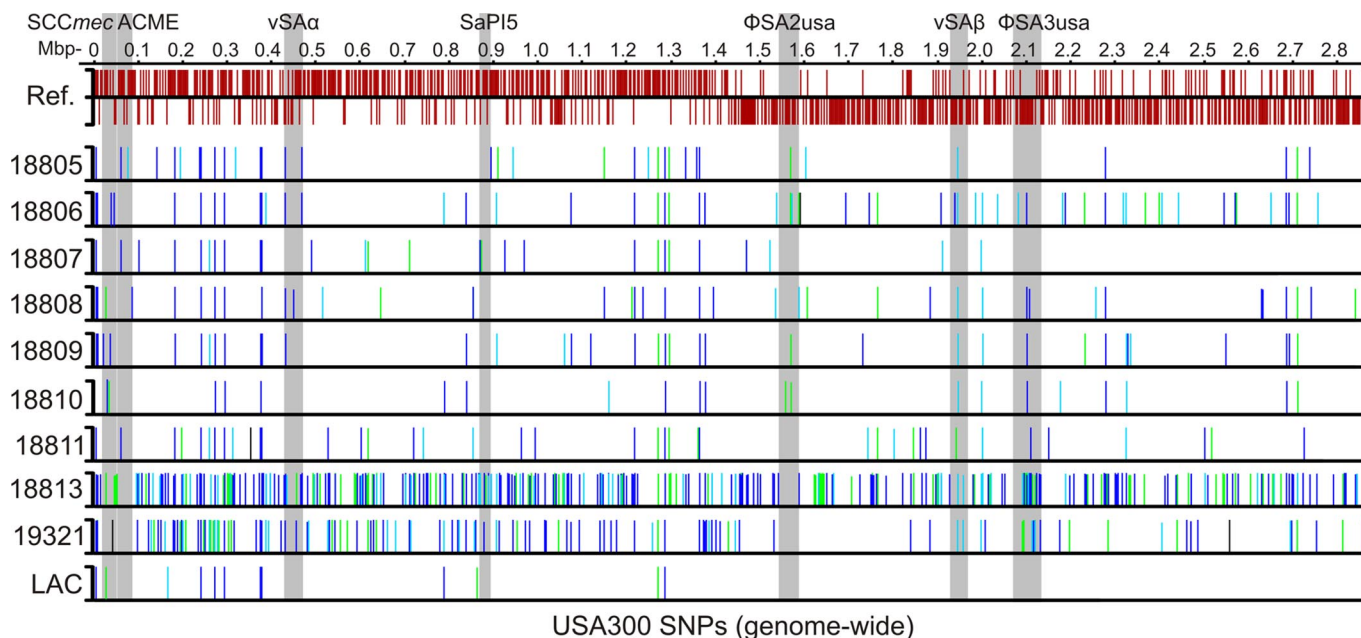
To determine overall genetic relationships among the core chromosome of these 11 isolates, we next conducted SNP-based phylogenetic analysis. As implied by the SNP data, 9 of the 11 isolates are closely related, whereas 18813 and 19321 are more divergent (Fig. 3B and C). Moreover, these 2 isolates were more similar to *S. aureus* strain COL, another reference strain whose genome also has been sequenced (11). These results suggest that COL, 18813, and 19321 diverged from a common ancestor before the genetic diversification events characterizing the cluster of USA300 isolates (Fig. 3D).

Importantly we found a relatively high overall ratio of non-synonymous-to-synonymous SNPs (1.9:1) in the 10 query genomes relative to FPR3757, indicating that the SNP diversity in these strains has accumulated recently (12, 13). Notably, the ratio of nonsynonymous to synonymous SNPs among the 8 most closely related isolates was 2.6:1, a value far greater than the ratio of 1.8:1 between divergent isolates 18813 and 19321. The only reasonable interpretation of these findings is that there has been very recent clonal expansion and geographic dissemination of the nine most closely related isolates (8 query organisms plus FPR3757). These results rule out the formal hypothesis of convergent evolution.



USA300 PFGE

**Fig. 1.** Analysis of USA300 isolates by pulsed-field gel electrophoresis (PFGE). PFGE was performed as described in *Materials and Methods*. Although there is variation among the strains, the PFGE patterns are typical of USA300. FPR3757 is the reference strain. The white line indicates a merge of two separate images.



**Fig. 2.** SNPs in the core genome of USA300 isolates. SNPs were identified by using comparative genome sequencing. Each bar represents a SNP. Green and dark blue vertical lines indicate synonymous and nonsynonymous nucleotide changes, respectively. Light blue indicates a SNP in an intergenic region. The gray shaded regions represent notable genetic elements. Forward (above the line) and reverse (below the line) strand ORFs of the FPR3757 reference strain (Ref.) are illustrated in red at the top of the aligned sequences.

#### Frequency of Core Genome SNPs Present in Genes Encoding Molecules Associated with Virulence.

When considering only the core genome, we found 472 SNPs in 350 unique ORFs (i.e., excluding intergenic regions) of the 10 isolates. These SNPs were distributed widely among ORFs assigned to many functional categories based upon Clusters of Orthologous Groups of proteins (COGs). Many of these 472 SNPs were present in more than one isolate, consistent with divergence from a common ancestor (SI Table 2). However, only 16 nonsynonymous SNPs were present in genes encoding proven or putative *S. aureus* virulence determinants, including exotoxin 7, clumping factor A and B (*clfA* and *clfB*), alpha-hemolysin (*hla*), gamma-hemolysin (*hlgC*), serine and cysteine proteases (*splF*, *sspA* and *sspB*), and fibronectin binding protein A and B (*fnbA* and *fnbB*) (SI Table 2). These genes encode proteins that participate in important pathogenesis processes such as adhesion, host cell lysis, and proteolysis. Isolate 18811 contained a missense mutation in *agrA* that lies within the putative DNA binding domain of the encoded protein, thereby potentially altering function of the accessory gene regulator (Agr) (SI Table 2). Agr is a major regulator of *S. aureus* virulence and controls production and secretion of virulence factors, including hemolysins, adhesions, and leukotoxins (14).

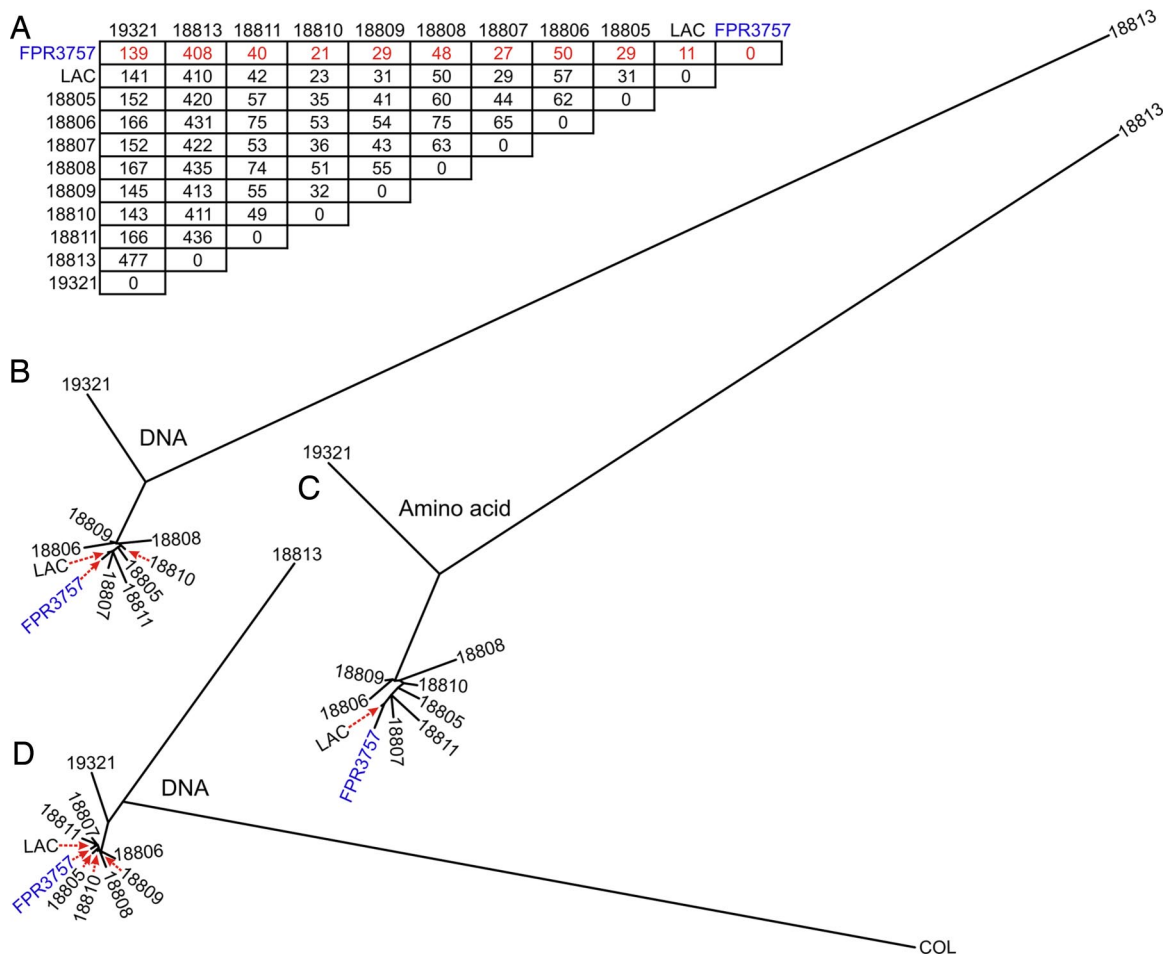
The occurrence of multiple nonsynonymous SNPs present in a single ORF can be a signature of elevated selective pressure or reflect a premium on diversification. Among all strains studied, 57 ORFs had more than one SNP, including *cap5E* (encoding capsular polysaccharide biosynthesis protein 5E), *fnbB*, *sdrC* and *sdrD* (encoding adhesins), *secF*, *tagX* (encoding teichoic acid biosynthesis protein X), and several cell surface or membrane proteins of unknown function (SI Table 2). Most of these SNPs occurred in isolate 18813 (SI Table 2).

**SNPs Present in Foreign Mobile Genetic Elements Support Recent Clonal Expansion of a Subset of Isolates Classified as USA300.** Many genes contributing to *S. aureus* virulence or antibiotic resistance are encoded by mobile elements, such as insertion sequences, transposons, plasmids, prophages, and genomic islands (9, 11, 15–17). In addition to SNPs in the core genome, we identified

365 SNPs in apparently mobile elements, including the *SCCmec* IV cassette (encoding the methicillin resistance genes), arginine catabolic mobile element (ACME), staphylococcal pathogenicity island 5 (SaPI5), prophages  $\phi$ Sa2usa and  $\phi$ Sa3usa, and plasmids pUSA01 and pUSA03 (SI Table 3). The two most genetically divergent isolates (19321 and 18813) had the greatest number of SNPs in mobile elements (83 and 262 of the 365 SNPs, respectively), compared with only 5–20 SNPs in all other organisms (some SNPs are present in more than one isolate) (SI Table 3). The ratio of nonsynonymous to synonymous SNPs in mobile elements was 1:1.7 for isolates 18813 and 19321 (relative to the reference strain), and this ratio was much different for the other isolates (1:0.9). Collectively, these findings are in accordance with the analysis of SNPs present in the core genome, which indicate earlier divergence of 18813 and 19321 followed by subsequent clonal expansion and diversification of isolates 18805–18811, LAC, and FPR3757.

**Identification of Non-SNP Polymorphisms in USA300 Isolates.** Compared with FPR3757, there were 17 unique RDs (excluding nonsense mutations, SI Table 4) among the 10 USA300 query isolates, mapping to seven loci and two intergenic regions within the core genome and 8 chromosomal RDs in foreign mobile elements (SI Tables 5 and 6 and SI Fig 7). There were 36 additional RDs identified in plasmids by CGS, but these RDs could not be confirmed by targeted conventional capillary DNA sequencing (SI Table 7). Consistent with data reported for group A *Streptococcus* (10), there was a high level of false-positive calls among the RDs ( $\approx 63\%$ ). This phenomenon may be due to segments of repetitive DNA, low DNA melting temperature, or DNA secondary structure. Three RDs were deletions within or encompassing ACME in isolates 18806, 18813, and 19321 (SI Table 6). Among *S. aureus*, ACME is primarily associated with isolates classified as USA300 and may promote survival within the host (9, 11, 15–17). Although 18806 lacks ACME, its core genome clearly places it among organisms involved in the recent clonal expansion of USA300 strains (Fig. 3B). In addition, 18806 has a *SCCmec* subtype IVa *mecA* gene versus subtype IVb in





**Fig. 3.** SNP matrix and SNP-based phylogenetic analysis. (A) SNP matrix. Red font indicates number of SNPs compared with reference strain FPR3757. (B) Phylogenetic analysis of the 11 USA300 isolates is based upon 578 concatenated core genomic SNPs and includes SNPs present in intergenic regions. (C) Phylogenetic analysis using 472 concatenated amino acid residues containing SNPs, synonymous, and nonsynonymous. (D) Phylogenetic analysis of USA300 isolates and strain COL is based upon 1,063 concatenated core genomic SNPs.

18813 and 19321 (Table 1). Isolate 18813 had the greatest number of non-SNP RDs (6 of the 17 total) and PCR analysis confirmed that the ACME and  $\phi$ SA2<sub>usa</sub> were absent in this isolate. Isolate 18805 had a 227-bp deletion that would produce a frameshift in *SAUSA300\_2589*, a gene encoding a secreted protein that has an LPXTG-cell wall anchor peptide (SI Table 5). It is also notable that this gene is juxtaposed to *secY*, which regulates secretion of exported proteins.

**Allelic Variation among USA300 Isolates Significantly Affects Virulence.** The restricted allelic variation identified in the core genome of USA300 isolates suggested there would be little or no difference in virulence among these organisms. However, there is an emerging literature that modest genetic changes such as SNPs can dramatically alter pathogen-host interaction and strain virulence (10, 18, 19). Understanding whether differences exist in virulence among clonally related organisms is a critical issue as the population size of USA300 isolates expands and genetic diversification increases. We used a mouse sepsis/bacteremia model to test the hypothesis that the 11 isolates we analyzed differed in virulence (Fig. 4). Consistent with this hypothesis, our analysis revealed significant variance in virulence among these 11 isolates (Fig. 4). Survival was significantly greater for animals infected with isolates 18805 and 18811 compared with the other 9 isolates (Fig. 4). Remarkably, the core genome of 18805 and 18811 was relatively similar to those of the most virulent isolates,

differing from the reference strain (FPR3757) by only 29 and 40 SNPs, respectively (Fig. 3A). There were a limited number of nonsynonymous SNPs or other polymorphisms unique to these two isolates, including those in the core genome and mobile elements excluding plasmids (7 and 14 SNPs in 18805 and 18811, respectively; and 2–3 non-SNP deletions in each isolate) (SI Tables 2, 3, and 5). None of the polymorphisms in 18805 were within putative or proven virulence factors, whereas 18811 had nonsynonymous SNPs in genes encoding AgrA and VraF (*SAUSA300\_0647*), each of which could negatively impact virulence (SI Table 2).

**Correlation of SNPs and Virulence with Patterns of Exoprotein Expression.** *S. aureus* produces many secreted exoproteins that contribute significantly to virulence, such as enterotoxins, cytolytic toxins, and protein A (Spa). Given the significant variance in mouse virulence among the highly related isolates in our sample, we tested the hypothesis that one or more of the SNPs we identified resulted in differences in exoprotein profiles (Fig. 5A–D). Nine of the 11 isolates had a very similar exoprotein profile, but this profile was remarkably different from those observed for 18805 and 18811, the two organisms of low virulence in the mouse model (Fig. 5A). The culture supernatant of these two isolates lacked  $\alpha$ -hemolysin (Hla) and had reduced levels of LukF-PV (a subunit of Panton-Valentine leukocidin, PVL) compared with the reference strain (Fig. 5C). There was



genome sequencing approaches have provided unprecedented detail about mechanisms of bacterial pathogen success or multidrug resistance. For example, the molecular basis for the origin and emergence of a highly successful clone of serotype M1 group A *Streptococcus* was made possible only by full genome resequencing (18), and studies by Mwangi *et al.* (19) used whole genome sequencing to identify steps in the evolution of antibiotic resistance in *S. aureus*.

Our studies reveal there has been recent clonal expansion (emergence) of a subset of USA300 isolates followed by diversification and thus provide insight into the molecular genetic basis of the current CA-MRSA epidemic. Previous work suggested CA-MRSA infections are caused by clonal groups of strains, but the strain typing methods used index genetic variation at a relatively few nucleotides and were not sufficient to demonstrate clonal emergence (5, 6).

Although our data unambiguously show that a number of USA300 isolates had a common ancestor, this finding should not be construed to mean that they are genetically identical nor have similar virulence characteristics. In fact, two of the clonal isolates analyzed in this study had a different virulence and exoprotein profile compared with the reference strain (Figs. 4 and 5). We acknowledge that relatively small changes in growth conditions in the laboratory or during processing may have a substantial impact on virulence. However, all strains tested had similar growth kinetics *in vitro* (SI Fig. 8) and phenotypic differences noted with the two isolates of low mouse virulence (e.g., absence of Hla) indicate the phenomenon is not likely related to variances in day-to-day growth *in vitro*. Taken together, these results are consistent with the idea that modest genetic changes in bacterial pathogens can have a dramatic impact on host-pathogen interaction, virulence and/or antibiotic resistance phenotype (10, 18, 19, 28). Therefore, it is reasonable to conclude that within a short period there will be an increasing pool of derivative USA300 isolates that differ in virulence potential and/or pathogenicity. A comprehensive understanding of genomic variation at the single nucleotide level is critical to elucidating the events that define bacterial virulence, multidrug resistance, and epidemics.

## Materials and Methods

For full details, see *SI Materials and Methods*.

**Bacteria and Culture Conditions.** Nine MRSA isolates (Table 1) classified as USA300 by PFGE were selected from among 864 MRSA isolates from invasive infections collected over an  $\approx$ 1-year period through the Active Bacterial Core Surveillance system (ABCs) at the Centers for Disease Control and Prevention (CDC) in Atlanta (7). Of these 864 isolates, 252 (29.5%) were classified as USA300. Nine study isolates were selected from among 252 USA300 isolates to yield broad geographic representation from across the U.S., a variety of clinical syndromes, and a range of patient ages. The isolates were drawn from eight surveillance sites (California, Colorado, Georgia, Minnesota, New York, Oregon, Pennsylvania, and Tennessee). Five isolates met criteria for community-associated MRSA infections (1, 5) and 4 were healthcare-associated infections. Strain FRP3757 was kindly provided by Binh An Diep (University of California, San Francisco) and is the sequenced reference strain used for these studies. LAC was selected for this study because it was recovered early during the CA-MRSA epidemic (2002), has been studied in virulence models, and was recovered from a skin and soft tissue infection, which represents the majority of CA-MRSA infections.

**Comparative Genome Sequencing (CGS).** The protocol used for comparative genome sequencing (NimbleGen, Inc.) is available at [www.nimblegen.com](http://www.nimblegen.com).

**Confirmation of SNPs and Analysis of RDs.** SNPs and RDs identified by CGS were verified by conventional capillary DNA sequencing using oligonucleotide primers (Sigma Genosys, The Woodlands, TX) designed with Vector NTI software (Invitrogen Corp., Carlsbad, CA).

**Phylogenetic Analyses.** Evolutionary relationships were reconstructed by using IQPNNI software (29). The DNA sequences analyzed were generated by concatenating nucleotides at all of the core genomic DNA SNP loci identified in sequential order relative to the FRP3757 reference genome sequence.

**ACKNOWLEDGMENTS.** We thank the Rocky Mountain Laboratories (RML) Genomics Unit, Research Technologies Section, Research Technologies Branch, National Institute of Allergy and Infectious Diseases (NIAID) for assistance with DNA sequencing and bioinformatics analyses; and the ABCs system of the Centers for Disease Control and Prevention (CDC, Atlanta, GA). The findings and conclusions in this publication are those of the authors and do not necessarily represent the views of the CDC. This work was supported by the Intramural Program of the NIAID, National Institutes of Health.

- Zetola N, Francis JS, Nuermberger EL, Bishai WR (2005) *Lancet Infect Dis* 5:275–286.
- Chambers HF (2005) *N Engl J Med* 352:1485–1487.
- Chambers HF (2001) *Emerg Infect Dis* 7:178–182.
- Moran GJ, Krishnadasan A, Gorwitz RJ, Fosheim GE, McDougal LK, Carey RB, Talan DA (2006) *N Engl J Med* 355:666–674.
- McDougal LK, Steward CD, Killgore GE, Chaitram JM, McAllister SK, Tenover FC (2003) *J Clin Microbiol* 41:5113–5120.
- Diep BA, Sensabaugh GF, Sombona NS, Carleton HA, Perdreau-Remington F (2004) *J Clin Microbiol* 42:2080–2084.
- Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, *et al.* (2007) *J Am Med Assoc* 298:1763–1771.
- King MD, Humphrey BJ, Wang YF, Kourbatova EV, Ray SM, Blumberg HM (2006) *Ann Intern Med* 144:309–317.
- Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, *et al.* (2006) *Lancet* 367:731–739.
- Beres SB, Richter EW, Nagiec MJ, Sumbly P, Porcella SF, DeLeo FR, Musser JM (2006) *Proc Natl Acad Sci USA* 103:7059–7064.
- Gill SR, Fouts DE, Archer GL, Mongodin EF, Deboy RT, Ravel J, Paulsen IT, Kolonay JF, Brinkac L, Beanan M, *et al.* (2005) *J Bacteriol* 187:2426–2438.
- Gutacker MM, Smoot JC, Migliaccio CA, Ricklefs SM, Hua S, Cousins DV, Gravis EA, Shashkina E, Kreiswirth BN, Musser JM (2002) *Genetics* 162:1533–1543.
- Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, Feil EJ (2006) *J Theor Biol* 239:226–235.
- Novick RP (2003) *Mol Microbiol* 48:1429–1449.
- Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, *et al.* (2002) *Lancet* 359:1819–1827.
- Holden MT, Feil EJ, Lindsay JA, Peacock SJ, Day NP, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, *et al.* (2004) *Proc Natl Acad Sci USA* 101:9786–9791.
- Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, *et al.* (2001) *Lancet* 357:1225–1240.
- Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, *et al.* (2005) *J Infect Dis* 192:771–782.
- Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de LH, Richardson P, Bruce D, Rubin E, Myers E, Siggia ED, *et al.* (2007) *Proc Natl Acad Sci USA* 104:9451–9456.
- Recei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A, Novick RP (1986) *Mol Gen Genet* 202:58–61.
- Voyich JM, Otto M, Mathema B, Braughton KR, Whitney AR, Welty D, Long RD, Dorward DW, Gardner DJ, Lina G, *et al.* (2006) *J Infect Dis* 194:1761–1770.
- Stewart GT, Holt RJ (1963) *Br Med J* 1:308–311.
- Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG (2002) *Proc Natl Acad Sci USA* 99:7687–7692.
- Robinson DA, Kearns AM, Holmes A, Morrison D, Grundmann H, Edwards G, O'Brien FG, Tenover FC, McDougal LK, Monk AB, *et al.* (2005) *Lancet* 365:1256–1258.
- Kirby WM (1944) *Science* 99:452–453.
- Barber M, Hayhoe FG, Whitehead JE (1949) *Lancet* 2:1120–1125.
- Jevons MP (1961) *Br Med J* 1:124–125.
- Highlander SK, Hulten KG, Qin X, Jiang H, Yerrapragada S, Mason EO, Shang Y, Williams TM, Fortunov RM, Liu Y, *et al.* (2007) *BMC Microbiol* 7:99.
- Vinh LS, von Haeseler A (2004) *Mol Biol Evol* 21:1565–1571.